Evolution and plasticity of the gene regulatory networks controlling sex determination: of masters, slaves, usual suspects and usurpators.

Amaury Herpin

UR/1037
FISH PHYSIOLOGY
AND GENOMICS [LPGP]

HABILITATION à DIRIGER les RECHERCHES (HDR)

Ecole Doctorale EGAAL

Amaury HERPIN

INRAE, Fish Physiology and Genomics Institute

Evolution and plasticity of the gene regulatory networks controlling sex determination: of masters, slaves, usual suspects and usurpators.

-Emphasis on fish sex determination, differentiation and maintenance-

Referees:

-Dr **Marie-Christine CHABOISSIER**, Institut de Biologie Valrose (iBV), Nice, France.

-Dr **Nathalie Di CLEMENTE**, Hôpital St-Antoine, Site Bâtiment Kourilsky, Paris, France.

-Pr **Lukas KRATOCHVIL Lukas**, Department of Ecology, Faculty of Science Charles University Vinicna, Praha, Czech Republic.

Examiners:

-Dr **Eric Pailhoux**, UMR BDR, INRAE, ENVA, Univerité Paris Saclay, France.

-Pr **Gilles Salbert**, Institut Génétique et Développement de Rennes, UMR 6290 CNRS, France.

Table of Contents

Presentation of the candidate

Amaury Herpin

Chargé de Recherches, INRA Rennes (France) Unit 1037, Fish Physiology and Genomics, team "Sexe, Ovogenèse et ComportementS".

Senior Scientist, Hunan Normal University (China), State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences.

Tel: +33 02 23 48 50 18
Email: amaury.herpin@.inra.fr
Profile: http://www.researcherid.com/rid/B-8950-2011

Citizenship: French
Date and place of Birth: 06/23/1975, Blois (41) Loir et Cher, France
Marital Status: married, two children (Nils, 2003 and Louison, 2008)
Languages: fluent in French and English (written and spoken), skills in Norwegian and German.

Education Background - Employments

Baccalaureat:
June 1993 – Section C (mathematics and physics) – Lycée Dessaignes – BLOIS (41), France.

Uinversity:

1994 – 1996 **DEUG** Sciences Cursus II; University of Orléans (45).
1996 – 1997 **Licence** of Biology, Geology, and Biotechnologies; University of Orléans.
1997 – 1998 **Master** in Biology and Cellular Physiology; University of Orléans.
1998 – 1999 **DEA** in Cellular Biology (Rouen-Caen).
Diploma thesis: « Identification and characterisation of a cDNA coding for a TGF-β receptor in the oyster Crassostrea gigas ».

1999-2003 **Joint Ph.D.**: Universities of Caen (France) and University of Bergen (Norway)
-Laboratoire de Biologie et Biotechnologies Marines (France, M. Mathieu)
-Sars International Centre for Marine Molecular Biology / EMBL (Norway, D. Chourrout)

Thesis: "Structural and functional characterisation of TGF-β family members in the oyster Crassostrea gigas."

2003 - 2005 **Post-Doc.** University of Wuerzburg (Germany), Group of Manfred Schartl.
2005 - 2006 **Scientific Assistant**. University of Wuerzburg (Germany), Physiological Chemistry I.
2006 - 2014 **Civil Servant**. University of Wuerzburg (Germany), Physiological Chemistry I.
2014 – present **Chargé de Recherches I**, INRA Rennes (France), Fish Physiology and Genomics Institute, SOCS team.
2020 – present **Senior Scientist**, Hunan Normal University (China), State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences.

Scientific Activities

-**Reviewer** (peer review) for: Development, Gene, PLoS One, PLoS Genetics, Marine Biotechnology, Protein and Cell, Journal of Genetics (…)

-**Reviewer** for the National Science Foundation (NSF), USA; Agence Nationale de la Recherche (ANR), France; and Czech Science Foundation, Czech Republic.

-Guest Editor (2007, 2011) for a FEBS journal mini review series dealing with BMPSignalling and Vertebrate Sex Determination.

-Editorial Board Member at "Reproduction and Breeding" (Elsevier).

-**Member of the exam committee** for bachelor and master students at the University of Wuerzburg, Germany (2003-2014).

-PhD referee for the University of Singapore (2014).

-**PhD committees** INRA (2014-present).

-Co-coordination of the "CRISPR/Cas9 club", INRA, CNRS, INSERM, Rennes (2015-2016).

-**Coordinator** of the scientific animation at Fish Physiology and Genomics Laboratory, Rennes, France.

-Member of a special-topic-network (Connecting population genetics and developmental biology to elucidate vertebrate sex evolution, CONGEN-DEVSEX), ESEB, European Society for Evolutionary Biology, 2017-2020).

-**Active Member** of a working group on Genome Editing, INRA (2019-2020).

Meetings (speaker/invited speaker):

* 2002	Norwegian Biochemical Society Meeting, Roros, Norway. Oral Communication.
* 2003	International Society of Developmental and Comparative Immunology meeting, Scotland. Oral Communication.
* 2003	Marine Biotechnology Conference, Chiba, Japan. Oral Communication: "*TGF-β superfamilly receptors in the bivalve mollusc Crassostrea gigas*"
* 2006	4th International Symposium on Vertebrate Sex Determination, Hawai, USA. Oral Communication: "*Inhibition of male germ cell proliferation by Dmrt1bY* ".
* 2008	International Symposium for Gonad and Brain Sex Differentiation, Fukuoka, Japan Oral Communication: "*Expression, Regulation and Sex Determining Function of Dmrt1bY in Medaka*".
*2014	5th EFOR Annual Meeting, Paris, France. Invited Speaker: "Transcriptional rewiring of the sex-determining dmrt1 gene: how Dmrt1bY was born in Medaka fish".
*2014	2nd Strategical Meeting for Medaka research (PI meeting), Sevilla, Spain. Invited speaker.

*2014 27th Conference of European Comparative Endocrinologists, Rennes, France. Invited speaker.

*2015 7th International Symposium on Vertebrate Sex Determination, Hawai, USA. Oral Communication: "A novel evolutionary conserved mechanism of RNA stability regulates dmrt1bY expression in PGCs prior to the sex determination stage". 1st *ex aequo* best talk.

*2017 1st European Symposium on sex determination in vertebrates, Dinard, France. Invited speaker. "Insertion of a transposable element rewires transcriptional regulation of a novel sex-determining gene".

*2018 European Society for Evolutionary Biology (ESEB) Special Topic Network, Prague, Czech Republic. Invited speaker. "Connecting genome-wide approaches and functional genomics for elucidating evolution in motion in fish."

*2018 8th International Symposium on Vertebrate Sex Determination, Hawai, USA. Oral Communication: "Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements".

*2019 Workshop: "Paradigm shift in sex chromosome evolution", Berlin, Germany. Invited speaker.

*2019 Workshop: "CRISPR/Cas9 improvement in bivalve molluscs, IFREMER", Nantes, France. Invited speaker.

Teaching Experience:

* 1999 Lab course (128 hours) Reproduction and development, University of Caen, France.

* 2004-2014 Lab course, seminar and tutorials, Developmental biology using Medaka, University of Wuerzburg. Medical students (2nd year, 22 hours per semester).

* 2015-2018 Master II, Biology of reproduction; University of Tours (4 hours).

Students:

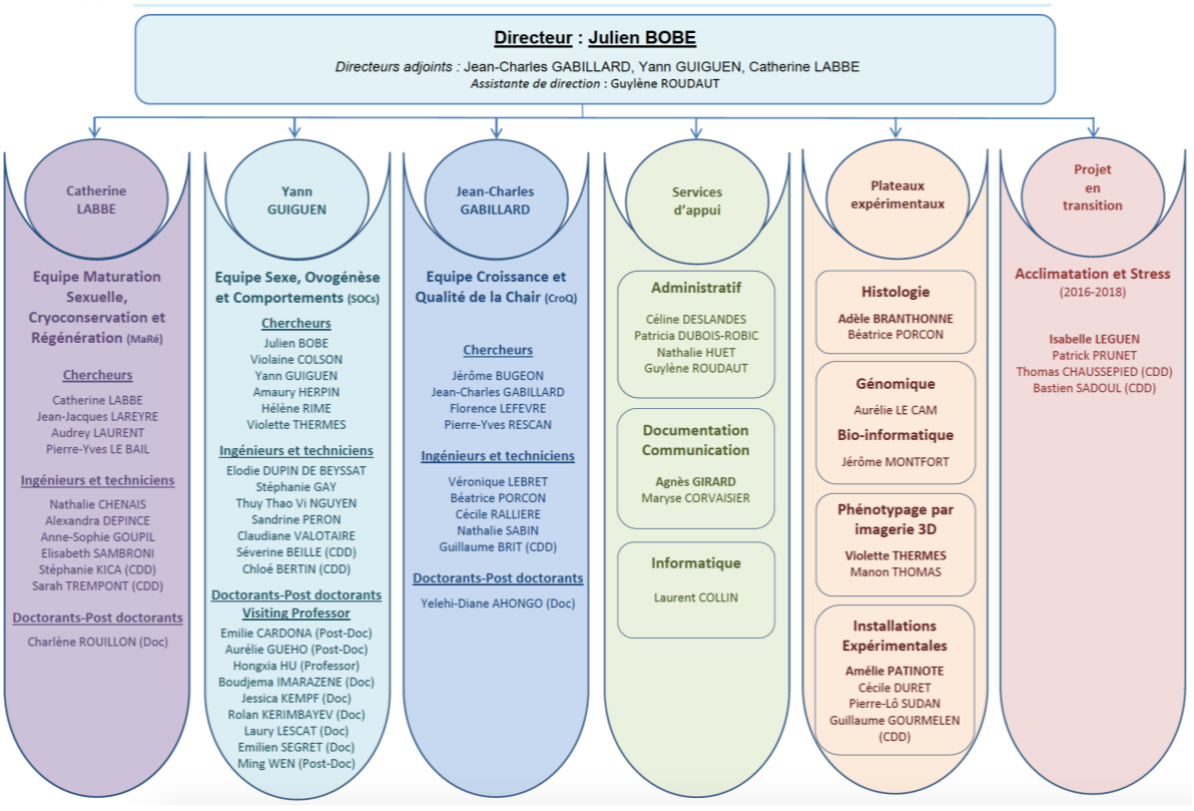
* 2008 J. Driessle, Bachelor Biomedicine (supervisor).
* 2010 J. Klughammer, Bachelor Biomedicine (supervisor).
* 2010 M. Engel, Bachelor Biomedicine. (supervisor).
* 2010-2012 M. Zehner, Ph.D. medical (supervisor).
* 2010-2012 M. Hinzmann, Bachelor/Diploma Biomedicine (supervisor).
* 2008-2014 P. Fischer, Ph.D. biology (co-supervisor).
* 2012- 2016 M. Adolfi, Ph.D. biology (co-supervision). Sao Paulo University, Brasil.
* 2012- 2016 S. Bertho, Ph.D. biology (co-supervision/Ph.D committee).
* 2014- 2017 M. Pan, PhD. biology (co-supervision/Ph.D committee).
* 2015-2019 A. Martinez, PhD. biology (supervisor).
* 2017-2020 B. Imarazene, PhD. biology (co-supervisor).
* 2017 C. Pagneux, DUT biology (supervisor).
* 2017 C. Lane, Master II biology (supervisor).
* 2016-2019 L. Lescat, Ph.D biology (co-supervisor).
* 2018-2020 J. Kempf, Ph.D biology (supervisor).
* 2020 S. Petit, DUT biology (supervisor).

Grants:

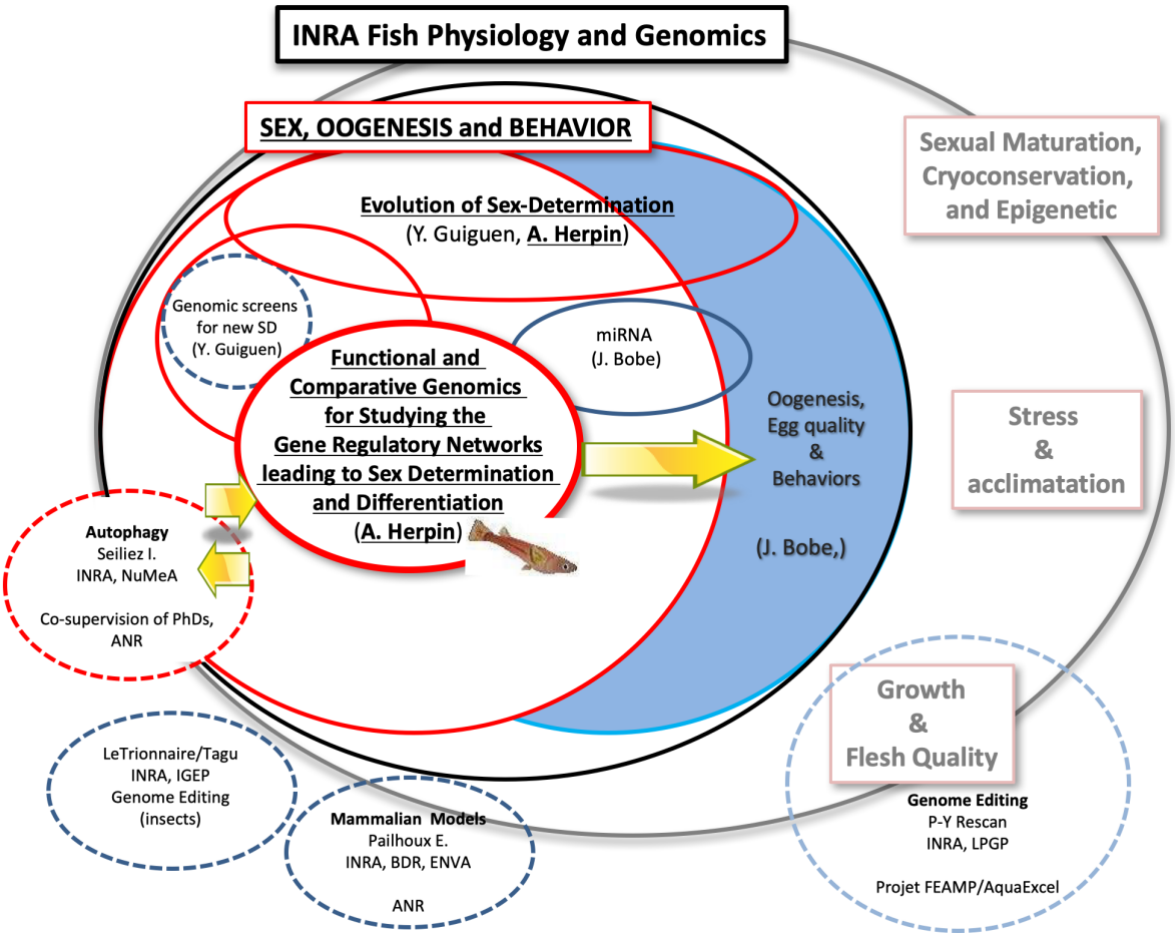
* 2007-2010 GRK1048, #7F57C Pluripotency and differentiation of the germ line. (**Partner**/ PI: M. Schartl)
* 2010-2013 NSF (1R01GM085318) Mechanisms of sex determination in zebrafish. (**Partner**/ PI: J. Postlethwait)
* 2012-2015 ANR (11-BSV7-0016) Role and evolution of SdY in salmonids. (**Partner**/ PI: Y. Guiguen)
* 2015-2019 DFG (HE7135/2-1) Post-transcriptional regulation of Dmrt1 during sexual development. (**Coordinator**)
* 2015-2017 ACI /dept. PHASE INRA (**Coordinator**)
* 2017-2021 COFASP ERA-net (ANR-16-COFA-0004-01) AquaCRISPR (**Co-coordinator**)
* 2018-2021 ANR Fish'n Chap (**Partner**)
* 2019-2021 AquaExcel 3.0, European Project (**Task leader**)
* 2020-2023 TUNESAL Robust Atlantic Salmon Through Fine-Tuned Genome Editing (Research Project- HAVBRUK2, PN: 294971, **Task leader**)
* 2020-2025 111 Project (China, Grant No. D20007, **Partner**)

Functional Organigram of my scientific Environment (INRA)

- Laboratory:



- Team:

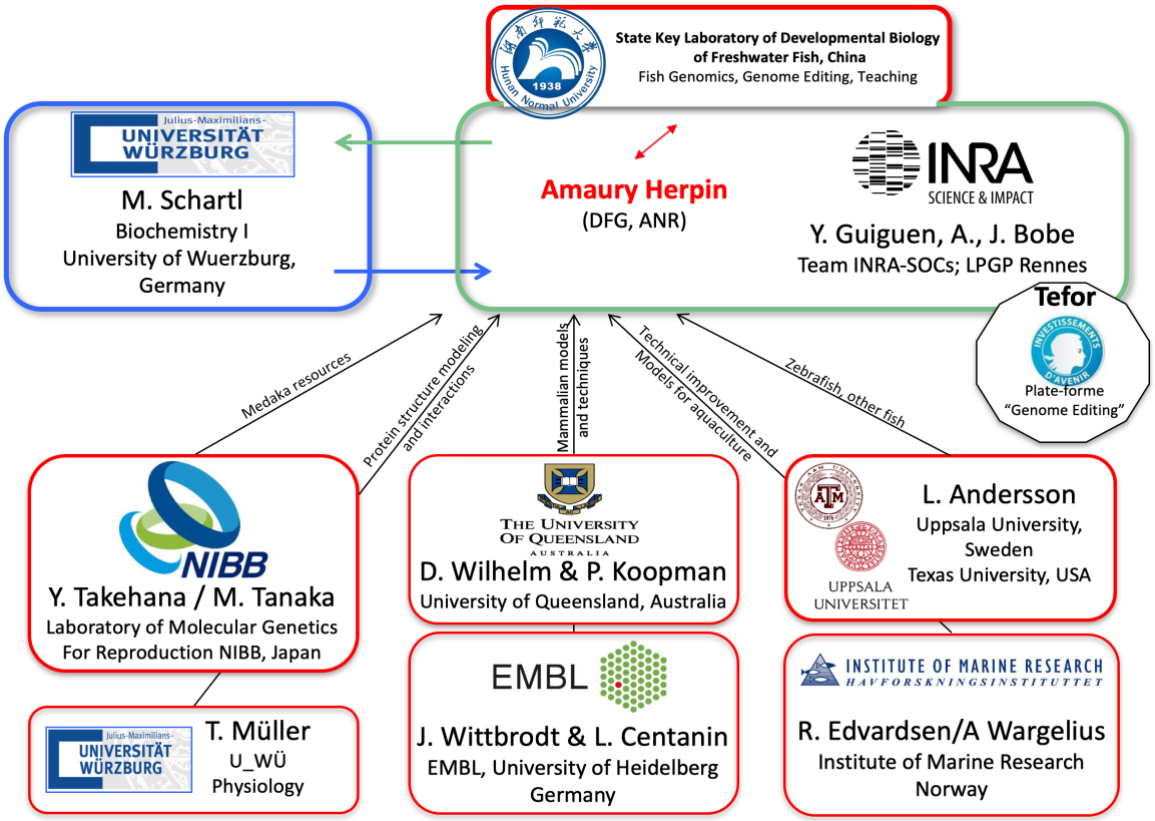


Functional diagram of my work interactions with my colleagues at INRA/LPGP

Within the "Fish Physiology and Genomics Institute" (LPGP), I belong to the "Sex Oogenesis and behaviours" (SOCs) team. There I have developed a thematic dealing with functional genomics approaches for studying sex determination (SD). In addition to students, I directly supervise one assistant engineer (Alexandra Depincé). Within the SOCs team I daily interact and collaborate with Yann Guiguen (evolution of sex determination through genomic approaches) and Julien Bobe (functional genomics applied to oocyte development). Within the laboratory I developed several collaborations mainly focusing on genome editing in salmonids and single cell transcriptomics (Pierre-Yves Rescan).

At INRA I am mainly collaborating with Iban Seiliez (INRA, NuMeA, see page 54). That long-term and fruitful collaboration, not centred on sex determination but rather on autophagy, nevertheless allows us to get fundings and students in co-supervision (two PhD students so far), and… good publications **[see papers #37, #46 and #47 in the publication list]**. I also collaborate with Gael LeTrionnaire at INRA for setting up genome editing in insects **[see papers #34, #41 and BC#5]**. Last, I also like to collaborate with Eric Paillhoux (INRA, BDR, ENVA). Working on mammalian species we always use that opportunity for conducting comparative approaches in order to confer a higher visibility to our respective studies (we wrote several ANR grants together) **[see papers #34, #39, #40]**.

- Ongoing international collaborations:



Since I have been mainly working abroad, I have also developed an international network of reliable collaborations. **Manfred Schartl** is the person with who I have been working for more than 10 years on sex determination, first as postdoc and then progressively independently. During that period, I could settle down and fine tune my own thematic, studying the evolution and plasticity of the underlying gene regulatory networks controlling and regulating sex determination (SD) in fish. Still, Manfred is my main collaborator and we regularly write grant applications together **[see papers #5, #9, #13 to #18, #20 to #31, #33 to #36, #38 to #40, #42 to #45, #48 and #49]**. **Minoru Tanaka** is also working on medaka SD, and we have been collaborating in a concerted manner for many years. Hence, we do not only share transgenic lines and molecular tools, but also data and information. Our collaborations have already resulted in common publications **[see papers #18, #21, #28 and #29, and #42]**. Currently, both of us have interest in autophagy and sexual development in medaka, topic for which we share mutant medaka lines. More recently I have been working with **Yusuke Takehana**, who is a young professor working on evolution of sex determination systems in the *Oryzias* lineage **[see paper #45]**. Thanks to Yusuke, I can get access to valuable genetic resources (different medaka species and populations). With the help of **Dagmar Wilhelm**, Institute for Molecular Bioscience, University of Queensland, Australia, we are also considering mammalian models, still having in mind a comparative and evolutionary view of sex determination mechanisms. This comparative approach will be extended to many aspects of our projects (pronephric contribution, gene regulatory network evolution, TGF-β…) using the mouse model in order to confer a higher scientific visibility to our research **[see papers #36 and #40]**. Dagmar is nowadays a very reliable collaborator. **Thomas Müller** (Julius-von-Sachs Institute, Germany) is a structural biologist/protein chemist internationally recognized in the field of TGFβ/BMP. Thomas usually brings support to our projects (sex determination, TGF-β and autophagy) giving his expertise on protein chemistry and structural biology to explore ligand/receptors and protein/protein interactions using molecular modelling, experimental structure analysis (X-ray crystallography) and protein chemistry. We know each other very well and have been collaborating already **[see paper #39]**. He will again collaborate with us in the future, especially with the TGF-β (ANR grant submitted) and autophagy projects (see below). **Lazaro Centanin/Jochen Wittbrodt** are good friends of mine also working on medaka. We regularly share mutant and transgenic (brainbow) lines and develop new technics together **[see paper #36]. Rolf Edvardsen/Anna Wargelius** are Norwegian colleagues I met during my PhD time in Bergen. We now collaborate to adapt technical protocols (KO, KI, allelic replacement…) from medaka to salmon and therefore wrote several successful grants together.

Publications:

50 publications, 5 book chapters; h-index: 24 (Thomson Reuters).

1. **Herpin, A**., Favrel, P. & Cunningham, C. (2002) Gene structure and expression of cg-ALR1, a type I activin-like receptor from the bivalve mollusc Crassostrea gigas, ***Gene***. *301*, 21-30. *(IF=2.778)*
2. **Herpin, A**., Badariotti, F., Rodet, F. & Favrel, P. (2004) Molecular characterization of a new leucine-rich repeat-containing G protein-coupled receptor from a bivalve mollusc: evolutionary implications, ***Biochim Biophys Acta***. *1680*, 137-44. *(IF=3.810)*
3. **Herpin, A**., Lelong, C. and Favrel, P. (2004) Transforming growth factor-beta-related proteins: an ancestral and widespread superfamily of cytokines in metazoans, ***Dev Comp Immunol***. *28*, 461-85. *(IF=2.652)*
4. Huvet, A., **Herpin, A**., Degremont, L., Labreuche, Y., Samain, J. F. and Cunningham, C. (2004) The identification of genes from the oyster Crassostrea gigas that are differentially expressed in progeny exhibiting opposed susceptibility to summer mortality, ***Gene***. *343*, 211-20. *(IF=2.705)*
5. Kluver, N., Kondo, M., **Herpin, A**., Mitani, H. and Schartl, M. (2005) Divergent expression patterns of Sox9 duplicates in teleosts indicate a lineage specific subfunctionalization, ***Dev Genes Evol***. *215*, 297-305. *(IF=2.549)*
6. **Herpin, A**., Lelong, C., Becker, T., Rosa, F. M., Favrel, P. and Cunningham, C. (2005) Structural and functional evidences for a type 1 TGF-beta sensu stricto receptor in the lophotrochozoan Crassostrea gigas suggest conserved molecular mechanisms controlling mesodermal patterning across bilateria, ***Mech Dev***. *122*, 695-705. *(IF=3.838)*
7. **Herpin, A**., Lelong, C., Becker, T., Rosa, F., Favrel, P. and Cunningham, C. (2005) Structural and functional evidence for a singular repertoire of BMP receptor signal transducing proteins in the lophotrochozoan Crassostrea gigas suggests a shared ancestral BMP/activin pathway, ***Febs J***. *272*, 3424-40. *(IF=3.415)*
8. **(Herpin, A., Gueguen, Y.)** ***equal contribution***, Aumelas, A., Garnier, J., Fievet, J., Escoubas, J. M., Bulet, P., Gonzalez, M., Lelong, C., Favrel, P. and Bachere, E. (2006) Characterization of a defensin from the oyster crassostrea gigas: Recombinant production, folding, solution structure, antimicrobial activities and gene expression, ***J Biol Chem.*** *281*, 313-323. *(IF=5.808)*
9. **Herpin, A**., Rohr, S., Riedel, D., Kluever, N., Raz, E., and Schartl, M. (2007) Specification of primordial germ cells in medaka (Oryzias latipes), ***BMC Dev Biol.*** *7*, 3. (Highly accessed, top 10 most viewed articles of all time in 2009, 2010 and 2011). *(IF=3.337)*
10. **Herpin, A**., Lelong, C., Becker, T., Favrel, P., and Cunningham, C. (2007) A tolloid homologue from the pacific oyster Crassostrea gigas, ***Gene Expr Patterns.*** *7*, 700-708. *(IF=2.238)*
11. **Herpin, A**. (2007) BMP signaling pathways, ***FEBS J.*** *274*, 2959. *(IF=3.396)* *** Cover picture of the issue**.
12. **Herpin, A**. and Cunningham, C. (2007) Cross-talk between the Bone Morphogenetic Protein pathway and other major signaling pathways results in tightly regulated cell-specific outcomes, ***FEBS J***. *274*, 2977-85. *(IF=3.396)*

*** Top-Cited Paper Award.**

13. Hornung, U., **Herpin, A**., and Schartl M. (2007). Expression of the male determining gene dmrt1bY and its autosomal coorthologue dmrt1a in Medaka, ***Sex Dev.*** *3, 197-206. (IF=1.000)*
14. **Herpin, A.**, Schindler, D., Kraiss, A., Hornung, U., Winkler, C. and Schartl, M. (2007) Inhibition of germ cell proliferation by the medaka male determining gene Dmrt1bY, ***BMC Dev Biol.*** *7*, 99. *(IF=3.337)* (Highly accessed).
15. **Herpin, A.**, Fischer, P., Liedtke, D., Kluever, N., Neuner, C., Raz, E. and Schartl, M. (2008) Sequential SDF1a and b – induced mobility guides Medaka PGC migration, ***Dev Biol.*** *320, 319-327. (IF=4.416)*

*** Cover picture of the issue**.

16. **Herpin, A.**, and Schartl, M. (2008) Regulatory putsches make new ways of determining sexual development, ***EMBO reports.*** *9(10) 966-8. (IF=7.099)*
17. Kluever, N., **Herpin, A**., Braasch, I., Driessle, J. and Schartl, M. (2009) Regulatory back-up circuit of Medaka Wt1 co-orthologs ensures PGC maintenance, ***Dev Biol.*** *325(1):179-88. (IF=4.379)*
18. **Herpin, A.**, Nakamura, S., Wagner, T., Tanaka, M. and Schartl, M. (2009) A highly conserved mRNA sequence motif directs differential gonadal regulation of mRNA during gonad development, ***Nucleic Acids Research.*** *37(5):1510-1520.* *(IF=7.479)*
19. Le Quere, H., **Herpin, A**., Huvet, A., Lelong, C. and Favrel, P. (2009) Structural and functional characterization of an Activin type II receptor from the pacific oyster Crassostrea gigas, ***Gene.*** 436, 1-2. *(IF=2.416)*
20. **Herpin, A.** and Schartl, M. (2009) Molecular mechanisms of sex determination and evolution of the Y chromosome: insights from the medakafish (*Oryzias latipes*), ***Molecular and Cellular Endocrinology*** *306(1-2):51-58.* *(IF=3.503)*
21. **Herpin, A.**, Braasch, I., Kraeussling, M., Schmidt, C., Thoma, E. C., Nakamura, S., Tanaka, M. and Schartl, M. (2010) Transcriptional rewiring of the sex determining *dmrt1* gene duplicate by transposable elements, ***PLoS Genetics***. *6(2).* *(IF=9.543)*
22. Research highlights: Male regulator switched (2010), ***Nature***. *463 (1003), doi:10.1038/4631003c. (IF=36.101)*
23. **Herpin, A**. and Schartl, M. (2011) Vertebrate sex determination: challenging the hierarchy, ***FEBS journal***. *278(7):1001.* *(IF=3.790)* *** Cover picture of the issue**.
24. **Herpin, A.** and Schartl, M. (2011) Dmrt1 at the crossroad: a widespread and central class of sexual development factors in fish, ***FEBS journal***. *278(7):1010-1019. (IF=3.790)* *** Top-Cited Paper Award.**
25. Thoma, E. C., Wagner, T. U., Weber, I. P., **Herpin, A.**, Fischer, A. and Schartl, M. (2011) Ectopic expression of transcription factors enables directed differentiation of a Medaka spermatogonial cell line, ***Stem Cells Dev***. *20(8): 1425-1438. (IF=4.459)*

26. He, X., Yan, Y-L., Eberhart, J. K., **Herpin, A**., Wagner, T. U., Schartl, M., and Postlethwait, J. H. (2011) miR-196 regulates axial patterning and pectoral appendage initiation, ***Dev Biol.*** 357(2): 463-477. *(IF=4.069)*

27. **Herpin, A**. and Schartl, M. (2011) Sex determination: switch and suppress, ***Curr Biol.*** *21(17): R656-659.* *(IF=9.647)*

28. **Herpin, A.**, Adolfi, M., Nicol, B., Hinzmann, M., Schmidt, C., Klughammer, J., Engel, M., Tanaka, M., Guiguen, Y. and Schartl, M. (2013) Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination, ***Mol Biol Evol.*** *30(10): 2328-46.* *(IF=14.308)*

29. Nishimura, T., **Herpin, A.**, Kimura, T., Hara, I., Kawasaki, T., Nakamura, S., Yamamoto, Y., Saito, T. L., Yoshimura, J., Morishita, S., Tsukahara, T., Kobayashi, S., Naruse, K., Shigenobu, S., Sakai, N., Schartl, M. and Tanaka, M. (2013) Cell autonomous sexual identity of germ cells by sex chromosomes prior to the gonadal formation in medaka, ***Development.****141(17):3363-9.* *(IF=6.273)*

30. **Herpin, A.**, Englberger, E., Zehner, M., Wacker, R., Gessler, M. and Schartl, M. (2015) Defective autophagy through *epg5* mutation results in germ plasm and mitochondria reduction failure during spermatogenesis, ***FASEB J.*** *29(10)4145-61.* *(IF=5.299)*

31. **Herpin, A.**, Schartl, M. (2015) Plasticity of the gene regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers and usurpators, ***EMBO reports.*** *16(10):1260-74.* *(IF=7.739)*

32. Zhang, X., Guan, G., Li, M., Zhu, F., Liu, Q., Naruse, K., **Herpin, A.**, Li, J., Nagahama, Y. and Hong, Y. (2016) Autosomal *gsdf* acts as a male sex initiator in the fish medaka, ***Sci Rep.*** *6, 19738.* *(IF=4.259)*

33. Pan, Q., Anderson, J., Bertho, S., **Herpin, A.**, Wilson, C., Postlethwait, J.H., Schartl, M., and Guiguen, Y. (2016) Vertebrate sex-determining genes play musical chairs, ***C R Biol*** 339, 258-262. *(IF=0.731)*

34. Bertho, S., Pasquier, J., Pan, Q., Le Trionnaire, G., Bobe, J., Postlethwait, J.H., Pailhoux, E., Schartl, M., **Herpin, A.**, and Guiguen, Y. (2016) Foxl2 and Its Relatives Are Evolutionary Conserved Players in Gonadal Sex Differentiation, ***Sex Dev*** *10, 111-129.* *(IF=1.972)*

35. Adolfi, M.C., **Herpin, A.**, Regensburger, M., Sacquegno, J., Waxman, J.S., and Schartl, M. (2016) Retinoic acid and meiosis induction in adult versus embryonic gonads of medaka. ***Sci Rep*** *6, 34281.* *(IF=4.259)*

36. Schartl, M., Schories, S., Wakamatsu, Y., Nagao, Y., Hashimoto, H., Bertin, C., Mourot, B., Schmidt, C., Wilhelm, D., Centanin, L., Guiguen, Y. and **Herpin, A**. (2018) Sox5 is involved in germ cell regulation and sex determination in medaka following co-option of nested transposable elements, ***BMC Biology*** *16(1):16.*

Included into the cross journal collection "In the Light of Evolution" (BMC Biology, BMC Evolutionary Biology and Biology Direct). *(IF=6.723)*

37. (**Herpin, A.**, Lescat, L.), Mourot, B., Véron, V., Guiguen, Y., Bobe, J. and Seiliez, I. (2018) CMA restricted to mammals and birds: myth or reality? ***Autophagy*** *20, 1-4.* *(IF=11.059)*

38. Regneri, J., Klotz, B., Wilde, B., Kottler, V., Hausmann, M., Kneitz, S., Regensburger, M., Maurus, K., Götz, R., Lu, Y., Walter, R., **Herpin, A**. and Schartl, M. (2018) Analysis of the putative tumor suppressor gene *cdkn2ab* in pigment cells and melanoma of *Xiphophorus* and medaka, ***Pigment cell and melanoma research*** *doi: 10.1111/pcmr.12729* *(IF=4,172)*

39. Bertho, S., **Herpin, A.**, Jouanno, E., Yano, A., Nicol, B., Müller, T., Pannetier, M., Pailhoux, E., Miwa, M., Yoshizaki, G., Schartl, M. and Guiguen, Y. (2018) The unusual rainbow trout sex determination gene hijacked the canonical vertebrate gonadal differentiation pathway, ***PNAS*** *115(50):12781-12786.* *(IF=9,58)*

40. **Herpin, A.**, Schmidt, C., Kneitz, S., Gobé, C., Regensburger, M., Le Cam, A., Montfort, J., Adolfi, M.C., Lillesaar, C., Wilhelm, D., Kraeussling, M., Mourot, B., Porcon, B., Pannetier, M., Pailhoux, E., Ettwiller, L., Dolle, D., Guiguen, Y. and Schartl, M. (2019) A novel evolutionary conserved mechanism of RNA stability regulates synexpression of primordial germ cell-specific genes prior to the sex-determination stage in medaka, ***PLoS Biology*** *Apr 4;17(4):e3000185* *(IF=7,076)*

41. Le Trionnaire, G., Tanguy, S., Hudaverdian, F., Gleonnec, F., Richard, G., Cayrol, B., Monsion, B., Pichon, E., Deshoux, M., Webster, C., Uzest, M., (Tagu, D. and **Herpin, A.**) (2019) An integrated protocol for targeted mutagenesis with CRISPR-Cas9 system in the pea aphid, ***Insect Biochemistry and Molecular Biology*** *Jul; 110:34-44; doi: 10.1016/j.ibmb.2019.04.016.* *(IF=3,827)*

42. Adolfi, M. C., Fisher, P., **Herpin, A.**, Regensburger, M., Kikuchi, M., Tanaka, M. and Schartl, M. (2019) Increase of cortisol levels after temperature stress activates dmrt1a causing female-to-male sex reversal and reduced germ cell number in medaka, ***Molecular Reproduction and Development*** *Oct; 86(10):1405-1417;* doi: 10.1002/mrd.23177. *(IF=2,823)*

43. Pan, Q., Feron, R., Yano, A., Guyomard, R., Jouanno, E., Vigouroux, E., Wen, M., Busnel, JM., Bobe, J., Concordet, JP., Parinello, H., Journot, L., Klopp, C., Luuch, J., Roques, C., Postlethwait, J., Schartl, M., **Herpin, A.** and Guiguen, Y. (2019) Identification of the master sex determining gene in the Northern pike (*Esox Lucius*) reveals restricted sex chromosome differentiation, ***PLoS genetics*** *Aug 22;15(8):e1008013.* *(IF=5.174)*

44. Feron, R., Zahm, M., Cabau, C., Klopp, C., Roques, C., Bouchez, O., Eché, C., Valière, S., Donnadieu, C., Haffray, P., Bestin, A., Morvezen, R., Acloque, H., Euclide, P., Wen, M., Jouano, E., Schartl, M., Postlethwait, J., Schraidt, C., Christie, M., Larson, W., **Herpin, A.** and Guiguen, Y. (2020) Characterization of a Y-specific duplication/insertion of the anti-Müllerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, *Perca flavescens*, ***Molecular Ecology Resources*** *doi: 10.1111/1755-0998.13133.* *(2017 IF: 7,049)*

45. Takehana, Y., Zahm, M., Cabau, C., Klopp, C., Roques, C., Bouchez, O., Donnadieu, C., Barrachina, C., Journot, L., Kawaguchi, M., Yasumasu, S., Ansai, S., Naruse, K., Inoue, K., Shinzato, C., Schartl, M., Guiguen, Y. and **Herpin, A.** (2020) Genome sequence of the euryhaline Javafish medaka, Oryzias javanicus: a small aquarium fish model for studies on adaptation to salinity, ***G3 Genes Genomes Genetics*** *doi: 10.1534/g3.119.400725.* *(2017 IF: 2,742)*

46. Lescat, L., Véron, V., Mourot, B., Dias, K., Riera, N., Beaumatin, F., Panserat, S., Guiguen, Y., Bobe, J., (**Herpin, A.** and Seiliez, I.) (2020) Chaperon-Mediated Autophagy in the light of evolution: insight from fish, ***Molecular Biology and Evolution***, https://doi.org/10.1093/molbev/msaa127. *(2018 IF: 14,797)*

47. **Herpin, A.**, Lescat, L., Bobe, J., Jenny, A. and Seiliez, I. (2020) Lighting Chaperone-Mediated_Autophagy (CMA)

evolution with an ancient LAMP: the existence of a functional CMA activity in fish, ***Autophagy*** *Aug 10:1-3. doi: 10.1080/15548627.2020.1797344. (2018 IF: 11,059)*

48. Wen, M., Feron, R., Pan, Q., Guguin, J., Jouanno, E., **Herpin, A.**, Klopp, C., Cabau, C., Zahm, M., Parrinello, H., Journot, L., Burgess, S., Omori, Y., Postlethwait, J., Schartl, M. and Guiguen, Y. (2020) Sex chromosome and sex locus characterization in goldfish, *Carassius auratus* (Linnaeus, 1758), ***BMC genomics***, *Aug 11;21(1):552. doi: 10.1186/s12864-020-06959-3.(2017 IF: 3,730)*

49. (**Herpin, A**., Rafati, N., Chen, J.), Petterson, M., Han, F., Feng, C., Wallerman, O., Rubin, CJ., Péron, S., Cocco, A., Larsson, M., Trötschel, C., Poetsch, A., Korsching, K., Bönigk, W., Körschen, HG., Berg, F., Folkvord, A., Kaupp, UB., Schartl, M. and Andersson, L. (2020) Reconstruction of the birth of an unexpected male sex chromosome in Atlantic herring, ***PNAS*** *(2018 IF=9,58)*

50. Imarazene, B., Beille, S., Jouanno, E., Branthonne, A., Thermes, V., Thomas, M., **Herpin, A.**, Retaux, S., Guiguen, Y. (2020) Histoligical and molecular characterization of gonadal differentiation in Pachon cavefish Astyanax mexicanus, ***Sex. Dev.*** *(2018 IF= 2.25)*

51. Pan, Q., Kay, T., Depincé, A., Adolfi, M., Schartl, M., Guiguen, Y. and **Herpin, A.** (2021) Evolution of master sex determiners: TGF-β signalling pathways at the crossroads, ***Phil. Trans. R. Soc. B.*** *(2018 IF= 6.139)*.

(submitted, under review, in revision)

1. Klionsky, D., (…), **A., Herpin**, (…) and Z., Zhang. (2020) Guidelines for the use and interpretation of assays for monitoring Autophagy, (in revision in ***Autophagy***, *(2018 IF: 11,059)*)

2. Bertho, S., **Herpin, A.**, Schartl, M. and Guiguen, Y. (2020) Lessons from an unusual vertebrate sex-determining gene, (in revision, ***Phil. Trans. R. Soc. B.*** (2018 IF= 6.139)).

3. Adolfi, M., **Herpin, A.** and Schartl, M. (2020) The replaceable master of sex determination: bottom-up hypothesis revisited, (in revision, ***Phil. Trans. R. Soc. B.*** (2018 IF= 6.139)).

4. Pan, Q., Feron, R., Jouanno, E., Darras, H., **Herpin, A.**, Koop, B., Rondeau, E., (…) Schartl, M., Postlethwait, J. and Guiguen, Y. (2020) The rise and fall of the ancient northern pike master sex determining gene, (under review, ***E-life*** (2018 IF= 7.08)).

5. Feron, R., Pan, Q., Wen, M., Imarazene, B., Jouanno, E., Anderson, J., Herpin, A., Journot, L., (…) Postlethwait, J., Schartl, M. and Guiguen, Y. (2020) RADSex: a computational workflow to study sex determination using Restriction Site-Associated DNA Sequencing data, (In revision, **Molecular Ecology Resources** (2018 IF=6.286)).

6. **Herpin, A.**, Schartl, M., Depincé, A., Guiguen, Y., Bobe, J., Hua-Van, A., Hayman, E., Octavera, A., Yoshizaki, G., Nichols, K., Goetz, G. and Luckenbach, A. (2020) Allelic diversification after transposable element exaptation into its proximal regulatory region promoted Gsdf as the master sex determining gene of sablefish *Anoplopoma fimbria*, (to be submitted to Genome Biology)

Publications in 32 different journals:

Autophagy, BMC Biology, BMC Developmental Biology, *BMC Genomics, Biochim Biophys Acta, Comptes Rendus Biologie*, *Current Biology, Development, Developmental and Comparative Immunology, Development Genes and Evolution, Developmental Biology, EMBO reports, FASEB Journal, FEBS Journal, G3 Genes Genomes Genetics, Gene, Gene Expression Patterns, Insect Biochemistry and Molecular Biology, Journal of Biological Chemistry, Mechanisms of Development, Molecular Biology and Evolution, Molecular and Cellular Endocrinology, Molecular Reproduction and Development, Molecular Ecology Resources, Nature, Nucleic Acids Research, Phylosophical Transactions of the Royal Society B, PLoS Biology, PLoS Genetics, Pigment cell and melanoma research, PNAS*, *Sexual Development, Scientific Reports, Stem Cells Dev.*

Book Chapters

1. Guiguen, Y., Fostier, A., and **Herpin, A.** (2017) **Book chapter**: Sex determination and differentiation in fish: genetic, genomic and endocrine aspects in "Sex control in aquaculture" (Wang, H.P., F. Piferrer & S.L Chen Ed.). ISBN: 978-1-119-12726-0. Wiley-Blackwell. DOI:10.1002/9781119127291. (cover picture)
2. Guiguen, Y., Bertho, S., **Herpin, A.**, and Fostier, A. (2017) **Book chapter**: Sex determination and control in salmonidae in "Sex control in aquaculture" (Wang, H.P., F. Piferrer & S.L Chen Ed.). ISBN: 978-1-119-12726-0. Wiley-Blackwell. DOI:10.1002/9781119127291.
3. Quiaowei, P., Guiguen, Y., and **Herpin, A.** (2017) **Book chapter:** Evolution of sex determining genes in fish. Encyclopedia of Reproduction, Second Edition. Elsevier.
4. Schartl, M. and **Herpin, A.** (2017) **Book chapter:** Sex determination systems and sex determiners in vertebrates. Encyclopedia of Reproduction, Second Edition. Elsevier.
5. Bertin, C. and **Herpin, A.** (2018) **Book chapter:** "Analyse de la fonction d'un gène par mutagénèse ou transgénèse chez les poisons" dans "Principes des techniques de biologie moléculaire et génomique". Editions QUAE. ISBN 978-2-7592-2885-0.

Cover Pictures & Highlights

I- Experience and Current Research Work

1- Past and Current Scientific Research Career

I have received my undergraduate degree in cellular biology from the University of Caen (France, Laboratory of Biology and Marine Biotechnologies / IFREMER), working on the early development of one of the world's most economically important bivalve species for aquaculture: The Pacific oyster (*Crassostrea gigas*).

To acquire a broader expertise, and as a follow up of this work, I have then joined the Sars International Centre for Marine Molecular Biology / EMBL in Bergen (Norway) where I have obtained my PhD degree (jointly between the Universities of Caen and Bergen) in 2003. Intending at gaining insight into the mechanisms triggering early development, reproduction, fertility and immunity, the subject of my PhD was dealing with the characterization of the morphogenetic roles of oyster TGF-β superfamily receptors during early development as well as their involvement as cytokine receptors. To circumvent experimental limitations inherent to the oyster model, for functional analyses, I have set up and used the zebrafish as a "reporter organism" **[see papers #1 to #4, #6 to #8, #10 to #12 and #19]**.

Thanks to the experience gathered with fish model organisms (zebrafish, salmon, hagfish) during my PhD, I have then taken the opportunity to join the group of Pr. M. Schartl at the University of Wuerzburg (Germany, Physiological Chemistry), one of the leading teams working on fish sex determination. There, over more than ten years, I have been studying the different aspects of molecular sex determination and gonad formation and differentiation using the medaka fish (*Oryzias latipes*) as a main model organism **[see papers #5, #9, #13 to #15, #16 and #17, #21, #25, #28 and #29, and #32]**. During that period, I have settle down my thematic and established my own group. I then got interest in studying, on a larger scale, the evolution and plasticity of the underlying gene regulatory networks controlling and regulating such mechanisms in metazoans. Using functional genomic approaches, this also brought us to face, challenge and discuss the current evolutionary concepts for the diversity of sex determining mechanisms and evolution of sex chromosomes **[see papers #16, #22 to #24, #27 and #28 and #31]**.

Being willing to pursue my career on that specific thematic for which I gained a kind of expertise in the field, in 2015 I have joined as a full-time researcher the INRA Fish Physiology and Genomics Laboratory (France, UR1037, LPGP, PHASE). There, in line with my past scientific background and still keeping the same "editorial policy", I have developed cutting edge techniques of genome editing (*CRISPR-engineered KO and KI, conditional KI...*) and other genomic (*high throughput genome sequencing, in-vivo ChIP-seq*) and functional genomic (*cell lineage tracing, topological single cell transcriptomiques*) approaches in fish to better tackle our current/long standing concerns **[see papers #33 to #49]**.

In the following I will first describe the current concerns in the vertebrate sex determination field, and how I have been trying to address them as a guiding thread for my research using teleost fish. Then, I will present the context and positioning of my research at INRAE and discuss the main results. Finally, after a general discussion, I will give some perspectives to illustrate the way forward I would like to conduct my research in the future.

2- Sex Determination in Vertebrates: Current Concerns

Sexual dimorphism is one of the most pervasive and diverse features of animal morphology, physiology, and behaviour. Despite the generality of the phenomenon itself, the mechanisms how sex is determined are very different among various organismic groups, have evolved repeatedly and independently across metazoans, and the underlying molecular pathways can change quickly during evolution **[see papers #23, #27, #31 and BC#1 to BC#4]**. Even within closely related groups of organisms for which the development of gonads on the morphological, histological and cell biological level is undistinguishable, the molecular control and the regulation of the factors involved in sex determination and gonad differentiation can be substantially different **[see paper #28]**. The biological meaning of the high molecular plasticity of an otherwise common developmental program is unknown. While phylogenetic studies suggest that the downstream "business" ends of sex-determining pathways do tend to be more stable than the triggering mechanisms at the top (1, 2), how similar are the downstream gene networks that respond to these triggers and actually do the work? In other words, did sex determination (SD) arise several times with different regulatory mechanisms slowly converging by recruiting conserved key effectors, or is it an ancient process with little surviving evidence of ancestral genes? This question becomes truly meaningful in the light of the recent identification of related sexual regulators in different phyla indicating that at least some aspects of the sexual core regulation might be anciently co-opted or preserved **[see papers #16, #24, #27, #31].** In the recent years, however, some common themes in the development of sex-specific traits in different animal lineages have started to emerge. Central to this stream has been the discovery of the DMRT1 factors, one of the most ancestral and widespread family of transcription factors involved in the molecular sex-determination cascades over phyla (3).

In stark contrast to birds and mammals, where the whole group has the same sex determination mechanisms with only enigmatic exceptions, but similar to the situation in amphibians and reptiles, the diversity of sex determination mechanisms is especially obvious in fish, where within groups of closely related species a wide spectrum of different systems can be found (4). Within teleosts, for instance in the poeciliid fish, which include guppies, mollies, platyfish and swordtails, there are reports on temperature-dependent SD (TSD) and various forms of genetic sex determination (GSD), ranging from polyfactorial SD to female and male heterogamety, multiple sex chromosomes and autosomal modifiers. Even within the same species several SD mechanisms can occur (5). The coexistence of two or more of these systems has also been reported within the same genus, like for instance, the XY/XX and ZW/ZZ genotypes that are found in different *Oreochromis* species (6) or even within the same species like in some platyfish populations (5).

3- Research Strategy: Animal Model and Scientific Value

For approaching numbers of questions dealing with sex determination / differentiation from either comparative evolutionary or functional point of views, I have mainly used the medaka fish, *Oryzias latipes,* as a model species. The medaka is a small freshwater fish species that lives in the small rivers and rice fields of East Asia. During the last decade, it became a widely used laboratory fish for developmental and biomedical research, comparable to the well-known zebrafish with which the medaka shares many advantages and characteristics. Medaka offers valuable genomic resources and the latest most advanced transgenic technologies, including genome editing, that are available for functional studies (7). With respect to sex determination, and in contrast to zebrafish, for which sex determination genetics and mechanisms still remain obscure, the medaka has a XX/XY genetic sex determination system with homomorphic sex chromosomes and the

master male sex determination gene (*dmrt1bY*) is known and has been characterized (8, 9). Interestingly this *dmrt1bY* gene is however absent from closely related species (10). *De facto* relying on different master sex determining genes, these evolutionary switches offer interesting opportunities for studying how transitions among sex-determining mechanisms can occur in closely related species of the same genus (Figure 1)…

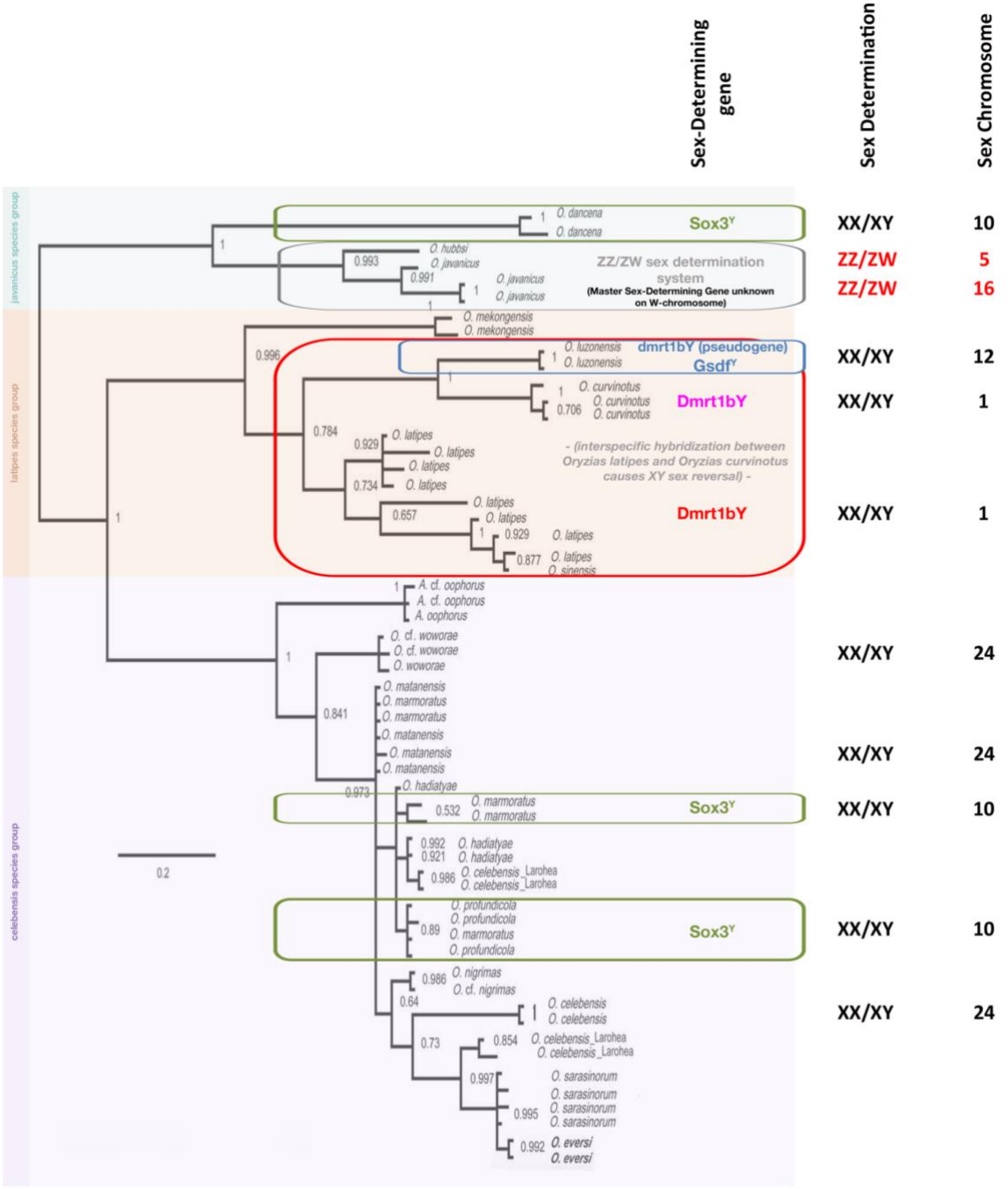


Figure 1. Phylogenetic relationships and sex determination systems in *Oryzias* fish species.
Medaka fish (*Oryzias* genus) show an amazing diversity regarding to their sex determination systems and sex chromosomes, providing an excellent model group for investigating the molecular mechanisms underlying the rapid turnover of sex chromosomes. They possess both XY and ZW systems, and their sex chromosomes differ from one species to another (11, 12).

…or in fish in general (Figure 2).

	Organism	Master Sex-Determining Gene	Sex-Determination System	Ancestral MSD gene	Establishment
	Therian mammals	SRY	*XY*	SOX3	*Allelic Diversification*
	Monotremes / *Platypus* (*Ornithhorhynchus anatinus*)	**AMH**	*Multiple different X and Y chromosomes*	**AMH**	*Chromosome fusion and allelic diversification*
	Birds	DMRT1	*ZW*	*Dmrt1*	*Allelic Diversification*
	African clawed frog (*Xenopus laevis*)	*DM-W*	*ZW*	*dmrt1*	*Gene Duplication (& truncation)*
Teleost fish	Atlantic herring (*Clupea harengus*)	***bmpr1bbY***	XY	*bmpr1bb*	Gene Duplication (& truncation)
	Characiformes species	***gdf6-bY***	XY	*gdf6*	Gene Duplication
	European pike (*Esox lucius*)	***amhbY***	XY	*amh*	Gene Duplication
	Rainbow trout (*Oncorhynchus mykiss*)	*sdY*	XY	*irf9*	Gene Duplication
	Seriola fishes	*hsd17b1*	ZW	*hsd17b1*	Allelic Diversification
	Chinese tongue sole (*Cynoglossus semilaevis*)	*dmrt1*	ZW	*dmrt1*	Allelic Diversification
	Nile tilapia (*Oreochromis niloticus*)	***amhY***	XY	*amh*	Gene Duplication
	Medaka (*Oryzias latipes*)	*dmrt1bY*	XY	*dmrt1*	Gene Duplication
	Luzon rice fish (*Oryzias luzonensis*)	***gsdf-Y***	XY	*gsdf*	Allelic diversification
	Indian rice fish (*Oryzias dancena*)	*sox3-Y*	XY	*sox3*	Allelic Diversification
	Pejerrey (*Odontesthes hatcheri*) and other *Odontesthes* species	***amhY***	XY	*amh*	Gene Duplication
	Turquoise Killifish (*Nothobranchius furzeri*)	***gdf6-Y***	XY	*gdf6*	Allelic Diversification
	Yellow perch (*Perca flavescens*)	***amhr2bY***	XY	*amhr2*	Gene Duplication (& truncation)
	Stickleback (*Gasterosterus aculeatus*)	***amhY***	XY	*amh*	Gene Duplication
	Lingcod (*Ophiodon elongatus*)	***amhbY***	XY	*amh*	Gene Duplication
	Sablefish (*Anoplopoma fimbria*)	***gsdf-Y***	XY	*gsdf*	Allelic Diversification
	Torafugu (*Takifugu rubripes*)	***amhr2-Y***	XY	*amhR2*	Allelic Diversification

Figure 2. Currently known or suspected master sex determining genes and sex determination systems in vertebrates.

Currently known and documented master sex determining (MSD) genes in vertebrates. Of special interest, MSD genes belonging to the TGF-β are highlighted in red (see perspectives). MSD genes that arose after allelic diversification or gene duplication are highlighted in green and blue respectively.

In that particular context of functional and comparative evolution and using the medaka fish as a model organism of reference, my past and current research projects on sex determination / sex differentiation in fish are articulated around two main general questions and subsequent precise research objectives:

I. **How stable are SD mechanisms in evolutionary terms? And what is the evolutionary meaning of the high variability of sex determination mechanisms?**

On these aspects, I have focused my research mainly on the sex determination function of *dmrt1bY* in medaka and also on the comparative functional analysis of other master sex-determining genes (*SdY*, *Amh*, *AmhR2*, *BMPR1*, *Gsdf*, *Gdf6*…) in different fish (*Rainbow trout*, *Northern pike*, *Killifish*, *Atlantic Herring*, *Goldfish*, *Java medaka*, *Yellow perch*, *Cave fish*…)

II. **Are different SD mechanisms triggering the same, related or totally specific molecular pathways during the process of the fate differentiation of the gonad anlage?**

To address that precise point, I have directed my research toward the characterization of the gene regulatory network triggering sex differentiation with further interest on the germ line/soma interactions and gonadogenesis in medaka and in fish in general.

Being the guidelines for conducting my research, my contributions in addressing these two major questions and objectives in the field of sex determination are further developed in the following.

3.1- Sex determination evolution and the central role of the dmrt1 gene in fish.

The occurrence of two different sexes and consequently the necessity to make a developmental decision for an embryo to become male or female (the so-called sex-determination process) and the further differentiation of the whole organism into two distinct phenotypes, are common throughout the animal, plant and fungi kingdoms. While developmental cascades are generally headed by highly conserved universal master regulators that univocally determine the developmental fate of a cell lineage to a tissue or organ during embryogenesis, all the evidence suggests that sex determination might disobey the conventional rules of evolutionary conservation **[see papers #16, #22 and #23, #27 and #28]**. Hence, decades of genetic studies have led to the global picture that the gene-regulatory cascades triggering sexual differentiation from worms and *Drosophila* to mammals bear little resemblance to each other. The remarkable diversity of 'master sex-determining genes' at the top of the genetic hierarchies now seems obvious, while downstream components surprisingly appear to be evolutionarily more conserved and tend to converge upon the regulation of few central common effectors **[see paper #16]**. Hence, a comparative view on genetic sex determination mechanisms led to the paradigm that "master change, slaves remain" (2). A well-known example illustrating this paradigm is the SRY gene – the master sex-determining gene of mammals- that has not been detected outside of the placental mammals. Conversely SRY subordinated genes (Sox9 **[see paper #28]**, Wt1 **[see paper #17]**, Dmrt1, AMH, SF1, FoxL2…) or signalling pathways (TGF-β, Wnt4/β-catenin, Hedgehog **[see paper #28]**) have homologues involved in gonadogenesis or gonadal differentiation in a much broader spectrum of species, including non-vertebrates and even protostomes lacking Sry (see Figure 3).

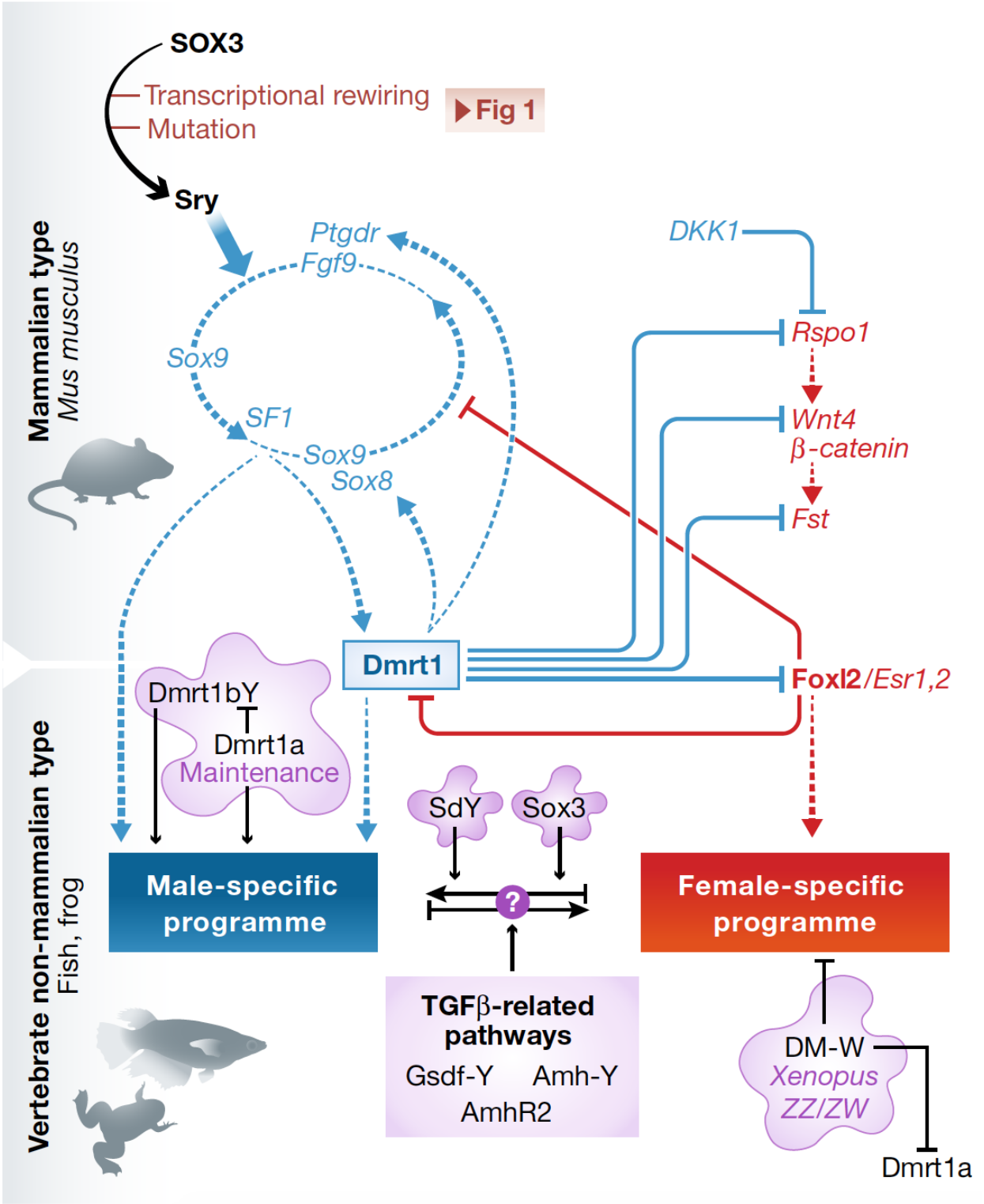


Figure 3. Gene-regulatory network of gonadal sex induction and maintenance in vertebrates.

Schematic representation of main interactions within the regulatory network. In gonadal fate determination of mammals, Sry initiates activation of the male pathway (blue) through up-regulation of Sox9. Dmrt1 is not only important for keeping the male pathway on but also in suppressing the two female networks (red). These two female networks involve Foxl2 as well as the Wnt/β-catenin signalling pathways. Maintenance of gonadal identity in the differentiated gonads is a result of the cross-inhibition activities of Dmrt1 and Foxl2. A critical equilibrium between these conflicting pathways underlies the bi-potentiality of the gonadal somatic cells. Tipping the balance into one direction or the other will regulate the gonadal fate as a consequence of the activation of the male or female pathways. Solid lines define negative regulations. Dashed lines designate positive regulations. Beside the Sry ancestor Sox3 and Dmrt1, other genes (pink) can become the master sex-determining genes by similarly impacting on the seesaw between the male and female programme **[see paper #31]**.

In contrast to this relatively good conservation at the bottom of the sex determination cascade, examples show that the master regulator is not necessarily elected -co-opted-, like SRY in mammals, from the outside to perform a new sex-determining function, but rather frequently recruited as a member of the existing cascade, usurping the position at the top **[see paper #16]**. Indeed, independent duplication events **[see papers #16]** or dosage effect led, for example, Dmrt1, AMH/AMHR or GSDF genes to take over the leadership as master sex determining gene in Medaka, *Xenopus*, Chicken or Pejerrey and *Fugu* or *Oryzias luzonensis* respectively (**[see paper #20]** and Figure 2). Beyond vertebrates, and definitively illustrating an evolutionary convergence driving that scheme, a similar duplication event was also described in the honeybee (3) (Figure 4, **[see paper #16]**).

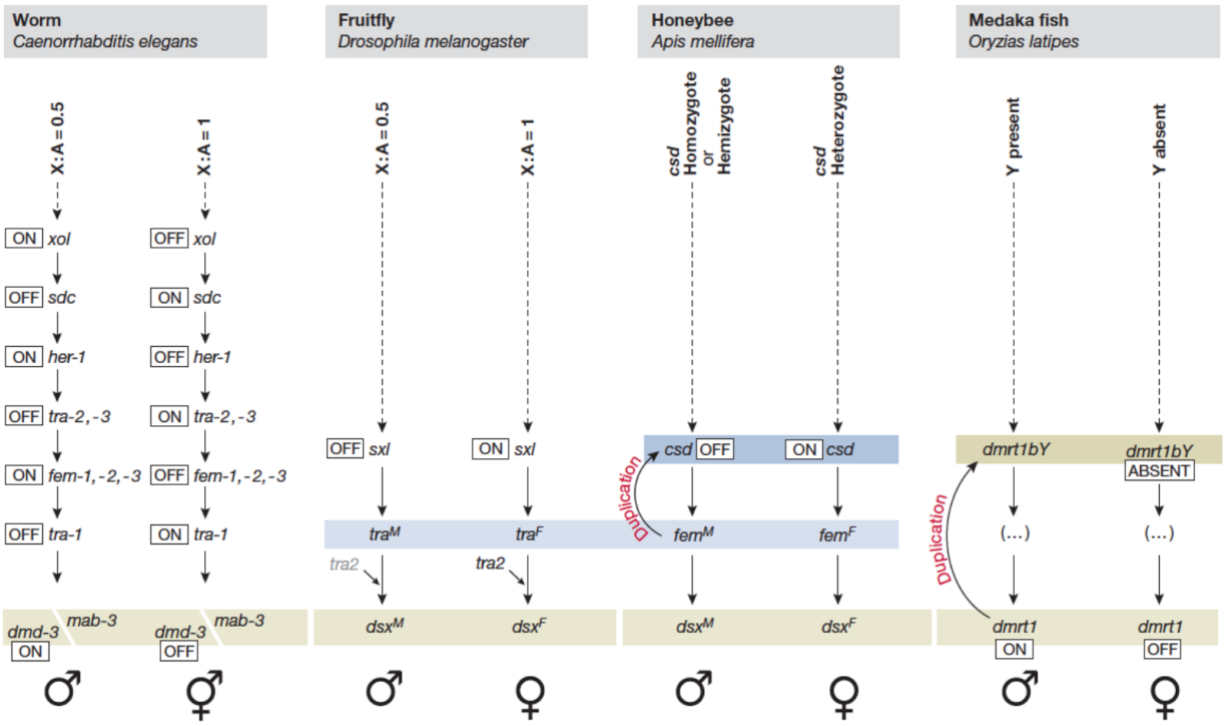


Figure 4. Sex-determining cascades in the worm, insects and medaka fish.

Molecular pathways leading to the formation of the gonad in *Caenorhabditis elegans*, *Drosophila melanogaster*, *Apis mellifera* and *Oryzias latipes*. Conserved doublesex/male abnormal 3/dsx and mab-3 related transcription factor 1 (dmrt1) or Drosophila transformer (tra)-like homologues are indicated with light-brown and light-blue boxes, respectively; duplicated factors in the cascades are also highlighted (brown and blue boxes, respectively). Note that tra of C. elegans is not related to tra of Drosophila. csd, complementary sex determiner; dmd3, doublesex/mab-3 domain family member 3; dsx, doublesex; fem, feminizer; her-1, hermaphroditization of XO-1; sdc, sex determination and dosage compensation defective; sx1, sex lethal; xol, XO lethal. **[see paper #31]**.

Additionally, it is now emerging that the primary sex-determining decision is not final but has to be constantly affirmed life-long by suppressing the opposing sex determination programs, thereby challenging our basic notions of the function and evolution of the sex-determining pathways and later on of the gonadal maintenance **[see papers #26 and #28]**. Interestingly, in addition to the primary function as sex determining gene in the above-mentioned species –and probably in others as well- Dmrt1 factors have been pointed out to have an even more universal evolutionary widespread role in male gonadal maintenance (Figures 3 and 4) **[see papers #22, #23 and #26]**.

Among the few highly evolutionary conserved downstream effector genes of the genetic sex determining cascades, a gene family involved in sex differentiation in organisms as phylogenetically divergent as *Caenorhabditis elegans*, *Drosophila*, frogs, fish, birds, mammals and corals is the *dmrt* gene family **[see papers #22 and #23]**. The Dmrt group of molecules is characterized by a conserved DNA-binding motif known as the DM (Doublesex and Mab-3 related) domain (3). Dmrts were originally described to play important roles during sex determination in flies and worms by regulating several aspects of somatic sexual dimorphism (13). They were also reported to be able to substitute for each other across species, indicating that their function is possibly interchangeable and that sex determination in invertebrates might rely on conserved molecules. Interestingly, recent studies have linked DM domain genes to primary sexual differentiation in many other species including medaka fish, frog and chicken, providing valuable entry points for understanding the control of sexual differentiation and evolution of regulatory networks in diverse animals **[see papers #22 and #23]**. But certainly, the deep interest of Dmrt1 in the field of sex determination in fish first came with the discovery -in my former laboratory in Germany- of a dmrt1 homologue on the Y-chromosome of the medaka fish: *dmrt1bY* (9). From then on, genetic evidences have suggested that the *dmrt1* gene is an important regulator of male development in vertebrates (**[see paper #24]** for review). In humans, haploinsufficiency of the genomic region that includes *DMRT1* and its paralogs *DMRT2* and *DMRT3* leads to XY male to female sex reversal (14). In chicken and other avian species *Dmrt1* is located on the Z chromosome, but absent from W, making it an excellent candidate for the male sex-determining gene of birds (15). Finally in frog it has been shown that DM-W, a W-linked truncated Dmrt1 homolog gene antagonizing the Dmrt1 transcriptional activity, participated in primary ovary development ((16) and **[see paper #23]**).

As mentioned above, the medaka male sex-determining gene has been identified in my former laboratory, making the medaka the first vertebrate species outside mammals where such gene has been described and characterized as being necessary and sufficient for directing testes development. With respect to its biochemical function, Dmrt1bY, but also the other Dmrt1s in fish, appear to act as transcription factors (Figure 5). This is evident from the nuclear localization of Dmrt1 fusion proteins **[see papers #13 and #14]** and studies showing direct effects of dmrt1 on reporter gene expression as well as binding to cognate motif in electric mobility shift assays **[see paper #21]**. Finally, linking the earliest sexual dimorphic trait to its expression dynamic, we could show Dmrt1bY to be possibly responsible for the male-specific primordial germ cell mitotic arrest **[see paper #9]** (Figure 5). Indeed, after cell transfection and FACS analyses, additional *in vivo* functional evidences showed that expression of Dmrt1bY leads to negative regulation of male primordial germ cell proliferation (G2/M arrest cell cycle) prior to sex differentiation at the sex determination stage **[see paper #9]**, suggesting that in XY medaka males, Dmrt1bY-driven primordial germ cell number regulation, as well as determination of pre-Sertoli cells, is the primary event by which the whole gonad (germ-line and soma) would be specified through a directional cross-talk from pre-Sertoli and Sertoli cells with the primordial germ cells **[see paper #9]** (Figure 5).

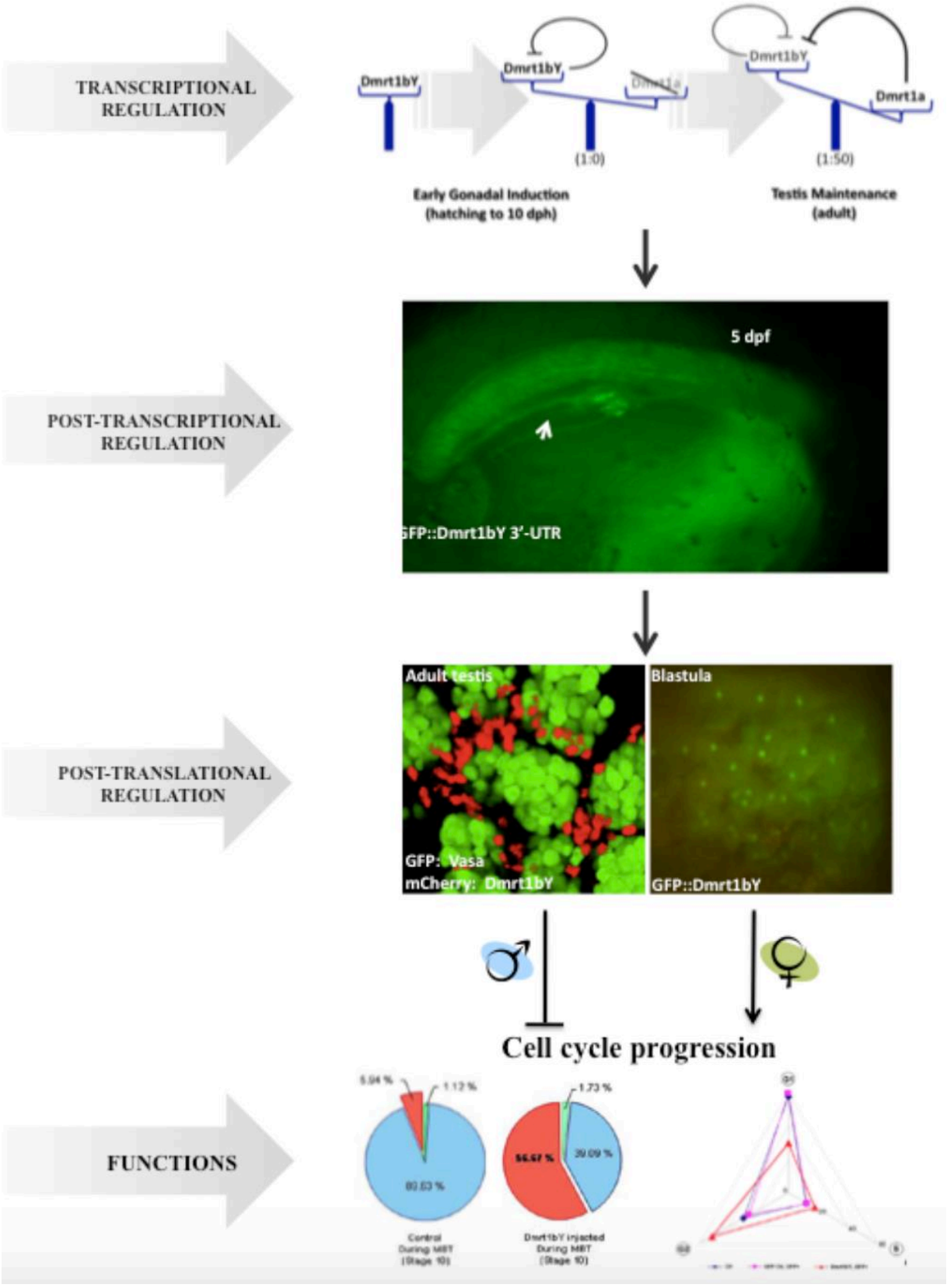


Figure 5. Medaka dmrt1a/dmrt1bY regulations and functions.

Grey arrows illustrate the different levels for which active dmrt1a/dmrt1bY regulation mechanisms could be shown. **Transcriptional regulation**: the feedback autoregulation of dmrt1bY promoter activity and transregulation by its paralogue Dmrt1a is a key mechanism of dmrt1bY transcriptional tuning. **Post-transcriptional regulation**: a highly conserved cis-regulatory motif directs differential gonadal synexpression of dmrt1 transcripts during gonadal development. **Post-translational regulation**: Dmrt1a and Dmrt1bY have a short half-life and consequently a high turnover. **Functions**: Dmrt1bY inhibition of germ cell proliferation might be part of its known male determining function. **[see papers #14, #18, #21, and #24]**.

3.2- Characterization of the molecular network regulating gonad differentiation in medaka.

Medaka gonad is formed by the coordinated development of two different cell lineages: the germ cells and the somatic gonadal mesoderm surrounding the germ cells. Using mRNA microinjection in one-cell staged embryos we were able to *in vivo* monitor medaka primordial germ cells (PGCs) specification and migration dynamics during gonadal formation **[see paper #9]**. Hence, confocal and electronic microscopic analyses revealed that PGCs can be first recognized as early as gastrulation stage. At that stage PGCs are aligned bilaterally along the trunk until somitogenesis completes and then finally migrate posteriorly in the ventrolateral region of the developing hindgut **[see paper #9]**. Together with the lateral plate mesoderm where they now lie, PGCs then move to the future gonadal region at the dorsal region of the hindgut where they combine to form a single undifferentiated gonadal primordium. Interestingly carrying on over-expression (capped mRNA) or transient knock-down (morpholino) experiments *via* microinjections we have additionally shown that the observed PGC migration dynamic was tightly regulated by cytokines interactions (Sdf1a/b and Cxcr4, (Figure 6)) **[see paper #15]** and that their maintenance was under the control of a regulatory back-up circuit mediated by Wt1 co-orthologs **[see paper #17]**. Shortly before hatching the germ cells of the female gonadal primordium actively proliferate and undergo meiosis while in male gonads, correlating with the male-specific onset of Dmrt1bY expression in the gonadal anlage, the PGCs remain quiescent **[see paper #14]**. Ten days later the first somatic gonadal dimorphisms are apparent with the formation of the acinus (the seminiferous tubule precursor) and the follicles in gonads of male and female respectively. Interestingly, during the process of gonadal differentiation in medaka, germ line soma interactions and cross talk have been shown to be of primary importance since germ cell depleted gonads develop as male while conversely an excessive number of germ cells lead to the development of a female-type gonad **[see paper #14]**.

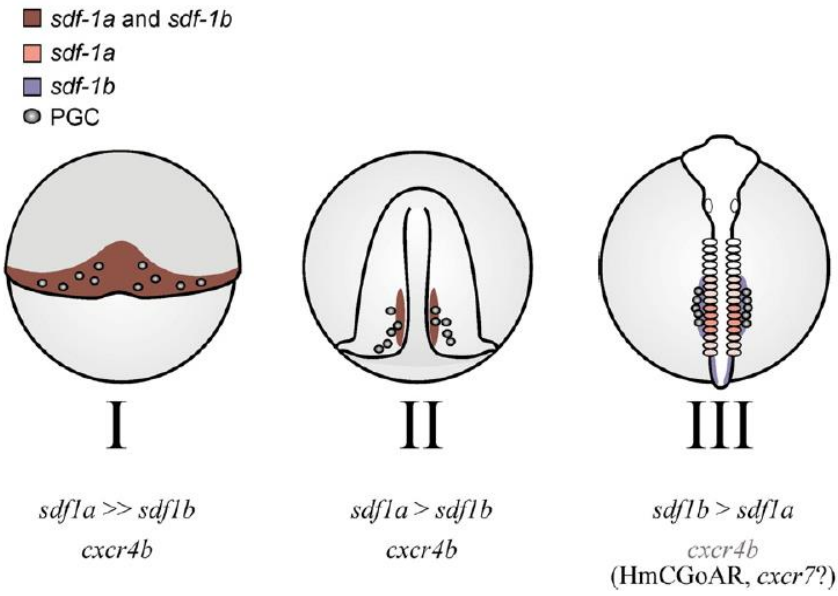


Figure 6. PGC migration in Medaka is the consequence of alternation of active OlaSDF1a/b-induced mobility.
During the process of early PGC migration (Step I), CXCR4 and Sdf1a are absolutely required for proper PGC migration. During neurulation/early somitogenesis (Step II), PGC migration is dependant of somatic movements, while arrest of PGCs in the lateral plate mesoderm along the embryo body is due to SDF1a action. During late PGC migration (step III), the role of Sdf1b is becoming predominant while the function of Sdf1a and CXCR4 appears to be much less important. The recent description of the sequential action of the CXCR4 and CXCR7 receptors in regulating PGC migration in Medaka (17), indicates that a CXCR7-dependent process would act after the CXCR4-dependent bilateral alignment of PGC and causes drift of PGC to the gonadal area. **[see papers #9 and #15]**.

The medaka male sex-determining gene is a duplicated version of the *dmrt1* gene. This duplicated copy was designated *dmrt1bY*. The ancestral copy, which is located on an autosome, was consequently named *dmrt1a*. The duplicated fragment, consisting of *dmrt1* and three neighbouring genes, was inserted into another chromosome while its homologue became the X (**[see paper #20]** and Figure 7). This duplication event occurred approximately 5-10 million years ago in the lineage leading to the medaka. Compared to the human Y, which is more than 200 million years old, the medaka Y is very young and allows studying the initial events of Y-chromosome evolution.

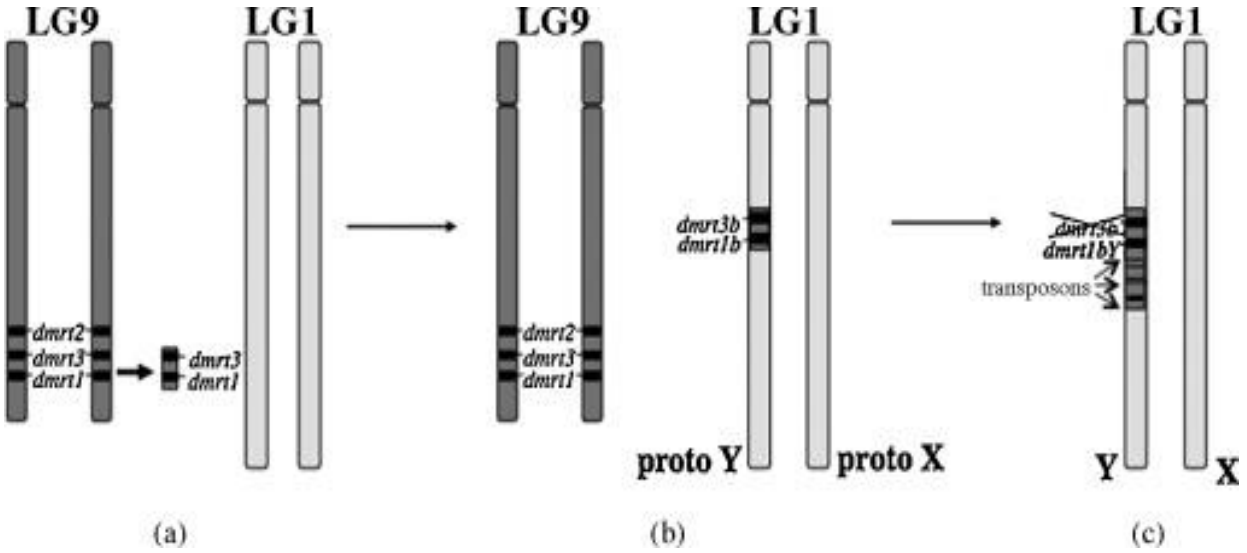


Figure 7. Origin and evolution of the sex chromosomes in medaka.
(a) A segment on linkage group (LG) 9 containing the dmrt1 gene and neighbouring genes including the paralogs dmrt3 and 2 is duplicated. **(b)** The duplicated segment is inserted into one of the chromosomes of LG1. This chromosome is the proto Y while its homologue becomes the proto X. **(c)** Due to lack of recombination between the duplicated fragment on the Y and the X genes not involved in sex determination degenerate and transposable elements and repetitive sequences accumulate. **[see paper #20]**.

Of special interest is the question how a gene, dmrt1 for instance, which is placed downstream in the majority of the sex determination cascades (see Figure 4), was brought about to take over the position and function of the gene at the top in medaka. So far, using phylogenetics we found that the Dmrt1 duplicated copy on the Y-chromosome (Dmrt1bY), which became the male determining gene, diverged considerably in the 5' flanking region from its ancestral gene, *dmrt1a* **[see paper #21]**. Further on, functional promoter analyses in different transfected cell lines (luciferase and promoter bashing assays) and *in vivo via* transgenic reporter fish, revealed that a transposable element (called "*Izanagi"*), inserted in dmrt1bY promoter shortly after the gene duplication event, brought in 'ready to use' transcription factor binding sites including a *dmrt1* target site **[see paper #21]**. Finally, using electromobility shift assay (EMSA) and *in vivo* chromatin immunoprecipitation (*in vivo* ChIP) we have demonstrated that the insertion of the "*Izanagi"* element results in a feedback (*cis*-) downregulation of *dmrt1bY* after the sex determination period and (*trans*-) downregulation of the sex determination gene by its autosomal ancestor, *dmrt1a,* in adult testes **[see papers #21 and #13]**. This new transcriptional regulatory element, nested within Dmrt1bY promoter, allowed neo-functionalization of the sex chromosomal copy and guaranteed the survival of the autosomal copy by a possible sub-functionalization in Sertoli cells of adult testes (**[see papers #20, #21, #23 and #24]** and Figure 8). This is a first mosaic stone of the evolutionary history how *dmrt1bY* became the master sex regulatory gene, but certainly much more has to be explained to understand this process in detail.

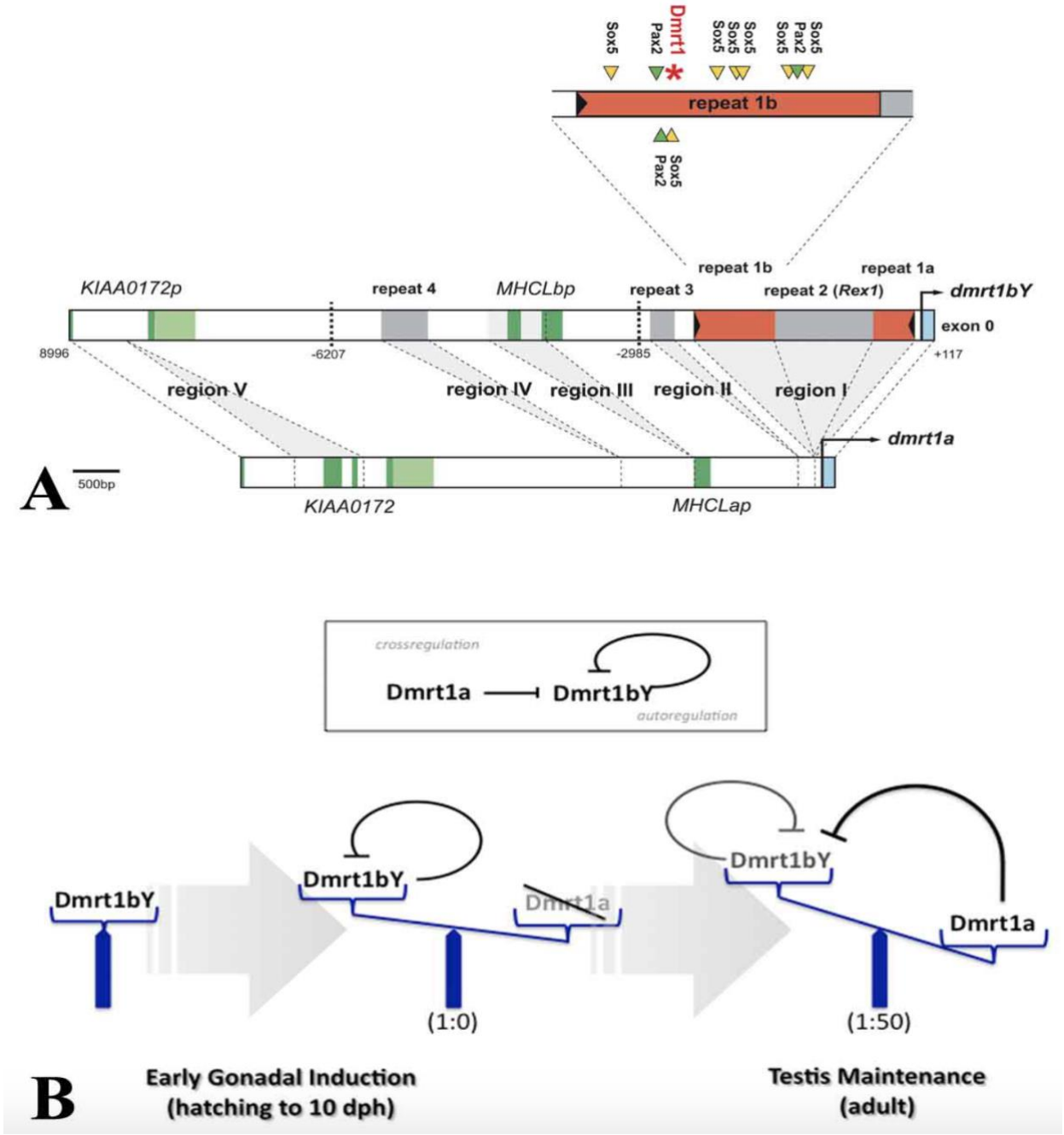


Figure 8. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements.
(A) Comparative analysis of the dmrt1bY promoter and its transcription factor binding sites. The analysed promoter region of dmrt1bY in comparison to its dmrt1a paralog. Length differences between the promoters are based on regions I–IV, of which three (I, II, IV) have been added to the dmrt1bY promoter after duplication, while two others have been lost from the dmrt1bY promoter (KIAA0172 region V) or from the dmrt1a promoter (MHCLp region III), respectively. Bold dashed lines indicate cutting sites for transcriptional regulation analyses. Region I contains a putative Izanagi DNA transposons (repeat 1), into which a Rex1 element (repeat 2) was inserted secondarily. The upstream part of split repeat 1 (repeat 1b) in the dmrt1bY promoter contains multiple Sox5 and Pax2 binding sites as well as a Dmrt1 binding site. Dark green indicates coding sequence of genes and pseudogenes, light green indicates their untranslated regions.
(B) Model for feedback and cross-regulation of the medaka dmrt1 paralogs. During sex determination stages only dmrt1bY is expressed and dmrt1a is off. Hence, the sex determining function of dmrt1bY is exerted. In adult testes, both paralogs are expressed notably in Sertoli cells, but the auto-repression of the dmrt1bY promoter by its own gene product and the cross-regulation by Dmrt1a lead to a predominant expression of dmrt1a compared to dmrt1bY (appox. 50-fold) **[see paper #21]**.

To further understand the function of Dmrt1bY in orchestrating such a complex gene regulatory network, I have then driven my research toward the identification of potential targets genes and regulated pathways downstream of the induction of Dmrt1bY. Hence, monitoring *in vivo* gene expression and interactions by transgenic fluorescent reporter fish lines, using *in vivo* tissue-chromatin immunoprecipitation and *in vitro* gene regulation assays, I already started to outline the genetic framework in which Dmrt1bY is interacting during medaka gonadal formation and maintenance **[see papers #27 and #29]** (Figure 9). The results of these analyses revealed several concordances but also major discrepancies between mammals and medaka, notably amongst spatial, temporal expression patterns and regulations of the canonical Hedgehog and R-spondin/Wnt/Follistatin signalling pathways (**[see paper #28]**, Figure 9). For example, unlike their mammalian counterparts, none of the different medaka gonadal marker genes analysed (*r-spo-1*, *fst* or *ptch-2*) were detected in germ cells at any time of gonadal development. This major inconsistency between mammals and fish certainly reflects intrinsically divergent modes of germ cell commitment and interaction between germ and somatic cells, possibly accounting for a higher sexual plasticity of germ cells in fish. In this context, our data reflect a profound reorganization of parts of the fish gonadal regulatory network compared to mammals. While some components like DMRT1, SOX9, FOXL2 and pathways such as Hedgehog or R-spo1/Wnt/Fst of the gonadal gene regulatory network are conserved on the DNA sequence level across phyla, their functions, regulation and interplays might be considerably different (see Figure 9 and **[papers #27 and #29]**).

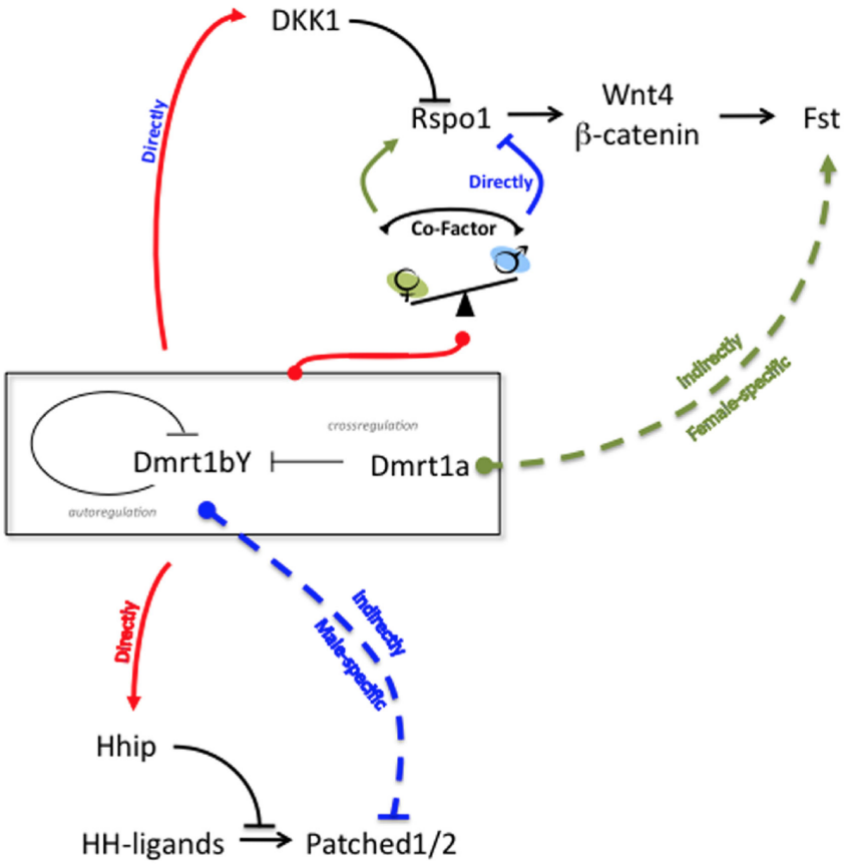


Figure 9. The *Dmrt1bY/Dmrt1a* gene regulatory networks during gonadal formation in medaka.
Interaction scheme of the possible Dmrt1bY/Dmrt1a-triggered regulations of the Hedgehog and Wnt4 pathways during gonadal formation in medaka. Solid red arrows indicate for both Dmrt1bY and Dmrt1a positive regulation while dashed lines indicate sex-specific regulations. The green dashed line indicates a Dmrt1a, female-specific, indirect positive regulation favoring the expression of *follistatin* while the blue dashed line reports a Dmrt1bY, male-specific, indirect repression of *patched-2* transcription. For the Dmrt1bY/Dmrt1a-triggered transcriptional regulation of *R-spondin1*, depending on the cellular contexts, the involvement of a sex-specific co-factor is proposed **[see paper #28]**.

Interestingly, examining early gene expression and behaviour of just-specified germ cells, we could show that medaka germ cells exhibit sexually different characters before the formation of the somatic gonadal primordium. Into that direction *in vitro* culture analysis additionally revealed that XY germ cells behave differently from XX germ cells with respect to their mitotic activity. Again, our results challenge the traditional view of the initial acquisition of sexual identity in the somatic cells and provide insights into the evolutional and developmental processes on cell-autonomous sexual identity at cellular levels **[see paper #29]**. In addition to the genes mentioned above, examination of Foxl2 protein distribution in the adult medaka ovary revealed a new sub-population of theca cells, where ovarian-type *aromatase* transcriptional regulation appears to be independent of Foxl2 (Figure 10).

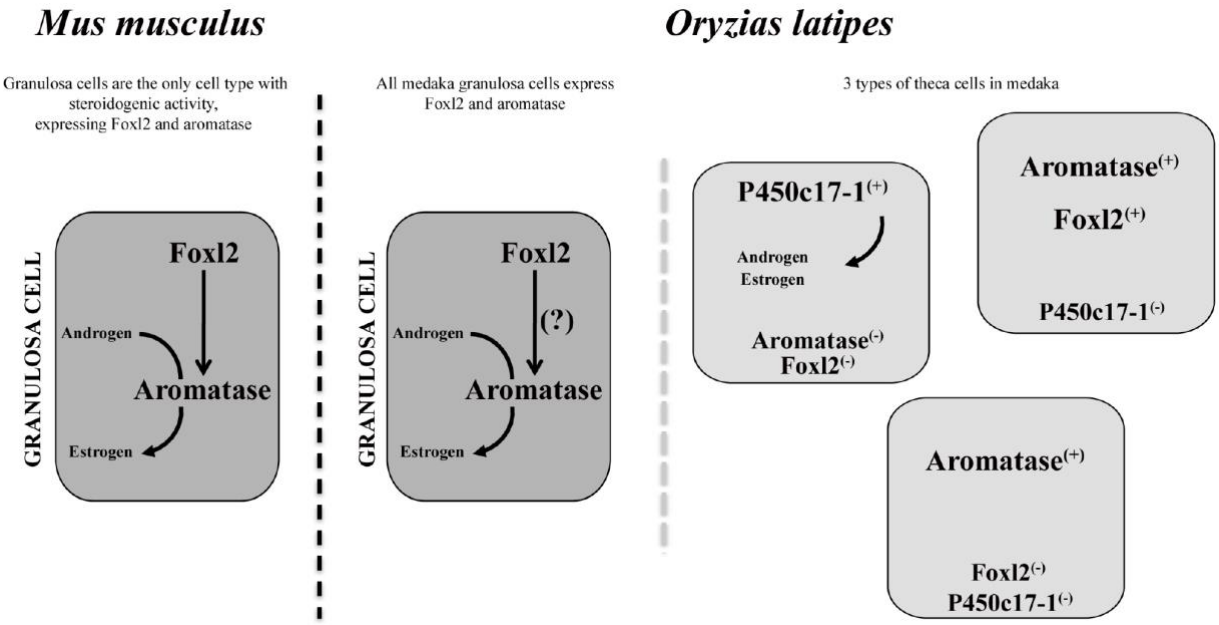


Figure 10. Schematic representation of granulosa and theca cell populations in mouse and medaka.
In mammals, granulosa cells are the only cell type with steroidogenic activity, expressing both foxl2 and aromatase. Aromatase expression is directly induced by foxl2. In medaka, like in mammals, granulosa cells express both Foxl2 and aromatase. Examination of Foxl2 protein distribution in the medaka ovary revealed a new subpopulation of theca cells expressing Foxl2 **[see paper #28]**.

As a result, we could show that the regulation of the downstream regulatory network of sex determination is less conserved than previously thought and emphasize the importance of the cellular context on modulating these regulations in Vertebrates.

II- Current Research Work, Objectives and Perspectives

1- Context and positioning of our current research work at INRAE

In stark contrast to the traditional ovine, bovine or porcine animal productions, most of the aquaculture fish species, due to the recent expansion of fisheries, are only slightly domesticated. While genetic improvement, as it is in other more mature forms of animal productions, could be also considered as a conventional contributing factor in developing efficient fish farming, it has only recently been applied to just a handful of species (carp, catfish, trout and tilapia). Whereas additionally the number of fish species suitable for aquaculture is steadily increasing worldwide, nearly all farmed aquatic species are still very similar to their wild relatives. As a matter of fact, most fish did/do not benefit from a deep basic understanding of their biology, often precluding development of a rational and sustainable aquaculture. Therefore, efforts should converge towards an improvement of production efficiency (growth, disease resistance, fertility…) and adaptation to new contexts (fluctuating environment, pollution, density…). Improvement in that direction, in addition to the high potential of "classical" methods, should also take advantage of the development of genomic technologies and biotechnologies, first to make the best use of molecular pedigrees and in a long-term vision to improve the knowledge in the genetic and biological bases of traits as well as in directly selecting on the genotype. Biotechnologies certainly provide powerful tools for sustainable development of aquaculture, fisheries, as well as food industry. Beside the "hot-seller" idea of universally increasing fish growth rates, more reasonable and need-driven genetic biotechnologies are already importantly contributing but also significantly challenging aquaculture and fisheries development. For instance, chromosome sex manipulations to induce polyploidy (triploidy and tetraploidy) and uniparental chromosome inheritance (gynogenesis and androgenesis) are important in the improvement of fish breeding as they provide a rapid approach for gonadal sterilization, sex control improvement of hybrid viability and clonation. But also, the development of embryonic stem cell technologies, cryopreservation, manipulation and transplantation of fish germ cells would allow the creation of broodstock systems in which target species can be produced from surrogate or allogenic parents. Undeniably the benefits offered by these new technologies couldn't be fulfilled without constant commitment for basic research.

Control of sex is an important issue of modern aquaculture as it allows the mass production of either all-female or all-male populations of fish that are often economically more interesting to breed than normal mixed-sex populations **[see BC #2]**. In salmonid species for instance, all-female populations are often preferred because males have the propensity to mature precociously, resulting in reduced growth rates, lower food conversion efficiency, lesser flesh quality and a high sensitivity to pathological problems. In other fish species, sex control can either ***i***) facilitate broodstock management, for instance by optimizing the ratio between males and females in hermaphrodite species, ***ii***) prevent uncontrolled reproduction that favours energy investment into the gonad instead of body growth like in the tilapias or ***iii***) allow the production of a sex specific product like the caviar in female sturgeon fish. A better knowledge on fish sex determination and sex differentiation mechanisms is then a pre-requisite towards a more rational and efficient control. This is especially essential for fish as the number of aquaculture fish species is important and diversification is still a current challenge in many countries. Akin to sex determination systems, genetic sex determinants and even their downstream regulations are not well conserved (see part I), the transfer of a sex-control technique from one species to another is often problematic. Even in species in which biotechnologies are available, and allowing mass production of all-male or all-female populations, a better knowledge of sex-determination and sex-differentiation mechanisms would still allow an evolution of the current biotechnologies

towards more sustainability. This could be done, for instance in salmonids, by finding alternative methodology for the production of neo-males that are currently obtained by (massive) steroid hormone treatment **[see BC #2]**.

Consequently, my project at INRAE was built into that perspective of bringing in and developing, in parallel to the existing fish models already well-established at the INRAE/LPGP laboratory for studying sex differentiation, a new -complementary- model for studying gonadal differentiation and development in order to bypass some limitations inherent to the salmonid models. These limitations are mainly related to the size and generation time of salmonids that prevent any development of easy and affordable cell lineage tracing, transgenesis, targeted inactivation and other genetic manipulations that are now needed to develop deep functional genomic approaches that are absolutely mandatory to gain further insight into the regulatory mechanisms of sex determination in fish. Medaka is certainly one of the best-understood fish model organisms regarding to early gonadal induction, differentiation and maintenance (morphogenesis, sex determination and differentiation, oogenesis and spermatogenesis (see part I). Its sex determination system relies on a simple male heterogametic system (XX/XY) and its master sex-determining gene (*dmrt1bY*) has been characterized. This species is also eligible for the most cutting-edge tools in terms of transgenesis (BAC clones recombination, meganuclease and transposon-mediated fluorescent reporters…), RNA injection, *in vivo* ChIP, *in vivo* inducible systems (on/off) after recombination, genome editing (TALENS, CRISPR), stem cells (embryonic, spermatogonia, and embryoid bodies) and imaging (live confocal imaging).

Hence, in that particular context of functional and comparative evolution and genomics, and using the medaka fish as a model organism of reference, my current research projects on sex determination / sex differentiation in fish are interdependently constructed around few main general questions that constitute my research objectives:

(i) Functions and Evolution of Master Sex Determining Genes and Systems

(ii) Functional Analysis of the Gene Regulatory Networks Underlying Sex Determination and Differentiation.

(iii) Gonadal Morphogenesis and Plasticity

(iv) Research and Development: Genome Editing / Single Cell Resolution

[v] Autophagy in Fish

(i) Functions and evolution of master sex determining genes and systems: cases studies

The high diversity of sex determination systems seen in teleost fish is linked to the high turnover rate of their sex chromosomes (Figure 2 **[see BC #1, #3, #4]**). As a result, most fish sex chromosomes are considered to be relatively young and are often described as being homomorphics, i.e., with little differentiation and no cytological difference. Sex chromosomes turnover is usually connected with the emergence of new master SD genes (4, 18). Indeed, in therian mammals and birds, master SD genes are highly conserved with the quasi-exclusive usage of SRY for mammals and DMRT1 for birds. But, a high diversity of SD genes (Figure 2) has been found in the past two decades in many teleost fish and this has greatly expanded our understanding of the SD mechanisms and SD evolution. Despite this great diversity, most of these newly

discovered fish master SD genes are classical "usual suspects" **[see paper #31]** that fall into three independent protein families implicated in the "canonical" vertebrate sex differentiation gene network, namely the Sry-related HMG box (SOX), the Doublesex and mab-3 related transcription factors (DMRT) and the transforming growth factors (TGF-β) family. Their recruitments as master SD genes do not follow any obvious phylogenetic relationship and are more likely to be independent and repeated events during teleost evolution (Figure 11).

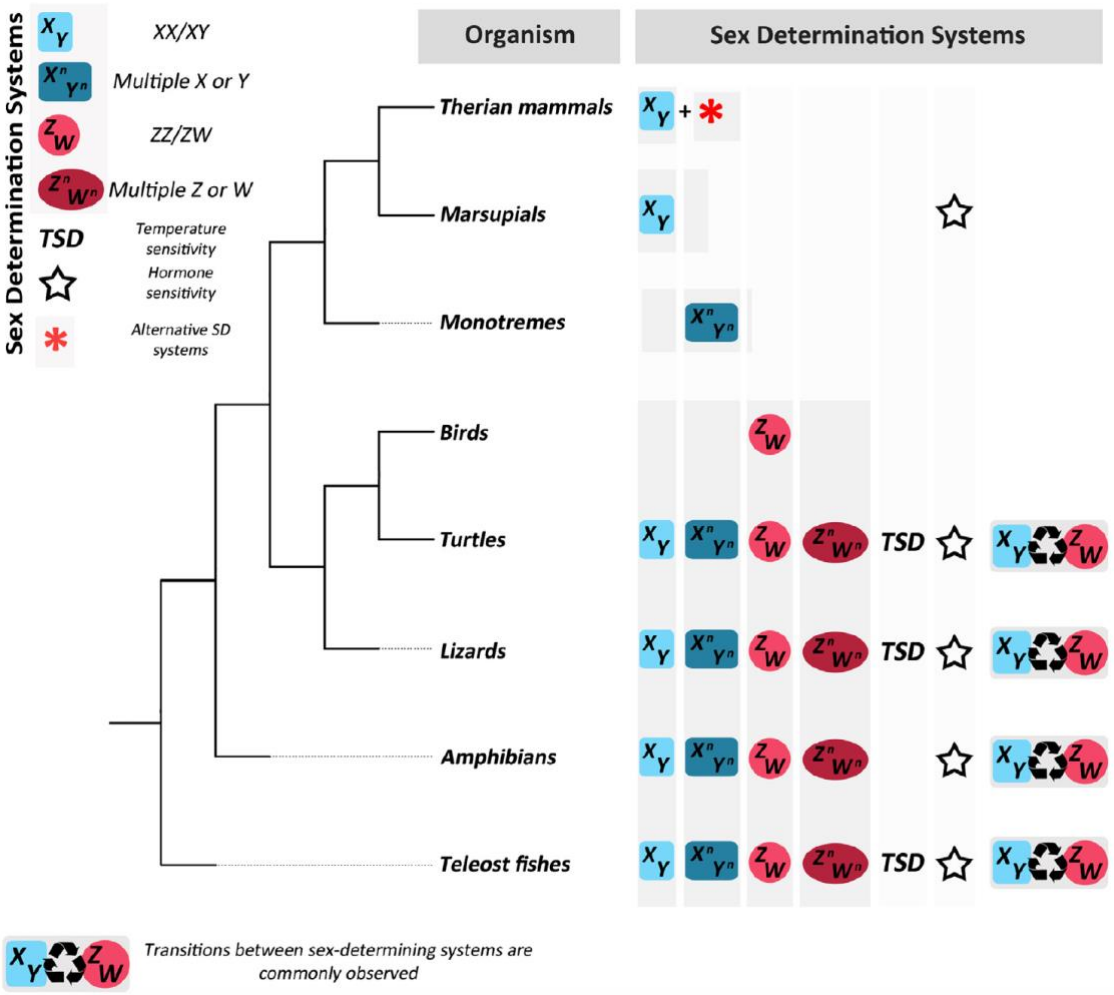


Figure 11. Diversity and evolution of sex determination systems and master sex-determining genes in vertebrates.

In contrast to therian mammals and birds, the other vertebrate groups including reptiles, amphibians and fish, exhibit plasticity of sex determination mechanisms (XX/XY, multiple X or Y, ZZ/ZW, multiple Z or W, autosomal modifiers and environmentally induced sex determination) and of master sex-determining genes.

Only one known exception to this "usual suspects" rule is the salmonid SD gene, sdY (sex determining region on the Y chromosome, (19)), which has derived from the duplication of an immune-related gene unsuspected to be involved in the sex differentiation pathway. The birth of a new master SD genes can start with a duplication of an autosomal gene that will acquire a new SD function leading to new sex chromosomes. On the other hand, it can begin with allelic diversification from a step-wise diversification of two alleles of the future sex chromosome pair. One allele then becomes the sex determiner, while the other allele either retains a non-SD function or favours the development of the opposite sex (20). The gene duplication and allelic diversification mechanisms have both been found in teleost fish to generate new master SD genes and drive the turnover of sex chromosomes (Figure 12). In the following part of this manuscript, I will review the cases of master sex determination genes in teleost fish we have been working with.

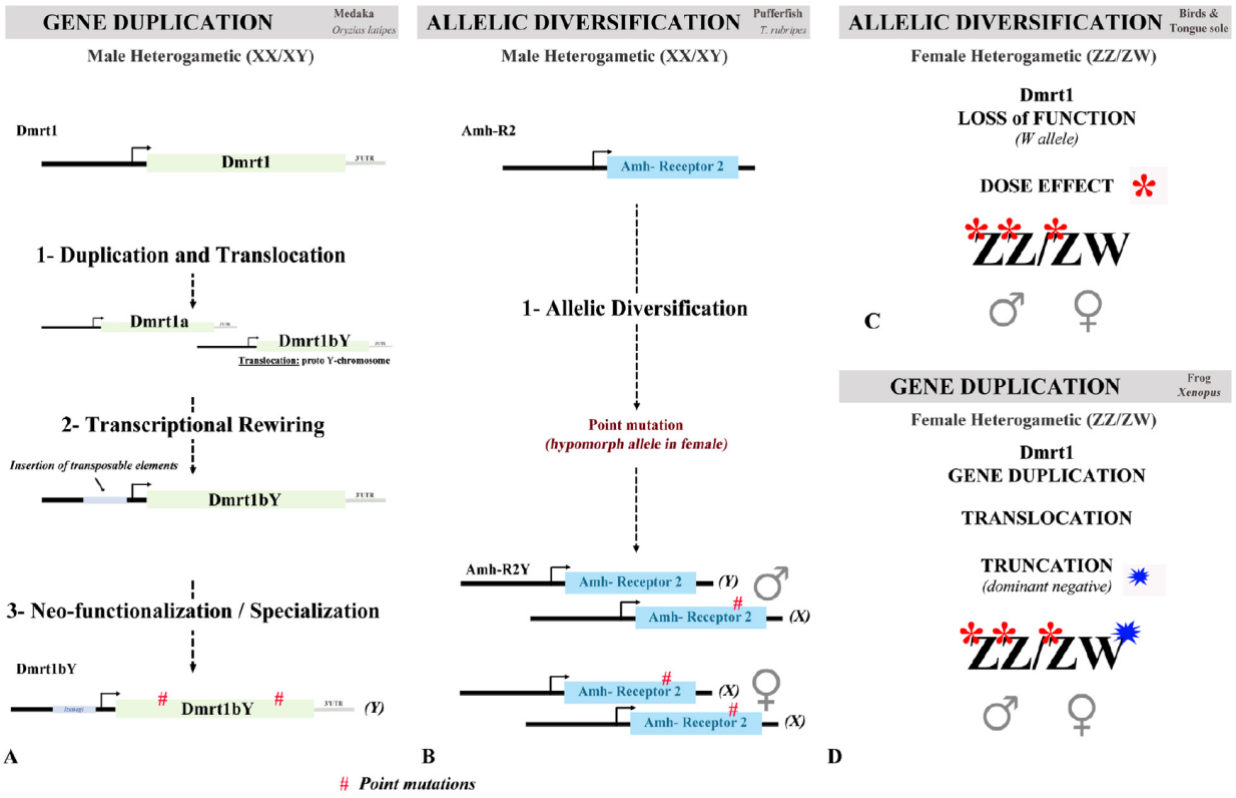


Figure 12. Mechanisms accounting for the emergence of new master sex-determining genes after either gene duplication (examples medaka and frog) or allelic diversification (examples torafugu, birds and tongue sole). **(A)** The medaka dmrt1 gene, together with neighbouring genes (including dmrt2 and 3), first underwent local gene duplication. This whole duplicated segment was then inserted into another chromosome, creating a proto Y-chromosome. Because of non-homology of the insert to the proto-X recombination is suppressed and this part of the proto Y segment degenerated. Transposable elements and repetitive sequences accumulated, resulting in transcriptional rewiring of the dmrt1bY gene and acquisition of an early gonadal expression pattern compatible with a master sex-determining function. Further on, point mutations accounted for neo-functionalization and specialization of the dmrt1bY gene **[see papers #16, #18, #21 and #22]**. **(B)** In the pufferfish as a result of allelic diversification two versions of the amh receptor II exist differing only by one aminoacid located in the kinase domain (H384D). This hypomorphic mutation conferring lower receptor activity is encoded on the X-chromosome. Quantitative variations in Amh signalling in females (homozygous for the hypomorphic amhrII allele) versus males (heterozygous for the wild type and hypomorphic alleles), account for male gonadal development (Reproduced from (21)). **(C)** Birds and the tongue sole and have a ZW female heterogametic system. A single loss of function event in the dmrt1 gene occurred (W allele). This resulted in ZW individuals having only one copy of the dmrt1 gene while ZZ individuals harbour two copies (red stars indicate functional dmrt1 genes). Here a dosage effect leads to either female (ZW, 1 "dose" of dmrt1) or male (ZZ, 2 "doses" of dmrt1) gonadal development respectively (15, 22). **(D)** Sex determination in the frog Xenopus laevis is female heterogamety. Similar to medaka the frog dmrt1 gene underwent gene duplication and translocation. Additionally, truncation of the dmrt1 gene occurred, generating a dominant negative version (DM-W). The DM-W protein product from the W-chromosome antagonizes the action of the wild type dmrt1 gene on the Z-chromosome, leading to female gonadal development. ZZ individuals having two copies of the dmrt1 gene (and no DM-W) develop as males (16). Red and blue stars indicate functional and dominant-negative version of the dmrt1 gene respectively.

(ia) Dmrt1bY the medaka master sex determining gene: function and evolution.

The dmrt1 gene is a transcriptional factor belonging to the evolutionarily well-conserved DM (Doublesex and Mad-3) domain DMRT gene family, which is involved in sexual development in many phylogenetically distant groups from worms to mammals **[see papers #23 and #24, #31, and BC03]**. In teleost fish, sexually dimorphic expression of dmrt1 has been correlated with gonadal identity in both gonochoristic and hermaphroditic fish species **[see paper #24]**.

Dmrt1bY, was the first non-mammalian SD gene found in vertebrates and it is a duplicated copy of autosomal dmrt1/dmrt1a gene found in two ricefish species of the genus *Oryzias*: the Japanese medaka (*O. latipes*) and the Malabar ricefish (*O. curvinotus*), both having an XX/XY sex determination system (8, 9) (Figure 1). Dmrt1bY arose about 5–10 million years ago from a local duplication of an autosomal fragment encompassing the dmrt1 gene (Figure 7). Consequently, this pair of X and Y chromosomes are very young in comparison with the chromosome pair found in therian mammals, and remain homomorphic **[see paper #20]**.
The expression pattern of the medaka dmrt1bY is typical of a master SD gene with an early and transient expression during gonadal primordium differentiation, and low persistence in the adult gonads **[see paper #13]**. Around the time of hatching dmrt1bY is first expressed in the nuclei of the somatic undifferentiated cells surrounding the primordial germ cells (PGC) and later on exclusively in the Sertoli cells in males **[see papers #13, #18 and #29]** (Figure 13). In comparison, the autosomal dmrt1/dmrt1a is expressed much later at 20 days after hatching in males and shows a testis-specific expression pattern in adult males with 50 times higher expression than that of dmrt1bY **[see paper #13]**.

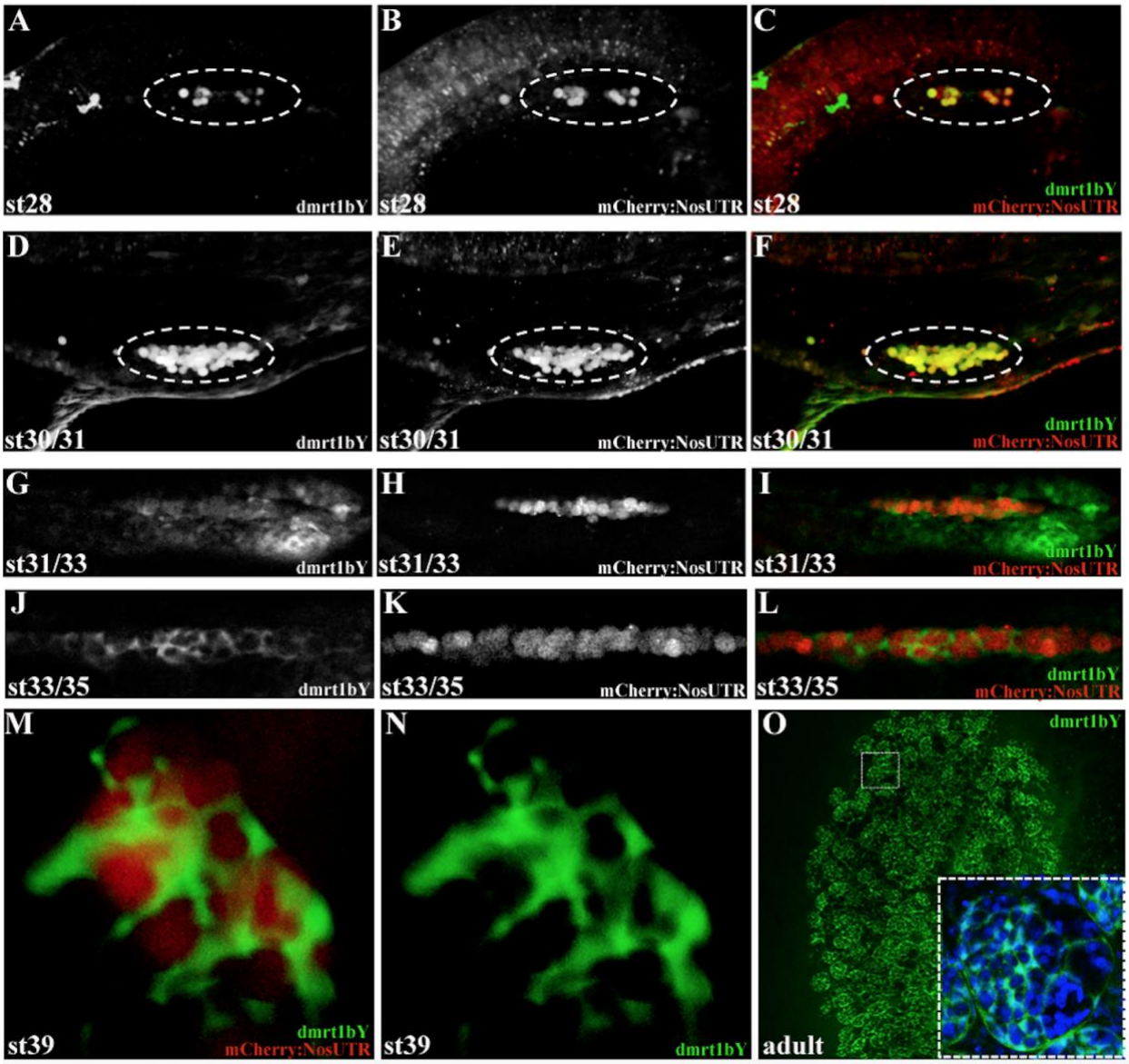


Figure 13. Detection of dmrt1bY expression in germ cells and Sertoli cells of dmrt1bY-eGFP reporter transgenic medaka at hatching and adult stages.

(A-C) Dmrt1bY expression is first detected in clustered germ cells located within the lateral plate mesoderm at stage 28. Expression is co-localized with the fluorescence of the germ cell-specific marker (mCherry:nos-3'UTR). **(D-F)** Germ cell-specific expression is observed until stages 30-31. **(G-L)** By stage 33, while germ cell expression becomes weaker **(G-I)**, additional dmrt1bY expression is detected in the somatic cells directly surrounding the germ cells **(J-L)**. **(M-O)** In adult testes dmrt1bY expression is restricted to Sertoli cells.

Concomitantly to the acquisition of a dominant position within the sex-determining network, we have previously shown that dmrt1bY was subjected to a profound rearrangement of its regulatory landscape brought about by exaptation of a "ready-to-use" cis-regulatory element from a transposable element (TE). This element, called *Izanagi*, recruits Dmrt1bY and Dmrt1a proteins to turn off the dmrt1bY gene after it has fulfilled its function as the primary male SD gene **[see paper #21]** (Figure 8).

Interestingly we have further reported that TE-mediated transcriptional rewiring can reach an unexpected level of complexity that exceeds this simple feedback regulation. Indeed, we found that another TE, Rex1, has jumped into Izanagi **[see paper #36]**. Through the disruption of Izanagi, Rex1 immobilized this TE and fixed the Dmrt1-mediated downregulation. Moreover, Rex1 brought in a preformed regulatory element for the transcription factor Sox5. We have then demonstrated that medaka Sox5 binds to the sox5-responsive elements of the dmrt1bY promoter and downregulates its transcriptional activity. Interestingly, *in vivo* analysis of double transgenic fluorescent reporter fish additionally revealed a complementary pattern of expression of both genes. The higher expression of sox5 correlates with a lower expression of dmrt1bY and *vice versa* (Figure 14). Our results underpin the importance of the Rex1 TE for the establishment of a new SD mechanism in medaka and likely contribute in establishing the temporal and cell-type specific expression pattern of dmrt1bY **[see paper #36]**.

Showing that sox5 was recruited to the very top of the primary SD cascade after insertion of Rex1 and that it controls the fine-tuning of dmrt1bY expression, our results provide evidence for a more general and ancestral SD function of Sox5 in regulating germ-cell number and, in consequence, gonadal identity.
Interestingly, the finding that a preformed transcription factor binding site contributed by the Rex1 transposon modulates the regulation of dmrt1bY promoter highlights the important role that mobile elements play in the genome for shaping the evolution of new functions. Intriguingly, although *bona fide* examples of this process are still rare (23), Rex1 is the second such event found in the same promoter. It will be interesting to analyse whether the other repeats present in the dmrt1bY, but not in the promoter region of dmrt1a, provide further instances of TE exaptation. Genes that arose by gene duplication such as dmrt1bY are primarily dispensable and can only escape degeneration through sub- or neo-functionalization. As dmrt1bY and dmrt1a both have exclusive functions in male sexual development in line with the highly conserved role of dmrt1 in invertebrates and vertebrates (3), a change in transcriptional control *via* the insertion of two different TEs might initially have led to sub-functionalization; dmrt1bY acquired its transient early expression, whereas the transcription of dmrt1a was pushed back to the later testis differentiation phase. It will be interesting to have a closer look at the SD genes of other fish that have been subject to fast evolutionary change and thus, might also be targets for TE exaptation.

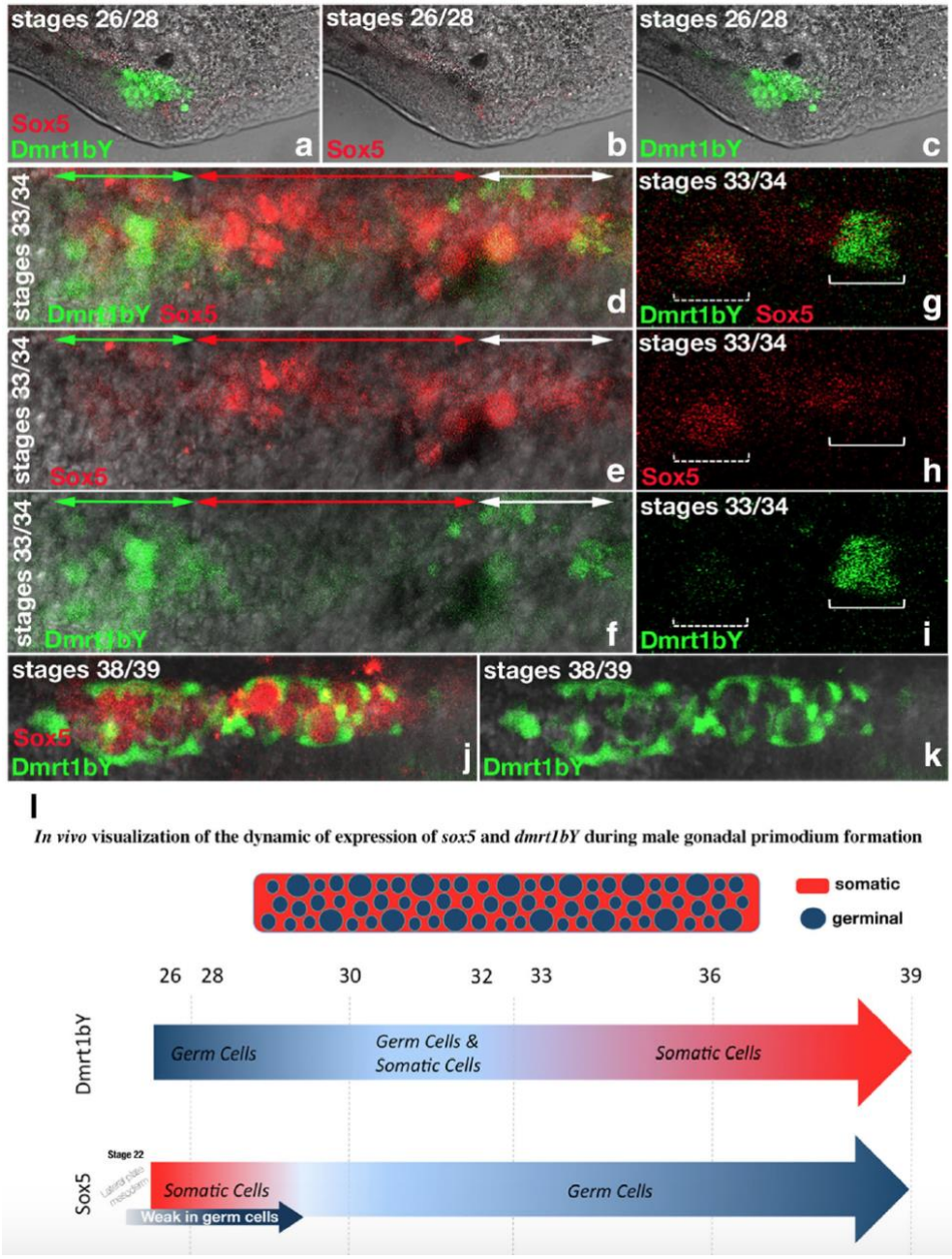


Figure 14. Comparative analysis of sox5 and dmrt1bY expression dynamics during gonadal primordium formation.

Expression of sox5 compared to dmrt1by in a double transgenic fluorescent reporter line. **(a–c)** During early gonadal formation, sox5 is first detected in the somatic tissues surrounding the germ cells at stages 26 to 28. At the same time, dmrt1bY is expressed in germ cells. **(d–i)** By stages 33 to 34, sox5 expression becomes restricted to the germ cells. dmrt1bY is also expressed in the germ cells at those specific stages of development. Variations within the respective levels of sox5 and dmrt1bY expression are clearly observable (**d** compared to **e** and **f** and **g** compared to **h** and **i**). **(j,k)** Around hatching (stages 38/39), the expression of sox5 strengthens in all germ cells while parallel dmrt1bY expression quickly switches from germ cells only to somatic germ-cell-surrounding cells only. **(l)** *In vivo* visualization of the dynamics of expression localization of sox5 and dmrt1bY during male gonadal primordium development. The expression of sox5 and dmrt1bY is highly dynamic during primordium gonadal formation, switching from somatic to germ cells and vice versa, respectively, from stage 26 until hatching. Being mutual repressors of each other, a seesaw of expression is observed, finally finely restricting dmrt1bY expression in the somatic part of the primordium gonad. Blue and red represent cellular expression localizations only and should not been interpreted as expression levels.

(ib) sdY: an unusual and conserved SD gene in Salmonids.

The recent discovery of a new type of SD gene in rainbow trout (*Oncorhynchus mykiss*) demonstrated an unexpected flexibility of the system in vertebrate SD (19). This gene, called sdY, came to the spotlight as it was found to be expressed only in embryonic testes and was also tightly linked to the sex-locus on the Y chromosome. It was further confirmed to be the trout SD gene based on functional proofs showing that sdY is by itself necessary and sufficient to induce testicular differentiation (19). One of the multiple unusual characteristics of sdY is that it evolved from the duplication of an immune-related gene, irf9 (interferon related factor 9). Compared to its ancestral protein, sdY lost the N-terminal DNA-binding domain, but preserved its C-terminal protein-protein interaction domain. Another interesting feature is that sdY is sex-linked in most salmonid species, indicating that it is a conserved SD gene in this teleost group. Yet, the SEX *locus* containing sdY is not located in the same chromosome in different salmonids species (24), suggesting a rapid turnover of sex chromosome while preserving what seems to be a jumping master SD gene (25).
Our additional studies strongly suggested that SdY could mediate its sex-determining effects through protein-protein interactions. Indeed, the search for potential interacting partners uncovered that SdY is interacting with the forkhead box L2 (Foxl2) protein **[see paper #39]**, a transcription factor well-known for its crucial role in ovarian differentiation **[see paper #34]**. By binding to Foxl2, SdY is able to prevent the Foxl2 and Nr5a1 (Nuclear Receptor Subfamily 5 Group A Member 1) synergistic activation of the aromatase (cyp19a1a) gene, blocking the positive feedback loop of regulation required for synthesis of the feminizing estrogenic steroids (26) in the early differentiating gonads. Thus, we could show that SdY is not a "genuine" male-promoting factor as, despite being a male MSD gene it acts as an anti-female factor by blocking a default female differentiation regulatory network, and thereby allowing testicular differentiation to proceed. Hence, SdY presents a paradigmatic case that un-related genes outside the "usual suspects" are able to acquire *de novo* sex determining functions in teleosts. These results additionally suggest that the evolution of unusual vertebrate master sex determination genes recruited from outside the classical pathway -like sdY- is nevertheless strongly constrained by their ability to interact with the canonical gonadal differentiation pathway (27) (Figure 3).

(ic) Esocidae species challenge the fitness advantage and sexually antagonistic models.

Classical models of sex chromosome evolution postulate that there first occurs an expansion of regions with reduced recombination near the MSD gene, and that the loss of recombination subsequently results in the gradual decay of the whole sex chromosome (28, 29). These models are based on the highly degenerated sex chromosomes of model species, and the universality of these models has recently been called into question by new findings in non-model species we have been looking at. For example, we could show that the MSD gene in Northern pike (Amhby) arose > 65 My ago **[see paper #43]** and that the surrounding Y-specific region remains restricted in size and without other protein coding genes (Figure 15). Furthermore, no male beneficial/female antagonistic genes have been found in this tiny SD region **[see paper #43]**. Interestingly, we pointed out that this very restricted sex-differentiated region encompassing *amhby* is a conserved feature of all Y-chromosomes in several species of this clade, illustrating that SD systems can be stable for over 50 million years **[see paper #43, and Pan *et al.* submitted]**. These examples temper the "need" for highly degenerated sex-chromosomes as 'evolutionary traps' for supposedly preventing the turnover of SD systems (4).

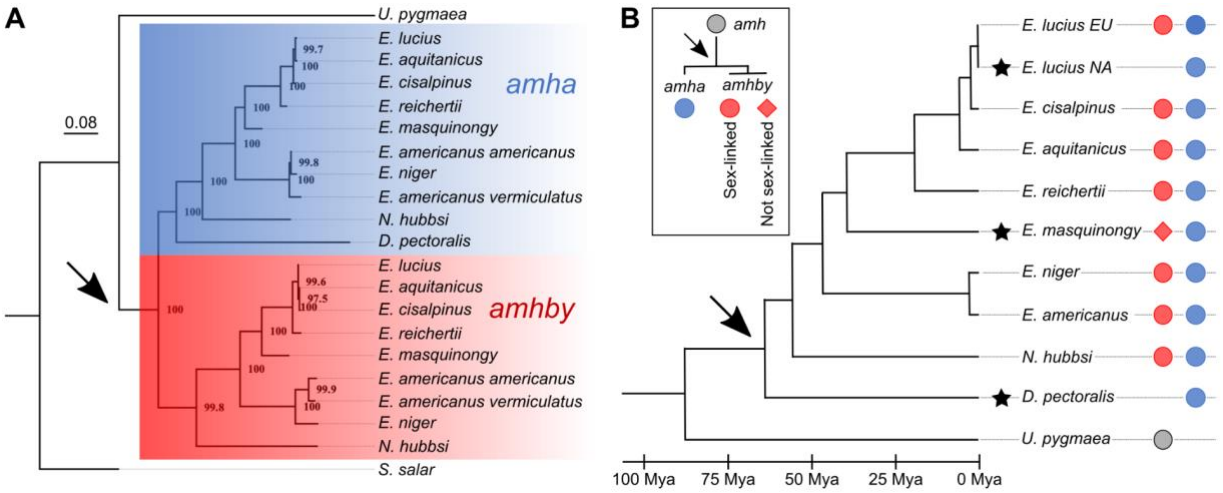


Figure 15. Phylogenetic analysis of amh homologs from the Esociformes revealed an ancient origin of amhby. **(A)** Phylogeny tree of amh-coding sequences built with the maximum likelihood method. Bootstrap values are given on each node of the tree. The amha ortholog cluster is highlighted with blue background and the amhby ortholog cluster is highlighted with red background. **(B)** Species phylogeny of the Esociformes. The three putative sex determination transitions are shown by a black star. The presence of pre-duplication amh, amha, and amhby along with its sex-linkage is represented by coloured dots at the end of each branch. The earliest duplication timing of amh is denoted by a black arrow at the root of the Esocidae lineage.

(id) The "improbable" sex determination in the Atlantic herring

Classically, the different sex determination triggers have been identified to be components of either the genome itself (genetic sex determination, GSD) or come from the environment (environmental sex determination, ESD). Adaptive hypotheses prevail in explaining the evolution of the various sex determining triggers, which finally will guarantee a stable sex ratio of the population. For ESD according to the Charnov-Bull model (30) an environmental trigger, e.g. the incubation temperature of the developing embryo, is favoured when the developmental environment differentially influences male *versus* female fitness, as for instance shown in the Jacky dragon (23). GSD is generally explained to be stabilized (if we except the just above-described situation in esociforms) through the evolution of sex chromosomes, when sexual selection links the genetic sex determiner to a *locus* that is beneficial to the same sex or even antagonistic to the opposite sex (31–35). Recently, a third general mechanism has been proposed, namely random sex determination (RSD) (36). Perrin (36) recognized that many ESD and GSD systems are not always perfectly reflecting the command of the trigger, when for instance in an ESD system at a transitional temperature range males and females are produced, or when GSD species without an obvious reason generate individuals which have the phenotype of the sex that is opposing their genotypes. This indicated stochastic processes resulting from developmental noise (36). Extrapolating this reasoning, a model was proposed that sex as a bistable equilibrium (the male and female phenotype) can be triggered by any random process which, from the undifferentiated state, tips the balance to the one or the other direction. If the population size is large enough RSD would guarantee an equal sex ratio as well (Figure 16).

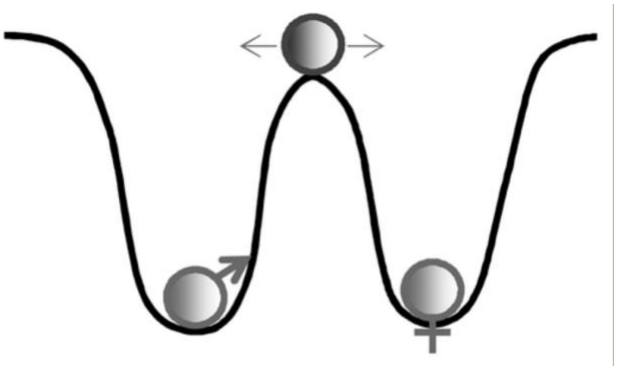


Figure 16. Sex as a bistable equilibrium.
Genomes are fundamentally bipotential, encoding developmental landscapes (black line) that allow for either male (left) or female (right) development. From an undifferentiated larval state (dark ball in unstable equilibrium), initial triggers tip the balance in one or the other direction (arrows), after which mutually antagonistic gene pathways ensure convergence toward fully fledged male or female phenotypes (36).

We have initially chosen to look at sex determination in the Atlantic herring (*Clupea harengus*) because this fish has characteristics that would allow RSD to act. Indeed, it is one of the most abundant vertebrates on earth (swarms of up to 4 billion individuals) with massive spawns (up to 200,000 eggs spawned per female), it has no intraspecific social interactions or a structured environment that would provide triggers for ESD in its natural range. There is no obvious sexual morphological dimorphism or courtship behaviour that would indicate sexual selection. And, neither the karyotype nor the male and female linkage map (28) indicated sex chromosomes.
Despite, we found SD in the Atlantic herring is ruled by a strong genetic determiner: a truncated BMPR1BBY receptor carried by a well differentiated sex chromosome. Interestingly, the herring Y-chromosome should be considered as a minimal differentiated sex chromosome: the male specific region of the Y contains just three genes – a testis-determining gene from the TGF-ß signalling pathway and two other male beneficial genes **[see paper #49]** (Figure 17).

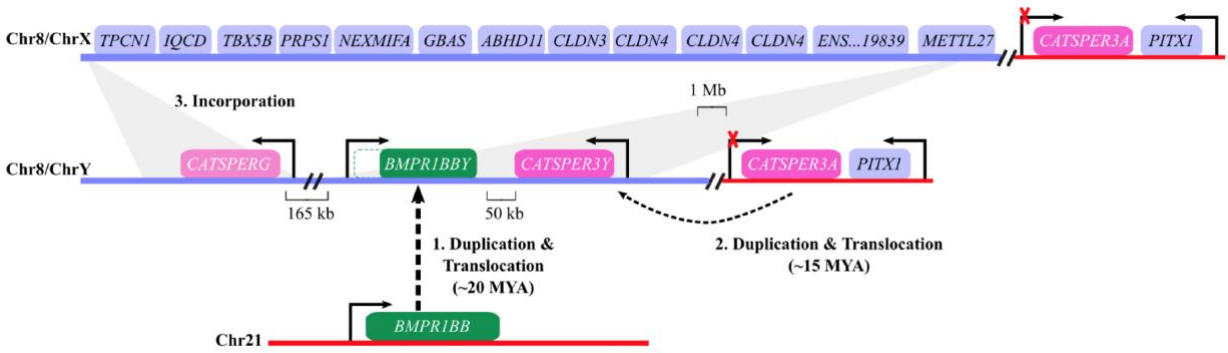


Figure 17. Reconstruction of the evolution of sex chromosomes in Atlantic herring.
The evolution of ChrY occurred in four steps, but the exact order of steps 2 to 4 is not yet known. **1)** Duplication and translocation of BMPR1BB from Chr21 to Chr8. **2)** Duplication and translocation of CATSPER3 within Chr8/ChrY. **3)** Incorporation of CATSPERG in ChrY and loss from ChrX. **4)** CATSPER3A becomes pseudogenized or evolved a new function.

Hence our work revealed that RSD processes might be sporadically resilient, but not perfectly random. Thus, even minimal deviations from optimal sex ratios might initiate evolutionary driftage towards GSD, particularly when population sizes are large and/or natural selection is effective.

Obviously, the evolution of genetic sex determination mechanisms is closely linked to the evolution of sex chromosomes. The discovery of multiple new master SD genes in teleosts provided invaluable insights on the evolution of SD and the processes of the establishment of new sex chromosomes. Before the identification of the first fish master SD gene, models of sex chromosome evolution predicted that SD genes should evolve from a stepwise allelic diversification process starting from a single gene on an autosome. After fixation of a sex-specific allele in that newly formed SD gene, recombination reduction around that new sex *locus* in the proto-sex chromosome pair would give rise to classical heteromorphic sex chromosomes.
Some newly described fish SD genes have evolved thanks to this allelic diversification mechanism, for instance sox3Y (37) and gsdfY (11) in some medaka species and amhr2 in Takifugu (21). But this mechanism did not lead to strongly differentiated heteromorphic sex chromosomes in these teleost species. This phenomenon can probably be explained if we consider that these SD genes are very young in term of evolution.

But the conservation of a single missense SNP in amhr2 acting as the sole signal to initiate sex determination (Figure 12) in at least three Takifugu species that diverged over 10 million years ago, shows that recombination reduction around a new sex locus in the proto-sex chromosome is *de facto* not necessary to fix a new SD gene.

The characterization of new teleost SD genes (Figure 2) also emphasized novel mechanisms for sex chromosome and SD gene evolution as many fish SD genes are the result of a duplication/insertion event. In these cases, a genomic area containing an ancestral gene is duplicated into a different chromosome that will initiate the formation of a new sex chromosome after subsequent functional and/or regulatory changes of the proto SD gene. This mechanism of gene duplication leading to sex chromosome formation first came in light for dmrt1bY in Japanese medaka, and has also been found to generate other teleost –and vertebrates- master SD genes (like in medaka, rainbow trout, European pike, Atlantic herring, Mexican tetra and Yellow perch, for the fish species we have been working with; **[see papers #20, #39, #43, #44, #49]**).
In addition, thanks to their high turnover rate, teleost sex chromosomes also provide many different models to study the process of sex chromosome formation, like for instance in the medaka **[see papers #20, #36, #40]** or Atlantic herring **[see paper #49].**

(ii) Functional analysis of the gene regulatory networks underlying sex determination and differentiation.

While many recent studies dealing with sex determination/differentiation in fish have been focusing on the quest for new master sex determining genes, the genetic architecture of such a complex phenotypic trait as sex cannot be simply limited -or restricted- to the action of unique and totipotent master sex determining triggers. Then, what happens when “masters change”? The classical view of sexual development suggests that not much would change downstream, since “slaves remain” (3). Potentially a new master would pop up at the top from a postulated conserved downstream gene regulatory network, possibly slightly adjusting (bottom-up theory (1, 38)). However, accumulating evidence, notably gathered within teleost fish species, indicates that the phenotypic expression of sex is a rather plastic trait, relying on a complex and unstable equilibrium of a constantly adjusting network of regulatory interactions (**[see papers #28 and #31]** for review). Hence, the main emerging idea is that sex determination gene regulatory cascades should no longer be seen as simply hierarchical but, rather, as a regulatory network or, even more, as connections of interdependent regulatory networks.

To explore more onto that direction, and because the amazing diversity of sex triggers in fishes emphasizes the many options possible at the sex determination stage (and possibly beyond) to switch and supervise over the destiny of the gonad, we have additionally been focusing our researches on the functional analysis of the gene regulatory network(s) underlying sex determination and differentiation. For that, we mainly used (***i***) medaka, which MSD gene arose after gene duplication, and (***ii***) sablefish, which MSD gene arose after allelic diversification. Our main idea is now to progress towards a comprehensive and comparative view of the evolution of sex determining genes and regulatory networks in relation to the emergence or turnover of master sex determining genes.

(iia) The dmrt1bY regulatory network of medaka

Dmrt1 is a highly conserved transcription factor, which is critically involved in regulation of gonad development of vertebrates **[see paper #24]**.

In medaka, a functional duplicate of the autosomal dmrt1a gene on the Y chromosome—dmrt1bY—became the master regulator of male sex determination (8, 9). This duplication event went along with the acquisition of a tightly timely and spatially controlled gonadal expression pattern. Indeed, dmrt1bY mRNA expression is very dynamic, occurring first in the PGCs prior to morphological somatic sex differentiation and then quickly switches to an exclusive Sertoli cell localisation **[see papers #21, #29, #36]**. Importantly, dmrt1bY is expressed in PGCs of male embryos much before its expression in the pre-Sertoli cells at the sex determination stage **[see papers #29 and #36]**. This early PGC expression is necessary for the later onset of dmrt1bY expression in the pre-Sertoli cells at the sex-determination stage of male development **[see papers #36]**. There, the level of dmrt1bY mRNA needs to reach a certain threshold to exert the sex-determining function (Figures 13 and 14). This suggested that medaka germ cells exhibit sexually different characters before the formation of the somatic gonadal primordium depending on dynamic and tightly timely regulated mechanisms of transcriptional/post-transcriptional regulations **[see papers #18 and #29]**. We had previously shown that transcriptional rewiring was brought about by exaptation of two transposable elements, *Izanagi* and *Rex1*, co-opted to act as silencers (see above in this manuscript, and **[see papers #21 and #40]**. These turn off the somatic and the germ cell–specific expressions of the dmrt1bY gene. Thus far, two factors, dmrt1 itself **[see papers #21]** and sox5 **[see papers #36]**, were identified, which turn off dmrt1bY expression after it has fulfilled its function in the early developing gonad.

Further on we identified a *cis*-regulatory 11-bp motif in the 3' UTR of dmrt1bY called the "D3U-box". This motif confers stability to the dmrt1bY mRNA in germ cells of the developing embryonic gonad, whereas in other tissues, the transcript is rapidly degraded **[see papers #18 and #40]**. The D3Ubox motif was found to be highly conserved in the dmrt1 3' UTR in the fish lineage (*O. latipes*, *O. curvinotus* , *Takifugu rubripes* , *Tetraodon nigroviridis* , *Epinephelus coioides* , and *Danio rerio*), as well as in other vertebrates, including *Mus musculus* , *Pan troglodytes* , *Macaca mulatta*, and *Homo sapiens*, and even in the ecdysozoan clade (*Anopheles gambiae* and *Bactocera oleae*), (Figure 18).

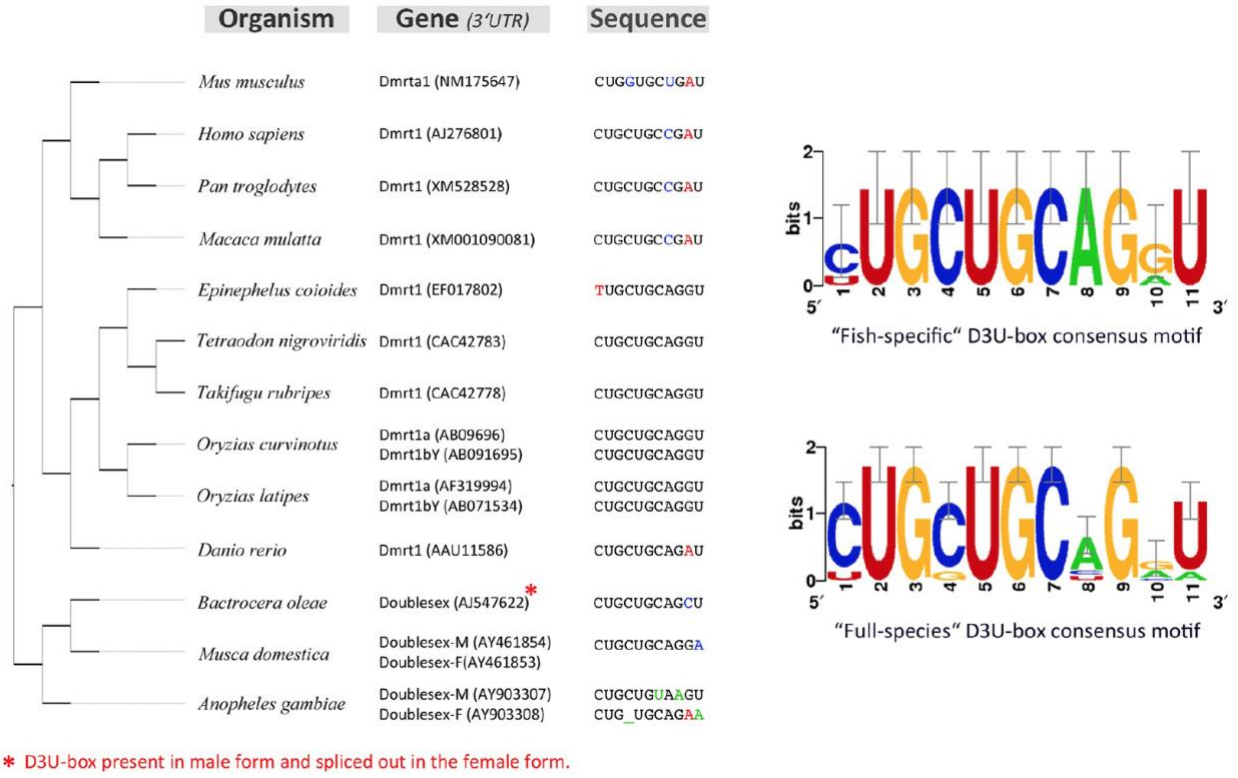


Figure 18. Conservation of the D3U-box motif from ecdysozoans to mammals.
Conservation of the D3U-box motif from *Drosophila* up to mammals. From the different D3U-box sequences among vertebrates, weight-position matrices were deduced and used for genome scans **[see paper #40]**.

Both, phylogenetic conservation together with the additional presence of the D3U-box in several germ cell-specific transcripts implied the existence of similarly conserved *trans*-acting factor(s) involved in the synexpression of those transcripts. To identify such factor(s), we undertook an unbiased approach centred on the D3U-box sequence and based on the evolutionary conservation of the 'split' motifs of the D3U-box, implying evolutionary conserved *trans*-acting factors. Further bioinformatics analyses and literature searches revealed that the D3U-box motif is a putative target for two RNA-binding proteins, namely cug-bp (39, 40) and bsf (also known as lrpprc in mammals (41–43). EMSAs indicated that Ol-bsf and Ol-cug-bp1, but not Ol-cug-bp2, specifically target and interact with the different parts of the D3U-box, the 3'and the 5' parts, respectively **[see paper #40]**. Additionally, our results suggest that the observed regulation of dmrt1bY transcript abundance is likely to be the result of a differential binding of the two RNA-binding proteins (Ol-bsf and Olcug-bp1) with antagonistic properties, *trans*-regulating RNA stability *via* the D3U-box (Figure 19), **[see paper #40]**.

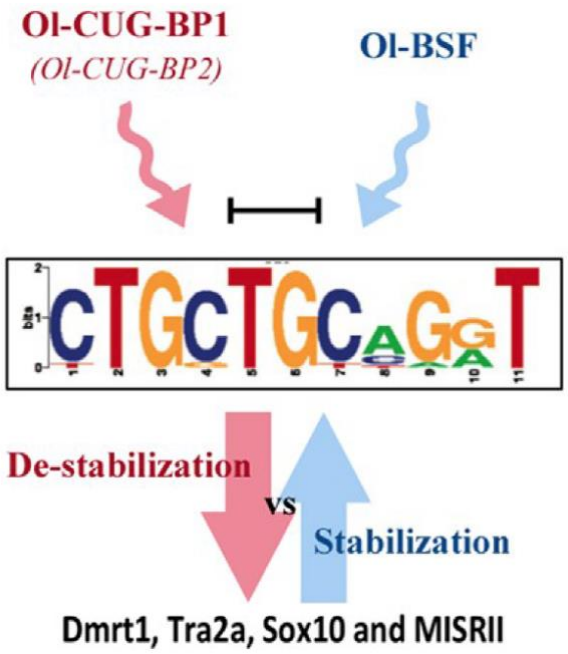


Figure 19. Model for D3U-box–mediated mRNA regulation.

Overall and in addition to a cytoplasmic localization of Ol-bsf, Ol-bsf and O-cug-bps might mutually antagonize toward the access to the D3U-box, resulting in either stabilisation (more Ol-bsf binding) or destabilisation (more Ol-cug-bp binding) of the transcripts harbouring the D3U-box.

Hence, using complementary approaches, our data suggests that the D3U-box motif is—depending on the cellular context—targeted by two antagonizing RNA binding proteins, promoting either RNA stabilization in germ cells or degradation in the soma. This new mechanism of dmrt1 RNA stability appears to also regulate the abundance of other transcripts specifically expressed in PGCs, depending of the preservation of the D3U-box motif.

Altogether the data we could gather in medaka allows to decipher and reconstruct the different evolutionary steps that were necessary to establish –and preserve- *dmrt1bY* ahead of the gonadal gene regulatory network after duplication (see Figure 20).

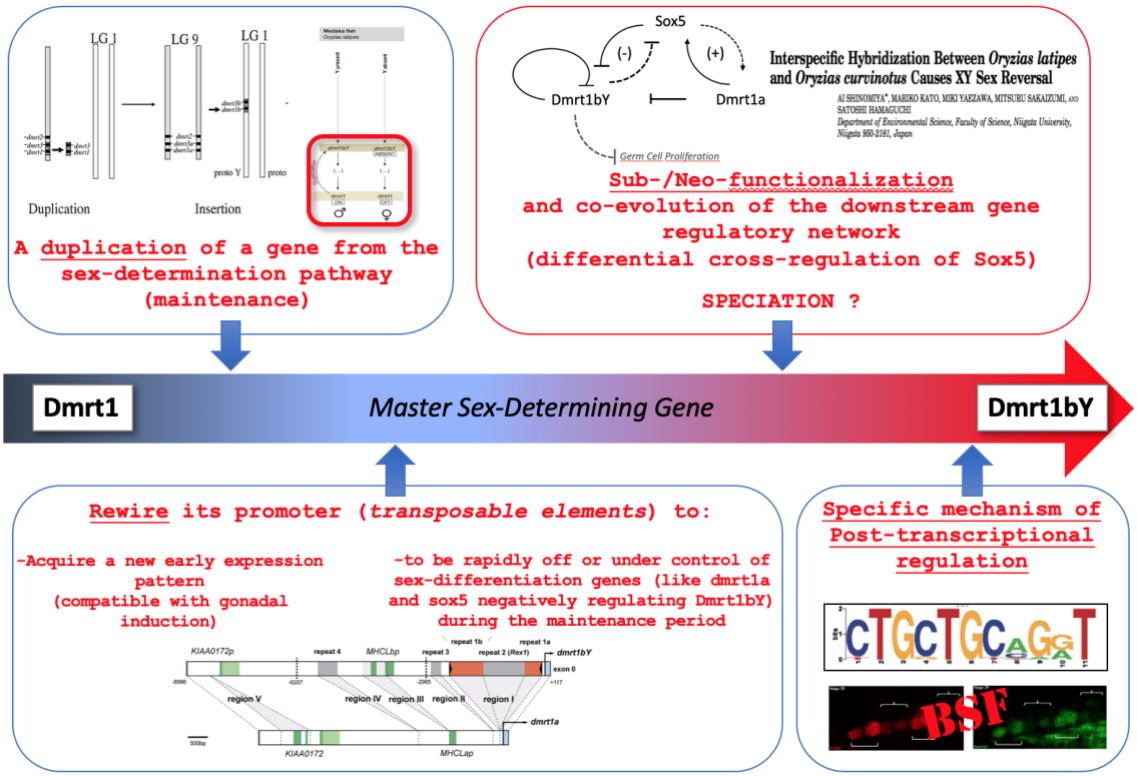


Figure 20. How Dmrt1bY was born in medaka?

This scheme depicts the different evolutionary steps that were necessary (and sufficient) to establish and preserve *dmrt1bY* ahead of the gonadal gene regulatory network in medaka. First a duplication/insertion of a gene (dmrt1) from the sex-determination pathway (maintenance) occurred. In a second step the region of insertion (proto Y) was kind of "corrupted" by insertion of transposable elements. For instance, in medaka this resulted in a transcriptional rewiring of the duplicated copy ((dmrt1bY) that led (***i***) to the acquisition of a new pattern of expression compatible with gonadal induction and (***ii***) to be rapidly off or under the control of sex-differentiating genes (like dmrt1a or Sox5). In parallel a specific mechanism of post-transcriptional set up, resulting in a tightly timely and spatially controlled gonadal expression pattern. Finally sub/neo-functionalization together with co-evolution of the downstream gene regulatory network processes developed.

(iib) Allelic diversification after transposable element exaptation into its proximal regulatory region promoted Gsdf as the master sex determining gene of sablefish (Anoplopoma fimbria)

New MSD genes primarily emanate from one of two evolutionarily conserved processes: (***i***) sporadic gene duplication and insertion, followed by sub- and/or neo-functionalization, or (***ii***) allelic diversification of a pre-existing *locus* (20) **[see BC #1 and BC#4]** (Figure 12). However, the molecular changes that allow new MSD genes to exert a novel function are not well understood, aside from those of a few model species.

In the marine teleost, sablefish (*Anoplopoma fimbria*), complementary genomic and genetic studies recently led to the identification of a sex *locus* on the Y-chromosome (44). Further characterization of this *locus* resulted in identification of the transforming growth factor β (TGF-β) gene, *gonadal soma-derived factor* (*gsdf*), as the main candidate for fulfilling the MSD function. The presence of different X- and Y-chromosome copies of this gene indicated that the male heterogametic (XY) system of sex determination in sablefish arose by allelic diversification.

In sablefish, *gsdfY*, in contrast to its X-variant counterpart, is specifically expressed in male (XY-genotype) fry earlier than any other male or female sex-related genes and prior to both molecular and morphological sexual differentiation of the gonads (45). Comparative analysis of the *gsdfX* and *gsdfY* expression patterns clearly showed that *gsdf-Y*, which is expressed much earlier than *gsdfX*, experienced transcriptional rewiring during the process of allelic diversification that ultimately gave rise to the Y- and X- chromosomes of sablefish.

Such acquisition of a new transcriptional context represented by a different spatio-temporal expression pattern, compatible with a sex-determining function, seems to be the main prerequisite in the process of establishment and fixation of a new MSD gene. In the two sister species, *Oryzias latipes* and *O. dancena*, and in mammals, either *dmrt1* or *sox3* genes respectively, were subjected to profound transcriptional rewiring for establishing either *dmrt1bY* (*O. latipes*; duplication / insertion, (46–48) (see Figure 12)), sox3Y (*O. dancena*; allelic diversification, (37)) or SRY (most mammals; allelic diversification, (49)) as master sex-determining genes respectively (see (50) for review, and Figure 21).

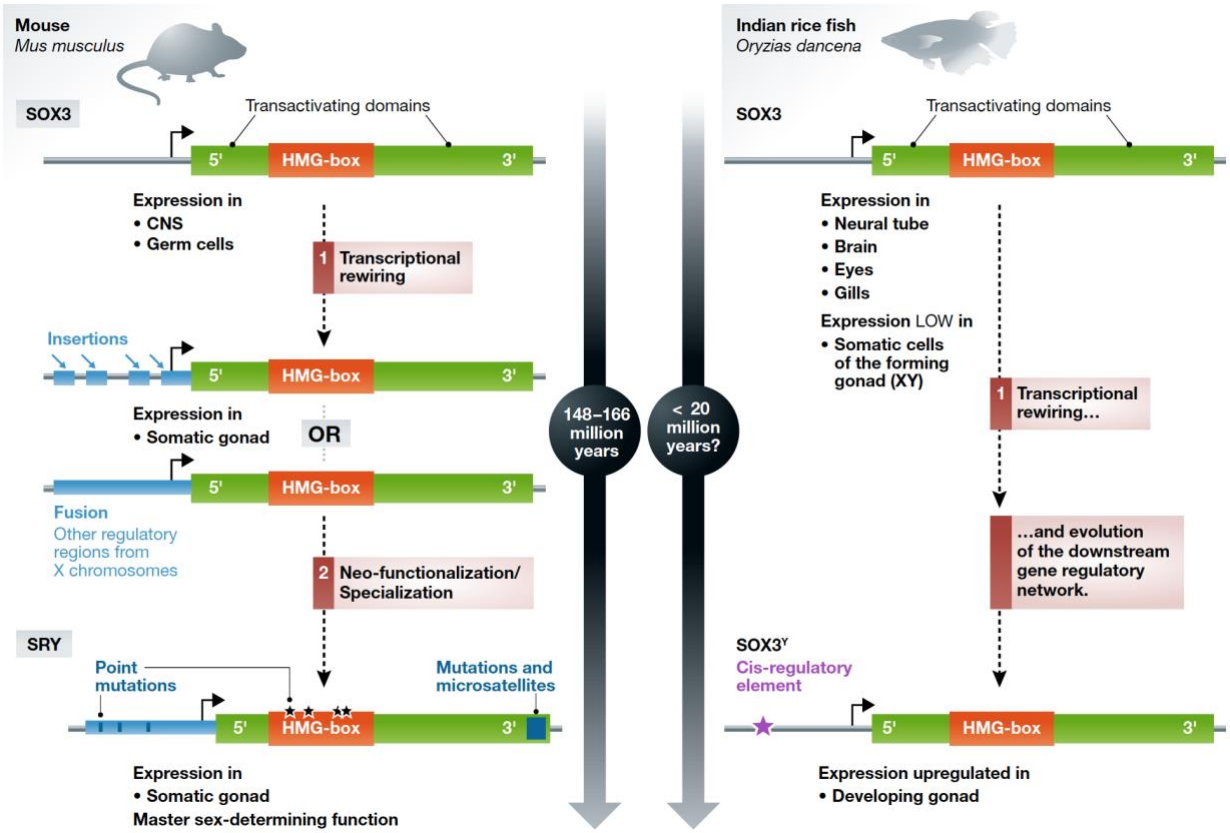


Figure 21. Independent evolution of SOX3 genes toward a master sex-determining function in mice and Indian rice fish.

While SRY appears to be restricted to the therian mammals, evidence accumulates that SOX3 has independently been recruited as a "precursor" of master sexdetermining genes also outside mammals. Hence, although not a priori destined to have a direct function during sex determination, common mechanisms of evolution seem to be repeatedly employed. Given that SOX3 is not generally expressed during gonadal induction or during gonadal development, the first step toward a sexdetermining function is a transcriptional rewiring in order to acquire a timed pattern of expression compatible with sex determination. Such transcriptional rewiring, although not unique to SOX3 (see Dmrt1bY in medaka fish for example, (Figure 12)), generally involves either fusion of the gene to new promoters or insertions of transposable elements into their pre-existing promoter, bringing in cis-regulatory elements compatible with the timing of gonadal induction. Interestingly and surprisingly, it seems that at least in mice and rice fish, this step alone was sufficient to endow SOX3 with a sex-determining function. Usually, the transcriptional rewiring steps seem to be accompanied by neo-functionalization or functional specialization processes. These include specialization of the protein activity itself in therian mammals (adapted from reference (49)) or more surprisingly adaptation of the downstream gene-regulatory network (target genes) in the Indian rice fish.

In sablefish, while a unique missense mutation between the two *gsdf* variants does not appear to drastically impact their physiological activity with regard to downstream activation of Smad's, any processes of functional divergence of the variants after allelic diversification might be reasonably excluded. Uniquely, the newly acquired MSD function of the *gsdfY* gene seems to be entirely ascribable to its new pattern of expression.

Into that direction we reported that the high expression of the *gsdfY* allelic copy during gonadal differentiation is largely imputable to a Y-specific insert derived from a TE of the hAT family incorporated into its promoter (Y-specific insert in Figure 22). Gsdf is an important downstream

component of the male sex determination regulatory network, which in sablefish, like in *Oryzias luzonensis* (11) acquired the role of the MSD. Interestingly, co-option of this TE within the *"neo"-gsdfY* promoter was likely sufficient for transforming and elevating a protein acting downstream in the sex determination network to a MSD gene expressed at the right time and right place.

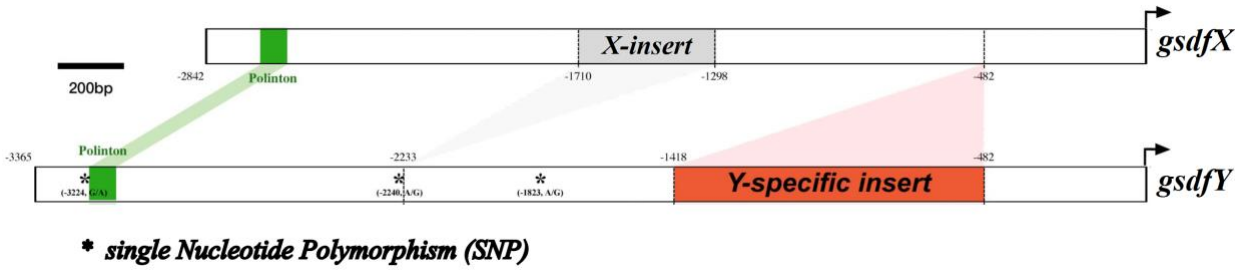


Figure 22. Comparative analysis of the *gsdfX* and *gsdfY* promoters.

The analysed promoter region of *gsdfX* in comparison to its *gsdfY* paralog. Length differences between the two *gsdfX* and *gsdfY* promoters are due to Y- and X- specific regions of which unique Y- (935 bp) and X- (412 bp) specific inserts have been respectively added and lost concomitantly during the allelic diversification event. The Y-specific insert is made of a transposable element of the hAT family.

Deciphering the mechanism by which the *neo-gsdfY* is transcriptionally controlled, we could further show that the hAT-type TE facilitated early up-regulation of *gsdf-Y* expression by Dmrt1 and Wt1 (Figure 23), two key genes of the canonical gonadal gene regulatory network, preventing any expression pattern redundancy between the two *gsdf* allelic copies, that therefore might constitute a reasonable evolutionary way for preserving both *gsdf* gene copies from any purification /degeneration processes after allelic diversification. Altogether our results demonstrate that allelic diversification of the *gsdf* gene gave rise to the sex determination system in sablefish. Importantly, the MSD function of *gsdfY* was not attained by acquisition of a new function of the protein itself, *via* amino acid changes, but rather through to the acquisition of elements in the promotor region resulting in a unique expression profile, which relocated *gsdfY* to the most upstream position in the sex-determining network.

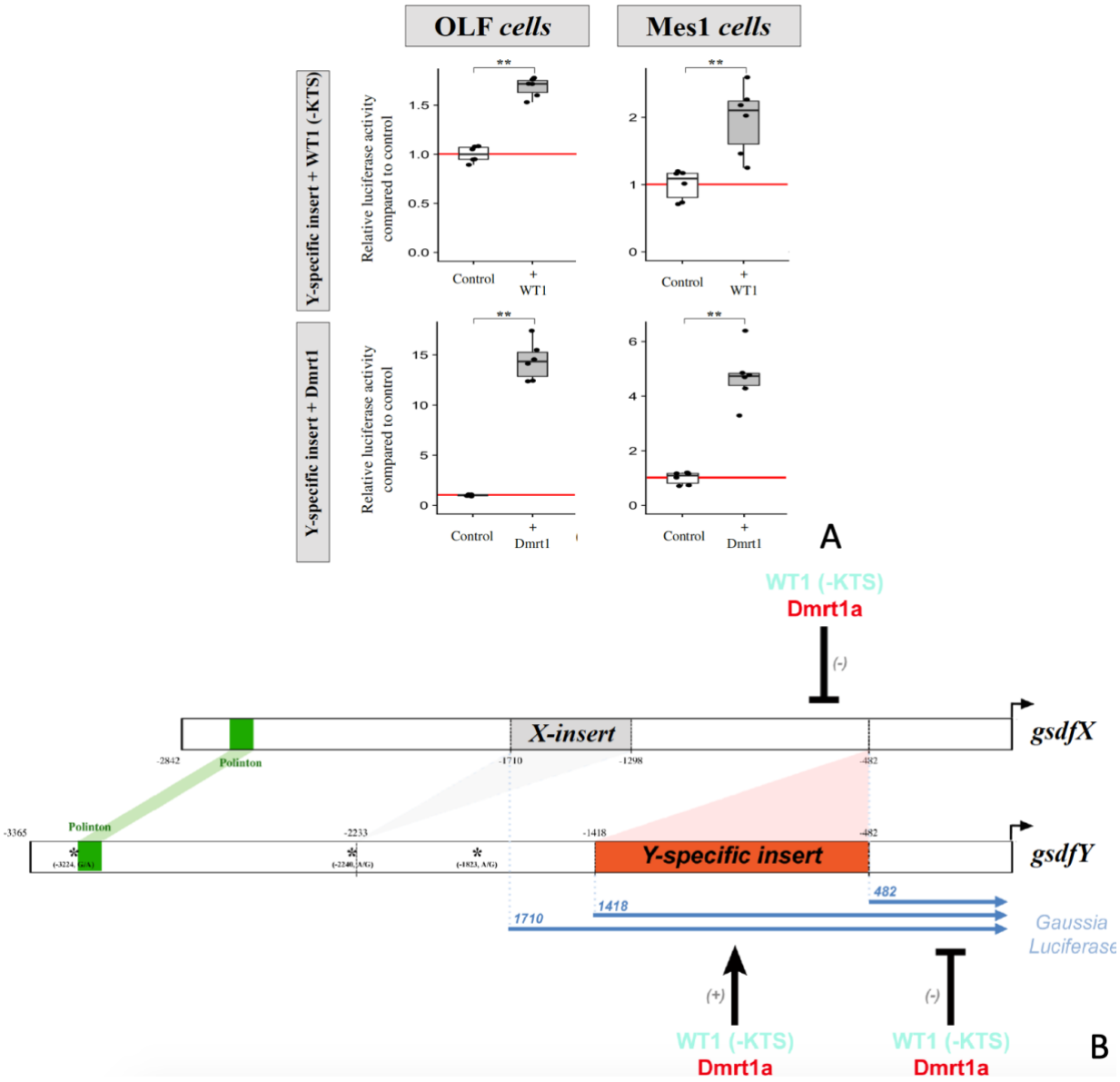


Figure 23. hAT-mediated male-specific transcriptional regulation of *gsdfY* promoter by Dmrt1 and Wt1.

(A) The hAT element (Y-specific insert) induces modulation of transcriptional activity of a thymidine kinase minimal promoter upon transient transfection of Dmrt1 or Wt1(-KTS). **(B)** Model for *gsdfX* and *gsdfY* transcriptional modulation by Dmrt1 and Wt1.

Evolution of new sex determination genes by allelic diversification has often intuitively been associated with gradual processes that proceed slowly over evolutionary timescales (34, 51). We found that in *A. fimbria*, allelic diversification of a sex determination gene initiated by the exaptation of a TE led to complete transcriptional rewiring of the allele on the proto-Y chromosome. This provides a unique functional example of a *bona fide* punctual process as an efficient alternative to the phyletic gradualism model (52) for the molecular evolution of a master sex determining gene.

(iii) Germline/soma interactions, gonadal morphogenesis and plasticity

(iiia) The autosomal gsdf gene acts as a male sex initiator in medaka fish.

As much as possible, we always have been trying to correlate/connect regulatory mechanisms and evolution with physiology. It is then not surprising that we also got interest for gonadal morphogenesis, notably *via* studying the function of the genes for which we studied regulation. I think this is, in the field, our specificity (strength?) in the way that we are able to report comprehensive "stories" about evolution of master sex determining genes, blending genomic plasticity, phylogeny, transcriptional and post-transcriptional regulations, and up to the physiology *in fine*.

Gonadal development is a multi-step process including sex determination, initiation, differentiation, maintenance, and culminating with the production of sperm or eggs for germline transmission. Like seen before, a hallmark of genetic mechanisms underlying sex determination and development is that they show a remarkable diversity and do not follow any evolutionary trajectories (Figure 11). For example, even master sex-determining genes or sex determiners (SDs) so far identified in different animal taxa (Figures 1 and 2) show considerable differences in sequence and activity of their products. The primary role of an SD is to determine the initial sex by triggering testicular or ovarian differentiation of a sexually bi-potential gonad. Hence, the presence of a SD determines the genetic sex, whereas the onset of gonadal differentiation towards a testis or an ovary delineates primary sex initiation. SDs act at the top of hierarchical networks to control sex differentiation. For example, sry in mammals initiates testicular differentiation through activating its direct target sox9 (Figure 3). Notably, we have previously shown that the networks downstream of SDs also vary enormously from one animal to another (Figure 9) **[see paper #28]**. The enormous diversity of genetic sex determination mechanisms is a long-standing mystery and also a major challenge for understanding sex development. Fish have sex-determination mechanisms ranging from environmental to different modes of genetic determination and thus provide a paradigm for studying sex plasticity and development. Particularly, medaka fish is the first vertebrate that showed crossing-over between X and Y chromosomes (53), induction of sex reversal (54), and most importantly, offered the first vertebrate SD besides the mammalian sry, namely dmrt1bY (8, 9). Most recent studies have revealed female germ stem cell markers capable of making intrinsic sperm-egg fate decision in medaka (55). It is known that dmrt1bY activates dmrt1, which in turn maintains testicular differentiation, as dmrt1 mutation causes male-to-female sex reversal after the initiation of testicular differentiation (56). However, beside mediating a mitotic arrest of PGCs in males prior to testes differentiation **[see paper #14]**, how dmrt1bY exerts its primary role in male decision *via* triggering testicular differentiation remains unknown.

Paradoxically, there are several cases where dmrt1bY is dispensable for maleness (9, 57), which points to the presence of autosomal essential gene(s) for male sex initiation in medaka (57, 58). Recently, the gsdf gene has emerged as a novel sex-related factor in several distantly related fish species (including in the sablefish, like discussed above). This gene (59) encodes the gonadal soma derived factor, which belongs to the transforming growth factor-β superfamily (60). In medaka, gsdf is located on chromosome 12 and is predominantly expressed in the Sertoli cells and granulosa cells in mature gonads (61).

Beside suggesting its potential role as an endogenous inducer of gonadal development, our studies have additionally provided sufficient evidence to firmly established gsdf as a male sex initiator acting downstream of dmrt1bY in medaka, thus serving as a prime candidate for the searched autosomal gene essential for maleness **[see paper #32]**: -First, gsdf addition is sufficient for masculinization in the absence of dmrt1bY, and gsdf disruption causes feminization without compromising dmrt1bY expression, pointing to its hypostasis to dmrt1bY in action. -Second,

feminization upon gsdf disruption results from the initiation of ovarian differentiation despite the presence of dmrt1bY, demonstrating that the earliest and perhaps the primary role of gsdf is to initiate testicular differentiation, which cannot be replaced by dmrt1bY. -Finally, the fact that dmrt1bY protein binds to the gsdf promoter and activates gsdf transcription convincingly reveals that gsdf mechanistically acts downstream of dmrt1bY (Figure 24, **[see paper #32]**).

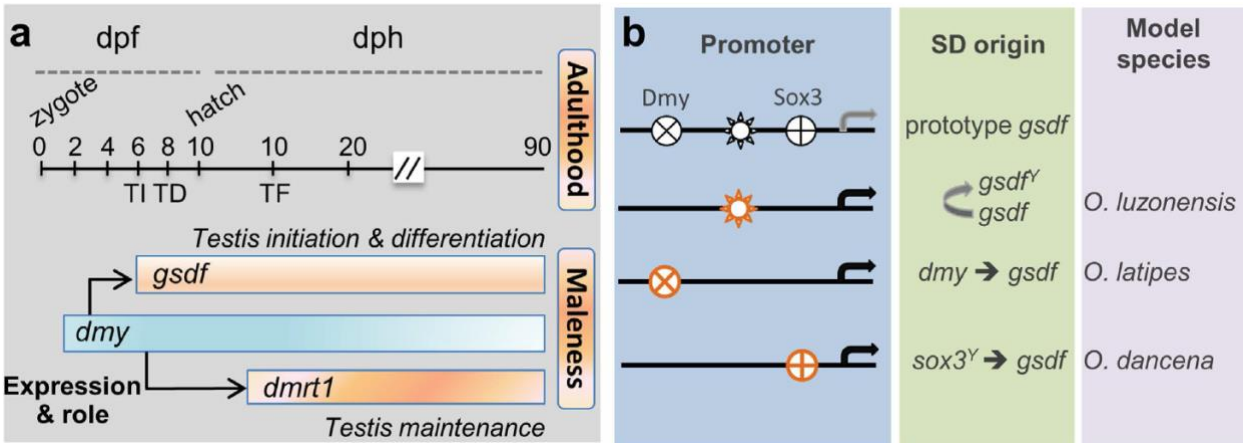


Figure 24. Sex determination and its evolution in Oryzias.
(A) The core cascade controlling sex development in medaka. The black horizontal line depicts developmental day(s) post fertilization (dpf) or hatching (dph). Shown are key developmental events including testis initiation (TI), testis differentiation (TD) and testis formation (TF). Horizontal boxes depict developmental RNA expression stages of dmrt1bY, gsdf and dmrt1. Arrows depict transcriptional activation of gsdf and dmrt1 by dmrt1bY. Indicated are the primary roles for gsdf in testis initiation and differentiation and for dmrt1 in testis maintenance. **(B)** Hypothetical evolution of sex determiners in the genus Oryzias. gsdf may easily become a sex determiner (namely gsdfY in *O. luzonesis*) *via* acquiring proper temporospatial expression (bent arrow) or preferentially recruit gsdf-regulating genes as new sex determiners (namely dmrt1bY in *O. latipes* and sox3Y in *O. dancena*). Shown are potential or putative binding sites (black) for Dmrt1, Sox3 and an unknown transcription factor in the prototype gsdf promoter. These binding sites may be used (orange) by transcription factors such as Dmrt1 and Sox3 for activation (horizontal arrows) of proper temporospatial gsdf expression to initiate testicular differentiation **[see paper #32]**.

Additionally, gsdf disruption causes the alteration of global gene expression and dmrt1 down-regulation in adult gonads, suggesting the potential involvement of gsdf in subsequent processes such as male sex maintenance besides acting as a male sex initiator (Figure 24, **[see paper #32]**).

(iiib) Of the pronephric contribution to the differentiating gonads in fish (focus on medaka).

Still aiming at deciphering how gonad morphogenesis is induced and sets up after the different master sex determining genes have fulfilled their function(s), we are currently addressing the question of a possible contribution of the pronephros in the differentiation of the teleost fish gonads (**[see paper #17]** and ongoing work). Although so far not investigated in fish, this question is of prime interest since pronephric contribution is a major component of gonadal differentiation in other vertebrates (62, 63).
Critical to our understanding of sex-determination processes are studies investigating the origin and development of cells involved in the formation of the primordial gonad. Despite the fact that different determinants and plastic gene regulatory networks **[see papers #28 and #33]** trigger gonadal formation amongst vertebrates, the adult gonads are however very similar in morphology,

cellular organization and physiology. In contrast to the ambiguities and complexities of the developing gonad in mammals and birds, the histological structures present during the formation of the gonadal primordium are relatively simple in medaka fish. The gonad of medaka is formed by the coordinated development of two different cell lineages: the germ cells and the somatic gonadal mesoderm surrounding the germ cells. Once specified, PGCs remain closely associated with endodermal tissues and migrate *via* the dorsal gut mesentery to the region of presumptive gonad (Figure 6 and **[see papers #9 and #15]**). Somatic cells of the presumptive gonad have an embryological origin distinct from PGCs. The external somatic layer of the developing gonad is derived from the genital ridge epithelium, but some somatic cells are probably also derived by the invasion of mesenchyme (similar to the primitive sex chords in mammals, (64)). However, unlike mammals, no clear indication of medullary tissue can be found in the teleost gonad. Prior to differentiation, all somatic cells appear to be derived from a cortex epithelial layer, and are similar in presumptive males and females (64). Following migration of PGCs into the germinal ridge, cell division occurs to form oogonia and spermatogonia, and differentiation and migration of somatic cells is initiated concomitant with this process (**[see papers #9 and #15]**. Shortly before hatching, at the time of expression of *dmrt1bY* in the male gonad primordium (Figure 14), the germ cells in the female gonad actively proliferate and undergo meiosis, while this is not observed in male gonads (65). It is only ten days later that the first somatic gonadal dimorphisms are apparent with the formation of the acinus (the seminiferous tubule precursor) and the follicles in gonads of male and female respectively (65). Interestingly, ovarian cords within the germinal epithelia of medaka ovaries have been recently characterized. These cords, composed of somatic *Sox9b*-expressing cells and mitotic *nos2*-expressing oogonia continually give rise to germ cells and form a stem cell niche within the ovary referred to as germinal cradle (66). These cradles, containing germline stem cells contribute to the production of fertile eggs and are reminiscent of the germanium of the *Drosophila* ovary, hence implying fundamental processes governing oogenesis across animal species to be conserved.

In mammals, anlagen of gonads, or genital ridges, are formed as two thickenings of coelomic epithelium protruding into the coelomic cavity and situated near the mesonephros, laterally to the dorsal mesentery. In most species, the genital ridges initially consist of somatic cells derived exclusively from coelomic epithelium, and later primordial germ cells (PGCs) invade this preformed gonadal mesoderm. Cells that delaminate from the coelomic epithelium seem to provide one source of cells for the growing genital ridges, while recruitment of underlying cells from the mesonephros to the epithelial population also augments the cell population in the gonadal primordium in males (62). Later, as the male gonad differentiates into a testis, the mesonephric duct develops into the Wolffian duct, and is also thought to contribute to the rete testis and, in the female to the rete ovarii. During that process of mammalian gonad differentiation, mesonephric cells make a substantial contribution to the structure of the ovary or testis itself (62). Hence in mammals a mesonephric origin for both Sertoli cells and interstitial cells, including peritubular myoid and Leydig cells, has been first postulated and later clearly demonstrated (63). Nevertheless, the source of medullar cell precursors in gonads remains a controversial aspect of vertebrate gonadal development.

Interestingly, in amphibians, while the mainstream studies report the origin of medullary cells from proliferating coelomic epithelium surrounding the gonad, suggesting that cells of the gonadal medulla are equivalent to Sertoli cells in the testis and follicular cells in the ovary (55–58), others point to the mesonephric blastema or mesonephric tubules as the origin of medullary cells (71–76), or even the interrenal blastema, constituting the primordium of interrenal glands homologous to the cortex of the adrenal glands in mammals (77, 78). While teleost gonads have many features in common with those of other vertebrates, the general understanding is that teleosts gonads only originate from one primordium, the cortex or peritoneal wall (79). Hence, unlike other vertebrates,

the early gonad does not –or would not- have any equivalent of the medulla that is derived from the interrenal or mesonephric blastema. However, some of our data using BAC transgenic fluorescent reporter medaka lines challenge this view and seriously point to a possible pronephric contribution to the gonad in medaka and consequently to other teleosts as well.

Indeed, while analysing the early expression patterns of an array of specific gonadal markers (BAC transgenic reporter lines for Dmrt1bY, Dmrt1a, Follistatin, Aromatase, Bicoid stability factor…) we observed that all these markers, prior to be expressed in the forming gonads, were firstly expressed within the pronephros (co-expression with Cdh17 a pronephros only marker, Figure 25) at hatching stage when the gonad is forming. Later on, from hatching stage and up to twenty days post hatching, the expression of these markers was progressively lost within the pronephros while on the other hand gradually rising in the forming somatic gonad. Therefore, the question relative to a possible pronephric contribution to the gonad formation conserved across vertebrates was legitimately addressed. In other words, does this apparent switch in expression of our markers between pronephros and gonad during the course of early gonadal formation result from two independent dynamics of expression or from actively migrating cells from the pronephros towards the forming gonad (Figure 25)?

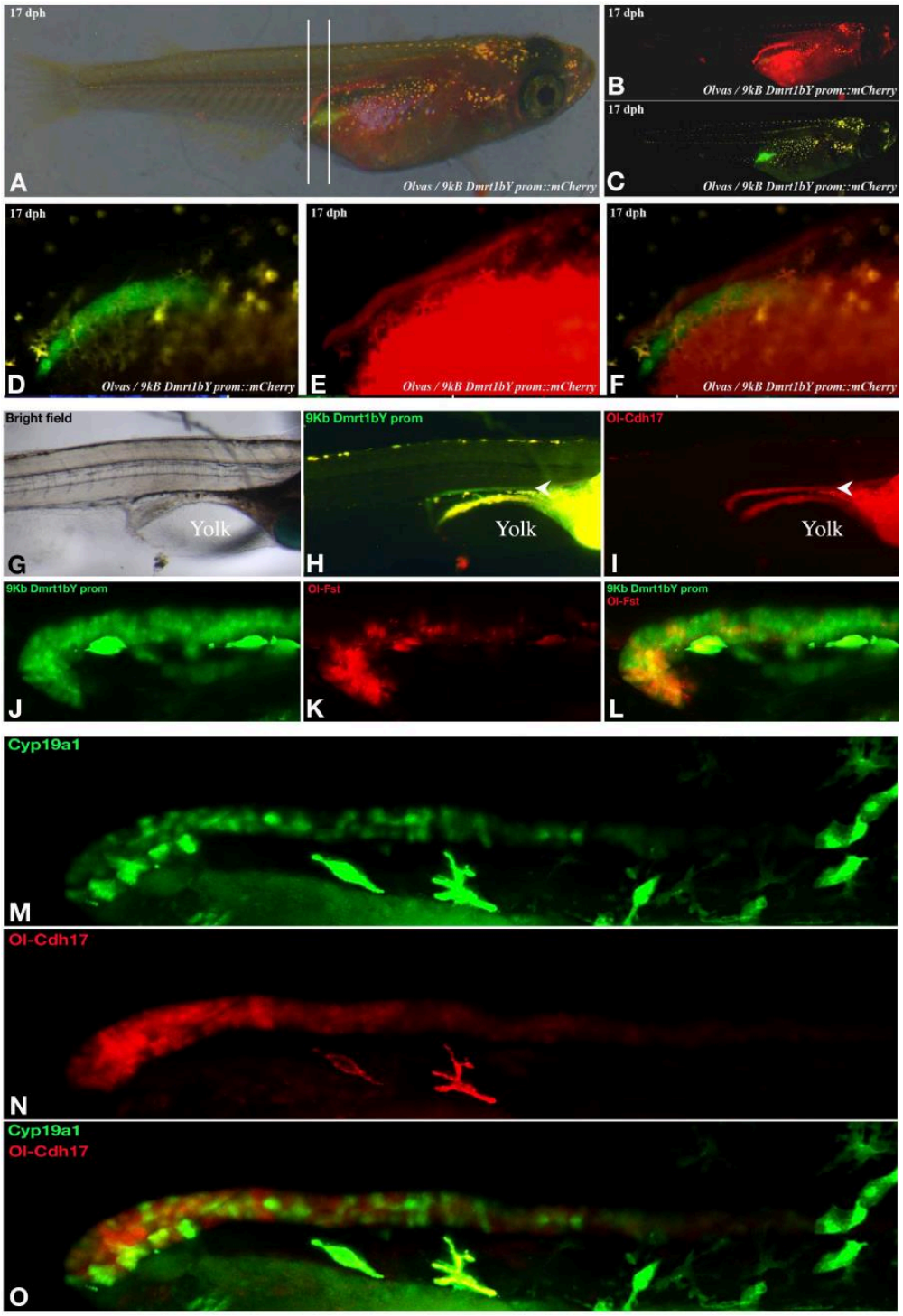


Figure 25. Examples of gonad-specific markers firstly expressed in the pronephric precursors during early primordial gonadal induction.

(A to I) dmrt1bY expression. **(J to L)** follistatin expression. **(M to O)** aromatase expression.

Ol-Cdh17 (medaka Cadherin 17), red fluorescence is a specific marker of the pronephros (80).

To address the question of a possible active pronephric contribution to the gonad in medaka and more generally in teleost fish, we next made use of the "brainbow" system to perform *in vivo* specific cell lineage tracing of pronephric precursor cells. Figure 26 gives an overview of the method that was already successfully used in mice (81) and already adapted for that precise application in zebrafish (82) and medaka (83).

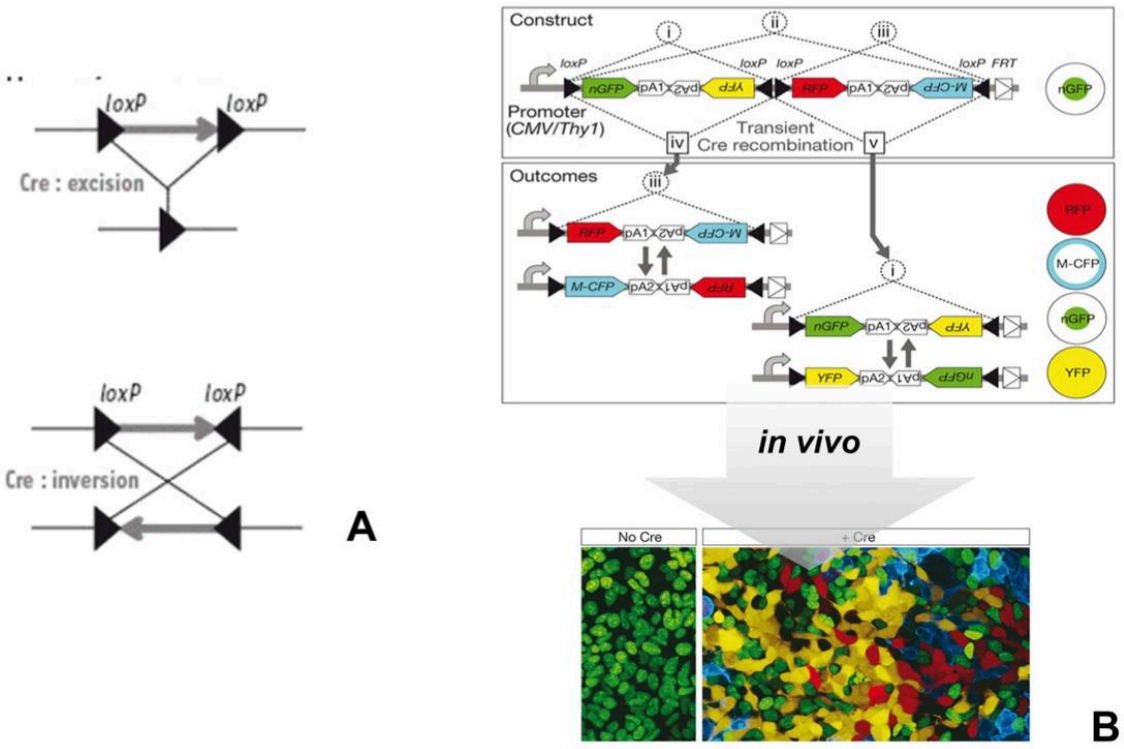


Figure 26. The multicolour "Brainbow" system: stochastic recombination using Cre-mediated excisions and inversions.

The brainbow system uses random Cre/Lox recombination to create varied combinations of red, blue and green fluorescent proteins in each cell. The differences in colour allow users to follow multiple cells, regardless of how closely they are positioned. For cell lineage analysis Cre and brainbow cassette promoter specific driven expressions allow to mark neighbouring clones with distinct colours for further tracing during development. **(A)** The *Cre/LoxP* system. *Cre* recombinase catalyses recombination between LoxP sites. Depending on LoxP site orientations Cre expression triggers intermediate DNA fragment excision or inversion. **(B)** The "Brainbow" system. The Brainbow 2.1 construct contains two tandem invertible DNA segments. Inversion (i-iii) and excision (iv, v) recombination events create four expression possibilities. In medaka we developed that system together with L. Centanin (EMBL, Eidelberg, Germany, see collaborations).

Briefly, a first transgenic line specifically expressing the Cre recombinase in the pronephros at hatching stage has been created. Such a line expresses a fluorescent Cre-recombinase driven by Cdh17 promoter, and show specific expression in the pronephros only throughout development. In a second step, this Cdh17:Cre line was crossed with another line expressing the "brainbow" cassette ubiquitously (ubiquitin promoter, (Figure 27)). After Tamoxifen administration (in order to nuclear localize the recombinase) and subsequent induced specific recombination exclusively in the pronephric cells, cell tracing is activated (recombination and fluorescent switch, Figure 27). It is then possible to live track any migration or commitment of the pronephric precursor cells during development.

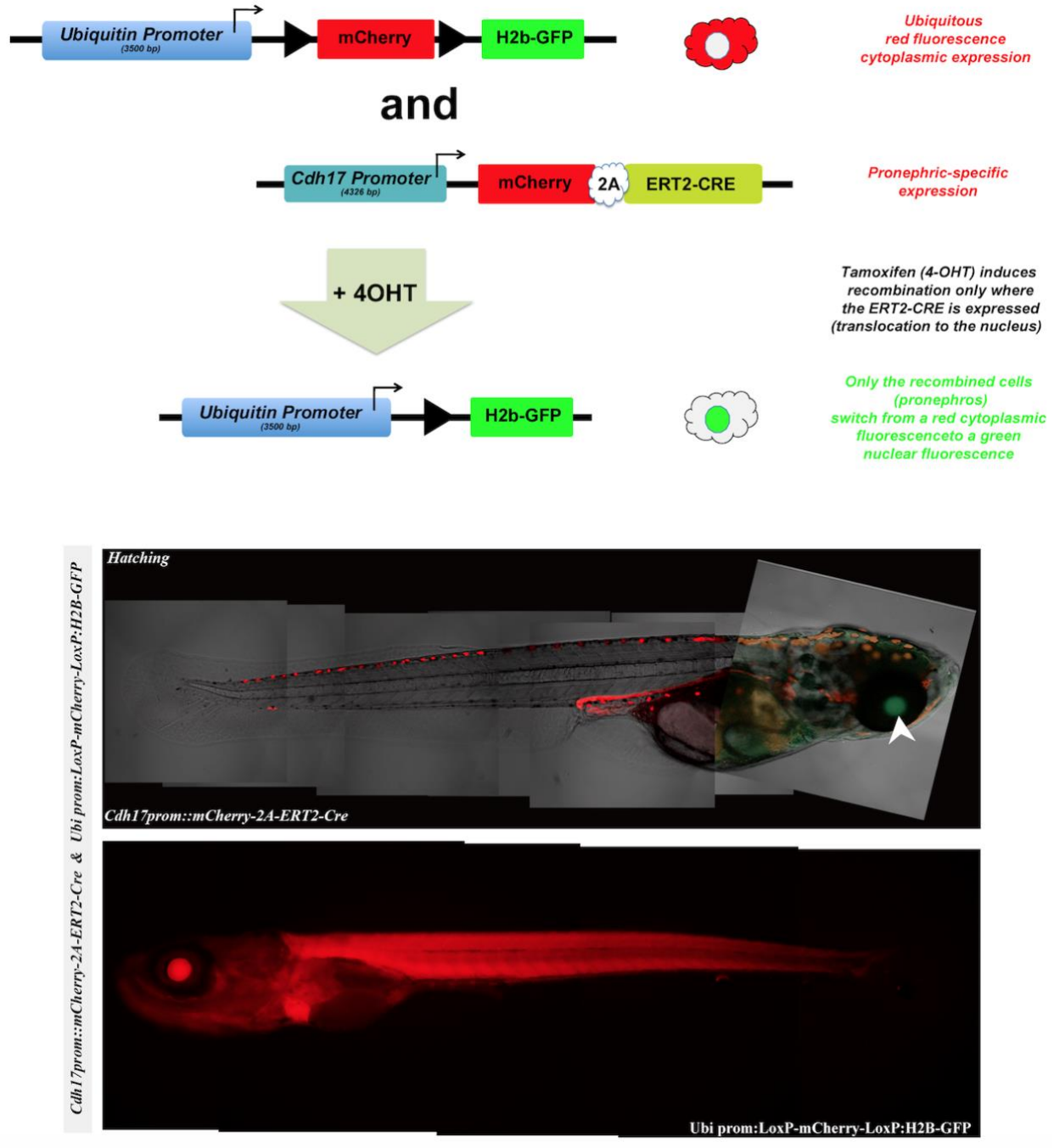


Figure 27. Cre-mediated excisions and inversions in medaka.

Example of strategy for tracking the contribution of pronephric progenitor cells to the adult gonad in medaka: first a reporter line expresses the "brainbow" cassette driven by an ubiquitin promoter (Red expression). A second line expresses an inducible Cre recombinase driven by Cdh17 promoter (pronephros-specific). After Tamoxifen treatment and translocation of the Cre to the nucleus, the recombination will occur, hence switching from red to green expression. Cell lineage tracing of the subset of cells that recombined is now possible.

Interestingly, following specific recombination (cytoplasmic red to nuclear green colour switching) in the pronephric cells at hatching stage, we were able, as early as 8 days post hatching and up to adulthood, to identify these very same cells in gonads of both sexes (Figure 28). This strongly accredits an early and effective pronephric field contribution to the gonad in medaka.

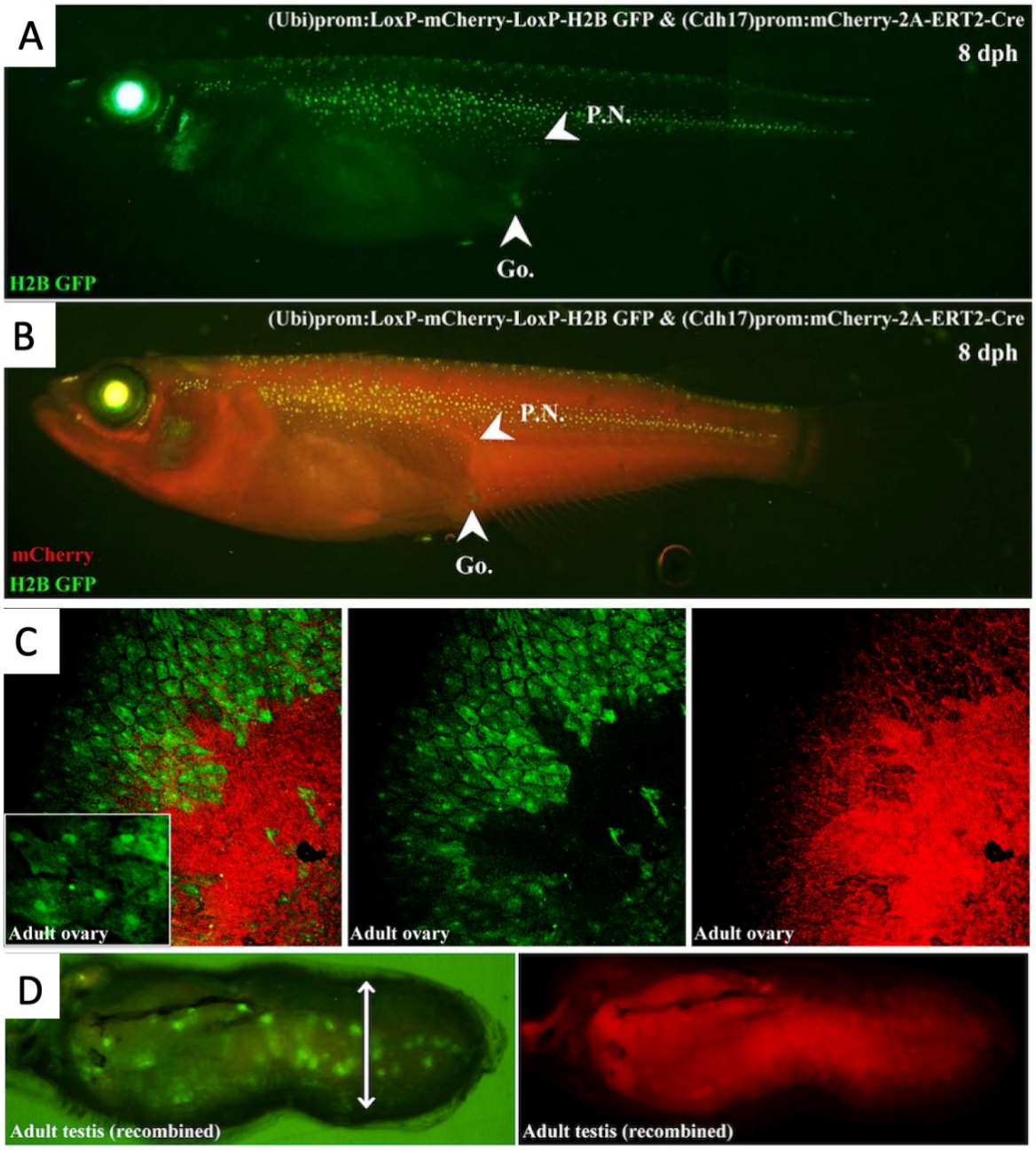


Figure 28. Cell lineage of pronephric cell in medaka.

Cell tracking of recombined pronephric cells at hatching stage allowed to identify cells of pronephric origin in gonads of both sexes as early as 8 dph **(A and B)**, in adult ovaries **(C)**, as well as in adult testes **(D)**.

For the functional part of this study two complementary approaches will be conducted. The first one relies on pronephros development interference *via* morpholino knockdown experiments. For these experiments expression of Cdh17, a gene we have shown to be essential for the pronephros formation, will be morpholino-knockdown with the perspective of interfering with gonadal development in case of an effective pronephric field contribution. In that case, gonadal development will be conjointly monitored through dmrt1bY expression (fluorescent transgenic line). The second one will rely on mapping single cell transcriptomes with the tissue context of the pronephros (Visium Spatial Gene Expression, 10X Genomics) in order to potentially identify different cell subtypes/populations/lineages within the pronephros at hatching stage.

[iv] Autophagy in fish

Few years ago, we published a paper entitled "Defective autophagy through epg5 mutation results in germ plasm and mitochondria failure during spermatogenesis" **[see paper #30]**. When I came to INRAE, I was contacted by a colleague who wanted to benefit of our experience in transgenesis and genome editing (…and autophagy) for his own project dealing with autophagy in fish (Iban Seiliez, INRA nutrition métabolisme et aquaculture, NuMeA). What was at first initiated as a punctual co-operation turned out to be a more serious and real collaboration. We then got fundings for it, students (two PhDs), and now, although this topic is not my core business, I am having pleasure at doing it. Following are some lines to contextualize that (recreational) work.

Chaperone-mediated autophagy (CMA) is an intracellular catabolic pathway that mediates the degradation of specific soluble proteins within lysosomes. Because CMA defects are associated to several human pathologies, including neurodegenerative diseases, cancers and immune disorders, research efforts over the past years have been undertaken to study that essential cellular function. Accordingly, recent findings emphasized the fundamental role of CMA for regulating numerous cellular functions including cellular energetics, transcriptional programs, cell death, and cell survival mechanisms or DNA repair.

Briefly, during CMA, cytosolic proteins containing a KFERQ-like motif are first recognized by HSPA8/HSC70 (heat shock protein family A [Hsp70] member 8) and co-chaperones. The substrate-chaperone complex then docks at the lysosomal membrane through specific binding to the cytosolic tail of LAMP2A (lysosomal associated membrane protein 2A). LAMP2A then organizes into a multimeric complex that allows the substrate to translocate across the lysosomal membrane where it is degraded by acid hydrolases. LAMP2A is one of the three variants (namely, LAMP2A, LAMP2B and LAMP2 C) that originate from alternative splicing of the LAMP2 gene. All three splice variants share a common lumenal domain and only differ by their cytosolic and transmembrane regions. CMA activity has been directly correlated to the amount of LAMP2A (but not those of LAMP2B and LAMP2 C) at the lysosomal membrane. As such, LAMP2A has been described as being the limiting component for CMA activity.

Until recently, the absence of any clearly identifiable LAMP2A protein outside of the mammalian and bird species raised concerns about the presence of CMA in other vertebrate lineages. However, using Basic Local Alignment Search Tools (BLASTs) against different expression databases of ray-finned fish species, we were able to identify several expressed sequence tags (ESTs) displaying high homology with mammalian LAMP2A (Figure 30 and **[paper #37]**). This suggested that this protein possibly appeared much earlier during evolution than initially thought, and provided the grounds for looking at the existence of a "genuine" CMA function in fish.

We then provided evidence in this direction. Next, we demonstrated that the LAMP2 gene and its structure containing the three alternatively spliced exons (B, A and C) encoding the transmembrane domain and cytoplasmic tail specific to each isoform (LAMP2B, LAMP2A and LAMP2 C, respectively), is definitively not restricted to mammals or birds, but is also present in the genome of different fish species (Figure 30). In contrast, no homologous sequence was found in invertebrate species, suggesting that CMA is indeed restricted to vertebrates. We could then show that the splice variant lamp2a is expressed from the earliest stages of development as well as in several adult tissues of medaka, supporting the idea that fish might exhibit CMA activity - or at least a CMA-like process- **[see paper #37]**. In order to firmly establish whether or not CMA exists in fish, a medaka fibroblast cell line was transfected with the photoactivable KFERQ-PA-mCherry construct, which has proven to be a reliable reporter for tracking and measuring CMA activity in mammalian cells **[see paper #46]**. Results clearly showed that, upon long-term starvation, this CMA reporter accumulates in characteristic *puncta* that co-localize with lysosomes and/or late endosomes, and that specific knockdown of lamp2a results in a significant

loss of these *puncta*, thereby providing functional evidence for the existence of CMA activity in fish. Finally, to address the physiological role of Lamp2a in fish, a medaka knockout for the splice variant lamp2a was generated. These KO fish display severe alterations in carbohydrate and fat metabolism, similar to what has been observed in the liver of mice deficient for CMA. These results further demonstrated that the CMA function is definitively not restricted to mammals and birds **[see paper #46]**.

Overall, these findings open up new and exciting perspectives to approach (or differently appreciate) CMA under a novel angle. For instance, the relative sequence variability observed within the different functional domains of LAMP2A between phylogenetically distant species will certainly be informative for identifying evolutionarily conserved, or species-dependent, key residues necessary for the structure function relationship of this protein (Figure 30). Beyond these perspectives on the structure-function relationship of LAMP2A, these new findings also emphasize the interest of teleost fish, which diverged from the tetrapod lineage early during vertebrate evolution, as attractive and unique models at the functional interface between invertebrates (assumed to lack any CMA activity, but relying on an endosomal microautophagy [eMI]-like system for targeting KFERQ-like-motif-containing proteins) and mammals (making use of both eMI and CMA functions), for studying the interplay between these two related pathways.

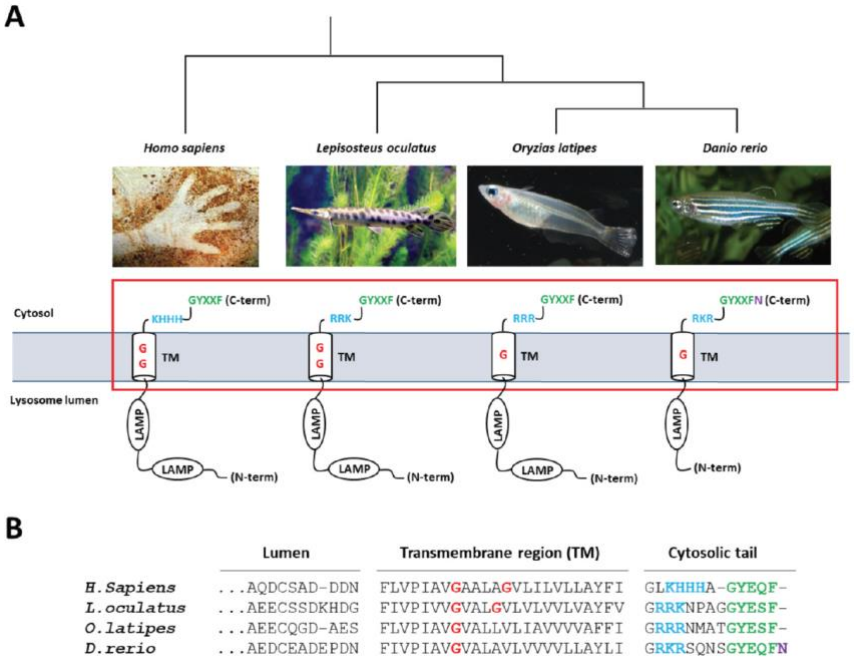


Figure 30. Sequence variability within different functional domains of LAMP2A between phylogenetically distant species.

(A) Schematic representation of selected LAMP2As. **(B)** Sequence alignment of the boxed region. Three positively charged amino acids (in blue) necessary for the binding of substrate proteins are present in fish sequences. However, whereas two glycine (G) residues (in red) located within the transmembrane (TM) region are essential for the multimerization of LAMP2A in rodents, only one G is found in that region in some fish species, including medaka. A GYXXF sequence at the C terminus (in green) is required for targeting LAMP2A to the lysosomal membrane. Although conserved in most teleosts, the divergence of that motif in zebrafish, encoding an additional asparagine residue (N, purple), raises a question about the ability of this species to perform CMA, and certainly deserves special attention.

Hence, much more than challenging the currently tetrapod-centred paradigm, our studies **[see papers #30, #37, #46 and #47]** tackle the urgent need of considering complementary and powerful alternative genetic models, for approaching the entirety of that fundamental catabolic process from an evolutionary perspective (like for sex determination I would say).

2- Discussion (*evolution of master sex determining genes*)

More and more, sex determination is now appreciated as a developmental perspective. As a matter of fact, sex determination is not anymore seen as one hierarchical cascade with the master SD gene on the top, but as a complex network with different types of genetic and environmental factors influencing cell proliferation and hormone levels to push sexual development over either a male or female threshold (84, 85) **[see papers #28, #31, #33]**.

Definitely, studies on the teleost SD support this view in which each node of the SD gene network has the potential of becoming the major-effect *loci*, not necessarily being on the top of the cascade. Archetype of this hypothesis would be amhr2 in Takifugu that modulates transduction efficiency of an upstream ligand signal, i.e., AMH acting on germ cell proliferation or regulating steroid synthesis activity (21), to direct male or female development. Nevertheless, many open questions still remain concerning teleost SD and probably most important, why teleost fish use so many SD genes and SD systems with such a high sex chromosomes turnover rate?

Rapid turnover of sex chromosomes is thought as a way to escape the decaying fate of sex chromosomes, assuming that the lack of recombination around the sex *locus* will lead to the accumulation of deleterious mutations. Alternatively, sexual antagonistic selection has also been put forward to explain the expansion of the sex *locus* and the accumulation of sex specific genes near it. Sex ratio selection provides a scenario that a new master SD gene could rise in frequency when it confers a more balanced sex ratio in the population. Other mechanisms such as random genetic drift and pleiotropic selection have also been used to explain the rise of new SD systems. Another hypothesis, explaining the lack of sex chromosome turnover in mammals and birds, proposes that these highly heteromorphic sex chromosomes might be acting as evolutionary traps that would stabilize SD genes for long span of evolutionary time (4). Interestingly, SD in many teleost species with a genetic SD system can still be influenced by environmental cues like for instance in the Nile tilapia (86), the half-smooth tongue sole (87), and medaka **[see paper #42]**. Transition stages between GSD and TSD are also well illustrated in the case of pejerrey species that conserves a SD gene (amhY) that is used alone in *O. hatcheri* and used in combination with temperature in the closely related species *O. bonariensis* (88). These transition stages could potentially facilitate the turnover of SD systems, leading to the emergence of new SD genes.

It is now clear that the phenotypic expression of sex translates from either genetic triggers, environmental triggers, endocrine triggers or a blend of all. The main emerging idea is that sex determination gene regulatory cascades should no longer be seen as simply hierarchical but, rather, as a regulatory network or, even more, as connections of interdependent regulatory networks. It also seems that the sex determination case should now be treated more like a developmental perspective, rather than a simple one-way top-down differentiation process (neither genetically nor physiologically). Indeed, the process of gonadal differentiation and maintenance is highly plastic lifelong, with formed gonads able to transdifferentiate after what was thought to be “final” differentiation (Figure 29).

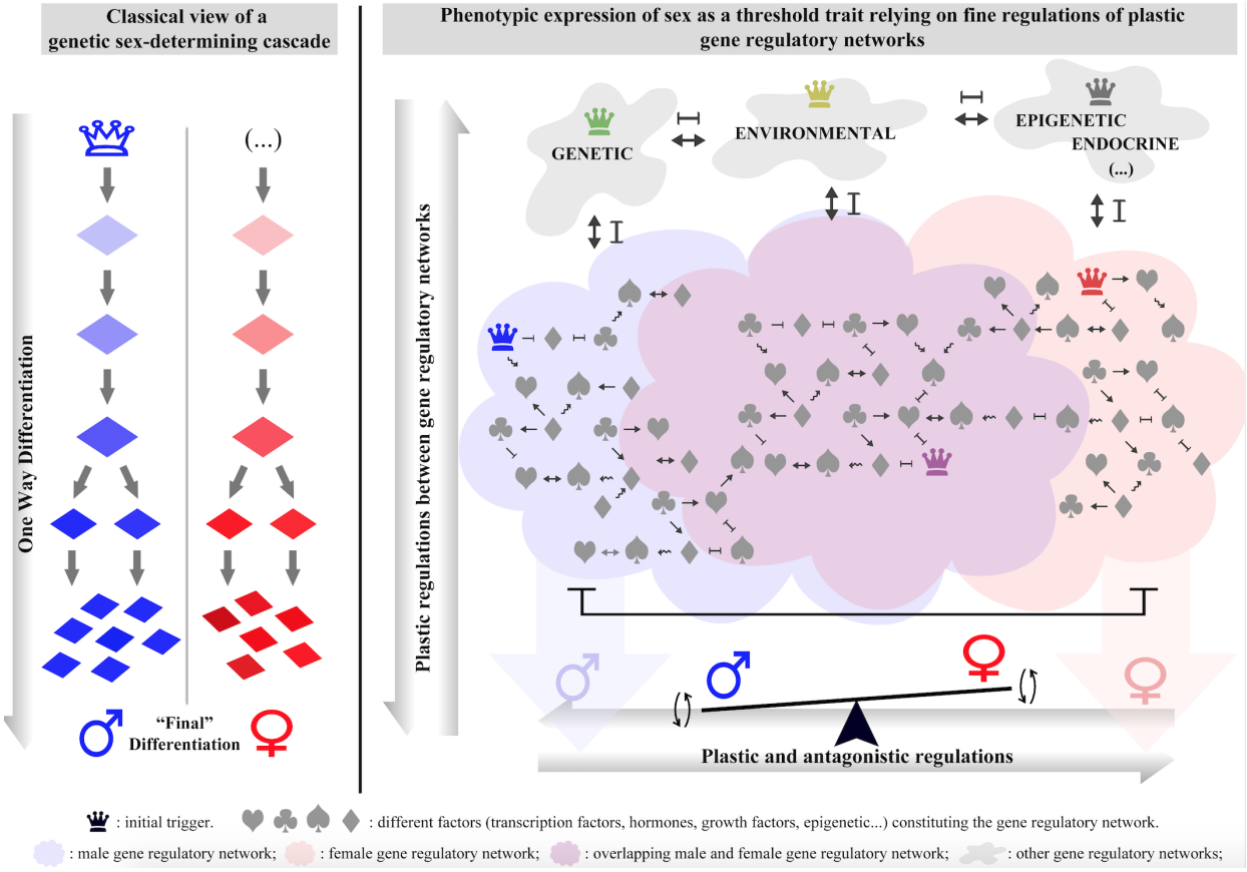


Figure 29. Sex differentiation as a threshold phenotype relying on fine regulations of plastic gene regulatory networks.

It is obvious that the phenotypic expression of sex cannot be seen any longer as a simple one-way top down differentiation process under the action of a unique and totipotent master sex determining trigger (left). Sex determination should, rather, be seen through the prism of a developmental perspective as the emanation of either genetic, environmental, endocrine triggers, or a blend of all, acting among interconnected gene regulatory networks (right).

Sex determination seen as a hierarchized cascade also led to the view that master sex determining genes were necessarily up-recruited from the pre-existing sex cascade (duplication, transcriptional rewiring and sub-functionalization), and then added to the top. Further on, but still as a variation on the theme, one could say that such up-recruitment is not necessary: the gene can stay at its place in the cascade and just become more powerful. But these views are rather mechanistic. Indeed, the underlying mechanisms are always the same: gene duplication, transcriptional rewiring, neo-/sub-functionalization (coding change, truncation …), and specialization. The translation of these mechanisms into physiology opens many more options for evolution (presence or absence does not make it all). Hence, mechanistically speaking, it might be much "easier" to find a gene already known for being able to influence gene regulatory network(s), although others, essentially any other gene, could do the same function. Sdy in salmonids, a gene *a priori* not primarily involved in sex determination, differentiation or maintenance, does not play any physiological role on its own but, rather, provokes a slight bend into the gonadal gene regulatory network **[see paper #39]**. This is enough for doing the job. Indeed, for such bending, maybe some signalling pathway components are better at doing it – for instance, as seen with the emergence of the TGF-β signalling pathway (Figure 2, and perspective part of this manuscript). The existence of such intricate and plastic regulatory networks has drastically changed our traditional perception of a standard linear developmental process for initiating and developing either a male or a female gonad and now opens up fascinating questions for future research (Figure 29).

3- Perspectives

(i) TGF-β molecules challenge the paradigms.

(ia) Emergence of the TGF-β signalling molecules ahead of the sex determining networks in fish.

To date, more than twenty different vertebrate MSD genes have been identified on different sex chromosomes of mammals, birds, frogs and fish (Figure 2). Interestingly, six of these genes are transcription factors (*Dmrt1*- or *Sox3*- related) and thirteen others belong to the TGF-β signalling pathway (*Amh*, *Amhr2*, *Bmpr1b*, *Gsdf* and *Gdf6*). This pattern suggests that only a limited group of factors/signalling pathways are prone to become top regulators, while other well-characterized and indispensable components of sex-determining pathways –e.g. *Sox9* or *Foxl2*- have apparently not been recruited as master regulators in any species studied so far. Moreover, while TGF-β members are clearly subordinate in the mammalian sex-regulatory network, they have independently and recurrently made it to the top in fish (Figure 2). Few years ago, we were among the first to describe and pinpoint the major role of Amh signalling during fish gonadal induction **[see paper #17]**. The biological importance of TGF-β members is unfortunately contrasted by the lack of information on how such signalling(s) is/are elicited and physiologically integrated during gonadal induction and development. Additionally, such a profusion of master regulators from TGF-β pathway-related members draws attention to the evolutionary meaning of this convergent evolution. Thus, outstanding issues remain unanswered:

1- What is the evolutionary meaning of this recurrent convergent evolution toward re-establishing TGF-β signaling pathways for controlling sex-determination in fish?

2- How are the respective sex determining function(s) of TGF-β signalling molecules (Amh, Gsdf and Gdf6) accomplished during fish gonadal induction? Do they converge to a general "TGF-β hub" that connects and integrates them all, or do they remain independent of each other?

3- How do(es) that gonadal TGF-β regulatory network(s) interact with the canonical gonadal gene regulatory network?

To address these questions, our project -for the next years- will be articulated around three main objectives that are detailed and contextualized below.

(ib) A central role for gonadal TGF-β signalling?

The TGF-β family of cell signalling active polypeptides have attracted much attention because of their ability, from nematodes to mammals, to control cellular functions that regulate embryonic development and tissue homeostasis. TGF-β molecules can be subdivided into (***i***) TGF-β *sensu stricto*, bone morphogenetic proteins (BMP), (***iii***) activins and most growth and differentiation factors (GDFs), and (***iv***) distant members of the TGFβ family not fitting into one of the other three subgroups. Despite the tremendous diversity of physiological responses that this family elicits, an astonishingly "simple" system is the core of this signalling pathway. Canonical TGF-β family members have initially been shown to transmit signals through heteromeric complexes consisting of type I and type II serine/ threonine kinase receptors. Ligand binding induces the formation of a heteromeric ternary complex, which is competent to initiate intracellular signal transduction. In this complex, the constitutively active serine/threonine kinase domain of type II receptors

transphosphorylate type I receptors. Activated type I receptors in turn activate specific receptor substrates (Smads) organized in multisubunit complexes that move into the nucleus to regulate transcription of target genes (see Figure 31 and **[papers #3, #11 and #12]** for review). Interestingly an emerging concept is that the distinctive activities of TGF-β family growth factors (GFs) are not only determined as simply explained by signalling type I and II receptors, -which can show varying degrees of promiscuity for GFs-, but also by other proteins binding either intracellularly to TGFβ receptors or when interacting in the extracellular space, e.g. soluble TGFβ/BMP antagonists or membrane-located co-receptors, can function as activating or inactivating modulator protein. They thereby greatly diversify signalling activity by adding another layer of signalling **[see papers #11 and #12]**. They do not only control whether GFs reach their receptors on cells, but also whether additional components are present within GF-receptor complexes (89). Thus, molecular recognition in the TGF-β family is not singularly achieved by GF-receptor interactions, but by a network of interactions with multiple partners (89).

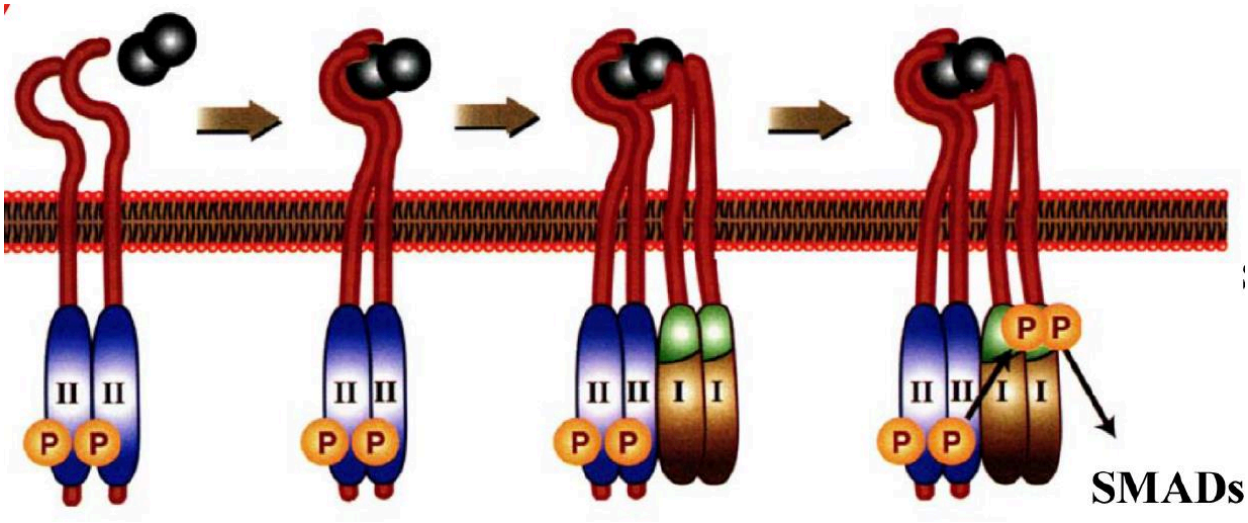


Figure 31. Cell signalling pathway of TGF-β superfamily of ligands.

(1) Ligands bind and interact with constitutively phosphorylated dimers of type II receptors. **(2)** Heterotetramers form by recruitment of dimers of type I receptors. **(3)** Transphosphorylation of type I receptors by type II receptors occurs. **(4)** Smads are recruited and phosphorylated by type I receptors, and then move from the cytoplasm to the nucleus where they act as transcription factors.

With respect to gonad development, the following factors are of special interest when investigating the TGF-β signalling pathway:

AMH. The anti-Müllerian hormone (Amh) is a distant member of the TGF-β family belonging rather to the forth subgroup which has limited homology to members of the activin/inhibin and *sensu stricto* TGFβs. In mammals *Amh* plays a major role for the regression of the Müllerian duct-forming part of the female reproductive tract during male embryo development (90, 91). Nevertheless, it is not required for mouse testis development although apparently playing a central role in testis formation in non-mammalian species. For instance, in chicken embryonic gonads, Amh is expressed much higher in males and is predicted in birds to be responsible for organizing the early testis structures (92). Fish do not have a Müllerian duct, but have an *Amh* homolog, as we showed for the first time (93). Although being clearly a subordinate member of the sex regulatory network in mammals, the Amh/ Amh-receptor system has made it to the top in several species (Figure 32).

In zebrafish (*Danio rerio*), a major model for vertebrate development and human health, the genetic basis of SD is still unknown. Interestingly, being expressed during the gonadal

differentiation period Amh has been shown to be implicated in zebrafish testis formation (94). Nevertheless, to date the presence of the *bona-fide* receptor (Amhr 2) could not be detected in the genome of zebrafish and other Cyprinids. This indicates that Amh signalling in zebrafish must employ a so far unknown or uncharacterized type of receptor. Certainly, this opens up questions regarding the exclusive specificity of the Amh with respect to "its own" type II receptor in zebrafish and cyprinids, but also in other fish species.

GSDF. Gonadal soma-derived factor (Gsdf) is another growth factor displaying key features of the TGF-β family (see Figure 32). During evolution, it has been lost in tetrapods (95), and its biochemical function is by far not well studied. It is assumed to have a role in male gonad development due to its exclusive expression in the early differentiating testis of all fish analysed so far. Besides its proposed role in the downstream regulatory network (96), *Gsdf* has made it up to the top in *Oryzias luzonensis* (12) (a sister species to medaka) and most likely also in the sablefish ((44) and **Herpin et al. in prep**). However, the receptors and Smads through which Gsdf elicits signalling are still completely unknown (Figure 32). Interestingly, its peculiar and composite nature, being most homologous to inhibins, but harbouring a C-terminal LEFTY/DAN-like extension and on the other hand lacking any α-helix type I receptors epitopes (***personal data***) make of Gsdf a real biochemical challenge for deciphering its likely atypical signalling modalities.

BMP/GDF. In the mouse embryo, several independent BMP signals are necessary for proper PGC induction: the primary induction of PGCs at the posterior proximal epiblast is driven by BMP4 (97), whereas the number of PGCs is guided by BMP2, BMP4 and BMP8b in a synergistic action (98). Hence, although not apparently involved during the early events of SD, the BMP signalling pathway is implicated during mammalian germ cell specification and gametogenesis (99). Through participation in the turquoise killifish (*Nothobranchius*) genome project, a well-established fish model for aging research we have recently shown that *Gdf6/Bmp13* is likely to be the master sex-determining gene of this species (100). Nevertheless, what the downstream targets of that signalling route are during fish sex-determination are still completely unknown (an ongoing collaboration has been set up with C. Englert (University of Jena, Germany)).

(ic) TGF-β signalling specificity

One most important open question is how signalling specificity is determined? Specificity of BMP signalling relies critically on the particular combination of type I and II receptors, as well as on the identity of the specific ligand associated with the active receptor complex **[see paper #3]**. It also involves the differential usage of several substrates (Smads), which are distinct between activated forms of different type I receptors (Figure 32, (101)). What is not so straightforward, however, is how such a simple core-system coordinates, integrates and elicits a plethora of physiological outcomes in a tightly regulated cell-specific context. Binding versatility and flexible receptor oligomerization patterns may lead in some cases to cross-talk between the TGF-β ligand family pathways. It could possibly result in direct competition at the membrane through shared receptors or at the level of transcription factors. This represents a first level of complexity that leads to the multiplicity of biological end-points compared to the relatively low number of primary interacting participants (ligands, receptors and Smads, **[see papers #3, #11 and #12]**). A second and more complex level of integration can be attributed to the ability of Smads to act as signalling platforms through interactions with a variety of different proteins. Thus, the composition of Smad transcriptional complexes and the resulting physiological activities are necessarily driven and modulated in a cell-specific fashion. An emerging concept is that the response of a given cell to

extrinsic signals relies not only on the effect of single, isolated pathways, but also more on the integration of multiple signals from a plethora of cross-talking pathways **[see paper #12]**.

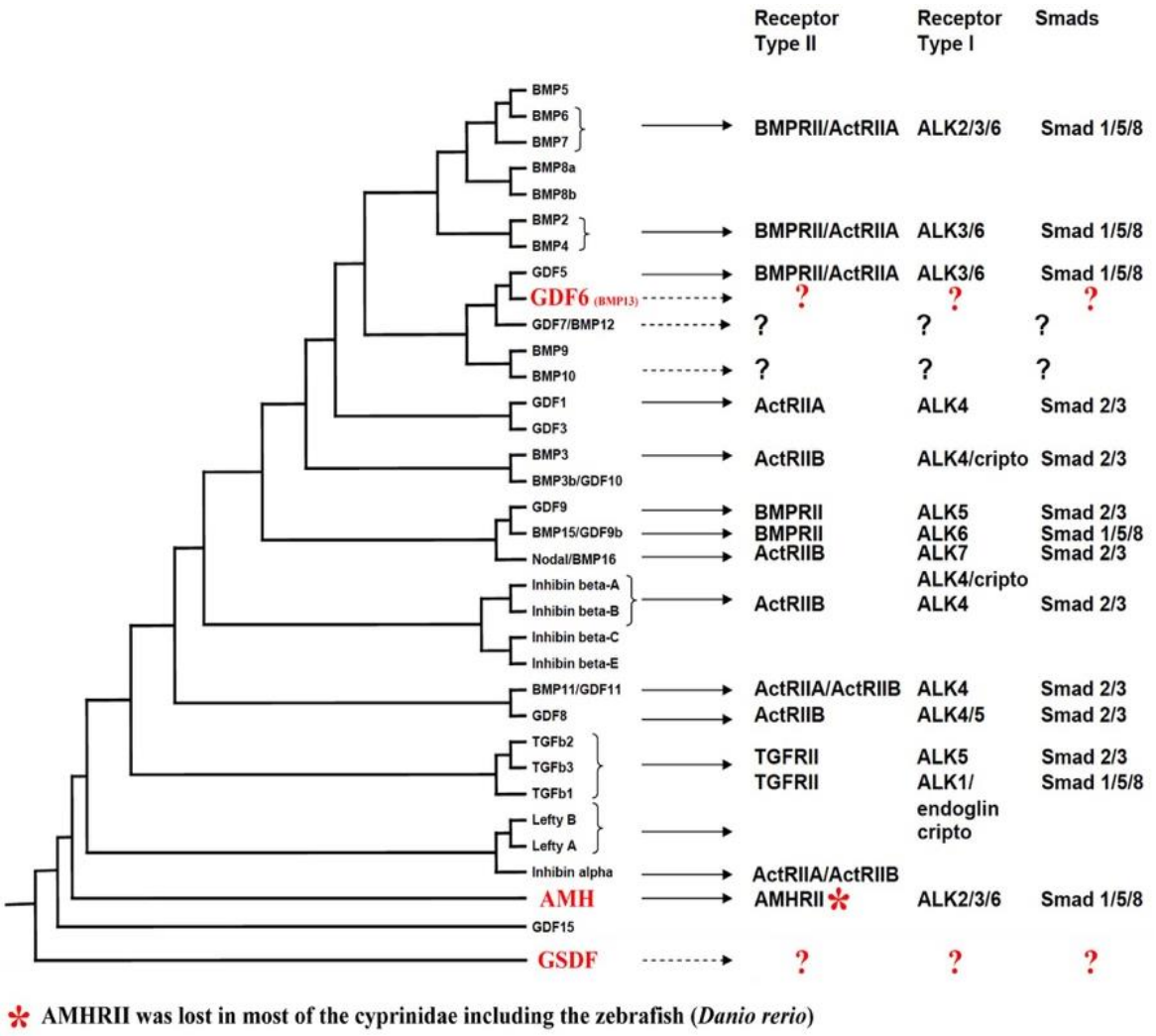


Figure 32. Phylogenetic relationship of paralogous TGF-β, GDF, BMP, AMH, and GSDF ligands, as well as characterized receptors and signalling pathways for individual ligands.

To draw near these questions, more information on the molecular mechanisms of TGF-β signalling involved during gonadal development is required. To this end, our perspectives aim at deciphering the importance of TGF-β signalling for SD and maintenance. We want to explore the nodes that create specificity in the signalling process, the ligand/receptor and receptor/Smad levels. Using the medaka fish as model organism, we will employ functional genomics approaches combining classical comparative investigations together with cutting edge methods for functional investigation. We will develop transgenic approaches and tools for testing, *in vivo*, the responsiveness dynamics of the different TGF-β signalling pathways at a cellular resolution. Synergistically, *in silico* modelling of predictable ligand/receptor interaction will be conducted thanks to our collaborations (Thomas Müller, see collaborators), in order to address our 3 main objectives:

<u>Objective 1:</u> Signalling specificity and crosstalk of different gonadal TGF-β signal transducing factors (Amh/AmhR2/Gsdf/GDF6) and inferred structure/function relationship.

This objective aims at connecting each ligands/receptor pairs (Amh/AmhR2, GDF6 and Gsdf with identified partners) to SMAD (1, 2, 3, 5 and 8) activation *via* phosphorylation. Specifically testing the relative degrees of phosphorylation for each Smads upon various ligand/receptors interactions will additionally be informative in terms of crosstalk and relative specificity between different TGF-β signalling pathways (see Gal4_UAS system in Figure 33). This will notably also allow us to tell whether (or not) a general gonadal TGF-β "hub" centralizes and converts signalling(s) from different ligand/receptor combinations into a common downstream outcome (similar combinations of Smads' activation).

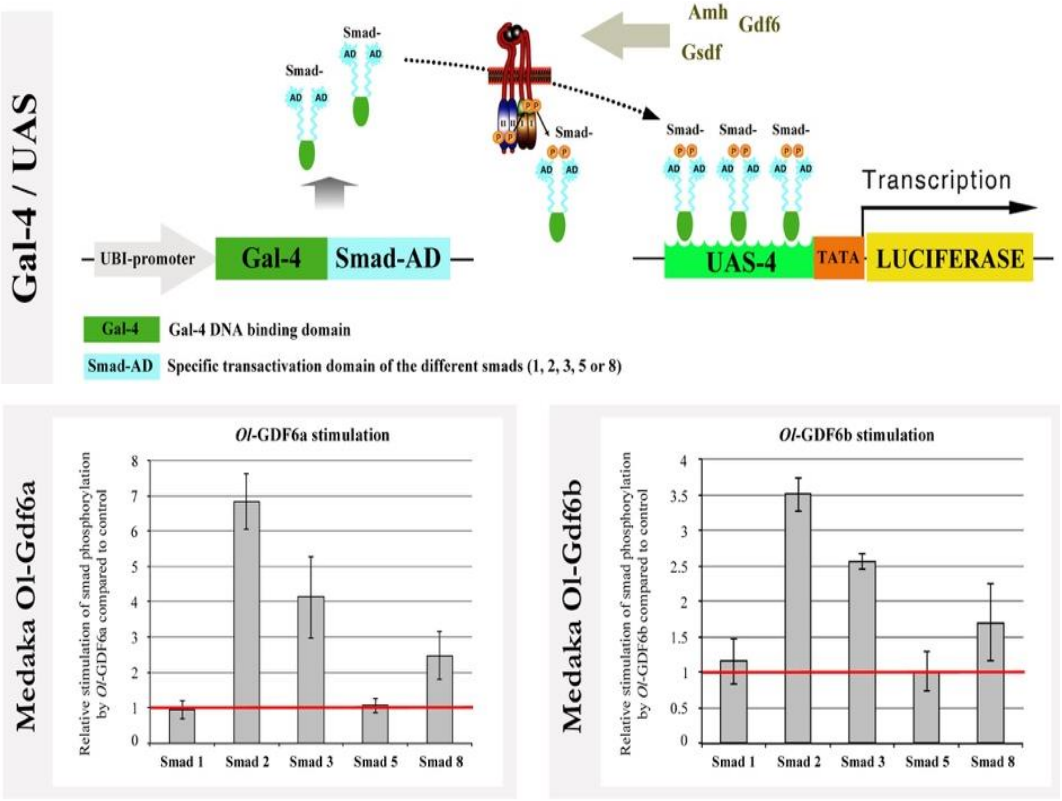


Figure 33.

<u>Upper:</u> Gal-4/UAS reporter system for TGF-ß signalling to detect specific Smad activation.

Our own system, based on a Gal4 transcriptional activator coupled with the transactivation domains of the different Smads (1, 2, 3, 5 and 8) and an UAS-Luc reporter expressing a luciferase will be employed in different medaka cell lines.

<u>Lower:</u> UAS/Gal4-mediated smad phosphorylation assay of medaka Gdf6a/b signal through phosphorylation of Smads 2, 3, and 8. Our system has been tested for sensitivity and specificity. For example, preliminary data using this system shows that GDF6a/b, in a cell-specific context, "signal" through phosphorylation of smads 2 and 3 and to a lesser extent through smad8.

This system was initially developed by Sedes *et al*. (102).

<u>Objective 2:</u> Sex-determining function of TGF-β pathway/signalling during sex-determination and maintenance in medaka and selected representative fish species.

Facing the lack of data on how TGF-β signalling is elicited and physiologically integrated during gonadal induction, development and maintenance in fish (and beyond), we first want to identify the components that are active during sexual development.

For that purpose, first, detailed comparative expression patterns analysis of Amh, AmhR2, Gsdf, and Gdf6 but also all other TGF-β-related components (including receptors) will be conducted in key phylogenetically fish species (23 ray-finned fish, including 2 holosteans, and 21 teleost species). Blast-based searches on our in-house developed PhyloFish database (http://phylofish.sigenae.org, (103)) providing consistent and exhaustive gene expression data from 23 different ray-finned fish species will be perform to determine the spatial and temporal collinearity expression relationship of these TGF-β components during SD stages and in adult gonads (see Figure 34 for preliminary data in medaka). This is a prerequisite for assigning and validating functional and physiologically relevant interactions between TGF-β ligands/receptors/smads.

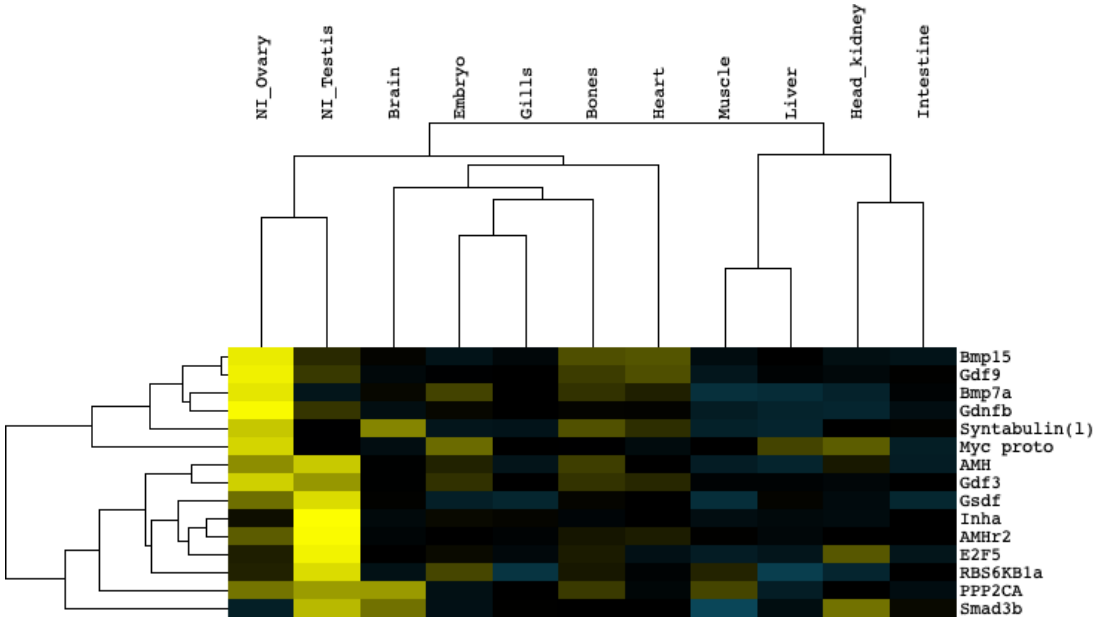


Figure 34. Expressional analysis of different TGF-β components shows collinearity of expression of certain members in medaka gonads.

Next, in order to better picture how the whole gonadal gene regulatory network is affected by TGFβ signalling, we will perform RNA-seq after TGFß modulation in different medaka cell lines (spermatogonia and embryonic) transfected with either Amh, Gsdf or Gdf6. In parallel, proliferation assays will be performed.

Functional assessment will then be also performed *in vivo* for Amh, Gdf6a and b. For this specific purpose, "knock out" medaka lines for Amh, Gdf6a and b have already been produced (CRISPR/Cas9) and will be analysed (on histological and transcriptomical levels) for deviations in the development of gonads and for possible sex-reversal. This will complement our knowledge previously generated for the role of Gsdf in male gonadal sexual differentiation initiation **[see paper #32]**.

<u>Objective 3:</u> Integration of the TGF-β signalling pathway(s) within the canonical sex-determination gene regulatory network.

In tilapia (104) and medaka (96), evidences suggest that Gsdf is a direct effector of DMRT1 and DMRT1bY respectively. In order to assign and estimate the importance of the different gonadal TGF-β signaling pathways within the canonical SD gene regulatory network we will investigate the regulation of the different essential TGF-β pathways components by DMRT1bY, the master sex-determining factor of medaka**.** We will determine the different relationships and transactivating properties among these factors by *in vivo* Chromatin Immuno-Precipitation of an *in vivo* tagged version of DMRT1bY (*in vivo* ChIP-seq). To this end we will make use of our already established transgenic reporter fish expressing a tagged version of Dmrt1bY (CRISPR/Cas9 mediated KI) and asses its binding to specific promoter region of TGF-β signalling factors involved in sexual development. We have validated this method in several publications (48, 96, 105), (Figure 35).

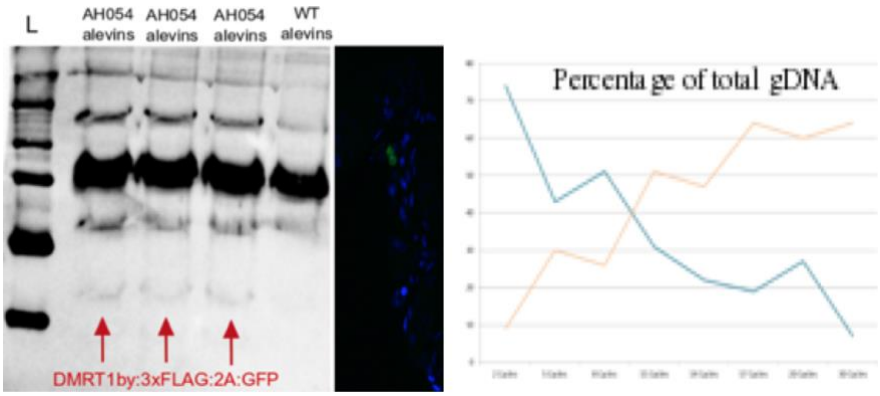


Figure 35. Western blot and fluorescent validation for expression of the *in vivo*-tagged Dmrt1bY gene; and set up conditions for DNA sonication (ChIP-seq).

The pleiotropic nature of the TGF-β signalling pathway (combining signalling molecules, ligands, receptors, effectors, and transcription factors) makes this pathway particularly capable of re-structuring and fine-tuning intricate networks of gene regulation. While TGF-β family members clearly play a crucial role in integrating a plethora of signals, and have been functionally and genomically proven to be MSD genes, environmental factors can still override them and cause sex reversal both in the lab and in the wild. This plasticity in species with clear sex chromosomes and a GSD system emphasizes that sex should no longer be viewed as a rigid and pre-determined path from genotype to phenotype, but rather as a multilayer reaction norm resulting from developmental noise, and which can be contingently modulated or totally ruled by genetic factors. In this perspective, sex is the net product of a variety of environmental, genomic and epigenomic, and stochastic determinants. The master genetic trigger therefore has to cope with influences from factors which may disturb its action directly, or disturb the downstream action. TGF-β, with its tight-knit regulatory network, and surplus copies in teleosts, is probably especially suitable for the production of a phenotype as plastic as sex.

(ii) Physiological readout of the neo-functionalization of the medaka Dmrt1 paralogs (Dmrt1a *vs* Dmrt1bY)

Like said before, in the medaka fish a duplicated copy of *dmrt1*, designated *dmrt1bY*, on the Y chromosome was shown to be the master regulator of male development. Further on we found that the *dmrt1bY* gene has acquired a new feedback downregulation of its expression **[see papers #21 and #36]**. Additionally, the autosomal *dmrt1a* gene is also able to regulate transcription of its duplicated paralog by binding to a unique target Dmrt1 site nested within the *dmrt1bY* proximal promoter region **[see paper #21]**. We could trace back this novel regulatory element to a highly conserved sequence within a new type of TE that inserted into the upstream region of *dmrt1bY* shortly after the duplication event (Figures 8 and 20). This new transcriptional regulatory element allowed neo-functionalization of the sex chromosomal copy and guaranteed the survival of the autosomal copy by a possible sub-functionalization in Sertoli cells of adult testes. In the particular case of *dmrt1bY*, this contributed to create new hierarchies of sex-determining genes (Figure 20). Of special interest is the question, how a gene, dmrt1 for instance, which is placed downstream in the sex determination cascade, can take over the position and function of the gene at the top. In other words, we now want to know whether -or not- the two dmrt1 paralogs have acquired different functions after duplication. Interestingly few lines of evidences suggest that it is indeed the case (Figures 36 to 38):

***(i)* Single aminoacid changes between dmrt1 paralogs from *O. latipes* and *curvinotus* within either the DM DNA-binding or transactivation domains suggest functional neo-functionalization** (Figure 36).

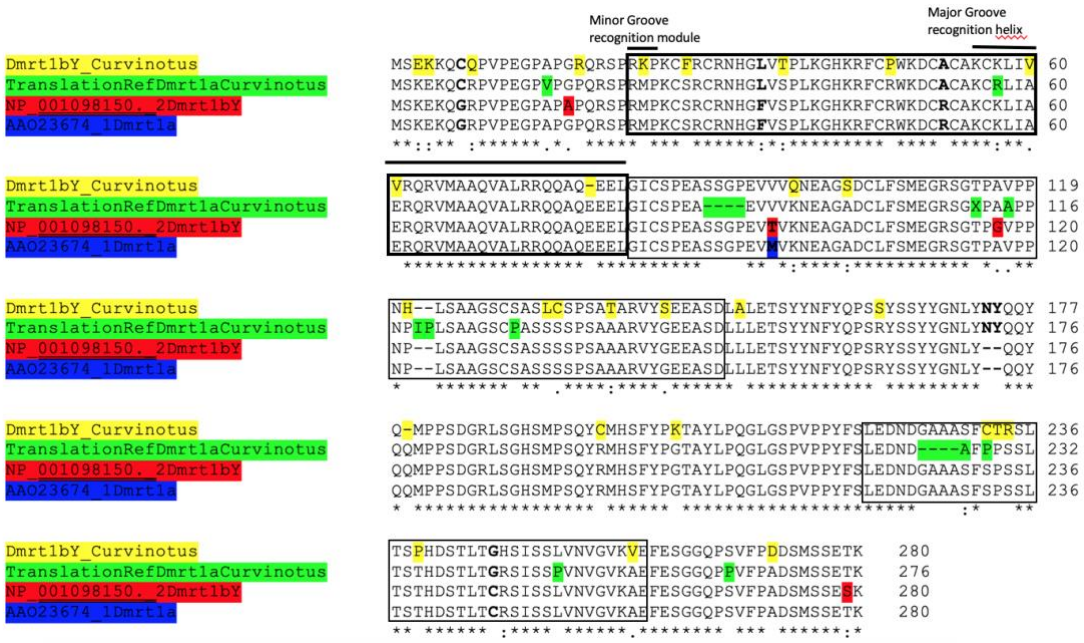


Figure 36. Sequence comparison of the dmrt1 paralog genes (dmrt1a and dmrt1bY) in *O. latipes* and *O. curvinotus*.

(ii) Positive selection is apparent between the two paralogs (Figure 37).

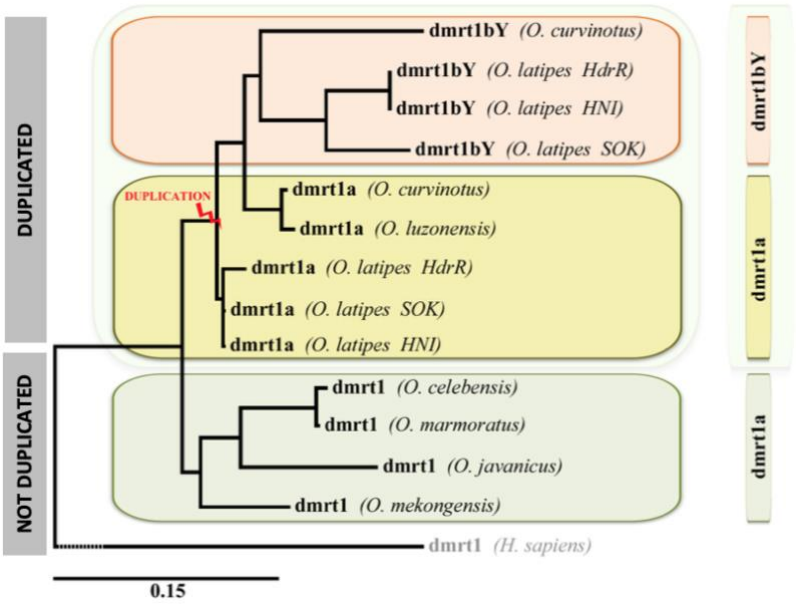


Figure 37. Evolution of the dmrt1 paralogs after duplication.

-Higher rate of synonymous substitution in Dmrt1bY than in Dmrt1a (higher mutation rate in Dmrt1bY than in Dmrt1a). –Higher rate of non-synonymous substitution in Dmrt1bY than in Dmrt1a (elevation of mutation rate and change in natural selection). –dn/ds ratios up to 3.25 for the DM DNA-binding domain (provide evidence for positive selection). See also (106).

(iii) Dmrt1a and Dmrt1bY differentially activate target genes and bind to different targets (Figure 38).

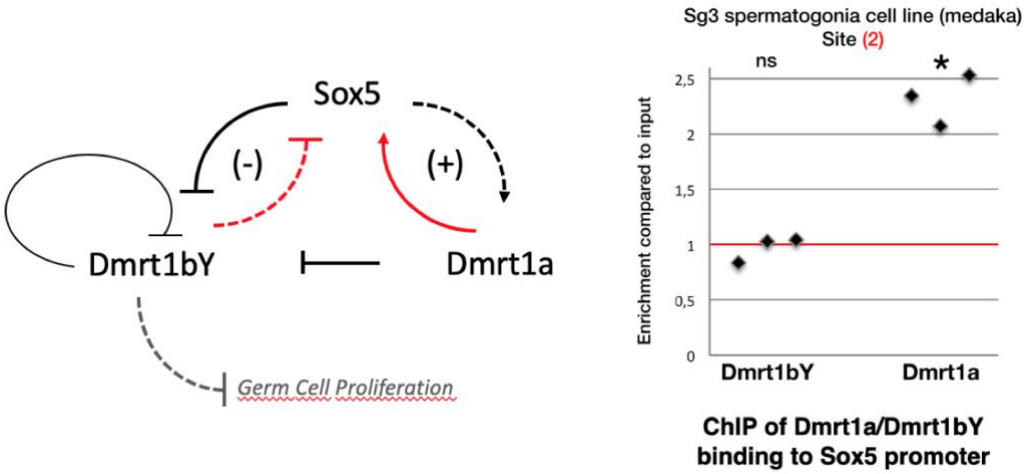


Figure 38. Activation and binding of target genes by the Dmrt1 paralogs.

While *in vitro* experiments show that Dmrt1a activates transcription of Sox5, on the other hand Dmrt1bY represses Sox5 transcription. *In vivo* ChIP (chromatin immunoprecipitation) reveals that Dmrt1a binds to Sox5 promoter, while Dmrt1bY does not.

Hence, using sequences analyses, phylogeny, promoter activity, ChIP-PCR and interspecific hybrids (107), all the evidence indeed suggests that the function of the dmrt1 paralogs diverged after duplication.
To decipher the physiological readout of that functional divergence between the two paralogs, *in vivo* Chromatin ImmunoPrecipitation, sequencing (*in vivo* ChIP-seq), and Co-ImmunoPrecipitation (Co-IP) will be employed to isolate target genes and co-factors interacting with the Dmrt1 paralogs. We will hence get a clearer picture of the Dmrt1bY-aheaded medaka gonadal regulatory network.
Because no specific antibodies could be produced to specifically recognize Dmrt1a or Dmrt1bY (which sequences are nevertheless too similar), we have tagged both paralogs *in vivo* using genome editing (Knock-in, Figure 39).

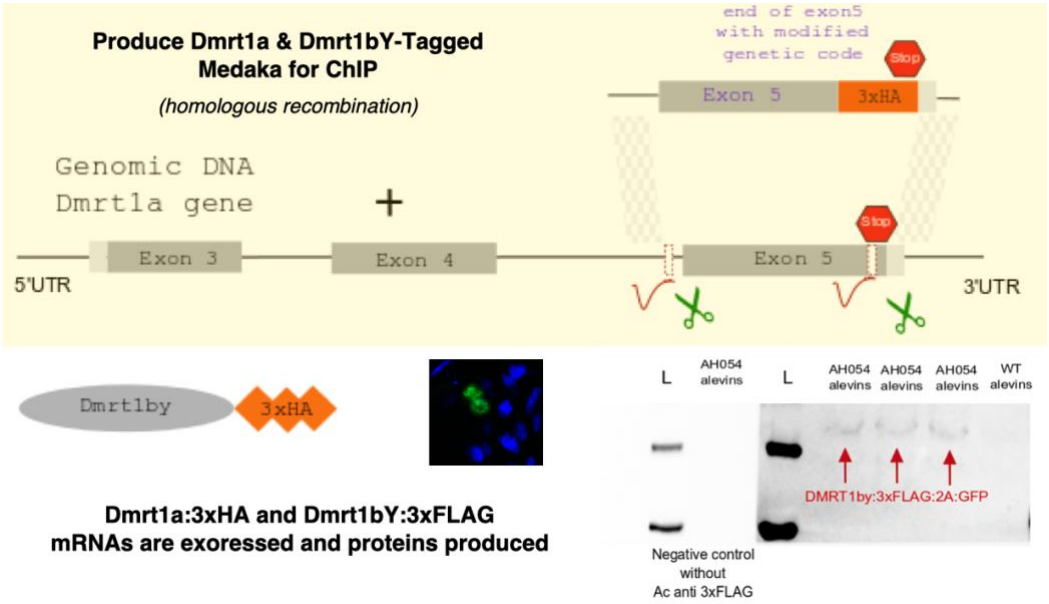


Figure 39. In vivo tagging of the two Dmrt1 paralogs.
Using guide RNAs specific of either dmrt1 paralogs together with the Cas9 enzyme, we were able to specifically remove the last exon (exon5) of Dmrt1a and Dmrt1bY. Thanks to micro-homology repair mechanism acting on provided homologous DNA fragment encompassing exons 5 fused to either FLAG- or HA- tags, we could *in vivo* tag both dmrt1 paralogs. Detection of either Dmrt1 paralogs was verified with specific antibodies (anti-HA or anti-FLAG) by mean of western-blotting or immunofluorescence.

After ChIP-seq, bioinformatic analysis should then give us an extensive list of genes, for which Dmrt1s are binding to their corresponding promoters. Validation of the data will be done by real time RT-PCR from the cell lines and *in-vivo* from different stages of male gonad development as well as adult testes (and ovary for control). Further on, interactions and transcriptional regulation will be tested in different cell lines. Genes proven to be critical in the genetic network governed by *dmrt1bY* will be targeted for functional analyses by genome editing (Knock-out).
Protein partners of both Dmrt1 paralogs will be isolated after Co-Imunoprecipitation and mass spectrometry analysis.
Eventually, knowing how Dmrt1a/Dmrt1bY function and target genes evolved after duplication will certainly give us the most comprehensive view how a master sex determining gene is spawn, evolve and maintained in a Vertebrate.

(iii) Of the pronephric contribution to the differentiating gonads in fish…at a cellular resolution.

As a follow-up of our pronephric field contribution to the gonad story (see Figures 25 to 28), and considering the fact that cell lineage tracing experiments strongly suggested that, indeed very early on during development, cells from the pronephros contribute to the gonads, we now want to focus our research on the molecular mechanisms underlying this phenomenon, but at a cellular level/resolution. For that purpose, our approach will rely on mapping single cell transcriptomes within the pronephros tissue context (Visium Spatial Gene Expression System, 10X Genomics, Figure 40).

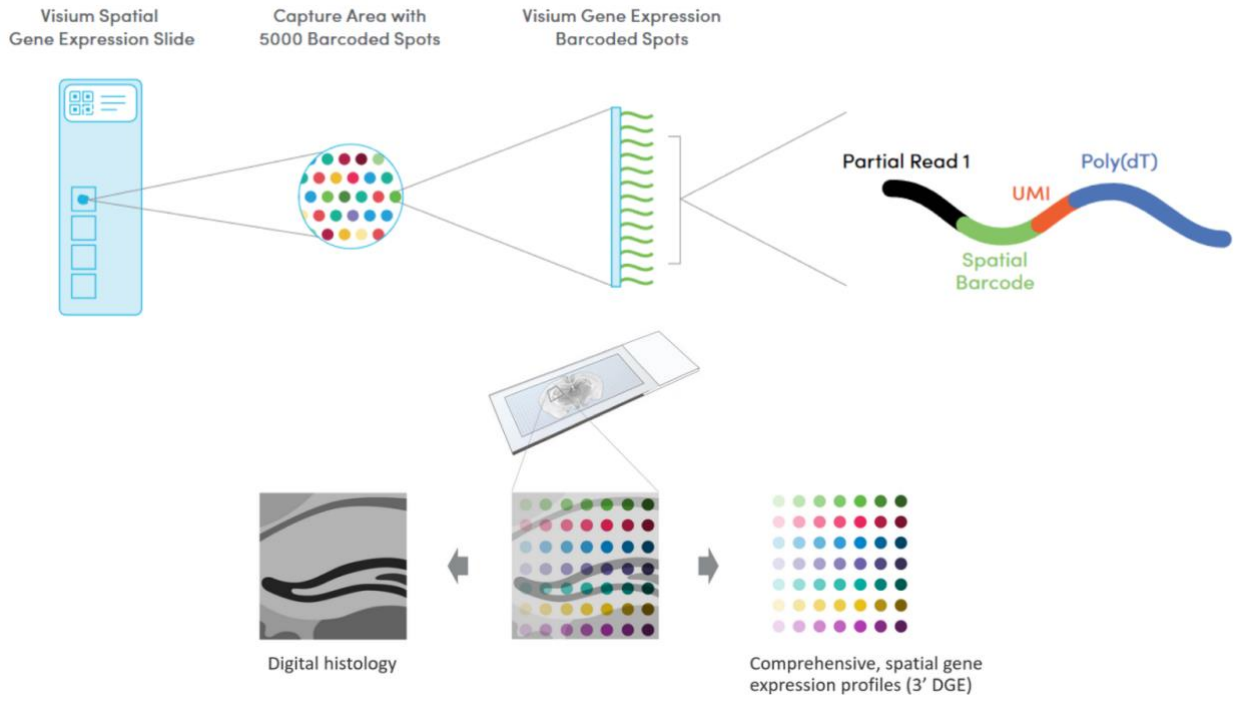


Figure 40. Spatial gene expression system workflow.

Each slide contains four capture areas with approximately 5000 barcoded spots, which in turn contain millions of spatially-barcoded capture oligonucleotides. From cryosections slices, tissue mRNAs are released and binds to the barcoded oligos, enabling capture of gene expression information. Overlaid histology together with gene expression information allows transcriptomic information to each cells/group of cells keeping tissue topology.

Hence, we will be able to identify different cell subtypes/populations/lineages within the pronephros at hatching stage and decipher which sub-lineage(s) (according to its/their transcriptomic signatures) indeed contribute to the gonads in medaka.

4- References

1. Wilkins,A.S. (1995) Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *Bioessays*, **17**, 71–77.

2. Graham,P., Penn,J.K.M. and Schedl,P. (2003) Masters change, slaves remain. *Bioessays*, **25**, 1–4.

3. Matson,C.K. and Zarkower,D. (2012) Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat. Rev. Genet.*, **13**, 163–174.

4. Bachtrog,D., Mank,J.E., Peichel,C.L., Kirkpatrick,M., Otto,S.P., Ashman,T.-L., Hahn,M.W., Kitano,J., Mayrose,I., Ming,R., *et al.* (2014) Sex determination: why so many ways of doing it? *PLoS Biol.*, **12**, e1001899.

5. Schultheis,C., Böhne,A., Schartl,M., Volff,J.N. and Galiana-Arnoux,D. (2009) Sex determination diversity and sex chromosome evolution in poeciliid fish. *Sex Dev*, **3**, 68–77.

6. Cnaani,A., Lee,B.-Y., Zilberman,N., Ozouf-Costaz,C., Hulata,G., Ron,M., D'Hont,A., Baroiller,J.-F., D'Cotta,H., Penman,D.J., *et al.* (2008) Genetics of sex determination in tilapiine species. *Sex Dev*, **2**, 43–54.

7. Wittbrodt,J., Shima,A. and Schartl,M. (2002) Medaka--a model organism from the far East. *Nat. Rev. Genet.*, **3**, 53–64.

8. Matsuda,M., Nagahama,Y., Shinomiya,A., Sato,T., Matsuda,C., Kobayashi,T., Morrey,C.E., Shibata,N., Asakawa,S., Shimizu,N., *et al.* (2002) DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*, **417**, 559–563.

9. Nanda,I., Kondo,M., Hornung,U., Asakawa,S., Winkler,C., Shimizu,A., Shan,Z., Haaf,T., Shimizu,N., Shima,A., *et al.* (2002) A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11778–11783.

10. Kondo,M., Nanda,I., Hornung,U., Asakawa,S., Shimizu,N., Mitani,H., Schmid,M., Shima,A. and Schartl,M. (2003) Absence of the candidate male sex-determining gene dmrt1b(Y) of medaka from other fish species. *Curr Biol*, **13**, 416–420.

11. Myosho,T., Otake,H., Masuyama,H., Matsuda,M., Kuroki,Y., Fujiyama,A., Naruse,K., Hamaguchi,S. and Sakaizumi,M. (2012) Tracing the emergence of a novel sex-determining gene in medaka, Oryzias luzonensis. *Genetics*, **191**, 163–170.

12. Myosho,T., Takehana,Y., Hamaguchi,S. and Sakaizumi,M. (2015) Turnover of Sex Chromosomes in Celebensis Group Medaka Fishes. *G3 (Bethesda)*, **5**, 2685–2691.

13. Raymond,C.S., Murphy,M.W., O'Sullivan,M.G., Bardwell,V.J. and Zarkower,D. (2000) Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev.*, **14**, 2587–2595.

14. Eser,M. and Ayaz,A. (2018) Haploinsufficiency of the DMRT Gene Cluster in a Case with 46,XY Ovotesticular Disorder of Sexual Development. *Balkan Med J*, **35**, 272–274.

15. Smith,C.A., Roeszler,K.N., Ohnesorg,T., Cummins,D.M., Farlie,P.G., Doran,T.J. and Sinclair,A.H. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature*, **461**, 267–271.

16. Yoshimoto,S., Okada,E., Umemoto,H., Tamura,K., Uno,Y., Nishida-Umehara,C., Matsuda,Y., Takamatsu,N., Shiba,T. and Ito,M. (2008) A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 2469–2474.

17. T,S., A,Y., K,A., H,M., M,F.-S., M,T. and H,K. (2008) Distinct contributions of CXCR4b and CXCR7/RDC1 receptor systems in regulation of PGC migration revealed by medaka mutants kazura and yanagi. *Developmental biology*, **320**.

18. Schartl,M., Schmid,M. and Nanda,I. (2016) Dynamics of vertebrate sex chromosome evolution: from equal size to giants and dwarfs. *Chromosoma*, **125**, 553–571.

19. Yano,A., Guyomard,R., Nicol,B., Jouanno,E., Quillet,E., Klopp,C., Cabau,C., Bouchez,O., Fostier,A. and Guiguen,Y. (2012) An immune-related gene evolved into the master sex-determining gene in rainbow trout, Oncorhynchus mykiss. *Curr. Biol.*, **22**, 1423–1428.

20. Kikuchi,K. and Hamaguchi,S. (2013) Novel sex-determining genes in fish and sex chromosome evolution. *Dev. Dyn.*, **242**, 339–353.

21. Kamiya,T., Kai,W., Tasumi,S., Oka,A., Matsunaga,T., Mizuno,N., Fujita,M., Suetake,H., Suzuki,S., Hosoya,S., *et al.* (2012) A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, Takifugu rubripes (fugu). *PLoS Genet.*, **8**, e1002798.

22. Chen,S., Zhang,G., Shao,C., Huang,Q., Liu,G., Zhang,P., Song,W., An,N., Chalopin,D., Volff,J.-N., *et al.* (2014) Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat. Genet.*, **46**, 253–260.

23. de Souza,F.S.J., Franchini,L.F. and Rubinstein,M. (2013) Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong? *Mol. Biol. Evol.*, **30**, 1239–1251.

24. Woram,R.A., Gharbi,K., Sakamoto,T., Hoyheim,B., Holm,L.-E., Naish,K., McGowan,C., Ferguson,M.M., Phillips,R.B., Stein,J., *et al.* (2003) Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome Res.*, **13**, 272–280.

25. Faber-Hammond,J.J., Phillips,R.B. and Brown,K.H. (2015) Comparative Analysis of the Shared Sex-Determination Region (SDR) among Salmonid Fishes. *Genome Biol Evol*, **7**, 1972–1987.

26. Guiguen,Y., Fostier,A., Piferrer,F. and Chang,C.-F. (2010) Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. *Gen. Comp. Endocrinol.*, **165**, 352–366.

27. Marshall Graves,J.A. and Peichel,C.L. (2010) Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biol*, **11**, 205.

28. Vicoso,B. and Charlesworth,B. (2006) Evolution on the X chromosome: unusual patterns and processes. *Nature Reviews Genetics*, **7**, 645–653.

29. Singh,N.D., Koerich,L.B., Carvalho,A.B. and Clark,A.G. (2014) Positive and Purifying Selection on the Drosophila Y Chromosome. *Mol Biol Evol*, **31**, 2612–2623.

30. Charnov,E.L. and Bull,J.J. (1989) The primary sex ratio under environmental sex determination. *Journal of Theoretical Biology*, **139**, 431–436.

31. Fisher,R.A. (1931) The Evolution of Dominance. *Biological Reviews*, **6**, 345–368.

32. Bull,J.J. (1983) Evolution of Sex Determining Mechanisms Benjamin/Cummings Publishing Company, Advanced Book Program.

33. Rice,W.R. (1987) The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination Between Primitive Sex Chromosomes. *Evolution*, **41**, 911–914.

34. Charlesworth,D. and Charlesworth,B. (2005) Sex chromosomes: evolution of the weird and wonderful. *Curr. Biol.*, **15**, R129-131.

35. Kirkpatrick,M. (2017) The Evolution of Genome Structure by Natural and Sexual Selection. *J. Hered.*, **108**, 3–11.

36. Perrin,N. (2016) Random sex determination: When developmental noise tips the sex balance. *BioEssays*, **38**.

37. Takehana,Y., Matsuda,M., Myosho,T., Suster,M.L., Kawakami,K., Shin-I,T., Kohara,Y., Kuroki,Y., Toyoda,A., Fujiyama,A., *et al.* (2014) Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish Oryzias dancena. *Nat Commun*, **5**, 4157.

38. Wilkins,A.S. (2005) Recasting developmental evolution in terms of genetic pathway and network evolution ... and the implications for comparative biology. *Brain Res. Bull.*, **66**, 495–509.

39. Timchenko,L.T., Miller,J.W., Timchenko,N.A., DeVore,D.R., Datar,K.V., Lin,L., Roberts,R., Caskey,C.T. and Swanson,M.S. (1996) Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.*, **24**, 4407–4414.

40. Moraes,K.C.M., Wilusz,C.J. and Wilusz,J. (2006) CUG-BP binds to RNA substrates and recruits PARN deadenylase. *RNA*, **12**, 1084–1091.

41. Mancebo,R., Zhou,X., Shillinglaw,W., Henzel,W. and Macdonald,P.M. (2001) BSF binds specifically to the bicoid mRNA 3' untranslated region and contributes to stabilization of bicoid mRNA. *Mol. Cell. Biol.*, **21**, 3462–3471.

42. De Renzis,S., Elemento,O., Tavazoie,S. and Wieschaus,E.F. (2007) Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. *PLoS Biol.*, **5**, e117.

43. Sterky,F.H., Ruzzenente,B., Gustafsson,C.M., Samuelsson,T. and Larsson,N.-G. (2010) LRPPRC is a mitochondrial matrix protein that is conserved in metazoans. *Biochem. Biophys. Res. Commun.*, **398**, 759–764.

44. Rondeau,E.B., Messmer,A.M., Sanderson,D.S., Jantzen,S.G., von Schalburg,K.R., Minkley,D.R., Leong,J.S., Macdonald,G.M., Davidsen,A.E., Parker,W.A., *et al.* (2013) Genomics of sablefish (Anoplopoma fimbria): expressed genes, mitochondrial

phylogeny, linkage map and identification of a putative sex gene. *BMC Genomics*, **14**, 452.

45. Hayman,E.S., Fairgrieve,W.T. and Adam Luckenbach,J. (2020) Molecular and morphological sex differentiation in sablefish (Anoplopoma fimbria), a marine teleost with XX/XY sex determination. *Gene*, 10.1016/j.gene.2020.145093.

46. Herpin,A. and Schartl,M. (2009) Molecular mechanisms of sex determination and evolution of the Y-chromosome: insights from the medakafish (Oryzias latipes). *Mol. Cell. Endocrinol.*, **306**, 51–58.

47. Herpin,A., Nakamura,S., Wagner,T.U., Tanaka,M. and Schartl,M. (2009) A highly conserved cis-regulatory motif directs differential gonadal synexpression of Dmrt1 transcripts during gonad development. *Nucleic Acids Res.*, **37**, 1510–1520.

48. Herpin,A., Braasch,I., Kraeussling,M., Schmidt,C., Thoma,E.C., Nakamura,S., Tanaka,M. and Schartl,M. (2010) Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. *PLoS Genet.*, **6**, e1000844.

49. Sekido,R. and Lovell-Badge,R. (2008) Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature*, **453**, 930–934.

50. Herpin,A. and Schartl,M. (2015) Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. *EMBO Rep.*, **16**, 1260–1274.

51. Charlesworth,B. (1991) The evolution of sex chromosomes. *Science*, **251**, 1030–1033.

52. Sheldon,P.R. (2001) Punctuated Equilibrium and Phyletic Gradualism. In *eLS*. American Cancer Society.

53. Aida,T. (1921) On the Inheritance of Color in a Fresh-Water Fish, APLOCHEILUS LATIPES Temmick and Schlegel, with Special Reference to Sex-Linked Inheritance. *Genetics*, **6**, 554–573.

54. Yamamoto,T.O. (1955) Progeny of Artificially Induced Sex-Reversals of Male Genotype (Xy) in the Medaka (Oryzias Latipes) with Special Reference to Yy-Male. *Genetics*, **40**, 406–419.

55. Nishimura,T., Sato,T., Yamamoto,Y., Watakabe,I., Ohkawa,Y., Suyama,M., Kobayashi,S. and Tanaka,M. (2015) Sex determination. foxl3 is a germ cell-intrinsic factor involved in sperm-egg fate decision in medaka. *Science*, **349**, 328–331.

56. Masuyama,H., Yamada,M., Kamei,Y., Fujiwara-Ishikawa,T., Todo,T., Nagahama,Y. and Matsuda,M. (2012) Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka. *Chromosome Res.*, **20**, 163–176.

57. Nanda,I., Hornung,U., Kondo,M., Schmid,M. and Schartl,M. (2003) Common spontaneous sex-reversed XX males of the medaka Oryzias latipes. *Genetics*, **163**, 245–251.

58. Schartl,M. (2004) A comparative view on sex determination in medaka. *Mech. Dev.*, **121**, 639–645.

59. Gautier,A., Le Gac,F. and Lareyre,J.-J. (2011) The gsdf gene locus harbors evolutionary conserved and clustered genes preferentially expressed in fish previtellogenic oocytes. *Gene*, **472**, 7–17.

60. Sawatari,E., Shikina,S., Takeuchi,T. and Yoshizaki,G. (2007) A novel transforming growth factor-beta superfamily member expressed in gonadal somatic cells enhances primordial germ cell and spermatogonial proliferation in rainbow trout (Oncorhynchus mykiss). *Dev. Biol.*, **301**, 266–275.

61. Shibata,Y., Paul-Prasanth,B., Suzuki,A., Usami,T., Nakamoto,M., Matsuda,M. and Nagahama,Y. (2010) Expression of gonadal soma derived factor (GSDF) is spatially and temporally correlated with early testicular differentiation in medaka. *Gene Expr. Patterns*, **10**, 283–289.

62. Karl,J. and Capel,B. (1995) Three-dimensional structure of the developing mouse genital ridge. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **350**, 235–242.

63. Yao,H.H.-C., DiNapoli,L. and Capel,B. (2003) Meiotic germ cells antagonize mesonephric cell migration and testis cord formation in mouse gonads. *Development*, **130**, 5895–5902.

64. Nakamura,S., Kobayashi,D., Aoki,Y., Yokoi,H., Ebe,Y., Wittbrodt,J. and Tanaka,M. (2006) Identification and lineage tracing of two populations of somatic gonadal precursors in medaka embryos. *Dev. Biol.*, **295**, 678–688.

65. Kobayashi,T., Matsuda,M., Kajiura-Kobayashi,H., Suzuki,A., Saito,N., Nakamoto,M., Shibata,N. and Nagahama,Y. (2004) Two DM domain genes, DMY and DMRT1, involved in testicular differentiation and development in the medaka, Oryzias latipes. *Dev. Dyn.*, **231**, 518–526.

66. Nakamura,S., Kobayashi,K., Nishimura,T., Higashijima,S. and Tanaka,M. (2010) Identification of germline stem cells in the ovary of the teleost medaka. *Science*, **328**, 1561–1563.

67. Falconi,R., Dalpiaz,D. and Zaccanti,F. (2004) Ultrastructural aspects of gonadal morphogenesis in Bufo bufo (Amphibia Anura) 1. Sex differentiation. *J. Exp. Zoolog. Part A Comp. Exp. Biol.*, **301**, 378–388.

68. Tanimura,A. and Iwasawa,H. (1988) Ultrastructural Observations on the Origin and Differentiation of Somatic Cells during Gonadal Development in the Frog Rana nigromaculata. *Development, Growth & Differentiation*, **30**, 681–691.

69. Merchant-Larios,R. and Villalpando,I. (1981) Ultrastructural events during early gonadal development in Rana pipiens and Xenopus laevis. *Anat. Rec.*, **199**, 349–360.

70. Tanimura,A. and Iwasawa,H. (1989) Origin of somatic cells and histogenesis in the primordial gonad of the Japanese tree frog Rhacophorus arboreus. *Anat. Embryol.*, **180**, 165–173.

71. Cheng,T.-H. (1932) The germ cell history of Rana cantabrigensis Baird: I. Germ cell origin and gonad formation Springer-Verlag, Berlin Heidelberg.

72. Christensen,K. (1930) Sex differentiation and development of oviducts in rana pipiens. *American Journal of Anatomy*, **45**, 159–187.

73. Lopez,K. (1989) Sex differentiation and early gonadal development in Bombina orientalis (anura: Discoglossidae). *Journal of Morphology*, **199**, 299–311.

74. Rugh,R. (1951) The frog: its reproduction and development McGraw-Hill Book Company, Inc., New York.

75. Swingle,W.W. (1926) The germ cells of anurans. II. An embryological study of sex differentiation in Rana catesbeiana. *Journal of Morphology*, **41**, 441–546.

76. Development of Vertebrates by Witschi Emil - AbeBooks.

77. Normal table of Xenopus laevis (Daudin) : a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis / edited by P.D. Nieuwkoop and J. Faber - Trove.

78. Vannini,E. and Sabbadin,A. The Relation of the Interrenal Blastema to the Origin of the Somatic Tissues of the Gonad in Frog Tadpoles.

79. Devlin,R.H. and Nagahama,Y. (2002) Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture*, **208**, 191–364.

80. Horsfield,J., Ramachandran,A., Reuter,K., LaVallie,E., Collins-Racie,L., Crosier,K. and Crosier,P. (2002) Cadherin-17 is required to maintain pronephric duct integrity during zebrafish development. *Mech Dev*, **115**, 15–26.

81. Livet,J., Weissman,T.A., Kang,H., Draft,R.W., Lu,J., Bennis,R.A., Sanes,J.R. and Lichtman,J.W. (2007) Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, **450**, 56–62.

82. Pan,Y.A., Livet,J., Sanes,J.R., Lichtman,J.W. and Schier,A.F. (2011) Multicolor Brainbow imaging in zebrafish. *Cold Spring Harb Protoc*, **2011**, pdb.prot5546.

83. Centanin,L., Hoeckendorf,B. and Wittbrodt,J. (2011) Fate restriction and multipotency in retinal stem cells. *Cell Stem Cell*, **9**, 553–562.

84. Heule,C., Salzburger,W. and Böhne,A. (2014) Genetics of sexual development: an evolutionary playground for fish. *Genetics*, **196**, 579–591.

85. Uller,T. and Helanterä,H. (2011) From the origin of sex-determining factors to the evolution of sex-determining systems. *Q Rev Biol*, **86**, 163–180.

86. Baroiller,J.F., D'Cotta,H., Bezault,E., Wessels,S. and Hoerstgen-Schwark,G. (2009) Tilapia sex determination: Where temperature and genetics meet. *Comp. Biochem. Physiol., Part A Mol. Integr. Physiol.*, **153**, 30–38.

87. Chen,S.-L., Tian,Y.-S., Yang,J.-F., Shao,C.-W., Ji,X.-S., Zhai,J.-M., Liao,X.-L., Zhuang,Z.-M., Su,P.-Z., Xu,J.-Y., *et al.* (2009) Artificial gynogenesis and sex determination in half-smooth tongue sole (Cynoglossus semilaevis). *Mar. Biotechnol.*, **11**, 243–251.

88. Yamamoto,Y., Zhang,Y., Sarida,M., Hattori,R.S. and Strüssmann,C.A. (2014) Coexistence of genotypic and temperature-dependent sex determination in pejerrey Odontesthes bonariensis. *PLoS ONE*, **9**, e102574.

89. Hinck,A.P., Mueller,T.D. and Springer,T.A. (2016) Structural Biology and Evolution of the TGF-β Family. *Cold Spring Harb Perspect Biol*, **8**.

90. Josso,N., Racine,C., di Clemente,N., Rodolfo Rey and Xavier,F. (1998) The role of anti-Müllerian hormone in gonadal development. *Molecular and Cellular Endocrinology*, **145**, 3–7.

91. di Clemente,N., Josso,N., Gouédard,L. and Belville,C. (2003) Components of the anti-Müllerian hormone signaling pathway in gonads. *Mol. Cell. Endocrinol.*, **211**, 9–14.

92. Cutting,A., Chue,J. and Smith,C.A. (2013) Just how conserved is vertebrate sex determination? *Dev. Dyn.*, **242**, 380–387.

93. Klüver,N., Pfennig,F., Pala,I., Storch,K., Schlieder,M., Froschauer,A., Gutzeit,H.O. and Schartl,M. (2007) Differential expression of anti-Müllerian hormone (amh) and anti-Müllerian hormone receptor type II (amhrII) in the teleost medaka. *Dev. Dyn.*, **236**, 271–281.

94. Schulz,R.W., Bogerd,J., Male,R., Ball,J., Fenske,M., Olsen,L.C. and Tyler,C.R. (2007) Estrogen-induced alterations in amh and dmrt1 expression signal for disruption in male sexual development in the zebrafish. *Environ. Sci. Technol.*, **41**, 6305–6310.

95. Forconi,M., Canapa,A., Barucca,M., Biscotti,M.A., Capriglione,T., Buonocore,F., Fausto,A.M., Makapedua,D.M., Pallavicini,A., Gerdol,M., *et al.* (2013) Characterization of sex determination and sex differentiation genes in Latimeria. *PLoS ONE*, **8**, e56006.

96. Zhang,X., Guan,G., Li,M., Zhu,F., Liu,Q., Naruse,K., Herpin,A., Nagahama,Y., Li,J. and Hong,Y. (2016) Autosomal gsdf acts as a male sex initiator in the fish medaka. *Sci Rep*, **6**, 19738.

97. Lawson,K.A., Dunn,N.R., Roelen,B.A., Zeinstra,L.M., Davis,A.M., Wright,C.V., Korving,J.P. and Hogan,B.L. (1999) Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.*, **13**, 424–436.

98. Ying,Y., Qi,X. and Zhao,G.Q. (2001) Induction of primordial germ cells from murine epiblasts by synergistic action of BMP4 and BMP8B signaling pathways. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 7858–7862.

99. Pangas,S.A. (2012) Bone morphogenetic protein signaling transcription factor (SMAD) function in granulosa cells. *Mol. Cell. Endocrinol.*, **356**, 40–47.

100. Reichwald,K., Petzold,A., Koch,P., Downie,B.R., Hartmann,N., Pietsch,S., Baumgart,M., Chalopin,D., Felder,M., Bens,M., *et al.* (2015) Insights into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish. *Cell*, **163**, 1527–1538.

101. Mazerbourg,S., Sangkuhl,K., Luo,C.-W., Sudo,S., Klein,C. and Hsueh,A.J.W. (2005) Identification of receptors and signaling pathways for orphan bone morphogenetic protein/growth differentiation factor ligands based on genomic analyses. *J. Biol. Chem.*, **280**, 32122–32132.

102. Sèdes,L., Leclerc,A., Moindjie,H., Cate,R.L., Picard,J.-Y., di Clemente,N. and Jamin,S.P. (2013) Anti-Müllerian hormone recruits BMPR-IA in immature granulosa cells. *PLoS ONE*, **8**, e81551.

103. Pasquier,J., Cabau,C., Nguyen,T., Jouanno,E., Severac,D., Braasch,I., Journot,L., Pontarotti,P., Klopp,C., Postlethwait,J.H., *et al.* (2016) Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. *BMC Genomics*, **17**, 368.

104. Kaneko,H., Ijiri,S., Kobayashi,T., Izumi,H., Kuramochi,Y., Wang,D.-S., Mizuno,S. and Nagahama,Y. (2015) Gonadal soma-derived factor (gsdf), a TGF-beta superfamily gene, induces testis differentiation in the teleost fish Oreochromis niloticus. *Mol. Cell. Endocrinol.*, **415**, 87–99.

105. Schartl,M., Schories,S., Wakamatsu,Y., Nagao,Y., Hashimoto,H., Bertin,C., Mourot,B., Schmidt,C., Wilhelm,D., Centanin,L., *et al.* (2018) Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. *BMC Biology*, **16**, 16.

106. Zhang,J. (2004) Evolution of DMY, a newly emergent male sex-determination gene of medaka fish. *Genetics*, **166**, 1887–1895.

107. Shinomiya,A., Kato,M., Yaezawa,M., Sakaizumi,M. and Hamaguchi,S. (2006) Interspecific hybridization between Oryzias latipes and Oryzias curvinotus causes XY sex reversal. *J. Exp. Zoolog. Part A Comp. Exp. Biol.*, **305**, 890–896.

4- Selected publications

Following are four selected publications that do not necessarily represent the "keystones of my scientific career", but are rather representative of "my" science the way I like it.

1- Herpin, A., Nakamura, S., Wagner, T., Tanaka, M. and Schartl, M. **(2009) A highly conserved mRNA sequence motif directs differential gonadal regulation of mRNA during gonad development,** ***Nucleic Acids Research.*** *37(5):1510-1520.*

This is a pure "hard-core" mechanistic paper, where we isolated a conserved *cis*-regulatory motif involved in the regulation of the Dmrt1 RNA stability in medaka. It is mainly molecular biology experiments (EMSA, cell transfections, DNA constructions and transgenic animals) that show how we can explain the physiology. I am particularly proud of this work because it later allowed clinicians to explain gonadal dysgenesis in a family of patients that were carrying mutation in this evolutionary conserved (from drosophila to mammals) cis regulatory motif we discovered and characterized in medaka ("Novel DMRT1 3'UTR+11 insT mutation associated to XY partial gonadal dysgenesis", Arq Bras Endocrinol Metab, 2010).

2- Herpin, A., Adolfi, M., Nicol, B., Hinzmann, M., Schmidt, C., Klughammer, J., Engel, M., Tanaka, M., Guiguen, Y. and Schartl, M. **(2013) Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination,** *Mol Biol Evol. 30(10): 2328-46.*

With this paper we challenge the paradigm that says that downstream effectors of sex determination are well conserved ("Masters change, slaves remain", (2)). By saying "we challenge the paradigm", I mean that for the first time we provide functional proof that the paradigm is not that universal. This is what I like, providing functional evidences to test evolutionary hypotheses.

3- Herpin, A., Schartl, M. **(2015) Plasticity of the gene regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers and usurpators,** ***EMBO reports.*** *16(10):1260-74.*

This is a review/perspective paper that sums up all we are working for. Here we challenge and explain most of the theories for sex determination and sex chromosome evolution in the light of functional studies, and give perspectives about it. We eventually conceptualize the notion of interacting gene regulatory networks for sex determination.

4- (Herpin, A., Rafati, N., Chen, J.), Petterson, M., Han, F., Feng, C., Wallerman, O., Rubin, CJ., Péron, S., Cocco, A., Larsson, M., Trötschel, C., Poetsch, A., Korsching, K., Bönigk, W., Körschen, HG., Berg, F., Folkvord, A., Kaupp, UB., Schartl, M. and Andersson, L. **(2020) Reconstruction of the birth of an unexpected male sex chromosome in Atlantic herring,** ***PNAS.***

I like this paper because, using genomics together with functional experiments (from genome to genes and function) we could reconstruct the birth of a sex chromosome…which was, according to some theories, predicted to be absent/unexpected.

Nucleic Acids Research Advance Access published January 12, 2009

Nucleic Acids Research, 2009, 1–11
doi:10.1093/nar/gkn1065

A highly conserved *cis*-regulatory motif directs differential gonadal synexpression of Dmrt1 transcripts during gonad development

Amaury Herpin[1,*], Shuhei Nakamura[2], Toni Wagner[1], Minoru Tanaka[2,*] and Manfred Schartl[1,3]

[1]Physiological Chemistry I, University of Würzburg, Biozentrum, Am Hubland, D-97074 Würzburg, Germany, [2]Laboratory of Molecular Genetics for Reproduction, National Institute for Basic Biology 5-1, Higashiyama, Okazaki, Aichi 444-8787, Japan and [3]Rudolf-Virchow-Center for Experimental Biomedicine (DFG Research Center), University of Würzburg, D-97078 Würzburg, Germany

Received November 17, 2008; Revised December 16, 2008; Accepted December 18, 2008

ABSTRACT

Differential gene expression largely accounts for the coordinated manifestation of the genetic programme underlying embryonic development and cell differentiation. The 3′ untranslated region (3′-UTR) of eukaryotic genes can contain motifs involved in regulation of gene expression at the post-transcriptional level. In the 3′-UTR of *dmrt1*, a key gene that functions in gonad development and differentiation, an 11-bp protein-binding motif was identified that mediates gonad-specific mRNA localization during embryonic and larval development of fish. Mutations that disrupt the 11-bp motif leading to *in vitro* protein-binding loss and selective transcript stabilization failure indicate a role for this motif in RNA stabilization through protein binding. The sequence motif was found to be conserved in most of the *dmrt1* homologous genes from flies to humans suggesting a widespread conservation of this specific mechanism.

INTRODUCTION

Assembly and formation of the gonad primordium is the first step towards gonad differentiation and subsequent sex differentiation (1). Primordial germ cells (PGCs) give rise to the gametes that are responsible for the development of a new organism in the next generation. In many organisms, following their specification the germ cells migrate towards the location of the prospective gonadal primordium (2–5). Similar to other vertebrates, the structure of fish gonads is composed of germ cells and associated supporting somatic cells (6). The precursors of the somatic cells originate from cells of the lateral plate mesoderm where the gonadal primordium develops, while germ cells are derived from the germline lineage (7,8). To carry out their highly specialized biological functions, together somatic gonadal primordium and germline cells must establish specialized programs of gene expression. However, the early transcriptional and post-transcriptional regulatory events underlying the differentiation of gonad precursor cells through crucial interactions of somatic and germline cells are barely understood.

The *dmrt1* gene is an important regulator of male development in vertebrates (9). It is a highly conserved gene involved in the determination and early differentiation phase of the primordial gonad in vertebrates. In the fish medaka *dmrt1bY*, a functional duplicate of the autosomal *dmrt1a* gene on the Y-chromosome, has been shown to be the master regulator of male gonadal development (10,11), comparable to *Sry* in mammals (12). In males mRNA and protein expression occur before morphological sex differentiation in the somatic cells surrounding PGCs of the gonadal anlage and later on exclusively in Sertoli cells (13). Here it is synexpressed with the autosomal *dmrt1a* (14,15). However, nothing is known about the mechanism(s) that bring about this highly restricted expression pattern.

The expression of most genes is dynamically regulated temporally and spatially. Spatial organization of cells and subcellular compartments arises in part from the sorting and subsequent localization of proteins and RNA. Evidence has been obtained that regulation occurs at multiple steps on the level of gene expression including

*To whom correspondence should be addressed. Email: amaury.herpin@biozentrum.uni-wuerzburg.de
Correspondence may also be addressed to Minoru Tanaka. Tel/Fax: +81 564 59 5851; Email: mtanaka@nibb.ac.jp

transcription, splicing, mRNA transport, mRNA stability, translation, protein stability and post-translational modifications (16,17). Selective advantages could have favored the evolution of regulatory mechanisms at the post-transcriptional level, such as speed of response, reversibility, fine-tuning of protein amounts, coordinated regulation of protein families, potential for spatial control, and efficacy in systems lacking transcriptional control mechanisms. Efficient cell-specific mRNA processing depends on a temporally and spatially orchestrated sequence of protein–protein, protein–RNA and RNA–RNA interactions (16).

Striking examples of localized messengers, transcriptional, post-transcriptional and translational regulations can be found among the maternal mRNAs of fly, fish and frog implicated in the establishment of axial polarity. For example, in the posterior part of the *Drosophila* embryo, Nanos (Nos) protein represses translation of maternal *hb* mRNA (18). Conversely, synthesis of both Bcd and Hb proteins in the anterior of the embryo requires that Nos is limited to the posterior (19). The restricted distribution of Nos is generated by selective translation of a subset of *nos* mRNA that is localized to the germ plasm at the posterior of the embryo coupled with translational repression of *nos* mRNA distributed throughout the whole embryonic cytoplasm (20). Both posterior localization and translational repression of *nos* RNA are mediated by the *nos* 3′-untranslated region (3′-UTR) (21). A nucleotide translational control element (TCE) within the *nos* 3′-UTR confers repression through formation of two stem-loop structures, whose functions are temporally distinct (22).

The zebrafish *nanos1* homologue which is required during germline development (23) has also been shown to be remarkably post-transcriptionally regulated. Here, microRNA miR-430 targets the 3′-UTR of *nanos1* during zebrafish embryogenesis in order to confer restriction of mRNA to PGCs (24). This miR-430 target site was shown to reduce poly(A) tail length, mRNA stability and translation, suggesting that differential susceptibility to microRNAs contributes to tissue-specific gene expression (24). Implicit in these mechanisms are the existence of *cis*-acting signals and *trans*-acting factors forming mRNA–protein complexes (mRNPs) that account for specificity and selectivity.

While the importance of complex post-transcriptional regulation—like in the case of *nanos*- has been widely demonstrated for the development of the germline, such mechanisms have not been uncovered so far for the development of the somatic part of the gonad, which determines the development towards testis or ovary. Studying mechanisms regulating localization and translation of gonad-specific genes during early gonad induction, we demonstrate that an 11-nt protein-binding motif located in the 3′-UTR of *dmrt1bY* mediates gonad-specific mRNA stability during embryonic and larval development. Interestingly, the sequence motif was found to be highly conserved in the homologous genes from flies to humans.

MATERIALS AND METHODS

Fish maintenance and breeding

Medaka were taken from closed breeding stocks of the Carbio strain and kept under standard conditions. Medaka embryos were staged according to Iwamatsu (25).

Whole-mount *in situ* hybridization

RNA whole-mount *in situ* hybridization using GFP digoxigenin (DIG)-labeled probe was performed as described for Medaka (26). Briefly, after capped mRNA injection, embryos of the desired stage were fixed with 4% paraformaldehyde and dehydrated with methanol. Antisense-DIG-labeled RNA probes were synthesized according to the manufacturers' instructions (Roche, Meylan). Hybridization and detection with alkaline phosphatase (AP)-coupled anti-DIG antibody (Roche, Meylan) were performed according to Thisse *et al*. (27).

Plasmid constructs and RNA injections

To obtain RNA transcripts of *eGFP* the *GFP* open reading frame (ORF) from *pEGFP-N1* (Clontech Laboratories) was inserted (*BamH1*/*Not1*) into *pCS2+* plasmid (*pCS2:GFP*). To produce *pCS2:GFP:dmrt1bY 3′-UTR*, a PCR product containing the entire 3′-UTR of *dmrt1bY* flanked by *Not1* sites, was amplified from medaka testes and inserted into the *Not1* sites of *pCS2:GFP*. Similarly, *pCS2:GFP:dmrt1a 3′-UTR, pCS2:GFP:d. rerio dmrt1bY 3′-UTR, pCS2:GFP:h. sapiens 3′-UTR, pCS2:GFP:O. curvinotus dmrt1bY 3′-UTR and pCS2:GFP:fugu 3′-UTR* and deletion constructs (Supplementary Figure 2) were constructed the same way. Corresponding RFP plasmids were constructed by replacing the *GFP* ORF by *RFP*. *Xenopus β-globin* constructs were produced by inserting the *xenopus β-globin* 3′-UTR (*Not1*/*Kpn1*) from plasmid pRN3 (28) into *pCS2:GFP* (*pCS2:GFP:xlβ-globin* 3′-UTR). For constructing the *pCS2:GFP:BOXxlβ-globin* 3′-UTR plasmid, the *dmrt1bY* box was inserted between the *Not1* sites. The GFP/RFP 3′-UTR constructs include the mmGFP5/RFP ORF cloned upstream of the 3′-UTR of the zebrafish nanos1 gene (23,29). All constructs were checked by restriction digests, diagnostic PCRs and sequencing. Sertoli cell-specific Ds-Red expressing sox9prom:DsRed transgenic medaka fish (30) was provided by Prof. Tanaka.

Capped RNAs for injections were transcribed from linearized vectors using the SP6/T3/T7 m MESSAGE mMACHINE Kit (Ambion). One nanoliter was injected into the cytoplasm of one-cell stage Medaka embryos as described (31).

Cell culture and transfection

Medaka spermatogonial (Sg3), embryonic stem (MES-1) or fibroblast (OL-17) cells were cultured as described (32–34). Cells were grown to 70–80% confluency in six-well plates and then transfected with 5 μg expression vector using GeneJuice reagent (Novagen) as described by the manufacturer. Luciferase activity was then quantified using the Luciferase Reporter Assay System from

Promega and luciferase activity was normalized against mRNA luciferase copy number. Transcriptional differences between luciferase constructs were evaluated statistically by paired Student's test.

Electromobility shift assay

Nearly confluent cells [Medaka spermatogonial (Sg3), mouse Sertoli (TM4) and Medaka embryonic stem (MES-1)] cell lines grown on plates were washed with ice-cold phosphate-buffered saline (PBS) and removed from the plates with 1 mM EDTA in PBS. The cells were centrifuged at low speed and then resuspended in Passive Lysis Buffer (Promega) supplemented with leupeptin (0.2 μg/ml) and aprotinin (10 μg/ml). Binding assays were carried out using the Gel Shift Assay System (Promega) using radiolabeled RNA oligonucleotides: (Y) UGGUUCACGU**CUGCUGCAGGU**CUCUGACUCU for the native box target sequence and Mut(2)-box UGG UUCACGU**UUGGUCGGGAT**CUCUGACUCU; Mut(3)-box UGGUUCACGU**UCUUCACAUGU**CUCU GACUCU; Mut(4)-box UGGUUCACGU**CUGCUGAG ACG**CUCUGACUCU; Mut(5)-box UGGUUCACGU**C UGCCAUAGGU**CUCUGACUCU as competitors.

RESULTS AND DISCUSSION

Dmrt1a/dmrt1bY 3′-UTRs regulate spatial and temporal expression during early Medaka development

While searching for potentially conserved regulatory sequences in genes involved in gonad induction and formation we analyzed the tightly regulated Medaka *dmrt1a* and *dmrt1bY* duplicated gene pair. Postulating that important *cis*-regulatory motifs required for mRNA regulation in the context of gonad formation might have been retained between the duplicates despite of the processes of co-ortholog gene specialization and subfunctionalization, we noticed that the *dmrt1a* and *dmrt1bY* 3′-UTRs appeared more conserved than expected for independently diverging genes. To test whether this conservation of *dmrt1a/dmrt1bY* UTRs implied common regulatory mechanisms, we examined GFP expression of reporter constructs that contained either the *dmrt1a* or *dmrt1bY* 3′-UTR or the *Xenopus* β-globin 3′-UTR as a control (Figure 1). After injection of the different constructs into one cell stage embryos we initially observed a high and uniformly distributed GFP expression in the whole embryo (Figure 1). After 2 days of development (stages 22–24; 12–16 somites), GFP fluorescence slowly vanished elsewhere except in the primordial gonad area of the fish injected with either *dmrt1a* or *dmrt1bY* UTR (Figures 1A, B, C, D compared to 1E, F, G, H).

GFP fluorescence was clearly detectable in the primordial gonad area until more than 7 days after hatching (stage 40; first fry stage, 2.5 weeks after fertilization) (Supplementary Figure 1). In controls with the β-globin 3′-UTR GFP remained ubiquitously expressed throughout the whole embryonic development. Obviously, the *dmrt1a/dmrt1bY* 3′-UTRs are responsible for specific expression of the GFP protein in the primordial gonad area.

Surprisingly, fusing either human (AJ276801) (Figure 1I, J and K) or takifugu (CAC42778) (Figure 1L, M and N) *dmrt1* 3′-UTRs to GFP mRNA also drove primordial gonad area-specific fluorescence in Medaka (Figure 1I–N compared to 1A–D). In addition, injection of GFP mRNA fused to Medaka *dmrt1bY* 3′-UTR in zebrafish resulted in a similar gonadal persistence of GFP fluorescence (Figure 1O, P and Q) indicating a functional cross-species conserved mechanism mediated by *cis*-regulatory element(s) in these *dmrt1* 3′-UTRs.

A short, highly conserved *cis*-regulatory motif located in *dmrt1a/dmrt1bY* 3′-UTR is responsible for gonadal differential regulation

To delineate the precise RNA sequence and/or secondary structures involved in gonadal-specific fluorescence, GFP expression of a series of reporters (Supplementary Figure 2) that contained deletion mutants of the *dmrt1a/dmrt1bY* 3′-UTRs was investigated (Supplementary Figure 2B, C and D). As a result a core 11-nt box located in the 5′ region of the Medaka *dmrt1a/dmrt1bY* UTRs was isolated and shown to be responsible for gonad-specific fluorescence. Consequently, when the box was inserted into the *Xenopus* β-globin 3′-UTR, GFP expression was identical to *GFP::dmrt1a* and *dmrt1bY* 3′-UTR constructs, namely gonad-specific expression (Figure 1R, S, T and U). Conversely, deletion of the box sequence from the *dmrt1bY* 3′-UTR drastically extinguished gonad-specific expression (Supplementary Figure 2D compared to 2B and C).

Considering the functional conservation of gonad-specific expression seen with other fish and human *dmrt1* UTRs we then searched for similar motifs in the 3′-UTR of *dmrt1* genes of other organisms (Table 1). The motif was found to be highly conserved in the fish lineage (*Oryzias latipes*, *Oryzias curvinotus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Epinephelus coioides* and *Danio rerio*), but as well in the *dmrt1* 3′-UTR of other vertebrates including man (*Mus musculus*, *Pan troglodytes*, *Macaca mulatta* and *Homo sapiens*) and most surprisingly even in the ecdysozoan clade. For the *doublesex* (*dsx*, the *dmrt1* orthologue) of *Anopheles gambiae* (Table 1), interestingly, the sex-specific differentially spliced anopheles *dsx* transcript results in a male *dsx* form where the box is largely conserved while this is not the case for the female splice form (Table 1). This situation is similar for *dsx* of the olive fruit fly (*Bactocera oleae*) for which a male-specific splicing leads to the preservation of a highly conserved box in the ORF while this fragment is spliced out in the female form (Table 1).

Taken together, a short highly conserved *cis*-regulatory motif located in *dmrt1a* and *dmrt1bY* 3′-UTRs (CUGCU GCAGGU) appears to be mostly responsible for differential expression of the transcripts.

The *dmrt1* box drives specific stability in the somatic mesoderm anlage of the gonadal primordium as well as in a sub-population of PGCs

Dmrt1 is expressed in most species specifically in Sertoli cells and PGCs. To find out the contribution of

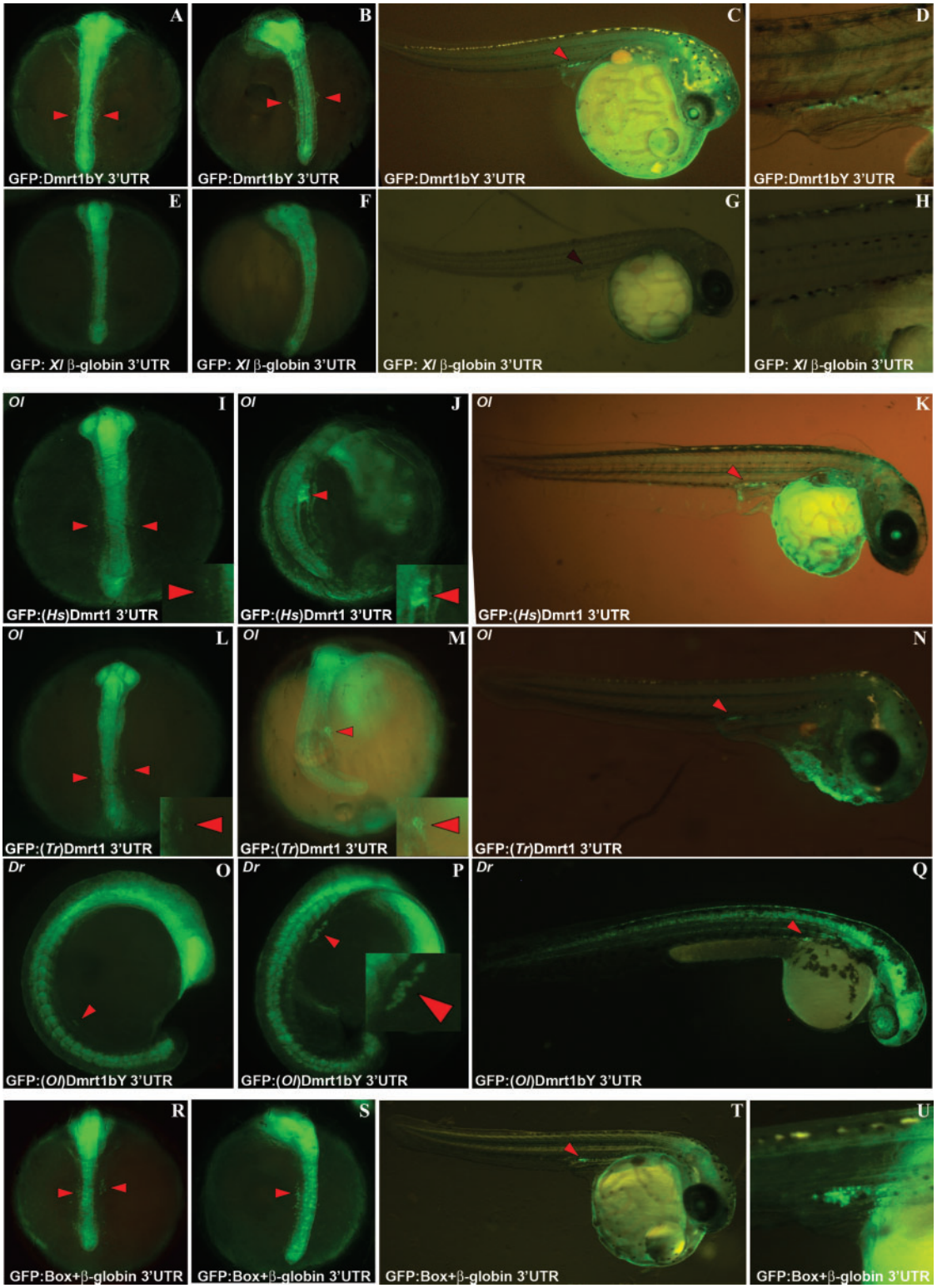


Figure 1. A short highly conserved *cis*-regulatory motif located in *dmrt1bY*/*dmrt1a* 3′-UTRs regulates spatial and temporal expression during early development. (**A–D**) GFP expression of a reporter construct that contains Medaka *dmrt1bY* 3′-UTR during somitogenesis (A and B) and at hatching stage (**C** and **D**). (**E–H**) GFP expression of a control reporter construct that contains *Xenopus* β-globin 3′-UTR during somitogenesis (E and F) and at hatching stage (G and H). (**I–Q**) GFP expression of reporter constructs that contain either human *dmrt1* 3′-UTR (I–K) or takifugu *dmrt1* 3′-UTR (**L**–**N**) in Medaka embryos during somitogenesis (I, J and L, M) and at hatching stage (K and N). (**O–Q**) GFP expression of a reporter construct that contains Medaka *dmrt1bY* 3′-UTR in zebrafish embryos during somitogenesis (O and P) and at hatching stage (Q). (**R–U**) GFP expression in Medaka embryos of a reporter construct that contains *Xenopus ß-globin* 3′-UTR in which the Box was inserted. Specific GFP expression in PGCs is indicated (arrow heads).

Table 1. The *dmrt1* 3′ UTR *cis*-regulatory motif is well conserved from ecdysozoans to mammals

Oryzias latipes	Dmrt1a3′-UTR(AF319994):	CUGCUGCAGGU
Oryzias latipes	DmrtlbY 3′-UTR (AB071534):	CUGCUGCAGGU
Oryzias curvinotus	DmrtlbY 3′-UTR (AB091695):	CUGCUGCAGGU
Oryzias curvinotus	Dmrt1ba 3′-TR (AB091696):	CUGCUGCAGGU
Takifugu rubripes	Dmrt1 3′-UR (CAC42778):	CUGCUGCAGGU
Tetraodon nigrovihdis	Dmrt1 3′-UTR (CAC42783):	CUGCUGCAGGU
Epinephelus coioides	Dmrt1 3′-UTR(EF017802):	TUGCUGCAGGU
Danio rerio	Dmrt1 3′-UTR(AAU11586):	CUGCUGCAGAU
Homo sapiens	Dmrt1 3′-UTR(AJ276801):	CUGCUGCCGAU
Pan troglodytes	Dmrt1 3′-UTR (XM528528):	CUGCUGCCGAU
Macaca mulatta	Dmrt1 3′-UTR (XM001090081):	CUGCUGCCGAU
Mus musculus	Dmrt1 3′-UTR (NM175647):	CUGGUGCUGAU
Anophele gambia	Dsx 3′-UR (Male-spe, AY903307):	CUGCUGUAAGU
Anophele gambia	Dsx 3′-UTR (Female-spe, AY903308):	CUG_UGCAGAA

Bactocera oleae (olive fruit fly) Doublesex ORF (AJ547622): 'Box' present in Male form (dsxm) but absent in Female form (dsxf)	CUGCUGCAGCU
Musca domestica Doublesex ORF (AY461854/AY461853): 'Box' present for both dsxm anf dsxf	CUGCUGCAGGA

post-transcriptional regulation by this *cis*-regulatory motif to the restricted expression pattern, Medaka *dmrt1bY* 3′-UTR was fused to the monomeric RFP mCherry and injected either together with *GFP::Nos* 3′-UTR into one-cell stage embryos of wild-type Medaka or of the Olvas transgenic strain. The *GFP::Nos* 3′-UTR construct was previously shown to drive PGC-specific fluorescence in medaka (35–37). Similarly, in Olvas fish the PGCs are marked by GFP expression from stage 25 [18–19 somite stage (25)] onwards (38). RFP expression was compared to germ cell-specific GFP expression due to the *nanos* 3′-UTR or the *vasa* promoter (Figure 2A–F). Starting at stages 14–18 fluorescence was exclusively observed from GFP in PGCs (data not shown; Figure 3A). Only by stages 22–24 (nine somite stage) red fluorescence could also be observed with specific gonadal localization (Figure 2D–F). Noteworthy, although at this stage both fluorescences (green and red) were confined to the same embryonic structures, they are clearly expressed from different cell populations (Figure 2A–C). Additionally, around stage 24–26 (16–22 somite stage) cells expressing both fluorochromes could be observed (Figure 2A–C). Subsequently, at hatching stage, gonadal *dmrt1bY* 3′-UTR-driven GFP expression was then investigated in Sertoli cell-specific RFP expressing sox9prom:DsRed transgenic fish (30) either at stage 34 (Figures 2G, I, H and J) or just after hatching (Figure 2K, L and M) in males and females, respectively. It revealed that although no more supporting cell expression was apparent for these later stages, two different populations of germ cells could be discriminated according to their high or just above background GFP fluorescence expression (arrowheads in Figure 2G, H, K and L). The fact that the RFP reporter gene product was only observed at later stages indicated that mRNAs with the *dmrt1bY* 3′-UTR are translationally repressed in migrating PGCs and in the somatic gonad precursor cells.

Interestingly, such variation and diversity in germ cell-specific Dmrt1 expression is also reported for mouse gonads for which two populations of germ cells could be observed according to Dmrt1 protein presence or absence (39). Consequently, akin to what was observed during medaka gonad formation (Figure 2G–N), it might be possible that mouse $Dmrt1^{+}$ and $Dmrt1^{-}$ expressing germ cells are the result of a similar mechanism since the 'box' motif is also found to be highly conserved in mouse dmrt1 3′-UTR (Table 1).

Embryonic expression of *dmrt1* mRNA has been examined in various vertebrates, including mammals (40–42), birds (41), reptiles (41,43) and fish (44). In most cases, *dmrt1* is expressed very early in the genital ridge–the structure from which the gonad derives–and later on in germ and Sertoli cells of the male gonadal primordium [see Zarkower *et al.* (45) for review]. It is then evident that

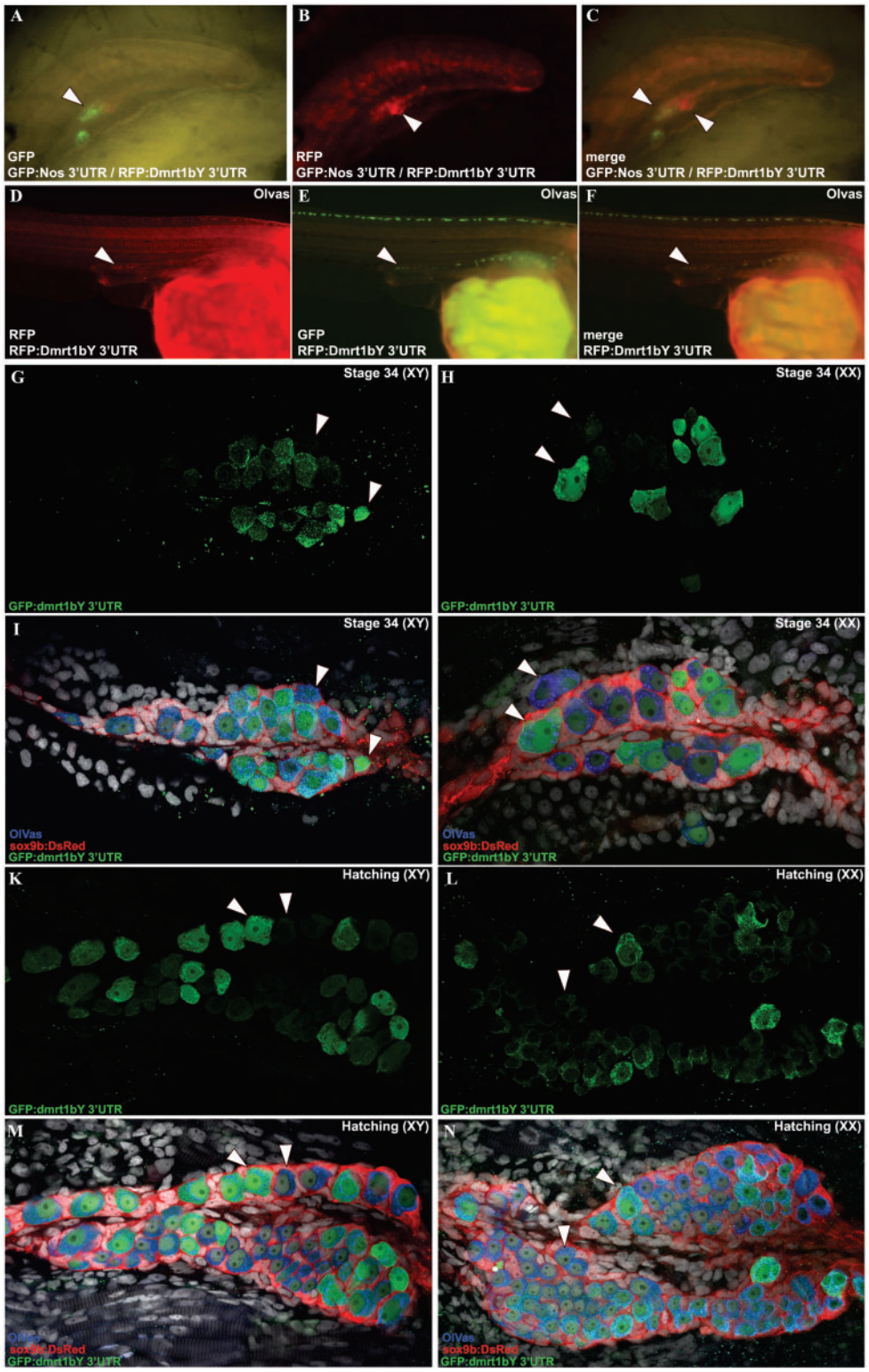


Figure 2. The *dmrt1* box drives specific stability in the somatic mesoderm of the gonadal primordium as well as in a sub-population of PGCs. (**A–F**) RFP expression from *dmrt1bY* 3′-UTR containing capped RNA compared to germ cell-specific GFP expression due to the *nanos* 3′-UTR at stages 24–26 (A–C). (D–F) RFP expression compared to germ cell specific GFP expression achieved with *vasa* promoter around hatching stage. (**G–N**) Gonadal *dmrt1bY* 3′-UTR-driven GFP expression investigated in Sertoli cell specific DsRed expressing sox9prom:DsRed transgenic fish either at stage 34 (G, I and H, J) or just after hatching (K, M and L, N) in males (XY) and females (XX). Arrow heads indicate either putative somatic gonadal precursor cells (C) or different sub-populations of germ cells (A, B and D–N). Blue: DAPI or Olvas; Red: Sertoli cell specific expression (sox9prom:DsRed transgenic fish) and Green: gonadal *dmrt1bY* 3′-UTR-driven GFP expression.

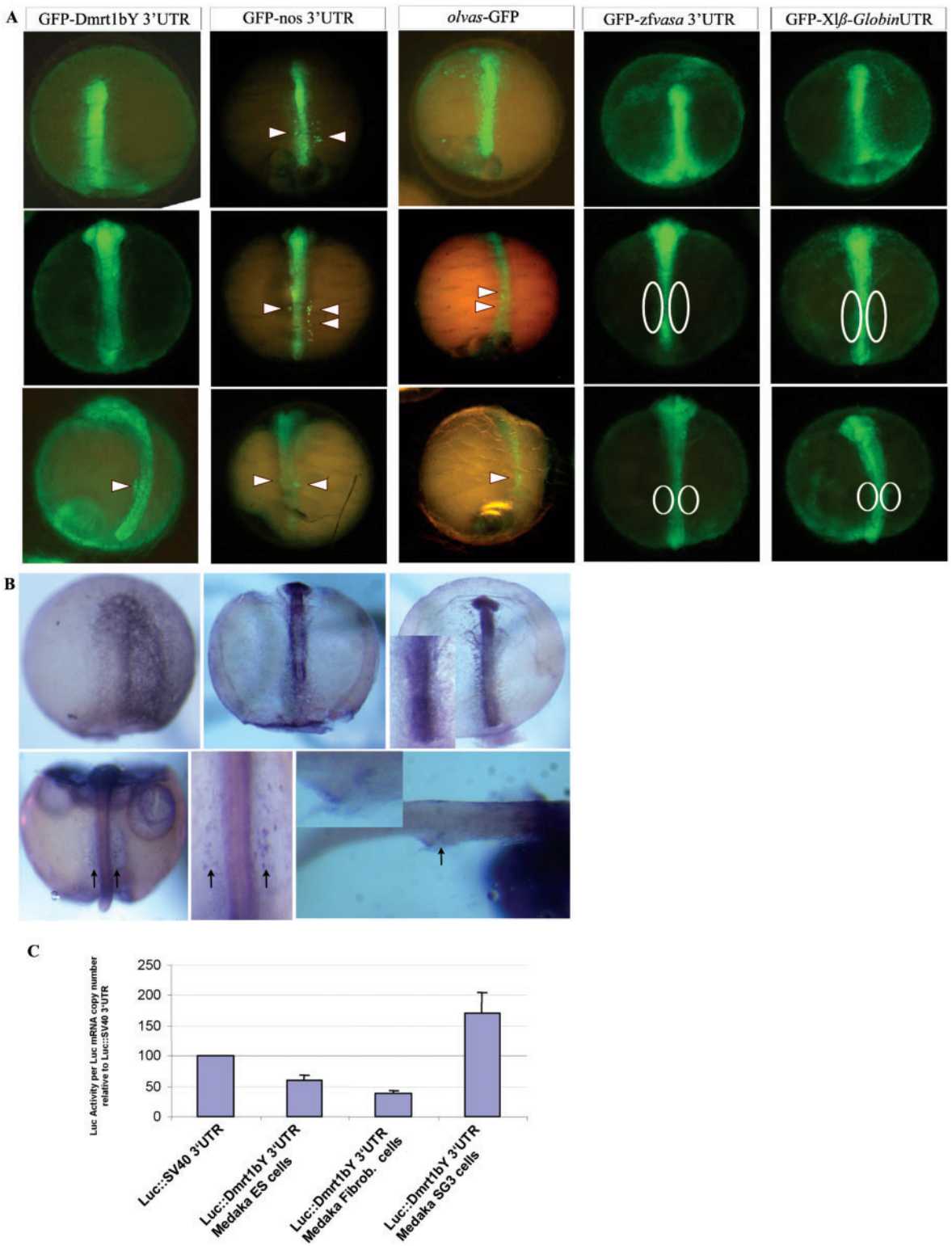


Figure 3. Tissue-specific and temporal-restricted expression by a combination of *dmrt1* 3′-UTR induced differential mRNA stability and translational regulation. (**A**) GFP:*dmrt1bY* 3′-UTR was injected. As controls, mRNA constructs such as GFP:*nos* 3′-UTR, *olvas*:GFP, GFP:*zfvasa* 3′-UTR and GFP:*xlß-globin* 3′-UTR were also injected to be able to compare the expression with the pattern of known post-transcriptional mechanisms such as micro-RNA mRNA induced decay, specific PGC translational regulation, and ubiquitous stability, respectively. (**B**) Subsequently *in situ* hybridization using an antisense GFP probe was performed at different stages of development to reveal the spatial distribution of the injected GFP:*dmrt1bY* 3′-UTR RNAs. GFP fluorescence and the distribution of RNA were followed at different stages of development. Arrowheads and circles indicate where the PGCs are or should be located respectively. (**C**) Luciferase expression in different cell lines transfected with a *dmrt1bY* 3′-UTR containing construct reveals that translation of the reporter gene was significantly enhanced by the presence of *dmrt1bY* 3′-UTR in Medaka spermatogonial cells in contrast to either Medaka embryonic stem or fibroblast cells.

GFP expressing cell are PGCs; likewise are the cells expressing jointly both GFP and RFP. The question remains what the identity of the RFP only expressing cells at early stages (stages 24–26) is. Considering the temporal appearance of these cells and a report identifying and tracing different populations of gonadal precursor cells in medaka (7), it is most likely that these cells are representatives of a subset of gonadal precursors.

Interestingly, like observed for mouse Dmrt1, these results point to the existence of at least two populations of PGCs within the primordial gonad of Medaka with possibly different regulations of *dmrt1* in relation to the above described *cis*-acting element located in its 3′-UTR.

Tissue-specific and temporal-restricted expression is caused by a combination of *dmrt1* 3′-UTR-induced differential mRNA stability and translational regulation

Independently of promoter-induced transcription striking examples of specifically localized messengers can be found among the mRNAs of fly, fish and frog, most of them being post-transcriptionally (dynamic relocalization and stabilization) and translationally regulated. To test whether *dmrt1bY* 3′-UTR was more involved in regulating mRNA stability or controlling its translational regulation, embryos were injected with *GFP:dmrt1bY* 3′-UTR and subsequently analyzed for the spatial distribution of the injected RNA by RNA *in situ* hybridization at different stages of development (Figure 3). As controls, *GFP:nos* 3′-UTR, *olvas:GFP*, *GFP:zfvasa* 3′-UTR and *GFP:xlß-globin* 3′-UTR mRNA constructs were also injected in order to compare the expression with the pattern of known post-transcriptional mechanisms such as microRNA mRNA-induced decay, specific PGC translational regulation and ubiquitous stability, respectively (Figure 3). GFP fluorescence and the distribution of RNA were followed at different stages of development. Remarkably, GFP expressed from the construct containing the *dmrt1bY* 3′-UTR, was shown to be quite stable in the whole body, including the primordial gonad area like with the GFP:*zfvasa* 3′-UTR and GFP:*xlß-globin* 3′-UTR mRNAs (Figure 3). This is in striking contrast to GFP:*nos 3′*-UTR construct for which mRNA underwent rapid somatic degradation [Figure 3; (22,24,36,46)]. This pattern of decay is typical for a microRNA-mediated process (24). Although an analogous germ cell-specific fluorescence was observed in germ cells of *olvas*:GFP injected embryos, *in situ* hybridizations showed homogenous olvas RNA distribution. This pattern implies germ cell-specific translational regulation (35). Accordingly GFP *in situ* hybridization after GFP:*dmrt1bY* 3′-UTR injection revealed that the fused mRNA is first homogenously distributed in the whole embryonic body until stage 24 and only then becomes progressively restricted to the gonad area to be exclusively present in the gonadal primordium around hatching (stage 39) (Figure 3). Hence, it can be inferred that the apparent specific GFP:*dmrt1*-UTR driven expression in primordial gonad is the result of differential RNA stabilization.

Interestingly these results reflect also the probability of another mechanism affecting mRNAs containing the *dmrt1* UTR. The fact that strong fluorescence is observed in the whole body (including PGCs) at early stages suggests that the *dmrt1bY* box here contributes to enhanced translation. To test whether the *dmrt1bY* 3′-UTR is able to enhance translation, a plasmid construct containing a thymidine kinase promoter driven luciferase fused either to SV40 3′-UTR or *dmrt1bY* 3′-UTR was transfected in different cell lines and translation efficiency measured. Translation of the reporter gene was significantly ($P < 0.01$) enhanced in the presence of *dmrt1bY* 3′-UTR in Medaka spermatogonial cells in contrast to either Medaka embryonic stem cells or fibroblast cells where protein production was even reduced (Figure 3). The suppression of translation in non-gonad cell types is interesting to note in relation to the non-gonad-specific expression of *dmrt1bY* transcripts during early Medaka embryogenesis and in adult male spleen (10,14).

In summary, taking into account the endogenous expression of *dmrt1a* exclusively in the developing gonad and the here observed late gonad-specific translation of injected RFP:*dmrt1bY* 3′-UTR mRNA, it appears that the UTR mediates a translational regulation process specifically in a subset of PGCs and certain somatic cells of the gonad. From RNA localization studies it is apparent that the later gonad-specific expression is then primarily due to the 11-bp *cis*-acting *dmrt1* motif conferring differential RNA stabilization in primordial gonad cells.

The *dmrt1* 3′-UTR box specifically binds a protein possibly involved in gonad-restricted expression

Obviously the 11-mer is involved in a cell-type-specific stability of *dmrt1* RNAs. One explanation could be that it is the target for a protein or small ribonucleoproteins particles. To test this hypothesis, electrophoretic mobility shift assays (EMSAs) were performed using the 11-mer box as a probe and different mutated boxes as competitors (Figure 4). In a first step cell extracts from three different cell lines were tested: (i) a medaka spermatogonia cell line, (ii) a medaka embryonic stem cell line and (iii) a mouse sertoli cell line (TM4) (Figure 4A). While no mobility shift was observed for the medaka embryonic stem cell-like line in all conditions tested, a weak shift was apparent for the TM4 cells and a quite robust one for the Medaka spermatogonial cell line (Sg3) (Figure 4A). As controls, using the medaka spermatogonial cell line, a mutated version of the box used as competitor (Mut(3)-Box) did not interfere significantly with the Dmrt1bY-Box binding, indicating the specificity of the interaction (Figure 4B). Furthermore, competition of radioactively with non-radioactively labeled box probes (Dmrt1bY-Box) resulted in progressive loss of the apparent shift (Figure 4C). Next, to better characterize the interaction domain(s) of the box, different point mutations were then introduced to the native *dmrt1bY*-box (Figure 4D).

The first mutated RNA oligonucleotide [Mut(2)-Box] regrouped the three main mutations observed for the different dmrt1 boxes within fish and mammals (see also Table 1). These introduced mutations did not significantly interfere for binding (Figure 4D). Interestingly, this result

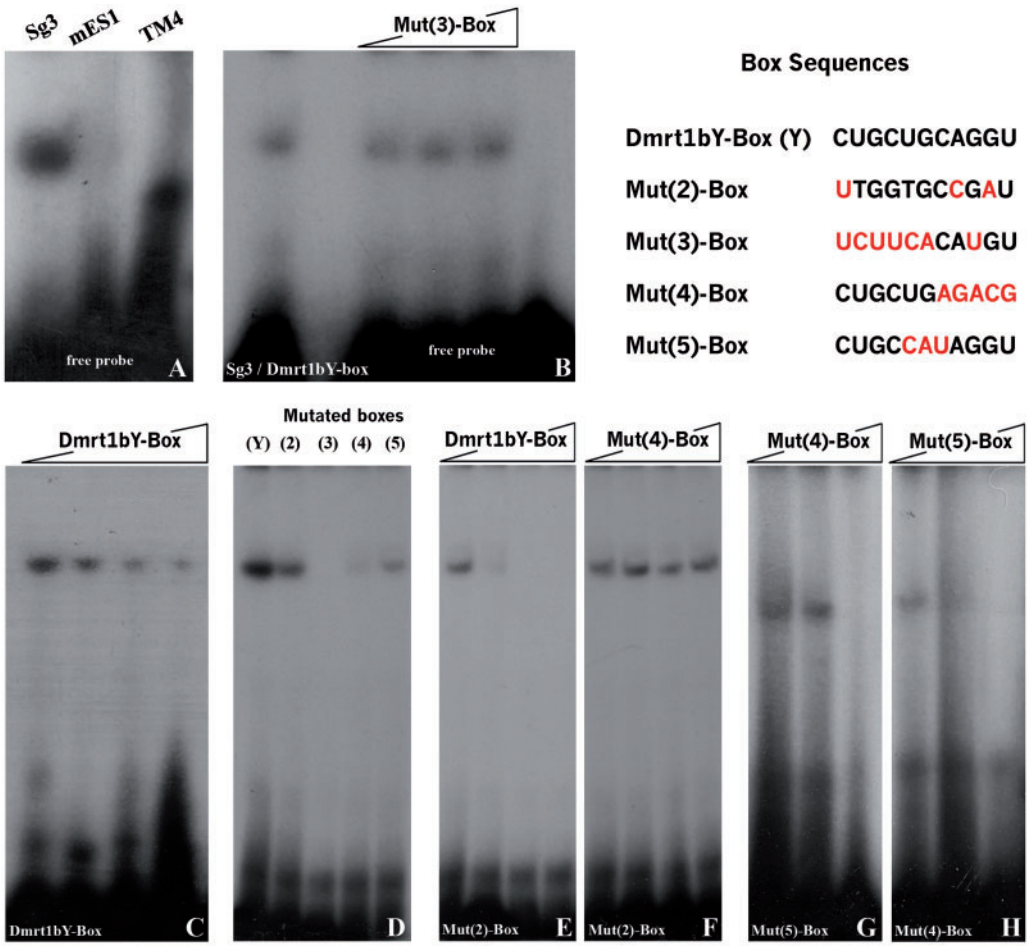


Figure 4. A RNA-binding factor that recognizes the Box could be responsible for primordial gonad restricted stability. (**A**) Electrophoretic Mobility Shift Assay (EMSA) using cell extracts from three different cell lines (i) a medaka spermatogonial cell line (Sg3), (ii) a medaka embryonic stem cell line (MES-1) and (iii) a mouse sertoli cell line (TM4) and the 11-mer box as a probe shows shift for Sg3 and TM4 cells. (**B**) A mutated version of the box [Mut(3)-Box] used as competitor (1–1, 1–5 and 1–10 ratios) did not interfere significantly with the binding, indicating the specificity of the binding. (**C**) As control, using the spermatogonial cell line, competition of radioactively and non-radioactively labeled box probes resulted in progressive loss of the apparent shift (1–1, 1–5 and 1–10 ratios). (**D**) Different mutated versions of the box [Mut(2–5)-Box] were then tested for binding, and resulted in apparent different binding affinities. (**E–H**) Relative robustness of the shift was then tested by mean of competition assays among the different mutated versions of the box (E: 1–1, 1–5 and 1–10; F: 1–1, 1–2 and 1–5; G and H: 1–1 and 1–5 ratios).

corroborates *in vivo* observations showing GFP stabilization after injection of either human or zebrafish dmrt1 3′-UTR containing constructs in medaka (Figure 4D). Nevertheless, this affinity seems to be significantly weaker when competing against the native Dmrt1bY-Box (Figure 4E). Noticeably, Mut(2)-Box competition against another box mutated in its 3′ region [Mut(4)-Box] could not appreciably interfere with the binding (Figure 4F). Similarly, using this box [Mut(4)-Box] and a box mutated in its core region [Mut(5)-Box] alternatively against each other, revealed that although the entire native box is required for efficient binding, its 3′ region is important while the 5′ and core regions are more likely involved in modulating the affinity (Figures 4G and 4H).

In summary these experiments suggest that the 11-mer box is a preferential target for RNA-binding proteins that may be involved in specifically regulating transcript stability.

CONCLUSIONS

The great majority of sequences that direct the subcellular localization, translation and degradation have been found within the 3′-UTRs [see (17,47) for review]. UTRs are highly diverse in sequence, but often contain regulatory motifs that are common to members of the same metabolic family (48). Nevertheless, the analysis of such *cis*-acting elements has been rather unsatisfying with respect to the identification of common sequences that direct localization of different RNAs (49).

We have identified an 11-nt motif in the Dmrt1 3′-UTR, CUGCUGCAGGU, common to the great majority of Dmrt1 orthologs genes from fly to mammals that is responsible for specific stabilization and translational control of *dmrt1* mRNA in the forming primordial gonad of fish and probably of other species. Interestingly, for the first time such stabilizing motif involving *cis*-regulatory actor(s) has been shown to be not only present in a

single organism [see Kloc *et al*. (49) for review], co-regulating a very specific pool of few synexpressed transcripts (50), but rather to be highly conserved across species for a given transcript family, namely the Doublesex/Mab-3/Dmrt1 genes. The occurrence of such an 11-mer in the 3′-UTR of orthologs genes over a wide range of organisms already indicates that different systems may employ common RNA regulatory mechanisms. Hence, other than a motif involved in specific localization of mRNAs to a well defined subtype of cells, this box would be responsible for mRNA-specific preservation of different subsets of cells all together involved in primordial gonad assembly and formation, namely putative gonadal precuror cells and a specific subclass of PGCs. Consequently, like synexpressed groups reflect the functional compartmentalization of the eukaryote genome (51) and have a striking parallel to the prokaryote operon, our finding might be of particular interest since it might reveal an otherwise hidden logic of cellular regulation where *cis*-regulatory motifs couple spatial and temporal gene expression in different subset of cells during organogenesis.

Finally, our data indicate that transcript stabilization is achieved by interaction of a specific protein with a *cis*-acting stability element located in the *dmrt1* 3′-UTR. Although the identity and the dynamic of action of this stabilizing factor has still to be resolved, our findings point to an obvious level of integrated regulation, namely the presence and the accessibility of this *cis*-regulatory element. The occurrence of multiple *dmrt1*-related spliced variants in corals (52), insects (53–55), lizards (56), fish (57,58), chicken (59), mice (60) and human (61) selecting or splicing out parts of the UTRs also argue for such way of regulation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

This work was supported by a grant of the Rudolf-Virchow-Zentrum for Experimental Medicine (DFG Forschungszentrum) and DFG-Graduiertenkolleg 1048 (Molecular Basis of Organ Development in Vertebrates) through a PhD fellowship to T.W. Funding for open access charge: XXX.

Conflict of interest statement. None declared.

REFERENCES

1. Wilhelm,D., Palmer,S. and Koopman,P. (2007) Sex determination and gonadal development in mammals. *Physiol. Rev.*, **87**, 1–28.
2. Santos,A.C. and Lehmann,R. (2004) Germ cell specification and migration in Drosophila and beyond. *Curr. Biol.*, **14**, R578–R589.
3. McLaren,A. (2001) Mammalian germ cells: birth, sex, and immortality. *Cell Struct. Funct.*, **26**, 119–122.
4. Raz,E. (2004) Guidance of primordial germ cell migration. *Curr. Opin. Cell Biol.*, **16**, 169–173.
5. Saito,D., Morinaga,C., Aoki,Y., Nakamura,S., Mitani,H., Furutani-Seiki,M., Kondoh,H. and Tanaka,M. (2007) Proliferation of germ cells during gonadal sex differentiation in medaka: Insights from germ cell-depleted mutant zenzai. *Dev. Biol.*, **310**, 280–290.
6. Delvin,R.H. and Nagahama,Y. (2002) Sex determination and sex differentiation in fish. *Aquaculture*, **208**, 191–364.
7. Nakamura,S., Kobayashi,D., Aoki,Y., Yokoi,H., Ebe,Y., Wittbrodt,J. and Tanaka,M. (2006) Identification and lineage tracing of two populations of somatic gonadal precursors in medaka embryos. *Dev. Biol.*, **295**, 678–688.
8. Kurokawa,H., Saito,D., Nakamura,S., Katoh-Fukui,Y., Ohta,K., Baba,T., Morohashi,K. and Tanaka,M. (2007) Germ cells are essential for sexual dimorphism in the medaka gonad. *Proc. Natl Acad. Sci. USA*, **104**, 16958–16963.
9. Morrish,B.C. and Sinclair,A.H. (2002) Vertebrate sex determination: many means to an end. *Reproduction*, **124**, 447–457.
10. Nanda,I., Kondo,M., Hornung,U., Asakawa,S., Winkler,C., Shimizu,A., Shan,Z., Haaf,T., Shimizu,N., Shima,A. *et al*. (2002) A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. *Proc. Natl Acad. Sci. USA*, **99**, 11778–11783.
11. Matsuda,M., Nagahama,Y., Shinomiya,A., Sato,T., Matsuda,C., Kobayashi,T., Morrey,C.E., Shibata,N., Asakawa,S., Shimizu,N. *et al*. (2002) DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*, **417**, 559–563.
12. Brennan,J. and Capel,B. (2004) One tissue, two fates: molecular genetic events that underlie testis versus ovary development. *Nat. Rev. Genet.*, **5**, 509–521.
13. Kobayashi,T., Matsuda,M., Kajiura-Kobayashi,H., Suzuki,A., Saito,N., Nakamoto,M., Shibata,N. and Nagahama,Y. (2004) Two DM domain genes, DMY and DMRT1, involved in testicular differentiation and development in the medaka, Oryzias latipes. *Dev. Dyn.*, **231**, 518–526.
14. Hornung,U., Herpin,A. and Schartl,M. (2007) Expression of the male determining gene dmrt1bY and its autosomal coorthologue dmrt1a in medaka. *Sex Dev.*, **1**, 197–206.
15. Winkler,C., Hornung,U., Kondo,M., Neuner,C., Duschl,J., Shima,A. and Schartl,M. (2004) Developmentally regulated and non-sex-specific expression of autosomal dmrt genes in embryos of the Medaka fish (Oryzias latipes). *Mech. Dev.*, **121**, 997–1005.
16. Moore,M.J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science*, **309**, 1514–1518.
17. Keene,J.D. (2007) RNA regulons: coordination of post-transcriptional events. *Nat. Rev. Genet.*, **8**, 533–543.
18. Irish,V., Lehmann,R. and Akam,M. (1989) The Drosophila posterior-group gene nanos functions by repressing hunchback activity. *Nature*, **338**, 646–648.
19. Struhl,G., Johnston,P. and Lawrence,P.A. (1992) Control of Drosophila body pattern by the hunchback morphogen gradient. *Cell*, **69**, 237–249.
20. Deshpande,G., Calhoun,G. and Schedl,P. (2004) Overlapping mechanisms function to establish transcriptional quiescence in the embryonic Drosophila germline. *Development*, **131**, 1247–1257.
21. Dahanukar,A. and Wharton,R.P. (1996) The Nanos gradient in Drosophila embryos is generated by translational regulation. *Genes Dev.*, **10**, 2610–2620.
22. Duchow,H.K., Brechbiel,J.L., Chatterjee,S. and Gavis,E.R. (2005) The nanos translational control element represses translation in somatic cells by a Bearded box-like motif. *Dev. Biol.*, **282**, 207–217.
23. Koprunner,M., Thisse,C., Thisse,B. and Raz,E. (2001) A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev.*, **15**, 2877–2885.
24. Mishima,Y., Giraldez,A.J., Takeda,Y., Fujiwara,T., Sakamoto,H., Schier,A.F. and Inoue,K. (2006) Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. *Curr. Biol.*, **16**, 2135–2142.
25. Iwamatsu,T. (2004) Stages of normal development in the medaka Oryzias latipes. *Mech. Dev.*, **121**, 605–618.
26. Kluver,N., Kondo,M., Herpin,A., Mitani,H. and Schartl,M. (2005) Divergent expression patterns of Sox9 duplicates in teleosts indicate a lineage specific subfunctionalization. *Dev. Genes Evol.*, **215**, 297–305.
27. Thisse,C., Thisse,B., Schilling,T.F. and Postlethwait,J.H. (1993) Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development*, **119**, 1203–1215.

28. Lemaire,P., Garrett,N. and Gurdon,J.B. (1995) Expression cloning of Siamois, a Xenopus homeobox gene expressed in dorsal-vegetal cells of blastulae and able to induce a complete secondary axis. *Cell*, **81**, 85–94.
29. Herpin,A., Fischer,P., Liedtke,D., Kluever,N., Neuner,C., Raz,E. and Schartl,M. (2008) Sequential SDF1a and b-induced mobility guides Medaka PGC migration. *Dev. Biol.*, **320**, 319–327.
30. Nakamura,S., Aoki,Y., Saito,D., Kuroki,Y., Fujiyama,A., Naruse,K. and Tanaka,M. (2008) Sox9b/sox9a2-EGFP transgenic medaka reveals the morphological reorganization of the gonads and a common precursor of both the female and male supporting cells. *Mol. Reprod. Dev.*, **75**, 472–476.
31. Koster,R., Stick,R., Loosli,F. and Wittbrodt,J. (1997) Medaka spalt acts as a target gene of hedgehog signaling. *Development*, **124**, 3147–3156.
32. Hong,Y., Liu,T., Zhao,H., Xu,H., Wang,W., Liu,R., Chen,T., Deng,J. and Gui,J. (2004) Establishment of a normal medakafish spermatogonial cell line capable of sperm production in vitro. *Proc. Natl Acad. Sci. USA*, **101**, 8011–8016.
33. Komura,J., Mitani,H. and Shima,A. (1988) Fish cell culture: Establishment of two fibroblast-like cell lines (OL-17 and OL-32) from fins of the medaka, Oryzias latipes. *In Vitro Cell Dev. Biol.*, **24**, 294–298.
34. Hong,Y., Winkler,C. and Schartl,M. (1996) Pluripotency and differentiation of embryonic stem cell lines from the medakafish (Oryzias latipes). *Mech. Dev.*, **60**, 33–44.
35. Herpin,A., Rohr,S., Riedel,D., Kluever,N., Raz,E. and Schartl,M. (2007) Specification of primordial germ cells in medaka (Oryzias latipes). *BMC Dev. Biol.*, **7**, 3.
36. Kurokawa,H., Aoki,Y., Nakamura,S., Ebe,Y., Kobayashi,D. and Tanaka,M. (2006) Time-lapse analysis reveals different modes of primordial germ cell migration in the medaka Oryzias latipes. *Dev. Growth Differ.*, **48**, 209–221.
37. Saito,T., Fujimoto,T., Maegawa,S., Inoue,K., Tanaka,M., Arai,K. and Yamaha,E. (2006) Visualization of primordial germ cells in vivo using GFP-nos1 3'UTR mRNA. *Int. J. Dev. Biol.*, **50**, 691–699.
38. Tanaka,M., Kinoshita,M., Kobayashi,D. and Nagahama,Y. (2001) Establishment of medaka (Oryzias latipes) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate. *Proc. Natl Acad. Sci. USA*, **98**, 2544–2549.
39. Lei,N., Hornbaker,K.I., Rice,D.A., Karpova,T., Agbor,V.A. and Heckert,L.L. (2007) Sex-specific differences in mouse DMRT1 expression are both cell type- and stage-dependent during gonad development. *Biol. Reprod.*, **77**, 466–475.
40. De Grandi,A., Calvari,V., Bertini,V., Bulfone,A., Peverali,G., Camerino,G., Borsani,G. and Guioli,S. (2000) The expression pattern of a mouse doublesex-related gene is consistent with a role in gonadal differentiation. *Mech. Dev.*, **90**, 323–326.
41. Smith,C.A., McClive,P.J., Western,P.S., Reed,K.J. and Sinclair,A.H. (1999) Conservation of a sex-determining gene. *Nature*, **402**, 601–602.
42. Moniot,B., Berta,P., Scherer,G., Sudbeck,P. and Poulat,F. (2000) Male specific expression suggests role of DMRT1 in human sex determination. *Mech. Dev.*, **91**, 323–325.
43. Kettlewell,J.R., Raymond,C.S. and Zarkower,D. (2000) Temperature-dependent expression of turtle Dmrt1 prior to sexual differentiation. *Genesis*, **26**, 174–178.
44. Marchand,O., Govoroun,M., D'Cotta,H., McMeel,O., Lareyre,J., Bernot,A., Laudet,V. and Guiguen,Y. (2000) DMRT1 expression during gonadal differentiation and spermatogenesis in the rainbow trout, Oncorhynchus mykiss. *Biochem. Biophys. Acta.*, **1493**, 180–187.
45. Zarkower,D. (2001) Establishing sexual dimorphism: conservation amidst diversity? *Nat. Rev. Genet.*, **2**, 175–185.
46. Aoki,Y., Nagao,I., Saito,D., Ebe,Y., Kinjo,M. and Tanaka,M. (2008) Temporal and spatial localization of three germline-specific proteins in medaka. *Dev. Dyn.*, **237**, 800–807.
47. Schier,A.F. (2007) The maternal-zygotic transition: death and birth of RNAs. *Science*, **316**, 406–407.
48. Mazumder,B., Seshadri,V. and Fox,P.L. (2003) Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem. Sci.*, **28**, 91–98.
49. Kloc,M., Zearfoss,N.R. and Etkin,L.D. (2002) Mechanisms of subcellular mRNA localization. *Cell*, **108**, 533–544.
50. Gao,F.B., Carson,C.C., Levine,T. and Keene,J.D. (1994) Selection of a subset of mRNAs from combinatorial 3' untranslated region libraries using neuronal RNA-binding protein Hel-N1. *Proc. Natl Acad. Sci. USA*, **91**, 11207–11211.
51. Niehrs,C. and Pollet,N. (1999) Synexpression groups in eukaryotes. *Nature*, **402**, 483–487.
52. Miller,S.W., Hayward,D.C., Bunch,T.A., Miller,D.J., Ball,E.E., Bardwell,V.J., Zarkower,D. and Brower,D.L. (2003) A DM domain protein from a coral, Acropora millepora, homologous to proteins important for sex determination. *Evol. Dev.*, **5**, 251–258.
53. Scali,C., Catteruccia,F., Li,Q. and Crisanti,A. (2005) Identification of sex-specific transcripts of the Anopheles gambiae doublesex gene. *J. Exp. Biol.*, **208**, 3701–3709.
54. Nagoshi,R.N. and Baker,B.S. (1990) Regulation of sex-specific RNA splicing at the Drosophila doublesex gene: cis-acting mutations in exon sequences alter sex-specific RNA splicing patterns. *Genes Dev.*, **4**, 89–97.
55. Burtis,K.C. and Baker,B.S. (1989) Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell*, **56**, 997–1010.
56. Sreenivasulu,K., Ganesh,S. and Raman,R. (2002) Evolutionarily conserved, DMRT1, encodes alternatively spliced transcripts and shows dimorphic expression during gonadal differentiation in the lizard, Calotes versicolor. *Mech. Dev.*, **119(Suppl. 1)**, S55–S64.
57. Guo,Y., Cheng,H., Huang,X., Gao,S., Yu,H. and Zhou,R. (2005) Gene structure, multiple alternative splicing, and expression in gonads of zebrafish Dmrt1. *Biochem. Biophys. Res. Commun.*, **330**, 950–957.
58. Huang,X., Guo,Y., Shui,Y., Gao,S., Yu,H., Cheng,H. and Zhou,R. (2005) Multiple alternative splicing and differential expression of dmrt1 during gonad transformation of the rice field eel. *Biol. Reprod.*, **73**, 1017–1024.
59. Zhao,Y., Lu,H., Yu,H., Cheng,H. and Zhou,R. (2007) Multiple alternative splicing in gonads of chicken DMRT1. *Dev. Genes Evol.*, **217**, 119–126.
60. Lu,H., Huang,X., Zhang,L., Guo,Y., Cheng,H. and Zhou,R. (2007) Multiple alternative splicing of mouse Dmrt1 during gonadal differentiation. *Biochem. Biophys. Res. Commun.*, **352**, 630–634.
61. Cheng,H.H., Ying,M., Tian,Y.H., Guo,Y., McElreavey,K. and Zhou,R.J. (2006) Transcriptional diversity of DMRT1 (dsx- and mab3-related transcription factor 1) in human testis. *Cell Res.*, **16**, 389–393.

Divergent Expression Regulation of Gonad Development Genes in Medaka Shows Incomplete Conservation of the Downstream Regulatory Network of Vertebrate Sex Determination

Amaury Herpin,*[1] Mateus C. Adolfi,[1] Barbara Nicol,[2] Maria Hinzmann,[1] Cornelia Schmidt,[1] Johanna Klughammer,[1] Mareen Engel,[1] Minoru Tanaka,[3] Yann Guiguen,[2] and Manfred Schartl[1]

[1]University of Wuerzburg, Physiological Chemistry, Biocenter, Am Hubland, Wuerzburg, Germany

[2]INRA, UR1037, LPGP, Fish Physiology and Genomics, Rennes, France

[3]Laboratory of Molecular Genetics for Reproduction, National Institute for Basic Biology, Okazaki, Japan

***Corresponding author:** E-mail: amaury.herpin@biozentrum.uni-wuerzburg.de.

Associate editor: Takashi Gojobori

Abstract

Genetic control of male or female gonad development displays between different groups of organisms a remarkable diversity of "master sex-determining genes" at the top of the genetic hierarchies, whereas downstream components surprisingly appear to be evolutionarily more conserved. Without much further studies, conservation of sequence has been equalized to conservation of function. We have used the medaka fish to investigate the generality of this paradigm. In medaka, the master male sex-determining gene is *dmrt1bY*, a highly conserved downstream regulator of sex determination in vertebrates. To understand its function in orchestrating the complex gene regulatory network, we have identified targets genes and regulated pathways of Dmrt1bY. Monitoring gene expression and interactions by transgenic fluorescent reporter fish lines, in vivo tissue-chromatin immunoprecipitation and in vitro gene regulation assays revealed concordance but also major discrepancies between mammals and medaka, notably amongst spatial, temporal expression patterns and regulations of the canonical Hedgehog and R-spondin/Wnt/Follistatin signaling pathways. Examination of Foxl2 protein distribution in the medaka ovary defined a new subpopulation of theca cells, where ovarian-type *aromatase* transcriptional regulation appears to be independent of Foxl2. In summary, these data show that the regulation of the downstream regulatory network of sex determination is less conserved than previously thought.

Key words: **gene regulatory network evolution, divergent expression regulation, gonadal development, adaptive evolution.**

Article

Introduction

Sex determination, the decision whether the bipotential gonad anlage will become a testis or an ovary, is a complex and tightly controlled developmental process. The fate determination and cell differentiation programs are regulated and tuned by cascades or networks of genes. Comparative studies on sex determination cascades of different organisms revealed a remarkable diversity of "master sex-determining genes" at the top of the genetic hierarchies, whereas downstream components surprisingly appeared to be evolutionarily more conserved and tend to converge upon the regulation of common effectors. Hence, a comparative view on genetic sex determination mechanisms led to the paradigm that "masters change, slaves remain" (Graham et al. 2003). A well-known example illustrating this paradigm is the *SRY* gene, the master sex-determining gene of mammals, which has not been detected outside of the therian mammals. However, its subordinated genes (*SOX9*, *WT1*, *DMRT1*, *AMH*, *SF1*, *FOXL2*) or signaling pathways (TGF-beta, WNT4/beta-catenin, Hedgehog) have homologs in a much broader spectrum of species, including nonvertebrates, where they apparently are also involved in sex determination. These observations led to the emergence of the subversive stereotype that master sex-determining genes individually spark and orchestrate the irreversible action of uniform and integrated gender-specific pathways. Conservation at the bottom and diversity at the top could be convincingly explained by an evolutionary scenario in which these hierarchies evolve from common core downstream components that acquire new upstream regulators (Wilkins 2007). Although the global rule of sex determination evolution is intuitively appealing and well accepted, only the variety at the top is well supported by comparative experimental data (Graham et al. 2003; Haag and Doty 2005; Herpin and Schartl 2008). The downstream conservation is less studied and relies only on a few gene expression studies. We have used the medaka as a versatile model system to study gene regulatory interactions and their evolutionary conservation (Wittbrodt et al. 2002; Herpin and Schartl 2009).

In the medaka fish, which has XY-XX sex determination, *dmrt1bY*, the duplicated copy on the Y-chromosome of *dmrt1a*, was shown to be the dominant master regulator of male development (Matsuda et al. 2002; Nanda et al. 2002), similar to *Sry* in mammals. Interestingly, *dmrt1*, the ancestor of *dmrt1bY*, is one of the downstream effectors of SRY in the

mammalian male pathway. The duplicated copy of *dmrt1* on the Y-chromosome has acquired an upstream position in the sex-determining cascade. Remarkably this evolutionary novelty, requiring a rewiring of the regulatory network, was brought about by co-optation of "ready-to-use" pre-existing *cis*-regulatory elements contributed by transposable elements (Herpin et al. 2010). With respect to their biochemical functions, both medaka *dmrt1* paralogs act as transcriptional regulators (Herpin et al. 2010). Dmrt1bY was shown to be responsible for male-specific primordial germ cell mitotic arrest in the developing gonad at the sex-determination stage (Herpin et al. 2007). In contrast, the autosomal *dmrt1a* medaka gene is essential for testis maintenance (Masuyama et al. 2012).

The medaka gonad is formed by the coordinated development of two different cell lineages: the germ cells and the somatic gonadal mesoderm surrounding the germ cells. Shortly before hatching, at the time of expression of *dmrt1bY* in the male gonad primordium, the germ cells in the female gonad actively proliferate and undergo meiosis, whereas this is not observed in male gonads (Kobayashi et al. 2004; Herpin et al. 2007). It is only 10 days later that the first somatic gonadal dimorphisms are apparent with the formation of the acinus (the seminiferous tubule precursor) and the follicles in gonads of male and female, respectively. Interestingly, ovarian cords within the germinal epithelia of medaka ovaries have been recently characterized (Nakamura et al. 2010). These cords composed of somatic *sox9b*-expressing cells and mitotic *nos2*-expressing oogonia continually give rise to germ cells and form a stem cell niche referred to as the germinal cradle (Nakamura et al. 2010). These cradles containing germline stem cells contribute to the production of fertile eggs during the life cycle of the adult ovary.

Like in other vertebrates studied so far, many components of the classical repertoire of mammalian sex-determining genes could be inventoried in medaka as well (Matsuda 2005; Siegfried 2010). To elucidate the gene regulatory network that controls specification and patterning of the gonads in the medaka fish, in this study we report the in vivo expression dynamics of classical mammalian markers within the forming and the adult gonads in medaka.

Because of their important roles in initiating male or female gonadal development in mammals, the respective implications of either Dmrt1 or Foxl2 transcription factors were examined. Further on, two of the major signaling pathways central for early gonadal induction and maintenance in mammals, namely the canonical Hedgehog and Wnt4/β-catenin signaling pathways (Wilhelm et al. 2007; Liu et al. 2010; Franco and Yao 2012), were investigated.

The Hedgehog (HH) signaling pathway plays an essential role in a wide variety of developmental processes (Ingham and McMahon 2001). Three HH proteins have been identified in mammals; SONIC (SHH), INDIAN (IHH), and DESERT (DHH) (Hooper and Scott 2005). *Dhh* and *Ihh* are co-expressed in the adult ovary where they stimulate proliferation and steroidogenesis of theca cells (Spicer et al. 2009). *Dhh* is also required for maintenance of the male germ line and spermatogenesis in mice (Bitgood et al. 1996; Clark et al. 2000). In mammals all three HH ligands signal by binding to one of two homologous transmembrane receptors, PATCHED homolog 1 and 2 (PTCH1 and 2). HH signaling is modulated by HH-induced transcription of the HH antagonistic interacting protein (HHIP) that binds to HH ligands and prevents their interaction with PTCH receptors (Chuang and McMahon 1999).

R-spondin1 (*Rspo1*), a member of a small family of secreted growth factors, is a key female-determining factor. RSPO protein operates through the canonical WNT signaling pathway (Tomizuka et al. 2008) to activate the β-catenin pathway as well as via upregulation of *Follistatin* (*Fst*) through WNT4 (Yao et al. 2004). It is well established that mammalian R-SPO-1, WNT4, β-catenin, and FST are components of a single pathway that promotes ovarian development and suppresses the formation of testis cord (Chassot et al. 2008).

The winged helix/forkhead transcription factor FOXL2 is mainly expressed in the somatic cells of the female gonad (Crisponi et al. 2001). The major role of *Foxl2* during gonadal differentiation and maintenance has recently been shown via the mutual antagonistic relationship of *Foxl2* and *Dmrt1*. FOXL2 suppresses expression of *Dmrt1* and vice versa for maintaining female or male gonadal fate, respectively (Uhlenhaut et al. 2009; Matson et al. 2011). Additionally it has been reported that FOXL2 and WNT4 (Yao 2005; Ottolenghi et al. 2007; Garcia-Ortiz et al. 2009) cooperate in regulating FST expression during ovarian development. Interestingly in the ovary the expression profiles of *Foxl2* highly correlate with that of *Aromatase* (*Cyp19*), suggesting that *Foxl2* is involved in the regulation of estrogen synthesis via direct transcriptional upregulation of *Aromatase* (Pannetier et al. 2006). On the other hand, several other factors (e.g., testosterone, TGF-β1, TNF-α, and glucocorticoids) have been shown to direct the expression of the aromatase gene in Sertoli, Leydig, and germ cells of rat testis (see Bourguiba et al. 2003 for review).

Analyzing several fluorescent reporter lines established from a bacterial artificial chromosome recombination (BAC recombination) method resulting in optimal spatial resolution and high reliability of gene expression (Giraldo and Montoliu 2001; Nakamura et al. 2008b; Suster et al. 2011), we find major discrepancies between mammals and medaka, notably amongst spatial and temporal expression patterns of the canonical signaling pathways. Using in vivo whole tissues chromatin immunoprecipitation and in vitro gene regulation assays, we can reveal possible interactions between these pathways that emphasize the importance of the cellular context on modulating these regulations and call into question a strict conservation of regulatory and functional interactions of sexual development genes in vertebrates.

Results

Patched-2 Expression and Hedgehog Pathway Regulation

To determine the role of the hedgehog signaling pathway in gonadal development of medaka, we recorded the temporal and spatial expression patterns of the key receptor, Ptch2.

Although expressed quite early on during somatogenesis (data not shown), *patched-2* was surprisingly neither detected throughout the early phase of gonadogenesis when the undifferentiated gonad anlage grows up to hatching stage (fig. 1*A*) nor further on up to the stage when the dimorphic gonad develops 10 days after hatching (fig. 1*B* and *C*). The first specific gonadal *ptch-2* expression was detected in the young ovary (fig. 1*D*–*H*). Here expression is restricted to the somatic cells that express *sox9b* and surround the germline stem cells of the germinal cradle within the ovarian cord (Nakamura et al. 2010) (fig. 1*G* and *H*). Interestingly, *r-spo-1* is also co-expressed with *sox9b* in these cells (figs. 1*I*–*K* and 2*A*–*D* compared with fig. 8*A*–*F*). In ovaries of fish of later reproductive phase, when the cradle number has declined, *ptch-2* expression could be only noticed in the interstitium (fig. 2*E*–*G*). Unexpectedly and in contrast to *patched* expression in mammals, *ptch-2* expression was only detected at background levels throughout stages of testis development (data not shown). This inconspicuous role for the Hedgehog signaling in gonad development and maintenance was also apparent in real-time polymerase chain reaction (PCR) analysis. Although similarly expressed in different tissues, both *patched* receptor transcripts (1 and 2) are only expressed at background levels in adult gonads of both sexes (fig. 3*A*).

In vivo Dmrt1bY Binding to hhip Promoter Region

Given the pivotal role of Dmrt1 transcription factors during medaka sex determination, *patched-2* and the antagonistic regulatory HH interacting protein (*hhip*) promoter regions were scanned for putative Dmrt1bY and Dmrt1a target sites. Although no Dmrt1-binding sites could be identified in the *patched-2* promoter region (10 kb upstream scanned), two sites were predicted with high fidelity in the *hhip* 5′ region (fig. 4*B*). To assess the in vivo relevance of the predicted Dmrt1 interaction, two stable transgenic lines expressing either the full-length Dmrt1bY protein (Dmrt1::GFP) or a truncated form lacking the DNA-binding domain (ΔDmrt1::GFP), both fused to GFP, were utilized (Herpin et al. 2010). These two lines were used for in vivo tissue chromatin immunoprecipitation (in vivo tissue ChIP) on testis tissue using GFP antibody for immunoprecipitation (fig. 4*A*). For the predicted two proximal Dmrt1-binding sites in the *hhip* promoter, a more than 2.3-fold enrichment after immunoprecipitation validated Dmrt1 binding (fig. 4*B*). Under the same conditions, no binding was detected in ovary (data not shown). This indicates that, in vivo, Dmrt1bY and/or Dmrt1a are potentially regulating the gonadal HH signaling through direct transcriptional regulation of the antagonist *hedgehog interacting protein*, *hhip*.

Dmrt1-Induced Hedgehog Pathway Regulation

Further functional characterization of Dmrt1-induced *hhip* transcriptional regulation was performed by overexpression of Dmrt1a or Dmrt1bY in spermatogonial (SG3) or fibroblast (OLF) cell lines of medaka. In both cell lines, *hhip* transcription was clearly induced (fig. 5*A*–*D*). Consistently, examination of the *hhip* expression pattern disclosed high and specific expression in testes (fig. 3*A*). Interestingly, at the receptor level (*patched 1* and *2*), although no direct interaction with Dmrt1bY could be demonstrated, transcriptional down regulation of *patched-2* was observed after Dmrt1bY overexpression in the spermatogonial cell line (fig. 5*E*–*H*). In line with the in vitro regulation data, *patched-1/2* expression in gonads of both sexes was not above background (fig. 3*A*). Due to the absence of the *dmrt1bY* gene in females, and since a background level of *dmrt1a* expression is detected in ovary (Hornung et al. 2007), the high expression of *hhip* and the suppression of *ptch-2* must be exclusively regulated by the autosomal *dmrt1a* ortholog in the ovary of medaka.

R-spondin 1 and *Follistatin* Expression and Wnt Pathway Regulation

Although ubiquitously expressed at early stages of development, neither medaka *r-spo-1* nor *fst* expressions could be detected in the presumptive gonadal mesoderm before the PGCs reach the undifferentiated gonadal primordium at stage 30 (fig. 7*A*–*C*; supplementary fig. S1, Supplementary Material online). Subsequently, in the dorsal region of the hindgut, a very furtive and time restricted pulse of *r-spo-1* and *fst* expression appears in the male gonadal primordium between stages 33 and 35 (fig. 7*D*–*F*). Interestingly, this very brief pulse of *r-spo-1*and *fst* expression occurs shortly before the rise of *dmrt1bY* expression at that stage (Kobayashi et al. 2004; Hornung et al. 2007). At hatching stage when sex determination occurs, no *r-spo-1*or *fst* expression could be detected in the gonadal primordia of males and females (fig. 7*G*–*I*). This lack of later gonadal expression is in line with similar findings in zebrafish (Zhang et al. 2011), turtle (Smith et al. 2008), chicken (Smith et al. 2008), and mice (Yao et al. 2004; Parma et al. 2006). Further on, while the ovary develops in juvenile females, *r-spo-1* expression is restricted to few somatic cells surrounding germline stem cells of the germinal cradle within the ovarian cord (fig. 8*A*–*F*). These somatic cells of the germinal cradle are the same that also co-express *sox9b* and *patched-2* (fig. 2*A*–*D* compared with fig. 8*A*–*F*). Similar to zebrafish (Zhang et al. 2011) and mouse (Smith et al. 2008), expression of medaka *r-spo-1* was also detected in granulosa cells around young oocytes (fig. 8*G*–*L* and table 1). Granulosa expression is then progressively lost while the oocytes are growing. In male gonads, only a low expression of *r-spo-1* is detected (fig. 3*B*–*D*). Medaka *fst* could not be detected in the ovary at this stage. In contrast to germ cell expression of *r-spo-1* in chicken (Smith et al. 2008), zebrafish (Zhang et al. 2011) and mouse (Smith et al. 2008) ovaries or zebrafish testis (Zhang et al. 2011) or *fst* in rat testis (Meinhardt et al. 1998), neither *r-spo-1* nor *fst* could be detected in gonadal germ cells or spermatogonia in medaka (fig. 8 and table 1). In older ovaries, *r-spo-1* and *fst* are also expressed in the interstitium as well as in the ovarian epithelium (fig. 8*M*, *N*, and *Q*–*S*). Of note, sparse clusters of *fst* expressing interstitial somatic cells were detected in testes (fig. 8*O* and *P*).

In vivo Dmrt1bY Binding to the r-spo-1 and dkk1 Promoter Regions and Regulation of the r-spo-1 Pathway

We next analyzed the capacity of Dmrt1 to transcriptionally regulate positive (*r-spo-1* and *fst*) and negative (*dkk1*)

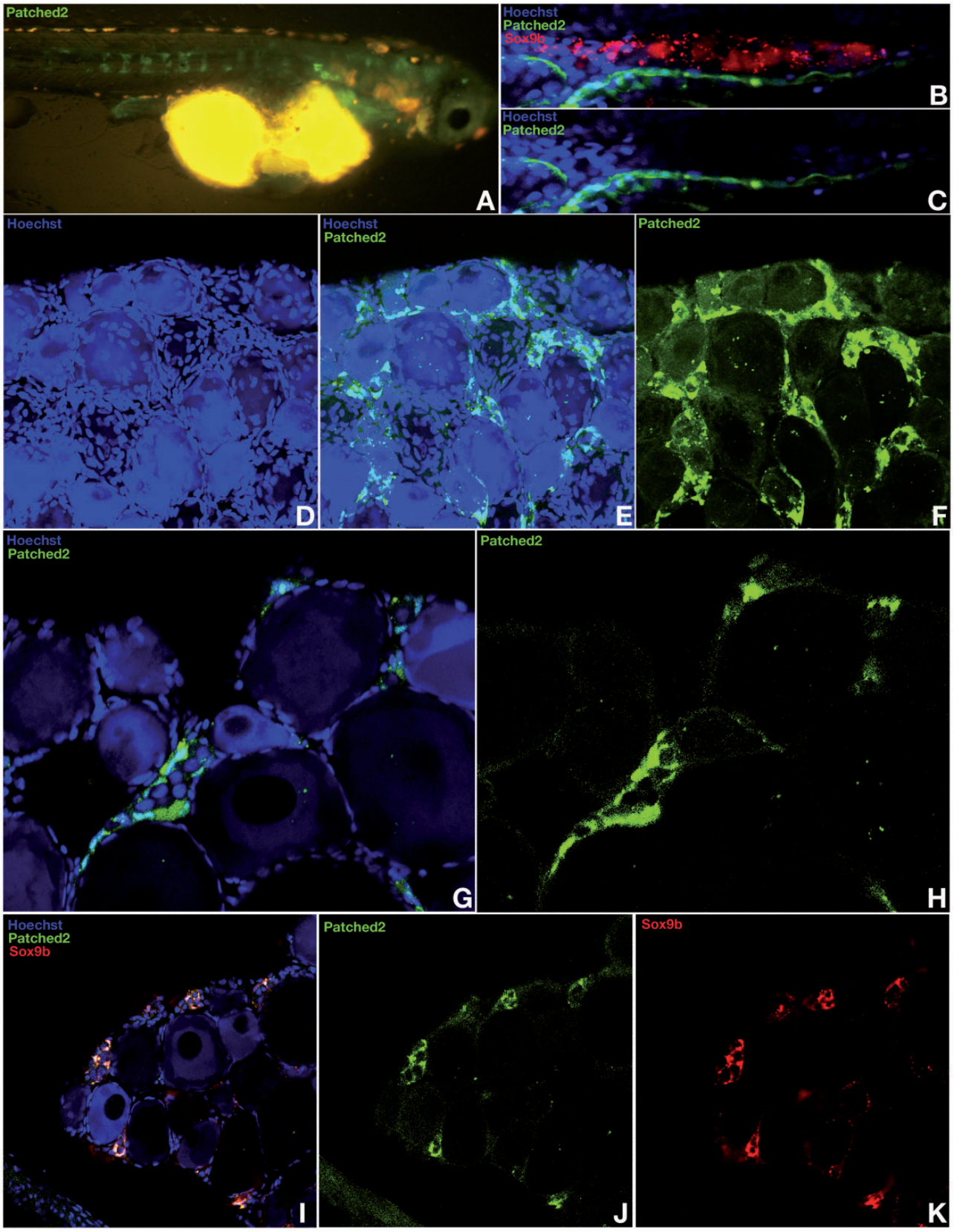


Fig. 1. Spatial and temporal expression patterns of *patched-2* during medaka early gonad formation and in adult gonads. A *patched-2* BAC reporter transgenic medaka line expressing GFP was established to follow *patched-2* expression dynamics in vivo during gonad formation. Although expressed early on during somatogenesis (*A*), *patched-2* expression was never detected during the early phase of gonadogenesis at hatching stage (*B* and *C*). In the young ovary, *patched-2* expression is restricted to somatic cells of the ovarian cord (*D*–*H*) where it is co-expressed together with *sox9b* (*I*–*K*).

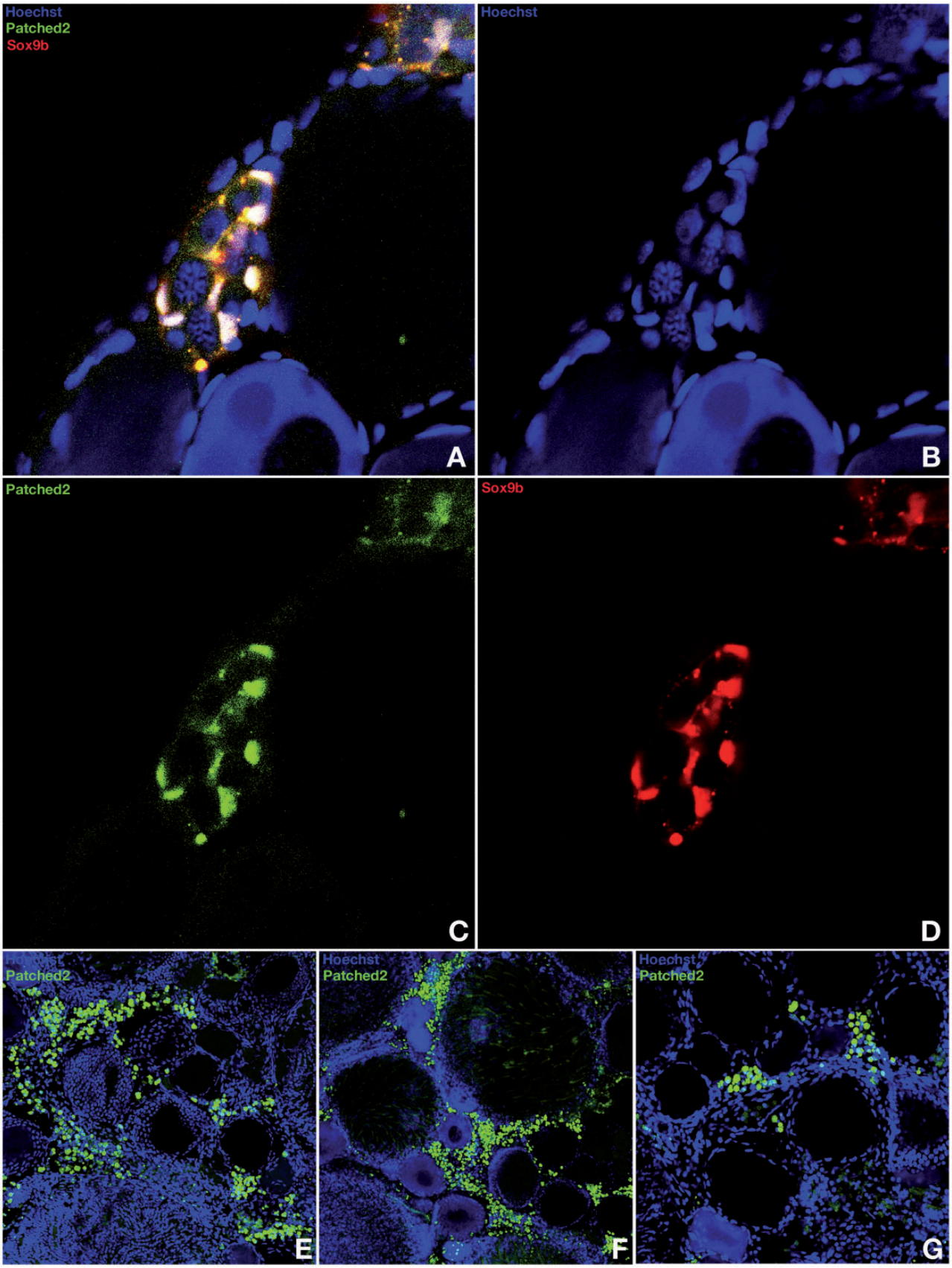


Fig. 2. Specific *patched-2* expression in the germinal cradle of the ovary. *Patched-2* expression is specifically restricted to the somatic cells that express *sox9b* and surround the germ line stem cells of the germinal cradle within the ovarian cord but is absent in the germ line stem cells (*A*–*D*). In ovaries of fish of later reproductive phase, *patched-2* expression is only apparent in the interstitium (*E*–*G*).

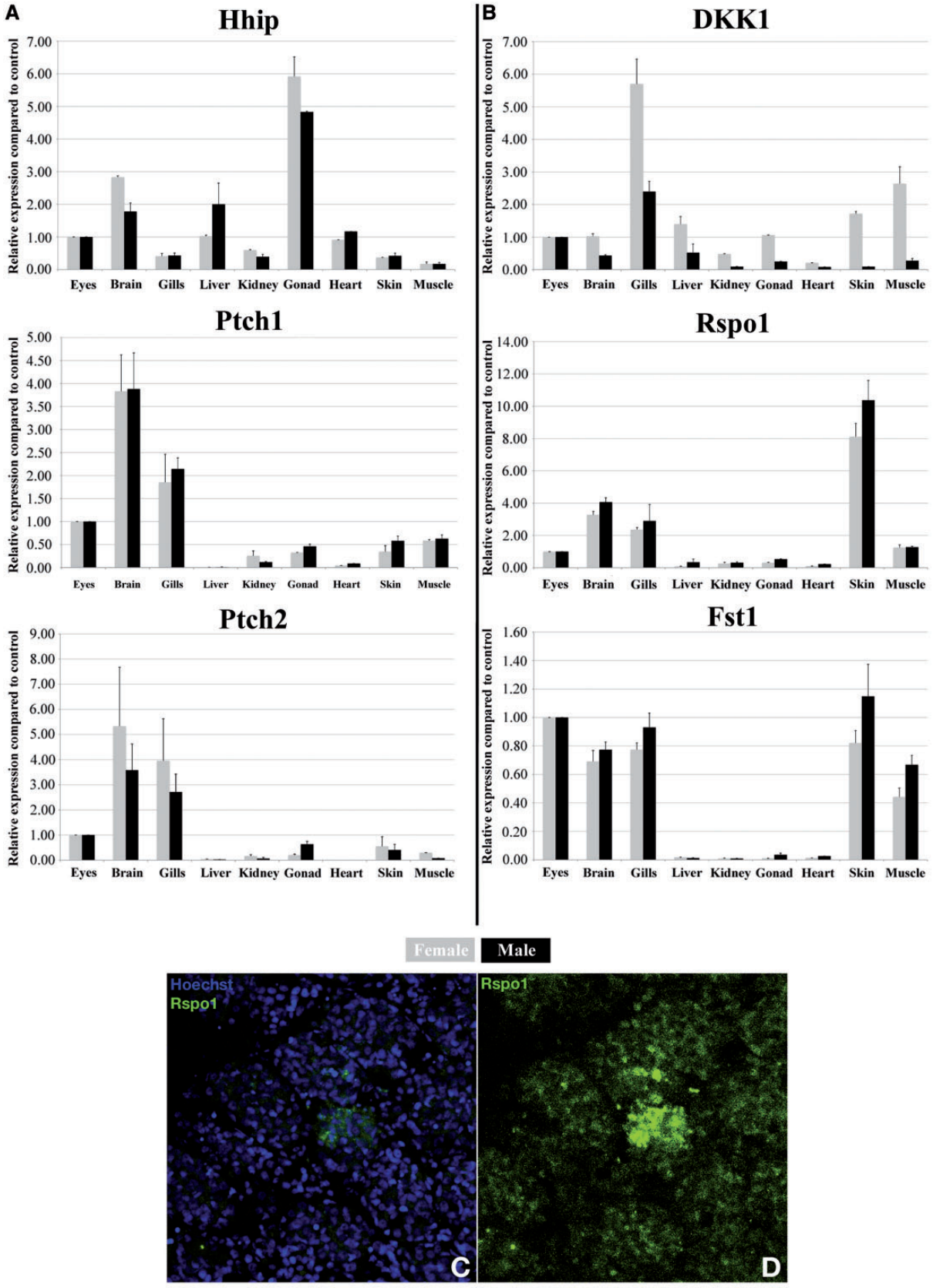


FIG. 3. Real-time PCR expression patterns of Hedgehog and Wnt pathway components and in vivo reporter expression of *R-spondin1* in medaka adult testis. Expression patterns of different components of the Hedgehog (*A*) and Wnt (*B*) pathways in organs of adult male and female medaka determined from pooled (3–4 animals) total RNA extracts. In adult testes, background levels of *R-spondin1* expression are detected either by real-time PCR (*B*) or BAC reporter fluorescence (*C* and *D*) methods.

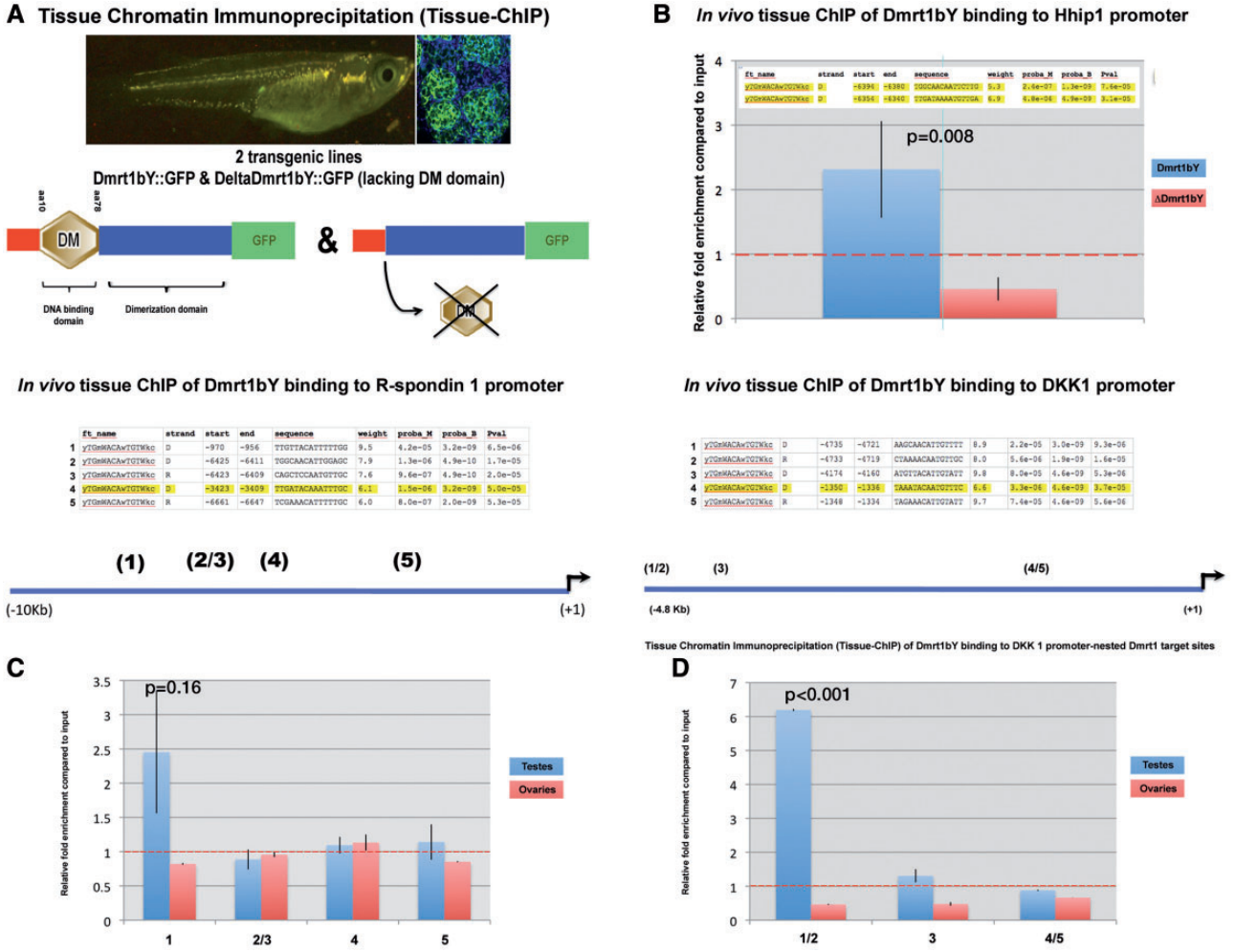


FIG. 4. In vivo tissue chromatin immunoprecipitation (in vivo tissue-ChIP) analysis of Dmrt1bY targets. Chromatin immunoprecipitation using both Dmrt1bY::GFP and deltaDmrt1bY transgenic lines respectively expressing either Dmrt1bY or as control a truncated Dmrt1bY (delta DM form lacking the DNA-binding domain) fused to GFP revealed in vivo specific Dmrt1bY protein affinities to target sites nested within *hedgehog interacting protein 1* (*hhip1*), *r-spondin-1*, and *dkk-1* respective promoter regions. (*A*) Transgenic lines established for in vivo tissue-ChIP. (*B*–*D*) Specific enrichment of *hhip-1* (*B*), *r-spondin-1* (*C*), and *dkk-1* (*D*) promoter-nested Dmrt1 binding sites subsequent to Dmrt1bY immunoprecipitation.

effectors of the Wnt/Rspo1/Fst pathway. Hence, the ability of Dmrt1 to directly bind in vivo to the *fst*, *r-spo-1*, and *dkk1* promoters was investigated. Out of several putative Dmrt1-binding sites within the *r-spo-1* and *dkk1* promoter regions, in vivo ChIP revealed direct Dmrt1bY interaction with both promoters (fig. 4*C* and *D*). Of note, robust binding of Dmrt1bY to *dkk1* promoter region was seen in testes, but similar interactions did not occur in ovaries (fig. 4*D*). Although not highly significant, an analogous trend is observed for Dmrt1bY binding to *r-spo-1* promoter (fig. 4*C*). Furthermore cell transfection experiments overexpressing either Dmrt1a or Dmrt1bY showed a Dmrt1bY-specific slight transcriptional upregulation of *dkk1* (fig. 6*A*–*D*). Interestingly, although *fst* does not seem to be under Dmrt1bY regulation (fig. 6*J* and *L*), Dmrt1a does upregulate *fst* transcription (fig. 6*I* and *K*). Astonishingly and in contrast to our in vivo expression data, Dmrt1-induced transcriptional upregulation was observed for *r-spondin-1* (fig. 6*E*–*H*). Real-time PCR quantification nevertheless revealed *r-spo-1*and *fst* to be only expressed at background levels in gonads (fig. 3*B*), surprisingly also including adult testis (fig. 3*C* and *D*).

Foxl2 Expression in the Adult Gonad

To investigate the role of Foxl2 in ovarian differentiation, we analyzed Foxl2 protein distribution in the medaka ovary (figs. 9 and 10). During the transition process of germ line stem cells to oocytes within the germinal cradle, medaka Foxl2 expression starts within the germ line stem cells and continues during meiosis until early oogenesis (fig. 9*A*–*L*). On the contrary, no Foxl2 protein could be detected in the interwoven threadlike ovarian cords of *sox9b*-expressing cells where the supporting follicular cells reside (fig. 9*A*–*L*). During the following steps of oogenesis, the accompanying cells of the supporting layer progressively loose *sox9b* expression while Foxl2 expression rises (fig. 9*A*–*L*). Consistent with mRNA localization (Nakamoto et al. 2006), Foxl2 protein in the medaka ovary was localized within the follicular cells of the previtellogenic and vitellogenic follicles and then gradually lost while maturation proceeds (fig. 10*A*).

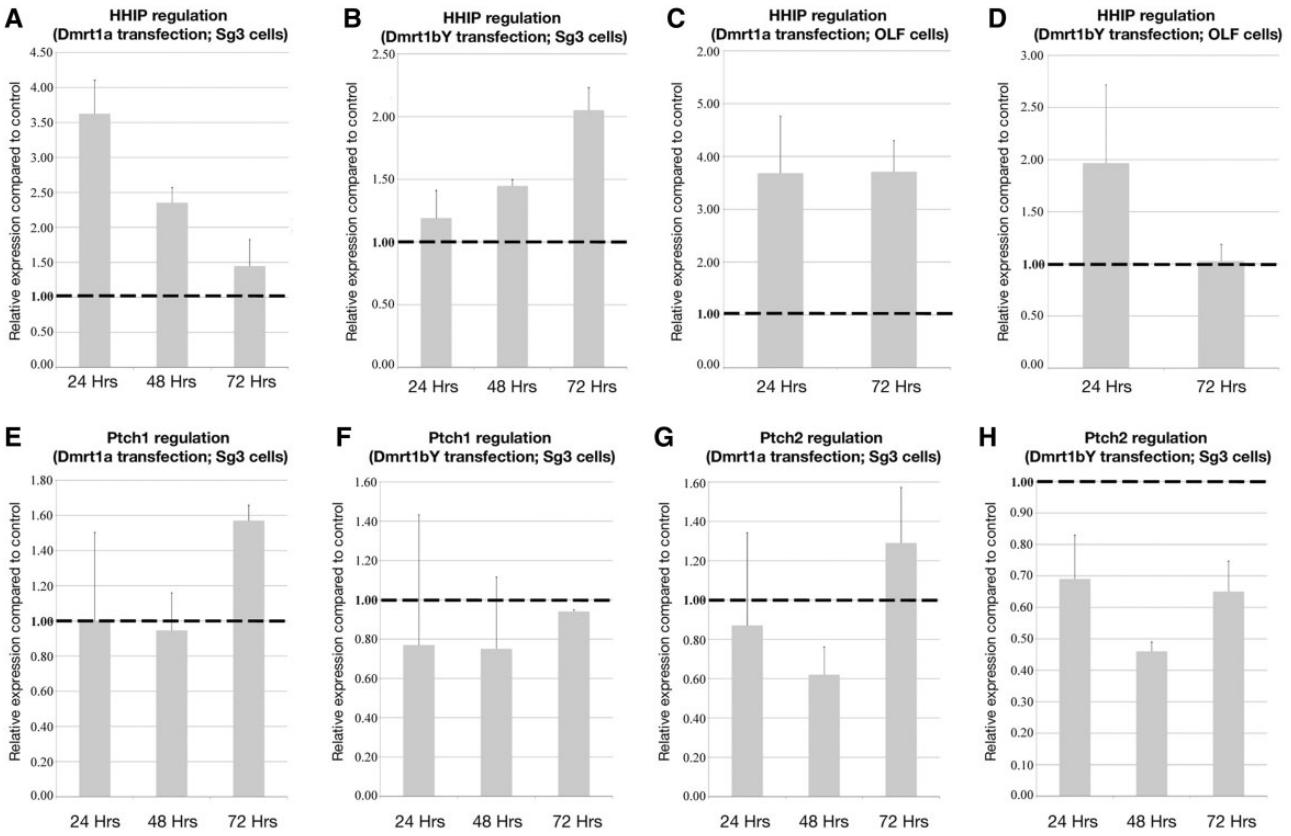


FIG. 5. Hedgehog pathway transcriptional regulation after Dmrt1a/Dmrt1bY overexpression in different cell lines. *hhip* (*A*–*D*), *patched-1* (*E* and *F*), and *patched-2* (*G* and *H*) expression were monitored in different cell lines (SG3 and OLF) after Dmrt1a or Dmrt1bY overexpression. *Hhip* transcription is mainly upregulated by both Dmrt1s (*A*–*D*). *Patched-2* transcription is only clearly downregulated by Dmrt1bY (*G* and *H*) and *patched-1* expression levels remain unaffected (*E* and *F*).

Particularly, Foxl2 is present in the nuclei of all granulosa cells (fig. 10*B*–*D*). Interestingly, while Foxl2 has been reported to be a strong inducer of the steroidogenic activity of granulosa cells via upregulation of the aromatase gene (the ovarian-type Cyp19a1) (Hudson et al. 2005; Wang et al. 2007; Guiguen et al. 2010), unexpectedly, and in contrast to mammals, a minority of theca cells do also express Foxl2 in medaka (fig. 10*E*–*G*). Our results reveal two subpopulations of *cyp19a1*-positive theca cells, which are either Foxl2 positive or do not express the transcription factor (fig. 10*H*–*K*).

Discussion

Comparative studies on sex determination cascades of different organisms revealed that the genetic control of male or female gonad development displays between different groups of organisms a remarkable diversity of "master sex-determining genes" at the top of the genetic hierarchies, whereas downstream components surprisingly were found more widespread and are evolutionarily more conserved. Without much further studies, these observations led to the reasoning that conservation of sequence equalizes to conservation of function. While vertebrates have at least a common set of transcriptional regulators, including DMRT1 and FOXL2, as well as some signaling molecules and pathways such as the Hedgehog and R-spo-1/Wnt4 pathways, their molecular interplay and epistatic relationships are nevertheless far from being understood. The purpose of our work was to examine this molecular interplay in fish.

Absence of *R-spo-1*, *fst*, and *ptch-2* in Medaka Germ Cells

In the mouse it was proposed that activation of the R-spo-1/Wnt/Fst signaling pathway in both somatic and germ cells, besides triggering meiosis in fetal germ cells, is required for ovarian differentiation and maintenance of ovarian cell identity (Chassot et al. 2008). Indeed inherent to their inductive epigenetic mode of specification, the germline sex is likely to be determined early on by *Sry* acting in the somatic cells (McLaren and Southee 1997; Sekido and Lovell-Badge 2008; Bowles et al. 2010) in a noncell-autonomous manner (see De Felici 2009 for review). *Sry* controls whether bipotential precursor cells differentiate into testicular Sertoli cells or ovarian granulosa cells (Koopman et al. 1991). This pivotal decision in a single gonadal cell type ultimately controls sexual differentiation throughout the body. Sex determination can be viewed as a battle for primacy in the fetal gonad between a male regulatory gene network in which *Sry* activates *Sox9* and a female network involving WNT/β-catenin signaling (Uhlenhaut et al. 2009; Herpin and Schartl 2011; Matson et al. 2011).

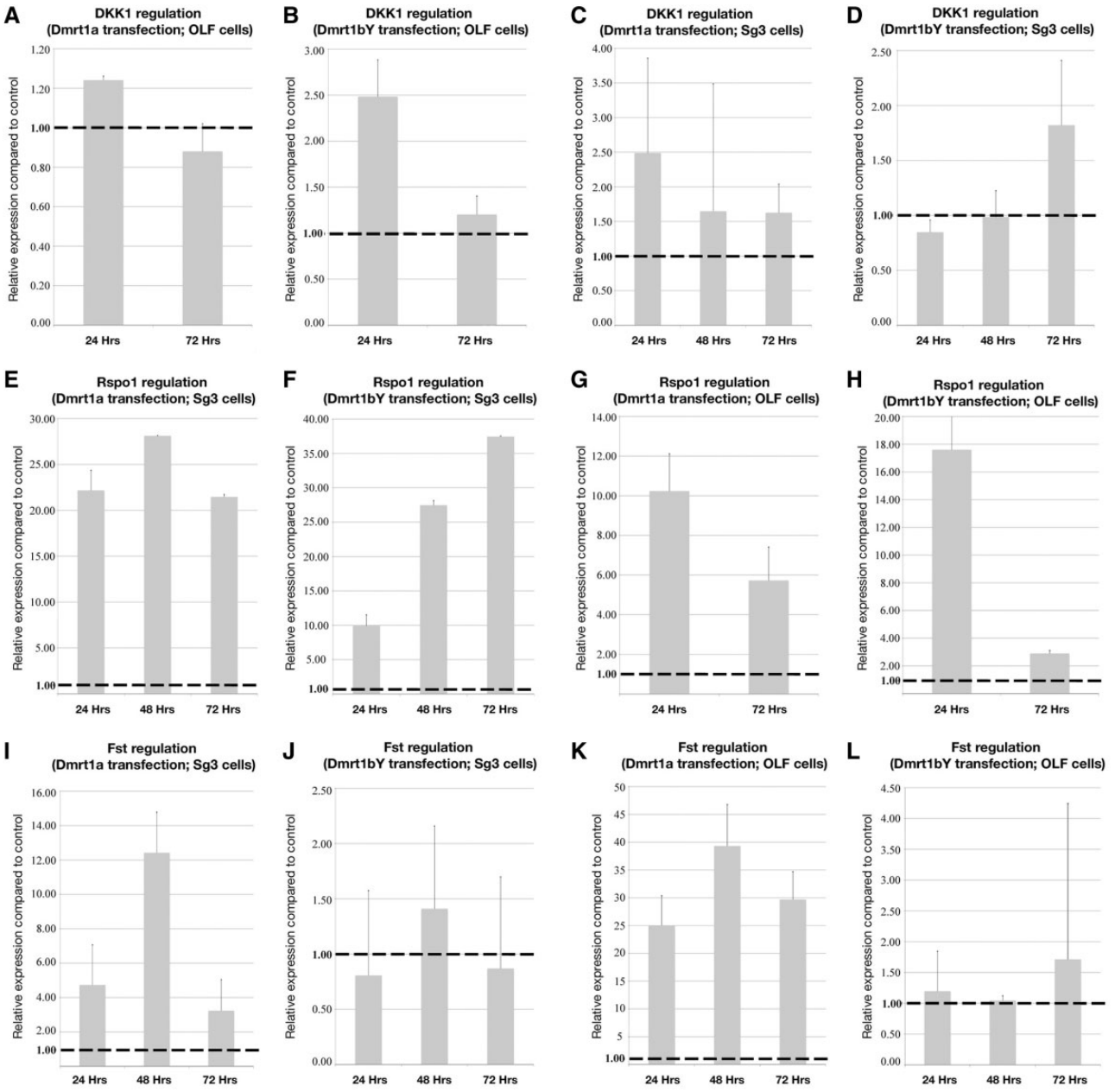


Fig. 6. Canonical *Wnt* pathway (*dkk-1*, *r-spondin-1*, and *fst*) transcriptional regulation after Dmrt1a/Dmrt1bY overexpression in different cell lines. *dkk-1* (*A*–*D*), *r-spondin-1* (*E*–*H*), and *fst* (*I*–*L*) expression was monitored in different cell lines (SG3 and OLF) following Dmrt1a or Dmrt1bY overexpression. Transcription of dkk-1 and r-spondin-1 is upregulated by both Dmrt1s (*A*–*H*). Differential regulation of *fst* is observed. While no regulation was observed after Dmrt1bY overexpression, *fst* transcription is clearly upregulated in both cell lines overexpressing Dmrt1a (*I*–*L*).

In contrast, the sexual plasticity of medaka germ cells seems to be retained much longer than in mammals as illustrated by XX/XY transplantation chimeras. Although XY somatic cells differentiate into male cells according to their sex chromosome composition, in this environment XX germ cells differentiate into male cells regardless of their sex chromosome composition (Shinomiya et al. 2002). Hence, while unlike their mammalian counterparts none of the different medaka marker genes analyzed (*r-spo-1*, *fst* or *ptch-2*) were detected in germ cells at any time of gonadal development, this major inconsistency between mammals and fish certainly reflects intrinsically divergent modes of germ cell commitment and interaction between germ and somatic cells possibly accounting for a higher sexual plasticity of germ cells in fish.

Expression of the *Rspo1*/*Fst* and Hedgehog Pathways

Role of the Rspo1/Wnt/Fst Pathway during Gonad Development and Maintenance

In line with observations made in zebrafish (Zhang et al. 2011), turtle (Smith et al. 2008), chicken (Smith et al. 2008), and mice (Yao et al. 2004; Parma et al. 2006), the absence in medaka of *r-spo-1* and *fst* dimorphic expression during sex determination stages does not support a role during gonad induction. Of particular interest and unlike in mammals, the lack of medaka *ptch-2* expression during the same

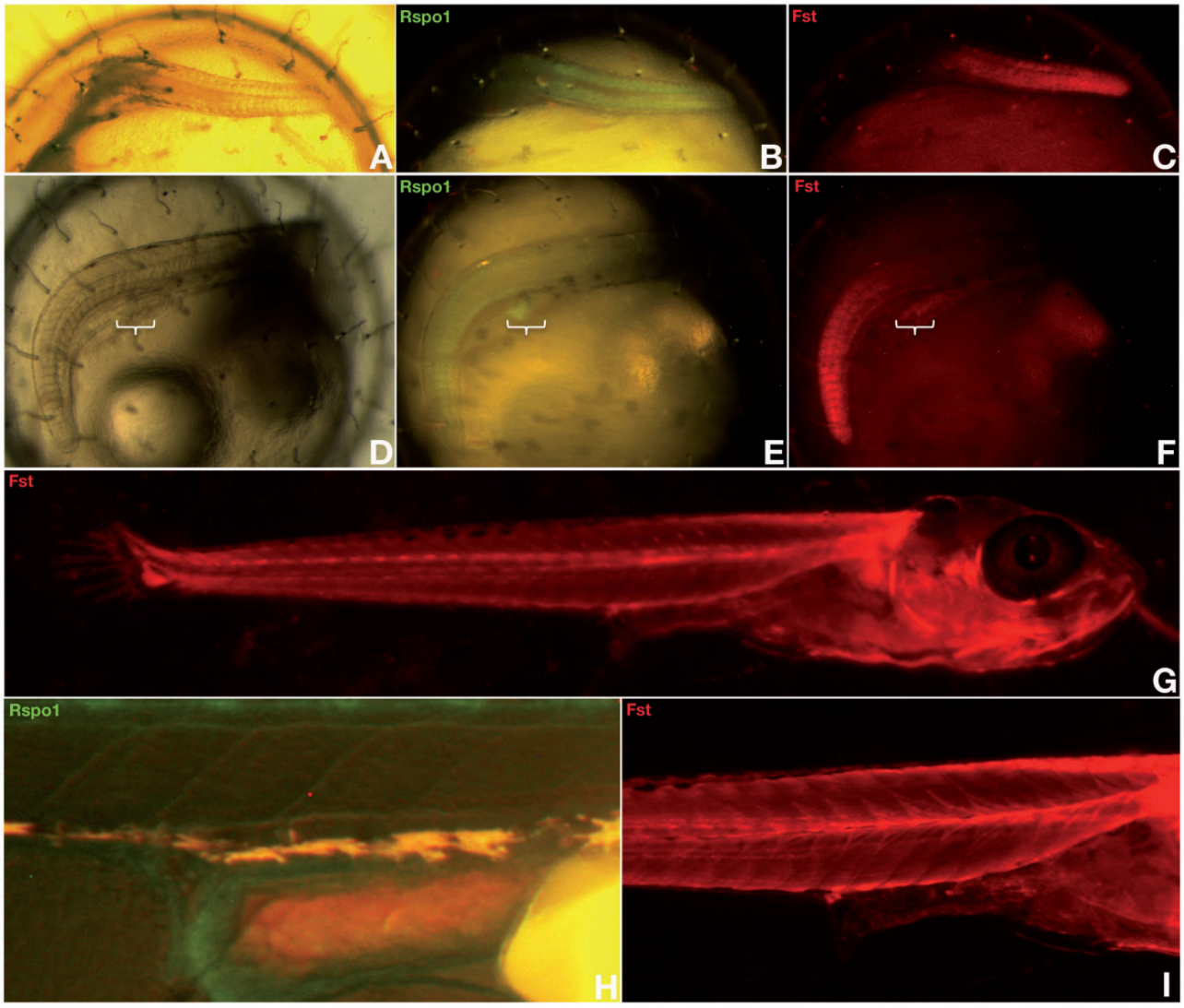


Fig. 7. *R-spondin-1* and *follistatin* spatial and temporal expression patterns during medaka gonad primordium formation. Two BAC reporter fish transgenic lines expressing either GFP or mCherry were established to follow r-spo-1 and fst expression dynamics during gonad formation in vivo. Neither *r-spo-1* nor *fst* expression was detected in the presumptive gonadal mesoderm at early stages of development (*A*–*C*). Between stages 33 and 35 in the dorsal region of the hindgut *r-spo-1* and *fst* are co-expressed (*D*–*F*). At hatching stage, although ubiquitously expressed, neither *r-spo-1* nor *fst* were detected in the gonadal primordium (*G*–*I*).

period additionally rules out any involvement of the Hedgehog pathway for gonad induction. Further on, while the ovary develops in the juvenile female, the strict co-expression of *r-spo-1* and *ptch-2*, together with *sox9b* in the somatic cells of the germinal cradles, likely indicates a role in differentiating and specifying the somatic supporting lineage of the ovary. Of note the early decoupling of *r-spo-1 and fst* expression patterns, although expected to be involved in the same signaling pathway in mammals (Yao et al. 2004), de facto invalidates such an interplay in the medaka ovarian cradle.

Our results support the view that *r-spo-1* has a globally conserved female-specific expression profile in vertebrate gonads despite some slight but intriguing divergences. In all analyzed vertebrate species it is apparent that *r-spo-1* expression goes along with the process of somatic cell organization within the young ovary (table 1). In mouse, chicken, and medaka, *r-spo-1* expression at that time, predominantly or exclusively in the somatic cells, suggests its implication in a conserved pathway leading to folliculogenesis. On the other hand, the absence of medaka *r-spo-1* expression in gonad embedded germ cells, unlike in zebrafish, mice, chicken, and turtle, reveals that a role for germ cell development might not be accordingly conserved in the adult gonad of medaka.

In mammals, R-spo-1 engages the effector pathway of Wnt signaling and β-catenin and thereby activates Fst expression (Carmon et al. 2011). Hence, also medaka *fst* spatial and temporal expression pattern was expected to overlap with *r-spo-1*. While medaka *r-spo-1* and *fst* indeed display similar expression in adult tissues (fig. 3*B*), medaka *fst* was, however, absent during the early and late stages of ovarian induction and development. *Fst* expression was perceptible only in the interstitium and epithelium of old ovaries together with *r-spo-1* (fig. 8*M*, *N*, and *Q*–*S*). Unexpectedly, *fst* was also

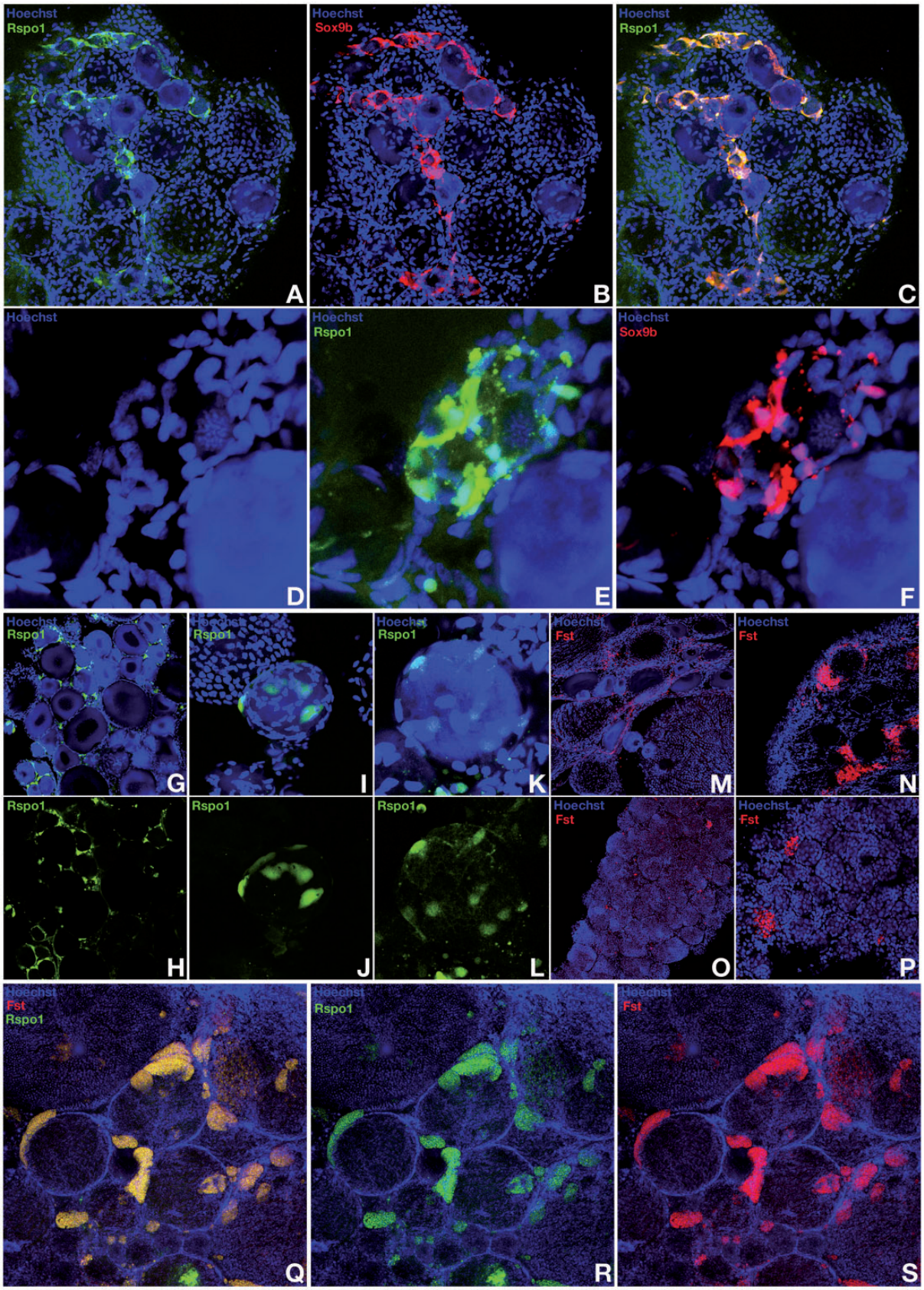


Fig. 8. *R-spondin-1* and *follistatin* expression in the adults gonads. In the ovary of juvenile females, *r-spondin* expression is restricted to somatic cells surrounding the germ line stem cells of the germinal cradle within the ovarian cord (*A*–*F*). Co-expression with *sox9b* (*A* to *F*) and *patched-2* (see fig. 2*A*–*D*) is observed. *R-spondin-1* expression in granulosa cells is also detected around young oocytes (*G*–*L*). In the ovary, *follistatin* expression was only detected in the interstitium (*M* and *N*) and in the ovarian epithelium together with *r-spondin-1* (*Q*–*S*). Sparse clusters of *follistatin* expressing interstitial somatic cells are also detected in adult testes (*O* and *P*).

Table 1. Comparative Analysis of Gonadal Expression of R-spondin1, Follistatin, Patched2, and Foxl2 in Vertebrates.

Genes		Gonadal Expression Patterns (Adult Ovaries/Testes)		
		Medaka	Mouse	Other Vertebrates
R-spondin 1	O	(A) Somatic cells surrounding the germ cells and germline stem cells of the ovarian cord. Granulosa cells of the young oocytes. (E) Not expressed during early gonadal development.	(E) Predominantly in the somatic cells of the developing ovary (Smith et al. 2008). (E) Germ cells (during meiosis at low levels) (Smith et al. 2008).	Chicken: (E) outer cortical zone of the developing ovary and germ cells during meiosis (Smith et al. 2008). *Danio*: (A) granulosa and theca cells. Premature germ cells, oogonia, primary oocytes (Zhang et al. 2011).
	T	N.D.	N.D. (Smith et al. 2008).	Chicken: (E) not detected (Smith et al. 2008). *Danio*: (A) Leydig cells, spermatogonia, and spermatocytes (Zhang et al. 2011).
Follistatin	O	(A) Ovarian eptithelium and interstitium of old ovaries.	(E) Somatic cells of the embryonic ovary (Menke and Page 2002; Yao et al. 2004), (A) co-localization with Foxl2 (Kashimada et al. 2011).	Sheep: (A) granulosa cells of the growing follicles II and III (Tisdall et al. 1994).
	T	N.D.	(E) Not detected in embryonic testes (Menke and Page 2002).	Rat: (A) Sertoli and endothelial cells, germ cells, spermatogonia, spermatocytes, and round spermatids (Meinhardt et al. 1998).
Patched 2	O	(E) Not expressed during early gonadal development. (A) Somatic cells surrounding the germ cells and germline stem cells of the ovarian cord (co-expression with r-spondin1). Additional expression in the interstitium of the old ovaries.	(A) Highly expressed (testis-specific splice variants) (Szczepny et al. 2006).	Tammar wallaby: (E) expressed throughout the development of the embryonic ovary. (A) Abundant in granulosa, cumulus, and theca cells of the adult ovary. Very weak in germ cells (O'Hara et al. 2011).
	T	N.D.	(A) Lowly expressed (Spicer et al. 2009).	Tammar: (E) Leydig cells in the interstitium of the developing testes. (A) Restricted to Sertoli cells of the adult testes (O'Hara et al. 2011).
FoxL 2	O	(A) Ovarian germline stem cells, initial stage of post meiotic oocytes. Sub-population of theca and granulosa cells together with aromatase (Cyp19a1) expression.	(E) From 12.5 dpc in mesenchymal pre-granulosa cells and (A) later in granulosa cells (Schmidt et al. 2004). (A) Small and medium size follicles (Pisarska et al. 2004).	Chicken: (A) medullar part of the ovary, maturing and ovulated oocytes. Granulosa cells, weak in theca cells layer (Govoroun et al. 2004).
	T	N.D.	N.D.	

NOTE.—(E) embryonic expression; (A) adult expression.

detected in testes (fig. 8*O* and *P*). This expression pattern indicates that in medaka and unlike in mammals, Fst is probably neither acting throughout early female gonad patterning nor during maintenance of cell identity in the adult ovary. Instead medaka *fst* expression appears to be more a marker of the aging ovary, probably acting during follicular atresia. Of interest, while exclusive ovarian expression is generally described for *R-spo-1* and *Fst* in vertebrates, such strict female dimorphism was not observed for *r-spo-1* and *fst* in zebrafish, rat, and medaka (table 1 and fig. 3).

Role of the Hedgehog Pathway during Gonad Development and Maintenance

The gonadal expression pattern of two components of the Hedgehog pathway in medaka is peculiar and different from what has so far been reported for other vertebrates including mammals (table 1). Unlike in mammals it appears that medaka gonadal HH signaling through the *patched-2* receptor is not involved in inducing and specifying the gonad primordia. It would rather act late exclusively in the process of somatic cell differentiation in the ovarian cradle. In strict contrast to its mammalian counterparts, the quasi-absence of medaka *patched-2* expression during testis formation in larvae and for testis cell identity maintenance in the adult rules out any functions during these processes (table 1). Certainly, the low testicular expression of *patched-2* together with Dmrt1a/1bY-induced transcriptional upregulation of the HH antagonist *hhip* indicates a general function of Dmrt1 in actively downregulating the Hedgehog pathway in medaka testes. Taken together, we can conclude that although apparently downregulated at the transcriptional level, a background expression of *patched-2* remains. This phenomenon is known as illegitimate transcription (Chelly et al. 1989; McLeod and Cooke 1989). We speculate that the high expression of the *hhip* hedgehog pathway antagonist in gonads (about 5 to 10 times higher than patched1/2

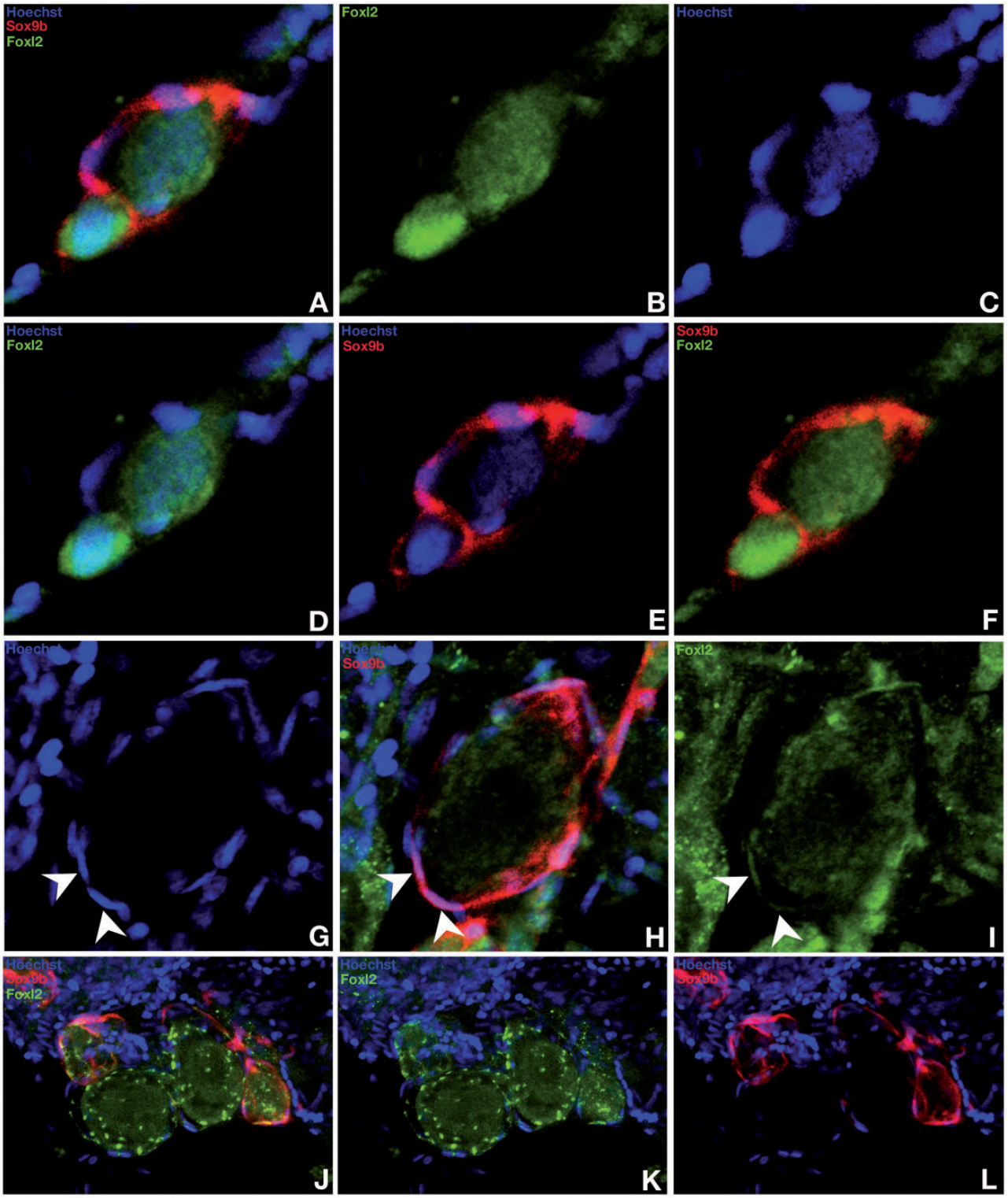


Fig. 9. *Foxl2* protein localization in the ovarian cradle. During development from germ line stem cells to oocytes within the germinal cradle, foxl2 localization is first detected in the germ line stem cells and remains during meiosis until early oogenesis (*A*–*L*). Concomitantly, the accompanying somatic cells of the supporting layer progressively loose *sox9b* expression while *foxl2* expression rises (*A*–*L*).

expression [fig. 3]) likely prevents any hedgehog activation resulting from leaky background expression of patched receptors. Interestingly, several putative Sox9 binding sites are present in the patched-2 promoter region and might explain the strict *sox9b*/*patched-2* co-expression observed in the supporting cells of the ovarian cradle.

A Newly Identified Subpopulation of Theca Cells Expressing *Aromatase* but Not Foxl2

Examination of Foxl2 protein distribution in the medaka ovary allowed us to define a new subpopulation of theca cells expressing Foxl2. Also expressing *cyp19a1* (aromatase) these cells are then suspected of having a steroidogenic

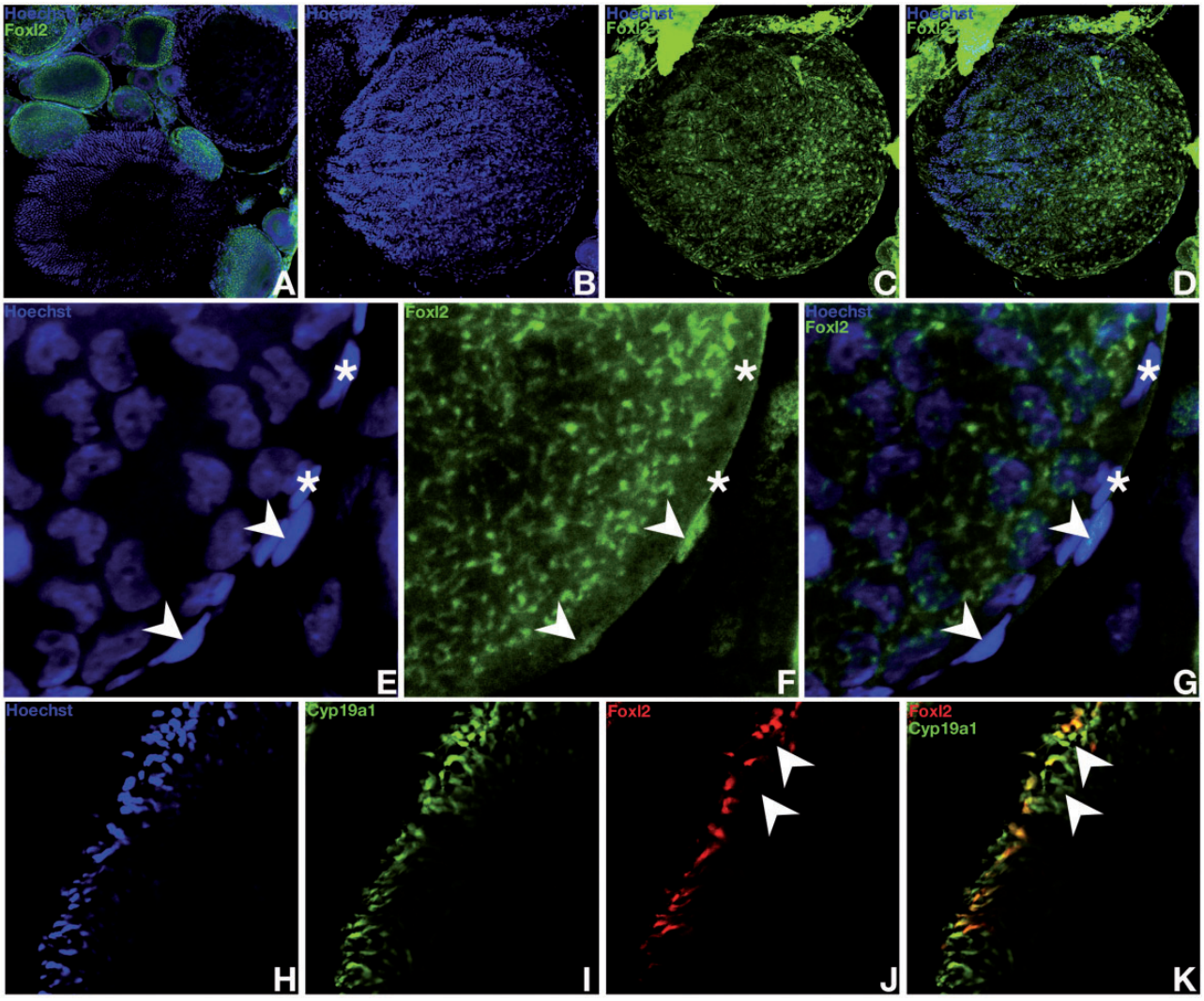


Fig. 10. Protein expression and localization of *foxl2* in the ovary. *foxl2* immunostaining is present in the nuclei of the follicular cells of the previtellogenic and vitellogenic follicles (*A*). In vitellogenic follicles, *foxl2* protein is detected in all granulosa cells (*B–D*). Nuclear localization of *foxl2* in few theca cells (*E–G*). *foxl2* expression occurs only in a sub-population of theca cells (arrow heads vs. asterisk in *E–G*) as shown by comparison with the thecal layer marker aromatase *cyp19a1* (*H–K*).

activity. In contrast to mammals where ovarian-type *aromatase* is only produced by granulosa cells, the biological significance of such cells expressing *cyp19a1* and Foxl2 within the thecal layer remains unclear. Of note, in ovaries of *cyp19a1/p450c17l* double transgenic medaka reporter lines, two subpopulations of theca cells were previously identified, being either *cyp19a1* or *p450c17l* positive in a mutually exclusive manner (Nakamura et al. 2009). The *cyp19a1* expressing subpopulation of theca cells was already considered as the precursors of the theca lineage (Nakamura et al. 2009).

The strict co-expression of *Foxl2* and *Aromatase* (*Cyp19*) in the mammalian ovary led to the further demonstration that *Foxl2* is involved in the regulation of estrogen synthesis via direct transcriptional upregulation of ovarian-type *Aromatase* (see Pannetier et al. 2006 for review). Surprisingly in medaka we found, within the thecal layer, aromatase-only positive theca cells that remained Foxl2-negative. In that perspective it is interesting to note that birds also have multiple populations of theca cells some of which are also steroidogenic (Nitta et al. 1991). In contrast to the main consensus, the discordance of spatial expression patterns of Foxl2 and ovarian-type *aromatase* (*cyp19a1*) calls into question an exclusive transcriptional regulation of *cyp19a1* by Foxl2 in the ovary of medaka (fig. 11). Although we cannot exclude that aromatase-only positive cells have not been previously also positives for Foxl2, implying the requirement of Foxl2 for the induction of the aromatase expression, our results indicate that foxl2 is nevertheless not required for the maintenance of the aromatase expression.

Variable Molecular Interplay among the Repertoire of Gonadal Markers during Medaka Gonad Formation and Maintenance

We could show a direct regulation of the Hedgehog and R-spo-1 pathways by Dmrt1bY (fig. 12). It is thus becoming apparent that despite its tangible requirement for mammalian testis formation and later on in regulating

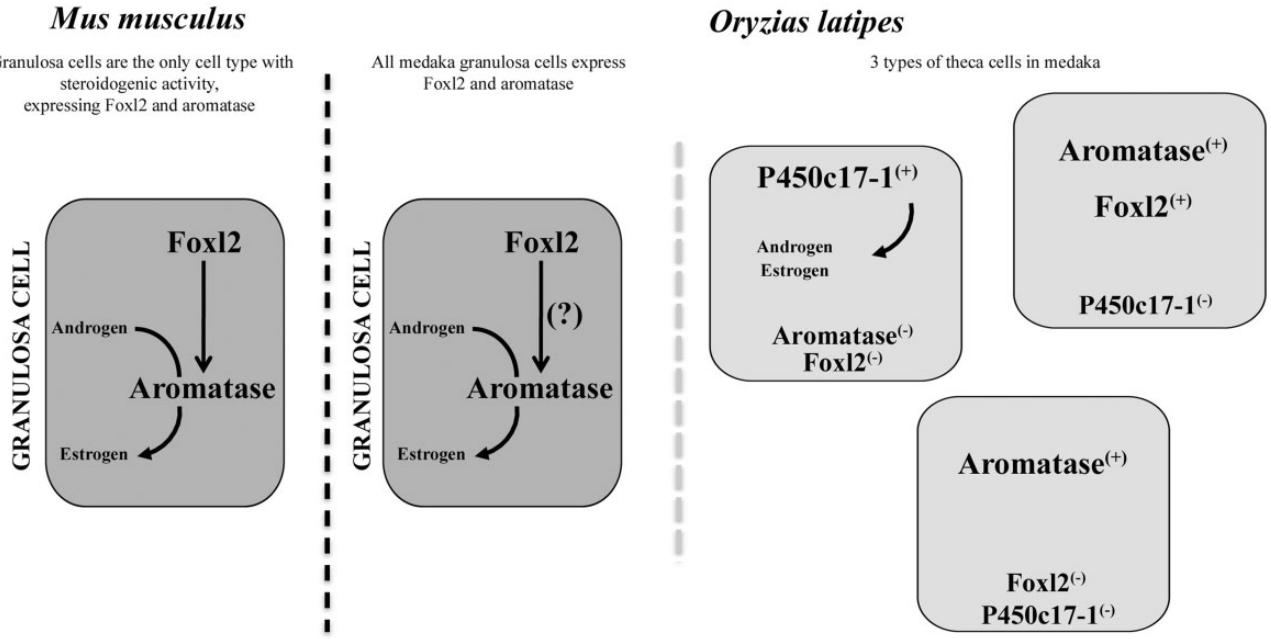


Fig. 11. Schematic representation of granulosa and theca cell populations in mouse and medaka. In mammals, granulosa cells are the only cell type with steroidogenic activity, expressing both *foxl2* and *aromatase*. *Aromatase* expression is directly induced by *foxl2*. In medaka, like in mammals, granulosa cells express both Foxl2 and *aromatase*. Examination of Foxl2 protein distribution in the medaka ovary revealed a new subpopulation of theca cells expressing Foxl2.

Leydig and myoid cell function (Clark et al. 2000; Pierucci-Alves et al. 2001; Canto et al. 2004), the Hedgehog pathway might not only be dispensable during medaka male gonadogenesis and maintenance but even needs to be suppressed as it is actively repressed by Dmrt1 genes. This would also explain the specific lack of *ptch-2* expression in medaka testes shown to be indirectly downregulated by Dmrt1bY (see fig. 12 for summary). In contrast the female-specific, Dmrt1a-triggerd upregulation of *follistatin* transcription might nevertheless point out the importance of the R-spo-1 pathway during female gonad differentiation although the upstream components of this same pathway are tightly regulated by the Dmrt1 co-orthologs (fig. 12).

Importantly, showing that under certain conditions Dmrt1 paralogs are able to strongly upregulate the female-specific r-spondin-1 gene expression, we could also demonstrate that Dmrt1bY/Dmrt1a-triggered regulations are highly dependent of the cellular context and might suggest requirement of co-factors (fig. 12). These findings are reminiscent of observations showing Stra8 transcription to be directly repressed by DMRT1 in mouse testes while activated in the fetal ovary (Krentz et al. 2011).

Interestingly, the canonical Wnt/ β-catenin (R-spo-1 and Fst) pathway strongly antagonizes FGF9, a robust component of the male sex-determining cascade in mammals (Kim et al. 2006; Matson et al. 2011). However, FGF9 is absent in the fish lineage and no indication of a redundant action of related FGF has been obtained, questioning the importance of FGF signaling in fish sex determination (Forconi et al. 2013). Into that direction the importance of the FGF9 signaling in vertebrates is actually also challenged by the finding that in chicken embryos *Fgf9* does not show any sexually dimorphic expression pattern during gonadal differentiation (Cutting et al. 2013). Similarly, although phylogenetically preserved, the Sox9 gene, a direct target of Sry in mammals, has been shown to be functionally dispensable for medaka testis determination (Nakamura et al. 2008a, 2012). In this context our data might reflect a profound reorganization of that part of the fish gonadal regulatory network compared with mammals. While some components such ase DMRT1, SOX9, FOXL2 and pathways such as Hedgehog or R-spo1/Wnt/Fst of the gonadal gene regulatory network are conserved on the DNA sequence level across phyla, their functions, regulation, and interplays might be considerably different.

Materials and Methods

BAC Recombination

Bacterial artificial chromosome clones encompassing medaka *patched-2* (ola1-199K19), *follistatin* (ola1-124N21), or *r-spondin-1* (ola1-158A23) genomic regions were obtained from NRBP Medaka (http://www.shigen.nig.ac.jp/medaka/, last accessed August 12, 2013). A BAC transgenic method using homologous recombination was employed to generate the reporter constructs as previously described (Nakamura et al. 2008b; Herpin et al. 2010). The following primers were used to amplify *eGFP/mCherry* fragments for homologous recombination into the different BAC clones: BAC-Ptchd2-GFP-Fw: GCTGAACTCGCACCGATTCTGCGTCGCCTCCTGTTACCCGTCTTTGGACTATGGATATCATTTCTGTCGCCTTAAAG, BAC-Ptchd2-GFP-Rv: CGCAAGCGGCTGGGAGCGCGTATAACTCGGGGGTAAATCTCCAAAGACGCCAGAACAAACGACCCAACACCGTGCG; BAC-Fst-Cherry-Fw: CTTTTGCGCTGCTTGTGTCAAATACGTGGCTCACTTTGCCTCTCCATCATGCTTGGGCCACCGGTCGCCACCATGGT, BAC-Fst-Cherry-Rv: CTTACC

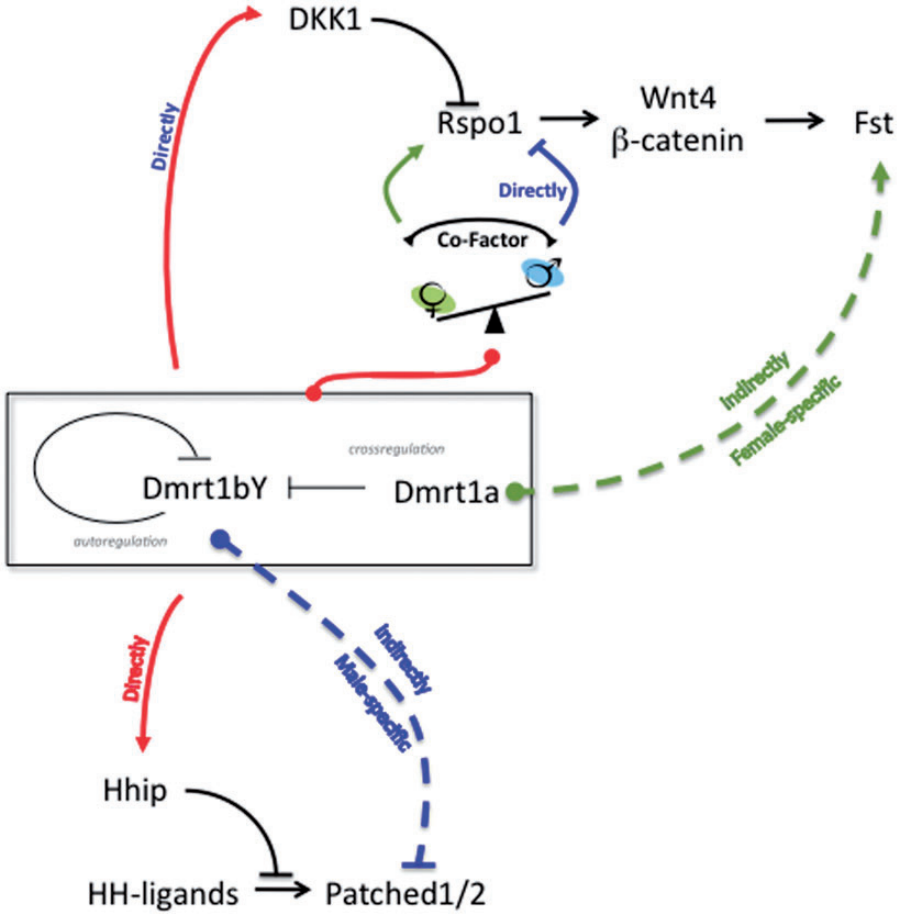


Fig. 12. The *Dmrt1bY/Dmrt1a* gene regulatory networks during gonadal formation in medaka. Interaction scheme of the possible Dmrt1bY/Dmrt1a-triggered regulations of the Hedgehog and Wnt4 pathways during gonadal formation in medaka. Solid red arrows indicate for both Dmrt1bY and Dmrt1a positive regulation while dashed lines indicate sex-specific regulations. The green dashed line indicates a Dmrt1a, female-specific, indirect positive regulation favoring the expression of *follistatin*, while the blue dashed line reports a Dmrt1bY, male-specific, indirect repression of *patched-2* transcription. For the Dmrt1bY/Dmrt1a-triggered transcriptional regulation of *R-spondin1*, depending on the cellular contexts, the involvement of a sex-specific co-factor is proposed.

TTGAACCTTCTGATGTTCCATGAGGTGACAAAGCCACATG AAGAAGAGAGTCGACCAGTTGGTGATTTTG; BAC-Rspo1-GFP-Fw: GATCCATCTGGTTGCAGGGGGGGACCTTGCACA GCCTGGAAGGCAGCAGGGACTCCACCGGTCGCCACCATG GTG, BAC-Rspo1-GFP-Rv: CTTCTCGCCTTGGAGAGTTTGAC AACATCGCTGTGACCCATGGAGCTGAGAATGAGTCGACC AGTTGGTATTTTG. After homologous recombination, the generated fragments were inserted into the BAC clones in frame downstream of the translation initiation site of the targeted genes.

Generation of BAC Transgenic Medaka Lines and Imaging Analyses

The Carbio (WLC# 2674) strain of medaka (*Oryzias latipes*) was used for establishment of the transgenic lines. Microinjection of DNA was performed as described previously (Herpin et al. 2009) using BAC clone DNA at a concentration of 50–100 ng/mL. Adult G0 fish were then screened for fluorescence, and positive individuals were raised to adulthood. Siblings from positive G0 fish were mated to each other and the offspring were again sorted for fluorescence. *Sox9b* and *cyp19a1* (*aromatase*) transgenic lines were described earlier (Nakamura et al. 2008a, 2009). For imaging embryos, hatchlings or tissues were mounted with 1.2% low melting temperature agarose. Confocal pictures and image stacks were acquired using a Nikon C1 (eclipse Ti) confocal laser scanning microscope and the NIS elements AR software.

Immunochemistry

Ovaries or testes from juvenile and adult fish were fixed with 4% paraformaldehyde/balanced salt solution (111 mM NaCl, 5.37 mM KCl, 1 mM $CaCl_2 \cdot H_2O$, 0.6 mM $MgSO_4 \cdot 7H_2O$, 5 mM Hepes, pH 7.3) for 30 min on ice. After fixation samples were washed three times for 10 min with MABT buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5, 0.1% Triton X-100) and subsequently twice for 30 min with MABDT buffer (MABT buffer complemented with 1% BSA and 1% DMSO). After blocking in MABDT-blocking buffer (MABDT buffer supplemented with 2% lamb or sheep serum), the tissues were incubated in MABDT-blocking buffer together with the primary antibody (1:150 dilution) overnight at 4 °C. Samples were then washed three times 5 min in MABDT buffer and

washed again four times for 30 min in MABDT-blocking buffer on ice. Thereafter samples were incubated overnight at 4 °C with the secondary antibody diluted at 1:600 in MABDT-blocking buffer. Finally the tissues were washed in PBS, stained with Hoechst solution for 3 h at 4 °C, mounted, and imaged with a confocal microscope (Nikon C1 confocal microscope).

In vivo Chromatin Immunoprecipitation

For in vivo chromatin immunoprecipitation, the EpiQuik Tissue Chromatin Immunoprecipitation kit (Epigentek) was used according to the manufacturer's instructions, using 20 mg of testis tissue samples either from *dmrt1bY::GFP* or *deltadmrt1bY::GFP* transgenic fish (Hornung et al. 2007; Herpin et al. 2010) (20 testes for each) and GFP antibody (3 µg, Upstate) for immunoprecipitation. After tissue disaggregation and cell re-suspension, DNA was sheared by sonication (9 pulses of 10 s with an amplitude of 10%). After immunoprecipitation ([DKK1-(1/2) Fw01]: 5′-GATAACTCC GGCTGGGACGTTGAC-3′/[DKK–(1/2) Rv01]: 5′-ACAACAC TGAAGTGCTACAGAAGTC-3′; [DKK1-(3) Fw02]: 5′-AGTAT CAAGTGCTCAAGACGATCC-3′/[DKK1-(3) Rv02]: 5′-TACGA GCTGACATGTTCACATCTGCC-3′; [DKK1-(4/5) Fw03]: 5′-GC TGCAAGACAGGAAGAC-3′/[DKK1-(4/5) Rv03]: 5′-GTTAAT AGTCATGCTCAGTCTG-3′; [R-spo1-(1) Fw01]: 5′-CATCGGA TTTAACAGTTATGATTGC-3′/[R-spo1-(1) Rv01]: 5′-CGATAG TGATTGGTCAGTTA-3′; [R-spo1-(2/3) Fw02]: 5′-CATCGTGC CAACTTACAGCCAATC-3′/[R-spo1-(2/3) Rv02]: 5′-CTACCA AGACACGCTAGAAGCTCC; [R-spo1-(4) Fw03]: 5′-AAGTTG CTCAACACTTGTACAC-3′/[R-spo1-(4) Rv03]: 5′-AAGCAGA GACAATAGAATGCATC-3′; [R-spo1-(5) Fw04]: 5′-ATAAAC ATGTACAACAGTCATCTG-3′/[R-spo1-(5) Rv04]: 5′-TTCCA CTCTCGGCAAGAAATCAG-3′; [HHIP-Fw01]: 5′-TAGAGTAC GTCCGTCTACTG-3′/[HHIP-Rv01]: 5′-TGACAACAAAGTCG CAA-3′) primer sets were used for enrichment quantification by real-time PCR.

Bioinformatic Analyses

Binding sites for Dmrt1bY were identified using the matrix provided by (Murphy et al. 2007) together with the Regulatory Sequence Analysis Tools portal; RSat (http://rsat.ulb.ac.be/rsat/, last accessed August 12, 2013).

In Vitro Expression Regulation Analyses and Real-Time PCR

Medaka spermatogonial (SG3) and fibroblast-like (OLF) cell lines were cultured as described (Etoh 1988; Hong et al. 2004). For transfection cells were grown to 80% confluency in 6-well plates and transfected with 5 µg expression vector using FuGene (Roche) reagent as described by the manufacturer.

Total RNA was extracted from fish tissues or transfected cells using the TRIZOL reagent (Invitrogen) according to the supplier's recommendation. After DNase treatment, reverse transcription was done with 2 µg total RNA using RevertAid First Strand Synthesis kit (Fermentas) and random primers. Real-time quantitative PCR was carried out with SYBR Green reagents and amplifications were detected with an i-Cycler (Biorad). All results are averages of at least two independent reverse transcription reactions. Error bars represent the standard deviation of the mean. Relative expression levels (according to the equation 2–DeltaCT) were calculated after correction of expression of elongation factor 1 alpha (ef1alpha) and brain expression was set to 1 as a reference.

Supplementary Material

Supplementary figure S1 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by a grant of the Graduate Training Program of the DFG (GK 1048, Molecular Basis of Organ Development in Vertebrates). Part of this study was also supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 222719-LIFECYCLE (to Y.G and B.N). This study was also partly funded by a grant dedicated to the "Mechanisms of sex determination in zebrafish" (UO Subaward Number 212791A).

References

Bitgood MJ, Shen L, McMahon AP. 1996. Sertoli cell signaling by Desert hedgehog regulates the male germline. *Curr Biol.* 6:298–304.

Bourguiba S, Genissel C, Lambard S, Bouraima H, Carreau S. 2003. Regulation of aromatase gene expression in Leydig cells and germ cells. *J Steroid Biochem Mol Biol.* 86:335–343.

Bowles J, Feng CW, Spiller C, Davidson TL, Jackson A, Koopman P. 2010. FGF9 suppresses meiosis and promotes male germ cell fate in mice. *Dev Cell.* 19:440–449.

Canto P, Soderlund D, Reyes E, Mendez JP. 2004. Mutations in the desert hedgehog (DHH) gene in patients with 46,XY complete pure gonadal dysgenesis. *J Clin Endocrinol Metab.* 89:4480–4483.

Carmon KS, Gong X, Lin Q, Thomas A, Liu Q. 2011. R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. *Proc Natl Acad Sci U S A.* 108:11452–11457.

Chassot AA, Gregoire EP, Magliano M, Lavery R, Chaboissier MC. 2008. Genetics of ovarian differentiation: Rspo1, a major player. *Sex Dev.* 2: 219–227.

Chelly J, Concordet JP, Kaplan JC, Kahn A. 1989. Illegitimate transcription: transcription of any gene in any cell type. *Proc Natl Acad Sci U S A.* 86:2617–2621.

Chuang PT, McMahon AP. 1999. Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* 397: 617–621.

Clark AM, Garland KK, Russell LD. 2000. Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. *Biol Reprod.* 63:1825–1838.

Crisponi L, Deiana M, Loi A, et al. (22 co-authors). 2001. The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat Genet.* 27:159–166.

Cutting A, Chue J, Smith CA. 2013. Just how conserved is vertebrate sex determination? *Dev Dyn.* 242:380–387.

De Felici M. 2009. Primordial germ cell biology at the beginning of the XXI century. *Int J Dev Biol.* 53:891–894.

Etoh. 1988. Establishment and characterization of various cell lines from Medaka (Teleostei) Invertebrate and fish tissue culture, p 266–269.

Forconi M, Canapa A, Barucca M, et al. (14 co-authors). 2013. Characterization of sex determination and sex differentiation genes in latimeria. *PLoS One* 8:e56006.

Franco HL, Yao HH. 2012. Sex and hedgehog: roles of genes in the hedgehog signaling pathway in mammalian sexual differentiation. *Chromosome Res.* 20:247–258.
Garcia-Ortiz JE, Pelosi E, Omari S, et al. (12 co-authors). 2009. Foxl2 functions in sex determination and histogenesis throughout mouse ovary development. *BMC Dev Biol.* 9:36.
Giraldo P, Montoliu L. 2001. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res.* 10:83–103.
Govoroun MS, Pannetier M, Pailhoux E, Cocquet J, Brillard JP, Couty I, Batellier F, Cotinot C. 2004. Isolation of chicken homolog of the FOXL2 gene and comparison of its expression patterns with those of aromatase during ovarian development. *Dev Dyn.* 231:859–870.
Graham P, Penn JK, Schedl P. 2003. Masters change, slaves remain. *Bioessays* 25:1–4.
Guiguen Y, Fostier A, Piferrer F, Chang CF. 2010. Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol.* 165:352–366.
Haag ES, Doty AV. 2005. Sex determination across evolution: connecting the dots. *PLoS Biol.* 3:e21.
Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, Tanaka M, Schartl M. 2010. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. *PLoS Genet.* 6:e1000844.
Herpin A, Nakamura S, Wagner TU, Tanaka M, Schartl M. 2009. A highly conserved cis-regulatory motif directs differential gonadal synexpression of Dmrt1 transcripts during gonad development. *Nucleic Acids Res.* 37:1510–1520.
Herpin A, Schartl M. 2009. Molecular mechanisms of sex determination and evolution of the Y-chromosome: insights from the medakafish (Oryzias latipes). *Mol Cell Endocrinol.* 306:51–58.
Herpin A, Schartl M. 2008. Regulatory putsches create new ways of determining sexual development. *EMBO Rep.* 9:966–968.
Herpin A, Schartl M. 2011. Sex determination: switch and suppress. *Curr Biol.* 21:R656–R659.
Herpin A, Schindler D, Kraiss A, Hornung U, Winkler C, Schartl M. 2007. Inhibition of primordial germ cell proliferation by the medaka male determining gene Dmrt I bY. *BMC Dev Biol.* 7:99.
Hong Y, Liu T, Zhao H, Xu H, Wang W, Liu R, Chen T, Deng J, Gui J. 2004. Establishment of a normal medakafish spermatogonial cell line capable of sperm production in vitro. *Proc Natl Acad Sci U S A.* 101:8011–8016.
Hooper JE, Scott MP. 2005. Communicating with Hedgehogs. *Nat Rev Mol Cell Biol.* 6: 306–317.
Hornung U, Herpin A, Schartl M. 2007. Expression of the male determining gene dmrt1bY and its autosomal coorthologue dmrt1a in medaka. *Sex Dev.* 1:197–206.
Hudson QJ, Smith CA, Sinclair AH. 2005. Aromatase inhibition reduces expression of FOXL2 in the embryonic chicken ovary. *Dev Dyn.* 233: 1052–1055.
Ingham PW, McMahon AP. 2001. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* 15:3059–3087.
Kashimada K, Pelosi E, Chen H, Schlessinger D, Wilhelm D, Koopman P. 2011. FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology* 152:272–280.
Kim Y, Kobayashi A, Sekido R, DiNapoli L, Brennan J, Chaboissier MC, Poulat F, Behringer RR, Lovell-Badge R, Capel B. 2006. Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biol.* 4:e187.
Kobayashi T, Matsuda M, Kajiura-Kobayashi H, Suzuki A, Saito N, Nakamoto M, Shibata N, Nagahama Y. 2004. Two DM domain genes, DMY and DMRT1, involved in testicular differentiation and development in the medaka, Oryzias latipes. *Dev Dyn.* 231:518–526.
Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R. 1991. Male development of chromosomally female mice transgenic for Sry. *Nature* 351:117–121.
Krentz AD, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. 2011. DMRT1 promotes oogenesis by transcriptional activation of Stra8 in the mammalian fetal ovary. *Dev Biol.* 356: 63–70.
Liu CF, Liu C, Yao HH. 2010. Building pathways for ovary organogenesis in the mouse embryo. *Curr Top Dev Biol.* 90:263–290.
Masuyama H, Yamada M, Kamei Y, Fujiwara-Ishikawa T, Todo T, Nagahama Y, Matsuda M. 2012. Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka. *Chromosome Res.* 20:163–176.
Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* 476:101–104.
Matsuda M. 2005. Sex determination in the teleost medaka, Oryzias latipes. *Annu Rev Genet.* 39:293–307.
Matsuda M, Nagahama Y, Shinomiya A, et al. (13 co-authors). 2002. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417:559–563.
McLaren A, Southee D. 1997. Entry of mouse embryonic germ cells into meiosis. *Dev Biol.* 187:107–113.
McLeod JF, Cooke NE. 1989. The vitamin D-binding protein, alpha-fetoprotein, albumin multigene family: detection of transcripts in multiple tissues. *J Biol Chem.* 264:21760–21769.
Meinhardt A, O'Bryan MK, McFarlane JR, Loveland KL, Mallidis C, Foulds LM, Phillips DJ, de Kretser DM. 1998. Localization of follistatin in the rat testis. *J Reprod Fertil.* 112:233–241.
Menke DB, Page DC. 2002. Sexually dimorphic gene expression in the developing mouse gonad. *Gene Expr Patterns.* 2:359–367.
Murphy MW, Zarkower D, Bardwell VJ. 2007. Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. *BMC Mol Biol.* 8:58.
Nakamoto M, Matsuda M, Wang DS, Nagahama Y, Shibata N. 2006. Molecular cloning and analysis of gonadal expression of Foxl2 in the medaka, Oryzias latipes. *Biochem Biophys Res Commun.* 344: 353–361.
Nakamura S, Aoki Y, Saito D, Kuroki Y, Fujiyama A, Naruse K, Tanaka M. 2008a. Sox9b/sox9a2-EGFP transgenic medaka reveals the morphological reorganization of the gonads and a common precursor of both the female and male supporting cells. *Mol Reprod Dev.* 75: 472–476.
Nakamura S, Kobayashi K, Nishimura T, Higashijima S, Tanaka M. 2010. Identification of germline stem cells in the ovary of the teleost medaka. *Science* 328:1561–1563.
Nakamura S, Kurokawa H, Asakawa S, Shimizu N, Tanaka M. 2009. Two distinct types of theca cells in the medaka gonad: germ cell-dependent maintenance of cyp19a1-expressing theca cells. *Dev Dyn.* 238:2652–2657.
Nakamura S, Saito D, Tanaka M. 2008b. Generation of transgenic medaka using modified bacterial artificial chromosome. *Dev Growth Differ.* 50:415–419.
Nakamura S, Watakabe I, Nishimura T, Toyoda A, Taniguchi Y, Tanaka M. 2012. Analysis of medaka sox9 orthologue reveals a conserved role in germ cell maintenance. *PLoS One* 7:e29982.
Nanda I, Kondo M, Hornung U, et al. (12 co-authors). 2002. A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. *Proc Natl Acad Sci U S A.* 99: 11778–11783.
Nitta H, Osawa Y, Bahr JM. 1991. Multiple steroidogenic cell populations in the thecal layer of preovulatory follicles of the chicken ovary. *Endocrinology* 129:2033–2040.
O'Hara WA, Azar WJ, Behringer RR, Renfree MB, Pask AJ. 2011. Desert hedgehog is a mammal-specific gene expressed during testicular and ovarian development in a marsupial. *BMC Dev Biol.* 11:72.
Ottolenghi C, Pelosi E, Tran J, Colombino M, Douglass E, Nedorezov T, Cao A, Forabosco A, Schlessinger D. 2007. Loss of Wnt4 and Foxl2 leads to female-to-male sex reversal extending to germ cells. *Hum Mol Genet.* 16:2795–2804.
Pannetier M, Fabre S, Batista F, Kocer A, Renault L, Jolivet G, Mandon-Pepin B, Cotinot C, Veitia R, Pailhoux E. 2006. FOXL2 activates P450 aromatase gene transcription: towards a better characterization of the early steps of mammalian ovarian development. *J Mol Endocrinol.* 36:399–413.

Parma P, Radi O, Vidal V, Chaboissier MC, Dellambra E, Valentini S, Guerra L, Schedl A, Camerino G. 2006. R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nat Genet.* 38: 1304–1309.

Pierucci-Alves F, Clark AM, Russell LD. 2001. A developmental study of the Desert hedgehog-null mouse testis. *Biol Reprod.* 65:1392–1402.

Pisarska MD, Bae J, Klein C, Hsueh AJ. 2004. Forkhead l2 is expressed in the ovary and represses the promoter activity of the steroidogenic acute regulatory gene. *Endocrinology* 145:3424–3433.

Schmidt D, Ovitt CE, Anlag K, Fehsenfeld S, Gredsted L, Treier AC, Treier M. 2004. The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. *Development* 131:933–942.

Sekido R, Lovell-Badge R. 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* 453: 930–934.

Shinomiya A, Shibata N, Sakaizumi M, Hamaguchi S. 2002. Sex reversal of genetic females (XX) induced by the transplantation of XY somatic cells in the medaka, Oryzias latipes. *Int J Dev Biol.* 46: 711–717.

Siegfried KR. 2010. In search of determinants: gene expression during gonadal sex differentiation. *J Fish Biol.* 76:1879–1902.

Smith CA, Shoemaker CM, Roeszler KN, Queen J, Crews D, Sinclair AH. 2008. Cloning and expression of R-Spondin1 in different vertebrates suggests a conserved role in ovarian development. *BMC Dev Biol.* 8:72.

Spicer LJ, Sudo S, Aad PY, Wang LS, Chun SY, Ben-Shlomo I, Klein C, Hsueh AJ. 2009. The hedgehog-patched signaling pathway and function in the mammalian ovary: a novel role for hedgehog proteins in stimulating proliferation and steroidogenesis of theca cells. *Reproduction* 138:329–339.

Suster ML, Abe G, Schouw A, Kawakami K. 2011. Transposon-mediated BAC transgenesis in zebrafish. *Nat Protoc.* 6:1998–2021.

Szczepny A, Hime GR, Loveland KL. 2006. Expression of hedgehog signalling components in adult mouse testis. *Dev Dyn.* 235: 3063–3070.

Tisdall DJ, Hudson N, Smith P, McNatty KP. 1994. Localization of ovine follistatin and alpha and beta A inhibin mRNA in the sheep ovary during the oestrous cycle. *J Mol Endocrinol.* 12:181–193.

Tomizuka K, Horikoshi K, Kitada R, et al. (12 co-authors). 2008. R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. *Hum Mol Genet.* 17:1278–1291.

Uhlenhaut NH, Jakob S, Anlag K, et al. (15 co-authors). 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* 139:1130–1142.

Wang DS, Kobayashi T, Zhou LY, Paul-Prasanth B, Ijiri S, Sakai F, Okubo K, Morohashi K, Nagahama Y. 2007. Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as interacting with ad4 binding protein/steroidogenic factor 1. *Mol Endocrinol.* 21:712–725.

Wilhelm D, Palmer S, Koopman P. 2007. Sex determination and gonadal development in mammals. *Physiol Rev.* 87:1–28.

Wilkins AS. 2007. Between "design" and "bricolage": genetic networks, levels of selection, and adaptive evolution. *Proc Natl Acad Sci U S A.* 104(Suppl. 1), 8590–8596.

Wittbrodt J, Shima A, Schartl M. 2002. Medaka—a model organism from the far East. *Nat Rev Genet.* 3:53–64.

Yao HH. 2005. The pathway to femaleness: current knowledge on embryonic development of the ovary. *Mol Cell Endocrinol.* 230: 87–93.

Yao HH, Matzuk MM, Jorgez CJ, Menke DB, Page DC, Swain A, Capel B. 2004. Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. *Dev Dyn.* 230:210–215.

Zhang Y, Li F, Sun D, Liu J, Liu N, Yu Q. 2011. Molecular analysis shows differential expression of R-spondin1 in zebrafish (*Danio rerio*) gonads. *Mol Biol Rep.* 38:275–282.

Published online: September 9, 2015

Review

EMBO reports

Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators

Amaury Herpin[1,2] & Manfred Schartl[1,3,*]

Abstract

Sexual dimorphism is one of the most pervasive and diverse features of animal morphology, physiology, and behavior. Despite the generality of the phenomenon itself, the mechanisms controlling how sex is determined differ considerably among various organismic groups, have evolved repeatedly and independently, and the underlying molecular pathways can change quickly during evolution. Even within closely related groups of organisms for which the development of gonads on the morphological, histological, and cell biological level is undistinguishable, the molecular control and the regulation of the factors involved in sex determination and gonad differentiation can be substantially different. The biological meaning of the high molecular plasticity of an otherwise common developmental program is unknown. While comparative studies suggest that the downstream effectors of sex-determining pathways tend to be more stable than the triggering mechanisms at the top, it is still unclear how conserved the downstream networks are and how all components work together. After many years of stasis, when the molecular basis of sex determination was amenable only in the few classical model organisms (fly, worm, mouse), recently, sex-determining genes from several animal species have been identified and new studies have elucidated some novel regulatory interactions and biological functions of the downstream network, particularly in vertebrates. These data have considerably changed our classical perception of a simple linear developmental cascade that makes the decision for the embryo to develop as male or female, and how it evolves.

Keywords Dmrt1; ovary; SRY; testis; transcription factor
DOI 10.15252/embr.201540667 | Received 13 May 2015 | Revised 28 July 2015 | Accepted 31 July 2015

See the Glossary for abbreviations used in this article.

Introduction

Developmental cascades are generally headed by evolutionary conserved master regulators that determine the developmental fate of a cell lineage toward distinct tissues or organs during embryogenesis. In contrast, determination of the development of the reproductive organs does not follow this rule. Studies over the last decades have revealed that the gene-regulatory cascades triggering sexual differentiation from worms and flies to mammals are composed of substantially different factors. In particular, a remarkable diversity of master sex-determining genes that govern the genetic hierarchies has become apparent. On the other hand, the downstream components seemed to be evolutionarily more conserved and appear to converge on the regulation of a few central common effectors. A well-known example illustrating this paradigm is the master sex-determining gene of mammals, the *SRY* gene. A corresponding homolog has not been detected outside of therian mammals (Marsupials and Placentalia). Conversely, those genes that act downstream of *SRY* as transcription factors (*SOX9*, *DMRT1*) or signaling pathways (TGF-β/Amh, Wnt4/β-catenin, Hedgehog), and genes involved in *SRY* regulation (*SF1*, *WT1*) have homologs with a known or presumed role in gonadogenesis or gonadal differentiation in many vertebrate species, and some even in non-vertebrate deuterostomes and protostomes. These findings suggested that a central paradigm of sex determination is that "masters change, slaves remain".

This appealing global rule was quickly commonly accepted, in particular as the diversity at the top was confirmed experimentally [1–3]. Remarkably, some master sex-determining genes were recurrently identified and became the "usual suspects" for future studies in the search for master regulators (Table 1). All of these are genes, or duplicates and paralogs of genes, which were previously known to act in the regulatory network of gonad development. Much progress has also been made in understanding some of the regulatory interactions of the networks or cascades governed by the long known master sex-determining genes as well as, although to a lower extent, for the newly detected ones. We review here the current knowledge about the different molecules that have been demonstrated

1 Department Physiological Chemistry, Biocenter, University of Würzburg, Würzburg, Germany
2 INRA, UR1037 Fish Physiology and Genomics, Sex Differentiation and Oogenesis Group (SDOG), Rennes, France
3 Comprehensive Cancer Center Mainfranken, University Clinic Würzburg, Würzburg, Germany
*Corresponding author. Tel: +49 931 318 4148; E-mail: phch1@biozentrum.uni-wuerzburg.de

Glossary

Amh
Anti-Müllerian hormone
Autosome
On contrary to a sex chromosome, autosomal chromosomes are chromosomes that are not involved in primary sex determination
Csd
Complementary sex determiner
CTD
C-terminal domain
DKK1
Dickkopf-related protein 1
Dmd3
Doublesex and Mab-3 domain family member 3
DMRT1 or 3
Doublesex and Mab-3 related transcription factor 1 or 3
Dosage sensitive gene
Gene where the amount of gene product that determines the phenotype is dependent on the number of copies. Two copies are usually sufficient to establish the phenotype, while one is not (haploinsuffiency). For example, in birds two copies of the *Dmrt1* gene trigger male gonadal development, while one copy is not sufficient to make a male and then leads to female development
Dsx
Doublesex
Environmental sex determination (ESD)
When the sex of an individual is driven by different external factors including temperature, pH, social interactions (dominance, stress. . .)
Esr1
Estrogen receptor 1 is the human estrogen receptor alpha
Fem
Feminizer
FGF9
Fibroblast growth factor 9
Foxl2
Forkhead box transcription factor L2
Fru
Fruitless
Fst
Follistatin
Gene regulatory network
Set of interactions between different regulators (DNA, RNA, proteins) leading to their interdependent modulation of expression and regulation
Genotypic sex determination (GSD)
When the sex of an individual is triggered by its genotype only (can be mono or polygenic)
Gonadal maintenance
Establishment of a genetic programm in order to maintain the fate and differentiation state of the different cellular types composing the gonad, keeping either the male or female identity
Gsdf
Gonadal soma derived factor
Her-1
Hermaphroditization of XO-1
Hetero-/homo- gamety
When individuals produce gametes with either different sex chromosomes (hetero-) or similar sex chromosomes (homo-). It is refered to male heterogamety when males produce X and Y chromosome-containing gametes or female homogamety for females producing only X chromosome-containing gametes (XX-XY sex determination system, like in most mammals). For instance in birds, snakes and butterflies males are (ZZ) homogametic and females (ZW) heterogametic (ZZ-ZW sex determination system)
Heteromorphic sex chromosomes
When sexual chromosomes are morphogically distinguishable (different degrees of heteromorphism exist, depending on the age of the sex chromosomes)
Hhip
Hedgehog-interacting protein
HMG
High mobility group
irf9
Interferon regulatory factor 9
Mab-3
Male abnormal 3
masc
Masculinizer
Master sex-determining gene
A gene (not necessarily coding for a protein) responsible for the initial trigger leading to sex determination
Neofunctionalization
The process by which a gene changes its function or adds a new one by mutations that change the structure of its gene product and/or its expression pattern
Nix
Male-determining factor in the mosquito *Aedes aegypti*
NTD
N-terminal domain
piRNA
PIWI-interacting RNA
Primordial germ cells
In the embryo the precursors of the stem cells that will give rise to the germ cell lineage. During sex determination and gonad differentiation they become committed to either produce male or female germ cells as spermatogonia or oogonia, which after meiosis will become the gametes. Primordial germ cells continiously express a certain set of genes in order to maintain their unique undifferentiated/pluripotent state
Ptch
Patched
Rspo1
R-spondin 1
Sdc
Sex determination and dosage compensation defective
SdY
Sexual dimorphic on the Y chromosome
Sex chromosome
Chromosome involved in the primary sex determination. They usually harbour a master sex determining gene/trigger
Sex determination
Primary mechanism leading to the expression of the phenotypic sex. Sex determination is mostly triggered either by the genome (genotypic sex determination) or by the environment (environmental sex determination)
Sexual differentiation
Developmental consequence of the sex determination process. Regroups the events dealing with internal and external genitalia and secondary sex characters
SF1
steroidogenic factor–1
Somatic gonad
The non-germ line component of the gonad. The somatic gonad consists of mainly two characteristic cell types in female: the granulosa and theca cells of the ovary and three specific cell types in the testis: Sertoli, Leydig and peritubular myoid cells
SOX9
Sry-related HMG box 9

SRY
Sex determining region Y
STRA8
Stimulated by retinoic acid gene 8
Sxl
Sex lethal
TAD
Transactivation domain
TESCO
Testis-specific enhancer core
TGF-β
Transforming growth factor beta
Therian mammals
Non-egg-laying = marsupials and placental mammals
TRA
Transformer
Wnt
Wingless-related MMTV integration site
WT1
Wilm's tumor gene 1
Xol
XO lethal

Table 1. Master sex-determining genes in vertebrates.

Master SD gene	Organism	SD system	SD gene ancestor	SD gene generated from ancestor by	Ancestor gene function
SRY	Therian mammals	XY	Sox3	Allelic diversification	Transcription factor, required in formation of the hypothalamo–pituitary axis, functions in neuronal differentiation, expressed in developing gonads
Dmrt1	Birds	WZ	Dmrt1	Allelic diversification	Transcription factor, key role in male sex determination and differentiation
DM-Y	*Xenopus laevis*	WZ	Dmrt1	Gene duplication	Transcription factor, key role in male sex determination and differentiation
Dmrt1bY	Medaka (*Oryzias latipes, O. curvinotus*)	XY	Dmrt1	Gene duplication	Transcription factor, key role in male sex determination and differentiation
SdY	Rainbow trout (*Oncorhynchus mykiss*)	XY	Irf9	Gene duplication	Interferon response factor, no gonadal function known
GsdfY	Luzon ricefish (*Oryzias luzonensis*)	XY	Gsdf	Allelic diversification	TGF-β factor, important role in fish gonad development
Sox3Y	Indian ricefish (*Oryzias dancena*)	XY	Sox3	Allelic diversification	Transcription factor, required in formation of the hypothalamo–pituitary axis, functions in neuronal differentiation, expressed in developing gonads
amhY	Perjerrey (*Odontesthes hatcheri*)	XY	Amh	Gene duplication	Anti-Muellerian hormone, growth factor
amhr2Y	Fugu (*Takifugu rubripes*)	XY	Amh receptor 2	Allelic diversification	Type II receptor for Amh, important function in gonad development, medaka mutant shows sex reversal
Dmrt1	Chinese tongue sole (*Cynoglossus semilaevis*)	WZ	Dmrt1	Allelic diversification	Transcription factor, key role in male sex determination and differentiation
GsdfY	Sablefish *(Anoplopoma fimbria*)	XY	Gsdf	Allelic diversification	TGF-β factor, important role in fish gonad development

to determine sex in a variety of animals and what has been learned about the maintenance of the sexual identity of ovary and testis.

Master sex-determining genes: case studies from Sox and DM domain factors to emerging "unusual" suspects

From Sry down to Sox3 across vertebrates

SRY belongs to a family of transcription factors, which are characterized by an evolutionary conserved high-mobility group (HMG box) DNA-binding domain flanked by weakly conserved N- and C-terminal sequences. In mice, both, gain- and loss-of-function studies have shown that *SRY* is not only sufficient but also necessary for triggering testis development [4,5]. With the exception of only two species (the mole vole *Ellobius* [6] and the spiny rat [7]) which have probably lost the gene), *SRY* is the universal master male sex regulator of all therian mammals [8]. Cytogenetic and comparative molecular studies of mammalian sex chromosomes provided evidence that SRY most probably arose after two major events: (i) a dominant mutation of the *SOX3* allele (giving rise to the proto-Y) as well as (ii) fusion of the gene with regulatory sequences from another gene already located on the X chromosome [9] (Fig 1). Necessarily occurring before the divergence of the therian lineage, these events could be estimated to have happened ~146–166 million years ago [10,11]. Sharing an overall identity of 67% at the amino acid level and up to 90% identity when specifically considering the HMG DNA-binding domain, the X-chromosomal *SOX3*-encoded protein is most similar to SRY [12]. Consistent with this hypothesis, the expression of *SOX3* has been documented in the developing gonads of mice, chicken [13], fish [14], and frog [15]. Only the absence of *SOX3* expression

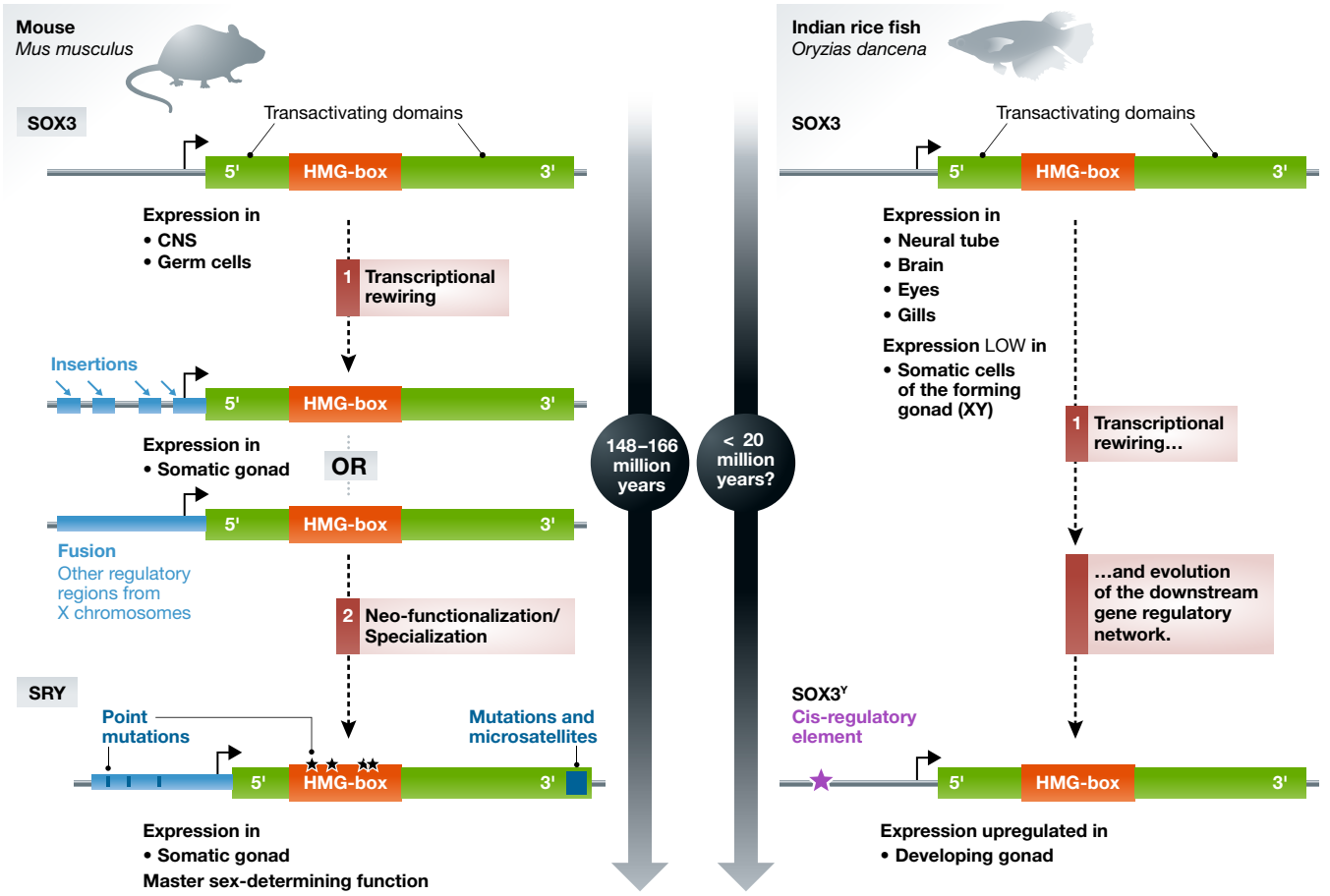


Figure 1. Independent evolution of *SOX3* genes toward a master sex-determining function in mice and Indian rice fish.
While SRY appears to be restricted to the therian mammals, evidence accumulates that SOX3 has independently been recruited as a "precursor" of master sex-determining genes also outside mammals. Hence, although not *a priori* destined to have a direct function during sex determination, common mechanisms of evolution seem to be repeatedly employed. Given that SOX3 is not generally expressed during gonadal induction or during gonadal development, the first step toward a sex-determining function is a transcriptional rewiring in order to acquire a timed pattern of expression compatible with sex determination. Such transcriptional rewiring, although not unique to SOX3 (see Dmrt1bY in medaka fish for example [56]), generally involves either fusion of the gene to new promoters or insertions of transposable elements into their pre-existing promoter, bringing in *cis*-regulatory elements compatible with the timing of gonadal induction. Interestingly and surprisingly, it seems that at least in mice and rice fish, this step alone was sufficient to endow SOX3 with a sex-determining function. Usually, the transcriptional rewiring steps seem to be accompanied by neo-functionalization or functional specialization processes. These include specialization of the protein activity itself in therian mammals (adapted from reference [20]) or more surprisingly adaptation of the downstream gene-regulatory network (target genes) in the Indian rice fish.

in the developing marsupial gonad is not consistent with a conserved role in mammalian sex determination [16,17]. Although SOX3 has no obvious primary function in sex determination, as the *Sox3* knockout mice have no gonadal phenotype [18], the clear evolutionary relationship between SOX3 and SRY raised the question whether gain-of-function point mutations may account for SOX3-induced XX male sex reversal in mice or humans. This has been shown only recently using a transgenic mouse model in which ectopic expression of SOX3 in the developing XX gonads resulted in complete XX female to male sex reversal [19]. Interestingly, the XX gonads of the transgenic hemizygous mice (Tg/+) did not only display an up-regulation of *Sox9* but also started to differentiate Sertoli cells, forming testis cords together with the appearance of a male-specific vasculature. Interestingly, using co-transfection assays it was shown that, similar to SRY, SOX3 only modestly transactivated the SOX9 testis-specific enhancer "TESCO" element [20] and synergistically interacted with steroidogenic factor-1 (SF1). Interestingly, the development of SOX3-triggered testes in XX animals was not possible in the absence of *Sox9*. In the same direction, patients displaying XX female to male sex reversal due to rearrangements of the genomic regions encompassing the regulatory sequences of *SOX3* have been reported [19]. Together, these data suggest that gain of function of SOX3 during gonadal development can in principle substitute for SRY to trigger testis development. These findings provide functional evidence supporting the long-standing hypothesis that SOX3 is the evolutionary precursor of SRY (Fig 1). It is also reasonable to postulate that rearrangements of the *SOX3* gene might be an underappreciated cause of XX female to male sex reversal in human patients [19].

While SRY appears to be specific to the therian mammals, there is accumulating evidence that *SOX3* has spawned independently other sex chromosomes outside mammals. Though being expressed in the ovary of frogs [21] without any sex-determining function determined so far, *sox3* might be involved in the switch responsible

for sex determination in the Japanese wrinkled frog (*Rana rugosa*). Members of this species are either ZW or XY depending on which side of the island they are located [22]. Curiously, the Z and X chromosomes are not only homologous but share many genes with the X chromosome of humans including the *sox3* gene. Further molecular characterization and genetic mapping could disclose the presence of a Y-specific allele for *sox3* [23,24]. So far, this is an intriguing finding, but further studies are needed to ascertain a function for *sox3* in the sex developmental decision process of the embryonic gonad. If *sox3* has such a function, then the next question would be how the different genetic systems (ZW or XY) impact on *sox3* function.

Stronger evidence comes from the Indian ricefish (*Oryzias dancena*) (Fig 1), in which the XY sex chromosome pair also shares homology with the human X, including the presence of the *sox3* gene [14]. Using positional cloning to identify the sex-determining locus, it was found that the male-specific region on the Y chromosome harbors a *cis*-regulatory DNA segment that up-regulates expression of the Y-chromosomal copy of *sox3* during gonadal development (Fig 1). Sex reversal of XX fish transgenic for the regulatory segment linked to *sox3* to become males, and fish with targeted deletion of the Y-chromosomal *sox3* gene developing as females confirmed its major role during sex determination. Furthermore, it was demonstrated that Sox3 initiated testicular differentiation by up-regulating expression of *gsdf*, a gene highly conserved in fish male sex differentiation pathways [14]. Interestingly, a BAC clone carrying the *sox3* gene of *O. dancena* was not able to induce male gonadal development in the closely related species O. *latipes*, which has a different male sex determination gene. This supports the hypothesis that the acquisition of Sox3 function as a master sex-determining gene has occurred with a concomitant change in the downstream gonadal gene-regulatory network (Fig 1). Taken together, the results provided strong evidence for the recruitment—even in distantly related species—of Sox3 into the pathway leading to male gonadal development.

SRY reveals plasticity of sex-determining mechanisms among mammals Despite substantial variations in expression profiles, structure, and amino acid sequences within mammals, the function of SRY to activate a conserved target gene—SOX9—during testis development appears to be conserved [20]. SRY directly binds to the TESCO sequence of the *SOX9* gene [20]. Once activated, the SOX9 protein initiates the differentiation of somatic precursors into Sertoli cells that will then coordinate the gonadal development toward testes [25]. In the absence of *SOX9* activation, the fetal gonad will develop toward ovaries. While the function of *SRY* as a regulator of *SOX9* appears to be conserved, the molecular details underlying transcriptional regulation of *SOX9* by *SRY* [26] are not fully known and their conservation among mammals has not been deeply investigated. Such information would be important to evaluate whether under a conserved master determiner, the subordinate network is strictly conserved as well or shows variation in its regulatory interactions.

In contrast to most known transcriptional activators, most SRY proteins that have been studied in different mammalian species do not exhibit a well-defined transactivation domain (TAD). For instance, the N- and C-terminal domains (NTD and CTD) flanking the evolutionary conserved DNA-binding domain of human SRY are poorly preserved and do not seem to display any intrinsic transactivation activity [27]. Hence, it is assumed that the transcriptional activation of the human *SOX9* gene by SRY is possible only after the recruitment of a transactivating protein partner through its NTD and/or CTD sequences [28]. However, mouse SRY does not only lack the NTD but also displays an unusual CTD made of a bridge domain together with a poly-glutamine (polyQ) tract encoded by a CAG-repeat microsatellite [27]. It has recently been shown that this poly-glutamine domain does not only prevent mouse SRY from proteasomal degradation, but additionally functions as a bona fide TAD. Due to the fact that it allows the direct transcriptional induction of Sox9, this poly-Q domain plays a central role for the male-determining function of SRY *in vivo* [27]. Such data suggest that during evolution, mouse SRY has gained a functional unit, which is absent in other mammals [27]. Given such important transactivating properties for that poly-Q CTD in mice, it is puzzling that SRY proteins from either human or goat lacking a TAD are able to induce testicular development in transgenic XX mice embryos [29,30]. It appears reasonable to consider that both human and goat SRY proteins are able to bind to the highly conserved mouse TESCO target sequence using their respective DNA-binding HMG boxes. For the activation of *SOX9* transcription, it is assumed that transactivation is then mediated after the recruitment of a third TAD-containing protein partner. It can be further hypothesized that acquisition of a poly-glutamine stretch after insertion of a CAG microsatellite in a rodent ancestor made the recruitment of a transactivating partner unnecessary. Consequently, it is assumed that mouse SRY's ability to employ such a transactivating partner was lost during evolution. This assumption is supported by the observation that the acquisition of the poly-glutamine stretch is concomitant with an increase of variation in different parts of the SRY protein. These include the loss of the NTD as well as accumulation of deleterious amino acid substitutions in the HMG box [31]. Though no longer required, the third partner protein—probably a pleiotropic effector—may still be expressed at the sex determination stage. It would then potentially enable human and goat SRYs to trigger male gonadal development when expressed in transgenic mice. This reveals an unanticipated level of plasticity of the molecular mechanisms in the implementation of the primary sex-determining signal even among mammals. Identification of such putative partners of SRY may help in understanding human primary sex reversal pathologies, which are not explained by alterations in the known players of male sexual development [32].

Roles of DM domain factors in sex determination, differentiation, and gonadal maintenance

DMRT1, wherever you look Among the evolutionary conserved downstream effector genes of genetic sex-determining cascades, the DMRT gene family holds an outstanding position. This family is involved in sexual development of organisms as phylogenetically diverse as mammals, birds, fish, frogs, flies, worms, and corals [33–38] (Figs 2 and 3). Characterized by a highly conserved DNA-binding core motif—known as the DM (Doublesex and Mab-3) domain—, DMRT proteins act as transcription factors. Initially described to be involved in sex determination in worms and flies, they have been shown to regulate diverse aspects of somatic sexual dimorphism in these organisms. The ability to functionally

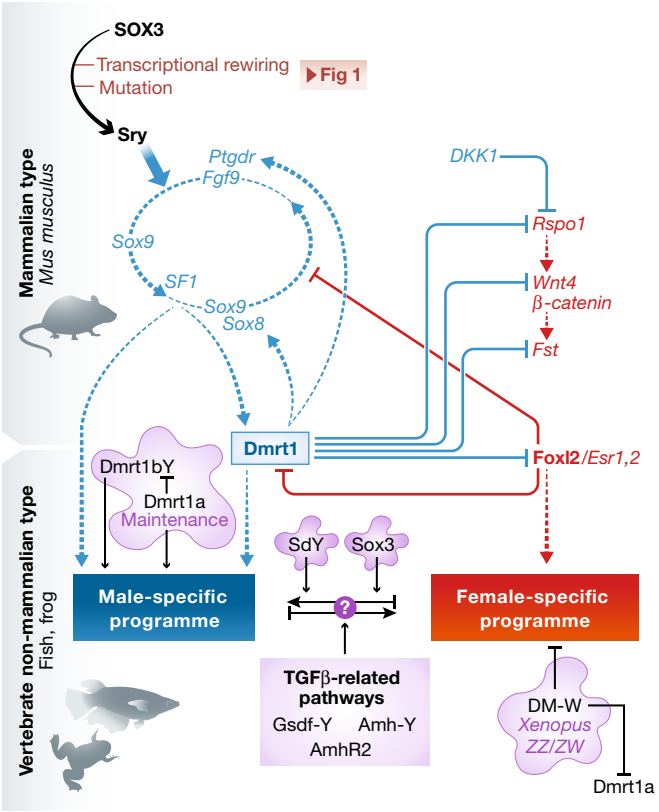


Figure 2. Gene-regulatory network of gonadal sex induction and maintenance in vertebrates.
Schematic representation of main interactions within the regulatory network. In gonadal fate determination of mammals, Sry initiates activation of the male pathway (blue) through up-regulation of Sox9. Dmrt1 is not only important for keeping the male pathway on but also in suppressing the two female networks (red). These two female networks involve Foxl2 as well as the Wnt/β-catenin signaling pathways. Maintenance of gonadal identity in the differentiated gonads is a result of the cross-inhibition activities of Dmrt1 and Foxl2. A critical equilibrium between these conflicting pathways underlies the bipotentiality of the gonadal somatic cells. Tipping the balance into one direction or the other will regulate the gonadal fate as a consequence of the activation of the male or female pathways. Solid lines define negative regulations. Dashed lines designate positive regulations. Beside the *Sry* ancestor *Sox3* and *Dmrt1*, other genes (pink) can become the master sex-determining genes by similarly impacting on the seesaw between the male and female programme.

substitute for each other across species led to the picture that sex determination cascades might—at least partially—rely on preserved molecules and pathways [37] (Figs 2 and 3; Table 1). Consistent with this, many of the DMRT homologs so far characterized among metazoans have been shown to be predominantly expressed during the development of the primordial gonad [35]. Interestingly, DM domain genes have also recently been described to be primarily involved in gonadal differentiation of the male flatworm (*Schmidtea mediterranea*) [39]. Similarly, in the water flea *Daphnia magna*, a crustacean with environmental sex determination, DMRT homologs have been found to trigger the switch in male versus female development of many dimorphic structures [40]. Thus, this widespread gene family appears to be directly involved in sexual development in all major animal groups. Nevertheless, DM domain factors were long considered as one of the underdogs of sexual determination because of their recurrent subordinate role in the cascade. A deeper interest in the field of sex determination for this group of genes only came with the discovery of a *dmrt1* homolog located on the Y chromosome of the medakafish (*Oryzias latipes*). Resulting from a gene duplication of the autosomal *dmrt1a* gene, it was designated *dmrt1bY* [41] or *dmy* [42]. It is the only functional gene in the Y-specific region of the sex chromosome, and it was shown to be not only necessary but also sufficient for triggering male development (see also Fig 2).

In humans, haploinsufficiency of the genomic region that includes *DMRT1* and its paralogs *DMRT2* and *DMRT3* leads to XY male to female sex reversal [43]. This suggested that the *DMRT1* gene is an important dosage-sensitive regulator of male development in vertebrates. In chicken and other avian species and in a fish, the smooth tongue sole (*Cynoglossus semilaevis*) [44]), *DMRT1* is located on the Z chromosome, but absent from W, and shows the expected expression pattern for a dosage-dependent male sex-determining gene of birds [45] and flatfish. In chicken, it was demonstrated through RNA interference experiments that *DMRT1* is indeed required for male gonad development [45]. While in these organisms *DMRT1* acts as a dosage-dependent male determiner, in *Xenopus laevis*, a duplicated copy of *dmrt1* on the W, which lacks the dimerization domain, appears to fulfill its function as a dominant-negative version. It is proposed to interfere with the transcriptional activation of the target genes of Dmrt1 and thus acts as a suppressor of male development [46].

Remarkably, all these *DMRT1* genes have acquired their new roles as master sex determination genes through different mechanisms: via gene duplication and translocation in medaka, duplication, translocation and truncation in *Xenopus*, or loss of function of the W allele in birds or tongue sole (Table 1).

In mice, it is apparent that *Dmrt1* is not required for male primary sex determination since newborn Dmrt1 mutants are males with testes [36]. However, *Dmrt1* is required for male gonadal differentiation of somatic cells and germ cells [47–49]. This is a parallel situation to mammalian *Foxl2* [50], which plays a conserved role in ovarian development but in mouse (opposed to some other mammals, including human and goat [51]) is not required for initiation of female development (see [52] for review). Targeted deletion of mouse *Dmrt1* and also of the autosomal *dmrt1a* of medaka, which is not involved in primary male sex determination, have revealed a major role in male gonad maintenance: when *Dmrt1* is lost, even in adults, this triggers sexual cell-fate reprogramming, in which male Sertoli cells trans-differentiate into their female counterparts, the granulosa cells [49]. This is accompanied by testicular reorganization toward a more ovarian morphology [49]. Ectopic DMRT1 expression in the ovary silenced the female sex-maintenance gene *Foxl2* and reprogrammed juvenile and adult granulosa cells into Sertoli-like cells, triggering formation of structures, which resemble male seminiferous tubules [53]. In the same direction, deletion of the *dmrt1* gene in medaka resulted in transition of the developing testis to ovary [54]. Hence, DMRT1's range of action is not limited to function in initiating the male gonadal phenotype during early development but also accounts for the livelong active repression of the two "anti-testis" pathways of FOXL2 and WNT4/β-catenin [49], and can do so even in the absence of the testis-determining genes *SOX8* and *SOX9* (Fig 2). Additionally, mRNA profiling revealed that DMRT1 activates many testicular genes and

down-regulates ovarian genes [53]. Interestingly, transient expression of DMRT1 has also been reported in the fetal gonad of both sexes. The involvement in the regulation of germ cell development in testes and ovaries indicates that DMRT1 has different functions in males and females [55].

DMRT1 is required in female germ cells for entry into meiotic prophase, and in male germ cells for the control of mitotic arrest until birth [55]. Control of the decision to enter meiosis versus mitotic arrest is mediated by the ability of DMRT1 to selectively modulate retinoic acid signaling through context-dependent regulation of STRA8. DMRT1, for example, directly represses STRA8 transcription during testicular differentiation [55]. Thus, a picture emerges where DMRT1 controls a regulatory network that on the one hand can drive sexual fate and on the other hand can maintain the program of sexually differentiated cells, depending on the cellular context.

DMRT1, a jack-of-all-trade From studies in mouse and medaka [49,53,54,56,57], it is emerging that *DMRT1* holds a key position as the master switch or gatekeeper controlling the cell fate of the somatic cells of the gonads in female and male [33,34,53,58,59]. If this is so, then one could ask, why such a complex regulatory network upstream of *DMRT1* would be necessary to flip the switch, because numerous examples indicate that DMRT1 can do it on its own as for instance in birds, *Xenopus* and medaka [41,42,45,46]. *DMRT1* orthologs in these species appear to have undergone mutational events causing either loss or gain of function. Such altered DMRT1 activity may have favored evolutionary transitions leading to new genetic sex determination systems (see [59] for review). The ability of DMRT1 to toggle Sertoli/granulosa cell fate supports the hypothesis that loss- or gain-of-function mutations in *DMRT1* can elevate it into a master sex-determining role. Such mutations would help to promote changes between genetic sex determination mechanisms that are commonly observed among vertebrates.

DMRT1 is one of the sex determination network genes that appears more often also as master regulator (Table 1). It can be hypothesized that its strategic position at the interface of sex determination and the process of sex-specific gonadal differentiation, integrating a developmental fate decision with activation of organ differentiation programmes (Fig 2), made *DMRT1* suitable to be selected either as new controller at the top or at least for being one of the few key genes to be regulated.

Emerging suspects from gonadal TGF-β signaling

The anti-Müllerian hormone (Amh) is a growth factor from the TGF-β family and plays a major role in mammals for the degradation of the Müllerian duct-forming part of the female reproductive tract in male embryos. It is not required for mouse testis development. However, in non-mammalian vertebrates, it appears to play a central role in testis formation. For instance, in chicken embryonic gonads, AMH is expressed much higher in males and is predicted to be responsible for organizing the early testis in birds [60]. In the medaka *hotei* mutant, Amh signaling is disrupted by a mutation in the type II receptor for Amh. As a consequence, a male to female sex reversal with an over-proliferation of germline stem cells occurs [61].

Although being clearly a subordinate member of the sex regulatory network in mammals and at least in those species that make use of *DMRT1* as master regulator of male development, the Amh/Amh-receptor system has, like *DMRT1*, sometimes made it to the top (Table 1). In the pejerrey, a freshwater fish species from Patagonia, a duplicated version of the *amh* gene became the male sex-determining gene on the Y chromosome [62], reminiscent of the situation for *dmrt1* in medaka fish. In the pufferfish, *Fugu rubripes*, the receptor for Amh exists in two versions that differ by one amino acid (H384D) in the kinase domain [63]. The 384^{His} allele is a Fugu-specific (conserved in several other pufferfishes) mutation that confers lower activity to the receptor and is encoded on the X chromosome [63]. Thus, a quantitative difference in Amh signal transduction in females, which are homozygous for the mutant, versus males, which have kept one allele of the wild-type receptor on their Y, is responsible for male development [63]. Like in the medaka *hotei* mutant [61], low signaling from the receptor is connected to feminization of the gonad.

Gonadal soma-derived factor (Gsdf) is another growth factor from the TGF-β family that is closely related to Amh. It is only found in fish, and its biochemical function is not well studied. It is assumed to have a role in male gonad development due to its exclusive expression in the early differentiating testis of all fish looked at so far [64–68]. Despite its proposed role in the downstream regulatory network, *gsdf* has made it up to the top in *Oryzias luzonensis* [69] a sister species to medaka, and most likely also in the sablefish [70].

Taken together, it appears that certain genes, which are members of the regulatory network, namely *sox3*, *dmrt1*, and TGF-β signaling components, can become the master sex-determining gene independently again and again, while other important components of the sex-determining pathways have not appeared as masters so far (Fig 2 and Table 1). Whether we just have to wait for the analyses of primary genes for sexual development in more species, in order to put genes like *foxl2*, *sox9*, *sox8*, *wnt4*, etc., on the list of usual suspects, or whether there is a biological reason that makes some genes more prone to become the top regulator, is currently unsolved. We could imagine that some genes remain "too difficult to recruit" as master regulators, for instance if they have also non-reproductive but vital functions in other organs. In such case, interferences between a duplicated new master gene and its homolog may not be tolerated, except for the case that the neo-gene would have an appropriate gonad-specific regulation as soon as the founder event occurs. Many of those genes that did not appear as master sex determiners so far indeed have important functions in other tissues and organs.

Recurrent actors in invertebrate sex determination

The invertebrate ancestors of DMRT1 DM domain-containing genes have been shown to be primarily involved in gonad differentiation in a flatworm [39] and to direct male versus female development of dimorphic structures in water flea [40]. Interestingly, this functional convergence is common among insects (see [3,71–73] for reviews). In *Drosophila*, the initial trigger of sex is dependent on the ratio of the number of X chromosomes versus the haploid autosome complement (X:A). In the female situation, an X:A ratio of one will enable the transcription of the *Sex lethal* gene (*Sxl*), a splicing regulator. The SXL protein will then promote the female-specific splicing of *Transformer* (Tra), a direct downstream target, and lead to the production of functional TRA proteins. Similarly, a complex made of TRA and TRA-2 proteins will then favor the female-specific

splicing of the *Doublesex* (*Dsx*, the *Dmrt1* homolog) gene transcripts. This results in the production of the female-type DSX protein DSX^F, which initiates up-regulation of the downstream gene-regulatory network for female development. In males, an X:A ratio of 0.5 will prevent the production of the SXL protein and, by default, results in the production of the male-specific splice form of the *Tra* gene. This splice variant translates into a non-functional protein due to a premature stop codon. In the absence of TRA, by default the male-specific splice form from the *Dsx* gene will be produced. The male-type DSX protein DSX^M will then orchestrate the downstream gene-regulatory network for male development [71,74] (Fig 3). Orthologs of *Drosophila dsx* have been identified and studied in a large number of insects [75–77]. Mediation of alternative sex-specific splicing of *dsx* by TRA and TRA2 is also widely conserved in insects although variations of the sex determination systems occur [3], suggesting that different molecular mechanisms involving splicing activators or repressors are employed to preferentially generate sex-specific variants of *dsx* mRNA [78].

Despite considerable efforts, similar sex-specific alternative splicing events in the molecular regulation of sex determination of vertebrates have not been shown. Conceptually similar is the fact that DSX translates the sexual determination process of a cascade of alternative splicing events into the transcriptional control of a large number of sex-specific effector genes. Similarly, DMRT1 in vertebrates appears to hold such a "translational" function at the interface where a fate-determining signal is put into effect at the level of sex-specific somatic cell differentiation (Figs 2 and 3).

In invertebrates, the homologs of vertebrate *Dmrt1* (e.g. *Dsx* in *Drosophila* and *Mab3* in *C. elegans*) are typical downstream factors of sex determination and so far, it is not reported that a DM domain

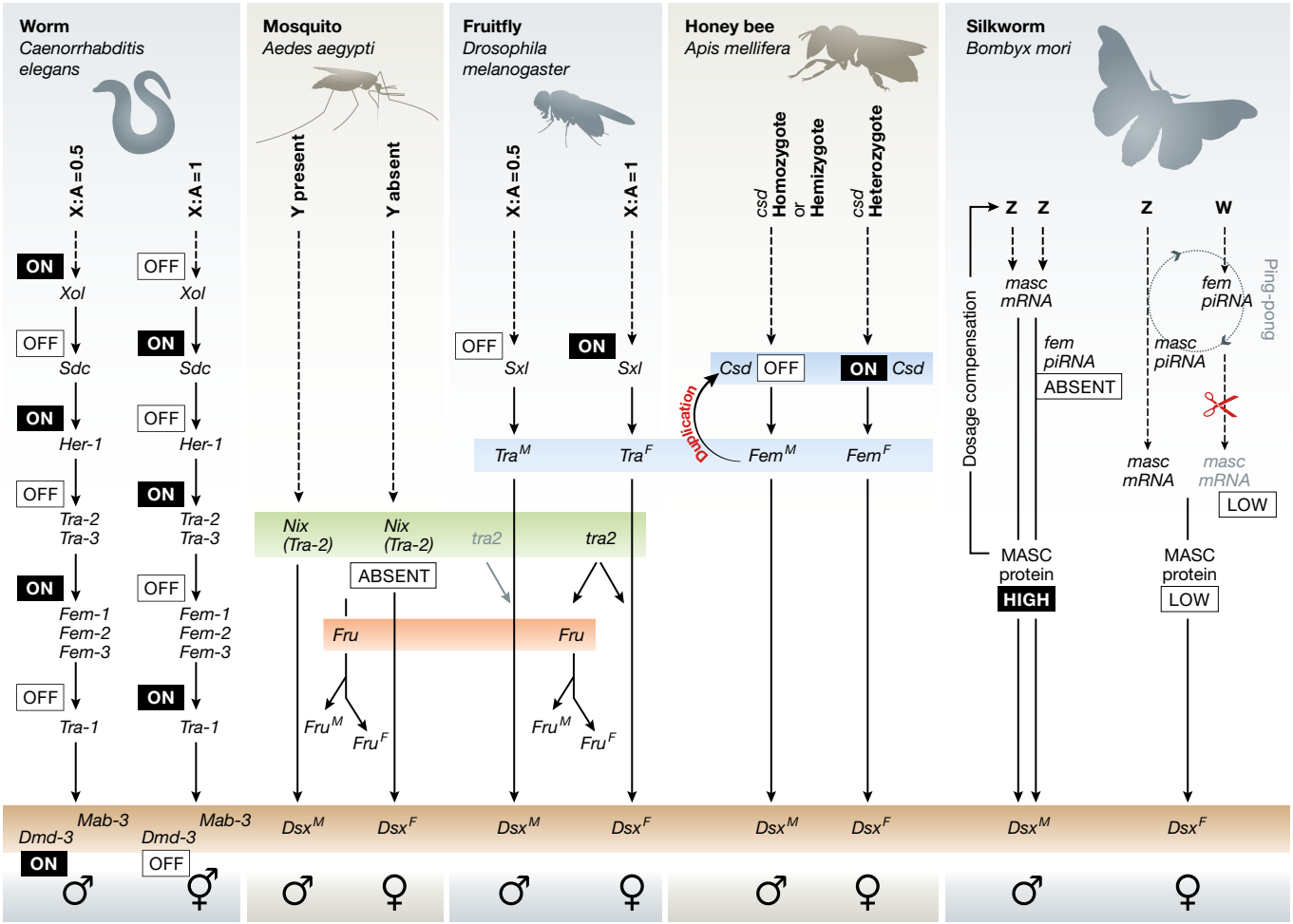


Figure 3. Sex-determining cascades in *C. elegans* and some insects.
Molecular and genetic pathways leading to the formation of the gonad in the worm *C. elegans*, the mosquito *Aedes aegypti*, the fly *D. melanogaster*, the honey bee *A. mellifera*, and the silkworm *B. mori*. Conservation of the *Dsx*, *Mab-3* and *Dmrt1*, *Tra*-like, (*Tra-2*), or *Fru* homologs is designated with either pale brown, pale blue, pale green, or pale orange boxes, respectively. *Tra-(1, 2* or *3)* of *C. elegans* are not phylogenetically related to *Tra* of *Drosophila*. *Fem-(1, 2* or *3)* of *C. elegans* are not phylogenetically related to *fem* of *Bombyx mori*. In *C. elegans* and *D. melanogaster*, a ratio between X chromosomes and autosomes determines the sex. This leads to the on/off state of *Xol* or *Sxl*, respectively. Heterozygosity turns on *Csd* in the honeybee *Apis melifera*, leading to female development, and hemizygosity or homozygosity leaves *Csd* unexpressed and produces a drone. In the mosquito (*Aedes aegypti*), sex determination is triggered by a dominant male determiner (*Nix*). *Nix* is a distant homolog of the splicing factor *Tra-2* of *Drosophila* and likely regulates the sex-specific splicing of *Fru* and *Dsx*. Sex in the silkworm *Bombyx mori* is controlled via a ZW sex chromosome system. Produced only from the sex-determining locus on the W, the piRNAs suppress the male sex-determining factor MASC.

gene has made it up to the top in any invertebrate species [3]. But like in vertebrates, genes that are known as downstream members in one species can also usurpate a position as an initial genetic trigger in another species [3]. In insects, paralogs of the gene *tra* that is a well-studied component of the sex determination cascade in *Drosophila*, evolved as the master sex-determining switch gene in the housefly (*Musca domestica*), a wasp (*Nasonia vitripennis*), and the honeybee (*Apis mellifera*) [72,79,80]. In this regard, studies about complementary sex determination in the honeybee give exciting insights into how molecular diversity of regulatory pathways can evolve [81,82], as discussed in more detail below.

Complementary sex determination in honeybees uses a conserved module from chromosomal sex determination Genetic sex determination in the honeybee does not depend on the presence of hetero- or homomorphic sex chromosomes with different genetic compositions but rather follows a haplodiploid mode. Males develop from haploid unfertilized eggs, while diploid fertilized eggs develop into females. Hence, male or female sexual development occurs as the result of a signal originating from either a single or two different alleles from one gene, called *complementary sex determiner* (*Csd*) (Fig 3). Consequently, maleness or femaleness is determined by either homo-, hemi-, or heterozygosity of the *Csd* locus. The *Csd* gene products belong to an arginine-/serine-rich protein family. Interestingly, the C-terminal end of Csd also displays high similarity with the TRA protein, an essential downstream genetic factor of the sex-determining pathway in *Drosophila* ([81] and Fig 3).

Intriguingly and in contrast to the situation in *Drosophila* with *Tra* and other downstream genes (see Fig 3), neither transcriptional nor splicing variations of the *Csd* gene could be detected as sex-specific triggers. It is currently presumed that the regulation of the downstream regulatory network is mediated by the tendency of the CSD proteins to form heterodimers. Interestingly, the sex determination locus of the honeybee harbors a second gene also required for sex determination: *feminizer* (*Fem*) [82]. Further, phylogenetic studies revealed that *Fem*—as *Csd*—is also a close homolog of the *Tra* gene from *Drosophila*. It has been shown that *Csd* arose after duplication of the *Fem* gene 10–70 million years ago while the honeybee lineage was specifying. Knockdown experiments using RNA interference (RNAi) of either *Csd* or *Fem* resulted in female to male phenotypic sex reversions, implying that both factors are required for sex determination in the honeybee downstream of sex-specific splicing of the *Fem* gene by the CSD protein ([81,83] and [3] for review).

The situation in the honeybee resembles the roles of *dmrt1* in medaka and *Xenopus* and of *amh* in the pejerrey: A highly conserved downstream component of the network underwent a gene duplication, and then, one of the duplicates evolved a new function at the top of the cascade (Figs 2 and 3).

Another usurpator in mosquito? In the yellow fever mosquito, *Aedes aegypti*, like *Drosophila* a member of the order Diptera, sex is dependent on the presence or absence of a Y chromosome. Recent work has uncovered the molecular nature of the male-determining gene [84]. Intriguingly, this gene, called *Nix*, shows some sequence similarity to the *Tra-2* gene. This gene in *Drosophila melanogaster* is a downstream member of the sex determination cascade. Further downstream in the fruitfly cascade are the *Fru* and *Dsx* genes, and also in *Aedes aegypti*, both genes are regulated by the *Tra-2* homolog *Nix* (Fig 3). It is tempting to propose that in the mosquito, we have another example of a subordinate sex determination gene that has made it to the top.

The "unusual" suspects

All the above discussed cases of turnovers and novel master sex determiners include genes that have been previously known as components of downstreams sex determination networks, for example, from mouse, human, *Drosphila*, and *C. elegans*. Unexpectedly, there are two recent reports on sex-determining genes which were neither known nor suspected to be involved in the molecular regulation of this process.

An immune-related gene evolved into the master sex-determining gene in rainbow trout In the rainbow trout *Oncorhynchus mykiss*, a gene expressed only in the testis, predominantly during testicular differentiation, was recently characterized [85]. Localized at the sex-determining locus, this gene was named *sdY* for sexual dimorphic on the Y chromosome. Astonishingly and unlike other master sex-determining genes characterized so far, *sdY* has no homology with any known gene in sex determination pathways but with an immunity-related gene, the interferon regulatory factor *irf9* [85]. *SdY* arose by duplication and truncation of the autosomal *irf9* gene (Table 1). It lost the DNA-binding domain but preserved its protein–protein interaction domain. So far, the molecular mechanism through which SdY triggers male gonad development is unknown.

A single female-specific piRNA is the primary determiner of sex in the silk worm Sex in the silkworm *Bombyx mori* and all butterflies is determined by a ZW sex chromosome system. The W chromosome lacks any protein-coding genes but consists predominantly of transposons and non-coding RNAs. The only transcripts produced from the sex-determining region on the W are PIWI-interacting RNAs (piRNAs). After deep sequencing and isolation of dimorphically expressed RNAs, the *Fem* piRNA (*Fem* standing for "feminizing factor") was shown to be specifically expressed in females at all stages of development [86]. Furthermore, *Fem* piRNA targets and cleaves the *Masculinizer* (*Masc*) RNA molecule transcribed from a gene located on the Z chromosome. Interestingly, MASC, a CCCH-type zinc finger protein, favors male-specific splicing of *Bm-dsx*, leading to male development [86]. Hence, in ZW embryos, *Masc* RNA level is down-regulated by *fem* piRNAs, inhibiting male development. By default, female-specific splicing of *Bm-dsx* then occurs, triggering female development [86] (Fig 3). Interestingly, genetic inhibition of *Masc* resulted in the premature death of ZZ embryos before they hatched. In light of this observation, it was shown that the MASC protein is necessary for dosage compensation in order to lower Z gene transcription in ZZ embryos to the same level as in ZW embryos [86]. Whether or not this sex determination pathway is conserved across all lepidopterans remains to be explored, but coupling two important mechanisms namely sex determination and dosage compensation within the same genetic pathway and additionally distributing their genes onto the sex chromosomes should strongly promote evolutionary conservation.

SdY from rainbow trout and *Fem* piRNA are paradigms showing that unrelated genes are able to acquire *de novo* sex-determining functions. It can, however, not be excluded that they are representing

factors of the sex determination regulatory network that have been overlooked so far.

Plasticity of the downstream sex determination regulatory network

What happens when "masters change"? The slogan "slaves remain" could imply that not much happens downstream of the changing master sex determiner. However, the findings on the diversity of SRY structure and its way to act as a transcriptional activator (see above) indicate that even under the same master gene, the regulatory interactions of the network undergo changes and that biology is not that simple.

In *Drosophila*, it has been shown that at the very downstream end of the sex determination, cascade pathways diverge by cooption of new effector genes [73] explaining the divergence of secondary sex characters between species. In vertebrates, some transcription factors like DMRT1, FOXL2, SOX9, and components of pathways such as Rspo1/Wnt/Fst or Hedgehog of the gonadal gene-regulatory network are well conserved on the DNA sequence level; however, their specific functions, regulations, and interplays can be substantially different. In medaka, down-regulation of the Hedgehog pathway by Dmrt1bY was shown [87]: Transcription of the Hedgehog receptor Ptch-2 in medaka testis is down-regulated by Dmrt1bY/Dmrt1a, while the antagonist Hhip is up-regulated [87]. The Hedgehog pathway is usually up-regulated by DMRT1 in mammals. It appears that despite its necessity for mammalian testis induction and development and later on in regulating Leydig and myoid cell function [88–90], the Hedgehog pathway might not only be dispensable during medaka male gonadogenesis and maintenance, but needs to be suppressed by *DMRT1* genes.

For R-spondin 1 (Rspo1), preferential ovarian expression is generally described. However, such strict female dimorphism was not observed in zebrafish [91], where the gene is also expressed in adult testes. Here, Rspo1 has a crucial role in testis cell proliferation [92] and it has further been shown to be involved in skin and mammary gland differentiation in mammals [93]. Follistatin (*Fst*) expression in the mouse co-localizes with *Foxl2* in the ovary [94], but in rat, it is expressed very broadly in germ and somatic cells of the testis [95]. Sparse expression of *fst* was also noted in the interstitial cells of the medaka testis, together with an up-regulation of *fst* expression *in vitro* after transfection of *dmrt1a* [87].

SOX9 has been shown to be expressed in the developing testes of all vertebrate embryos examined so far (see [60] for review). However, whereas SOX9 is upstream of AMH in mammals, the reverse applies in birds, and in medaka, Sox9 even appears to be not involved in primary sex determination at all [96,97]. In mammals, the current understanding is that SRY acts together with SF1 to activate *SOX9*, while in return, *SRY* is turned off by SOX9. *SOX9* further maintains its expression in an autoregulatory loop. SF1 is still required, but SRY becomes dispensable later during development [20]. In non-mammalian vertebrates, Sox9 activation must then rely on other factors than Sry. Intuitively, one could think that DM domain genes might have taken over. However, in chicken embryos, *DMRT1* expression is occurring at least 2 days before that of *SOX9* [60], implying that other genes mediating the DMRT1 signal to *SOX9* are involved. In medaka *sox9b*, the homolog of tetrapod *sox9* genes is rather involved in germ cell function than gonad determination although being expressed in the somatic part of the primordial gonads [96]. In addition, while in mammals, SOX9 activates the expression of *FGF9* [98], the gene does not exhibit any sexually dimorphic expression in chicken [60] and has even been lost in fish [99]. It is obvious that the gonadal function of *SOX9* underwent several changes during vertebrate evolution.

Genetic networks are indeed more complex than a straight top-down scenario. We have to add now that the differences in gene expression do not only reflect differences in cell biology and morphogenesis of the gonads but definitively are also the consequences of changes in the initial trigger for activating the network. That master sex-determining genes are prone to regulatory putsches in order to acquire an upstream position might only be possible because of the flexibility of the downstream gene-regulatory network. Hence, while Graham proposed a few years back that "Masters change, slaves remain" [1], it is now time to change this paradigm: "When masters change, some slaves remain, others are dismissed or acquire new tasks, and new ones can be hired".

Conclusions and perspectives

The variability and plasticity of the mechanisms that govern the development of the gonads is unmet by any other organ systems or tissues. While for instance the *Pax6* gene that is a master regulator of mammalian eye development is highly conserved (ectopic expression of human *PAX6* is able to induce eye development in *Drosophila* [100]), the downstream components of this cascade are not conserved (the induced eye is a typical composite insect eye). Surprisingly, it appears to be the other way round for sex determination genes. The evolution of genetic interactions in the sex-determining pathways and cascades is characterized by a relative conservation at the bottom and an apparent diversity at the top. This was explained in a classical hypothesis by A. Wilkins with an evolutionary scenario in which these hierarchies during evolution build up from a common downstream component (Sox or DM domain factors for instance), which acquires new upstream regulators. Those new additions would naturally vary in different evolutionary lineages [101]. Recent studies on the molecular identification of such upstream regulators and the downstream regulatory network, some of which provided the backbone for this review, brought new insights into how sexual development is regulated in different organisms, and how new sex determiners have evolved.

The "bottom-up hypothesis" formulated by Wilkins has to be revisited now taken into account the discoveries of the new master regulators. It seems that the master regulator/switch is not necessarily elected from the existing cascade usurping the top position but could be equally recruited from outside to accomplish a new sex-determining function after neo-functionalization. We also have to modify the hypothesis as we now know that in vertebrates, unlike in invertebrates, sex determination is not brought about by a simple linear cascade, but by a complex network of multiple regulatory interactions. Such a network might offer multiple opportunities where a newly added factor can trigger the outcome of the network signal toward male or female. There is also evidence accumulating

that regulatory cascades can become shorter, rather than being topped up, when a new sex determiner appears, for example, in honeybees [72,102].

Gonad development appears to cope well with such changes of primary triggers as the many examples of different master sex regulators show, which finally all guarantee the developmental switch to either a testis or ovary. An intriguing situation has been recently reported for zebrafish, where the laboratory strains used worldwide have all lost their original sex-determining chromosome, but still produce normal males or females [103]. New upstream sex determiners appear to evolve quickly in those domesticated strains—similar to a situation in the other small aquarium fish model, the medaka [104]—which might take care in the future of the current sex bias observed at present for many laboratory strains. These are instances of "evolution in action," which offer prospects to observe in the laboratory how new sex determiners evolve and to obtain insights into the underlying molecular mechanisms. Certainly, we also need more information from different species about their master sex-determining gene and how it acts on the downstream regulatory network to obtain a reasonable understanding of the variety of sex-determining mechanisms.

Somehow unexpected are the accumulating findings that also the downstream network is not as strictly conserved as the "masters change, slaves remain" paradigm was imposing. Whether these differences in the expression pattern and function are related to specific adaptations of varying reproductive biology is a challenging question for the future. On the other hand, such changes may be due to the impact of the new upstream regulator. Intriguingly, even in a setting of the same master sex-determining genes, intricate differences downstream can be found, as seen for SRY in different mammals. It has also been argued that genetic networks, including sex determination, in general can change randomly without necessarily impacting on the final phenotype and thus evolve neutrally (see Sidebar A). Again, we need more details on the molecular biology of the sex-determining networks from different organisms; for instance, on a comparative basis from birds, *Xenopus* and those fish that all use *dmrt1* as their common master sex-determining gene.

Unexpectedly, it turned out that sex determination is not only needed as the molecular switch for the undifferentiated gonad primordium to develop either as testis or ovary, but that the sexual identity of the gonadal soma needs to be maintained as long as the organ has to provide its function(s). In vertebrates, two genes that appear to have a more downstream function in the determination network of the embryo are the top players here: *DMRT1* and *FOXL2*. The dichotomous developmental potency of the gonadal soma is apparently kept throughout the entire life. The reason for this is unknown. In particular among fishes, hermaphroditic species are common. Those fish can switch during their reproductive life from one sex to the other. Whether these organisms have found a way to make a controlled use of the livelong plasticity of the gonad or whether the plasticity seen even in the mammalian gonads is a relic of an evolutionary past are just two questions that emerge from those new findings.

The recent progresses reviewed here have considerably increased our understanding of the diverse molecular mechanisms underlying the amazing variation and plasticity of sexual development, and we might so far just only see the tip of the iceberg.

Sidebar A: Evolutionary concepts for the diversity of sex determination mechanisms

Sex determination is a very basal and ubiquitous developmental process, and the fact that it is so variable even between closely related organisms poses many fascinating questions. Molecular biologists are most interested to understand how these different mechanisms work, what factors are involved, upstream and downstream, and how they are regulated to bring about the amazing plasticity of the respective genetic cascades and networks. These are the so-called proximate causes of the observed variability. Organismic biologists focus more on the "ultimate" causes that lead to the changes from one to the other sex determination mechanism within and between certain lineages. A number of scenarios and hypotheses have been put forward to explain which evolutionary forces could favor such transitions and turnovers [105].

One explanation is that a mutation, which creates a new sex determination mechanism, gives a fitness advantage to its carriers. Then, by natural selection, this mutation will sweep through the population and take over, while the previous mechanism is lost [106]. Such new mutations could for instance alter the sex ratio, and if the ecological conditions favor such a bias, this mutation will be beneficial. As another example, a new sex determination mechanism might for instance be more efficient under certain ecological conditions, for example, works faster or is less or more susceptible to environmental influences.

If sex is determined through sex chromosomes, a common feature is the reduction of recombination around the sex-determining gene, which spreads out from there over almost the entire chromosome and finally fully arrests. As a consequence, deleterious loss-of-function mutations will accumulate in genes on the chromosomes carrying the sex locus [107]. Hence, such a chromosome will become less fit in evolutionary terms because of its mutational load, and once these disadvantages accumulate to a critical level, an emerging "younger" and less degenerated sex chromosome can take over [108].

Another hypothesis is based on linkage of sex-determining genes to other genes that favor one sex or are antagonistic to the other sex [109]. Many examples exist for such genes, which for instance are involved in gonad development or sexual dimorphism. If such a gene is closely linked to a gene that can influence the developmental decision toward male or female, the sex-determining gene will be co-selected as a hitchhiker and enjoy the fitness advantage that the linked sex beneficial or sexually antagonistic gene has under conditions of natural or sexual selection.

Rather than postulating a fitness advantage for the emerging novel sex determination mechanism, it is also considered that neutral or non-adaptive processes of genetic drift, mutation, and recombination can be instrumental. Such hypotheses are based on an analysis by M. Lynch how in general genetic networks can evolve [110]. He pointed out that only the final gene product of a genetic network or cascade produces a phenotype, which is exposed to selection. Thus, many changes in the upstream system can occur without necessarily altering the finally expressed phenotype. These changes can become fixed in a population by random genetic drift. As a result, the regulatory network has changed, but the phenotype will be constant. Such considerations were then applied to the genetic cascades and networks that govern sex determination [102]. Indeed, the final outcomes of the sex determination process are morphologically and functionally surprisingly similar in related groups of organisms, which have very different master sex regulators [111].

For all of these theoretical explanations, which appear to be to a certain extent opposing or even contradictory, examples to support them can be found. A single one obviously cannot explain all the different cases of sex determination systems and the multitude of turnovers and transitions. Rather than being alternatives, they may be complementary to explain the biodiversity of mechanisms that make the undifferentiated gonad anlage of an embryo to develop toward testis or ovary. To further our understanding of the trajectories that lead to the evolution of diverse mechanisms, we need not only detailed molecular knowledge about the proximate causes of such diversity but also more information about the ecology and population genetics under which they occur.

Sidebar B: In need of answers

(i) What are the protein partners of SRY in human and goat that directly activate Sox9 expression?
(ii) Are the differences in the expression pattern and function of the genes in the downstream cascades or networks related to specific adaptations of varying reproductive biology? Or are they the result of neutral evolution and genetic drift?
(iii) Have the naturally occurring hermaphroditic species of fish found a way to make a controlled use of the livelong plasticity of the gonad? Or is the plasticity seen in the mammalian gonads a relic of an evolutionary past?
(iv) What are the evolutionary forces driving the outstanding high variability of molecular and genetic mechanisms of sex determination? Is this all due to stochastic variation? Or is there a global (so far unknown?) reason? Or do all evolutionary mechanisms postulated so far cooperate, with differing importance depending on the species or phylogenetic lineage?
(v) Are Sox3 and Irf9 in vertebrates and *Fem* piRNA components of the downstream sex determination cascades or networks that have been overlooked so far?
(vi) Why do some members of the regulatory networks of sexual development frequently become master sex-determining genes while others never appear at the top position?

Acknowledgements
We thank Yann Guiguen (INRA, LPGP, Rennes), Mateus Adolfi, Sylvain Bertho, and Alvaro Roco (Biocenter Würzburg) for reading the manuscript and helpful discussions and Monika Niklaus-Ruiz for help in preparing the manuscript. Work of the authors was supported by the Deutsche Forschungsgemeinschaft (Scha408/12-1, 10-1; He7135/2-1) and the ANR (ANR-13-ISV7-0005 PHYLOSEX; Crédits Incitatifs Phase 2015/Emergence to A.H.).

Conflict of interest
The authors declare that they have no conflict of interest.

References

1. Graham P, Penn JK, Schedl P (2003) Masters change, slaves remain. *BioEssays* 25: 1 – 4
2. Haag ES, Doty AV (2005) Sex determination across evolution: connecting the dots. *PLoS Biol* 3: e21
3. Herpin A, Schartl M (2008) Regulatory putsches create new ways of determining sexual development. *EMBO Rep* 9: 966 – 968
4. Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R (1991) Male development of chromosomally female mice transgenic for Sry. *Nature* 351: 117 – 121
5. Lovell-Badge R, Robertson E (1990) XY female mice resulting from a heritable mutation in the primary testis-determining gene, Tdy. *Development* 109: 635 – 646
6. Just W, Rau W, Vogel W, Akhverdian M, Fredga K, Graves JA, Lyapunova E (1995) Absence of Sry in species of the vole Ellobius. *Nat Genet* 11: 117 – 118
7. Soullier S, Hanni C, Catzeflis F, Berta P, Laudet V (1998) Male sex determination in the spiny rat *Tokudaia osimensis* (Rodentia: Muridae) is not Sry dependent. *Mamm Genome* 9: 590 – 592
8. Waters PD, Wallis MC, Marshall Graves JA (2007) Mammalian sex–Origin and evolution of the Y chromosome and SRY. *Semin Cell Dev Biol* 18: 389 – 400
9. Sato Y, Shinka T, Sakamoto K, Ewis AA, Nakahori Y (2010) The male-determining gene SRY is a hybrid of DGCR8 and SOX3, and is regulated by the transcription factor CP2. *Mol Cell Biochem* 337: 267 – 275
10. Foster JW, Graves JA (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. *Proc Natl Acad Sci USA* 91: 1927 – 1931
11. Graves JA (2006) Sex chromosome specialization and degeneration in mammals. *Cell* 124: 901 – 914
12. Stevanovic M, Lovell-Badge R, Collignon J, Goodfellow PN (1993) SOX3 is an X-linked gene related to SRY. *Hum Mol Genet* 2: 2013 – 2018
13. Caetano LC, Gennaro FG, Coelho K, Araujo FM, Vila RA, Araujo A, de Melo Bernardo A, Marcondes CR, Chuva de Sousa Lopes SM, Ramos ES (2014) Differential expression of the MHM region and of sex-determining-related genes during gonadal development in chicken embryos. *Genet Mol Res* 13: 838 – 849
14. Takehana Y, Matsuda M, Myosho T, Suster ML, Kawakami K, Shin IT, Kohara Y, Kuroki Y, Toyoda A, Fujiyama A *et al* (2014) Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish Oryzias dancena. *Nat Commun* 5: 4157
15. Oshima Y, Naruse K, Nakamura Y, Nakamura M (2009) Sox3: a transcription factor for Cyp19 expression in the frog *Rana rugosa*. *Gene* 445: 38 – 48
16. Bowles J, Schepers G, Koopman P (2000) Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* 227: 239 – 255
17. Pask AJ, Harry JL, Renfree MB, Marshall Graves JA (2000) Absence of SOX3 in the developing marsupial gonad is not consistent with a conserved role in mammalian sex determination. *Genesis* 27: 145 – 152
18. Rizzoti K, Brunelli S, Carmignac D, Thomas PQ, Robinson IC, Lovell-Badge R (2004) SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* 36: 247 – 255
19. Sutton E, Hughes J, White S, Sekido R, Tan J, Arboleda V, Rogers N, Knower K, Rowley L, Eyre H *et al* (2011) Identification of SOX3 as an XX male sex reversal gene in mice and humans. *J Clin Invest* 121: 328 – 341
20. Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* 453: 930 – 934
21. Koyano S, Ito M, Takamatsu N, Takiguchi S, Shiba T (1997) The Xenopus Sox3 gene expressed in oocytes of early stages. *Gene* 188: 101 – 107
22. Miura I, Ohtani H, Nakamura M, Ichikawa Y, Saitoh K (1998) The origin and differentiation of the heteromorphic sex chromosomes Z, W, X, and Y in the frog *Rana rugosa*, inferred from the sequences of a sex-linked gene, ADP/ATP translocase. *Mol Biol Evol* 15: 1612 – 1619
23. Graves JA (2013) How to evolve new vertebrate sex determining genes. *Dev Dyn* 242: 354 – 359
24. Miura IET, Ohtani H, Uno Y, Nishida C, Matsuda Y, Graves JAM (2009) The W chromosome evolution and sex-linked gene expression in the frog *Rana rugosa*. In *Sex Chromosomes: Genetics, Abnormalities and Disorders*, Weingarten CN, Jefferson SE (eds), pp 123 – 140. Scottsdale, AZ: Nova Science Inc
25. Wilhelm D, Palmer S, Koopman P (2007) Sex determination and gonadal development in mammals. *Physiol Rev* 87: 1 – 28
26. Kashimada K, Koopman P (2010) Sry: the master switch in mammalian sex determination. *Development* 137: 3921 – 3930
27. Zhao L, Ng ET, Davidson TL, Longmuss E, Urschitz J, Elston M, Moisyadi S, Bowles J, Koopman P (2014) Structure-function analysis of mouse Sry reveals dual essential roles of the C-terminal polyglutamine tract in sex determination. *Proc Natl Acad Sci USA* 111: 11768 – 11773

28. Dubin RA, Ostrer H (1994) Sry is a transcriptional activator. *Mol Endocrinol* 8: 1182 – 1192
29. Lovell-Badge R, Canning C, Sekido R (2002) Sex-determining genes in mice: building pathways. *Novartis Found Symp* 244: 4 – 18; discussion 18–22, 35–42, 253–257
30. Pannetier M, Tilly G, Kocer A, Hudrisier M, Renault L, Chesnais N, Costa J, Le Provost F, Vaiman D, Vilotte JL *et al* (2006) Goat SRY induces testis development in XX transgenic mice. *FEBS Lett* 580: 3715 – 3720
31. Chen YS, Racca JD, Sequeira PW, Phillips NB, Weiss MA (2013) Microsatellite-encoded domain in rodent Sry functions as a genetic capacitor to enable the rapid evolution of biological novelty. *Proc Natl Acad Sci USA* 110: E3061 – E3070
32. Zhao L, Koopman P (2012) SRY protein function in sex determination: thinking outside the box. *Chromosome Res* 20: 153 – 162
33. Herpin A, Schartl M (2011) Vertebrate sex determination: questioning the hierarchy. *FEBS J* 278: 1001
34. Herpin A, Schartl M (2011) Dmrt1 genes at the crossroads: a widespread and central class of sexual development factors in fish. *FEBS J* 278: 1010 – 1019
35. Thoma EC, Wagner TU, Weber IP, Herpin A, Fischer A, Schartl M (2011) Ectopic expression of single transcription factors directs differentiation of a medaka spermatogonial cell line. *Stem Cells Dev* 20: 1425 – 1438
36. Raymond CS, Murphy MW, O'Sullivan MG, Bardwell VJ, Zarkower D (2000) Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev* 14: 2587 – 2595
37. Raymond CS, Shamu CE, Shen MM, Seifert KJ, Hirsch B, Hodgkin J, Zarkower D (1998) Evidence for evolutionary conservation of sex-determining genes. *Nature* 391: 691 – 695
38. Miller SW, Hayward DC, Bunch TA, Miller DJ, Ball EE, Bardwell VJ, Zarkower D, Brower DL (2003) A DM domain protein from a coral, *Acropora millepora*, homologous to proteins important for sex determination. *Evol Dev* 5: 251 – 258
39. Chong T, Collins JJ III, Brubacher JL, Zarkower D, Newmark PA (2013) A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nat Commun* 4: 1814
40. Kato Y, Kobayashi K, Watanabe H, Iguchi T (2011) Environmental sex determination in the branchiopod crustacean *Daphnia magna*: deep conservation of a Doublesex gene in the sex-determining pathway. *PLoS Genet* 7: e1001345
41. Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A, Shan Z, Haaf T, Shimizu N, Shima A *et al* (2002) A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc Natl Acad Sci USA* 99: 11778 – 11783
42. Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, Morrey CE, Shibata N, Asakawa S, Shimizu N *et al* (2002) DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417: 559 – 563
43. Veitia R, Nunes M, Brauner R, Doco-Fenzy M, Joanny-Flinois O, Jaubert F, Lortat-Jacob S, Fellous M, McElreavey K (1997) Deletions of distal 9p associated with 46, XY male to female sex reversal: definition of the breakpoints at 9p23.3–p24.1. *Genomics* 41: 271 – 274
44. Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P, Song W, An N, Chalopin D, Volff JN *et al* (2014) Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet* 46: 253 – 260
45. Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ, Sinclair AH (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* 461: 267 – 271
46. Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C, Matsuda Y, Takamatsu N, Shiba T, Ito M (2008) A W-linked DM-domain gene, DM-W, participates in primary ovary development in *Xenopus laevis*. *Proc Natl Acad Sci USA* 105: 2469 – 2474
47. Kim S, Bardwell VJ, Zarkower D (2007) Cell type-autonomous and non-autonomous requirements for Dmrt1 in postnatal testis differentiation. *Dev Biol* 307: 314 – 327
48. Matson CK, Murphy MW, Griswold MD, Yoshida S, Bardwell VJ, Zarkower D (2010) The mammalian doublesex homolog DMRT1 is a transcriptional gatekeeper that controls the mitosis versus meiosis decision in male germ cells. *Dev Cell* 19: 612 – 624
49. Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D (2011) DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* 476: 101 – 104
50. Uhlenhaut NH, Jakob S, Anlag K, Eisenberger T, Sekido R, Kress J, Treier AC, Klugmann C, Klasen C, Holter NI *et al* (2009) Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* 139: 1130 – 1142
51. Boulanger L, Pannetier M, Gall L, Allais-Bonnet A, Elzaiat M, Le Bourhis D, Daniel N, Richard C, Cotinot C, Ghyselinck NB *et al* (2014) FOXL2 is a female sex-determining gene in the goat. *Curr Biol* 24: 404 – 408
52. Georges A, Auguste A, Bessiere L, Vanet A, Todeschini AL, Veitia RA (2014) FOXL2: a central transcription factor of the ovary. *J Mol Endocrinol* 52: R17 – R33
53. Lindeman RE, Gearhart MD, Minkina A, Krentz AD, Bardwell VJ, Zarkower D (2015) Sexual cell-fate reprogramming in the ovary by DMRT1. *Curr Biol* 25: 764 – 771
54. Masuyama H, Yamada M, Kamei Y, Fujiwara-Ishikawa T, Todo T, Nagahama Y, Matsuda M (2012) Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka. *Chromosome Res* 20: 163 – 176
55. Krentz AD, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D (2011) DMRT1 promotes oogenesis by transcriptional activation of Stra8 in the mammalian fetal ovary. *Dev Biol* 356: 63 – 70
56. Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, Tanaka M, Schartl M (2010) Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. *PLoS Genet* 6: e1000844
57. Herpin A, Schindler D, Kraiss A, Hornung U, Winkler C, Schartl M (2007) Inhibition of primordial germ cell proliferation by the medaka male determining gene Dmrt I bY. *BMC Dev Biol* 7: 99
58. Herpin A, Schartl M (2011) Sex determination: switch and suppress. *Curr Biol* 21: R656 – R659
59. Matson CK, Zarkower D (2012) Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat Rev Genet* 13: 163 – 174
60. Cutting A, Chue J, Smith CA (2013) Just how conserved is vertebrate sex determination? *Dev Dyn* 242: 380 – 387
61. Nakamura S, Watakabe I, Nishimura T, Picard JY, Toyoda A, Taniguchi Y, di Clemente N, Tanaka M (2012) Hyperproliferation of mitotically active germ cells due to defective anti-Mullerian hormone signaling mediates sex reversal in medaka. *Development* 139: 2283 – 2287
62. Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T, Fernandino JI, Somoza GM, Yokota M, Strussmann CA (2012) A Y-linked anti-Mullerian hormone duplication takes over a critical role in sex determination. *Proc Natl Acad Sci USA* 109: 2955 – 2959

63. Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, Fujita M, Suetake H, Suzuki S, Hosoya S *et al* (2012) A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu). *PLoS Genet* 8: e1002798
64. Crespo B, Gomez A, Mazon MJ, Carrillo M, Zanuy S (2013) Isolation and characterization of Ff1 and Gsdf family genes in European sea bass and identification of early gonadal markers of precocious puberty in males. *Gen Comp Endocrinol* 191: 155 – 167
65. Gautier A, Le Gac F, Lareyre JJ (2011) The gsdf gene locus harbors evolutionary conserved and clustered genes preferentially expressed in fish previtellogenic oocytes. *Gene* 472: 7 – 17
66. Horiguchi R, Nozu R, Hirai T, Kobayashi Y, Nagahama Y, Nakamura M (2013) Characterization of gonadal soma-derived factor expression during sex change in the protogynous wrasse, *Halichoeres trimaculatus*. *Dev Dyn* 242: 388 – 399
67. Sawatari E, Shikina S, Takeuchi T, Yoshizaki G (2007) A novel transforming growth factor-beta superfamily member expressed in gonadal somatic cells enhances primordial germ cell and spermatogonial proliferation in rainbow trout (*Oncorhynchus mykiss*). *Dev Biol* 301: 266 – 275
68. Shibata Y, Paul-Prasanth B, Suzuki A, Usami T, Nakamoto M, Matsuda M, Nagahama Y (2010) Expression of gonadal soma derived factor (GSDF) is spatially and temporally correlated with early testicular differentiation in medaka. *Gene Expr Patterns* 10: 283 – 289
69. Myosho T, Otake H, Masuyama H, Matsuda M, Kuroki Y, Fujiyama A, Naruse K, Hamaguchi S, Sakaizumi M (2012) Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*. *Genetics* 191: 163 – 170
70. Rondeau EB, Messmer AM, Sanderson DS, Jantzen SG, von Schalburg KR, Minkley DR, Leong JS, Macdonald GM, Davidsen AE, Parker WA *et al* (2013) Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene. *BMC Genom* 14: 452
71. Cho S, Huang ZY, Zhang J (2007) Sex-specific splicing of the honeybee doublesex gene reveals 300 million years of evolution at the bottom of the insect sex-determination pathway. *Genetics* 177: 1733 – 1741
72. Bopp D, Saccone G, Beye M (2014) Sex determination in insects: variations on a common theme. *Sex Dev* 8: 20 – 28
73. Williams TM, Carroll SB (2009) Genetic and molecular insights into the development and evolution of sexual dimorphism. *Nat Rev Genet* 10: 797 – 804
74. Cline TW, Meyer BJ (1996) Vive la difference: males versus females in flies versus worms. *Annu Rev Genet* 30: 637 – 702
75. Scali C, Catteruccia F, Li Q, Crisanti A (2005) Identification of sex-specific transcripts of the *Anopheles gambiae* doublesex gene. *J Exp Biol* 208: 3701 – 3709
76. Lagos D, Koukidou M, Savakis C, Komitopoulou K (2007) The transformer gene in *Bactrocera oleae*: the genetic switch that determines its sex fate. *Insect Mol Biol* 16: 221 – 230
77. Geuverink E, Beukeboom LW (2014) Phylogenetic distribution and evolutionary dynamics of the sex determination genes doublesex and transformer in insects. *Sex Dev* 8: 38 – 49
78. Suzuki MG, Ohbayashi F, Mita K, Shimada T (2001) The mechanism of sex-specific splicing at the doublesex gene is different between *Drosophila melanogaster* and *Bombyx mori*. *Insect Biochem Mol Biol* 31: 1201 – 1211
79. Verhulst EC, Beukeboom LW, van de Zande L (2010) Maternal control of haplodiploid sex determination in the wasp Nasonia. *Science* 328: 620 – 623
80. Hediger M, Henggeler C, Meier N, Perez R, Saccone G, Bopp D (2010) Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. *Genetics* 184: 155 – 170
81. Beye M, Hasselmann M, Fondrk MK, Page RE, Omholt SW (2003) The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* 114: 419 – 429
82. Hasselmann M, Gempe T, Schiott M, Nunes-Silva CG, Otte M, Beye M (2008) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature* 454: 519 – 522
83. Beye M (2004) The dice of fate: the csd gene and how its allelic composition regulates sexual development in the honey bee, *Apis mellifera*. *BioEssays* 26: 1131 – 1139
84. Hall AB, Basu S, Jiang X, Qi Y, Timoshevskiy VA, Biedler JK, Sharakhova MV, Elahi R, Anderson MA, Chen XG *et al* (2015) A male-determining factor in the mosquito *Aedes aegypti*. *Science* 348: 1268 – 1270
85. Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, Cabau C, Bouchez O, Fostier A, Guiguen Y (2012) An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. *Curr Biol* 22: 1423 – 1428
86. Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S, Sugano S, Shimada T *et al* (2014) A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* 509: 633 – 636
87. Herpin A, Adolfi MC, Nicol B, Hinzmann M, Schmidt C, Klughammer J, Engel M, Tanaka M, Guiguen Y, Schartl M (2013) Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination. *Mol Biol Evol* 30: 2328 – 2346
88. Canto P, Soderlund D, Reyes E, Mendez JP (2004) Mutations in the desert hedgehog (DHH) gene in patients with 46, XY complete pure gonadal dysgenesis. *J Clin Endocrinol Metab* 89: 4480 – 4483
89. Clark AM, Garland KK, Russell LD (2000) Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. *Biol Reprod* 63: 1825 – 1838
90. Pierucci-Alves F, Clark AM, Russell LD (2001) A developmental study of the Desert hedgehog-null mouse testis. *Biol Reprod* 65: 1392 – 1402
91. Zhang Y, Li F, Sun D, Liu J, Liu N, Yu Q (2011) Molecular analysis shows differential expression of R-spondin1 in zebrafish (*Danio rerio*) gonads. *Mol Biol Rep* 38: 275 – 282
92. Chassot AA, Bradford ST, Auguste A, Gregoire EP, Pailhoux E, de Rooij DG, Schedl A, Chaboissier MC (2012) WNT4 and RSPO1 together are required for cell proliferation in the early mouse gonad. *Development* 139: 4461 – 4472
93. Chadi S, Buscara L, Pechoux C, Costa J, Laubier J, Chaboissier MC, Pailhoux E, Vilotte JL, Chanat E, Le Provost F (2009) R-spondin1 is required for normal epithelial morphogenesis during mammary gland development. *Biochem Biophys Res Commun* 390: 1040 – 1043
94. Kashimada K, Pelosi E, Chen H, Schlessinger D, Wilhelm D, Koopman P (2011) FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology* 152: 272 – 280
95. Meinhardt A, O'Bryan MK, McFarlane JR, Loveland KL, Mallidis C, Foulds LM, Phillips DJ, de Kretser DM (1998) Localization of follistatin in the rat testis. *J Reprod Fertil* 112: 233 – 241
96. Nakamura S, Watakabe I, Nishimura T, Toyoda A, Taniguchi Y, Tanaka M (2012) Analysis of medaka sox9 orthologue reveals a conserved role in germ cell maintenance. *PLoS ONE* 7: e29982

97. Yokoi H, Kobayashi T, Tanaka M, Nagahama Y, Wakamatsu Y, Takeda H, Araki K, Morohashi K, Ozato K (2002) Sox9 in a teleost fish, medaka (*Oryzias latipes*): evidence for diversified function of Sox9 in gonad differentiation. *Mol Reprod Dev* 63: 5 – 16
98. Colvin JS, Green RP, Schmahl J, Capel B, Ornitz DM (2001) Male-to-female sex reversal in mice lacking fibroblast growth factor 9. *Cell* 104: 875 – 889
99. Forconi M, Canapa A, Barucca M, Biscotti MA, Capriglione T, Buonocore F, Fausto AM, Makapedua DM, Pallavicini A, Gerdol M *et al* (2013) Characterization of sex determination and sex differentiation genes in Latimeria. *PLoS ONE* 8: e56006
100. Halder G, Callaerts P, Gehring WJ (1995) Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science* 267: 1788 – 1792
101. Wilkins AS (1995) Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *BioEssays* 17: 71 – 77
102. Gempe T, Beye M (2011) Function and evolution of sex determination mechanisms, genes and pathways in insects. *BioEssays* 33: 52 – 60
103. Wilson CA, High SK, McCluskey BM, Amores A, Yan YL, Titus TA, Anderson JL, Batzel P, Carvan MJ III, Schartl M *et al* (2014) Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* 198: 1291 – 1308
104. Nanda I, Hornung U, Kondo M, Schmid M, Schartl M (2003) Common spontaneous sex-reversed XX males of the medaka *Oryzias latipes*. *Genetics* 163: 245 – 251
105. Beukeboom LW, Perrin N (2014) *The Evolution of Sex Determination*. New York: Oxford University Press
106. Bull JJ, Charnov EL (1977) Changes in the heterogametic mechanism of sex determination. *Heredity* 39: 1 – 14
107. Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* 355: 1563 – 1572
108. Blaser O, Grossen C, Neuenschwander S, Perrin N (2013) Sex-chromosome turnovers induced by deleterious mutation load. *Evolution* 67: 635 – 645
109. van Doorn GS, Kirkpatrick M (2007) Turnover of sex chromosomes induced by sexual conflict. *Nature* 449: 909 – 912
110. Lynch M (2007) The evolution of genetic networks by non-adaptive processes. *Nat Rev Genet* 8: 803 – 813
111. DeFalco T, Capel B (2009) Gonad morphogenesis in vertebrates: divergent means to a convergent end. *Annu Rev Cell Dev Biol* 25: 457 – 482

Schartl et al. BMC Biology (2018) 16:16
DOI 10.1186/s12915-018-0485-8

BMC Biology

RESEARCH ARTICLE Open Access

Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements

Manfred Schartl[1,2,3], Susanne Schories[1], Yuko Wakamatsu[1], Yusuke Nagao[4], Hisashi Hashimoto[4], Chloé Bertin[5], Brigitte Mourot[5], Cornelia Schmidt[1], Dagmar Wilhelm[6], Lazaro Centanin[7], Yann Guiguen[5] and Amaury Herpin[1,5*]

Abstract

Background: Sex determination relies on a hierarchically structured network of genes, and is one of the most plastic processes in evolution. The evolution of sex-determining genes within a network, by neo- or sub-functionalization, also requires the regulatory landscape to be rewired to accommodate these novel gene functions. We previously showed that in medaka fish, the regulatory landscape of the master male-determining gene *dmrt1bY* underwent a profound rearrangement, concomitantly with acquiring a dominant position within the sex-determining network. This rewiring was brought about by the exaptation of a transposable element (TE) called *Izanagi*, which is co-opted to act as a silencer to turn off the *dmrt1bY* gene after it performed its function in sex determination.

Results: We now show that a second TE, *Rex1*, has been incorporated into *Izanagi*. The insertion of *Rex1* brought in a preformed regulatory element for the transcription factor Sox5, which here functions in establishing the temporal and cell-type-specific expression pattern of *dmrt1bY*. Mutant analysis demonstrates the importance of Sox5 in the gonadal development of medaka, and possibly in mice, in a *dmrt1bY*-independent manner. Moreover, Sox5 medaka mutants have complete female-to-male sex reversal.

Conclusions: Our work reveals an unexpected complexity in TE-mediated transcriptional rewiring, with the exaptation of a second TE into a network already rewired by a TE. We also show a dual role for Sox5 during sex determination: first, as an evolutionarily conserved regulator of germ-cell number in medaka, and second, by de novo regulation of *dmrt1* transcriptional activity during primary sex determination due to exaptation of the Rex1 transposable element.

Keywords: Exaptation, Master sex-determining gene, Transcriptional rewiring, Medaka, Dmrt1bY, Sox5

Background

Sex determination (SD) is one of the most plastic processes in evolution. The trigger for the bipotential undifferentiated embryonic gonad anlage to develop into either testis or ovary can be provided by various signals from the environment, the genetic constitution of the individual, or a mixture of both [1, 2]. Studies of the modes of genetic SD revealed that the genes at the top of the regulatory network and the genes of the network itself are subject to rapid changes in evolution. New master SD genes evolved repeatedly and independently [3]. This situation is particularly obvious in fish, since closely related sister species can have different genetic SD systems or master SD genes [3–7].

Clearly, such a high turnover of genetic determinants can work only if the evolutionary innovations are accompanied by the ability of the respective genes to neo-functionalize or sub-functionalize quickly and efficiently [8, 9]. In addition to changes in protein structure, differences in gene regulation have an important role in evolution and are considered a quick and effective way to adapt gene functions to novelty [10–12]. Hence, the necessity for the transcriptional rewiring of the architecture of the SD regulatory network and connecting novel master SD genes to it requires high-capacity and fast mechanisms. Such a mechanism was proposed by Britten and Davidson

* Correspondence: amaury.herpin@inra.fr
[1]Physiological Chemistry, Biocenter, University of Würzburg, 97074 Würzburg, Germany
[5]INRA, UR1037 Fish Physiology and Genomics, F-35000 Rennes, France
Full list of author information is available at the end of the article

almost 50 years ago [13, 14]. They hypothesized that transposable elements (TEs) carry preformed transcription factor binding sites, which, after mobilization and insertion into novel locations of the genome, would contribute novel regulatory features to nearby genes through these motifs. However, examples of authentic co-option, or exaptation [15], of TEs, where most or all gene transcriptions initiate within a TE, remain sparse (see [16] for a review).

Intriguingly, of the handful of examples of this process, one of the best documented comes from a novel SD gene. The master male-determining gene *dmrt1bY* of the teleost fish *Oryzias latipes* (medaka) arose approximately 5 to 10 million years ago from an autosome encompassing the *dmrt1* gene. Dmrt1 is a highly conserved transcription factor that usually functions at a downstream position of the sexual regulatory cascade. In medaka, the *dmrt1* gene was locally duplicated and inserted into another chromosome that became the Y-chromosome [5]. To exert its novel function at its new upstream position, *dmrt1bY* acquired a divergent expression pattern and effector gene profile compared to its autosomal ancestor, *dmrt1a* [5, 17]. We previously showed that this evolutionary innovation, which required a complete rewiring of the regulatory network, was partly brought about by exaptation of a ready-to-use pre-existing *cis*-regulatory element contributed by a TE, called *Izanagi* [17]. This element acts as a silencer. It recruits proteins Dmrt1bY and Dmrt1a to turn off the *dmrt1bY* gene after it has fulfilled its function as the primary male SD gene.

We report here that TE-mediated transcriptional rewiring can reach an unexpected level of complexity that exceeds this simple feedback regulation. We find that another TE, *Rex1*, has jumped into *Izanagi*. Through the disruption of *Izanagi*, *Rex1* immobilized this TE and fixed the Dmrt1-mediated downregulation. Moreover, *Rex1* brought in a preformed regulatory element for the transcription factor Sox5. We show that medaka Sox5 binds to the *sox5*-responsive elements of the *dmrt1bY* promoter and downregulates its transcriptional activity. Interestingly, in vivo analysis of double transgenic fluorescent reporter fish additionally revealed a complementary pattern of expression of both genes. The higher expression of *sox5* correlates with a lower expression of *dmrt1bY* and vice versa. Our results underpin the importance of the *Rex1* TE for the establishment of a new SD mechanism in medaka and likely contribute in establishing the temporal and cell-type specific expression pattern of *dmrt1bY*.

Several transcription factors of the Sox family (Sox3, SRY, Sox 9, and sox8) play crucial roles in SD, but Sox5 has not been previously implicated in SD in any metazoan so far (neither Sox5 in vertebrates, nor its *Drosophila* homologue Sox102F). Interestingly, in medaka, disruption of *sox5* leads to XX female-to-male sex reversal. From an analysis of mutants, we find the critical involvement in gonadal development in medaka by regulating primordial germ cells (PGCs). In overexpression experiments, there is an ectopic induction of germ-cell markers including *dmrt1*. With all necessary notes of caution, our preliminary expression pattern data, also detecting SOX5 expression in the fetal gonad of mice, may indicate an evolutionarily conserved role for SOX5 during early mammalian gonad development.

Our work reveals a dual role for *sox5* during SD: (i) first being an evolutionarily conserved important regulator of germ-cell number in medaka and possibly beyond and (ii) second, de novo regulation of medaka *dmrt1* transcriptional activity during primary SD after it has been recruited for transcriptional rewiring of the *dmrt1* promoter due to exaptation of a TE.

Results

Identification of putative Sox5 transcription factor binding sites in the *dmrt1bY* promoter

In an initial analysis of the promoter of the medaka male-determining gene *dmrt1bY*, we found that after duplication from its autosomal progenitor *dmrt1a*, an insertion of an *Izanagi* DNA transposon brought in a novel transcriptional repressor element [17]. It functions by binding Dmrt1a and Dmrt1bY transcription factors and is essential for the downregulation of *dmrt1bY* after fulfilling its SD function in the male gonad.

In addition to this repressor element [17], the promoter region contains an unexpectedly high density of putative binding sites for Sox5 (see the [β] region in Fig. 1a and [17]). It harbors seven Sox5 binding sites; a random prediction would expect only 0.46 sites over the whole sequence. Interestingly, a unique putative Sox5 binding site is also found within a *Rex1* TE [α], and two within repeat 3 [γ] of the *dmrt1bY* proximal promoter region (see [α], [β], or [γ] in Fig. 1a and Additional file 1: Figure S1 for sox5 binding site locations). These regions were all inserted into the promoter after the duplication event and, thus, during the evolution of the novel male-determining function of *dmrt1bY*.

In particular, the Sox5 binding site nested within the *Rex1* TE of the *dmrt1bY* proximal active promoter showed high prediction probability (weight 9.4, p value 5.4×10^{-5}, lnPval -9;831; [α] in Fig. 1a, see also "Methods" for the positional weight matrix employed). To address the question whether this Sox5 binding site has evolved de novo after insertion or was already an integral part of the *Rex1* element that was inserted into the *dmrt1bY* promoter, we blasted the *Rex1* sequence of the *dmrt1bY* promoter against the medaka and other fish genomes (Fig. 1b,c). *Rex1* elements are present in

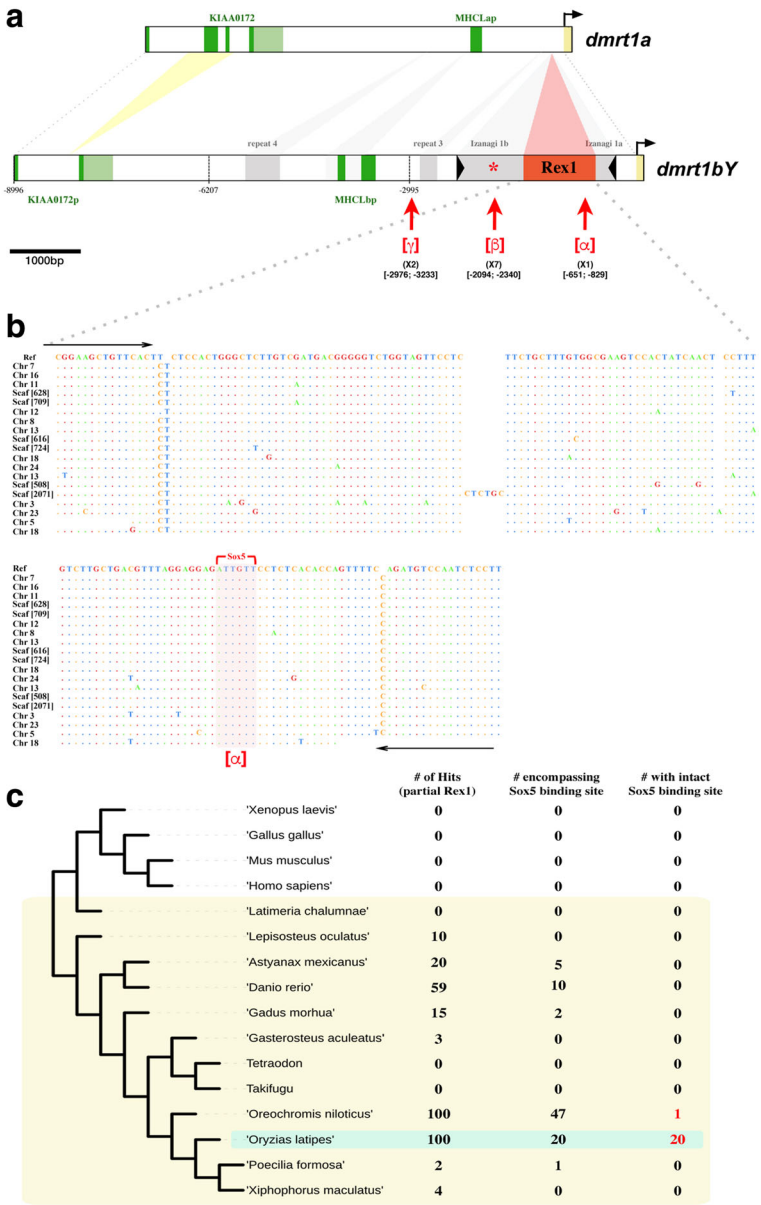


Fig. 1 Comparative analysis of the *dmrt1a* and *dmrt1bY* co-ortholog promoters and presence of *Rex1* element sequences in the genomes of selected fish species. **a** Comparative analysis of the medaka *dmrt1* co-ortholog promoter regions. Differences in length for the two promoter regions are caused by *Rex1* and *Izanagi* transposable elements as well as repeats 3 and 4 that were inserted into the *dmrt1bY* promoter after the duplication event that gave rise to the *dmrt1bY* gene approximately 10 million years ago [17]. Regions α, β, or γ (brackets [] underlined in red) contain multiple Sox5 binding sites within *Rex1*, *Izanagi*, and repeat 3, respectively, that have been subjected to chromatin immunoprecipitation (ChIP) (see also Additional file 1: Figure S1). The red star (*) identifies the Dmrt1 binding site described in [17]. **b** Alignment of the Y-chromosomal *Rex1* element together with the 19 remaining *Rex1* copies encompassing the sox5 binding site in the medaka genome. Dots indicate conserved nucleotides. Black arrows define primers used for chromatin immunoprecipitation. **c** Presence of *Rex1* element (i) partial sequences, (ii) sequences encompassing the dmrt1bY-nested sox5 binding site, and (iii) sequences encompassing the dmrt1bY-nested sox5 binding site with the intact sox5 binding site in the genomes of medaka (*Oryzias latipes*), tilapia (*Oreochromis niloticus*), zebrafish (*Danio rerio*), cavefish (*Astyanax mexicanus*), cod (*Gadus morhua*), gar (*Lepisosteus oculatus*), stickleback (*Gasterosteus aculeatus*), platy (*Xiphophorus maculatus*), Amazon molly (*Poecilia formosa*), fugu, tetraodon, and coelacanth (*Latimeria chalumnae*)

100 copies in the genome of medaka. Many copies are also found in tilapia (100) and zebrafish (59), but there are fewer in cavefish (20), cod (15) and gar (10). They are scarce in stickleback (3), platyfish (4), and Amazon molly (2), and absent in fugu, tetraodon, and coelacanth (Fig. 1c). Interestingly, among Rex*1* elements, the region encompassing the predicted *sox5*-binding site is very poorly conserved despite being part of the reverse transcriptase-coding region (region 6) of the *Rex1* element (Additional file 2: Figure S2). It can be detected with

some divergence to the consensus sequence in only 47, 20, 10, 5, 2, and 1 copies in tilapia, medaka, zebrafish, cavefish, cod, and Amazon molly, respectively (Fig. 1c). An intact sox5 binding site is predicted in only two species. In tilapia, a single copy has a high-fidelity site, whereas 20 copies in medaka, including the one in the *dmrt1bY* promoter, have putatively intact *sox5* binding sites (Fig. 1b,c and Additional file 3: Table S1). Hence, *Rex1*-nested *sox5* binding sites appear to be a medaka-specific feature.

Sox5 has been correlated with *dmrt1* expression in zebrafish [18] and the wrasse, *Halichoeres tenuispinis* [19], in in vitro promoter studies. Thus, we hypothesized that the identified *sox5* binding sites could be involved in the transcriptional rewiring of *dmrt1bY*.

Sox5 binds to the putative Sox5-responsive elements of the dmrt1bY promoter with different affinities

To assess the relevance of the predicted Sox5 binding sites [α], [β], and [γ] (see Fig. 1a,b) in the *dmrt1bY* promoter, two different medaka cell lines, *Oryzias latipes* spermatogonial (Sg3) and fibroblast (OLF) cells, were transfected with a FLAG-tagged version of Sox5 and then subjected to chromatin immunoprecipitation (ChIP) using an anti-FLAG antibody. DNA amplification with specific sets of primers from the Sox5 immunoprecipitates revealed that the Sox5 protein binds to the predicted sites ([α], [β], and [γ]) of the *dmrt1bY* promoter. However, much stronger binding is apparent for the proximal site [α] (up to tenfold higher enrichment, Fig. 2a) located within the *Rex1* element. Of note, although two DNA-binding sequences for Sox5 were predicted within the *dmrt1a* proximal promoter (at positions [-286/-300] and [-1100/-1116] upstream of the ATG), they do not appear to be functional, as Sox5 does not target them in the ChIP experiments (data not shown).

Of note, the predicted *sox5* binding region [β] has previously been shown to overlap with a high affinity binding site for Dmrt1 [17]. Hence, competition between Sox5 and Dmrt1 for access to this sequence motif cannot be excluded. If Dmrt1 already occupies the motif, this might explain the low amount of recovery in the Sox5 ChIP experiment.

Sox5 downregulates the activity of the dmrt1bY promoter

Next, we examined (i) the respective contributions of each part of the promoter for *dmrt1bY* transcriptional regulation and (ii) the direction of that regulation (up- or downregulation) using transcriptional reporter assays. Thus, 5′ deletions of the *dmrt1bY* promoter ([0/-1593], [0/-2995], [0/-6207], and [0/-8996]) were produced and analyzed after transient transfection for their ability to drive luciferase expression (Fig. 2b).

In the medaka spermatogonial cell line, the highest promoter activity was detectable for the [0/-2995] proximal promoter region encompassing sites [α] and [β] (Fig. 2b2). Promoter activity was significantly lower (Fig. 2b3,4), when more distal sequences ([-2996/-6207]) containing site [γ] were present in the construct. The shortest proximal promoter region (encompassing site [α] only) had intermediate transcriptional activity (Fig. 2b1). Interestingly, the transcriptional activity of the most proximal parts of the *dmrt1bY* promoter—encompassing the sites [α] in *Rex1* (Fig. 2b1) and [β] in *Izanagi* (Fig. 2b2)—was reduced by between 43% and 68% when Sox5 was overexpressed (Fig. 2b1,2). This effect of Sox5 overexpression was not apparent for the longer constructs including further distal sequences displaying strong constitutive repression (Fig. 2b3,4). Of note, the highest repression of *dmrt1bY* promoter transcriptional activity was observed for the proximal promoter region encompassing the [α] site in *Rex1* (68% in Fig. 2b1). Next, modulation of transcriptional activity was tested for the [α], [α]-MUT, and [β] regions alone (Fig. 2c). Interestingly, only the [α] region was able to downregulate promoter activity (by about 40%) when fused to a minimal thymidine kinase promoter. Neither the [β] or a mutated version of the [α] region ([α]-MUT) were able to modulate the activity of the minimal thymidine kinase promoter (Fig. 2c).

To obtain a more precise readout for the regulation of medaka *dmrt1bY* expression by Sox5, spermatogonial and fibroblast medaka cell lines were transiently transfected with a Sox5-expressing construct and endogenous *dmrt1bY* expression was quantified by the real-time polymerase chain reaction (RT-PCR) at different time points post-transfection (24, 48, and 72 h). Our findings reveal that the transcription of *dmrt1bY* is highly repressed in the presence of Sox5 (up to 90% 48 h after transfection, Fig. 3a) in both the OLF and Sg3 medaka cell lines.

Next, to validate our in vitro results, Sox5 coding mRNAs were microinjected into one-cell-stage embryos, and the expression of *dmrt1bY* and several germ-cell markers (*nanos2*, *nanos3*, *dead-end*, *vasa*, *tra2a*, and *piwi*) was monitored (Fig. 3b). The results confirm the transient cell transfection experiments and demonstrated that in vivo Sox5 can act as a negative modulator of *dmrt1bY* expression (Fig. 3b). Overexpression of Sox5 also resulted in the repression of all analyzed germ-cell genes regardless of their intrinsic endogenous expression levels (Fig. 3b). Our in vivo results identify Sox5 as a strong negative regulator of germ-cell gene expression, including *dmrt1bY*. Interestingly, although most of the germ-cell marker genes (*nanos3*, *dead-end*, *vasa*, and *piwi*) are maternally deposited, their lower relative abundances compared to controls after Sox5 overexpression

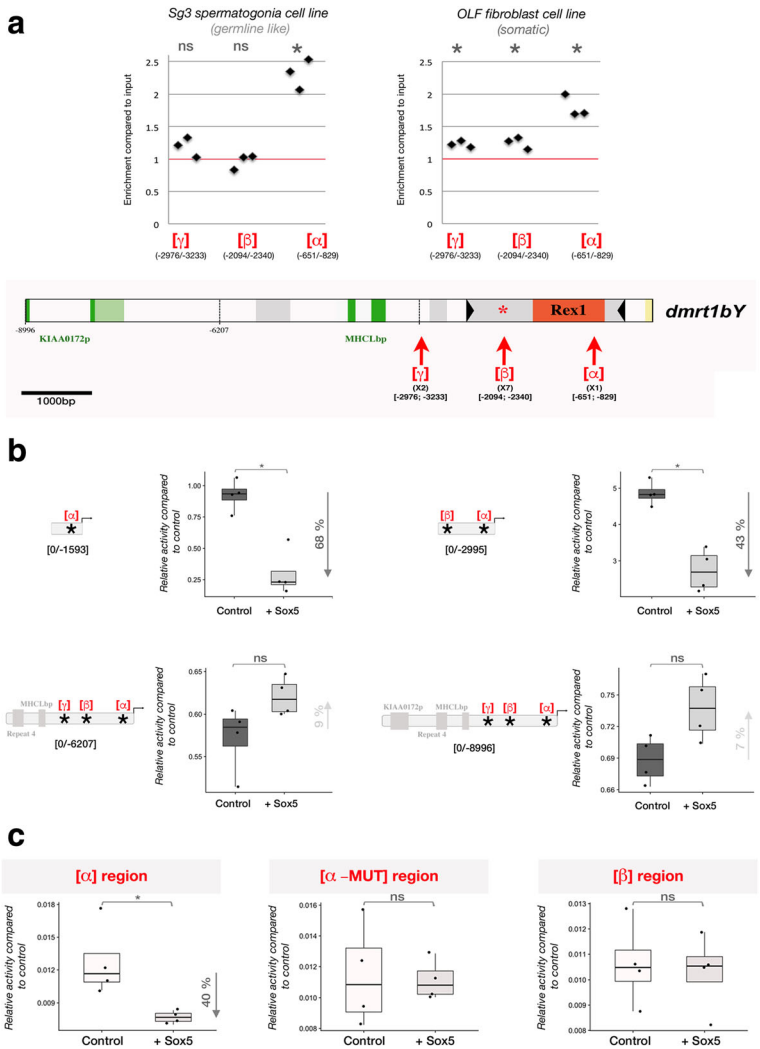


Fig. 2 Analysis of Sox5 binding to the *dmrt1bY* promoter and regulation of *dmrt1bY* promoter activity upon modulation of Sox5 expression. **a** Chromatin immunoprecipitation (ChIP) of Sox5 binding to regions α, β, or γ of the *dmrt1bY* proximal promoter. Transient transfection of a flagged and tagged version of Sox5 into either medaka spermatogonial or fibroblast cell lines and subsequent immunoprecipitation (FLAG antibody) followed by the quantitative real-time polymerase chain reaction. Results are presented as enrichment compared to the input and correspond to three independent immunoprecipitations for each region (α, β, or γ). Statistical significance was assessed with the Wilcoxon–Mann–Whitney test ($N = 3$). **b1–b4** Analysis of *dmrt1bY* proximal promoter activity after Sox5 transient transfection into the medaka spermatogonial cell line (Sg3). Deletions of the 5′ *dmrt1bY* promoter were generated: **b1** α region, **b2** α and β regions, **b3, b4** α, β, and γ regions. Transcriptional activity was quantified in the absence (control, -Sox5) or presence (+Sox5) of Sox5. Statistical significance was assessed with the Wilcoxon–Mann–Whitney test ($N = 4$). **c** Detailed analysis of the transcriptional activity of the alpha (α), alpha-mutant (α-MUT), and beta (β) fragments. Statistical significance was assessed with the Wilcoxon–Mann–Whitney test ($N = 4$). * p value ≤ 0.05, ** p value ≤ 0.01. ns non-significant, OLF *Oryzias latipes* fibroblast, Sg3 *Oryzias latipes* spermatogonial cell

are likely attributable to a total arrest of background transcription after zygotic transcription started, or possibly accelerated mRNA decay.

Expression of sox5 during early gonad primordium formation

Medaka *sox5* mRNA, which is expressed in embryonic and early larval development, has a distinct spatially and temporarily restricted expression pattern (Fig. 4). Between stages 18 and 22, *sox5* transcripts localize mainly in the head and tail bud regions of the embryos (Fig. 4a–d). At stage 22, expression is additionally detected in the lateral plate mesoderm (arrowheads in Fig. 4c,d), from which the somatic gonadal primordium will develop shortly thereafter [20]. Later, *sox5* expression is observed over the entire dorsal neural tube (Fig. 4e,e',f,f') and pre-migratory neural crest cells (NCCs, Fig. 4f'). At stage 29, *sox5* transcripts are present in migrating NCCs ventrally [21] (Fig. 4g and arrowheads in g'). A higher resolution of the dynamic

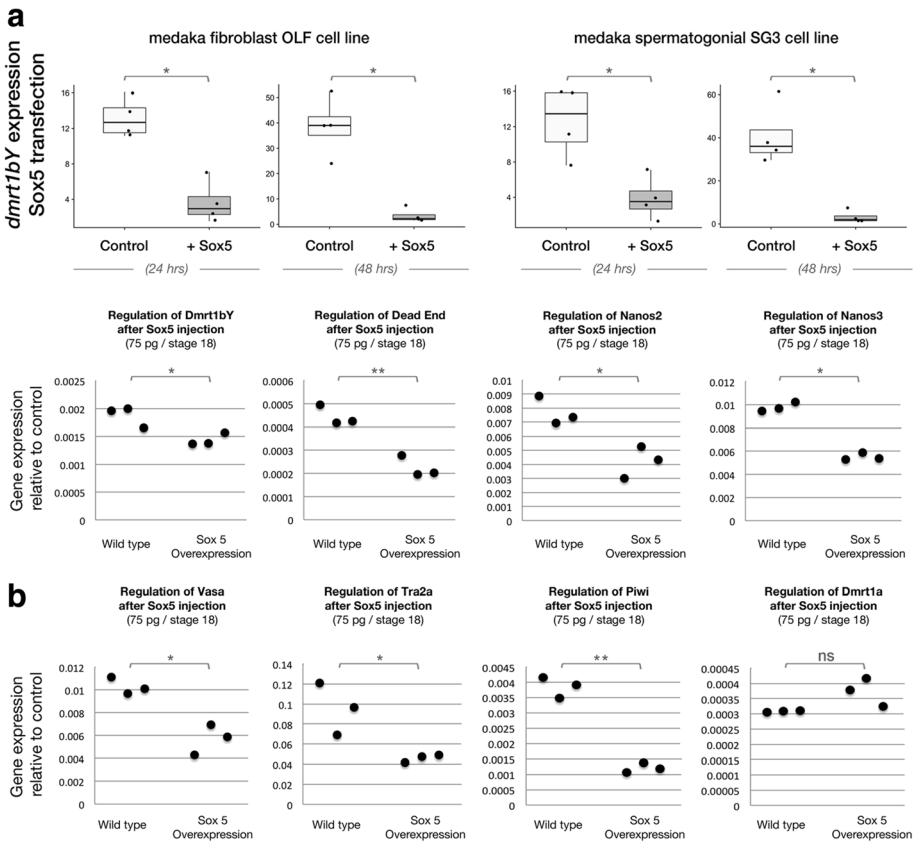


Fig. 3 Effects of Sox5 modulation on *dmrt1bY* gene expression. Analysis of *dmrtbY* **a** transcriptional regulation after *sox5* transient transfection in either spermatogonial or fibroblast medaka cell lines at 24 or 48 h post-transfection. Dataset results of four independent transfections. Statistical significance was assessed with the Wilcoxon–Mann–Whitney test ($N = 4$; * p value ≤ 0.05, ** p value ≤ 0.01). **b** Overexpression of sox5 was stimulated in fish eggs. The expression of *dmrt1bY* and germ-cell markers (*nanos2*, *nanos3*, *dead-end*, *vasa*, *tra2a*, and *piwi*) were monitored at stage 18 of development and compared to wild-type fish embryos. Dataset results of three independent transfections. Statistical significance was assessed with the *t*-test ($N = 3$ and each replicate is a pool of 25 eggs; * p value ≤ 0.05, ** p value ≤ 0.01). ns non-significant, OLF *Oryzias latipes* fibroblast, Sg3 *Oryzias latipes* spermatogonial cell

expression pattern of *sox5* in vivo was obtained with a transgenic line, which has the 3288-bp upstream promoter region of *sox5* fused to a fluorescent (*mCherry*) reporter. The reliability of *sox5* gene expression was confirmed by comparison of the observed fluorescence with the in situ hybridization data (Fig. 4c and f' compared to d, h, and i, respectively). *Sox5* promoter-driven fluorescence was detected as early as stage 22 in the lateral plate mesoderm (Fig. 4d). This region has been shown to have already the properties of a gonadal field because somatic gonadal precursors arise from the most posterior part of the lateral plate mesoderm [20]. Consistently, at stages 26/28, when the gonadal primordium just has formed, faint *sox5* expression is still observable in the somatic tissues surrounding the germ cells that express Dmrt1bY (Fig. 5a–c).

At stages 33 to 34, the gonadal expression of *sox5* is restricted to the germ cells (Fig. 5d–i). Of note, variations in the levels *of sox5* expression are clearly visible between different germ cells (Fig. 5d,e). Interestingly, *dmrt1bY* (Dmrt1bY:GFP) is also expressed in germ cells at that specific stage of development [22] and displays variations in expression between individual germ cells [23, 24] (Fig. 5d,f). Analysis of the [Sox5:mCherry and Dmrt1bY:GFP] double transgenic line revealed a complementary pattern of expression of both genes: those germ cells that have a higher expression of *sox5* have a lower expression of *dmrt1bY* and vice versa (Fig. 5d–i).

During the following developmental stages (stages 38/39), the expression of *sox5*:*mCherry* increases in all germ cells, whereas *dmrt1bY:GFP* concurrently switches from germ cells to the somatic, germ-cell-surrounding cells (Fig. 5j,k). Taken together, these results show that expression of *sox5* and *dmrt1bY* is highly dynamic during gonadal primordium formation, switching from germ cells to somatic cells (Fig. 5l).

To determine whether the expression of sox5 during early gonadal development is a medaka-specific feature or is more widely conserved, we used immunofluorescence on 13.5 and 14.5 days post coitum (dpc) mouse embryos.

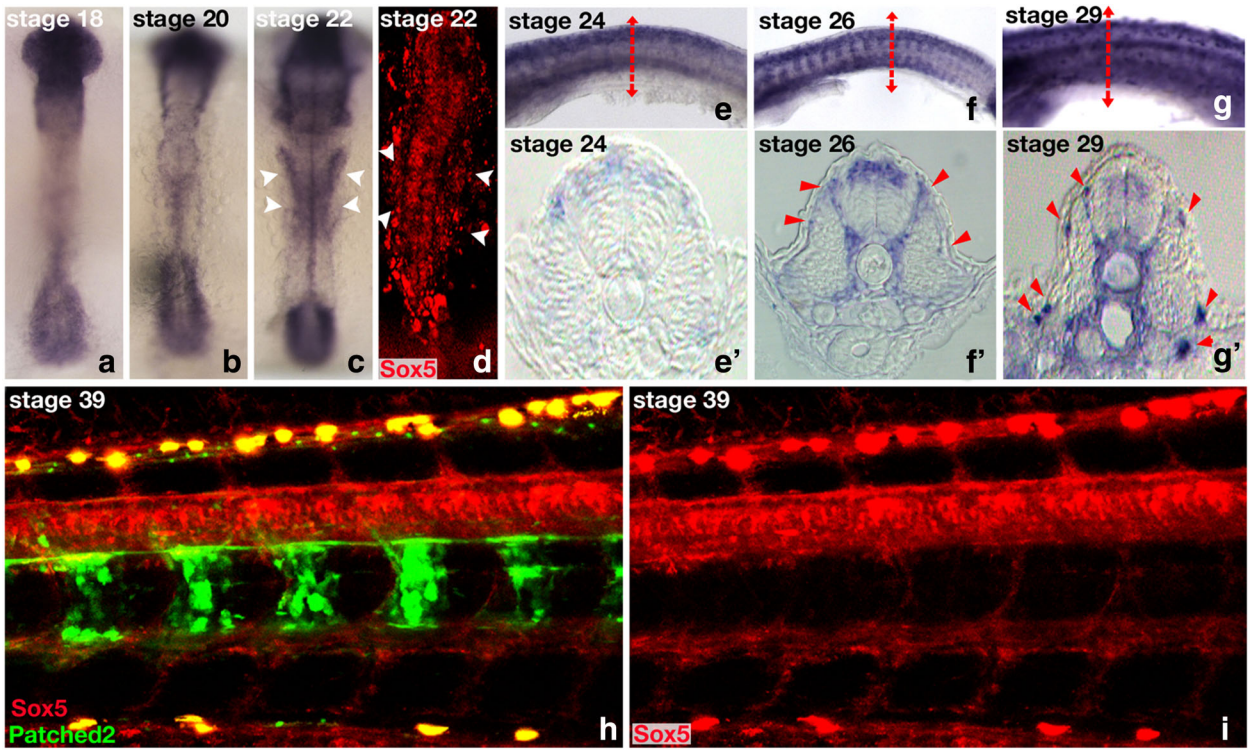


Fig. 4 Expression of medaka *sox5* during embryogenesis. **a** to **c** and **e** to **g** Medaka *sox5* expression investigated by whole-mount in situ hybridization or **d**, **h**, and **i** fluorescence using a transgenic reporter line for which a 3288-bp *sox5* promoter fragment drives the expression of mCherry. **a-c** Between stages 18 and 22, *sox5* mRNA localizes predominantly in the head and tail bud regions of the embryos. **c,d** At stage 22, additional expression is detected in the lateral plate mesoderm of the embryos (arrows). **e**–**g'** From stage 24 onward, sox5 expression spans over the dorsal neural tube and pre-migratory neural crest cells (arrowheads). **g,g'** At stage 29, *sox5* expression is also seen in ventral migrating neural crest cells (arrowheads). **h** and **i** Fluorescent *sox5* expression is monitored in the neural tube and neural crest cells of hatching embryos (stages 38/39). **h** For comparison, *patched2* highlights the notochord at stage 39 [11]

This revealed that SOX5 was expressed in peritubular myoid cells surrounding cords in the fetal testis (Additional file 4: Figure S3, left and middle panels) and in a subset of somatic and germ cells in the fetal ovary (Additional file 4: Figure S3, right panels). Of note, a substantial fraction of SOX5 expression also displayed cytoplasmic localization. Although reported for other SOX factors (see SOX2 [25], SOX9 [26, 27], and [28] for reviews), the functional significance of the SOX5 non-exclusive nuclear localization in mice gonads remains to be investigated.

Expression of sox5 in adult gonads

Given reduced *dmrt1bY* expression in the fully developed testis [29], we also monitored the expression of *sox5* in fully mature gonads of both sexes (Fig. 6). In adult testes, *sox5* fluorescence was mainly restricted to the interlobular cells (see Fig. 6a–d). In contrast, *dmrt1bY* expression is clearly localized within the Sertoli cells [17] of the testicular lobules (Fig. 6a). No co-localization of *dmrt1bY* and *sox5* transcripts whatsoever was observed (Fig. 6a). To define the nature of the interlobular *sox5*-positive cells better, immunofluorescence of 11-β-hydroxylase protein, a specific marker of Leydig cells, was performed (Fig. 6b). Interstitial cells of Leydig are found adjacent to the seminiferous lobules in the testes and produce androgens [30]. A perfect co-localization of *sox5* transcripts and 11-β-hydroxylase protein confirmed that the interlobular *sox5*-positive cells are indeed Leydig cells (Fig. 6c,d). Interestingly, another discrete population of *sox5*-positive cells is discernable between but close to the lobules (Fig. 6e–g). These cells are very small compared to their neighboring germ cells, and their nuclei appear compact (Fig. 6h–j). The expression of the germ-cell marker *vasa* (Fig. 6e–j) assigns these cells to the germ-cell lineage, which are probably at a very early stage of differentiation. These germ cells do not express *dmrt1bY* (Fig. 6k–n). The adult ovary displays only a very few s*ox5*-positive cells with small and condensed nuclei (Fig. 6o–q). At present, the identity of these cells is difficult to ascertain, but oocytes and somatic supporting cells (granulosa or theca cells) can be excluded from their morphology and location.

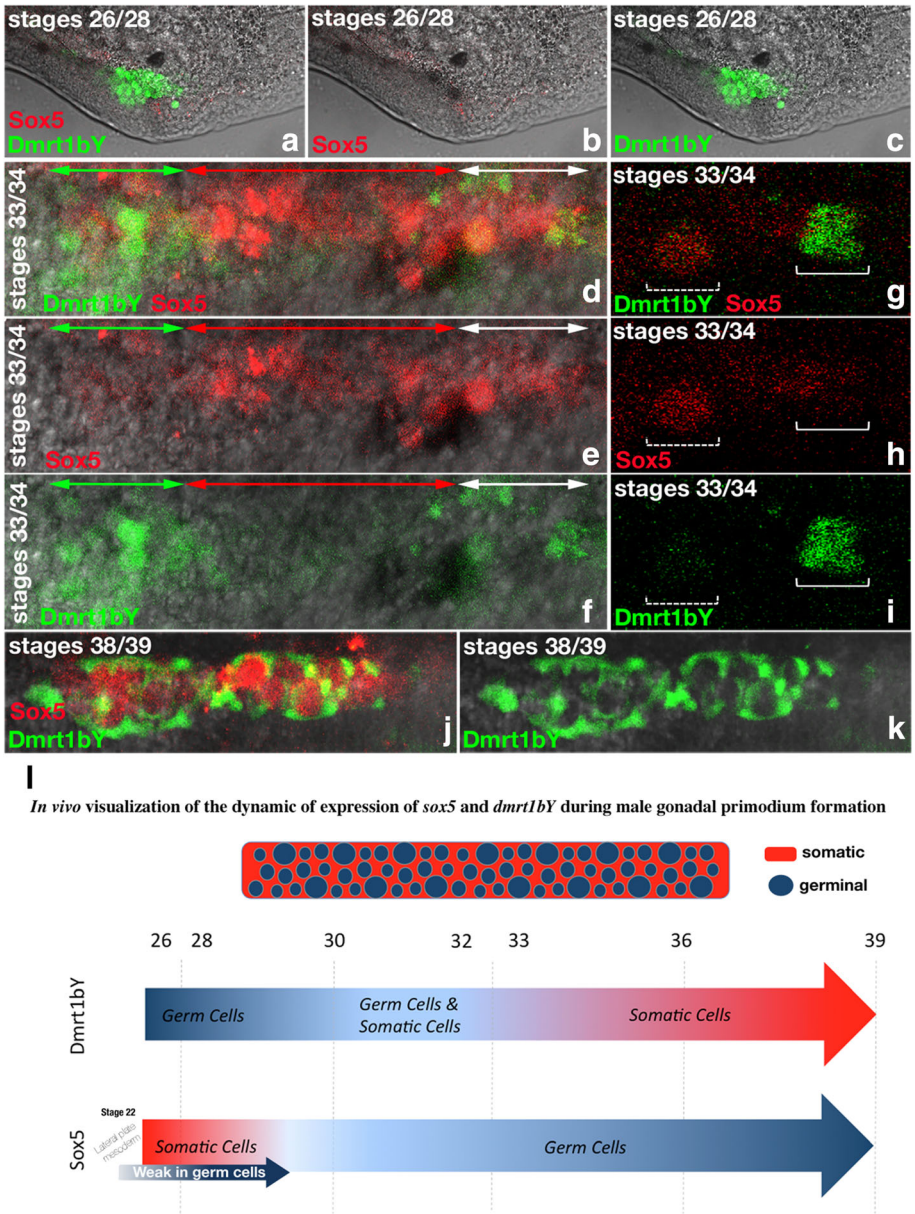


Fig. 5 Comparative analysis of *sox5* and *dmrt1bY* expression dynamics during gonadal primordium formation. Expression of *sox5* compared to *dmrt1by* in a double transgenic fluorescent reporter line. **a**–**c** During early gonadal formation, *sox5* is first detected in the somatic tissues surrounding the germ cells at stages 26 to 28. At the same time, *dmrt1bY* is expressed in germ cells. **d**–**i** By stages 33 to 34, *sox5* expression becomes restricted to the germ cells. *dmrt1bY* is also expressed in the germ cells at those specific stages of development. Variations within the respective levels of sox5 and dmrt1bY expression are clearly observable (**d** compared to **e** and **f** and **g** compared to **h** and **i**). **j**,**k** Around hatching (stages 38/39), the expression of *sox5* strengthens in all germ cells while parallel *dmrt1bY* expression quickly switches from germ cells only to somatic germ-cell-surrounding cells only. **l** In vivo visualization of the dynamics of expression localization of *sox5* and *dmrt1bY* during male gonadal primordium development. The expression of *sox5* and *dmrt1bY* is highly dynamic during primordium gonadal formation, switching from somatic to germ cells and vice versa, respectively, from stage 26 until hatching. Being mutual repressors of each other, a seesaw of expression is observed, finally finely restricting *dmrt1bY* expression in the somatic part of the primordium gonad. Blue and red represent cellular expression localizations only and should not been interpreted as expression levels

PGC number is decreased in sox5 mutants

To obtain functional data on the role of *sox5* in gonadal development, we next analyzed early gonadal development in mutants. The *ml-3* mutant (N541S) is a naturally occurring mutation, for which a premature stop codon results in the production of a truncated Sox5 protein lacking the HMG box domain and causing Sox5 loss of function [21, 31, 32]. Considering that *sox5* is first expressed in the lateral plate mesoderm and then in germ cells during

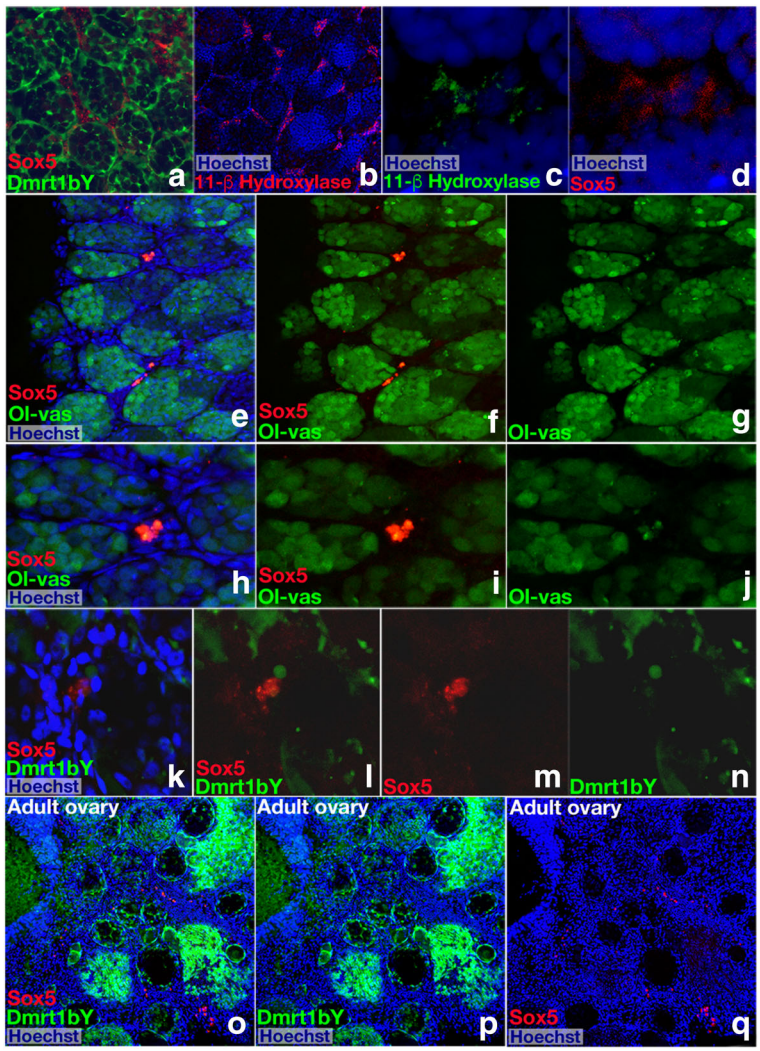


Fig. 6 Expression of *sox5* in adult gonads. **a**–**d** In adult testes, sox5 fluorescence is restricted to the cells located between the lobules where the germ and Sertoli cells lie. **a**,**b** In double transgenic reporter fish, fluorescence of sox5 and dmrt1bY (marking the Sertoli cells) is distinct. **c, d** The interlobular expression of *sox5* co-localizes with 11-β-hydroxylase, a marker of Leydig cells. **e**–**j** Expression of sox5 is also detected in another discrete population of cells between but always close to the germinal lobules. **h**-**j** Small in size, these *sox5*-positive cells also express *vasa*, a specific marker of germ cells. **k**–**n** The sox5- and vasa-positive cells do not express *dmrt1bY*. **o**–**q** In adult ovaries, only extremely few sox5 positive cells are detected.

primordial gonad formation, we investigated whether Sox5 plays a role in regulating PGC number. Whole-mount in situ hybridizations utilizing the PGC-specific marker *vasa* were performed in wild-type and *sox5*$^{-/-}$ mutant fish embryos (Fig. 7a,b). At stage 22, during the formation of the primordial gonad when *sox5* has a first expression peak in wild-type embryos, a drastic reduction in the PGC number is evident in mutants (Fig. 7a compared to b). This emphasizes a possible role for Sox5 as being a regulator of PGC proliferation although such a reduction in PGC numbers might also be ascribed to reduced proliferation, reduced survival, or defects in fate specification.

PGC number is rescued in sox5 mutants by in vivo conditional knock-in of Sox5

Given that absence of functional Sox5 expression in the *sox5*$^{-/-}$ mutant resulted in a reduced number of PGCs, we next attempted to rescue the gonadal phenotype by wild-type *sox5* expression to show that this gene is crucial in regulating PGC numbers. Thus, we established a transgenic line that expresses *sox5* after controlled homologous recombination in the *sox5* mutant genetic background (Fig. 7c). In this line, the *sox5* promoter drives the expression of a mCherry-stop cassette flanked by *LoxP* sites. This cassette is followed by the wild-type *sox5* open reading frame (ORF) fused to a 2A self-cleaving system [33] with nuclear-addressed green

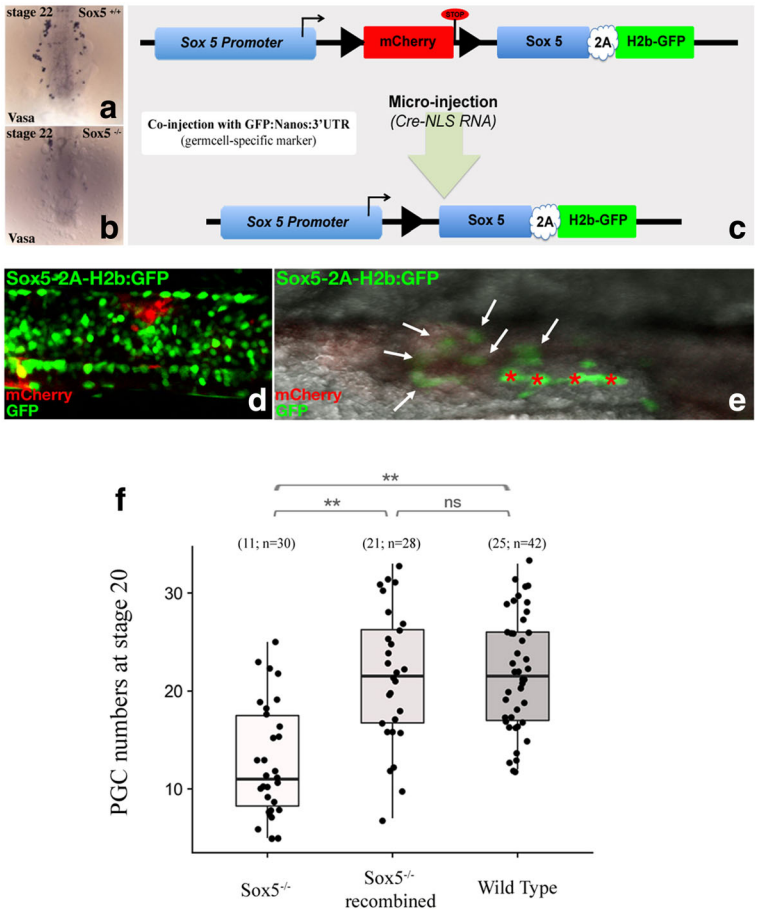


Fig. 7 Regulation of PGC numbers by Sox5. **a**,**b** As early as stage 22, a drastic reduction of the germ-cell number (*vasa* in situ hybridization) is observed in sox5$^{-/-}$ mutants compared to wild-type embryos. **c** For conditional knock-in and rescue of the *sox5*$^{-/-}$ mutant fish, a transgenic line expressing *sox5* was produced. See "Methods" for details. **d**,**e** In vivo visualization of the effective recombination and expression of *sox5* is apparent after a switch from red cytoplasmic to green nuclear-localized fluorescence. After *Cre recombinase* injection at the one-cell stage, an almost total recombination is observed, leading to the expression of the *sox5* transcript as monitored by green fluorescence. Stars indicate auto-fluorescent pigment cells and arrows indicate recombined germ cells. **f** Germ-cell numbers in *sox5* mutant medaka after in vivo recombination and expression of *sox5*. Of note, for that specific experiment, embryos were additionally injected with a GFP-Nanos 3′ untranslated region (UTR) mRNA construct allowing effective PGC monitoring. Statistical significance was assessed with the Wilcoxon–Mann–Whitney test ($N = 30$, 28, and 42 for Sox5$^{-/-}$, Sox5$^{-/-}$ recombined, and wild-type embryos, respectively; * p value ≤ 0.05, ** p value ≤ 0.01). GFP green fluorescent protein, ns non-significant, PGC primordial germ cell

fluorescent protein (GFP) (H2B-GFP) (Fig. 7c). After injection of mRNAs encoding for the *Cre-NLS* protein and effective recombination, the mCherry-stop cassette is excised, and the *sox5* ORF is expressed together with nuclear-localized GFP (see recombination in Fig. 7d for the trunk and 7e for germ cells). We employed this system as being preferable to conventional overexpression because it allows us to bypass the deleterious effect of overexpressing the pleiotropic *sox5* gene during early development. Using this system, we show that in terms of germ-cell numbers, a partial rescue of the Sox5$^{-/-}$ phenotype at stage 20 occurred in the *sox5*$^{-/-}$ mutant genetic background (Fig. 7f). This provides evidence that Sox5 is indeed required for controlling PGC numbers during the formation of the early gonadal primordium as early as stage 20.

Female-to-male sex reversal of Sox5 mutant fish

Since germ-cell number was reduced in the *sox5*$^{-/-}$ mutant fish, we further investigated whether sexual development was affected. Phenotypic and genotypic sex was determined in a *sox5*$^{-/-}$ mutant line [31, 34]. A complete XX female-to-male sex reversal (up to 95%) was recorded (Fig. 8a). The female-to-male sex-reversed phenotypic males were fully fertile, so this *sox5*$^{-/-}$ mutant line could be maintained exclusively on a XX genotypic background (see Fig. 8b for phenotypes). Genotypic XY *sox5*$^{-/-}$ mutant fish were never detected in the progeny (Fig. 8b) nor in outcrosses.

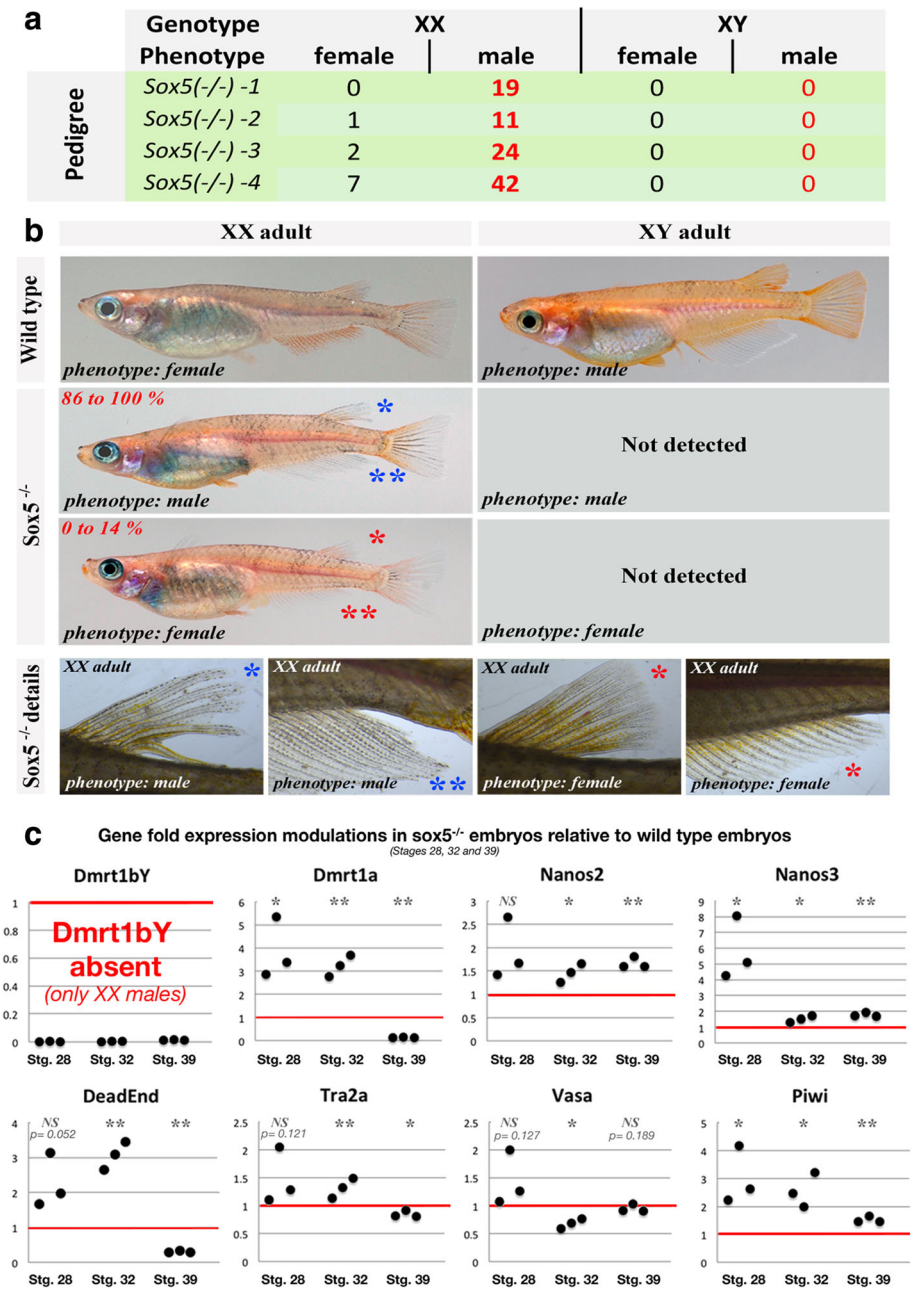


Fig. 8 Phenotypic versus genotypic sex of *sox5-/-*mutants and regulation of the *dmrt1* co-orthologs and a set of germ-cell markers in *sox5-/-* mutant embryos. **a** Phenotypic versus genotypic sex of *sox5*$^{-/-}$ mutant fish. Complete XX female-to-male sex reversion was obtained. **b** Sexual phenotype of the adult medaka. Wild-type XX females have a triangular-shaped anal fin as well as fused dorsal fin rays. Wild-type XY males have a parallelogram-shaped anal fin as well as split dorsal fin rays. **c** Regulation of the expression of *dmrt1bY* and other germ-cell markers (*nanos2*, *nanos3*, *dead-end*, *vasa*, *tra2a*, and *piwi*) in *sox5*$^{-/-}$ mutants compared to wild-type embryos at different stages of development (stages 28, 32, and 39). Dataset results of three different batches of eggs obtained from different couples. Statistical significance was assessed with the *t*-test ($N = 3$ and each replicate is a pool of 25 eggs; * p value ≤ 0.05, ** p value ≤ 0.01). ns non-significant

The absence of XY *sox5*$^{-/-}$ male-to-female sex-reversed fish, predicted to overexpress *dmrt1bY* according to our results, is in line with the scarcity of surviving YY zygotes—carrying two copies of the *dmrt1bY* gene—reported in the literature [35].

Monitoring gene expression in the XX embryos of that line, we find the upregulation of PGC marker genes, including *nanos2*, *nanos3*, *dead-end*, and *piwi*, whereas *tra2a* and *vasa* did not show significant changes compared to wild-type embryos (Fig. 8c). Of note is the

upregulation of the autosomal *dmrt1a* in the Sox5 mutant background. The precocious expression of this gene at early stages of development (stages 28 and 32 in Fig. 8c) is intriguing because *dmrt1a* expression is not expected before 10 days after hatching. This untimely expression may be related to the XX female-to-male sex reversal, because a similar untimely expression of dmrt1a has been seen in high-temperature-induced XX male-to-female sex reversals [36] (and our own unpublished data).

Discussion

SD relies on the proper control of a hierarchically structured, multilayered network of genes. The genes at the top orchestrate complex transcriptional and post-transcriptional regulations (see [3, 4, 6, 37] for reviews). Despite such a critical function, they appear to be dispensable in evolutionary terms and can be quickly replaced with the emergence of new lineages. Our present analysis provides evidence that concomitantly to the acquisition of a dominant position within the SD network, the medaka master male determiner, *dmrt1bY*, was subjected to a profound rearrangement of its regulatory landscape. We found that sequential insertions of both *Izanagi* and the *Rex1* transposon were instrumental for rewiring the *dmrt1bY* promoter in the process of diversification from its autosomal progenitor *dmrt1a*.

First, the integration site itself appears to be highly relevant. We have previously reported that a *P*-element-like DNA transposon, *Izanagi*, brought in a regulatory sequence that mediates specific transcriptional regulation of *dmrt1bY*, which was important for the Y-chromosomal duplicate to evolve its new function [17]. A common feature of class II transposons is that they can excise. The insertion of the *Rex1* element in the *dmrt1bY* promoter occurred in the DNA-binding domain of the transposase and thereby fixed the *Izanagi* element and the contained Dmrt1 transcription factor binding motif to the promoter of the new SD gene.

Second, the *Rex1* transposon contributed a functional high-affinity binding site for the transcription factor Sox5 as a novel regulatory element for *dmrt1bY* expression. Thus far, neither Sox5 in vertebrates nor its *Drosophila* homolog Sox102F [38] have been shown to be implicated in SD. In medaka, the expression pattern of *sox5* already indicated a function in gonad formation. Indeed, independently of *dmrt1bY*, *sox5* is expressed in the lateral plate mesoderm that later gives rise to the gonad. Thereafter, during gonadogenesis stages, *sox5* expression switches toward the germ-cell lineage. There the *sox5* expression pattern in PGCs is mutually exclusive with the expression of the master SD gene *dmrt1bY*, which, during the formation of the gonad primordium, concurrently changes from germ cell to somatic cell expression during this period. This is consistent with the in vitro findings of a suppressive action of Sox5 on the *dmrt1bY* promoter, although the early expression of *sox5* in the lateral plate mesoderm as well as in the adult gonads suggests other additional gonadal functions decorrelated from *dmrt1bY* activity.

It has been shown that Dmrt1bY has a suppressive effect on cell proliferation by mediating a G2 arrest [39]. Thus, continued expression of *dmrt1bY* in PGCs during early embryonic stages, which precede the actual SD stage at hatching, could have a negative effect on the number of PGCs that is presumed to be decisive at the SD stage [40, 41]. Clearly the observed suppressive action of sox5 toward *dmrt1bY* expression in vitro and in vivo at stages 33/34, down-modulates this negative effect.

Later, the persistence of *sox5* expression—independently of *dmrt1bY* expression—in germ cells within the gonadal primordium and in the early-differentiated germ cells of mature gonads indicates another independent major role in germ-cell physiology from gonad induction to adult development and maintenance. Hence, it is likely that following the transcriptional rewiring of *dmrt1bY* first by *Izanagi* and then by *Rex1* TEs, Sox5 has been hijacked in the primary SD cascade for controlling and fine-tuning *dmrt1bY* expression during the male-determining period. Independent of the regulatory function of *dmrt1bY*, Sox5 appears to have a more general involvement during gonadal formation (visible by expression in the lateral plate mesoderm) and germ-cell physiology (apparent from persistence of *sox5* expression in germ cells).

Although not directly related to *dmrt1bY* regulation during the formation of the male gonadal primordium, the most convincing evidence for a sexual development function of Sox5 comes from medaka strains that carry knockout alleles for this gene. We find that lack of Sox5 leads to a decrease in PGC numbers, which is rescued by re-introducing the wild-type version of the gene in mutant embryos. Strikingly, at the molecular level, we found an upregulation of several germ-cell markers in the mutants, even though the germ-cell number is sensibly reduced. It can be assumed that such overexpression in mutants is an indication of an insufficient compensatory mechanism needed to rescue germ-cell numbers properly. These findings demonstrate that germ-cell marker expression levels upon Sox5 modulation are primarily the result of gene expression regulation and are not due to the number of cells that express these genes.

The XX female-to-male sex reversal in the Sox5 mutant is in line with an inferred important function of maintaining the appropriate number of PGCs. The number of PGCs at the SD stage is critical for determining

male or female sex in medaka [39, 42]. It is higher in females at the SD stage. Lowering the number of germ cells in medaka or zebrafish results in female-to-male sex reversal [43, 44]. Hence, when the number of PGCs falls below the threshold in sox5$^{-/-}$ mutant XX fish, it will permit male development. Interestingly, we find an ectopic and earlier than normal expression of *dmrt1a* in the primordial gonad of the mutants. This has also been observed in environmentally induced XX female-to-male sex reversal in medaka and has been interpreted as a compensatory mechanism to supply the necessary trigger for testis development in the absence of *dmrt1bY* [36, 45].

Sox5 regulation of *dmrt1bY* and the importance of this gene for sexual development in medaka raise the question of whether this co-option of Sox5 regulation through *Rex1* insertion brought a novel member into the SD regulatory network as a medaka-specific evolutionary innovation or whether this event provided a necessary connection to an indispensable gonad-development downstream pathway. This is difficult to answer at present but should motivate further studies on the role of Sox5 in the formation of ovaries and testes in other species.

The in vitro data of *sox5* effects on *dmrt1* transcriptional regulation in zebrafish and wrasse [18, 19] point to an evolutionarily conserved function of Sox5 that have been unnoticed so far. Members of the SOX family of transcription factors play essential roles during SD in mammals. Both the founding member of the SRY family and the closely related factor SOX9 have been shown to be necessary and sufficient for mammalian male SD [46, 47]. In addition, other SOX family members, such as SOX3 and SOX10, can take over this role if they are expressed ectopically in the developing testis at the time of SD, as demonstrated in transgenic mice and in human patients with duplications in these genes [48–53]. In contrast, *Sox5* has been implicated in spermatogenesis in the adult [54, 55] but not in embryonic gonad development and/or SD in mouse. With all necessary notes of caution, our preliminary data detecting SOX5 expression in the fetal gonad of mice may indicate an additional role for SOX5 during embryonic gonad development in mice after SD.

The finding that a preformed transcription factor binding site contributed by the *Rex1* transposon modulates the regulation of *dmrt1bY* promoter highlights the important role that mobile elements play in the genome for shaping the evolution of new functions. Intriguingly, although bona fide examples of this process are still rare [16], *Rex1* is the second such event found in the same promoter. It will be interesting to analyze whether the other repeats present in the *dmrt1bY*, but not in the promoter region of *dmrt1a*, provide further instances of TE exaptation. Genes that arose by gene duplication such as *dmrt1bY* are primarily dispensable and can only escape degeneration through sub- or neo-functionalization. As *dmrt1bY* and *dmrt1a* both have exclusive functions in male sexual development in line with the highly conserved role of *dmrt1* in invertebrates and vertebrates [56], a change in transcriptional control via the insertion of two different TEs might initially have led to sub-functionalization; *dmrt1bY* acquired its transient early expression, whereas the transcription of *dmrt1a* was pushed back to the later testis differentiation phase. In other fish species, and in mouse and chicken, *dmrt1*, which represents the evolutionary precursor of the two genes in medaka, is expressed starting in very early male SD stages and continues to be expressed during testis differentiation and specialization and in the post-pubertal reproductively active organ [6, 37, 57–59].

Conclusion

In summary, the evolutionary history of the promoter of a newly arising SD gene in medaka not only provides a new example for TE-mediated rewiring that created evolutionary novelty but also shows the unexpected complexity and richness of such elements. It will be interesting to have a closer look at the SD genes of other fish that have been subject to fast evolutionary change and thus, might also be targets for TE exaptation.

In addition to showing that *sox5* was recruited—or more exactly promoted thanks to neo-functionalization—to the very top of the primary SD cascade after insertion of *Rex1* and that it controls the fine-tuning of *dmrt1bY* expression, our results provide evidence for a more general and ancestral SD function of Sox5 in regulating germ-cell number and, in consequence, gonadal identity.

Methods

Immunohistochemistry

Testes from adult fish were fixed with 4% paraformaldehyde in balanced salt solution (111 mM NaCl, 5.37 mM KCl, 1 mM $CaCl_2H_2O$, 0.6 mM $MgSO_47H_2O$, and 5 mM Hepes, pH 7.3) for 30 minutes on ice. After fixation, samples were washed three times for 10 minutes with MABT buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5, and 0.1% Triton X-100) and subsequently twice for 30 minutes with MABDT buffer (MABT buffer complemented with 1% bovine serum albumin and 1% dimethyl sulfoxide). After blocking in MABDT-blocking buffer (MABDT buffer supplemented with 2% lamb or sheep serum), the tissues were incubated in MABDT-blocking buffer together with anti-11-β-hydroxylase primary antibody (1:150 dilution) overnight at 4 °C. Samples were then washed three times for 5 minutes in MABDT buffer and washed again four times for 30 minutes in MABDT-blocking buffer on ice. Thereafter, samples were incubated overnight at 4 °C with

the secondary antibody diluted at 1:600 in MABDT-blocking buffer. Finally, the tissues were washed in phosphate-buffered saline, stained with Hoechst solution for 3 hours at 4 °C, mounted and imaged with a confocal microscope (Nikon C1 confocal microscope). For Additional file 4: Figure S3, immunohistochemistry was performed according to [60] using the mouse anti-SOX5 antibody from Abcam (ab26041).

Chromatin immunoprecipitation

For ChIP, the EpiQuik ChIP kit (Epigentek) was used according to the manufacturer's instructions, using 1 million cells (Sg3 spermatogonial or OLF fibroblast cell lines) of transiently transfected cells with FLAG-tagged Sox5 and 3XFLAG antibodies for immunoprecipitation. After fixation and cell re-suspension, DNA was sheared by sonication (nine pulses of 10 seconds with an amplitude of 10%). After immunoprecipitation, specific primer sets were used for enrichment quantification by real-time PCR. For controls, primer sets encompassing regions without any sox binding sites were used. The results are presented as enrichment compared to input. All primer sets were checked for the specificity of the amplifications.

Bioinformatic analyses

Binding sites for Dmrt1bY were identified using the matrix provided by [61] together with the Regulatory Sequence Analysis Tools portal (RSat) [62]. Sox5 transcription factor binding sites were determined using MatInspector from the Genomatix portal [63] using the following positional weight matrix: [A, C, G, T: (4, 6, 3, 9), (7, 4, 3, 8), (21, 0, 1, 1), (22, 0, 1, 0), (0, 22, 0, 1), (23, 0, 0, 0), (22, 1, 0, 0), (0, 0, 0, 23), (10, 3, 6, 4), (5, 7, 6, 5)].

In vitro expression regulation analyses and real-time PCR

Medaka spermatogonial (Sg3) and fibroblast-like (OLF) cell lines were cultured as previously described [64–66]. For transfection, cells were grown to 80% confluency in six-well plates and transfected with 5 μg of expression vector using FuGene (Roche) reagent as described by the manufacturer.

Total RNA was extracted from fish tissues or transfected cells using the TRIZOL reagent (Invitrogen) according to the supplier's recommendation. After DNase treatment, reverse transcription was performed with 2 μg of total RNA using a RevertAid First Strand Synthesis kit (Fermentas) and random primers. Real-time quantitative PCR was carried out with SYBR Green reagents and amplifications were detected with an i-Cycler (Biorad). All results are averages of at least two independent real-time reactions. Error bars represent the standard deviation of the mean. Relative expression levels were calculated (according to $2^{-\Delta CT}$ where CT is the number of cycles) after correction of the expression of elongation factor 1 alpha (ef1alpha).

Luciferase assay

For promoter analysis, a 9107-bp fragment upstream of the Dmrt1bY ORF was isolated by restriction enzyme digestion (*XhoI/EcoRI*) from BAC clone Mn0113N21 [17] and cloned into pBSII-ISceI plasmid (pBSII-ISceI::[0/-8927] Kb Dmrt1bY. Subsequently, the Gaussia luciferase gene from the pGLuc-basic plasmid (New England Biolabs) was inserted between *EcoRI* and *NotI* sites of pBSII-ISceI:: [0/-8927] Kb Dmrt1bY prom (pBSIIISceI:: [0/-8927] Kb Dmrt1bY prom::GLuc plasmid, Fig. 2). pBSII-ISceI:: [0/-1593] Kb Dmrt1bY prom::GLuc, pBSII-ISceI:: [0/-2963] Kb Dmrt1bY prom::GLuc and pBSII-ISceI:: [0/-6162] Kb Dmrt1bY prom::-GLuc plasmids were constructed in the same way, removing 5′ fragments of the 9107-bp Dmrt1bY promoter region using *Kpn1*, *Eco47III*, and *HindIII* restriction enzyme digestion, respectively, and re-ligation.

Gaussia luciferase activity was quantified using the Luciferase Reporter Assay System from Promega and normalized against co-transfected firefly luciferase expressing plasmid (ptkLUC+). For Fig. 2c, [α], [α]-MUT, and [β] fragments (same as used for the ChIP assay shown in Fig. 2a) were PCR amplified and cloned into ptkLUC+ plasmid (accession number AF027128) between *HindIII* and *BamHI* restriction sites.

Establishment of transgenic reporter lines and in vivo recombination and imaging

A transgenic line was created for the in vivo visualization of endogenous sox5 expression as well as in vivo functional knock-in of sox5. The Sox5 upstream promoter region of the Sox5 gene was cloned in front of an [mCherry-stop] cassette flanked with LoxP sites (Fig. 7c). In detail, in a first line, the sox5 promoter region drives the expression of an mCherry-Stop cassette. This cassette is followed by a Sox5 ORF fused (2A self-cleaving system [33]) with a nuclear-addressed GFP (H2B-GFP). For recombination, direct microinjection of one-cell-stage embryos with mRNA encoding for the *Cre-NLS* protein was performed (Fig. 7d–f). Sox5 $^{-/-}$ homozygosity of the fish was determined according to the pigmentation pattern of the embryos (see [21]).

To generate stable transgenic lines, the meganuclease protocol was used [67]. Briefly, approximately 10 to 15 pg of total vector DNA in a volume of 500 pl injection solution containing *I-SceI* meganuclease was injected into the cytoplasm of one-cell-stage medaka embryos (Carbio strain). Adult F0 fish were mated to each other and the offspring were tested for the presence of the transgene by checking for fluorescence. Siblings

from positive F1 fish were raised to adulthood and tested again for fluorescence.

For PGC visualization and counting, the GFP-nos1 3′ UTR construct that includes the mmGFP5 ORF cloned upstream of the 3′ UTR of the zebrafish nanos1 gene [68, 69] was injected at the one-cell stage (Fig. 7f).

For imaging, embryos, hatchlings, or tissues were mounted with 1.2% low melting temperature agarose. Confocal pictures and image stacks were acquired using Nikon C1 (eclipse Ti) confocal laser scanning and the NIS element AR software.

Whole-mount in situ hybridization

RNA whole-mount in situ hybridizations were performed as previously described [70]. Hybridization signals were detected using alkaline phosphatase conjugated anti-DIG antibody (Roche) and BM-purple (Roche) as chromogen.

Bioresources and animals

The Sox5 medaka mutant strain (N541S) has been deposited with the National Bioresource Center [71].

Additional files

Additional file 1: Figure S1. Annotation of the *dmrt1bY* promoter. Bold underlined: *Rex1* element. Bold red: location of sox5 binding sites. *Dmrt1bY* exon 0 is in blue letters. Primer sets used for ChIP are provided. (DOCX 169 kb)

Additional file 2: Figure S2. Comparison of medaka Rex1 reverse-transcriptase (RT) sequence with other non-LTR retrotransposons. RT conserved domains are given according to Malik, Burke, and Eickbush [72]. RT sequences are CR1 from *Gallus gallus* (U88211); Maui from *Fugu rubripes* (AF086712); Jockey, R1, and I from *Drosophila melanogaster* (P21328, X51968, and M14954, respectively); and Tad1 from *Neurospora crassa* (L25662). The degree of amino acid conservation between sequences is shown at the foot of the alignment. (JPG 811 kb)

Additional file 3: Table S1. Location and adjacent genes of *Rex1* elements containing Sox5 binding sites in the medaka genome. (PDF 70 kb)

Additional file 4: Figure S3. SOX5 protein expression in fetal mouse gonads. Double immunofluorescence of SOX5 (green) and MVH (red) on sagittal sections of 13.5 dpc (left panel) and 14.5 dpc (middle panel) mouse testes, as well as 14.5 dpc mouse ovaries (right panel). Gonads are demarcated with dotted lines. The lower panels are a higher magnification image of the area marked by a square in the upper panels. Scale bars 100 μm (upper panels) and 30 μm (lower panels). (JPG 1710 kb)

Additional file 5: Supporting data. Raw supporting data for Figs. 2A,B1–B4,C, 3A,B, 7 F, and 8C. (XLS 215 kb)

Acknowledgements

The authors thank the fish facilities at the University of Würzburg and at Institut national de la recherche agronomique (INRA) Rennes (C. Duret and F. Borel).

Funding

This work was supported by the Deutsche Forschungsgemeinschaft by a grant (SCHA 408/12-1; HE 7135/2-1) to AH and MS as well as Crédits Incitatifs Phase 2015/Emergence to AH. AH was additionally funded by the project AquaCRISPR (ANR-16-COFA-0004-01). This publication was funded by the German Research Foundation (DFG) and the University of Würzburg in the funding program Open Access Publishing.

Availability of data and materials

Datasets generated and analyzed during the current study are included in this published article and its Additional files 1, 2, 3, 4, and 5. The Sox5 medaka mutant strain (N541S) has been deposited with the National Bioresource Center (http://www.nbrp.jp/localeAction.do?lang=en). All materials generated during the current study are available from the corresponding author on reasonable request.

Authors' contributions

AH, MS, HH, and YG undertook the conceptualization. SS, YW, YN, DW, LC, CB, BM, and CS developed the methodology. AH, MS, DW, and YG did the formal analysis. AH, SS, YW, CB, BM, and CS performed the investigation. AH, MS, DW, and YG wrote the manuscript. AH and MS were responsible for the supervision, project administration, and funding acquisition. All authors read and approved the final manuscript.

Ethics approval

All animals used in this research were treated following the guidelines of the Comité d'Ethique en Expérimentation Animale (CEEA).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Physiological Chemistry, Biocenter, University of Würzburg, 97074 Würzburg, Germany. [2]Comprehensive Cancer Center Mainfranken, University Hospital, 97080 Würzburg, Germany. [3]Texas Institute for Advanced Study and Department of Biology, Texas A&M University, College Station, TX 77843, USA. [4]Bioscience and Biotechnology Center, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan. [5]INRA, UR1037 Fish Physiology and Genomics, F-35000 Rennes, France. [6]Department of Anatomy & Neuroscience, University of Melbourne, Parkville, Victoria 3010, Australia. [7]Centre for Organismal Studies (COS), University of Heidelberg, Heidelberg, Germany.

Received: 9 October 2017 Accepted: 11 January 2018
Published online: 29 January 2018

References

1. Bachtrog D, et al. Sex determination: why so many ways of doing it? PLoS Biol. 2014;12:e1001899.
2. Lin Y-T, Capel B. Cell fate commitment during mammalian sex determination. Curr Opin Genet Dev. 2015;32:144–52.
3. Herpin A, Schartl M. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 2015;16:1260–74.
4. Herpin A, Schartl M. Regulatory putsches create new ways of determining sexual development. EMBO Rep. 2008;9:966–8.
5. Herpin A, Schartl M. Molecular mechanisms of sex determination and evolution of the Y-chromosome: insights from the medakafish (*Oryzias latipes*). Mol Cell Endocrinol. 2009;306:51–8.
6. Herpin A, Schartl M. Dmrt1 genes at the crossroads: a widespread and central class of sexual development factors in fish. FEBS J. 2011;278:1010–9.
7. Martínez P, et al. Genetic architecture of sex determination in fish: applications to sex ratio control in aquaculture. Front Genet. 2014;5:340.
8. Pers D, et al. Global analysis of dorsoventral patterning in the wasp *Nasonia* reveals extensive incorporation of novelty in a regulatory network. BMC Biol. 2016;14:63.
9. Herpin A, Schartl M. Sex determination: switch and suppress. Curr Biol. 2011; 21:R656–9.
10. Rice AM, McLysaght A. Dosage-sensitive genes in evolution and disease. BMC Biol. 2017;15:78.
11. Herpin A, et al. Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream

regulatory network of vertebrate sex determination. Mol Biol Evol. 2013;30:2328–46.
12. Franchini LF, Pollard KS. Human evolution: the non-coding revolution. BMC Biol. 2017;15:89.
13. Britten RJ, Davidson EH. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q Rev Biol. 1971;46:111–38.
14. Britten RJ, Kohne DE. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science. 1968;161:529–40.
15. Gould SJ, Vrba ES. Exaptation—a missing term in the science of form. Paleobiology. 1982;8:4–15.
16. de Souza FSJ, Franchini LF, Rubinstein M. Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong? Mol Biol Evol. 2013;30:1239–51.
17. Herpin A, et al. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet. 2010;6:e1000844.
18. Gao S, et al. Molecular cloning, expression of Sox5 and its down-regulation of Dmrt1 transcription in zebrafish. J Exp Zoolog B Mol Dev Evol. 2005;304:476–83.
19. Jeong H-B, et al. Isolation and characterization of Dmrt1 and its putative regulatory region in the protogynous wrasse, *Halichoeres tenuispinis*. Gene. 2009;438:8–16.
20. Nakamura S, et al. Identification and lineage tracing of two populations of somatic gonadal precursors in medaka embryos. Dev Biol. 2006;295:678–88.
21. Nagao Y, et al. Sox5 functions as a fate switch in medaka pigment cell development. PLoS Genet. 2014;10:e1004246.
22. Nishimura T, et al. Analysis of a novel gene, Sdgc, reveals sex chromosome-dependent differences of medaka germ cells prior to gonad formation. Dev Camb Engl. 2014;141:3363–9.
23. Herpin A, Nakamura S, Wagner TU, Tanaka M, Schartl M. A highly conserved cis-regulatory motif directs differential gonadal synexpression of Dmrt1 transcripts during gonad development. Nucleic Acids Res. 2009;37:1510–20.
24. Saito D, et al. Proliferation of germ cells during gonadal sex differentiation in medaka: insights from germ cell-depleted mutant zenzai. Dev Biol. 2007;310:280–90.
25. Baltus GA, et al. Acetylation of Sox2 induces its nuclear export in embryonic stem cells. Stem Cells. 2009;27:2175–84.
26. Gasca S, et al. A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of Sox9 during sexual determination. Proc Natl Acad Sci U S A. 2002;99:11199–204.
27. Topol L, Chen W, Song H, Day TF, Yang Y. Sox9 inhibits Wnt signaling by promoting β-catenin phosphorylation in the nucleus. J Biol Chem. 2009;284:3323–33.
28. Malki S, Boizet-Bonhoure B, Poulat F. Shuttling of SOX proteins. Int J Biochem Cell Biol. 2010;42:411–6.
29. Hornung U, Herpin A, Schartl M. Expression of the male determining gene Dmrt1bY and its autosomal coorthologue Dmrt1a in medaka. Sex Dev. 2007;1:197–206.
30. Wainwright EN, Wilhelm D. The game plan: cellular and molecular mechanisms of mammalian testis development. Curr Top Dev Biol. 2010;90:231–62.
31. Tomita H. The lists of the mutants and strains of the medaka, common gambusia, silver crucian carp,goldfish, and golden venus fish maintained in the Laboratory of Freshwater Fish Stocks, Nagoya University. The Fish Biology Journal Medaka. 1992;v.4:5–47.
32. Kelsh RN, et al. The Tomita collection of medaka pigmentation mutants as a resource for understanding neural crest cell development. Mech Dev. 2004;121:841–59.
33. Szymczak AL, Vignali DAA. Development of 2A peptide-based strategies in the design of multicistronic vectors. Expert Opin Biol Ther. 2005;5:627–38.
34. Romanish MT, Lock WM, van de Lagemaat LN, Dunn CA, Mager DL. Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution. PLoS Genet. 2007;3:e10.
35. Yamamoto TO. The problem of viability of YY zygotes in the medaka, *Oryzias latipes*. Genetics. 1964;50:45–58.
36. Hattori RS, et al. Temperature-dependent sex determination in Hd-rR medaka *Oryzias latipes*: gender sensitivity, thermal threshold, critical period, and Dmrt1 expression profile. Sex Dev. 2007;1:138–46.
37. Herpin A, Schartl M. Vertebrate sex determination: questioning the hierarchy. FEBS J. 2011;278:1001.
38. Li A, et al. Silencing of the *Drosophila* ortholog of Sox5 leads to abnormal neuronal development and behavioral impairment. Hum Mol Genet. 2017;26:1472–82.
39. Herpin A, et al. Inhibition of primordial germ cell proliferation by the medaka male determining gene Dmrt I bY. BMC Dev Biol. 2007;7:99.
40. Tzung K-W, et al. Early depletion of primordial germ cells in zebrafish promotes testis formation. Stem Cell Rep. 2015;5:156.
41. Tanaka M, Saito D, Morinaga C, Kurokawa H. Cross talk between germ cells and gonadal somatic cells is critical for sex differentiation of the gonads in the teleost fish, medaka (*Oryzias latipes*). Dev Growth Differ. 2008;50:273–8.
42. Herpin A, et al. Specification of primordial germ cells in medaka (*Oryzias latipes*). BMC Dev Biol. 2007;7:3.
43. Kurokawa H, et al. Germ cells are essential for sexual dimorphism in the medaka gonad. Proc Natl Acad Sci U S A. 2007;104:16958–63.
44. Slanchev K, Stebler J, de la Cueva-Méndez G, Raz E. Development without germ cells: the role of the germ line in zebrafish sex differentiation. Proc Natl Acad Sci U S A. 2005;102:4074–9.
45. Sato T, Endo T, Yamahira K, Hamaguchi S, Sakaizumi M. Induction of female-to-male sex reversal by high temperature treatment in medaka, *Oryzias latipes*. Zoolog Sci. 2005;22:985–8.
46. Kashimada K, Koopman P. SRY: the master switch in mammalian sex determination. Development. 2010;137:3921–30.
47. Zhao L, Quinn A, Ng ET, Veyrunes F, Koopman P. Reduced Activity of SRY and its target enhancer Sox9-TESCO in a mouse species with X*Y sex reversal. Sci Rep. 2017;7:41378.
48. Laronda MM, Jameson JL. Sox3 functions in a cell-autonomous manner to regulate spermatogonial differentiation in mice. Endocrinology. 2011;152:1606–15.
49. Mizuno K, et al. Gene expression profile during testicular development in patients with SRY-negative 46,XX testicular disorder of sex development. Urology. 2013;82:1453.e1–7.
50. Moalem S, et al. XX male sex reversal with genital abnormalities associated with a de novo Sox3 gene duplication. Am J Med Genet A. 2012;158A:1759–64.
51. Polanco JC, Wilhelm D, Davidson T-L, Knight D, Koopman P. Sox10 gain-of-function causes XX sex reversal in mice: implications for human 22q-linked disorders of sex development. Hum Mol Genet. 2010;19:506–16.
52. Sukumaran A, Desmangles J-C, Gartner LA, Buchlis J. Duplication of dosage sensitive sex reversal area in a 46, XY patient with normal sex determining region of Y causing complete sex reversal. J Pediatr Endocrinol Metab. 2013;26:775–9.
53. Sutton E, et al. Identification of Sox3 as an XX male sex reversal gene in mice and humans. J Clin Invest. 2011;121:328–41.
54. Denny P, Swift S, Connor F, Ashworth A. An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. EMBO J. 1992;11:3705–12.
55. Daigle M, Roumaud P, Martin LJ. Expressions of Sox9, Sox5, and Sox13 transcription factors in mice testis during postnatal development. Mol Cell Biochem. 2015;407:209–21.
56. Matson CK, Zarkower D. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. Nat Rev Genet. 2012;13:163–74.
57. Ayers KL, Cutting AD, Roeszler KN, Sinclair AH, Smith CA. Dmrt1 is required for Müllerian duct formation in the chicken embryo. Dev Biol. 2015;400:224–36.
58. Lindeman RE, et al. Sexual cell-fate reprogramming in the ovary by Dmrt1. Curr Biol. 2015;25:764–71.
59. Shan Z, et al. Sex-specific expression of an evolutionarily conserved male regulatory gene, Dmrt1, in birds. Cytogenet Cell Genet. 2000;89:252–7.
60. Rastetter RH, et al. Marker genes identify three somatic cell types in the fetal mouse ovary. Dev Biol. 2014;394:242–52.
61. Murphy MW, Zarkower D, Bardwell VJ. Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. BMC Mol Biol. 2007;8:58.
62. Regulatory Sequence Analysis Tools portal. http://www.rsat.eu/. Accessed 19 Jan 2018.
63. MatInspector of the Genomatix portal. http://www.genomatix.de/. Accessed 19 Jan 2018.
64. Zhang X, et al. Autosomal Gsdf acts as a male sex initiator in the fish medaka. Sci Rep. 2016;6:19738.
65. Thoma EC, et al. Ectopic expression of single transcription factors directs differentiation of a medaka spermatogonial cell line. Stem Cells Dev. 2011;20:1425–38.

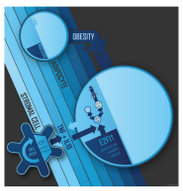
66. Hong Y, et al. Establishment of a normal medakafish spermatogonial cell line capable of sperm production in vitro. Proc Natl Acad Sci U S A. 2004; 101:8011–6.
67. Grabher C, Wittbrodt J. Meganuclease and transposon mediated transgenesis in medaka. Genome Biol. 2007;8 Suppl 1:S10.
68. Köprunner M, Thisse C, Thisse B, Raz E. A zebrafish nanos-related gene is essential for the development of primordial germ cells. Genes Dev. 2001;15: 2877–85.
69. Herpin A, et al. Sequential Sdf1a and b-induced mobility guides medaka PGC migration. Dev Biol. 2008;320:319–27.
70. Klüver N, Herpin A, Braasch I, Driessle J, Schartl M. Regulatory back-up circuit of medaka Wt1 co-orthologs ensures PGC maintenance. Dev Biol. 2009;325: 179–88.
71. National BioResource Project. http://www.nbrp.jp/localeAction.do?lang=en. Accessed 19 Jan 2018.
72. Malik HS, Burke WD, Eickbush TH. The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol. 1999;16:793–805.

Autophagy

Volume 11 • Issue 11 • November 2015

Editor-in-Chief
Daniel J. Klionsky
University of Michigan

Taylor & Francis
Taylor & Francis Group



Autophagy

ISSN: 1554-8627 (Print) 1554-8635 (Online) Journal homepage: http://www.tandfonline.com/loi/kaup20

CMA restricted to mammals and birds: myth or reality?

Laury Lescat, Amaury Herpin, Brigitte Mourot, Vincent Véron, Yann Guiguen, Julien Bobe & Iban Seiliez

To cite this article: Laury Lescat, Amaury Herpin, Brigitte Mourot, Vincent Véron, Yann Guiguen, Julien Bobe & Iban Seiliez (2018): CMA restricted to mammals and birds: myth or reality?, Autophagy, DOI: 10.1080/15548627.2018.1460021

To link to this article: https://doi.org/10.1080/15548627.2018.1460021

Accepted author version posted online: 21 Jun 2018.

Publisher: Taylor & Francis & Taylor and Francis Group, LLC

Journal: *Autophagy*

DOI: 10.1080/15548627.2018.1460021

CMA restricted to mammals and birds: myth or reality?

Laury Lescat[1,†], Amaury Herpin[2,†], Brigitte Mourot[2], Vincent Véron[1], Yann Guiguen[2], Julien Bobe[2], Iban Seiliez[1,*]

[1] INRA-UPPA, UMR1419 Nutrition Metabolisme Aquaculture, F-64310 St-Pée-sur-Nivelle, France.

LPGP, INRA, 35000 Rennes, France.

***Corresponding author:**

Iban Seiliez

INRA-UPPA, UMR 1419 Nutrition, Métabolisme, Aquaculture

F-64310 Saint Pée sur Nivelle, France.

Tel: (33) 5 59 51 59 99; Fax: (33) 5 59 54 51 52

e-mail : iban.seiliez@inra.fr

[†]These authors contributed equally to this work

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

ABSTRACT

Chaperone-mediated autophagy (CMA) is a major pathway of lysosomal proteolysis essential for the control of intermediary metabolism. So far, the absence of any identifiable LAMP2A—a necessary and limiting protein for CMA—outside of the tetrapod clade, led to the paradigm that this cellular function was (presumably) restricted to mammals and birds. However, after we identified expressed sequences displaying high sequence homology with the mammalian *LAMP2A* in several fish species, our findings challenge that view and suggest that CMA likely appeared much earlier during evolution than initially thought. Hence, our results do not only shed an entirely new light on the evolution of CMA, but also bring new perspectives on the possible use of complementary genetic models, such as zebrafish or medaka for studying CMA function from a comparative angle/view.

Keywords: chaperone-mediated autophagy; evolution; fish; LAMP2A; RNA-seq

Being one of the main pathways involved in lysosomal proteolysis, **c**haperone-**m**ediated **a**utophagy (CMA) has been described as a selective mechanism for the degradation of specific soluble proteins within lysosomes [1]. Besides a well described role in protein quality control (resulting from its ability to selectively target damaged or non-functional proteins for degradation), the diversity of the sub-proteome degraded by CMA also associates this function with the regulation of transcriptional programs [2], cell death and cell survival mechanisms [3–5], DNA repair and cell cycle progression [6], as well as a variety of intracellular processes related to the control of cellular energetics [7–10]. Over the last few years, CMA has thus emerged as a major core component in the control of cellular homeostasis [11].

In detail, cytosolic proteins bearing a KFERQ-like motif are first recognized by the chaperone HSPA8/HSC70 [12]. The substrate/chaperone complex then docks at the lysosomal membrane through specific binding to the cytosolic tail of LAMP2A (lysosome-associated membrane protein 2A). Multimerization of LAMP2A will then result in the formation of a translocation complex, and promote translocation of substrate proteins [13]. Following unfolding and internalization, substrate proteins are then rapidly degraded by lysosomal proteases. Next, LAMP2A disassembles from the translocation complex, allowing a new cycle of substrate binding and translocation [14].

LAMP2A originates from the alternative splicing of the *LAMP2* gene, giving rise to 3 different splice variants. These splice variants all share a common lumenal domain but display different cytosolic and transmembrane regions. CMA activity is tightly correlated with (*i*) the level of LAMP2A (and not those of the 2 other splice variants) at the lysosomal membrane [15] and (*ii*) the efficiency of assembly/disassembly of LAMP2A in this compartment [14]. As such, LAMP2A is considered to be the necessary and limiting component for CMA activity [15].

In this context, because LAMP2A has, so far, been characterized only in birds and mammals, and not in other clades, functional CMA is thought to be restricted to tetrapods [16]. Interestingly, in *Drosophila*, selective endosomal microautophagy (eMI), a recently identified form of microautophagy which shares together with CMA the dependence on KFERQ-like motifs and HSPA8/HSC70 for substrate targeting, has been suggested to constitute an alternative to CMA [16,17]. Although it is tempting to speculate that in non-tetrapod species eMI might be an ancestral form of selective autophagy for the degradation of substrates that in tetrapods are shared between eMI and CMA, our data reveal that the picture is probably much more complex.

Indeed, homology-based searches on in house-developed RNA-seq databases (PhyloFish), providing consistent and exhaustive gene expression data from 23 different ray-finned fish species [18], resulted in the identification of several contigs displaying high sequence homology with mammalian *LAMP2A*. In detail, the inferred amino acid sequences of the fish Lamp2A contigs not only displayed high homology when compared to the transmembrane and cytoplasmic domains of mammalian LAMP2A, but also high conservation of key motifs shown to be essential for proper function of the protein (Figure 1A). Hence, fish Lamp2A sequences display the typical GY dipeptide conserved in all forms of LAMP2 and required for targeting to lysosomes [19]. Fish Lamp2A present also the hydrophobic phenylalanine (F) residue (not conserved in the other forms of Lamp2) important for lysosomal targeting [20]. Additionally, 3 of the 4 positively charged amino acids necessary for binding substrate proteins [15] are also present in fish sequences, as are the 2 glycine residues (G) involved in the multimeric pattern of Lamp2A [13], at least in some fish species (Figure 1A).

Furthermore, sequence analysis of the complete genomes of several fish species also revealed high conservation of the genomic organization of the *LAMP2* gene across vertebrates

with notably the presence of the 3 alternative exons (B, A and C) encoding the transmembrane domain and cytoplasmic tail specific of each isoform (LAMP2B, LAMP2A and LAMP2C, respectively) (Figure 1B). However, the number and size of exons encoding the lumenal region are only moderately conserved among species (Figure 1B), leading to lower homology of the corresponding region (from 15 to 20% between analyzed species) compared to the cytoplasmic tail (from 21 to 85%) (Table S1), thus accounting for different structure (Figure 1A) and possibly functional variations of Lamp2A among species.

Finally, fine expression analysis of the newly characterized *lamp2a* splice variants in different fish species shows that they are expressed in different tissues of a large variety of ray-finned fish species, including the medaka (*Oryzias latipes*), a model species widely used in biomedical research (Figure 1C). Although no *lamp2a* transcripts could be identified in zebrafish (*Danio rerio*), RT-PCR analysis using specific primers targeting a conserved region of the *lamp2a* exon from fish nevertheless revealed significant zebrafish *lamp2a* expression in different tissues, including intestine, kidney and liver (data not shown).

Overall, our data show for the first time the existence as well as the expression of *lamp2a* transcripts in different tissues of a large variety of ray-finned fish species, and therefore imply that CMA function might have appeared much earlier during evolution than initially thought. A number of issues remain now to be addressed about the functionality of the Lamp2A sequences found in fish. This includes (i) whether or not in fish *lamp2a* mRNA can be successfully translated to protein, (ii) whether or not these Lamp2A proteins properly localize to lysosomes, and (iii) whether or not the polymerized fish Lamp2A regulates the relocation of the substrates. Further understanding of the structure-function relationship between fish Lamp2A displaying differences in key motifs will also help to complete this picture. The presence of Lamp2A outside of the tetrapod clades opens up new perspectives in autophagy research. Comparative approaches across phylogenetic distant species (fish *versus*

tetrapods for instance) will certainly provide new insights on selective autophagy by exploring the extent to which CMA, but also the crosstalk between CMA and other components of the cellular proteostasis networks (in particular eMI), diverged during vertebrate evolution.

ACKNOWLEDGEMENTS

This study was supported by the INRA "Animal Physiology and Livestock Systems" Division, and the French National Research Agency (ANR-17-CE20-0033 "Fish-and-Chap"). We thank AM Cuervo (Albert Einstein College of Medicine, NY) for helpful advice.

REFERENCES

[1] Kaushik S, Cuervo AM. Chaperone-mediated autophagy: a unique way to enter the lysosome world. Trends Cell Biol 2012; 22(8):407-17; PMID: 22748206; doi: 10.1016/j.tcb.2012.05.006.

[2] Valdor R, Mocholi E, Botbol Y, et al. Chaperone-mediated autophagy regulates T cell responses through targeted degradation of negative regulators of T cell activation. Nat Immunol 2014; 15(11):1046-54; PMID: 25263126; doi: 10.1038/ni.3003.

[3] Ferreira JV, Fofo H, Bejarano E, et al. STUB1/CHIP is required for HIF1A degradation by chaperone-mediated autophagy. Autophagy 2013; 9(9):1349-66; PMID: 23880665; doi: 10.4161/auto.25190.

[4] Xie W, Zhang L, Jiao H, et al. Chaperone-mediated autophagy prevents apoptosis by degrading BBC3/PUMA. Autophagy 2015; 11(9):1623-35; PMID: 26212789; doi: 10.1080/15548627.2015.1075688.

[5] Yang Q, She H, Gearing M, et al. Regulation of neuronal survival factor MEF2D by chaperone-mediated autophagy. Science 2009; 323(5910):124-7; PMID: 19119233; doi: 10.1126/science.1166088.

[6] Park C, Suh Y, Cuervo AM. Regulated degradation of Chk1 by chaperone-mediated autophagy in response to DNA damage. Nat Commun 2015; 6:6823; PMID: 25880015; doi: 10.1038/ncomms7823.

[7] Kaushik S, Cuervo AM. Degradation of lipid droplet-associated proteins by chaperone-mediated autophagy facilitates lipolysis. Nat Cell Biol 2015; 17(6):759-70; PMID: 25961502; doi: 10.1038/ncb3166.

[8] Lv L, Li D, Zhao D, et al. Acetylation targets the M2 isoform of pyruvate kinase for degradation through chaperone-mediated autophagy and promotes tumor growth. Mol Cell 2011; 42(6):719-30; PMID: 21700219; doi: 10.1016/j.molcel.2011.04.025.

[9] Schneider JL, Suh Y, Cuervo AM. Deficient chaperone-mediated autophagy in liver leads to metabolic dysregulation. Cell Metab 2014; 20(3):417-32; PMID: 25043815; doi: 10.1016/j.cmet.2014.06.009.

[10] Xia HG, Najafov A, Geng J, et al. Degradation of HK2 by chaperone-mediated autophagy promotes metabolic catastrophe and cell death. J Cell Biol 2015; 210(5):705-16; PMID: 26323688; doi: 10.1083/jcb.201503044.

[11] Tasset I, Cuervo AM. Role of chaperone-mediated autophagy in metabolism. FEBS J 2016; 283(13):2403-13; PMID: 26854402; doi: 10.1111/febs.13677.

[12] Chiang HL, Terlecky SR, Plant CP, et al. A role for a 70-kilodalton heat shock protein in lysosomal degradation of intracellular proteins. Science 1989; 246(4928):382-5; PMID: 2799391; doi: 10.1126/science.2799391.

[13] Bandyopadhyay U, Kaushik S, Varticovski L, et al. The chaperone-mediated autophagy receptor organizes in dynamic protein complexes at the lysosomal membrane. Mol Cell Biol 2008; 28(18):5747-63; PMID: 18644871; doi: 10.1128/MCB.02070-07.

[14] Bandyopadhyay U, Sridhar S, Kaushik S, et al. Identification of regulators of chaperone-mediated autophagy. Mol Cell 2010; 39(4):535-47; PMID: 20797626; doi: 10.1016/j.molcel.2010.08.004.

[15] Cuervo AM, Dice JF. Unique properties of lamp2a compared to other lamp2 isoforms. J Cell Sci 2000; 113 Pt 24:4441-50; PMID: 11082038.

[16] Tekirdag KA, Cuervo AM. Chaperone-mediated autophagy and endosomal microautophagy: joint by a chaperone. J Biol Chem 2017; PMID: 29247007; doi: 10.1074/jbc.R117.818237.

[17] Mukherjee A, Patel B, Koga H, et al. Selective endosomal microautophagy is starvation-inducible in Drosophila. Autophagy 2016; 12(11):1984-1999; PMID: 27487474; doi: 10.1080/15548627.2016.1208887.

[18] Pasquier J, Cabau C, Nguyen T, et al. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC Genomics 2016; 17:368; PMID: 27189481; doi: 10.1186/s12864-016-2709-z.

[19] Williams MA, Fukuda M. Accumulation of membrane glycoproteins in lysosomes requires a tyrosine residue at a particular position in the cytoplasmic tail. J Cell Biol 1990; 111(3):955-66; PMID: 2391371.

[20] Guarnieri FG, Arterburn LM, Penno MB, et al. The motif Tyr-X-X-hydrophobic residue mediates lysosomal membrane targeting of lysosome-associated membrane protein 1. J Biol Chem 1993; 268(3):1941-6; PMID: 8420968.

[21] Wilke S, Krausze J, Bussow K. Crystal structure of the conserved domain of the DC lysosomal associated membrane protein: implications for the lysosomal glycocalyx. BMC Biol 2012; 10:62; PMID: 22809326; doi: 10.1186/1741-7007-10-62.

FIGURE LEGEND

Figure 1. Protein Structure, gene organization and mRNA expression of Lamp2A in fish

(**A**) Schematic drawing of the structure of vertebrate LAMP2A. The human LAMP2A is used as reference. It shows a lumenal region comprising 2 N-glycosylated LAMP domains of approximately 160 residues (each with 2 disulfide bonds, S-S) separated by a proline-rich, O-glycosylated 'hinge' region of approximately 30 amino acid residues [21]. O- and N-linked glycosylation are indicated in green and red, respectively. The transmembrane (TM) domain, harboring 2 glycine residues (red G) involved in the multimeric pattern of LAMP2A, is followed by a short, C-terminal cytosolic tail that is comprised of 4 positively-charged amino acids (KHHH in blue) required for the binding of substrate proteins as well as motifs for lysosomal targeting (in green). The schematic drawing of fish Lamp2A has been done on the basis of sequence complementarity with the human LAMP2A. Potential O- and N-linked glycosylation are indicated in light green and pink, respectively. Sequence alignment of the boxed region of the 3 LAMP2/Lamp2 variants is shown below. The positively-charged amino acids required for the binding of substrate proteins are colored in blue. The GY dipeptide as well as the hydrophobic F required for targeting of LAMP2A to lysosomes are in green. The glycine residues (G) involved in the multimeric pattern of LAMP2A are in red. (**B**) The genomic structure of *LAMP2*/*lamp2* is conserved in vertebrates and contains the 3 alternative exons (B, A and C) encoding the transmembrane domain and cytoplasmic tail specific for each isoform. The size of exons (in base pairs) is shown in italics below or above each exon. (**C**) Data from RNAseq show that *lamp2a* is expressed in different tissues of a large number of ray-finned fish. Relative expression of *lamp2a* was expressed in number of reads per kilobase per million reads per species, after normalization of data by the total number of sequences obtained for each tissue and species. The obtained values were then log

transformed and centered to the median (set at 0.00). Ac, *Amia calva* (bowfin); Lo, *Lepisosteus oculatus* (spotted gar); Aa, *Anguilla anguilla* (European eel); Gp, *Gnathonemus petersi* (elephantnose fish); Aal, *Alosa alosa* (allis shad); Ph, *Pangasianodon hypophthalmus* (striped Catfish); Am, *Astyanax mexicanus* (cave Mexican tetra); El, *Esox lucius* (northern pike); Up, *Umbra pygmae* (eastern mudminnow); Tt, *Thymallus thymallus* (grayling); Cl, *Coregonus lavaretus* (European whitefish); Cc, *Coregonus clupeaformis* (American whitefish); St, *Salmo trutta* (brown trout); Om, *Oncorhynchus mykiss* (rainbow trout); Sf, *Salvelinus fontinalis* (brook trout); Gm, *Gadus morhua* (Atlantic cod); *Ol*, *Oryzias latipes* (medaka).

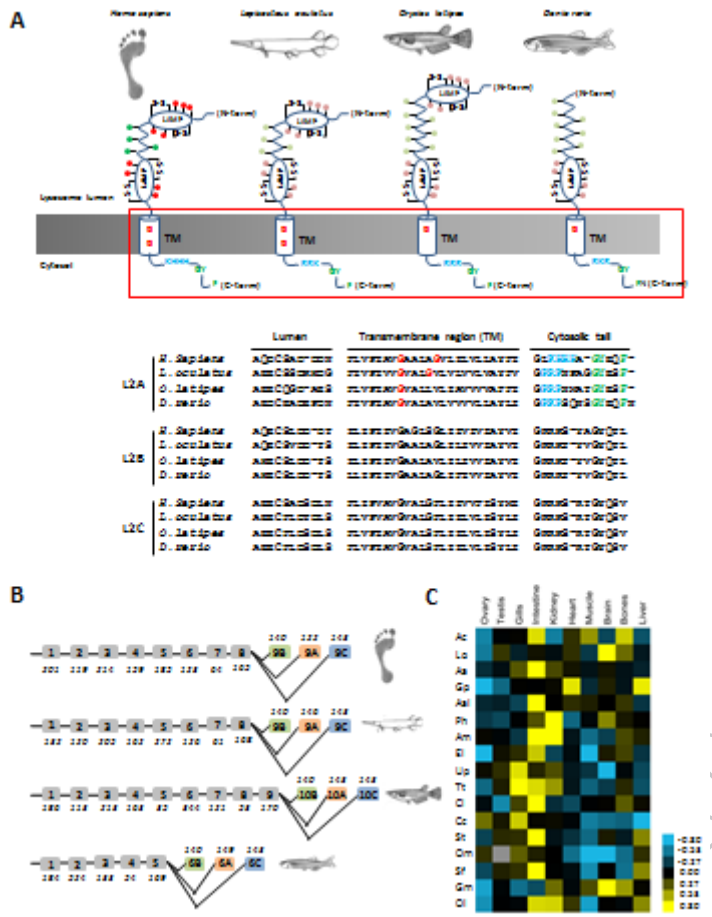


Figure 1

	H. sapiens	*L. oculatus*	*O. latipes*	*D. rerio*
Lumenal region				
H. sapiens				
L. oculatus	17.89			
O. latipes	15.24	19.76		
D. rerio	17.03	19.21	15.24	
Cytoplasmic tail				
H. sapiens				
L. oculatus	46.15			
O. latipes	38.46	84.62		
D. rerio	21.43	64.29	71.43	

table1

The unusual rainbow trout sex determination gene hijacked the canonical vertebrate gonadal differentiation pathway

Sylvain Bertho[a,b], Amaury Herpin[b], Adèle Branthonne[b], Elodie Jouanno[b], Ayaka Yano[b], Barbara Nicol[b,1], Thomas Muller[c], Maëlle Pannetier[d], Eric Pailhoux[d], Misako Miwa[e], Goro Yoshizaki[e], Manfred Schartl[a,f,g,h,2,3], and Yann Guiguen[b,2,3]

[a]Physiological Chemistry, Biocenter, University of Wuerzburg, D-97074 Wuerzburg, Germany; [b]Institut National de la Recherche Agronomique, UR1037 Fish Physiology and Genomics, F-35000 Rennes, France; [c]Julius-von-Sachs-Institute, Department of Molecular Plant Physiology and Biophysics, University of Wuerzburg, D-97082 Wuerzburg, Germany; [d]UMR Biologie du Développement et Reproduction, Institut National de la Recherche Agronomique, L'École Nationale Vétérinaire d'Alfort, Université Paris Saclay, F-78350 Jouy en Josas, France; [e]Department of Marine Biosciences, Tokyo University of Marine Science and Technology, 108-8477 Tokyo, Japan; [f]Comprehensive Cancer Center Mainfranken, University Hospital, D-97080 Wuerzburg, Germany; [g]Hagler Institute for Advanced Study, Texas A&M University, College Station, TX 77843; and [h]Department of Biology, Texas A&M University, College Station, TX 77843

Edited by Catherine L. Peichel, University of Bern, Bern, Switzerland, and accepted by Editorial Board Member Hopi E. Hoekstra October 23, 2018 (received for review March 5, 2018)

Evolutionary novelties require rewiring of transcriptional networks and/or the evolution of new gene functions. Sex determination (SD), one of the most plastic evolutionary processes, requires such novelties. Studies on the evolution of vertebrate SD revealed that new master SD genes are generally recruited from genes involved in the downstream SD regulatory genetic network. Only a single exception to this rule is currently known in vertebrates: the intriguing case of the salmonid master SD gene (*sdY*), which arose from duplication of an immune-related gene. This exception immediately posed the question of how a gene outside from the classical sex differentiation cascade could acquire its function as a male SD gene. Here we show that SdY became integrated in the classical vertebrate sex differentiation cascade by interacting with the Forkhead box domain of the female-determining transcription factor, Foxl2. In the presence of Foxl2, SdY is translocated to the nucleus where the SdY:Foxl2 complex prevents activation of the aromatase (*cyp19a1a*) promoter in cooperation with Nr5a1 (Sf1). Hence, by blocking a positive loop of regulation needed for the synthesis of estrogens in the early differentiating gonad, SdY disrupts a preset female differentiation pathway, consequently allowing testicular differentiation to proceed. These results also suggest that the evolution of unusual vertebrate master sex determination genes recruited from outside the classical pathway like *sdY* is strongly constrained by their ability to interact with the canonical gonadal differentiation pathway.

sex determination | Forkhead box proteins | sex differentiation | fish | evolution

Sexual development is a fundamental process that shapes animal morphology, physiology, and behavior. The development of the undifferentiated embryonic gonad toward a testis or an ovary is regulated by a complex network of genes, where the initial triggers for male or female sex differentiation can come from the environment or the genome (1). A great number of studies revealed that the chromosomal, molecular, and cellular mechanisms of genetic sex determination (SD) are highly variable (1–4). It is now particularly clear that SD mechanisms evolved frequently and independently, leading to a high turnover of the genes governing sexual development (1, 5), even between closely related organisms. For instance, the therian sex-determining gene *SRY* is not found in other vertebrates (6), and recent studies have identified many different master SD genes in birds, amphibians, and fish (7–14). In these species, known members of the downstream regulatory sex differentiation network usurped the position at the top of the sex determination cascade to become the master SD gene. However, not all downstream sex differentiation genes are equally able to take the lead as sexual master switches, and currently, only the genes encoding transcription factors *Sox3* and *Dmrt1* and several components of TGF-β signaling have been identified as master SD genes in vertebrates. This frequent reuse of the same SD genes led to the hypothesis that there are limited options in becoming a master sex-determining gene, which can be met by only a very limited number of genes from the sex differentiation network. However, this "limited option" hypothesis (15) was challenged by the discovery of the unusual salmonid sex-determining gene (16, 17). This gene, called *sdY* for "sexually dimorphic on the Y," turned out to be a duplicated and truncated version of a gene encoding IFN regulatory factor 9 (*irf9*), which functions in the immune response of vertebrates. Upon IFN binding to its receptor and activation of STAT signaling, IRF9 complexes with both STAT1 and STAT2 in the cytoplasm and then translocates to the nucleus to activate effector genes of the antiviral response through its DNA-binding domain (18). IRF9 nonimmunity

Significance

Sex determination is one of the most fundamental but also extraordinary plastic processes in nature. Many different master sex-determining genes have been characterized in vertebrates, and most of them are known to fulfill essential functions during sexual development and thus are already tightly linked to the process that they now govern. Only one exception is currently known: the salmonid master sex-determining gene (*sdY*), which arose from the duplication of an immune-related gene. Here we show that SdY prevents female differentiation by interacting and blocking the action of a key ovarian differentiation factor. These results suggest that the evolution of unusual vertebrate master sex determination genes is strongly constrained by their ability to interact with the canonical gonadal differentiation pathway.

Author contributions: M.S. and Y.G. designed research; S.B., A.H., A.B., E.J., A.Y., B.N., T.M., M.P., E.P., M.M., G.Y., M.S., and Y.G. performed research; M.P. and E.P. contributed new reagents/analytic tools; S.B., A.H., A.B., E.J., A.Y., B.N., T.M., E.P., M.M., G.Y., M.S., and Y.G. analyzed data; and S.B., A.H., M.S., and Y.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. C.L.P. is a guest editor invited by the Editorial Board.

[1]Present address: Reproductive and Developmental Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, Durham, NC 27709.

[2]M.S. and Y.G. contributed equally to this work.

[3]To whom correspondence may be addressed. Email: phch1@biozentrum.uni-wuerzburg.de or yann.guiguen@inra.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1803826115/-/DCSupplemental.

roles have been described in neurons and liver and heart pathophysiology (19), but so far no evidence implicates it in sex determination or sex differentiation processes. The birth of such a master SD gene recruited from outside of the classical sex differentiation cascade raises the intriguing question of its functional evolution and how this unusual SD gene determines sex. Did SdY evolve a new function to be able to interact directly with the classical gonadal sex differentiation cascade, or does it use part of its ancestral pathway, that is, the IFN immune-related response, for its action?

Results and Discussion

During evolution of SdY, the DNA-binding domain and the nuclear localization signals of Irf9 were lost, while the protein–protein interaction domain (IRF association domain, or IAD) was preserved and underwent some sequence diversification (16). To test the hypothesis that the IAD domain of SdY still functions in protein binding, we first performed molecular modeling and found that the 3D structure of SdY strongly overlaps the IAD domain of IRF proteins (Fig. 1*A*). As the IAD domain is the only domain of known function predicted from the primary sequence of the *sdY* gene (16, 17), we hypothesized that SdY could still exert its function based on protein–protein interactions. We thus searched for SdY interacting proteins using a yeast two-hybrid (Y2H) screen with SdY used as bait and with a rainbow trout prey cDNA library prepared from late differentiating testes sampled when *sdY* expression is still high (16). Among the 46 different putative interacting proteins there were none of the known Irf9 partners like Stat1 or Stat2. Instead, we found a very strong enrichment of many members of the Forkhead box (FOX) family (11 FOX proteins, *SI Appendix*, Tables S1–S3). The Forkhead box, a highly conserved DNA-binding domain (20) common to all FOX proteins, was identified as the minimum domain needed for an effective interaction with SdY (Fig. 1*B* and *SI Appendix*, Table S3). Interestingly, among all of the FOX proteins interacting with SdY in yeast, we found the well-known female sex differentiation protein Foxl2 (21, 22). Taking into account the importance of Foxl2 in vertebrate sex differentiation, we reasoned that this would be an interesting and biologically relevant SdY partner. We then explored the interaction of SdY with trout Foxl2 in a direct yeast interaction assay and confirmed that SdY and Foxl2 can interact together (*SI Appendix*, Fig. S1).

To better characterize this interaction, an in vitro approach was developed using cell transfection assays. In HEK 293T cells transfected only with *sdY* plasmid, SdY protein was localized predominantly in the cytoplasm (Fig. 2 *A–A″* and *SI Appendix*, Fig. S2). However, when cotransfected along with Foxl2, SdY was completely translocated into the nucleus (Fig. 2 *B–B″* and *SI Appendix*, Fig. S2). Such a complete SdY nuclear translocation was observed only with fish Foxl2 proteins (Fig. 2*H*), including the two rainbow trout paralogous gene products (Foxl2b1 and Foxl2b2) resulting from the salmonid whole-genome duplication (23) (Fig. 2 *I–O*) and the medaka, *Oryzias latipes*, Foxl2 (*SI Appendix*, Figs. S3 and S4). No complete nuclear translocation was observed with some other rainbow trout Fox proteins (Fig. 2 *C–H* and *SI Appendix*, Fig. S2), with some mammalian Foxl2, that is, mouse and goat (*SI Appendix*, Fig. S3), and with the rainbow trout Foxl2b2 containing a modified mouselike Forkhead box domain (*SI Appendix*, Fig. S4). This complete nuclear relocalization of SdY with trout and medaka Foxl2s indicated some specific protein–protein interaction and that this interaction required the conformation of a fish Forkhead domain. This interaction was also confirmed in vitro by co-immunoprecipitation experiments (Fig. 2 *P* and *Q* and *SI Appendix*, Fig. S5*A*) and in vivo by showing that SdY was also translocated into the nucleus following coinjection with Foxl2 in medaka embryos (*SI Appendix*, Fig. S5 *B* and *C*).

To obtain further insights into the physiological relevance of the SdY and Foxl2 interaction in vivo, a gene expression time course of *foxl2*, *nr5a1*, and *sdY* genes in differentiating trout gonads was performed. In agreement with its male-determining role, *sdY* expression was detected only in male gonads, with a peak of expression around 45 d postfertilization (dpf) (Fig. 3*A*). In contrast, *foxl2b1*, *foxl2b2*, and *nr5a1* were not expressed in a sexually dimorphic fashion before the time point at which *sdY* peaks in males; after this time point *foxl2b1* and *foxl2b2* are markedly up-regulated in females and down-regulated in males (Fig. 3 *A* and *B* and *SI Appendix*, Fig. S6). We also explored expression of gonadal aromatase (*cyp19a1a*), as this gene is a well-known direct target of Foxl2 (24). We found that *cyp19a1a* is expressed only in female gonads and its expression parallels the expression of the trout *foxl2* genes (Fig. 3*A*). These *cyp19a1a*, *nr5a1*, and *foxl2* expression patterns are consistent with the critical role of Foxl2 in the up-regulation of *Cyp19a1*/*cyp19a1a* (24, 25), in cooperation with Nr5a1 (steroidogenic factor 1, Sf1) for driving ovarian differentiation (26). In addition, *sdY* and *foxl2* are colocalized in some somatic cells of the early differentiating gonad in the male rainbow trout (Fig. 3*C*), suggesting that they could interact to modulate *cyp19a1a* expression. Interestingly, trout *foxl2* genes are also strongly and positively regulated by estrogens (Fig. 3*B*), the steroid end products of the aromatase enzyme (21). This points to a positive regulatory loop with Foxl2 inducing *cyp19a1a* expression and thus increasing estrogen synthesis that will, in return, stimulate the expression of *foxl2* (21). Taking into account the pivotal role of Cyp19a1a and estrogens in fish ovarian differentiation (27) and our results on a specific interaction of SdY with Foxl2, we proposed that SdY exerts its sex-determining function by suppressing this positive regulatory loop through its interaction with Foxl2. To evaluate this, we first confirmed, using a luciferase reporter assay, that activation of the medaka *cyp19a1a* promoter requires the presence of both Foxl2 and Nr5a1, which work in synergy to activate *cyp19a1a* (25) (Fig. 4*A*). We then tested the effect of SdY on the transcriptional activity of the *cyp19a1a* promoter and demonstrated that SdY strongly represses the synergistic Foxl2- and Nr5a1-induced *cyp19a1a* expression (Fig. 4*B*) but not the Foxl2-alone- or Nr5a1-alone-induced *cyp19a1a* expression (Fig. 4*C*).

In our attempt to understand how *irf9*, a transcriptional regulator of the immune system, has evolved into a master SD gene whose expression is necessary and sufficient to drive testicular differentiation (16), we obtained multiple sets of evidence supporting the hypothesis that SdY exerts its sex-determining function by interacting with the female-determining factor Foxl2. An interaction of SdY with Foxl2 was not anticipated, but some Forkhead box proteins such as the FOXO proteins have been shown to have direct protein–protein interactions with a variety of unrelated transcription

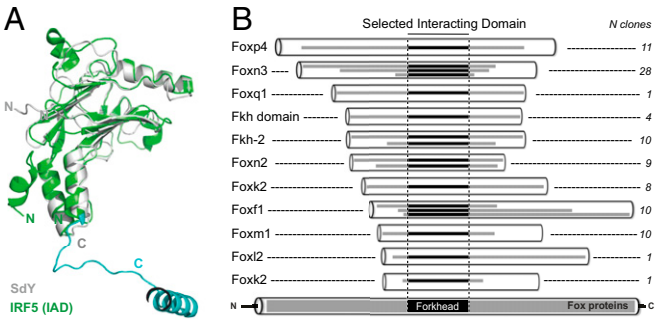


Fig. 1. SdY conserves the structure of the IRF protein–protein interaction domain and interacts with the Forkhead box domain of Fox proteins. (*A*) SdY shares structural homologies with IAD, a protein–protein interaction domain. The structure of SdY (in gray) was modeled using the crystal structure of IRF5 as a template (in green). This SdY structure reveals eight β-sheets forming a β-sandwich and three α-helices that are highly conserved with IRF5. (*B*) SdY interacts in yeast with Fox proteins through their highly conserved DNA-binding domain. The alignments of the SdY-Fox interacting clone sequences (gray lines) delineate the minimum domain or selected interacting domain needed for an effective interaction with SdY in yeast, which is the Forkhead box domain (110 aa, black lines). The 11 Fox proteins characterized in the Y2H screen are represented by open cylinders with numbers of interacting clones indicated on the right side.

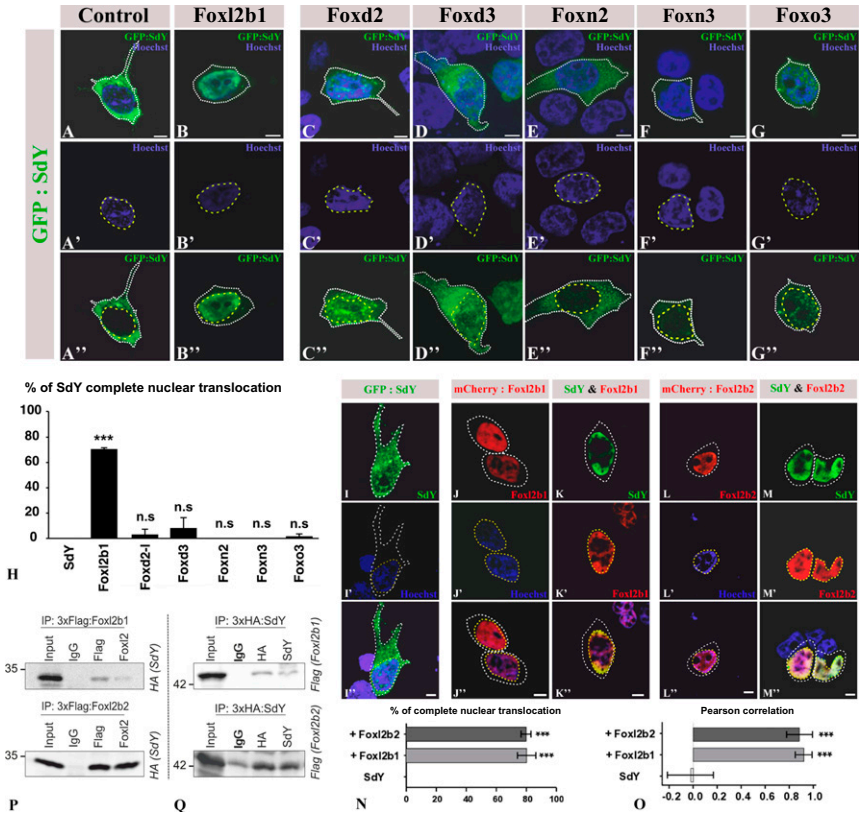


Fig. 2. SdY interacts with Foxl2, resulting in its nuclear translocation. (*A–H*) GFP:SdY alone (*A–A″*) and GFP:SdY in combination with different trout Fox proteins, Foxl2 (Foxl2b1) (*B–B″*), Foxd2 (*C–C″*), Foxd3 (*D–D″*), Foxn2 (*E–E″*), Foxn3 (*F–F″*), Foxo3 (*G–G″*), were cotransfected in HEK 293T cells (delimited by white dotted lines). GFP:SdY is translocated in the nucleus (delimited by yellow dotted lines and stained in blue with Hoechst staining) only in the presence of Foxl2 (*B–B″*). (Scale bar, 5 μm.) (*H*) Percentage of transfected cells (mean ± standard deviation on 200 cells) in which SdY is completely translocated in the nucleus after three independent cotransfection experiments with different trout Fox proteins. Significant differences compared with SdY alone were calculated using an unpaired two-tailed Student's *t* test, ***$P < 0.001$; ns, nonsignificant. (*I–O*) Foxl2b1 and Foxl2b2 are both able to drive SdY complete nuclear translocation (delimited by yellow dotted lines and stained in blue with Hoechst staining). Confocal images of HEK 293T cells (delimited by white dotted lines) transiently transfected with *sdY* (*I–I″*), mCherry:Foxl2b1 alone (*J–J″*), SdY and mCherry:Foxl2b1 (*K–K″*), mCherry:Foxl2b2 alone (*L–L″*), SdY and mCherry:Foxl2b2 (*M–M″*). (Scale bar, 10 μm.) (*N*) Quantitative analysis in the presence or absence of Foxl2b1 and Foxl2b2. Percentage of complete SdY nuclear translocation (mean ± standard deviation on 100 cells) after three independent cotransfection experiments. Statistical significances compared with SdY alone were calculated using an unpaired two-tailed Student's *t* test. (*O*) SdY colocalizes with Foxl2 in the nucleus. SdY, SdY-Foxl2b1, and SdY-Foxl2b2 colocalizations were measured in the nucleus for SdY ($n = 5$), SdY and Foxl2b1 ($n = 5$), and SdY and Foxl2b2 ($n = 5$) with Pearson's correlation. Statistical significance was calculated using an unpaired two-tailed Student's *t* test, ***$P < 0.001$. (*P* and *Q*) SdY binds with Foxl2 in co-immunoprecipitation (IP) experiments. HEK 293T cells were transiently transfected with expression plasmids for SdY fused to a hemagglutinin tag (3xHA:SdY) and for Foxl2 fused to a 3xFlag tag (3xFlag:Foxl2b1 or 3xFlag:Foxl2b2). Whole-cell lysates were used for immunoprecipitation with anti-Flag or anti-Foxl2 (*P*) and with anti-HA or anti-SdY (*Q*) followed by immunoblotting with the appropriate antibodies. Input represents 10% whole-cell lysate. IgG mouse antibody was used as the control. In *P*, 3xFlag:Foxl2b1 or 3xFlag:Foxl2b2 was immunoprecipitated with either Flag (*Top*) or FoxL2 (*Bottom*) antibodies followed by immunoblotting with an antibody against the HA tag to reveal the interaction with 3xHA:SdY (SdY). In *Q*, 3xHA:SdY was immunoprecipitated with an HA or SdY antibody, followed by immunoblotting with an antibody against the Flag tag to reveal 3xFlag:Foxl2b1 (Foxl2b1) (*Top*) or 3xFlag:Foxl2b2 (Foxl2b2) (*Bottom*).

factors (28). In zebrafish the Foxo3b protein is even able to repress the transcriptional activity of *irf* genes by a direct IRF/FOX protein–protein interaction (29). This may suggest that SdY acquired its ability to specifically bind to fish Foxl2 proteins based on a preexisting possibility of interaction of Irf9 with the FOX protein.

The specific interaction of SdY with Foxl2 carries SdY into the nucleus, where it can prevent, by blocking the synergistic action of Nr5a1 (Sf1) and Foxl2, the implementation of a positive regulatory loop controlling the expression of *cyp19a1a*. By doing so very early in the differentiation process of the gonads, i.e., long before implementation of the *cyp19a1a* loop of regulation in females (Fig. 3*A*), SdY would completely prevent estrogen production in the differentiating male gonads. The absence of estrogen production subsequently triggers masculinization as it has been demonstrated in many fish species, including rainbow trout, using treatments with an aromatase inhibitor (30) or, more recently, by direct inactivation of the *cyp19a1a* gene (26, 31, 32). Such masculinization following the blockade of estrogen production is even effective in adult females (27), showing that estrogens in fish are needed not only for ovarian differentiation but also for ovarian maintenance. However, the fact that *foxl2b1* and *foxl2b2* gene expression is not down-regulated in the male gonad before 45–50 dpf suggests that the inhibition of the positive regulatory loop between *cyp19a1a* expression, estrogen production, and *foxl2* gene expression is not active at these early testicular differentiation stages. This absence of inhibition of *foxl2* expression in the early male differentiating gonad could suggest that there is additional regulation of this positive loop or that expression of *foxl2* is not sensitive to estrogen at these early developmental stages. Such a mechanism of action through the blockage of *cyp19a1a* and estrogen production assigns to SdY an activity as an antiovarian determining factor directly preventing the ovarian differentiation pathway instead of activating the male pathway. However, it cannot be totally excluded that SdY, besides suppressing the female pathway, may also affect directly the activation of the male pathway. Nevertheless, known important male developmental actors such as Dmrt1, Amh, and Sox9 (1) were not identified in the Y2H screen, and this along with their late expression during rainbow trout male gonad development (33) compared with *sdY* expression suggests that they are not direct interacting partners of SdY.

In summary, we provide strong evidence that SdY determines sex in rainbow trout not by using part of its ancestral Irf9 pathway but by directly interacting with Foxl2, an important member of the classical gonadal sex differentiation cascade. This suggests that innovation at the top of the vertebrate sex determination cascade may be constrained because novel master SD genes have to cope with the regulation of the conserved vertebrate sex differentiation cascade. The "limited option" hypothesis is mainly based on the idea that only a small subset of genes and chromosomes, because they are better at doing the job, would be independently and repeatedly selected as new vertebrate master SD genes (15). We now propose that the limited option is actually more constrained by the conservation of the sex differentiation pathway and that evolution of SD genes may include some innovations like SdY if these unusual SD genes evolved a means to build an interface with the sex differentiation cascade.

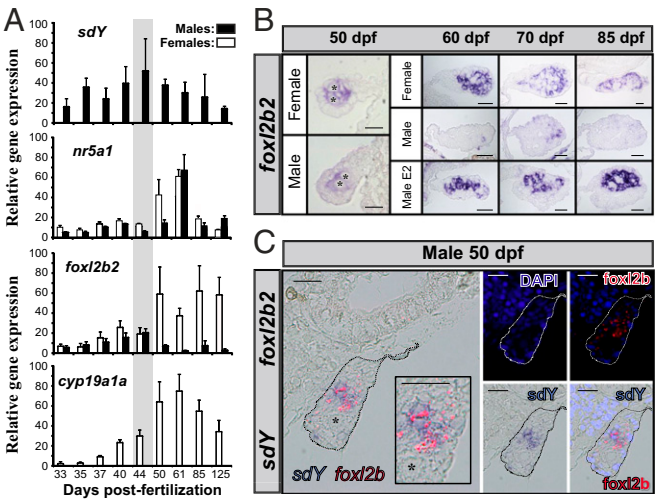


Fig. 3. The gonadal expression patterns of *sdY*, *foxl2*, *nr5a1*, and *cyp19a1a* (*A* and *B*) are in agreement with a repressive effect of SdY on the Foxl2 positive regulation of the aromatase promoter (*C*). (*A*) Gene expression profiles of *sdY*, *nr5a1*, *foxl2b2*, *and cyp19a1a* in male and female gonads from 33 to 125 dpf. All values represent the mean ± standard deviation of three biological replicates (in percentage of the highest measured value). The gray area highlights the period before the sexually dimorphic expression of trout *foxl2* genes during gonadal differentiation. (*B*) Gonadal localization of *foxl2b2* transcripts (NBT/BCIP signal in blue) in male and female gonads. *foxl2b2* is expressed in somatic cells of both female and male gonads at 50 dpf and only in female gonads at later stages (60, 70, and 85 dpf). In males fed with estrogens (male E2) *foxl2b2* is strongly up-regulated compared with control males quickly after (60 dpf) the application of the treatment (55 dpf). (Scale bar, 20 μm.) (*C*) Colocalization by double in situ hybridization of *sdY* (NBT/BCIP signal in blue) and *foxl2b2* (HNPP/Fast Red signal in red fluorescence) in somatic cells of a rainbow trout male differentiating gonad at 50 dpf. Cell nuclei are shown in the dark-field panels stained with DAPI either with or without the HNPP fluorescent detection of *foxl2b2*. Germ cells are shown by an asterisk. (Scale bar, 20 μm.)

Materials and Methods

Protein Structure Prediction. Three-dimensional homology modeling of SdY was predicted with the software SWISS-MODEL (https://swissmodel.expasy.org/) using the structure of the dimeric IRF5 (PDB ID 3DSH) transactivation domain (34) as a template. The resulting model was obtained by the superposition of the template and SdY. The 3D views of SdY were made with PyMOL (molecular graphics system, version 1.7.4; Schrödinger).

Yeast Two-Hybrid Screen. Yeast two-hybrid screening was performed by Hybrigenics Services (https://www.hybrigenics-services.com). The coding sequence for SdY (amino acids 1–215) (GenBank accession number GI:392583258) was PCR amplified and cloned into pB27_A as a C-terminal fusion to LexA (*N*-LexA-SdY-C) and into pB66_A as a C-terminal fusion to the Gal4 DNA-binding domain (*N*-Gal4-SdY-C). These constructs were checked by sequencing and used as baits to screen a random-primed *Oncorhynchus mykiss* immature male gonad (gonads sampled around 75 dpf) cDNA library; 112 million interactions were tested with pB27_SdY, and 71.3 million interactions were tested with p66_SdY, leading to the detection of 24 and 178 processed clones, respectively. The prey fragments of these 202 positive clones were amplified by PCR and sequenced at their 5′ and 3′ junctions. Each fragment corresponding to an interacting protein was identified using GenBank release 192 (National Center for Biotechnology Information). The common sequence shared by all prey fragments of the same protein defines the selected interacting domain containing all of the structural determinants required for a given interaction to occur. A confidence score [predicted biological score (PBS)] that outlines the reliability of the interaction is given to each interaction as previously described (35). PBS scores were divided into four categories, from A (highest confidence) to D (lowest confidence).

Genomic DNA Extraction. To clone rainbow trout *foxl2b1* and *foxl2b2*, genomic DNA was extracted from liver tissue with lysis buffer (10 mM Tris·HCl, 100 mm EDTA, 0.5% SDS, and 20 μg/mL RNase A). Proteinase K was added to 150 mg/mL, and the sample was incubated at 56 °C overnight. A double-extraction phenol–chloroform (1:1) followed by chloroform–isoamyl alcohol (24:1) extraction was done. DNA precipitation was performed with an equal volume of isopropanol (1:1). The precipitate was pelleted by centrifugation at 16,000 × *g* and washed twice in 70% ethanol, dried at room temperature, and dissolved in 2 mL of distilled water.

Cloning. Plasmids and primers used are listed in the *SI Appendix*, Tables S4 and S5. The coding sequence of SdY was amplified from the psdy:sdy-pcry:cfp plasmid (16) and inserted into pCS2$^+$. From this plasmid, the PCR-amplified fragment was inserted into pCS2$^+$:HA:mCherry (gift from Manfred Gessler, University of Wuerzburg, Wuerzburg, Germany), pCS2$^+$-emGFP, and pCS2$^+$-3xHA expression plasmids. Rainbow trout *foxl2b1* and *foxl2b2* and medaka *foxl2* (*Ol-foxl2*) were cloned by PCR amplification on genomic DNA and inserted in pCS2$^+$ plasmid, pCS2$^+$:HA:mCherry, and pCS2$^+$:3xFlag between the EcoRI-XhoI restriction sites. The constructs 3xHA-pCS2$^+$ and 3xFlag-pCS2$^+$ were obtained by concatemerization of three single HA sequences (3xHA) or three single FLAG sequences (3xFLAG) flanked by HindIII restrictions sites. pCS2$^+$-emGFP:SdY was obtained by inserting a PCR-amplified fragment corresponding to emGFP in frame into the EcoRI site. To explore the hypothesis that SdY could be able to interact with all Fox proteins through an interaction with their highly conserved Forkhead domain, we selected different rainbow trout Fox proteins from an EST resource collection in which ESTs were cloned into a CMV expression (pCMV-Sport6) plasmid (36). Five trout cDNA clones encoding for Foxd2, Foxd3, Foxl3, Foxn3, and Foxn2 were found to have a complete open coding frame, including Foxn3 and Foxn2, which were identified as SdY partners in the Y2H screen. The rainbow trout Foxl2b2 sequence with threonine 64 and threonine 79 replaced by adenines (T64A; T79A) was synthesized (Genscript) by replacing in the *foxl2b2* cDNA sequence the threonine codon sequence ACC or ACT with an adenine codon sequence GCC at amino acid positions 64 and 79, respectively. The corresponding sequence was inserted in a pcDNA3.1 plasmid between the EcoR1 and XhoI restriction enzyme sites and checked by sequencing.

Cell Culture. HEK 293T cells were cultured and maintained in DMEM (PAN Biotech), supplemented with 10% FCS (PAN Biotech) and 1% penicillin–streptomycin (PAN Biotech) at 37 °C with 5% CO_2. Transfections for HEK 293T cells were performed by incubating cells with polyethylenimine (PEI) (100 mg/mL PEI diluted 1:100 in 150 mM NaCl) and respective plasmids (10 μg for 10-cm dishes, 2 μg for six-well plates) for 6–8 h in fresh medium. Then, the medium was discarded, and fresh medium was added.

Rainbow trout gonadal (RTG2) cells were cultured and maintained in L15 medium, 20 mM glutamine (PAN Biotech), supplemented with 10% FCS (PAN Biotech) and 1% penicillin–streptomycin (PAN Biotech), at 20 °C in an atmosphere of air. For transfection, RTG2 cells were detached by Trypsin-EDTA (P0781; Sigma-Aldrich) and pelleted by centrifugation at 1,000 × *g* for 5 min, washed once with medium and once with PBS. The pellet was drained and resuspended in solution V (Amaxa Kit) at a density of 10^6 cells/mL; 2 μg of plasmid were added to the suspension. After mixing, the suspension was transferred to a cuvette (Kit V; Amaxa). Program D-23 was used to electroporate the cells. After transfection, cells were immediately transferred to six-well plates filled with medium. All experiments were performed 72 h after transfection.

Immunofluorescence. HEK 293T cells were seeded in a 6-well plate containing coverslips. After pCS2$^+$-meGFP-SdY and pCS2$^+$-mCherry-Foxl2b1 (or pCS2$^+$-3xFLAG:Foxl2b2 or pCS2$^+$-3xFLAG-Ol-Foxl2) cotransfection for 48 h, cells were fixed in 4% fresh paraformaldehyde for 15 min, extensively washed, and permeabilized with 0.1% Triton X-100 in PBS for 10 min. Then cells were blocked with 1% BSA for 20 min. The primary antibody (*SI Appendix*, Table S6) was incubated overnight at 4 °C. After extensive washing with PBS, cells were incubated with Alexa 488 or Alexa 594 conjugated secondary antibodies in 1% BSA for 1 h, followed by Hoechst 33342 (Invitrogen) staining for 5 min (1 μg/mL final concentration). Cells were mounted using Mowiol 4-88 (Roth). Confocal images were acquired using a Nikon Eclipse C1 laser-scanning microscope (Nikon), fitted with a 60× Nikon objective (PL APO, 1.4 N.A.), and Nikon image software. Images were collected at 1,024 × 1,024 pixel resolution. The stained cells were optically sectioned in the *z* axis. The step size in the *z* axis varied from 0.2 to 0.25 mm to obtain 50 slices per imaged file. All experiments were independently repeated at least three times.

Colocalization Analyses. The Nikon NIS-Elements imaging analysis software was used for the colocalization analyses. Confocal images of double-stained sections were first subjected to background correction. SdY nuclear translocation was counted as complete translocation when the majority of the GFP:SdY signal was found in the nucleus. Pearson's correlation was calculated and used to obtain the colocalization values as percentages of SdY overlapping with Foxl2b1 or Foxl2b2 for a minimum of five cells (n = 5). The Pearson's coefficient values were defined as very strong colocalization

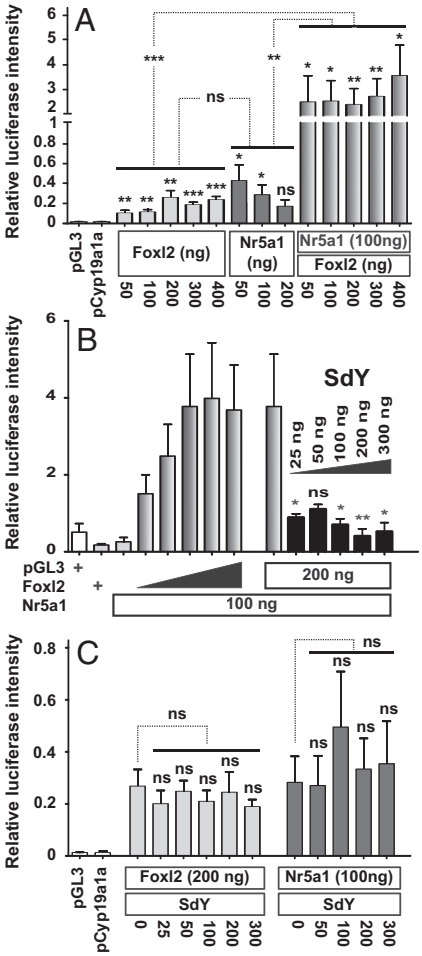


Fig. 4. SdY prevents the Foxl2/Nr5a1 positive regulation of the *cyp19a1a* promoter. Medaka *cyp19a1a* promoter activity (pGL3-*cyp19a1a* promoter coupled to a firefly luciferase) was measured using a luciferase reporter assay after HEK 293T cell cotransfection with either medaka *nr5a1* and/or *foxl2* expression plasmids and variable concentrations of the rainbow trout *sdY* expression plasmid. All results were calculated as the mean ± SEM of three biological replicates in two independent experiments. (*A*) Foxl2 and Nr5a1 act in synergy to induce *cyp19a1a* expression. Medaka *cyp19a1a* luciferase assay with variable concentrations of *foxl2* (50–400 ng) and *nr5a1* (50–200 ng) or 100 ng of *nr5a1* with variable concentrations of *foxl2* (50–400 ng). Statistical significances of luciferase activity within treatments were tested using an unpaired two-tailed Student's *t* test. Effects of *foxl2* alone and *nr5a1* alone compared with their synergistic effect (shown by asterisks on the dotted lines joining the different groups) were tested by a one-way ANOVA with a post hoc Dunnett test. (*B*) SdY prevents Foxl2/Nr5a1 positive regulation of the *cyp19a1a* promoter. Medaka *cyp19a1a* luciferase assay with variable concentrations of *foxl2* (50–400 ng) combined with a fixed concentration of *nr5a1* (100 ng) and a fixed concentration of *foxl2* (200 ng) and *nr5a1* (100 ng) combined with variable concentrations of sdY (25–300 ng). Empty plasmid control (pGL3) and Foxl2 alone (200 ng) are depicted by a + sign. Statistical significances of luciferase activity were tested using a Mann–Whitney *U* test. (*C*) SdY does not repress *cyp19a1a* promoter expression induced by medaka Foxl2 alone or Nr5a1 alone. Medaka *cyp19a1a* luciferase assay with fixed *foxl2* (200 ng) or *nr5a1* (100 ng) concentrations combined with variable concentrations of *sdY* (25–300 ng). Statistical significances of luciferase activity within treatments were tested using an unpaired two-tailed Student's *t* test. The effect of *foxl2* or *nr5a1* alone compared with *foxl2* or *nr5a1* and variable concentrations of *sdY* was tested by a one-way ANOVA with post hoc Tukey tests. ***$P < 0.001$; **$P < 0.01$; *$P < 0.1$; ns, $P > 0.05$ (nonsignificant).

between 0.85 and 1, strong colocalization between 0.5 and 0.85, and weak or no colocalization between −1 and 0.5.

Co-Immunoprecipitation. HEK-293T cells were transfected with pCS2$^+$-3xHA:SdY and pCS2$^+$-3xFlag-Foxl2b1(or pCS2$^+$-3xFlag:Foxl2b2 or pCS2$^+$-3xFlag-Ol-Foxl2) constructs to be assessed for their ability to co-immunoprecipitate. After 48 h cells were scraped and resuspended in 50 μL lysis buffer [20 mM Hepes (pH 7.8), 500 mM NaCl, 5 mM $MgCl_2$, 5 mM KCl, 0.1% deoxycholate, 0.5% Nonidet-P40, 10 mg/mL aprotinin, 10 mg/mL leupeptin, 200 mM Na_3VO_4, 1 mM phenylmethanesulfonyl fluoride, and 100 mM NaF]. Cells were incubated in lysis buffer for 30 min at 4 °C and then cleared by high-speed centrifugation for 20 min. After Bradford protein concentration measurement, HNTG buffer (20 mM Hepes pH 7.5; 150 mM NaCl; 10% glycerol; 0.1% Triton X-100) was added (1:1) to 250 μg of the whole-cell lysate. After preclearing with IgG antibodies for 1 h at 4 °C, whole-cell lysates were used for immunoprecipitation with the corresponding antibodies. One microgram of anti-Flag, anti-HA, or IgG antibody was added to 500 μL of cell lysate or 5 μg of anti-SdY or anti-FoxL2 (*SI Appendix*, Table S6) and then incubated at 4 °C overnight. After the addition of washed protein G agarose beads (Pierce, 20398), incubation in HNTG buffer was continued for another 2 h. Immunoprecipitates were washed (five times with centrifugation at 1,000 × *g*, supernatant discarded, HNTG lysis buffer added) and eluted with SDS/PAGE loading buffer by boiling for 10 min. Co-immunoprecipitation was detected by standard Western blot analysis procedure.

Western Blotting. Cells were lysed in a Hepes-based lysis buffer [20 mM Hepes (pH 7.8), 500 mM NaCl, 5 mM $MgCl_2$, 5 mM KCl, 0.1% deoxycholate, 0.5% Nonidet-P40, 10 mg/mL aprotinin, 10 mg/mL leupeptin, 200 mM Na3VO4, 1 mM phenylmethanesulfonyl fluoride, and 100 mM NaF] for 3 h. Cells debris was pelleted by centrifugation for 15 min at 16,000 × *g*. Cell lysate protein concentration was measured with a Bradford assay (Cary 50 Spectrophotometer; Varian). The protein lysates (30–50 μg) were resolved by SDS/PAGE on 12% Tris·glycine gels followed by transfer to nitrocellulose membranes. Unspecific binding was blocked with 5% BSA in Tris buffered saline with Tween-20 (TBST) [10 mM Tris (pH 7.9); 150 mM NaCl; 0.1% Tween] for 1 h at room temperature. Incubation with primary antibodies was performed overnight at 4 °C. After three washes with TBST, HRP conjugated antibodies were incubated with blocking solution for 1 h. Following the washes, membranes were incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) for 1 min. The signal from the membrane was detected using a Photo Image Station 4000MM (Kodak). At least two independent experiments were performed, and representative protein blot images are shown.

Quantitative PCR. Expression levels of *sdY*, *foxl2b1*, *foxl2b2*, *nr5a1*, and *cyp19a1a* were measured by qPCR as previously described (37). Gonads (15–20 pairs of gonads per time points) were sampled in triplicate at 33, 35, 37, 40, 44, 50, 61, 85, and 125 dpf in both genetic all-male (XY) and all-female (XX) populations of rainbow trout. Total RNA was extracted using the RNAqueous-Micro Kit (Ambion, Life Technologies) for the 33- to 50-dpf samples and the RNAqueous-Kit (Ambion, Life Technologies) for the 61- to 125-dpf samples. All samples were then treated with the TURBO DNA-free Kit (Ambion, Life Technologies) to remove any leftover genomic DNA. Reverse transcriptions were carried out using 150 ng of total RNA as the starting material with the Ovation RNA Amplification System (NuGEN Technologies), following the manufacturer's recommendations. Quantitative PCR was performed using the StepOnePlus system (Applied Biosystems) using 4 μL of reverse-transcribed cDNA (single tube quantification per sample, with three biological replicates for each sex and time point) diluted to 1:90, the Fast SYBR Green Master Mix (Applied Biosystem), and 600 nmol of each primer listed in *SI Appendix*, Table S7. The enzyme was activated for 20 s at 95 °C, followed by 40 cycles of denaturation at 95 °C for 5 s and annealing and elongation at 60 °C for 30 s.

Whole-Mount in Situ Hybridization. In situ hybridization was performed as previously described (38). RNA probes were produced from PCR products obtained by amplification of *foxl2b2*. Ten nanograms of the PCR product were used as a template for digoxigenin-labeled RNA probe synthesis using digoxigenin 11-UTP (Roche Diagnostics Corp.) and T3 or T7 RNA polymerase (Promega) following standard protocols. Whole-mount in situ hybridization was carried out using an in situ Pro, Intavis AG robotic station. Male and female embryos were fixed overnight in 4% paraformaldehyde at 4 °C, dehydrated in 100% methanol, and stored at −20 °C. Before in situ hybridization they were rehydrated, permeabilized by proteinase K treatment (25 μg/mL, 30 min, at room temperature), and postfixed (4% paraformaldehyde and glutaraldehyde 0.2%, for 20 min). Prehybridization and hybridization media contained 50% formamide, 5XSSC, 0.1% Tween 20, 0.005% heparin, 0.1 mg/mL tRNA. Hybridization was carried out at 65 °C for 16 h. After post-hybridization washes, embryos were incubated in blocking buffer (PBS/Triton 0.1%/Tween 20 0.2%, containing 2% serum) for 2 h before the addition of the alkaline phosphatase coupled anti-digoxigenin antibody (1:2,000; Roche Diagnostics Corp.) for 6 h. After washing, the color reaction was performed in

the presence of nitro-blue tetrazolium/5-Bromo-4-chloro-3′-indoly phosphate (NBT/BCIP) (Roche). Briefly, dehydration and paraffin infiltration were performed in a Citadel 1000 tissue processor (Shandon). Dehydrated tissues were embedded in plastic molds in paraffin using a HistoEmbedder (TBS88; Medite). Each embedded sample was sectioned 5 μm thick on a MICRO HM355 (Thermo Fisher Scientific).

Colocalization of sdY and foxl2 by in Situ Hybridization. Fifty-dpf male rainbow trout were fixed in Bouin's fixative at 4 °C. One hour after fixation and dehydration, the tissues were embedded in paraffin and cut into 4-μm sections. Antisense *sdY* and *foxl2b2* RNA probes were synthesized using in vitro transcription with a fluorescein RNA labeling mix (Roche) and a DIG RNA labeling mix (Roche), respectively. Sections were deparaffinized, rehydrated, and treated with 1% H_2O_2 in TBST buffer at room temperature for 30 min and 1 μg/mL proteinase K (Roche) at 37 °C for 13 min. After the enzymatic treatment, sections were dehydrated with ethanol and chloroform and then hybridized with *sdY* and *foxl2* probes simultaneously at 60 °C for 18 h. Fluorescein was visualized by using an anti-fluorescein–alkaline phosphatase (anti-fluorescein–AP) Fab fragment (Roche) (1:1,000) and the HNPP Fluorescent Detection Set (Roche). Digoxigenin (DIG) was visualized by using an anti-digoxigenin–AP Fab fragment (Roche) (1:500) and NBT/BCIP. Before the DIG visualization, alkaline phosphatase was inactivated in 0.1 M glycine/0.1% Tween 20 at room temperature for 30 min. DAPI staining was performed to visualize nuclei.

Luciferase Assay. HEK 293T cells were transfected using PEI with the following plasmids: 0.3 μg of pGL3-OlaCyp19a1a sequence (kindly provided by D. Wang Deshou, Key Laboratory of Aquatic Science of Chongqing, School of Life Sciences, Southwest University, Chongqing, China), 0.05–0.4 μg of $pCS2^+$-SdY expression plasmid, 0.05–0.4 μg of $pCS2^+$-OlaFoxl2, 0.1 μg of pcDNA3.1-OlaNr5a1, and 0.1 μg of plasmid thymidine kinase-*Renilla* used for calibration. Each experiment was performed with a 1.0-μg final amount. Adjustments were made with empty plasmid (pCS2+) accordingly. Firefly luciferase and *Renilla* luciferase readings were obtained using the Dual-Luciferase Reporter Assay System (Promega) and LUMAT LB 9501 luminometer (Berthold Technologies).

One-by-one Yeast Two-Hybrid Assay. Diploid cells containing the same bait construct of the yeast two-hybrid assay (p66_SdY) and a prey plasmid construct coding for Foxl2b2 cloned in frame with the activation domain of GAL4 (p14-*N*-GAL4-Foxl2b1-C) were mated and spotted on selective media. The medium lacking tryptophan and leucine was used as a control for the yeast growth test and to check for the presence of the bait or the prey. The assay is based on the histidine reporter gene. A triple-negative medium (tryptophan, leucine, and histidine) selects yeast growth if interaction occurs. Interaction pairs were tested at decreasing concentrations (10^-1, 10^-2, 10^-3, 10^-4) from two independent clones. An inhibitor of the histidine gene product 3-AT was used to increase stringency at four different concentrations (1, 5, 10, 50 mM).

RNA Injections. For injections capped RNA *GFP-sdY*, *mCherry-foxl2b2*, *Olafoxl2* from $pCS2^+$-meGFP:SdY, $pCS2^+$-HA:mCherry:Foxl2b2, and $pCS2^+$-OlaFoxl2, respectively, was transcribed from linearized pCS2+ plasmid using the SP6/T3/T7 m MESSAGE mMACHINE Kit (Ambion). One nanoliter was injected into the cytoplasm of one-cell stage Medaka embryos.

Statistical Analysis. Most of the data were analyzed using an unpaired Student's *t* test. A Mann–Whitney *U* test was then used to compare the median value when the Kolmogorov–Smirnov test was negative. Statistical analysis of multiple groups was performed using an ANOVA one-way test with post hoc Dunnett or Tukey tests for multiple comparison (a control group compared with an experimental group or a control sample compared to an experimental group, respectively). All statistical analyses were performed with GraphPad Prism (version 5; GraphPad Software). Significant differences are symbolized in figures by asterisks if $P < 0.001$ (***), $P < 0.05$ (**), or $P < 0.01$ (*) or indicated by ns if not significant.

ACKNOWLEDGMENTS. We acknowledge Wang Deshou for the medaka *cyp19a1a* promoter and Reiner Veitia for helpful discussions. This work was supported by Agence Nationale de la Recherche (ANR) Grant ANR-11-BSV7-0016 (SDS project) and grants from the Deutsche Forschungsgemeinschaft (Scha408/12-1, 10-1; to M.S.).

1. Herpin A, Schartl M (2015) Plasticity of gene-regulatory networks controlling sex determination: Of masters, slaves, usual suspects, newcomers, and usurpators. *EMBO Rep* 16:1260–1274.
2. Graves JAM (2013) How to evolve new vertebrate sex determining genes. *Dev Dyn* 242:354–359.
3. Lin Y-T, Capel B (2015) Cell fate commitment during mammalian sex determination. *Curr Opin Genet Dev* 32:144–152.
4. Windley SP, Wilhelm D (2015) Signaling pathways involved in mammalian sex determination and gonad development. *Sex Dev* 9:297–315.
5. Kikuchi K, Hamaguchi S (2013) Novel sex-determining genes in fish and sex chromosome evolution. *Dev Dyn* 242:339–353.
6. Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R (1991) Male development of chromosomally female mice transgenic for Sry. *Nature* 351:117–121.
7. Matsuda M, et al. (2002) DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417:559–563.
8. Nanda I, et al. (2002) A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. *Proc Natl Acad Sci USA* 99:11778–11783.
9. Smith CA, et al. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* 461:267–271.
10. Kamiya T, et al. (2012) A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, Takifugu rubripes (fugu). *PLoS Genet* 8:e1002798.
11. Yoshimoto S, et al. (2008) A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis. *Proc Natl Acad Sci USA* 105:2469–2474.
12. Hattori RS, et al. (2012) A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. *Proc Natl Acad Sci USA* 109:2955–2959.
13. Chen S, et al. (2014) Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet* 46:253–260.
14. Myosho T, et al. (2012) Tracing the emergence of a novel sex-determining gene in medaka, Oryzias luzonensis. *Genetics* 191:163–170.
15. Marshall Graves JA, Peichel CL (2010) Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biol* 11:205.
16. Yano A, et al. (2012) An immune-related gene evolved into the master sex-determining gene in rainbow trout, Oncorhynchus mykiss. *Curr Biol* 22:1423–1428.
17. Yano A, et al. (2013) The sexually dimorphic on the Y-chromosome gene (sdY) is a conserved male-specific Y-chromosome sequence in many salmonids. *Evol Appl* 6:486–496.
18. Fu XY, Kessler DS, Veals SA, Levy DE, Darnell JE, Jr (1990) ISGF3, the transcriptional activator induced by interferon alpha, consists of multiple interacting polypeptide chains. *Proc Natl Acad Sci USA* 87:8555–8559.
19. Yu W-L, Sun Y (2015) Interferon regulatory factor 9 plays a dual function in health and disease. *J Hepatol* 62:1446.
20. Clark KL, Halay ED, Lai E, Burley SK (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* 364:412–420.
21. Baron D, et al. (2004) An evolutionary and functional analysis of FoxL2 in rainbow trout gonad differentiation. *J Mol Endocrinol* 33:705–715.
22. Boulanger L, et al. (2014) FOXL2 is a female sex-determining gene in the goat. *Curr Biol* 24:404–408.
23. Bertho S, et al. (2016) Foxl2 and its relatives are evolutionary conserved players in gonadal sex differentiation. *Sex Dev* 10:111–129.
24. Pannetier M, et al. (2006) FOXL2 activates P450 aromatase gene transcription: Towards a better characterization of the early steps of mammalian ovarian development. *J Mol Endocrinol* 36:399–413.
25. Wang D-S, et al. (2007) Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as interacting with ad4 binding protein/steroidogenic factor 1. *Mol Endocrinol* 21:712–725.
26. Li M-H, et al. (2013) Antagonistic roles of Dmrt1 and Foxl2 in sex differentiation via estrogen production in tilapia as demonstrated by TALENs. *Endocrinology* 154:4814–4825.
27. Guiguen Y, Fostier A, Piferrer F, Chang C-F (2010) Ovarian aromatase and estrogens: A pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol* 165:352–366.
28. van der Vos KE, Coffer PJ (2008) FOXO-binding partners: It takes two to tango. *Oncogene* 27:2289–2299.
29. Liu X, Cai X, Zhang D, Xu C, Xiao W (2016) Zebrafish *foxo3b* negatively regulates antiviral response through suppressing the transactivity of *irf3* and *irf7*. *J Immunol* 197:4736–4749.
30. Guiguen Y, et al. (1999) Involvement of estrogens in the process of sex differentiation in two fish species: The rainbow trout (Oncorhynchus mykiss) and a tilapia (Oreochromis niloticus). *Mol Reprod Dev* 54:154–162.
31. Dranow DB, et al. (2016) Bmp15 is an Oocyte-produced signal required for maintenance of the adult female sexual phenotype in zebrafish. *PLoS Genet* 12:e1006323.
32. Lau ES-W, Zhang Z, Qin M, Ge W (2016) Knockout of zebrafish ovarian aromatase gene (cyp19a1a) by TALEN and CRISPR/Cas9 leads to all-male offspring due to failed ovarian differentiation. *Sci Rep* 6:37357.
33. Vizziano D, Randuineau G, Baron D, Cauty C, Guiguen Y (2007) Characterization of early molecular sex differentiation in rainbow trout, Oncorhynchus mykiss. *Dev Dyn* 236:2198–2206.
34. Chen W, et al. (2008) Insights into interferon regulatory factor activation from the crystal structure of dimeric IRF5. *Nat Struct Mol Biol* 15:1213–1220.
35. Formstecher E, et al. (2005) Protein interaction mapping: A Drosophila case study. *Genome Res* 15:376–384.
36. Rexroad CE, 3rd, et al. (2003) Sequence analysis of a rainbow trout cDNA library and creation of a gene index. *Cytogenet Genome Res* 102:347–354.
37. Marivin E, et al. (2014) Sex hormone-binding globulins characterization and gonadal gene expression during sex differentiation in the rainbow trout, Oncorhynchus mykiss. *Mol Reprod Dev* 81:757–765.
38. Yano A, Nicol B, Guerin A, Guiguen Y (2011) The duplicated rainbow trout (Oncorhynchus mykiss) T-box transcription factors 1, tbx1a and tbx1b, are up-regulated during testicular development. *Mol Reprod Dev* 78:172–180.

PLOS BIOLOGY

RESEARCH ARTICLE

A novel evolutionary conserved mechanism of RNA stability regulates synexpression of primordial germ cell-specific genes prior to the sex-determination stage in medaka

Amaury Herpin[1,2]*, Cornelia Schmidt[2], Susanne Kneitz[2], Clara Gobé[3], Martina Regensburger[2], Aurélie Le Cam[1], Jérome Montfort[1], Mateus C. Adolfi[2], Christina Lillesaar[2], Dagmar Wilhelm[4], Michael Kraeussling[2¶], Brigitte Mourot[1], Béatrice Porcon[1], Maëlle Pannetier[3], Eric Pailhoux[3], Laurence Ettwiller[5], Dirk Dolle[5], Yann Guiguen[1], Manfred Schartl[2,6,7]

1 INRA, UR1037 Fish Physiology and Genomics, Rennes, France, **2** University of Wuerzburg, Physiological Chemistry, Biocenter, Wuerzburg, Germany, **3** UMR BDR, INRA, ENVA, Université Paris Saclay, Jouy en Josas, France, **4** University of Melbourne, Department of Anatomy & Neuroscience, Parkville, Victoria, Australia, **5** University of Heidelberg, Centre for Organismal Studies (COS), Department of Developmental Biology, Heidelberg, Germany, **6** Comprehensive Cancer Center Mainfranken, University Hospital, Wuerzburg, Germany, **7** Hagler Institute for Advanced Study and Department of Biology, Texas A&M University, College Station, Texas, United States of America

¶Authorship confirmed by corresponding author.
* amaury.herpin@inra.fr

OPEN ACCESS

Citation: Herpin A, Schmidt C, Kneitz S, Gobé C, Regensburger M, Le Cam A, et al. (2019) A novel evolutionary conserved mechanism of RNA stability regulates synexpression of primordial germ cell-specific genes prior to the sex-determination stage in medaka. PLoS Biol 17(4): e3000185. https://doi.org/10.1371/journal.pbio.3000185

Academic Editor: Yukiko M Yamashita, University of Michigan, UNITED STATES

Received: September 28, 2018

Accepted: March 4, 2019

Published: April 4, 2019

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Deutsche Forschungsgemeinschaft by a grant (SCHA 408/12-1; HE 7135/2-1) to AH and MS as well as Crédits Incitatifs Phase 2015/Emergence to AH. AH was additionally funded by the project AquaCRISPR (ANR-16-COFA-0004-01). The

Abstract

Dmrt1 is a highly conserved transcription factor, which is critically involved in regulation of gonad development of vertebrates. In medaka, a duplicate of dmrt1—acting as master sex-determining gene—has a tightly timely and spatially controlled gonadal expression pattern. In addition to transcriptional regulation, a sequence motif in the 3′ UTR (D3U-box) mediates transcript stability of dmrt1 mRNAs from medaka and other vertebrates. We show here that in medaka, two RNA-binding proteins with antagonizing properties target this D3U-box, promoting either RNA stabilization in germ cells or degradation in the soma. The D3U-box is also conserved in other germ-cell transcripts, making them responsive to the same RNA binding proteins. The evolutionary conservation of the D3U-box motif within *dmrt1* genes of metazoans—together with preserved expression patterns of the targeting RNA binding proteins in subsets of germ cells—suggest that this new mechanism for controlling RNA stability is not restricted to fishes but might also apply to other vertebrates.

Author summary

The development of the gonads in vertebrates is mainly regulated by *dmrt1*, a master sex-determining gene that has a timely and spatially controlled gonadal expression pattern. In addition to transcriptional regulation, a sequence motif located in the 3′ UTR (D3U-box) mediates transcript stability of dmrt1 mRNAs. However, this regulation is complex, and

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: BAC, bacterial artificial chromosome; BSA, bovine serum albumin; BSF, bicoid stability factor; Carbio, Carolina Biological Supplies; celf2, CUGBP Elav-like family member 2; CRISPR/Cas9, clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9; D3U-box, dmrt1 3′ UTR box; dph, days post hatching; EMSA, electrophoretic mobility shift assay; GFP, green fluorescent protein; GO, gene ontology; Lrpprc, leucine rich pentatricopeptide repeat containing; MES-1, medaka embryonic stem cells; miR-224, microRNA 224; Ol-BSF, Oryzias latipes Bicoid Stability Factor; Ol-CUG-BP, Oryzias latipes CUG-binding protein; ORF, open reading frame; PARN, poly(a)-specific ribonuclease; PBS, phosphate-buffered saline; PFA, paraformaldehyde; PGC, primordial germ cell; WNT4, Wnt family member 4.

the mechanisms are barely understood. Here, we analyse further dmrt1 regulation in the fish model medaka and show that two RNA-binding proteins with antagonizing properties target the D3U-box and promote either RNA stabilization in germ cells or degradation in the soma. The evolutionary conservation of the D3U-box motif within metazoan *dmrt1* genes—together with preserved expression patterns of the targeting RNA binding proteins in subsets of germ cells—suggest that this new mechanism for controlling RNA stability is not restricted to fishes but might also apply to other vertebrates.

Introduction

The gonads of vertebrates are characterized by the intimate association of germ cells and supporting somatic cells [1–4]. The precursor cells of the soma are derived from the embryonic lateral plate mesoderm, whereas germ cells originate from the germline lineage [5–9]. To carry out their highly specialized biological functions, the somatic gonadal primordium and the germline cells together must establish timely regulated programs of gene expression [1,10,11].

The *mab-3*/*doublesex*/*dmrt1* gene orthologs are, among metazoans, the most evolutionary conserved key regulators of the earliest phases of gonad development. They control complex gene regulatory networks specifying male gonadal primordium development as well as gonadal maintenance [12–15]. Remarkably, besides being firmly anchored within the regulatory network at critical nodes, *dmrt1* genes were found to act as upstream male sex determiners in organisms as phylogenetically diverse as flatworm [16], water flea [17], frog [18], flatfish [19], birds [20] and medaka [21,22].

During the last decade, much has been learned about how *dmrt1* as the most versatile sex gene triggers and controls gonad development. In human, it is a critical dosage-sensitive sex-determining gene, such that haploinsufficiency leads to XY male-to-female sex reversal and infertility [23,24]. In mice, it is required for male gonadal differentiation of somatic and germ cells [25–27], although *Dmrt1* appears to be dispensable for primary sex determination [15]. Dmrt1 also plays the decisive role in maintaining the cellular identity of the adult testis, most obvious from the fact that its malfunction in adult mutant mice gonads leads to transdifferentiation of Sertoli to granulosa-like cells and feminization of a fully developed testis [27]. Consequently, the action range of Dmrt1 is not restricted to initiation of the male gonadal phenotype during early development but also contributes to the active suppression of the female networks via repression of two 'anti-testis' pathways, Foxl2 and Wnt family member 4 (WNT4)/β-catenin (see [28] for review).

In sex determination model fish medaka, male sex determination is implemented by a male-specific primordial germ cell (PGC) mitotic arrest due to the activity of a Y-chromosome–specific duplicate version of *dmrt1*, designated *dmrt1bY* [29]. In *dmrt1* knockout mice, germ cells fail to arrest mitosis [30]. Further work on dmrt1 has shown it to be a transcriptional gatekeeper controlling mitosis versus meiosis decision in male germ cells [26]. Thus, dmrt1 in mice and dmrt1bY in medaka appear to be regulators of germ cell proliferation.

Despite its well-characterized crucial functions for gonad development in many vertebrates, the mechanisms that regulate the complex temporal and spatial expression pattern and guarantee precise levels of dmrt1 transcripts are only barely understood. Diverse regulatory mechanisms have been occasionally reported. Indirect transcriptional regulation of *dmrt1* upon steroid treatments has been described in several fish species (see [13] for review). Gonadal dimorphic expression of *dmrt1* has been suggested to be possibly under the control of differential CpG methylation of its promoter in two different flatfish species [31,32]. Similarly, in the

red-eared slider turtle (*Trachemys scripta*), DNA methylation dynamics accounting for dmrt1 sexual dimorphic expression are tightly correlated with temperature [33]. In vitro transcriptional regulation assays revealed that binding sites for Sp1, Egr1 [34], and Gata4 [35] factors, which are present in the promoters of many genes, are also involved in transcriptional regulation of the rat *dmrt1* gene. And finally, evidence was presented that microRNA 224 (miR-224) promotes differentiation of mouse spermatogonial stem cells via direct targeting of *dmrt1*, decreasing its expression in testes [36]. Certainly in the context of 'indirect' regulation, *dmrt1* is one of the most prominent examples.

In medaka, for which a functional duplicate of the autosomal *dmrt1a* gene on the Y chromosome—*dmrt1bY*—became the master regulator of male sex determination [22,37], transcriptional rewiring was brought about by exaptation of two transposable elements, *Izanagi* and *Rex1*, co-opted to act as silencers. These turn off the somatic and the germ cell–specific expressions of the *dmrt1bY* gene [38, 39]. Thus far, two factors, dmrt1 itself [38] and sox5 [39], were identified, which turn off *dmrt1bY* expression after it has fulfilled its function in the early developing gonad [38, 39].

We previously identified a 11-bp sequence motif in the 3′ UTR of dmrt1bY (D3U-box, for dmrt1 3′ UTR box). This motif confers stability to the mRNA in the developing embryonic gonad, whereas in other tissues, the transcript is rapidly degraded [40], indicating that a post-transcriptional regulation mechanism could play a role in germline expression of dmrt1 in medaka.

Here, we show that the dmrt1 11-nucleotide *cis*-regulatory D3U-Box motif is a target for two antagonizing RNA binding proteins, *Oryzias latipes* CUG-binding protein (Ol-cug-bp) and *Oryzias latipes* Bicoid Stability Factor (Ol-bsf)—also known as cugbp Elav-like family member (celf) and leucine rich pentatricopeptide repeat containing (lrpprc), respectively, in mammals. In *Drosophila*, the bicoid stability factor (bsf) has initially been shown to be involved in regulating the stability of *bicoid* transcripts during oogenesis through binding structures within the 3′ UTR of transcripts that resemble CUG hairpins [41]. Later, bsf was also reported to have a role in regulation of early zygotic genes by binding a short consensus sequence (CAGGUA) in the 5′ UTR of genes expressed in the early zygote [42]. Cug-bp is the human homolog of the *Xenopus* eden-bp, which was shown to bind to mRNAs, such as *c-mos*, that exhibit rapid deadenylation following fertilization of oocytes [43]. Previous studies of cug-bp function have focused mainly on the roles of this protein in regulating alternative splicing [44] and also on its ability to modulate translation of several mRNAs [45]. However, as cug-bp is able to functionally substitute for eden-bp to induce deadenylation in *Xenopus* oocyte extracts [46], it seems likely that it also plays a similar role in regulating poly(A) shortening in mammalian cells. Indeed, it was shown that cug-bp can interact with poly(a)-specific ribonuclease (PARN) deadenylase to promote deadenylation of its substrate RNAs [47].

We find that in medaka, the D3U-box is targeted by these two different RNA binding proteins, with Ol-cug-bp1 leading to dmrt1bY degradation unless Ol-bsf is present in germ cells. Moreover, this new mechanism of dmrt1 RNA stability appears to regulate also the abundance of other transcripts specifically expressed in PGCs.

Results

Specific enrichment of a conserved *cis*-regulatory motif (D3U-box) in vertebrate genomes

In an initial analysis of dmrt1 post-transcriptional regulation [40], we found that an 11-bp long *cis*-regulatory motif in the D3U-box confers transcript stability in PGCs (Fig 1A). *In vitro* evidence was obtained that the D3U-box possibly mediates its function through protein

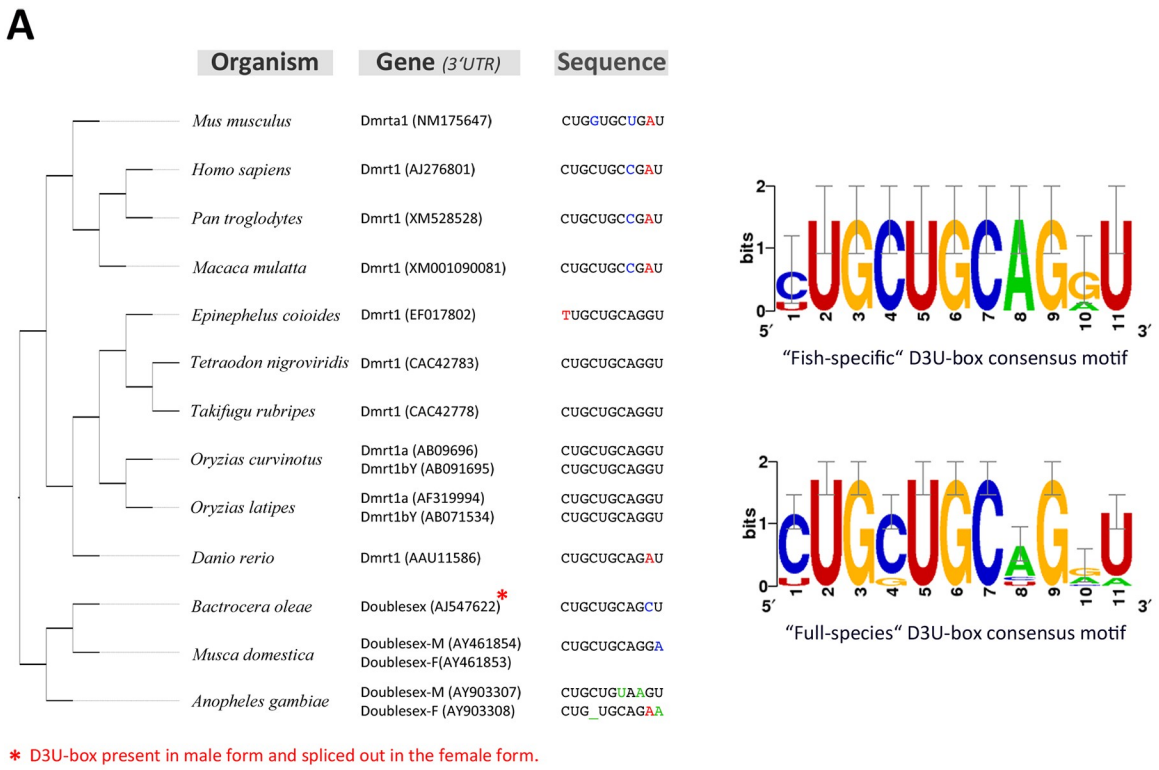


Fig 1. Conservation of the D3U-box motif from ecdysozoans to mammals. (A) Conservation of the D3U-box motif from *Drosophila* up to mammals. From the different D3U-box sequences among vertebrates, weight-position matrices were deduced and used for genome scans. (B) Occurrences of the D3U-box motif compared to control motifs within the 3′ UTR sequences of human, mouse, medaka, or the green spotted puffer (*Tetraodon nigroviridis*). For comparison, scrambled control motifs were used. Absolute values of the occurrences depend on the depth of annotation for every genome. Underlying data for (B) can be found in S1 Data. D3U-box, dmrt1 3′ UTR box.

https://doi.org/10.1371/journal.pbio.3000185.g001

binding [40]. Further on this sequence motif was found to be evolutionary conserved in *dmrt1* genes from flies to mammals (Fig 1A; [40]).

To investigate whether this motif is specific for the *dmrt1* genes or also present in other genes, we performed genome-wide searches for human, mouse, medaka, and the green spotted puffer (Fig 1B). The D3U-box was found in more than 250 3′ UTRs of genes from human and mice and is also abundant in medaka and green spotted puffer (Fig 1B). Notably, also, a 'split' version of the D3U-box (see Fig 1B) displayed specific enrichment in 3′ UTRs (Fig 1B), suggesting that the D3U-box motif might be articulated around two independent *cis*-regulatory sequences, hence putatively targeted by 2 different RNA binding proteins.

Medaka whole transcriptome scans (3′ UTR and coding sequences) using the vertebrate D3U-box consensus motif matrix (Fig 1A: 3′ UTR sequences; and S1 Fig: 3′ UTR and coding sequences) resulted in several hits, including tra2, sox10, misr2, dead end, and vasa (S1 Fig). Like dmrt1 [14,29,48,49], these RNAs are critically involved in germ cell development and maintenance in medaka and many other organisms ([36,50–55] and [56,57] for review).

Furthermore, bioinformatics analyses and literature searches [45,47] revealed that the D3U-box *cis*-regulatory motif is a putative target for 2 evolutionary conserved RNA binding proteins involved in either mRNA degradation or stabilization. These 2 proteins, Ol-cug-bp (also known as CELF in mammals) and Ol-bsf (also known as LRPPRC in mammals), have been shown to specifically recognize CUG repeats and the CAGGU(AG) motif, respectively, which constitute the D3U-box (see S2 Fig for phylogeny and synteny analysis of *Ol-bsf* and *Ol-cug-bps*).

Ol-BSF and Ol-CUG-BP1 specifically bind to the D3U-box motif

To confirm our bioinformatics prediction, Ol-bsf and Ol-cug-bp1 and 2 proteins were subjected to electrophoretic mobility shift assay (EMSA) using the D3U-box motif as target and different competitors (Fig 2). The *in vitro*–translated proteins (Fig 2A and 2F) were assayed for binding with radioactively labelled RNA probes. Using the D3U-box motif, mobility shifts were detected for the 2 proteins tested: Ol-bsf (Fig 2B to 2E) and Ol-cug-bp1 (Fig 2G to 2I), indicating that Ol-bsf as well as Ol-cug-bp1 are, in principle, able to bind the D3U-box in vitro. Binding specificities were confirmed by competition of the medaka D3U-box motif for Ol-bsf or Ol-cug-bp1 interactions with either a scrambled D3U-box–derived motif (Fig 2B and 2I) or a minus CUG repeat motif competitor (Fig 2H). The absence of any significant interference with the D3U-box binding indicated the specificity of the observed interactions (Fig 2B, 2I and 2H). Furthermore, competition experiments between radioactively and nonradioactively labelled D3U-boxes resulted in progressive loss of the apparent shifts (Fig 2C and 2G). Notably, a clearly visible shift was also observed when using the *Drosophila* D3U-box sequence together with the medaka Ol-bsf protein (Fig 2D). Altogether, these experiments suggest that the D3U-box is a preferential target for Ol-bsf and Ol-cug-bp1 binding. Of note, performing the very same set of experiments together with the Ol-cug-bp2 protein did not result in any convincing evidence for specific binding to the D3U-box.

Ol-bsf and Ol-cug-bps antagonistically regulate the expression and stability of reporter constructs harbouring the D3U-box motif and of dmrt1bY transcripts

To monitor a possible effect of Ol-bsf on regulation of the male sex-determination gene in medaka, we generated a dmrt1bY reporter line by introducing the green fluorescent protein (GFP) open reading frame (ORF) fused to the dmrt1bY 3′UTR (including the D3U-box) into exon 1 of a bacterial artificial chromosome (BAC) clone containing the *dmrt1bY* gene and

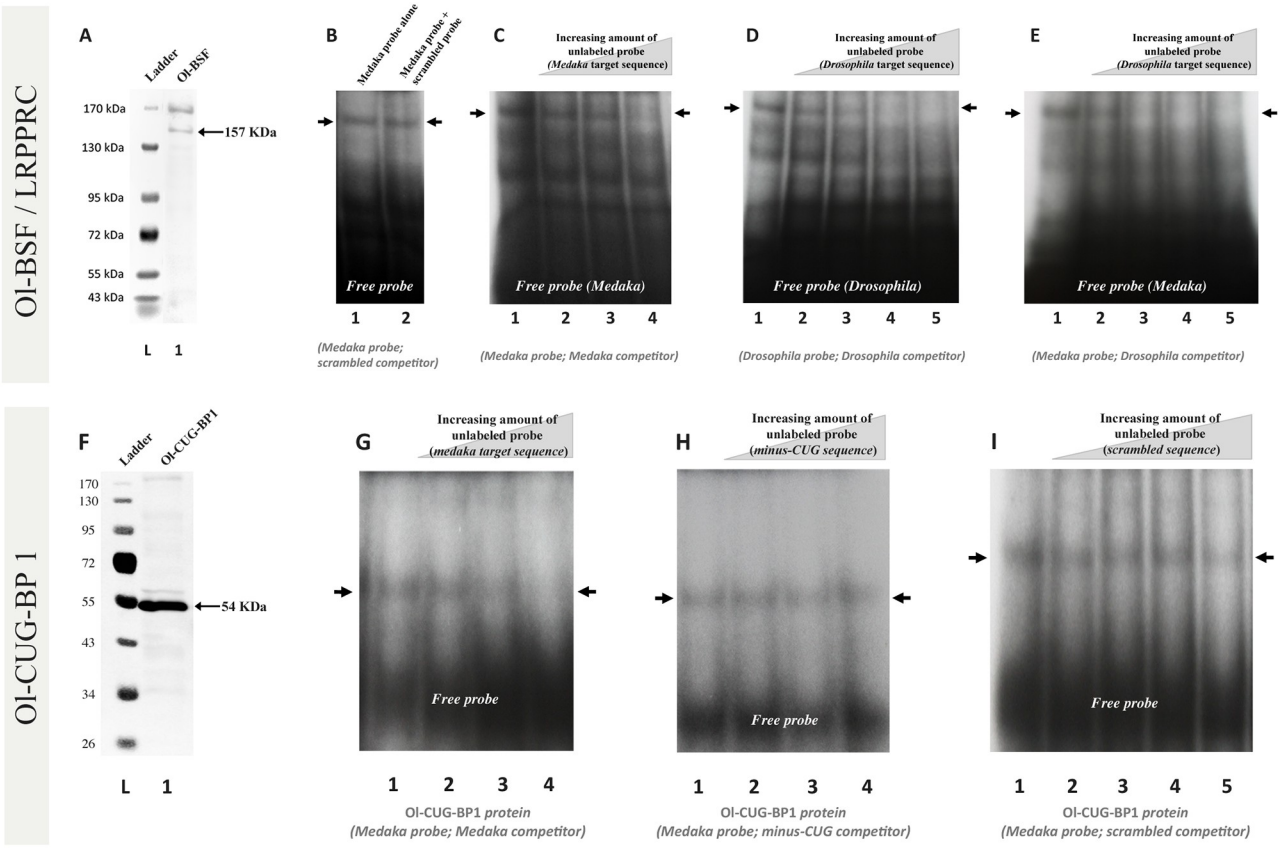


Fig 2. EMSA of in vitro–translated Ol-bsf, Ol-cug-bp1, and Ol-cug-bp2 proteins indicating interactions with the medaka D3U-box target sequence. (A and F) Recombinant protein production of FLAG-tagged versions of Ol-bsf and Ol-cug-bp1 and western blotting detection with an anti-FLAG antibody. (B–E and G–I) EMSA using the recombinant medaka Ol-bsf and Ol-cug-bp1 proteins (in vitro translated) and either the medaka D3U-box sequence (B, C, E, G, H, I), or the *Drosophila* D3U-box sequence (D) as radioactively labelled RNA probes. Increasing amounts of unlabelled probes (scrambled [B, I], medaka [C, G, K], *Drosophila* [D, E], minus CUG [H, L] target sequences) were used as competitors. (B–E and G–I) Apparent shifts are observed for the labelled RNA target probes with the Ol-bsf and Ol-cug-bp1 RNA-binding proteins, likely indicating direct interactions. Ratios of RNA probe to RNA competitor: (B, lane 2: 1/5; C, lane 2: 1/1, lane 3: 1/2, lane 4: 1/4; D, lane 2: 1/1, lane 3: 1/2, lane 4: 1/5, lane 5: 1/10; E, lane 2: 1/1, lane 3: 1/2, lane 4: 1/5, lane 5: 1/10; G, lane 2: 1/1, lane 3: 1/2, lane 4: 1/5; H, lane 2: 1/1, lane 3: 1/2, lane 4: 1/5; I, lane 2: 1/1, lane 3: 1/2, lane 4: 1/5, lane 5: 1/10). D3U-box, dmrt1 3′ UTR box; EMSA, electrophoretic mobility shift assay; lrpprc, leucine rich pentatricopeptide repeat containing; Ol-bsf, Oryzias latipes Bicoid Stability Factor; Ol-cug-bp, Oryzias latipes CUG-binding protein.

https://doi.org/10.1371/journal.pbio.3000185.g002

flanking regions (Fig 3A and 3B). The recombined BAC was then used for establishing a stable transgenic line in which GFP expression most reliably indicates endogenous *dmrt1bY* expression [48,58,59]. Expression of dmrt1bY is highly dynamic during primordial gonad formation, progressively switching from germ cell expression only between stages 26 (1.25 dpf) and 29 (3.1 dpf) to an exclusive somatic expression from stages 33/34 (5 dpf) up to hatching (9 dpf) when the gonad is formed [39,59]. After injection of the Ol-bsf morpholino (see S3 Fig for validation of the morpholino), we found a significant reduction of GFP expression (Fig 3B) and, after Ol-bsf overexpression, a strong increase of the reporter at both mRNA and protein (fluorescence) levels (Fig 3A and 3B).

Next, to obtain a more physiological readout of the role(s) of Ol-bsf and Ol-cug-bps for RNA stability in vivo, the relative abundances of endogenous dmrt1bY transcripts were monitored after modulation of Ol-bsf/Ol-cug-bp1/Ol-cug-bp2 expression in medaka embryos (Fig 3C and 3D). First, we checked for changes in dmrt1bY transcript levels after overexpression of the 2 medaka Ol-cug-bp ohnologs (Ol-cug-bp1 and Ol-cug-bp2, Fig 3C). This resulted in

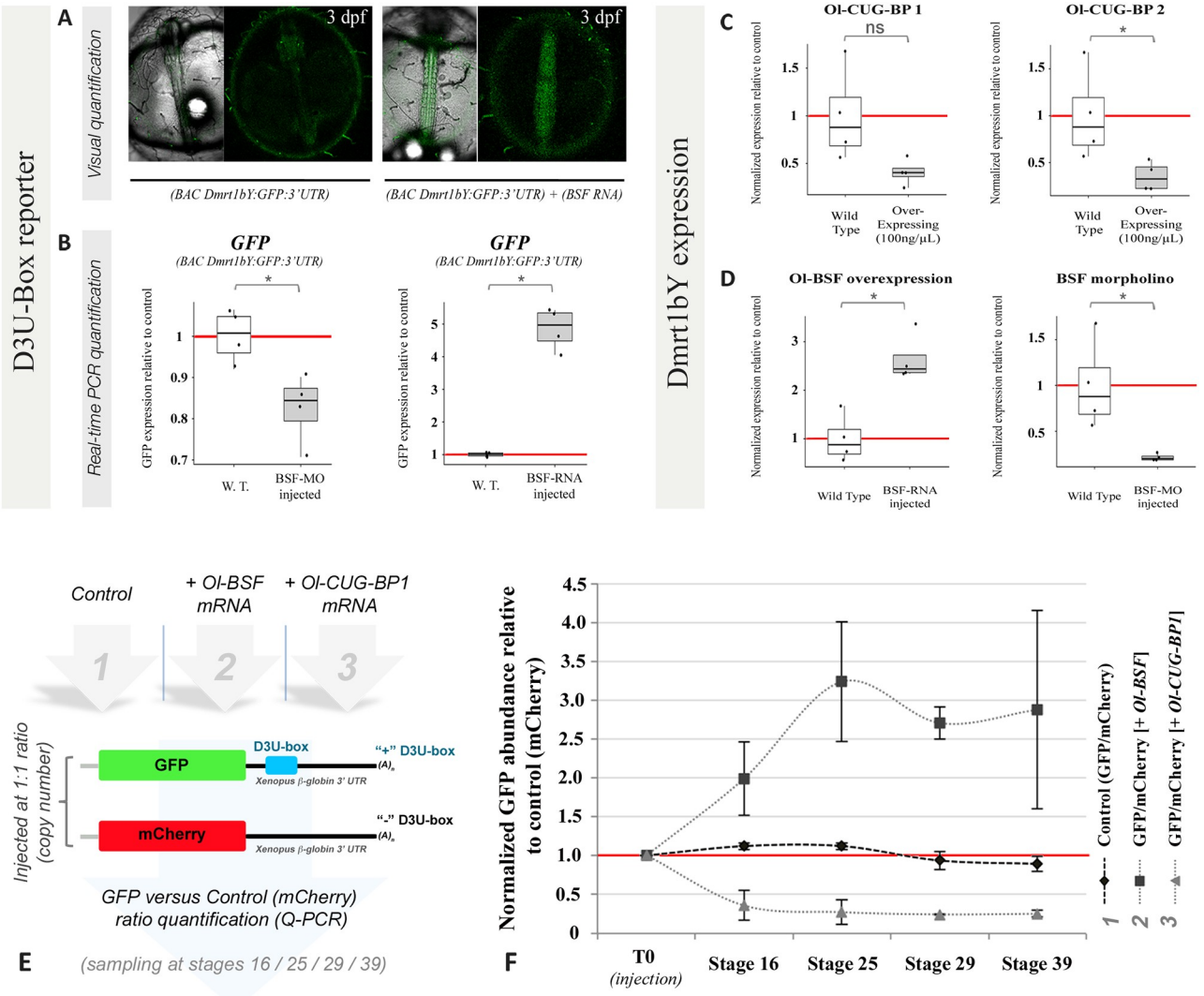


Fig 3. In vivo modulation of expression and mRNA stability of reporters and dmrt1bY transcripts via the D3U-box. (A and B) Effect of the modulation of the Ol-bsf expression on a GFP reporter for dmrt1bY expression. Injections of either Ol-bsf morpholino or capped mRNA result in the subsequent modulations of expression of a GFP:(dmrt1bY 3′ UTR) reporter/sensor construct in a transgenic line (BAC Dmrt1bY:GFP:3′UTR) as monitored by fluorescence (A) and RNA quantifications (B). Because endogenous GFP is lowly expressed already before experimental down-regulation, the more sensitive qPCR method was employed to monitor the morpholino effect (B). Dataset results of 4 independent reverse transcription reactions, resulting in 4 different batches of eggs obtained from different couples. Statistical significance has been assessed by mean of the Wilcoxon-Mann-Whitney test ($N = 4$). (C and D) Real-time qPCR determination of *dmrt1Y* abundances after Ol-cug-bp1 and Ol-cug-bp2 (C) or Ol-bsf (D) modulation of expression in embryos at stage 18 after injection of either capped mRNAs with the full Ol-bsf/Ol-cug-bps ORFs or a splice morpholino targeting the exon2-intron2 splice junction of *Ol-bsf* (S3 Fig). Dataset results of the analysis of 4 batches of eggs injected with either Ol-cug-bp1, Ol-cug-bp2, Ol-bsf, or BSF-MO. Statistical significance has been assessed by mean of the Wilcoxon-Mann-Whitney test ($N = 4$). $^{*}p \leq 0.05$, $^{**}p \leq 0.01$. (E and F) Real-time qPCR determination of the kinetics of RNA stability for D3U-box-containing reporters in presence or in absence of either Ol-cug-bp1 or Ol-bsf *coding mRNAs*. Datasets are results of 3 independent reverse transcription reactions, resulting from three different batches of injected eggs. Underlying data for (B, C, D, and F) can be found in S1 Data. BAC, bacterial artificial chromosome; D3U-box, *dmrt1* 3′ UTR box; GFP, green fluorescent protein; MO, Morpholino; ns, nonsignificant; Ol-BSF, *Oryzias latipes* Bicoid Stability Factor; Ol-CUG-BP, *Oryzias latipes* CUG-binding protein; ORF, open reading frame; qPCR, quantitative PCR; W.T., wild type.

https://doi.org/10.1371/journal.pbio.3000185.g003

decreased dmrt1bY mRNA (Fig 3C). Second, the relative abundances of dmrt1bY transcripts were recorded after either overexpression or morpholino knockdown of Ol-bsf in medaka embryos (Fig 3D). It revealed that higher Ol-bsf expression correlates with an increased abundance, while lowering Ol-bsf expression resulted in a reduction of dmrt1bY transcripts

(Fig 3D). In vivo, D3U-box–induced modulation of RNA stability was further investigated in embryos injected with either control RNAs or RNAs harbouring the D3U-box (Fig 3E). Ratios between control and D3U-box–containing mRNAs were then quantified in absence or in presence of either l-cug-bp1 or Ol-bsf mRNAs (Fig 3F). It revealed that, over time, overexpression of Ol-bsf correlates with an increased stabilization of the D3U-box–containing mRNAs, while overexpression of Ol-cug-bp1 correlates with a decreased stabilization of the D3U-box–containing mRNAs (Fig 3F).

Expression of Ol-cug-bp1, Ol-cug-bp2, and Ol-bsf during embryonic development and in adult tissues

During the embryonic developmental period, both Ol-cug-bp ohnologs display complementary patterns of expression. Ol-cug-bp1 is expressed at early stages and Ol-cug-bp2 only later when dmrt1bY transcripts appear in the germ cells around stage 25 (S4A and S4C Fig). In adult tissues, Ol-cug-bp1 and Ol-cug-bp2 are expressed at high levels in brain and gonads (S4B and S4D Fig), while Ol-cug-bp2 is additionally expressed in eyes, muscles, and skin (S4D Fig). Of note, both ohnologs are always more expressed in testes compared to ovaries (S4B and S4D Fig). Expression profiling of Ol-bsf revealed that in adult medaka, it is ubiquitously expressed in all adult tissues, with particular high levels in gonads of both sexes (S4E and S4F Fig).

Ol-bsf is specifically expressed in the germ cells during early gonad primordium formation with correlated levels of expression between Ol-bsf and dmrt1bY at hatching stage

For bioimaging analyses of protein localization over time, we used expression reporter lines for vasa [11] and Ol-bsf (see Materials and methods), respectively. During embryonic development, we noted a distinct spatially and temporarily restricted expression pattern (Fig 4). From fertilization up to stages 16/17, Ol-bsf is expressed throughout the embryo (Fig 4A). Of note, cell transfection of a tagged version of OL-bsf and subsequent immunohistochemistry revealed that bsf protein is localized in the cytoplasm (insert in Fig 4A). From stage 25 onwards—when germ cells line up on both sides of the embryo within the lateral plate mesoderm—progressively, Ol-bsf expression becomes restricted to the PGCs (Fig 4B to 4G) where it is co-expressed with Ol-vas, a specific germ cell marker in medaka [60]. During the following developmental stages (stages 33/34), expression heterogeneity for Ol-bsf between germ cells became obvious (Figs 4H to 4J and 5C and 5D). This heterogeneity was particularly apparent between 4 to 10 days post hatching (dph), when Ol-bsf is higher expressed at the tips of the forming gonads (curly brackets in Fig 4K to 4P and square brackets in Fig 5C and 5D). In summary, Ol-bsf has a highly dynamic expression pattern, switching from an early somatic to a progressively restricted germ cell expression. Within the germ cell pool, the levels of expression show a significant heterogeneity.

Furthermore, dmrt1bY expression shows heterogeneity between individual germ cells (Fig 4Q and 4R). Being also higher expressed at the tip of the primordial gonads, the expression levels of Ol-bsf and dmrt1bY show a clear positive correlation (Fig 4Q and 4R).

Expression of Ol-bsf in the germ-line stem cells of adult gonads

Given the high abundance of Ol-bsf transcripts in adult gonads detected by qPCR (S4E and S4F Fig), we next monitored expression of Ol-bsf in fully mature gonads of both sexes at cellular resolution (Fig 6). In adult testes, Ol-bsf fluorescence is restricted to two distinct

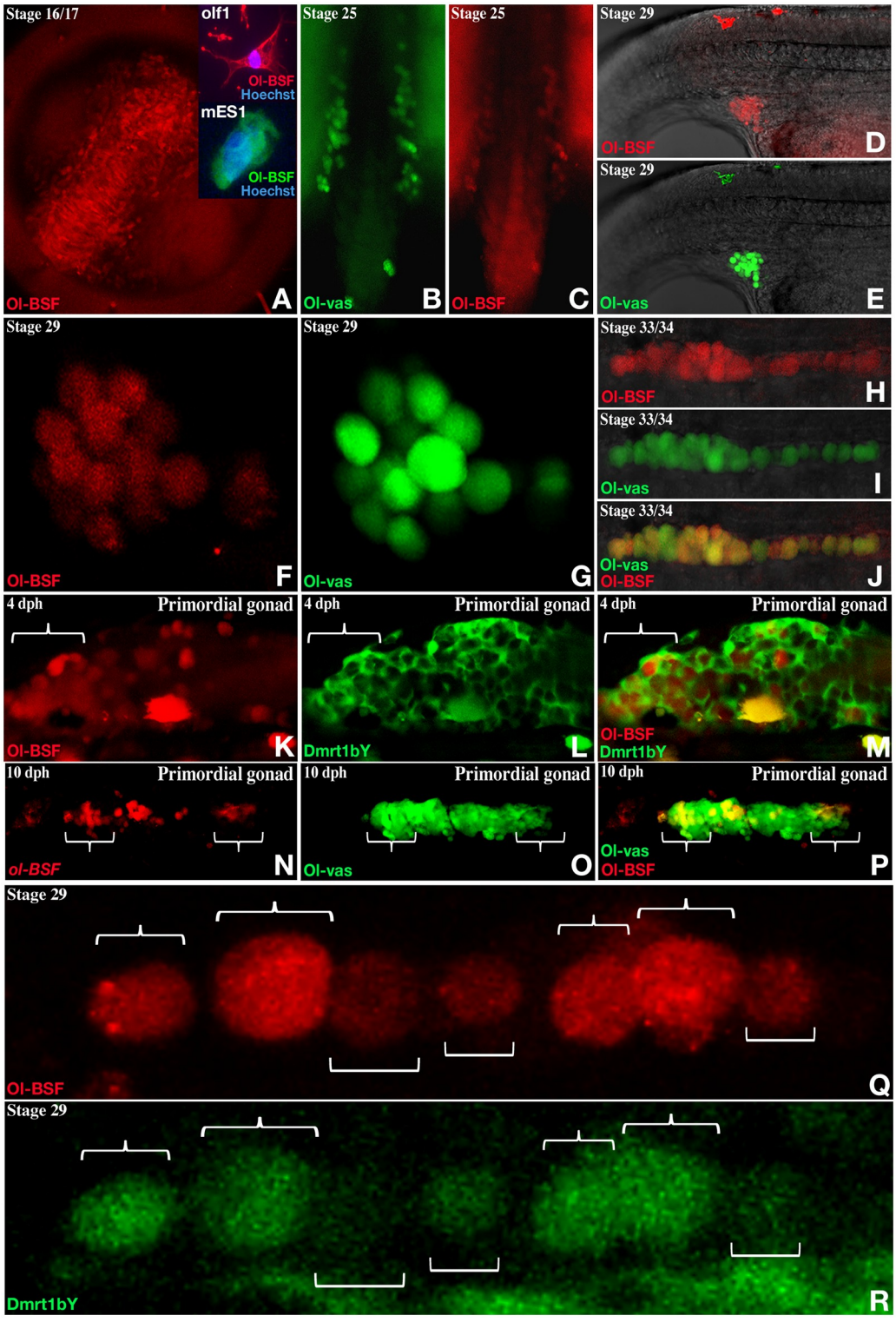


Fig 4. Embryonic expression of medaka Ol-bsf and correlated expression dynamics together with dmrt1bY during gonadal primordium formation. (A to R) Medaka Ol-bsf expression investigated by fluorescence using a transgenic reporter line for which a 732-bp *Ol-bsf* promoter fragment (up to the next upstream gene) drives the expression of the mCherry. Medaka Ol-vas and dmrt1bY expressions were investigated by fluorescence using previously described transgenic reporter lines ([11] and [61], respectively). (A) During neurulation (stages 16/17), Ol-bsf is ubiquitously expressed. (A insert) Cytoplasmic localisation of Ol-bsf after transient transfection of a 3-times FLAG version of Ol-bsf in

a medaka embryonic stem cell line and immunohistochemistry. (B to P) From stage 25 on, Ol-bsf expression is restricted to the germ cells and colocalize with the germ cell marker Ol-vas. (K to P) Four-dph expression of Ol-bsf is apparently heterogeneous within the population of PGCs while dmrt1bY is now expressed in the surrounding somatic cells of the primordial gonad (K to M). (Q and R) By stage 29, Ol-bsf and dmrt1bY are both expressed in the germ cells. Variations within the respective levels of Ol-bsf and dmrt1bY are clearly observable amongst different germ cells although always correlated between the 2 fluorescences (curly brackets compared to square brackets). dph, days post hatching; MES-1, medaka embryonic stem cells; Ol-BSF, Oryzias latipes Bicoid Stability Factor; olf1, *Oryzias latipes* fibroblast-1; PGC, primordial germ cell.

https://doi.org/10.1371/journal.pbio.3000185.g004

subpopulations of germ cells, which are also positive for Ol-vas (Fig 6A to 6C). Diagnosed by condensed nuclear morphology and size [62] and localization [39], these first Ol-bsf–positive cells represent the earliest step of germ cell differentiation while another subpopulation of more mature germ cells is also observed (Fig 6C and 6F). In ovaries, Ol-bsf fluorescence is restricted to the germinal cradle [63] located in the interwoven threadlike ovarian cords at the periphery of the ovary (Fig 6G to 6N). These Ol-bsf–positive cells, representing the smaller-size subpopulation of Ol-vas fluorescent cells, are assigned to germline stem cells and early dividing germ cell lineage [63] (Fig 6O to 6Q).

Furthermore, in mice both bsf (lrpprc) and cug-bp1 (celf1) are expressed in the germ cells within the testis cords and germ cells of the ovary (S5A and S5L Fig). Immunohistochemistry revealed that bsf/lrpprc is expressed only in a subpopulation of germ cells in mice (S5M to S5R Fig).

Ol-bsf mutant fish display gonadal phenotypes

To delineate the physiological role of Ol-bsf during gonad formation and maintenance, we generated medaka *Ol-bsf* knockout lines after genome editing using the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) technology (S6 Fig). Homozygous *Ol-bsf* knockout larvae display reduced swimming (S7 Fig) and die within the first 2 weeks after hatching. Heterozygote mutant fish develop normally and produce mature gametes. Histological analysis of gonads of heterozygous mutants of both sexes revealed, however, that ovaries had an accumulation of small-sized oocytes compared to wild type (Fig 7A to 7D, S8 and S9 Figs), whereas testes of heterozygous mutants exhibited reduced amounts of spermatogonial stem cells, with germ cells in advanced stages of spermatogenesis located close to the periphery of the organ (Fig 7E and 7F and S10 Fig). In both sexes, *Ol-bsf* heterozygote mutants present an increase of germ cell committing to gametogenesis. While adult mutant males do not display any observable bias of fertility, adult mutant females have reduced egg production together with lower fertilization rates (S11 Fig).

Ol-bsf and Ol-cug-bps antagonistically orchestrate expression levels of several germ cell transcripts selectively harbouring the D3U-box

Our matrix scan bioinformatic analysis had revealed the presence of the D3U-box either in 5′, 3′, or coding regions of several germ cell transcripts (S1 Fig). To find out whether Ol-bsf and Ol-cug-bp might regulate stability of these RNAs in a similar way, like for dmrt1bY, during gonadal development the relative abundances of these transcripts harbouring the box motif (tra2, sox10, misr2, dead end, and vasa; see S1 Fig) were first monitored after either overexpression or morpholino knockdown of Ol-bsf in medaka embryos (Fig 5A1 to 5A10). It revealed that for the majority of these transcripts (sox10, misr2, tra2, and vasa), higher Ol-bsf expression correlated with an increased abundance (Fig 5A1 to 5A5 odd numbers and 5A19), while reduced Ol-bsf expression resulted in lower levels of most of these transcripts (Fig 5A2 to 5A6 even numbers). As exceptions, dead end (Fig 5A7 and 5A8) transcript abundance

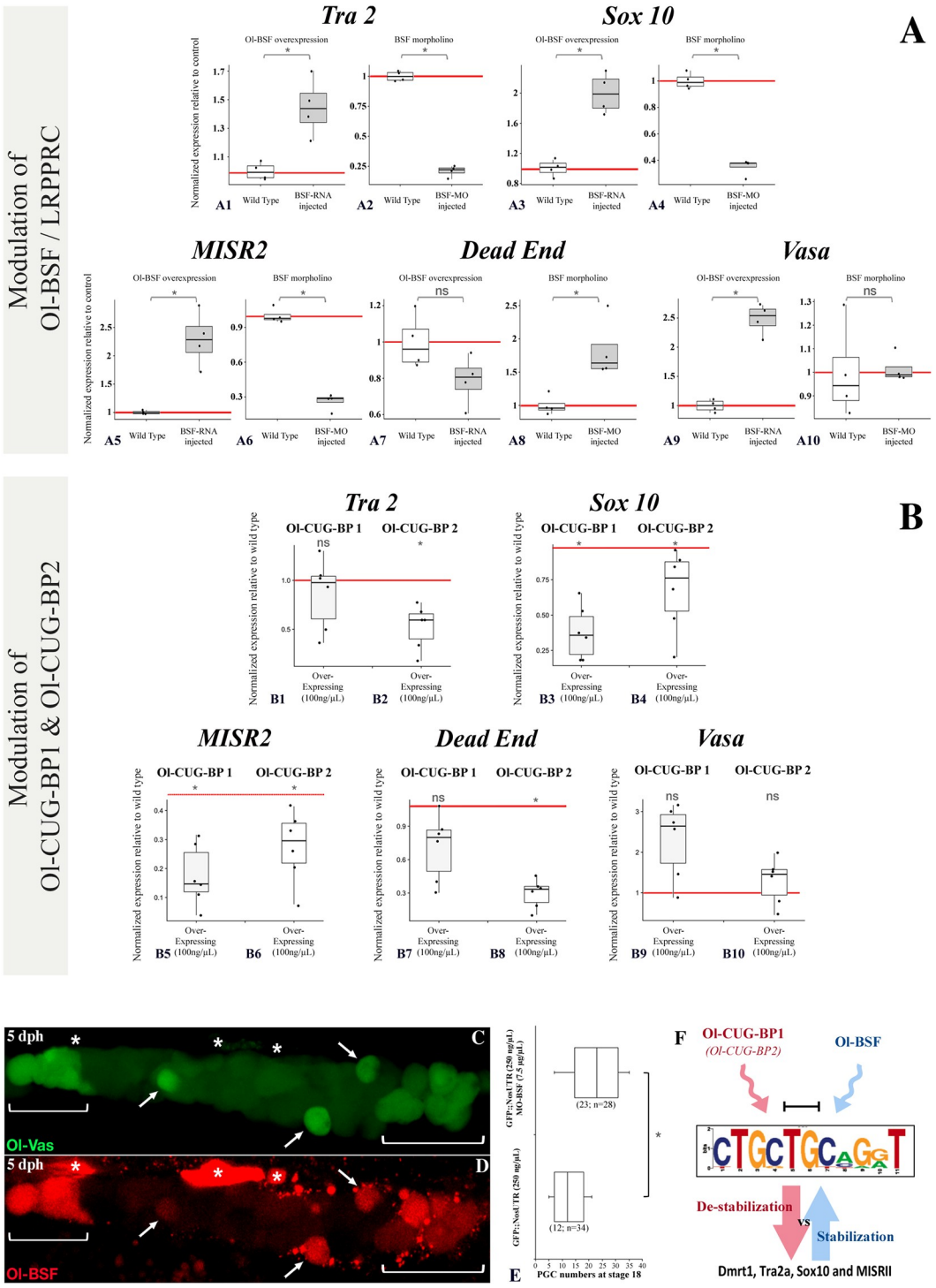


Fig 5. In vivo transcriptional regulation of putative gene candidates harbouring the D3U-box motif in their transcripts after modulation of the Ol-bsf or Ol-cug-bp expressions and correlation between Ol-bsf and Ol-vas expression in germ cells. (A1 to A10) Real-time qPCR determination of expression of putative Ol-BSF-regulated transcripts in embryos at stage 18 after injection of either capped mRNAs with the full Ol-bsf ORF (left odd-numbered panels) or a splice morpholino targeting the exon2-intron2 splice junction (S3 Fig) of the *Ol-bsf* gene (right even-numbered panels) at one cell stage. Ol-bsf overexpression results in an apparent increase of most RNA target candidates, while reduction of the Ol-bsf expression correlates with a reduction of expression. Dataset results of the analysis of 4 batches of eggs injected with either Ol-bsf or Ol-bsf-MO. Statistical significance has been assessed by mean of the Wilcoxon-Mann-Whitney test ($N = 4$). (B1 to B10) Real-

time qPCR determination of expression of putative Ol-cug-bp–regulated transcripts in embryos at stage 18 after overexpression of either Ol-cug-bp1 or Ol-cug-bp2. Results are presented as normalized expressions compared to wild type using 3 different housekeeping genes. Dataset results of the analysis of 2 batches of eggs injected with either Ol-cug-bp1 or Ol-cug-bp2. Statistical significance has been assessed by mean of the Wilcoxon-Mann-Whitney test ($N = 4$). (C and D) Germ cell expression of Ol-bsf in comparison to Ol-vas in double transgenic fluorescent reporter lines. After hatching expression levels of Ol-bsf and Ol-vas become more heterogeneous among germ cells (brackets at the tips of the primordial gonad) although tightly correlated between each other within individual germ cells (arrows; '*' = autofluorescent pigment cells). (E) In vivo modulation of PGC number after Ol-bsf morpholino injection. An apparent increase in PGC number is observed after negative regulation of Ol-bsf expression in early embryos (stage 18). Statistical significance has been assessed by means of the Wilcoxon-Mann-Whitney test ($N = 34$ and 28 for wild-type and MO-bsf-injected embryos, respectively). (F) Model for D3U-box–mediated mRNA regulation. Overall and in addition to a cytoplasmic localization of Ol-bsf (Fig 4A), Ol-bsf and O-cug-bps might mutually antagonize toward the access to the D3U-box, resulting in either stabilisation (more Ol-bsf binding) or destabilisation (more Ol-cug-bp binding) of the transcripts harbouring the D3U-box. $^{*}p \leq 0.05$; $^{**}p \leq 0.01$. Underlying data for (A, B, and E) can be found in S1 Data. D3U-box, *dmrt1* 3′ UTR box; MO, Morpholino; ns, nonsignificant; Ol-BSF, Oryzias latipes Bicoid Stability Factor; Ol-CUG-BP, Oryzias latipes CUG-binding protein; PGC, primordial germ cell; qPCR, quantitative PCR.

https://doi.org/10.1371/journal.pbio.3000185.g005

showed either no or an opposite pattern following Ol-bsf expression modulation, while vasa (Fig 5A10) was unaffected in the Ol-bsf-morpholino–treated embryos.

We next checked for the relative abundances of the very same transcripts harbouring the D3U-box (tra2, sox10, misr2, dead end, and vasa) after overexpression of the 2 medaka Ol-cug-bp ohnologs (Ol-cug-bp1 and Ol-cug-bp2, Fig 5B1 to 5B10). With the exception of vasa (Fig 5B9 and 5B10), the majority of the transcripts analysed had lower abundance, whereas medaka Ol-cug-bp ohnologs were overexpressed (Fig 5B1 to 5B8).

In vivo, an apparent correlation between Ol-bsf levels and Ol-vas expression could be visualized in the germ cells of the forming gonadal primordium using fluorescent reporter lines (Fig 5C and 5D). Reduced levels of Ol-bsf expression after morpholino injection led to a significant increase in PGC number at stage 18, (Fig 5E). In line with this observation, microarray data comparing *Ol-bsf*$^{(-/+)}$–deficient testes (displaying reduced levels of Ol-bsf; S12A Fig) to wild type revealed a general up-regulation of genes involved in germ cell proliferation or differentiation. A significant proportion (10.1%) of the down-regulated genes codes for proteins localized in the mitochondria. Finally, gene ontology (GO) term analysis revealed that, in mutant testes partially depleted for the *Ol-bsf* gene, rRNA processing is particularly affected (S12A Fig).

Discussion

The expression of most genes is dynamically tightly regulated, temporally and spatially. Such regulations occur at multiple steps, including transcription, splicing, mRNA transport, mRNA stability, translation, protein stability, and post-transcriptional modifications [65,66]. While the importance of complex post-transcriptional regulations—like in the case of nanos, oskar, or bicoid, e.g. [61,67–69]—has been mainly identified through genetic approaches for the development of the germline or oocyte, respectively, such approaches have proven to be much less valuable for finding the expected regulatory proteins that bind specifically to these mRNAs. In medaka fish, expression of the master sex determiner dmrt1bY mRNA is very dynamic, occurring first in the PGCs prior to morphological somatic sex differentiation and then quickly switches to an exclusive Sertoli cell localisation [59,70]. Importantly, dmrt1bY is expressed in PGCs of male embryos much before its expression in the pre-Sertoli cells at the sex determination stage [59]. This early PGC expression is necessary for the later onset of dmrt1bY expression in the pre-Sertoli cells at the sex-determination stage of male development [59]. There, the level of dmrt1bY mRNA needs to reach a certain threshold to exert the sex-determining function [21]. This suggested that medaka germ cells exhibit sexually different

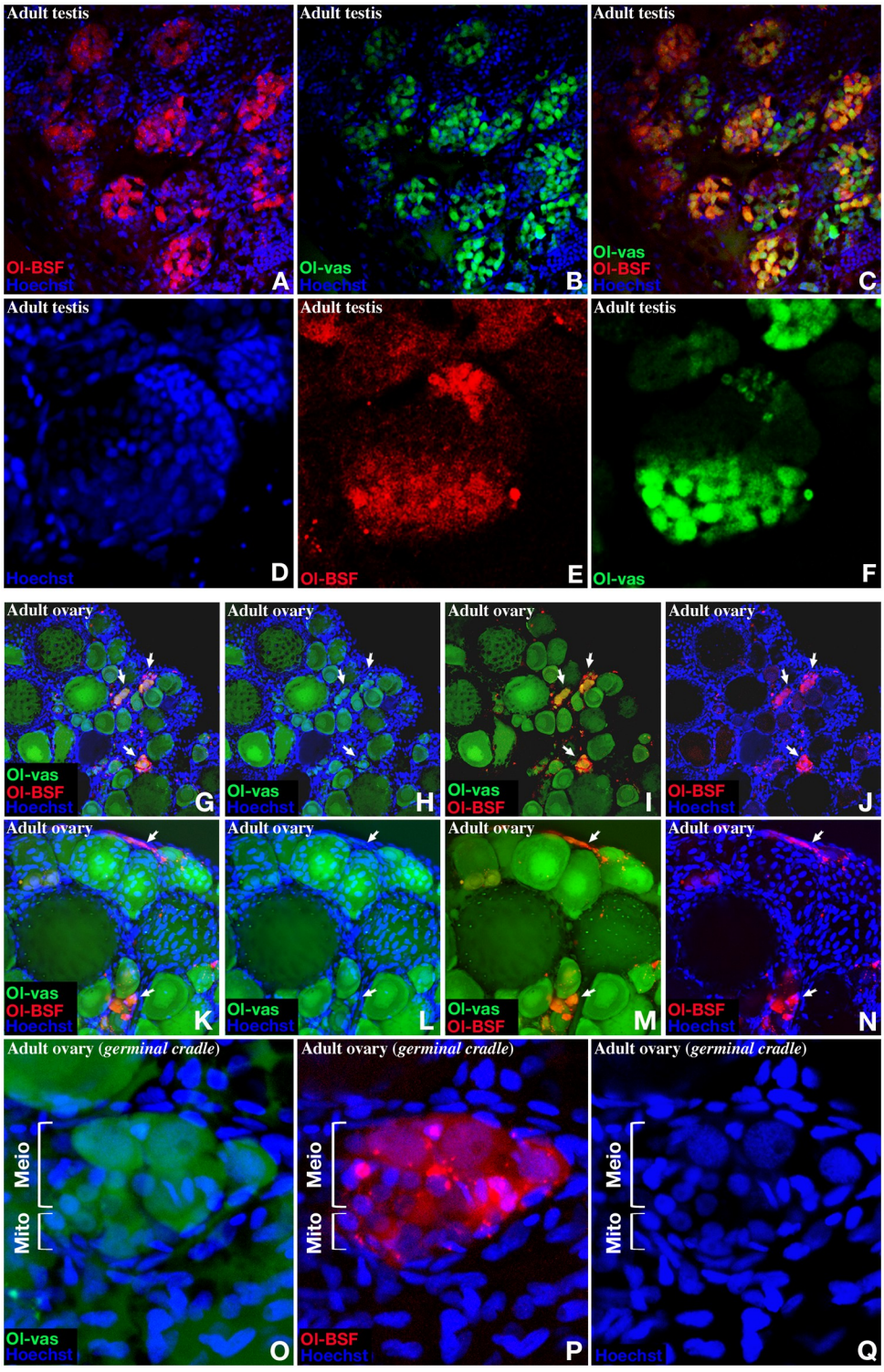


Fig 6. Expression of Ol-bsf in adult gonads. (A to F) In adult testes, Ol-bsf fluorescence (same transgenic lines as in Fig 3) is restricted to the cells located within the lobules and colocalize with a subpopulation of Ol-vas–positive cells. (G to Q) Within the ovary, Ol-bsf fluorescence is restricted to the germinal cradle located in the interwoven threadlike ovarian cords at the periphery of the ovary. The Ol-bsf–positive cells represent the smaller-size subpopulation of Ol-vas fluorescent cells (G to N) as well as early dividing germ cells (O to Q). Arrowheads indicate the germinal cradles (G to N), and brackets indicate early dividing mitotic or meiotic germ cell (O to Q). Ol-BSF, *Oryzias latipes* Bicoid Stability Factor.

https://doi.org/10.1371/journal.pbio.3000185.g006

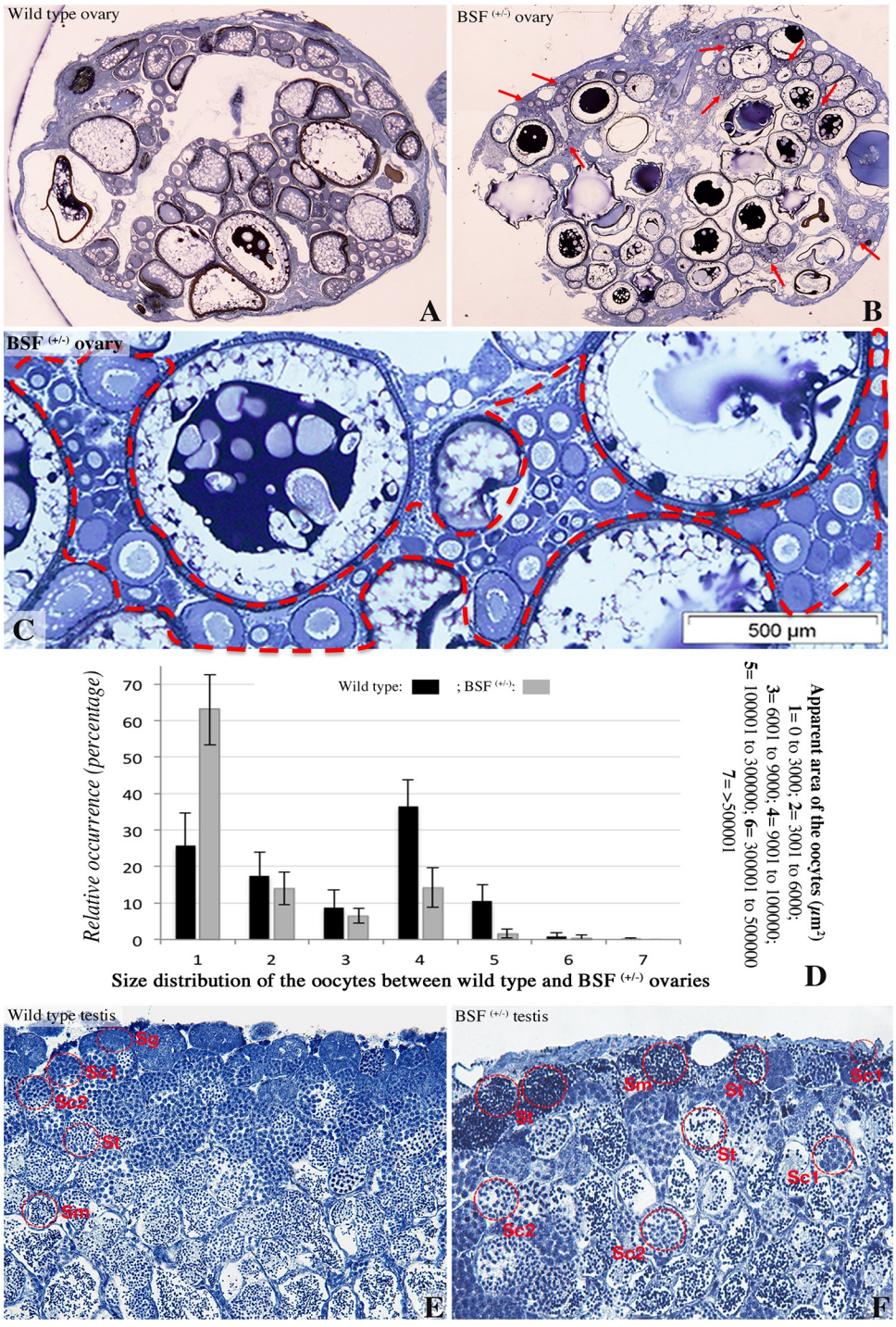


Fig 7. Ovarian and testicular phenotypes of the *Ol-bsf* mutant fish. Heterozygote mutant fish develop normally. (A to D) In-depth morphological inspection of heterozygote mutant ovaries discloses a significant accumulation of small-sized oocytes compared to wild type (dotted lines in [C] and red arrows in [B] and S8 and S9 Figs). (D) Size distribution of the oocytes in 2 wild-type and 2 *Ol-bsf*$^{(+/-)}$ adult ovaries (see also S8 Fig for details). (E and F) Heterozygote mutant testes exhibit a decreased number of spermatogonia with accumulation of type 2 spermatocytes, spermatids, and sperm within the most external layers of the seminiferous epithelium (see also S10 Fig). The different stages of spermatogenesis were determined according to [64]. Each gonad was sectioned through the mid-sagittal plane. Underlying data for (D) can be found in S1 Data. Ol-BSF, *Oryzias latipes* Bicoid Stability Factor; Sc1, type 1 spermatocytes; Sc2, type 2 spermatocytes; Sg, spermatogonia; Sm, sperm; St, spermatids.

https://doi.org/10.1371/journal.pbio.3000185.g007

characters before the formation of the somatic gonadal primordium depending on dynamic and tightly timely regulated mechanisms of post-transcriptional regulations [59,70].

We have identified in medaka a *cis*-regulatory 11-bp motif in the 3′ UTR of dmrt1bY called D3U-box. This motif confers stability to the dmrt1bY mRNA in germ cells of the developing embryonic gonad, whereas in other tissues, the transcript is rapidly degraded [40]. The D3U-box motif was found to be highly conserved in the dmrt1 3′ UTR in the fish lineage (*O. latipes*, *O. curvinotus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Epinephelus coioides*, and *Danio rerio*), as well as in other vertebrates, including *Mus musculus*, *Pan troglodytes*, *Macaca mulatta*, and *Homo sapiens*, and even in the ecdysozoan clade (*Anopheles gambiae* and *Bactocera oleae*).

Both phylogenetic conservation and presence of the D3U-box in several germ cell transcripts implied the existence of similarly conserved *trans*-acting factor(s) involved in the synexpression of those transcripts. To identify such factor(s), we undertook an unbiased approach centred on the D3U-box sequence and based on the evolutionary conservation of the 'split' motifs of the D3U-box, implying evolutionary conserved *trans*-acting factors. Further bioinformatics analyses and literature searches revealed that the D3U-box motif is a putative target for 2 RNA-binding proteins, namely cug-bp [45,47] and bsf (also known as lrpprc in mammals [41,42,71]).

EMSAs indicated that Ol-bsf and Ol-cug-bp1, but not Ol-cug-bp2, specifically target and interact with the different parts of the D3U-box, the 3′ and the 5′ parts, respectively. Additionally, our results suggested that the observed regulation of *dmrt1bY* transcript abundance is likely to be the result of a differential binding of the 2 RNA-binding proteins (Ol-bsf and Ol-cug-bp1) with antagonistic properties, *trans*-regulating RNA stability via the D3U-box.

Being—like its mammalian counterparts—ubiquitously expressed, Ol-cug-bps are highly expressed in the gonads of both sexes in medaka (S4C to S4F Fig) and mice ([72] and S5 Fig). Interestingly, Ol-bsf is specifically expressed in the germ cells during medaka early gonadal primordium formation as well as in the adult ovary and is cytoplasmically localized. This expression pattern resembles the subcellular localization of *Drosophila* bsf, which is present in cytoplasmic particles in oocytes and surrounding nurse cells [41] and in the cytoplasms and nuclei in early embryos [42].

In vivo, using fluorescent reporter medaka fish lines, we could show that, besides obvious correlated expression levels between Ol-bsf, dmrt1bY, and Ol-vas in a subpopulation of germ cells of the forming gonadal primordium, medaka Ol-bsf is also preferentially expressed in adult germline stem cells. This restricted and up to now unreported expression pattern might reveal another so far underappreciated role for bsf/lrpprc in germ cell physiology. Similarly, immunohistochemistry localization of mouse bsf/lrpprc revealed its presence in a subpopulation of germ cells in mice (S5 Fig), suggesting an evolutionary conserved function that is not restricted to exclusive regulation of dmrt1.

Reduced expression of bsf/lrpprc in mammalian cell lines [73,74] or flies [75] or because of a missense mutation in human (French Canadian Leigh Syndrome [76]) resulted in decreased levels of mitochondrial mRNAs. This led to respiratory chain dysfunction and increased lactate levels in flies and humans [75,76]. Similarly, in medaka, many transcripts with mitochondrial function displayed decreased steady-state levels when Ol-bsf expression was reduced (S12B Fig). Although knockdown of *bsf* in flies affects climbing ability, fecundity, and life span [75], mutant medaka hatchlings for *Ol-bsf* comparably displayed a significantly reduced swimming ability (S7 Fig), suggesting—with all the necessary notes of caution, such as in *Drosophila*—possible muscle weakness in relation to mitochondrial dysfunction and energy metabolism failure [75].

Ol-bsf–heterozygous mutant fish develop normally. Although adult mutant males do not seem to display any observable bias of fertility, adult mutant females have reduced egg production together with significantly lower fertilization rates. In-depth morphological inspection of heterozygote mutant gonads of both sexes revealed that ovaries display an important accumulation of small-sized oocytes compared to wild type. Mutant testes exhibited a discontinuous spermatogenetic flux, likely reflecting uncontrolled spermatogenesis, independent of the seminiferous epithelial cycle. Such a testicular phenotype can be interpreted in the light of dmrt1 loss of expression as a result of low Ol-bsf expression. Similarly, in mice, loss of dmrt1 in germ cells uncouples meiotic initiation from the seminiferous epithelial cycle, resulting in uncontrolled spermatogenesis, too [26]. The mouse celf1 (also known as cugbp1) is predominantly expressed in testis ([72] and S5E and S5F Fig). There, it was demonstrated that celf1 post-transcriptionally represses cyp19a1 (*aromatase*) mRNA translation, by direct binding, to maintain high concentrations of testosterone compatible with spermiogenesis [77]. This situation is reminiscent of the situation we observed in medaka with a strong repression of aromatase expression when both Ol-cug-bps are overexpressed (S12C1 and S12C2 Fig). Additionally, as observed in medaka, mouse celf/cug-bps and bsf/lrpprc are expressed at relevant levels in germ cells, possibly implying a functional conservation across vertebrates.

Whole transcriptome scans using either the medaka or *Drosophila* D3U-box sequences as query resulted in hits enriched for genes specifically expressed in the germ cells, including tra2, sox10, misr2, dead end, and vasa. We provide evidence that Ol- and Ol-cug-bps antagonistically regulate the expression of germ cell transcripts harbouring the D3U-box motif. Certainly most of these regulations occur via an mRNA decay versus stabilization equilibrium after Ol-cug-bps and Ol-bsf targeted the D3U-box motif.

Lastly, to explain the observed differences in the degree of regulation of D3U-box containing transcripts after either Ol-bsf or Ol-cug-bp1/2 modulations, we consider that this follows the degree of conservation of the D3U-box (S1 Fig). It is also intuitive that the location of the D3U-box (5′, 3′ or coding regions) within the transcripts is of relevance (S1 Fig). For efficient regulation, reasonably high (tra2) or moderate (sox10) conservation and location within the UTRs appears to be more effective than a moderately conserved motif located in the coding sequences (vasa and dead end). However, multiple, highly conserved D3U-boxes nested within the coding region seem to be efficient as well (misr2). Dead end transcripts, for which the D3U-box was identified within the 5′ UTR, however, has only strict conservation for the 3′ part (CUGCUG) and is only regulated by Ol-cug-bp1 and Ol-cug-bp2 (Fig 5B7 and 5B8) while it expectedly escapes Ol-bsf regulation (Fig 5A7 and 5A8). Altogether, using complementary approaches, our data suggest that the D3U-box motif is—depending on the cellular context—targeted by 2 antagonizing RNA binding proteins, promoting either RNA stabilization in germ cells or degradation in the soma. This new mechanism of dmrt1 RNA stability appears to also regulate the abundance of other transcripts specifically expressed in PGCs, depending of the preservation of the D3U-box motif.

Materials and methods

Bioinformatic analyses

Gene, transcript, and UTR annotation, coordinates, and sequence for human, mouse, medaka, and *Tetraodon* were retrieved from EnsEMBL using the EnsEMBL API (version 54). UTR regions spread across several exons were stitched together per transcript, and the resulting sequence was scanned for the presence of the D3U-box consensus matrix and the other motifs.

Fish maintenance and breeding

All medaka fish used in this study were taken from closed breeding stocks of the wild-type Carolina Biological Supplies (Carbio) strain or transgenic lines produced on a wild-type Carbio background and were kept under standard conditions [5]. Medaka embryos were staged according to Iwamatsu [78]. The fish used in this study from aquaria housed stocks were kept and sampled in accordance with the applicable EU and national German legislation governing animal experimentation. We hold an authorization (568/300-1870/13) of the Veterinary Office of the District Government of Lower Franconia, Germany, in accordance with the German Animal Protection Law (TierSchG).

Cell culture, transient cell transfection, and immunohistochemistry

Medaka embryonic stem cells (MES-1) were cultured as described [79]. For transfection, cells were grown to 80% confluence in 6-well plates and transfected with 5 μg expression vector using FuGene (Roche, Germany) reagent as described by the manufacturer. After *pCS2*::*OL-LRPPRC*:*3XFLAG* transfection for 48 hours, cells were fixed in 4% paraformaldehyde (PFA) for 15 minutes, washed with phosphate-buffered saline (PBS) buffer, and then permeabilized with 0.1% Triton X-100 in PBS for 10 minutes. After blocking in 1% Bovine Serum Albumin (BSA) for 20 minutes, cells were incubated overnight at 4˚C in blocking buffer (1% BSA) together with the primary antibody (3-times FLAG, monoclonal anti-FLAG M2, category number F1804; Sigma-Aldrich). After extensive washes in PBS, cells were then incubated with Alexa 488 conjugated secondary antibody in 1% BSA for 1 hour. Cell nuclei were stained with Hoechst 33343 (Invitrogen) for 5 minutes (1 μg/mL final concentration) and subsequently mounted using Mowiol 4–88 (Roth). Confocal images were acquired using a Nikon Eclipse C1 laser-scanning microscope (Nikon) and were fitted with a 60× Nikon objective (PL APO, 1.4 NA) and Nikon image software.

Phylogenetic analysis

The lrpprc phylogenetic tree was built using the online phylogeny.fr automatic pipeline [80]. *lrpprc* sequences were retrieved from public database sequences in the following species and lrpprc homologs were retrieved by tblastn searches on the PhyloFish [81] database (http://phylofish.sigenae.org/) using medaka protein (*O. latipes*, XP_011482612.1) as bait. Sequences were aligned with MUSCLE (version 3.8.31) configured with default settings. After alignment, ambiguous regions were removed with Gblocks (version 0.91b) using the following parameters: minimum length of a block after gap cleaning of 10, no gap positions allowed in the final alignment, rejection of all segments with contiguous nonconserved positions bigger than 4, and a minimum number of sequences for a flank position of 85%. The phylogenetic tree was reconstructed using the neighbour joining method implemented in the BioNJ program [82] with $N = 100$ bootstrapping. The resulting phylogenetic tree was exported as a Newick file and edited in Evolview [83]. The public database for lrpprc sequences is as follows: *Cyprinodon variegatus* (XP_015233698.1), *Haplochromis burtoni* (XP_005930623.2), *Fundulus heteroclitus* (XP_012720452.1), *Xiphophorus maculatus* (XP_005804231.1), *Poecilia reticulata* (XP_008426945.1), Lates calcarifer (XP_018552533.1), *Salmo salar* (XP_014035445.1), *Kryptolebias marmoratus* (XP_017287636.1), *Larimichthys crocea* (KKF31900.1), *Pygocentrus nattereri* (XP_017546072.1), *Ictalurus punctatus* (XP_017319936.1), *Austrofundulus limnaeus* (XP_013886428.1), Xenopus tropicalis (NP_001039203.1), *Cynoglossus semilaevis* (XP_008319858.1), *Callorhinchus milii* (XP_007895038.1), *Astyanax mexicanus* (XP_007255829.1), *Chrysemys picta bellii* (XP_005296166.1), *Scleropages formosus* (KPP63655.1), *Bos taurus* (XP_005212770.1), *Cyprinus carpio* (KTG42350.1 and

XP_018937975.1), *Equus caballus* (XP_005600080.1), *H. sapiens* (BAF82705.1), *Rattus norvegicus* (NP_001008519.1), *M. musculus* (AAH59862.1), *Coturnix japonica* (XP_015712995.1), *Serinus canaria* (XP_009094627.1), *Lepisosteus oculatus* (XP_015218688.1), *Amia calva* (AAC_LOC100694568.1), *Esox lucius* (XP_012989242.1), *Latimeria chalumnae* (XP_005999623.1), *Sinocyclocheilus anshuiensis* (XP_016313114.1 and XP_016355337.1), *Sinocyclocheilus rhinocerous* (XP_016417572.1 and XP_016400889.1), *Oreochromis niloticus* (XP_003438484.3), *D. rerio* (NP_001136064.1), *Stegastes partitus* (XP_008296301.1), *T. rubripes* (XP_011601434.1), *Nothobranchius furzeri* (XP_015800103.1), and *Gallus* (XP_001234903.3). The PhyloFish species are as follows:

- (DAA_LPPRC.1).2_*Anguilla_anguilla*;
- Y_Up_12_k31_Locus_13642_*Umbra_pygmaea*;
- W2_Dp_10_k25_Locus_11398_*Dallia_pectoralis*; S2_Pb_10_k31_Locus_2910_*Pantodon_bulchozi*;
- W_Ha_12_k43_Locus_3235_*Hiodon_alosoides*;
- D2_Om_14_k31_Locus_6093_*Oncorhynchus_mykiss*;
- N_St_1_k25_Locus_84_*Salmo_trutta*;
- O_Ot_10_k25_Locus_4385_*Oncorhynchus_tshawytscha*;
- Q_Sf_1_k25_Locus_1875_*Salvelinus_fontinalis*;
- P_Cl_2_k25_Locus_10853_*Coregonus_lavaretus*;
- I_Tt_3_k25_Locus_2938_*Thymallus_thymallus*;
- E_Aa_2_k25_Locus_4938_*Alosa_alosa*;
- Z_Pa_1_k25_Locus_2290_*Plecoglossus_altivelis*;
- U_Cc_1_k25_Locus_265_*Coregonus_clupeaformis*;
- R_Ok_10_k31_Locus_25108_*Oncorhynchus_keta*;
- G_Ph_4_k37_Locus_379_*Pangasianodon hypophthalmus*;
- K2_Cb_1_k25_Locus_1206_*clarias_batrachus*;
- M_Pf_1_k25_Locus_1736_*Perca_fluviatilis*;
- G2_Ps_1_k65_Locus_195_*Polypterus_senegalensis*;
- K_Gm_1_k43_Locus_6389_*Gadus_morhua*;
- C_Ob_2_k31_Locus_7388_*Osteoglossum bicirrhosum*;
- B2_Gp_10_k25_Locus_1874_*Gnathonemus petersii*;
- V_Sc_10_CL11520Contig1_*Scyliorhinus caniculata*

The celf2 phylogenetic tree was built using the online phylogeny.fr automatic pipeline [80]. *Celf2* sequences were retrieved from Ensembl. Sequences were aligned with MUSCLE (version 3.8.31) configured with default settings. After alignment, ambiguous regions were removed with Gblocks (version 0.91b) using the following parameters: minimum length of a block after gap cleaning of 5, gaps were allowed in the final alignment if they were within an appropriate block, all segments with contiguous nonconserved positions bigger than 8 were rejected, and a

minimum number of sequences for a flank position of 55%. The phylogenetic tree was reconstructed using the neighbour joining method implemented in the BioNJ program [82] with $N = 100$ bootstrapping. The resulting phylogenetic tree was exported as a Newick file and edited in Evolview [83]. The genomic context around *celf1* and *celf2* was analysed using the Genomicus website [84]. A few additional genomes were analysed manually by blasting (tblastn) some fish NCBI genomes with the spotted gar Celf protein and by manually extracting the NCBI gene annotation around these corresponding *celf* loci.

Locomotor activity

Locomotor activity was determined at hatching stage (stage 39, 9 dpf) either under dark conditions or with 10% light (approximately equal to 100 lux) by measuring the total distance swum during a period of 5 minutes (described in [85]) for 12 hatchlings in each condition. Briefly, the larvae were transferred into 12-well plates containing 1 mL Danieau's solution, and the plate was placed in a Zebrabox equipped with a video camera, infrared light, and filter (ViewPoint Life Sciences, Lyon, France) and the ZebraLab Videotrack software (ViewPoint Life Sciences) for tracking. Following a 5-minute-long habituation period with the same light conditions as for the experimental period, the locomotor activity was recorded. The detection threshold was 11, the inactive/small threshold was 0.5 cm/s, and the small/large threshold was 1.0 cm/s. The total distance swum for each individual is equal to the sum of distances reached during inactivity, small and large movements. *p*-Values were calculated using two-tailed unpaired *t* test with Excel 14.4.8 (Microsoft), and $p < 0.05$ was considered significant. Bars and error bars indicate mean ± standard deviation.

EMSA

Binding assays were carried out as previously described [38,40]. In detail, 5′-UGGUUCACGU**CUGCUGCAGGU**CUCUGACUCU-3′ for the native D3U-box, 5′-UGGUUCACGU**CUGCUGCAGCU**CUCUGACUCU-3′ for the Olive fruit fly box (Off-box), 5′-UGGUUCACGU**UCUUCAAGACG**CUCUGACUCU-3′ for the D3U-scrambled box (S-D3U-box), and 5′-UGGUUCACGU**UCUUCACAGGU**CUCUGACUCU-3′ for the D3U-minus CUG box (CUGminus-D3U-box) were synthesized and end-labelled. For radioactive labelling, 50 pmol of the duplex 5′ termini were used, together with 50 pmol of gamma-[^{32}P]-ATP and 20 units of T4 PNK in 1×-adjusted T4 PNK buffer, and were incubated for 20 minutes at 37 ˚C. Unincorporated nucleotides were removed through a Sephadex G-50 spin column. For producing Ol-lrpprc, Ol-cug-bp1, and Ol-cug-bp2 proteins, *pCS2*::*OL-LRPPRC*, *pCS2*::*OL-CUG-BP1* or *pCS2*::*OL-CUG-BP2* plasmids were linearized using *KpnI* and then in vitro transcribed using mMessage mMachine kit (Ambion). Finally, Ol-bsf, Ol-cug-bp1, and Ol-cug-bp2 proteins were in vitro translated using Ambion's Retic Lysate Kit from the previously in vitro–transcribed Ol-lrpprc, Ol-cug-bp1, and Ol-cug-bp2 capped RNAs. DNA binding reaction contained 10 mMTris-HCl (pH 7.9), 100 mM KCl, 10% glycerol, 5 mM $MgCl_2$, 1 μg torula rRNA, 0.075% Triton X-100, 1 mM DTT, 1 μg BSA, 0.5 ng radiolabelled probe, and 2 or 4 μL in vitro translation mix in a total volume of 20 μL. The amount of 1/10 volume heparin (50 mg/mL) was added just before loading the binding reaction. For control, reticulocyte lysate alone together with radiolabelled duplex probe was used and did not result in any shift. Binding reactions were performed on ice for 10 minutes, and complexes were resolved on a 5% native acrylamide (37.5:1) gel in 0.5 X TBE and then directly subjected to autoradiography. Of note, due to different exposure times—which can be appreciated through the different intensities of the free probes in the respective figures—autoradiographies should not be compared between each other.

Establishment of transgenic fluorescent reporter lines and imaging

For a dynamic and in vivo visualization of endogenous Ol-bsf expression, a transgenic fluorescent reporter line was created. The *Ol-bsf* upstream promoter region of the *Ol-bsf* gene (732 bp up to the next upstream gene) was cloned (*BamH1* sites) in front of the mCherry ORF of a meganuclease plasmid (BamH1-BSFp-Fw: AAAGGATCCAGTGTGAGTTCTATCAAGCC TGG; BamH1-BSFp-Rv: AAAGGATCCTTCTGTAGCTGCGTAGAGGAAGATC). For the generation of a stable transgenic line, the meganuclease protocol was used [86]. Briefly, approximately 10 to 15 pg of total vector DNA in a volume of 500 pL injection solution containing *I-SceI* meganuclease was injected into the cytoplasm of 1-cell–staged medaka embryos (Carbio strain). Adult F0 fish were mated to each other, and the offspring were tested for the presence of the transgene by fluorescence check. Siblings from positive F1 fish were raised to adulthood and tested again for fluorescence. Tg[*vasa*:*GFP*] and Tg[*dmrt1bY*:*GFP*] transgenic lines were described earlier [11,29,48,59].

Visualization of PGCs

For PGC visualization and counting, the GFP-nos1 3′ UTR construct that includes the mmGFP5 ORF cloned upstream of the 3′ UTR of the zebrafish *nanos1* gene [5,87] was injected at 1 cell stage. $N = 34$ and 28 for wild-type and MO-bsf–injected embryos, respectively. For imaging, embryos, hatchlings or tissues were mounted with 1.2% low melting temperature agarose.

In vivo expression regulation analyses and real-time PCR

For producing Ol-bsf, Ol-cug-bp1, and Ol-cug-bp2 capped mRNAs, *pCS2*::*OL-BSF*, *pCS2*::*OL-CUG-BP1*, or *pCS2*::*OL-CUG-BP2* plasmids were linearized using *KpnI* and then in vitro transcribed using mMessage mMachine kit (Ambion). For overexpression, 1 nanolitre was injected into the cytoplasm of 1-cell–stage medaka embryos. For mRNA stability assays, equimolar amounts of control and D3U-box–containing RNAs were injected.

Total RNAs were extracted from fish tissues or embryos using the TRIZOL reagent (Invitrogen) according to the supplier's recommendation. After DNase treatment, reverse transcription was performed with 2 μg total RNA using RevertAid First Strand Synthesis kit (Fermentas) and random primers. Real-time quantitative PCR was carried out with SYBR Green reagents, and amplifications were detected with an i-Cycler (Biorad). All results are averages of at least 2 independent reverse transcription reactions. Error bars represent the standard deviation of the mean. Relative expression levels (according to the equation 2–DeltaCT) were calculated after correction of expression of elongation factor 1 alpha (ef1alpha).

Morpholino injections

For Ol-bsf knockdown experiments, embryos were injected with a splice morpholino: 5′-TTGATGACTGGCCTGCCAACCTGTC-3′ targeting the 3′ splice junction of *Ol-bsf* exon 2 (see S3 Fig). The most efficient dose (4 mg/mL) was experimentally determined and the specificity of the targeting confirmed in control experiments (see S3 Fig).

Microarray

Total RNAs were extracted from adult medaka gonads using the Tri-reagent (Sigma-Aldrich) according to the supplier's recommendation. The total RNA yield was estimated using a Nanodrop ND-1000 spectrophotometer (Labtech, Palaiseau, France), and RNA integrity was checked by means of an Agilent Bioanalyzer (Agilent Technologies, Massy, France). Medaka

gene expression profiling was conducted using an Agilent 8x60K high-density oligonucleotide microarray (GEO platform GPL24100). Labelling and hybridization steps were performed following the 'One-Color Microarray-Based Gene Expression Analysis (Low Input Quick Amp labelling)' Agilent protocol. Briefly, for each sample, 150 ng of total RNA was amplified and labelled using Cy3-CTP. Yield (>825 ng cRNA) and specific activity (>6 pmol of Cy3 per μg of cRNA) of Cy3-cRNA produced were checked with the Nanodrop. The amount of 600 ng of Cy3-cRNA was fragmented and hybridized on a sub-array. Hybridization was carried out for 17 hours at 65 ˚C in a rotating hybridization oven prior to washing and scanning with an Agilent Scanner (Agilent DNA Microarray Scanner, Agilent Technologies, Massy, France) using the standard parameters for a gene expression 8x60K oligo-array (3 μm and 20 bits). Data were then obtained with the Agilent Feature Extraction software (10.7.3.1) according to the appropriate GE protocol (GE1_107_Sep09). The arrays were normalized (scale normalization) and log-transformed using Genespring Software (version 14.5). A *t* test analysis ($p < 0.05$) was employed to determine the genes that were the most differentially expressed between the 2 conditions. Microarray data sets have been deposited to the GEO-NCBI with the accession number GSE 104726. GO was performed using the panther program (http://geneontology.org/).

CRISPR/Cas9 genome editing

Identification of CRISPR/Cas9 target sites and design of oligonucleotides of were performed by the use of the ZiFiT software (http://zifit.partners.org/ZiFiT/Disclaimer.aspx). For preparation of sgRNAs, the DR274 plasmid (Addgene number 42250) was first linearized with *Bsa1*, electrophoresed in a 2% agarose gel, and purified. Pairs of complementary oligonucleotides were annealed (40 mM Tris–HCl [pH 8.0], 20 mM $MgCl_2$, and 50 mM NaCl buffer) by heating at 95 ˚C for 2 minutes and then cooled down slowly to 25 ˚C within 1 hour. The double-stranded oligonucleotides were then ligated into the linearized pDR274 vector. Different sgRNAs were designed to target several sites within the Ol-BSF gene (see S6 Fig) in order to create deletions. After linearization with *Dra1* and *Not1*, respectively, pDR274 and pCS2-nCas9n plasmids were used for generating either sgRNAs or *Cas9* RNAs. The synthesized RNAs were then injected into 1-cell–staged embryos at the following concentrations: 25 ng/μL for each sgRNAs and 100 ng/μL for the *Cas9* mRNA. CRISPR-positive fish were then screened for mutations using PCR primers flanking the site of deletion (see S6 Fig). The inferred mutant protein is presented in S6C Fig. Mutant fish used in this study have been outcrossed for at least 5 generations.

Histology

Gonads were fixed for 48 hours in Bouin-Holland fluid and then dehydrated serially in aqueous 70% and 95% ethanol, ethanol/butanol (5:95), and butanol. Tissues were embedded in paraffin, and 5-μm mid-sagittal gonad sections were stained with Regaud's haematoxylin and haematoxylin–eosin–safran.

Supporting information

S1 Fig. Presence of the D3U-box in different UTRs and coding sequence of gonadal genes. Presence of the D3U-box (matrix-scan) was evaluated within the transcriptome of the medaka fish. D3U-box, *dmrt1* 3′ UTR box.
(DOCX)

S2 Fig. Phylogeny and synteny analysis of *Ol-bsf* and *Ol-cug-bp* ohnologs. (A) Circular cladogram representation of the phylogenetic tree of lrpprc proteins in jawed vertebrates (gnatosthomes). This phylogeny shows that *lrpprc* genes were retained as single copies in most jawed vertebrates even following whole genome duplications (red stars), e.g., the teleost-specific duplication or the salmonid–specific duplication. *Lrpprc* is, however, present in duplicated copies in the Cyprininae (tree branches in red). Bootstraps ($N = 100$) values are indicated in each tree node when judged significant (>0.75). Tree branches are depicted in blue for lobefin vertebrates and cartilaginous fish and in black for teleosts with the exception of Cyprininae in red. (B) Gene evolution of *celf2* genes in some teleosts. The phylogeny on the left is a dendogram representation of *celf2* gene phylogeny in teleosts given as an indication as only a few nodes are supported by good bootstraps' values ($N = 100$, mentioned in each tree nodes when judged significant, i.e., >0.7). The teleost fish whole genomic duplication (3R) is indicated by a red star. The left part of the figure is a representation of the evolution of the genomic context around the *celf2* gene. After the 3R whole genome duplication, *celf2*—which is a single copy gene on the Chr 8 of the spotted gar genome—was duplicated in two 3R ohnologs, *celf2a* and *celf2b*, that were not retained as 2 copies in all teleost fish. The genomic context of the *celf2a* and *celf2b* paralogous regions clearly indicates a partition of the ancestral region found in spotted gar. The *celf2a* gene was retained in all species investigated, but the *celf2b* gene seems to have been lost in Otophysi or at least in *D. rerio* (Cypriniformes), *Astyanax mexicanus* (Characiformes), and *Ictalurus punctatus* (Siluriformes). *celf2*, CUGBP Elav-like family member 2; Chr 8, Chromosome 8; Lrpprc, leucine rich pentatricopeptide repeat containing; Ol-BSF, *Oryzias latipes* Bicoid Stability Factor; Ol-CUG-BP, *Oryzias latipes* CUG-binding protein.
(TIF)

S3 Fig. Analysis of morpholino efficiency and level of Ol-bsf down-regulation. For in vivo transient down-regulation of Ol-bsf, a splice morpholino was designed to encompass the splice junction between exon 2 and intron 2 of the *Ol-bsf* gene in order to induce aberrant splicing and frame shit of the ORF. To show to what extend the splicing/activity of *Ol-bsf* was impacted, RT-PCR using exons 1, 2, and 3 spanning primers together with cDNAs from different stages of morpholino-injected embryos was achieved. E2, exon 2; i2, intron 2; Ol-BSF, *Oryzias latipes* Bicoid Stability Factor; RT-PCR, Reverse Transcription-Polymerase Chain Reaction.
(TIF)

S4 Fig. Real-time PCR quantification of Ol-cug-bp1, Ol-cug-bp2, and Ol-bsf during embryogenesis and in adult tissues. (A and C) During embryonic development, both Ol-cug-bp ohnologs are expressed in a complementary manner. Being likely maternally deposited the expression of Ol-cug-bp1 rapidly decreases after mid-blastula transition (stage 10) to remain virtually off up to hatching stage. On the other hand, low expression of Ol-cug-bp2 is detected until stage 25 and rapidly increases by stage 33. (B and D) In adult tissues, both Ol-cug-bp ohnologs are expressed in brain, muscles, and gonads; ol-cug-bp2 is additionally expressed in eyes and skin. Both ohnologs are higher expressed in male gonads than in female gonads. (E and F) In adult tissues, Ol-bsf is ubiquitously present although higher expression is observed in gonads of both sexes. Underlying data for (A to F) can be found in S1 Data.
(TIF)

S5 Fig. Lrrprc and celf1, but not celf2, are expressed in mouse embryonic gonads. (A to H) ISHs on sagittal sections of 14.5 dpc mouse embryos showed expression of lrrprc (A to D) and celf1 (E to H) most likely in germ cells within testis cords (B and F) and germ cells within the ovary (D and H). In contrast, no celf2 expression was detected in developing gonads (I–L).

However, celf2 expression was detected in other tissues, such as part of the brain and dorsal root ganglia. Scale bars: 1 mm for A, C, E, G, I, and K; 10 mm for B, D, F, H, J, and L. (M–R) Immunofluorescent detection of LRPPRC (M, N, P, Q) and DDX4/VASA (O, R) in adult mouse testes (M–O) and ovaries (P–R). In adult testes, lrpprc is expressed in one subpopulation of germ cells; compared lrpprc staining on (M) and (N) with vasa staining on (O) where most of the germ cells (except some spermatogonia) remain stained by vasa. According to the position of lrpprc-positive cells (arrowheads in M and N) in the seminiferous tubule (not basal and below round spermatids) and to the fact that lrpprc-positive germ cells are those with the largest nucleus, lrpprc-positive cells seem to be spermatocytes at the pachytene stage. In adult ovaries (P–R), lrpprc is mainly expressed into the oocytes of primordial, primary and young secondary follicles (see arrows on [P] and [Q]). Lrpprc staining disappears from the oocyte of secondary follicles that are clearly stained for vasa in (R) (compared stars in [Q] and [R]). Scale bars: 200 μm for M to R. dph, days post hatching; ISH, in situ hybridisation; LRPPRC, leucine rich pentatricopeptide repeat containing.
(TIF)

S6 Fig. Generation *Ol-bsf* knockout medaka fish after genome editing by CRISPR/Cas9 method. (A) Several guide RNA were designed in order to target different locations on the *Ol-bsf* gene (targets 1, 2, and 6). (B) After injection of different combinations of guide RNAs together with the Cas9 mRNA, putative edited fish were subjected to RT-PCR using primer sets flanking the cutting sites. Lines displaying deletions within the *Ol-bsf* gene (red stars) were kept for further investigations. (C) Deletions obtained within the *Ol-bsf* gene (left panel) result in a truncated translated Ol-bsf protein (right panel). CRISPR-Cas9, clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9; Ol-BSF, *Oryzias latipes* Bicoid Stability Factor; RT-PCR, Reverse Transcription- Polymerase Chain Reaction.
(TIF)

S7 Fig. Locomotor activity of *Ol-bsf* mutant fish versus wild type. Locomotor activity (*Ol-bsf* mutants versus wild type) was determined at hatching stage (stage 39, 9 dpf) either under dark conditions (A) or with 10% light (approximately equal to 100 lux [panel B]) by measuring the total distance swum during a period of 5 minutes. (C) The total distance swum for each individual is equal to the sum of distances reached during inactivity, small and large movements. Bars and error bars indicate mean ± standard deviation. $N = 12$ for each condition. Underlying data for (C) can be found in S1 Data. dpf, days post fertilization; Ol-BSF, *Oryzias latipes* Bicoid Stability Factor.
(TIF)

S8 Fig. Ovarian phenotypes of the *Ol-bsf* mutant fish. Morphological inspection of heterozygote mutant ovaries discloses a significant accumulation of small sized-oocytes compared to wild type (A) and (A1 to A5 for details and statistical analyses). (A) Overall size distribution of the oocytes in 9 wild-type and 9 Ol-BSF$^{(+/-)}$ adult ovaries. Each gonad (testes or ovaries) was sectioned through the mid-sagittal plan (see also Materials and methods). Underlying data for (A) can be found in S1 Data. Ol-BSF, *Oryzias latipes* Bicoid Stability Factor.
(TIF)

S9 Fig. Ovarian phenotypes of the Ol-BSF mutant fish. (A to R) Mid-sagittal sections of the ovaries utilized for counting the oocytes in S8 Fig. Each gonad (testes or ovaries) was sectioned through the mid-sagittal plan (see also Materials and methods). Underlying data for (A) can be found in S1 Data. Ol-BSF, *Oryzias latipes* Bicoid Stability Factor.
(TIF)

S10 Fig. Testicular phenotypes of the *Ol-bsf* mutant fish. (A–J compared to K–T) Heterozygote mutant testes (A–J) exhibit a decreased number of spermatogonia with accumulation of type 2 spermatocytes, spermatids, and sperm within the most external layers of the seminiferous epithelium (arrowheads) compared to wild-type testes (K–T). Either 10 different wild-type (A–J) or *Ol-bsf*-deficient (K–T) testes were analysed. Mid-sagittal gonad sections were stained with haematoxylin–eosin–safran. Each gonad (testes or ovaries) was sectioned through the mid-sagittal plan (see also Material and methods). Ol-BSF, *Oryzias latipes* Bicoid Stability Factor.
(TIF)

S11 Fig. Fertility test. Egg numbers and fertilization rates were recorded over a period of 9 days for the following crosses: (A) male *Ol-bsf* (−/+) × female *Ol*-bsf (−/+); (B) male *Ol-bsf* (−/+) × female wild type; (C) male wild type × female *Ol-bsf* (−/+). Underlying data for (A to C) can be found in S1 Data. BSF, bicoid stability factor.
(TIF)

S12 Fig. Microarray data and mitochondrial gene quantification. (A) Microarray. Adult testes of either $bsf^{+/-}$ or wild-type animals were subjected to microarray (see Materials and methods). GO term analysis reveals that in mutant testes partially depleted for the *bsf* gene, rRNA processing is particularly affected. Ol-bsf and Ol-cug-bp2 are down- and up-regulated, respectively, in mutant animals compared to wild type. Of note, and in accordance with the literature, a significant proportion (10.1%) of the down-regulated genes code for proteins localized in the mitochondria. Finally, supporting our observations that lowering *ol-bsf* transcription (morpholino injection in Fig 7E) resulted in up-modulation of germ cell number and that mutant gonads presented an increase of germ cells committing to gametogenesis (Fig 6), our microarray analysis reveals a general up-regulation of genes involved in germ cell proliferation or differentiation. (B) RNA levels of different mitochondrial genes (Cox1, Cox2, ND1, ND5, and CytB) were quantified by real-time PCR after BSF-morpholino injections and compared to wild type. Most of the mitochondrial genes are down-regulated when the level of Ol-bsf decreases. (C) Modulation of RNA levels of the cyp19a1 (aromatase) gene after overexpression of Ol-cugbp1 or Ol-cug-bp2. Underlying data for (B and C) can be found in S1 Data. GO, gene ontology.
(TIFF)

S1 Table. Gene evolution of *cugbp elav-like family member 2* genes in some teleosts. *celf2*, *cugbp elav-like family member 2*.
(DOCX)

S1 Data. Underlying data.
(XLSX)

AcknowledgmentsAcknowledgements

The authors thank the fish facilities at the University of Würzburg and at the INRA Rennes (C. Duret, F. Borel, and A. Patinote).

Author Contributions

Conceptualization: Amaury Herpin, Manfred Schartl.

Formal analysis: Amaury Herpin, Clara Gobé, Aurélie Le Cam, Yann Guiguen.

Funding acquisition: Amaury Herpin.

Investigation: Amaury Herpin, Cornelia Schmidt, Clara Gobé, Martina Regensburger, Aurélie Le Cam, Jérome Montfort, Mateus C. Adolfi, Christina Lillesaar, Dagmar Wilhelm, Michael Kraeussling, Brigitte Mourot, Béatrice Porcon, Laurence Ettwiller, Dirk Dolle, Yann Guiguen, Manfred Schartl.

Methodology: Amaury Herpin, Christina Lillesaar, Maëlle Pannetier, Eric Pailhoux, Laurence Ettwiller, Dirk Dolle.

Project administration: Amaury Herpin.

Software: Susanne Kneitz, Laurence Ettwiller, Dirk Dolle.

Supervision: Amaury Herpin, Manfred Schartl.

Validation: Amaury Herpin.

Writing – original draft: Amaury Herpin, Manfred Schartl.

Writing – review & editing: Amaury Herpin, Manfred Schartl.

References

1. Brennan J, Capel B. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. Nat Rev Genet. 2004; 5: 509–521. https://doi.org/10.1038/nrg1381 PMID: 15211353
2. Devlin RH, Nagahama Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture. 2002; 208: 191–364. https://doi.org/10.1016/S0044-8486(02)00057-1
3. Lin Y-T, Capel B. Cell fate commitment during mammalian sex determination. Curr Opin Genet Dev. 2015; 32: 144–152. https://doi.org/10.1016/j.gde.2015.03.003 PMID: 25841206
4. Rios-Rojas C, Bowles J, Koopman P. On the role of germ cells in mammalian gonad development: quiet passengers or back-seat drivers? Reprod Camb Engl. 2015; 149: R181–191. https://doi.org/10.1530/REP-14-0663 PMID: 25628441
5. Herpin A, Fischer P, Liedtke D, Kluever N, Neuner C, Raz E, et al. Sequential SDF1a and b-induced mobility guides Medaka PGC migration. Dev Biol. 2008; 320: 319–327. https://doi.org/10.1016/j.ydbio.2008.03.030 PMID: 18440502
6. Kurokawa H, Saito D, Nakamura S, Katoh-Fukui Y, Ohta K, Baba T, et al. Germ cells are essential for sexual dimorphism in the medaka gonad. Proc Natl Acad Sci U S A. 2007; 104: 16958–16963. https://doi.org/10.1073/pnas.0609932104 PMID: 17940041
7. Marlow F. Primordial Germ Cell Specification and Migration. F1000Research. 2015; 4. https://doi.org/10.12688/f1000research.6995.1 PMID: 26918157
8. Nakamura S, Kobayashi D, Aoki Y, Yokoi H, Ebe Y, Wittbrodt J, et al. Identification and lineage tracing of two populations of somatic gonadal precursors in medaka embryos. Dev Biol. 2006; 295: 678–688. https://doi.org/10.1016/j.ydbio.2006.03.052 PMID: 16682019
9. Paksa A, Raz E. Zebrafish germ cells: motility and guided migration. Curr Opin Cell Biol. 2015; 36: 80–85. https://doi.org/10.1016/j.ceb.2015.07.007 PMID: 26232877
10. Johnson AD, Richardson E, Bachvarova RF, Crother BI. Evolution of the germ line-soma relationship in vertebrate embryos. Reprod Camb Engl. 2011; 141: 291–300. https://doi.org/10.1530/REP-10-0474 PMID: 21228047
11. Tanaka M, Kinoshita M, Kobayashi D, Nagahama Y. Establishment of medaka (Oryzias latipes) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate. Proc Natl Acad Sci U S A. 2001; 98: 2544–2549. https://doi.org/10.1073/pnas.041315498 PMID: 11226275
12. Herpin A, Schartl M. Dmrt1 genes at the crossroads: a widespread and central class of sexual development factors in fish. FEBS J. 2011; 278: 1010–1019. https://doi.org/10.1111/j.1742-4658.2011.08030.x PMID: 21281449
13. Herpin A, Schartl M. Dmrt1 genes at the crossroads: a widespread and central class of sexual development factors in fish. FEBS J. 2011; 278: 1010–1019. https://doi.org/10.1111/j.1742-4658.2011.08030.x PMID: 21281449
14. Matson CK, Zarkower D. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. Nat Rev Genet. 2012; 13: 163–174. https://doi.org/10.1038/nrg3161 PMID: 22310892

15. Raymond CS, Murphy MW, O'Sullivan MG, Bardwell VJ, Zarkower D. Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev. 2000; 14: 2587–2595. PMID: 11040213

16. Chong T, Collins JJ, Brubacher JL, Zarkower D, Newmark PA. A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. Nat Commun. 2013; 4: 1814. https://doi.org/10.1038/ncomms2811 PMID: 23652002

17. Kato Y, Kobayashi K, Watanabe H, Iguchi T. Environmental sex determination in the branchiopod crustacean Daphnia magna: deep conservation of a Doublesex gene in the sex-determining pathway. PLoS Genet. 2011; 7: e1001345. https://doi.org/10.1371/journal.pgen.1001345 PMID: 21455482

18. Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C, et al. A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis. Proc Natl Acad Sci U S A. 2008; 105: 2469–2474. https://doi.org/10.1073/pnas.0712244105 PMID: 18268317

19. Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P, et al. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 2014; 46: 253–260. https://doi.org/10.1038/ng.2890 PMID: 24487278

20. Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ, et al. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature. 2009; 461: 267–271. https://doi.org/10.1038/nature08298 PMID: 19710650

21. Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, et al. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. Nature. 2002; 417: 559–563. https://doi.org/10.1038/nature751 PMID: 12037570

22. Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A, et al. A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, Oryzias latipes. Proc Natl Acad Sci U S A. 2002; 99: 11778–11783. https://doi.org/10.1073/pnas.182314699 PMID: 12193652

23. Ledig S, Hiort O, Wünsch L, Wieacker P. Partial deletion of DMRT1 causes 46,XY ovotesticular disorder of sexual development. Eur J Endocrinol. 2012; 167: 119–124. https://doi.org/10.1530/EJE-12-0136 PMID: 22573722

24. Raymond CS, Parker ED, Kettlewell JR, Brown LG, Page DC, Kusz K, et al. A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. Hum Mol Genet. 1999; 8: 989–996. PMID: 10332030

25. Kim S, Bardwell VJ, Zarkower D. Cell type-autonomous and non-autonomous requirements for Dmrt1 in postnatal testis differentiation. Dev Biol. 2007; 307: 314–327. https://doi.org/10.1016/j.ydbio.2007.04.046 PMID: 17540358

26. Matson CK, Murphy MW, Griswold MD, Yoshida S, Bardwell VJ, Zarkower D. The mammalian doublesex homolog DMRT1 is a transcriptional gatekeeper that controls the mitosis versus meiosis decision in male germ cells. Dev Cell. 2010; 19: 612–624. https://doi.org/10.1016/j.devcel.2010.09.010 PMID: 20951351

27. Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature. 2011; 476: 101–104. https://doi.org/10.1038/nature10239 PMID: 21775990

28. Herpin A, Schartl M. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 2015; 16: 1260–1274. https://doi.org/10.15252/embr.201540667 PMID: 26358957

29. Herpin A, Schindler D, Kraiss A, Hornung U, Winkler C, Schartl M. Inhibition of primordial germ cell proliferation by the medaka male determining gene Dmrt I bY. BMC Dev Biol. 2007; 7: 99. https://doi.org/10.1186/1471-213X-7-99 PMID: 17760954

30. Krentz AD, Murphy MW, Kim S, Cook MS, Capel B, Zhu R, et al. The DM domain protein DMRT1 is a dose-sensitive regulator of fetal germ cell proliferation and pluripotency. Proc Natl Acad Sci U S A. 2009; 106: 22323–22328. https://doi.org/10.1073/pnas.0905431106 PMID: 20007774

31. Shao C, Li Q, Chen S, Zhang P, Lian J, Hu Q, et al. Epigenetic modification and inheritance in sexual reversal of fish. Genome Res. 2014; 24: 604–615. https://doi.org/10.1101/gr.162172.113 PMID: 24487721

32. Wen AY, You F, Sun P, Li J, Xu DD, Wu ZH, et al. CpG methylation of dmrt1 and cyp19a promoters in relation to their sexual dimorphic expression in the Japanese flounder Paralichthys olivaceus. J Fish Biol. 2014; 84: 193–205. https://doi.org/10.1111/jfb.12277 PMID: 24372528

33. Ge C, Ye J, Zhang H, Zhang Y, Sun W, Sang Y, et al. Dmrt1 induces the male pathway in a turtle species with temperature-dependent sex determination. Dev Camb Engl. 2017; 144: 2222–2233. https://doi.org/10.1242/dev.152033 PMID: 28506988

34. Lei N, Heckert LL. Sp1 and Egr1 regulate transcription of the Dmrt1 gene in Sertoli cells. Biol Reprod. 2002; 66: 675–684. PMID: 11870074

35. Lei N, Heckert LL. Gata4 regulates testis expression of Dmrt1. Mol Cell Biol. 2004; 24: 377–388. https://doi.org/10.1128/MCB.24.1.377-388.2004 PMID: 14673170

36. Cui J, Shen X, Zhao H, Nagahama Y. Genome-wide analysis of Sox genes in Medaka (Oryzias latipes) and their expression pattern in embryonic development. Cytogenet Genome Res. 2011; 134: 283–294. https://doi.org/10.1159/000329480 PMID: 21720154

37. Matsuda M, Shinomiya A, Kinoshita M, Suzuki A, Kobayashi T, Paul-Prasanth B, et al. DMY gene induces male development in genetically female (XX) medaka fish. Proc Natl Acad Sci U S A. 2007; 104: 3865–3870. https://doi.org/10.1073/pnas.0611707104 PMID: 17360444

38. Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, et al. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet. 2010; 6: e1000844. https://doi.org/10.1371/journal.pgen.1000844 PMID: 20169179

39. Schartl M, Schories S, Wakamatsu Y, Nagao Y, Hashimoto H, Bertin C, et al. Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. BMC Biol. 2018; 16: 16. https://doi.org/10.1186/s12915-018-0485-8 PMID: 29378592

40. Herpin A, Nakamura S, Wagner TU, Tanaka M, Schartl M. A highly conserved cis-regulatory motif directs differential gonadal synexpression of Dmrt1 transcripts during gonad development. Nucleic Acids Res. 2009; 37: 1510–1520. https://doi.org/10.1093/nar/gkn1065 PMID: 19139075

41. Mancebo R, Zhou X, Shillinglaw W, Henzel W, Macdonald PM. BSF binds specifically to the bicoid mRNA 3' untranslated region and contributes to stabilization of bicoid mRNA. Mol Cell Biol. 2001; 21: 3462–3471. https://doi.org/10.1128/MCB.21.10.3462-3471.2001 PMID: 11313472

42. De Renzis S, Elemento O, Tavazoie S, Wieschaus EF. Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS Biol. 2007; 5: e117. https://doi.org/10.1371/journal.pbio.0050117 PMID: 17456005

43. Paillard L, Omilli F, Legagneux V, Bassez T, Maniey D, Osborne HB. EDEN and EDEN-BP, a cis element and an associated factor that mediate sequence-specific mRNA deadenylation in Xenopus embryos. EMBO J. 1998; 17: 278–287. https://doi.org/10.1093/emboj/17.1.278 PMID: 9427761

44. Philips AV, Timchenko LT, Cooper TA. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. Science. 1998; 280: 737–741. PMID: 9563950

45. Timchenko LT, Miller JW, Timchenko NA, DeVore DR, Datar KV, Lin L, et al. Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. Nucleic Acids Res. 1996; 24: 4407–4414. PMID: 8948631

46. Paillard L, Legagneux V, Maniey D, Osborne HB. c-Jun ARE targets mRNA deadenylation by an EDEN-BP (embryo deadenylation element-binding protein)-dependent pathway. J Biol Chem. 2002; 277: 3232–3235. https://doi.org/10.1074/jbc.M109362200 PMID: 11707455

47. Moraes KCM, Wilusz CJ, Wilusz J. CUG-BP binds to RNA substrates and recruits PARN deadenylase. RNA N Y N. 2006; 12: 1084–1091. https://doi.org/10.1261/rna.59606 PMID: 16601207

48. Hornung U, Herpin A, Schartl M. Expression of the male determining gene dmrt1bY and its autosomal coorthologue dmrt1a in medaka. Sex Dev Genet Mol Biol Evol Endocrinol Embryol Pathol Sex Determ Differ. 2007; 1: 197–206. https://doi.org/10.1159/000102108 PMID: 18391530

49. Webster KA, Schach U, Ordaz A, Steinfeld JS, Draper BW, Siegfried KR. Dmrt1 is necessary for male sexual development in zebrafish. Dev Biol. 2017; 422: 33–46. https://doi.org/10.1016/j.ydbio.2016.12.008 PMID: 27940159

50. Shiraishi E, Imazato H, Yamamoto T, Yokoi H, Abe S, Kitano T. Identification of two teleost homologs of the Drosophila sex determination factor, transformer-2 in medaka (Oryzias latipes). Mech Dev. 2004; 121: 991–996. https://doi.org/10.1016/j.mod.2004.04.013 PMID: 15210204

51. Amrein H, Gorman M, Nöthiger R. The sex-determining gene tra-2 of Drosophila encodes a putative RNA binding protein. Cell. 1988; 55: 1025–1035. PMID: 3144434

52. Shiraishi E, Yoshinaga N, Miura T, Yokoi H, Wakamatsu Y, Abe S-I, et al. Mullerian inhibiting substance is required for germ cell proliferation during early gonadal differentiation in medaka (Oryzias latipes). Endocrinology. 2008; 149: 1813–1819. https://doi.org/10.1210/en.2007-1535 PMID: 18162517

53. Weidinger G, Stebler J, Slanchev K, Dumstrei K, Wise C, Lovell-Badge R, et al. dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. Curr Biol CB. 2003; 13: 1429–1434. PMID: 12932328

54. Gross-Thebing T, Yigit S, Pfeiffer J, Reichman-Fried M, Bandemer J, Ruckert C, et al. The Vertebrate Protein Dead End Maintains Primordial Germ Cell Fate by Inhibiting Somatic Differentiation. Dev Cell. 2017; 43: 704–715.e5. https://doi.org/10.1016/j.devcel.2017.11.019 PMID: 29257950

55. Hartung O, Forbes MM, Marlow FL. Zebrafish vasa is required for germ-cell differentiation and maintenance. Mol Reprod Dev. 2014; 81: 946–961. https://doi.org/10.1002/mrd.22414 PMID: 25257909

56. Siegfried KR, Nüsslein-Volhard C. Germ line control of female sex determination in zebrafish. Dev Biol. 2008; 324: 277–287. https://doi.org/10.1016/j.ydbio.2008.09.025 PMID: 18930041

57. Siegfried KR. In search of determinants: gene expression during gonadal sex differentiation. J Fish Biol. 2010; 76: 1879–1902. https://doi.org/10.1111/j.1095-8649.2010.02594.x PMID: 20557645

58. Nakamura S, Watakabe I, Nishimura T, Toyoda A, Taniguchi Y, Tanaka M. Analysis of medaka sox9 orthologue reveals a conserved role in germ cell maintenance. PLoS ONE. 2012; 7: e29982. https://doi.org/10.1371/journal.pone.0029982 PMID: 22253846

59. Nishimura T, Herpin A, Kimura T, Hara I, Kawasaki T, Nakamura S, et al. Analysis of a novel gene, Sdgc, reveals sex chromosome-dependent differences of medaka germ cells prior to gonad formation. Dev Camb Engl. 2014; 141: 3363–3369. https://doi.org/10.1242/dev.106864 PMID: 25078651

60. Herpin A, Rohr S, Riedel D, Kluever N, Raz E, Schartl M. Specification of primordial germ cells in medaka (Oryzias latipes). BMC Dev Biol. 2007; 7: 3. https://doi.org/10.1186/1471-213X-7-3 PMID: 17217535

61. Herpin A, Rohr S, Riedel D, Kluever N, Raz E, Schartl M. Specification of primordial germ cells in medaka (Oryzias latipes). BMC Dev Biol. 2007; 7: 3. https://doi.org/10.1186/1471-213X-7-3 PMID: 17217535

62. Herpin A, Adolfi MC, Nicol B, Hinzmann M, Schmidt C, Klughammer J, et al. Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination. Mol Biol Evol. 2013; 30: 2328–2346. https://doi.org/10.1093/molbev/mst130 PMID: 23883523

63. Nakamura S, Kobayashi K, Nishimura T, Higashijima S, Tanaka M. Identification of germline stem cells in the ovary of the teleost medaka. Science. 2010; 328: 1561–1563. https://doi.org/10.1126/science.1185473 PMID: 20488987

64. Uribe MC, Grier HJ, Mejía-Roa V. Comparative testicular structure and spermatogenesis in bony fishes. Spermatogenesis. 2015; 4. https://doi.org/10.4161/21565562.2014.983400 PMID: 26413405

65. Keene JD. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet. 2007; 8: 533–543. https://doi.org/10.1038/nrg2111 PMID: 17572691

66. Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science. 2005; 309: 1514–1518. https://doi.org/10.1126/science.1111443 PMID: 16141059

67. Arvola RM, Weidmann CA, Tanaka Hall TM, Goldstrohm AC. Combinatorial control of messenger RNAs by Pumilio, Nanos and Brain Tumor Proteins. RNA Biol. 2017; 1–12. https://doi.org/10.1080/15476286.2017.1306168 PMID: 28318367

68. Lasko P. Posttranscriptional regulation in Drosophila oocytes and early embryos. Wiley Interdiscip Rev RNA. 2011; 2: 408–416. https://doi.org/10.1002/wrna.70 PMID: 21957026

69. Lehmann R. Germ Plasm Biogenesis—An Oskar-Centric Perspective. Curr Top Dev Biol. 2016; 116: 679–707. https://doi.org/10.1016/bs.ctdb.2015.11.024 PMID: 26970648

70. Herpin A, Adolfi MC, Nicol B, Hinzmann M, Schmidt C, Klughammer J, et al. Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination. Mol Biol Evol. 2013; 30: 2328–2346. https://doi.org/10.1093/molbev/mst130 PMID: 23883523

71. Sterky FH, Ruzzenente B, Gustafsson CM, Samuelsson T, Larsson N-G. LRPPRC is a mitochondrial matrix protein that is conserved in metazoans. Biochem Biophys Res Commun. 2010; 398: 759–764. https://doi.org/10.1016/j.bbrc.2010.07.019 PMID: 20633537

72. Kress C, Gautier-Courteille C, Osborne HB, Babinet C, Paillard L. Inactivation of CUG-BP1/CELF1 causes growth, viability, and spermatogenesis defects in mice. Mol Cell Biol. 2007; 27: 1146–1157. https://doi.org/10.1128/MCB.01009-06 PMID: 17130239

73. Gohil VM, Nilsson R, Belcher-Timme CA, Luo B, Root DE, Mootha VK. Mitochondrial and nuclear genomic responses to loss of LRPPRC expression. J Biol Chem. 2010; 285: 13742–13747. https://doi.org/10.1074/jbc.M109.098400 PMID: 20220140

74. Sasarman F, Brunel-Guitton C, Antonicka H, Wai T, Shoubridge EA, LSFC Consortium. LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. Mol Biol Cell. 2010; 21: 1315–1323. https://doi.org/10.1091/mbc.E10-01-0047 PMID: 20200222

75. Bratic A, Wredenberg A, Grönke S, Stewart JB, Mourier A, Ruzzenente B, et al. The bicoid stability factor controls polyadenylation and expression of specific mitochondrial mRNAs in Drosophila melanogaster. PLoS Genet. 2011; 7: e1002324. https://doi.org/10.1371/journal.pgen.1002324 PMID: 22022283

76. Sasarman F, Nishimura T, Antonicka H, Weraarpachai W, Shoubridge EA, LSFC Consortium. Tissue-specific responses to the LRPPRC founder mutation in French Canadian Leigh Syndrome. Hum Mol Genet. 2015; 24: 480–491. https://doi.org/10.1093/hmg/ddu468 PMID: 25214534

77. Boulanger G, Cibois M, Viet J, Fostier A, Deschamps S, Pastezeur S, et al. Hypogonadism Associated with Cyp19a1 (Aromatase) Posttranscriptional Upregulation in Celf1 Knockout Mice. Mol Cell Biol. 2015; 35: 3244–3253. https://doi.org/10.1128/MCB.00074-15 PMID: 26169831

78. Iwamatsu T. Stages of normal development in the medaka Oryzias latipes. Mech Dev. 2004; 121: 605–618. https://doi.org/10.1016/j.mod.2004.03.012 PMID: 15210170

79. Hong Y, Winkler C, Schartl M. Production of medakafish chimeras from a stable embryonic stem cell line. Proc Natl Acad Sci U S A. 1998; 95: 3679–3684. PMID: 9520425

80. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008; 36: W465–469. https://doi.org/10.1093/nar/gkn180 PMID: 18424797

81. Pasquier J, Cabau C, Nguyen T, Jouanno E, Severac D, Braasch I, et al. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC Genomics. 2016; 17: 368. https://doi.org/10.1186/s12864-016-2709-z PMID: 27189481

82. Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol. 1997; 14: 685–695. https://doi.org/10.1093/oxfordjournals.molbev.a025808 PMID: 9254330

83. Zhang H, Gao S, Lercher MJ, Hu S, Chen W-H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. Nucleic Acids Res. 2012; 40: W569–572. https://doi.org/10.1093/nar/gks576 PMID: 22695796

84. Louis A, Muffato M, Roest Crollius H. Genomicus: five genome browsers for comparative genomics in eukaryota. Nucleic Acids Res. 2013; 41: D700–705. https://doi.org/10.1093/nar/gks1156 PMID: 23193262

85. Lange M, Norton W, Coolen M, Chaminade M, Merker S, Proft F, et al. The ADHD-linked gene Lphn3.1 controls locomotor activity and impulsivity in zebrafish. Mol Psychiatry. 2012; 17: 855. https://doi.org/10.1038/mp.2012.119 PMID: 22918194

86. Grabher C, Wittbrodt J. Meganuclease and transposon mediated transgenesis in medaka. Genome Biol. 2007; 8 Suppl 1: S10. https://doi.org/10.1186/gb-2007-8-s1-s10 PMID: 18047687

87. Köprunner M, Thisse C, Thisse B, Raz E. A zebrafish nanos-related gene is essential for the development of primordial germ cells. Genes Dev. 2001; 15: 2877–2885. https://doi.org/10.1101/gad.212401 PMID: 11691838

Reconstruction of the birth of a male sex chromosome present in Atlantic herring

Nima Rafati[a,b,1], Junfeng Chen[a,1], Amaury Herpin[c,d,1], Mats E. Pettersson[a,1], Fan Han[a], Chungang Feng[a], Ola Wallerman[a], Carl-Johan Rubin[a], Sandrine Péron[c], Arianna Cocco[a], Mårten Larsson[a], Christian Trötschel[e], Ansgar Poetsch[e,f,g], Kai Korsching[h], Wolfgang Bönigk[h], Heinz G. Körschen[h], Florian Berg[i,j], Arild Folkvord[i,j], U. Benjamin Kaupp[h,k], Manfred Schartl[l,m,n], and Leif Andersson[a,o,p,2]

[a]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, 75123 Uppsala, Sweden; [b]Science for Life Laboratory, National Bioinformatics Infrastructure Sweden, Uppsala University, 75237 Uppsala, Sweden; [c]UR 1037 Fish Physiology and Genomics, INRAE, F-35000 Rennes, France; [d]State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences, Hunan Normal University, Changsha 410081, China; [e]Fakultät für Biologie und Biotechnologie, Ruhr-Universität Bochum, 44801 Bochum, Germany; [f]Center for Marine and Molecular Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), 266237 Qinqdao, China; [g]College of Marine Life Sciences, Ocean University of China, 266100 Qingdao, China; [h]Molecular Sensory Systems, Center of Advanced European Studies and Research, 53175 Bonn, Germany; [i]Department of Biological Sciences, University of Bergen, 5020 Bergen, Norway; [j]Institute of Marine Research, 5817 Bergen, Norway; [k]Life & Medical Sciences Institute, University of Bonn, 53115 Bonn, Germany; [l]Developmental Biochemistry, Biocenter, University of Würzburg, 97074 Würzburg, Germany; [m]Hagler Institute for Advanced Study, Texas A&M University, College Station, TX 77843; [n]Xiphophorus Genetic Stock Center, Texas State University, San Marcos, TX 78666; [o]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843; and [p]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

Contributed by Leif Andersson, August 11, 2020 (sent for review May 20, 2020; reviewed by Henrik Kaessmann and Catherine L. Peichel)

The mechanisms underlying sex determination are astonishingly plastic. Particularly the triggers for the molecular machinery, which recalls either the male or female developmental program, are highly variable and have evolved independently and repeatedly. Fish show a huge variety of sex determination systems, including both genetic and environmental triggers. The advent of sex chromosomes is assumed to stabilize genetic sex determination. However, because sex chromosomes are notoriously cluttered with repetitive DNA and pseudogenes, the study of their evolution is hampered. Here we reconstruct the birth of a Y chromosome present in the Atlantic herring. The region is tiny (230 kb) and contains only three intact genes. The candidate male-determining gene *BMPR1BBY* encodes a truncated form of a BMP1B receptor, which originated by gene duplication and translocation and underwent rapid protein evolution. BMPR1BBY phosphorylates SMADs in the absence of ligand and thus has the potential to induce testis formation. The Y region also contains two genes encoding subunits of the sperm-specific Ca^{2+} channel CatSper required for male fertility. The herring Y chromosome conforms with a characteristic feature of many sex chromosomes, namely, suppressed recombination between a sex-determining factor and genes that are beneficial for the given sex. However, the herring Y differs from other sex chromosomes in that suppression of recombination is restricted to an ∼500-kb region harboring the male-specific and sex-associated regions. As a consequence, any degeneration on the herring Y chromosome is restricted to those genes located in the small region affected by suppressed recombination.

sex determination | BMPR1 | CatSper | gene duplication | molecular evolution

The mechanisms underlying sex determination are astonishingly plastic. Particularly the triggers for the molecular machinery, which recalls either the male or female program, are highly variable and have evolved independently and repeatedly (1–8). Fish show a rich diversity of sex determination systems, including both genetic and environmental triggers (1, 9, 10). Adaptive hypotheses prevail in explaining the evolution of the various sex-determining triggers that guarantee a stable sex ratio of the population. Genetic sex determination is generally explained to be stabilized through the evolution of sex chromosomes, when sexually antagonistic selection links the genetic sex determiner to a locus that is beneficial to the same sex or even antagonistic to the opposite sex (11–15).

The Atlantic herring (*Clupea harengus*) is one of the most abundant vertebrates on Earth. Neither its karyotype (16) nor the male and female linkage map (17) indicated the presence of sex chromosomes. Unexpectedly, our analysis of the Atlantic herring genome revealed a Y chromosome having a minute male-specific region (∼230 kb), a small sex-associated region (∼300 kb), and a large pseudoautosomal region (∼30 Mb) corresponding to chromosome 8 in the reference assembly (17). Here we reconstruct the birth of this chromosome involving two gene duplications followed by rapid protein evolution of the candidate male-determining gene *BMPR1BBY*.

Results

Identification of the Sex-Determining Region. The sex-associated region was identified by genome-wide association analysis using whole-genome sequencing data aligned to the reference genome (17). We compared SNP (single nucleotide polymorphism) allele

Significance

Understanding the evolution of sex determination mechanisms and sex chromosomes is of fundamental importance in biology. Here we have reconstructed the evolution of the sex-determining region in the Atlantic herring. The region is small and contains only three intact genes. The candidate sex-determining factor *BMPR1BBY* is an evolutionary innovation in the herring lineage. It encodes a truncated form of a BMP type I receptor, which originated by gene duplication and underwent rapid protein evolution. The receptor has maintained its kinase activity and has the potential to induce development of testis. The other two genes in the sex-determining region, *CATSPERG* and *CATSPER3Y*, are male beneficial genes because they encode proteins predicted to be essential for sperm to fertilize the egg.

Author contributions: N.R., A.H., U.B.K., M.S., and L.A. designed research; N.R., J.C., A.H., M.E.P., F.H., C.F., O.W., C.-J.R., S.P., A.C., M.L., C.T., A.P., K.K., W.B., and H.G.K. performed research; F.B. and A.F. contributed new reagents/analytic tools; N.R., J.C., A.H., M.E.P., F.H., O.W., C.-J.R., M.L., C.T., A.P., K.K., W.B., H.G.K., U.B.K., M.S., and L.A. analyzed data; and N.R., J.C., A.H., M.E.P., U.B.K., M.S., and L.A. wrote the paper.

Reviewers: H.K., Heidelberg University; and C.L.P., University of Bern.

The authors declare no competing interest.

[1]N.R., J.C., A.H., and M.E.P. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: leif.andersson@imbim.uu.se.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2009925117/-/DCSupplemental.

frequencies between eight males and eight females by χ^2 analysis (Dataset S1); 765 SNPs spanning ~160 kb on chromosome 8 showed significant allele frequency differences ($P \leq 1.0 \times 10^{-10}$) (*SI Appendix*, Fig. S1*A*). Analysis of SNP genotypes across the sex-associated region for the 16 fish revealed that males were heterozygous (XY), whereas females were homozygous (XX) for most sex-associated SNPs (*SI Appendix*, Fig. S1*B*). Thus, males are the heterogametic sex. An additional 17 Atlantic herring of unknown sex followed this pattern perfectly with individuals either being heterozygous or homozygous across the sex-associated region. The sequence of this region in the current assembly (17) represents the female haplotype (*SI Appendix*, Fig. S1*B*). In conclusion, herring chromosome 8 can now be split into chromosome (Chr) X and Y.

To characterize the male-specific region on ChrY, we examined the sequence of all unplaced scaffolds from the current assembly (*Methods*) (17). Two unplaced scaffolds, 18 and 229, showed regions of high identity with the female-specific sequence on ChrX (*SI Appendix*, Fig. S1 *C* and *D*). These regions were interrupted by male-specific sequences that are present in only one copy in males (Fig. 1 *A* and *B*). We also analyzed data from six individually sequenced Pacific herring of unknown sex (18) and found that four had no coverage in the male-specific regions, whereas two had a single copy of the male-specific regions (Fig. 1*B*). The results provide evidence that the male-specific region evolved prior to the divergence of the Atlantic and Pacific herring that separated ~2 My before present (19).

We used the Xdrop method (20) for targeted resequencing of long DNA molecules (up to ~100 kb) from the borders between the male-specific and the flanking region. The alignment of these reads to our ChrY assembly confirmed a correct orientation and did not reveal any gaps in the alignment (*SI Appendix*, Fig. S1*D*).

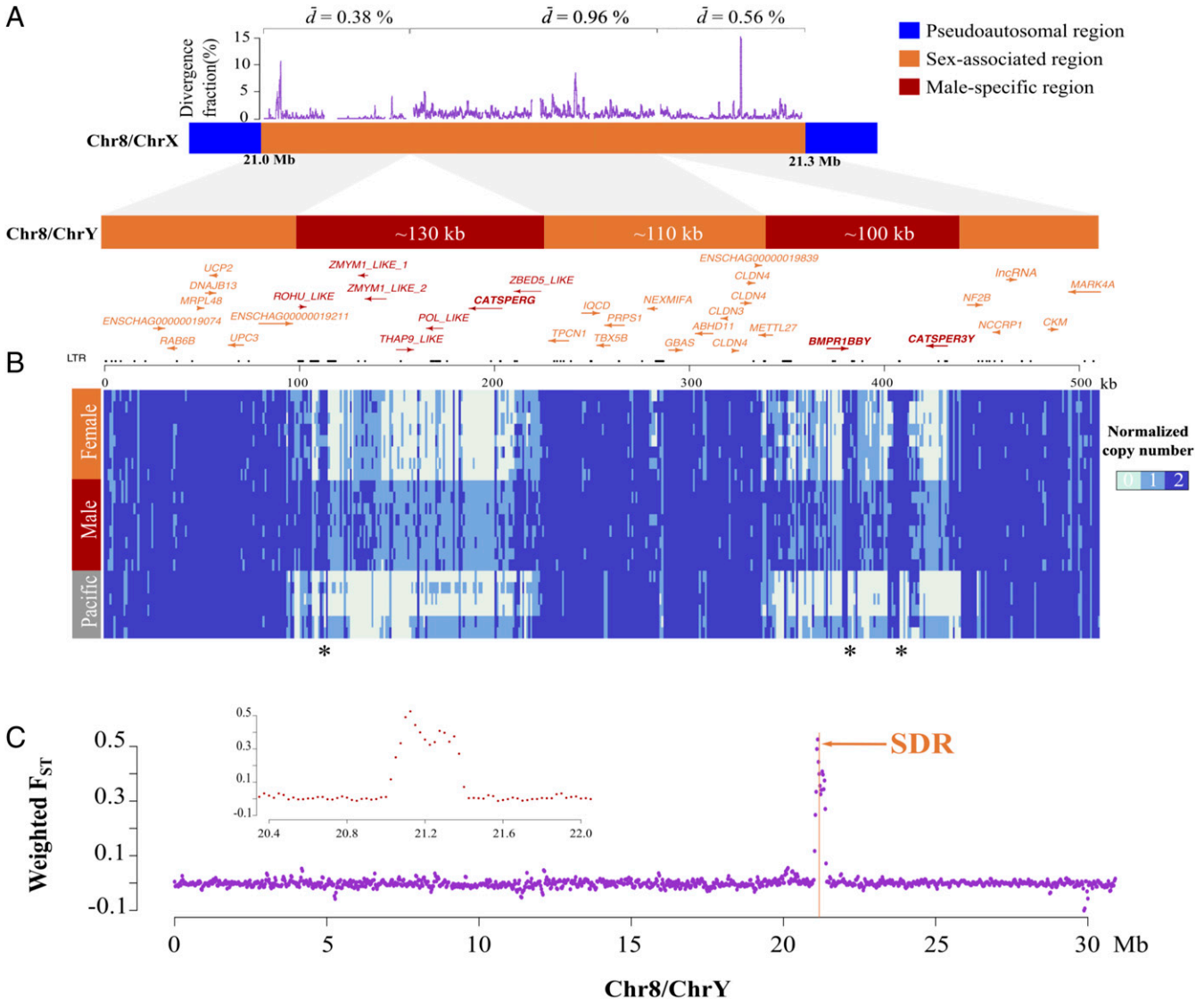


Fig. 1. Assembly of chromosome Y in Atlantic herring. (*A*) Assembly and annotation of male-specific region (red), sex-associated region (orange), and flanking pseudoautosomal region (blue). Some genes (annotated as "_like") within the male-specific region show partial similarity with known proteins. LTR track, long terminal repeats. d, nucleotide sequence divergence between ChrX and Y in 500-bp windows. (*B*) Normalized copy number of sequenced individuals (eight female and eight male Atlantic herring and six Pacific herring of unknown sex) across the male-specific and sex-associated regions. Regions with high coverage marked with an asterisk overlap repetitive sequences. (*C*) Genetic differentiation (F_{ST}) in 50-kb overlapping windows between males (n = 8) and females (n = 8) across ChrY. The SDR is indicated by a vertical line. The male-specific region was not included in this analysis, since it is missing on ChrX. (*Inset*) F_{ST} values in the near vicinity of the SDR.

Our assembly of the ~230-kb male-specific region of ChrY shows that the male-specific sequence is split into two parts (Fig. 1*A*, in red) separated by an ~110-kb sex-associated region (Fig. 1*B*, in orange) that is well assembled and annotated (17). It contains 13 genes that are present on both X and Y, as indicated by equal sequence depth in males and females (Fig. 1*B*). The region showing strong genetic differentiation between males and females is remarkably sharp and only extends for ~300 kb from ~21.0 to ~21.3 Mb on ChrX (Fig. 1*C*), and the male-specific region together with the sex-associated region extends for ~500 kb on ChrY (*SI Appendix*, Fig. S1). These results imply that there is no meiotic recombination in the male-specific regions (red) and no or at least suppressed recombination in the sex-associated regions (orange) (Fig. 1*A*). The genes located in the region showing strong genetic differentiation between X and Y are listed in *SI Appendix*, Table S1. The nucleotide distance between X and Y in the sex-associated region located between the male-specific regions (d = 0.96%; Fig. 1*A*) is only 2.5-fold higher than the estimated average nucleotide distance between Atlantic and Pacific herring (~0.4%), implying either a divergence time of ~5 My or ongoing genetic exchange in the sex-associated region.

The male-specific region contains only three intact genes—*CATSPERG*, *CATSPER3Y*, and *BMPR1BBY*—that all are of high interest for male function (Fig. 1*A*). In addition, we identified other genes annotated ab initio, but these were not well supported by RNA-seq data from gonadal tissue; one of these, POL-like, is of retroviral origin. *CATSPERG* and *CATSPER3Y* encode subunits of the sperm-specific Ca^{2+} channel CatSper that is required for sperm function and male fertility in mouse, humans, and sea urchin (21–25). A phylogenetic tree analysis shows that the substitution rates of *CATSPERG* in Atlantic herring and its orthologs in other teleost are similar, consistent with a maintained function as a CatSper subunit (*SI Appendix*, Fig. S2). *CATSPER3* exists in two copies, *CATSPER3Y* in the male-specific region and a pseudoautosomal copy, *CATSPER3A*, ~1 Mb apart from the Y copy (Chr8, 22,259,380 to 22,268,493 bp). We estimate that *CATSPER3Y* arose by duplication about 15 Mya (*SI Appendix*, Fig. S3). Although the two *CATSPER3* genes in herring are closely related, *CATSPER3Y* is more well conserved compared with paralogs in other species (*SI Appendix*, Fig. S3), suggesting that its function is maintained. By contrast, two findings suggest that *CATSPER3A* is nonfunctional: its exon–intron organization is not intact, and the dN/dS ratio of nonsynonymous (dN) to synonymous (dS) substitutions is close to 1.0, which is characteristic for pseudogenes.

The third gene in the male-specific region *BMPR1BBY* encodes a severely truncated form of BMPR1B (bone morphogenetic protein receptor 1B), a transmembrane protein that hosts an intracellular Ser/Thr kinase domain and is a member of the TGFβ receptor family. Components of TGFβ signaling control the development of male gonads (4). Sequence comparison of *BMPR1BBY* with the paralogs *BMPR1BA* and *BMPR1BB* on chromosomes 7 and 21, respectively, demonstrated that *BMPR1BBY* is more closely related to *BMPR1BB* (*SI Appendix*, Fig. S4). Transcriptome analysis of testis RNA revealed two isoforms with different 5′ start sites (S- and L-BMPR1BBY) containing only coding sequences corresponding to exon 7 to 11 of the autosomal paralogs (*SI Appendix*, Fig. S4). In summary, *BMPR1BBY* is a truncated copy of *BMPR1BB* that has been translocated to chromosome 8; it is a strong candidate for being the male sex-determining factor.

We genotyped 200 individuals with known sex for the presence of *BMPR1BBY*. The high consistency between the X/Y genotype and sex phenotype (94%; *SI Appendix*, Table S2) demonstrates strong association to sex determination. Eleven out of 12 discrepancies between genotype and phenotype concerned XY females carrying *BMPR1BBY*, suggesting that the herring Y may not show full penetrance. Such incomplete penetrance of a young genetic sex determination system has also been observed in the northern pike (26).

Tissue Expression of Genes from the Male-Specific Region. We compared the tissue expression of *BMPR1BBY*, *CATSPER3Y*, *CATSPERG*, and the autosomal copies *BMPR1BB* and *CATSPER3A*, as well as two other *TGFβ* signaling genes, the anti-Müllerian hormone (*AMH*) and its cognate receptor (*AMHR2*), involved in vertebrate sexual development. No expression of *CATSPER3A* was detected. The comparison revealed that the

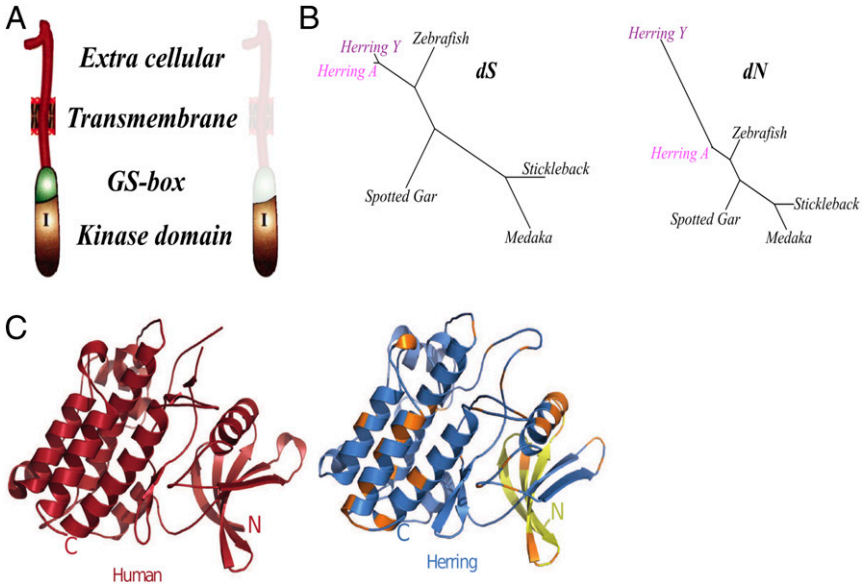


Fig. 2. Evolution of the ChrY-specific *BMPR1BBY* gene and protein. (*A*) Schematic representation of herring BMPR1BB and BMPR1BBY proteins showing the architecture of BMP receptors and highlighting the absence of the extracellular, transmembrane and GS-box domains in the truncated BMPR1BBY. (*B*) dS and dN phylogenetic trees of *BMPR1BB* (Herring A) and *BMPR1BBY* (Herring Y) compared to *BMPR1BB* in other species. (*C*) Crystal structure of the intracellular domain of human BMPR1B (PDB 3MDY, amino acids 206 to 500, red) and a model of herring BMPR1BBY (blue). Yellow color represents the N-terminal extension of L-BMPR1BBY compared to S-BMPR1BBY. Orange color indicates amino acid substitutions between BMPR1BB and BMPR1BBY. N, N terminus; C, C terminus.

three Y-specific genes were only expressed in males as expected (*SI Appendix*, Fig. S5). *BMPR1BBY* and, surprisingly, also *CATSPER3Y* were expressed in both testis and brain, whereas *CATSPERG* showed testis-specific expression. The autosomal copy of *BMPR1BB* was more broadly expressed with the highest level in gonads; female gonads showed higher transcript levels than male gonads (*SI Appendix*, Fig. S5*B*). As expected, *AMH* and *AMHR2* were highly expressed in male and female gonads (*SI Appendix*, Fig. S5 *D* and *F*), consistent with their important roles in sex determination and gonad differentiation.

Rapid Protein Evolution of BMPR1BBY. The predicted structure of the autosomal herring BMPR1BB and the canonical vertebrate BMPR1 receptor is similar (Fig. 2*A*) (27). The *BMPR1BBY* transcripts are lacking exons 1 to 6 of the autosomal *BMPR1BB* (*SI Appendix*, Fig. S4) that encode the extracellular, transmembrane, and part of the intracellular domain (Fig. 2*A*). These domains mediate ligand binding, receptor activation, and membrane localization and entail the phosphorylation sites of the GS loop. Amino acid alignment of BMPR1BBY and BMPR1BB from herring and other fish reveals more than 40 changes that have occurred since *BMPR1BBY* arose (*SI Appendix*, Fig. S4*C*). The rate of dS and dN substitutions between *BMPR1BBY* and *BMPR1BB* is 0.12 and 0.07, respectively, yielding a dN/dS ratio of 0.58, a dramatic and highly significant ($P < 10^{-21}$) increase compared with dN/dS ratios of 0.01 to 0.02 in other branches of the tree (Fig. 2*B* and *SI Appendix*, Fig. S6). We estimate that *BMPR1BBY* arose by duplication about 20 Mya (*SI Appendix*, Fig. S6).

We built a structural model of the intracellular domain of herring BMPR1BBY and BMPR1BB based on human BMPR1B (Fig. 2*C*). Both herring BMPR1BB protein folds are essentially identical to that of the human paralog (Cα root-mean-square deviation of 0.008, 0.077, and 0.073 Å for BMPR1BB, L-BMPR1BBY, and S-BMPR1BBY, respectively; *SI Appendix*, Fig. S7 and Table S5), suggesting that the truncated form may still display Ser/Thr kinase activity.

To conclude, *BMPR1BBY* shows an accelerated protein evolution typical for sex-determination genes (28) but appears to maintain the BMPR1B protein fold.

Physiological Activity of BMPR1BBY. TGFβ signaling by the anti-Müllerian hormone (AMH) system is critically involved in sex determination (29). BMPR1BB, a type-I TGFβ receptor, is activated by ligand-bound type-II receptor and, in turn, phosphorylates the transcription factor SMAD. AMH-induced signaling occurs primarily via SMAD5 phosphorylation and to some extent via SMAD1 and SMAD8 (30). To probe the kinase activity of truncated BMPRBBY, we established an in vitro reporter system for SMAD activation (Fig. 3*A* and *Methods*). The system consists of a Gal4 transcriptional activator coupled to the transactivation domains of SMAD1 and SMAD5, which are specific for AMH signaling (31), and a UAS-Luc reporter driving luciferase expression. We tested three different scenarios: 1) the canonical AMH-signaling machinery including herring AMH, AMHR2, and BMPR1BB; 2) the canonical AMH machinery for which BMPR1BB has been replaced by either S-BMPR1BBY or L-BMPR1BBY; and 3) the canonical AMH machinery supplemented by either S-BMPR1BBY or L-BMPR1BBY (Fig. 3 *B*–*E*). Phosphorylation levels of SMAD1 did not differ significantly when BMPR1BB was replaced by either S-BMPR1BBY or L-BMPR1BBY (Fig. 3 *B* and *D*). By contrast, S-BMPR1BBY is fourfold more active than BMPR1BB in SMAD5 phosphorylation (Fig. 3*C*). Both SMAD1 and SMAD5 activity were enhanced when both BMPR1BB and S-BMPR1BBY were present (Fig. 3 *B* and *C*). No significant increase of SMAD1 and only low increase of SMAD5 phosphorylation were recorded after addition of L-BMPR1BBY (Fig. 3 *D* and *E*). Thus, S- and L-BMPR1BBY can substitute for BMPR1BB in activating SMAD1 and SMAD5.

Next, we examined whether S- and L-BMPR1BBY synergize with canonical AMH signaling. To this end, AMHR2, BMPR1BB, and S- or L-BMPR1BBY were tested either in the presence or absence of the AMH ligand. While canonical AMH signaling (AMH, AMHR2, and BMPR1BB) was clearly stimulated by AMH (*SI Appendix*, Fig. S8*A*), AMH did not modulate S- or L-BMPR1BBY-mediated signaling when associated with AMHR2 (*SI Appendix*, Fig. S8 *B* and *C*). Moreover, S- or L-BMPR1BBY–mediated signaling via SMAD1 phosphorylation was unaffected when AMH was expressed without its cognate receptor AMHR2 (*SI Appendix*, Fig. S8 *D* and *E*), indicating that BMPR1BBY-mediated signaling is not further modulated by the AMH ligand. Finally, in the absence of AMH, the presence or absence of AMHR2 did not affect S- or L-BMPR1BBY–induced signaling (Fig. 3*F*), while differential synergistic effects were observed for BMPR1BBYs when expressed together with the whole AMH machinery (AMH ligand, AMHR2, and BMPR1BB; Fig. 3 *B*–*E*).

In summary, BMPR1BBY has maintained its enzymatic activity despite the severe truncation and the accelerated evolution at the protein level. S- and L-BMPR1BBY–mediated signaling requires neither the presence of the AMH ligand nor the AMHR2 receptor or the autosomal BMPR1BB, yet is able to enhance canonical AMH signaling. Thus, the Y-linked receptor has the enzymatic features to act as an upstream regulator of male gonad development.

CatSper3Y and CatSperG Are Essential Components of the Herring Sperm Proteome. The location of two CatSper genes on the herring Y chromosome is intriguing because CatSper has been thought to be lost in several metazoan lineages including fish (32). We revisited this issue and identify here CatSper genes in many but not all fish genomes (Fig. 4 *A* and *B* and Dataset S2). CatSper is a uniquely complex Ca^{2+} channel that consists of at least nine different subunits CatSper1–CatSper4, CatSperB, CatSperG, CatSperD, CatSperE, and CatSperZ. The Atlantic herring genome contains all CatSper genes except CatSperZ (33) (Dataset S2). Using mass spectrometry (MS), seven CatSper subunits including CatSper3 and CatSperG were identified in mature herring sperm (Datasets S3 and S4). Signaling pathways control CatSper and, thereby, sperm motility, navigation, and fertilization from marine invertebrates to mammals (34–36). The CatSper channel is gated open by membrane voltage and alkaline pH_i. Three signaling events are key to CatSper activation in sperm of the sea urchin *Arbacia punctulata*: hyperpolarization of the membrane potential V_m by a K^+ channel CNGK (37, 38) followed by repolarization via hyperpolarization-activated and cyclic nucleotide-gated (HCN) channels, so-called pacemaker channels (39, 40), and alkalization by a sperm-specific voltage-gated sodium/proton exchanger (sNHE or SLC9C1) (41). The activity of the SLC9C1 exchanger and HCN channel is modulated by cAMP, which is synthesized by a soluble adenylate cyclase (sAC or *adcy10* gene). Remarkably, all these signaling proteins upstream of CatSper are present in herring sperm, suggesting that CatSper and Ca^{2+} also control motility of herring sperm (Fig. 4*C* and Dataset S4). Two functionally important variations are striking. First, herring sperm carry an HCN-like (HCNL) channel that is distinct from classic HCN channels and falls into a new gene subfamily (42). HCNL1 in zebrafish sperm, unlike classic HCN channels, is proton-selective and not cAMP-sensitive (42). Although herring sperm carry the HCNL2 isoform that has not yet been characterized functionally, structural features underlying proton permeation are conserved between HCNL1 and HCNL2 (34). Second, the zebrafish CNGK channel is also gated not by cyclic nucleotides but by alkaline pH_i, indicating that the cellular messengers in fish sperm are protons rather than cAMP or cGMP (42, 43). These results suggest that herring sperm employ a signaling pathway for fertilization that combines signaling proteins and mechanisms from marine invertebrates, zebrafish, and mammals. As more signaling molecules

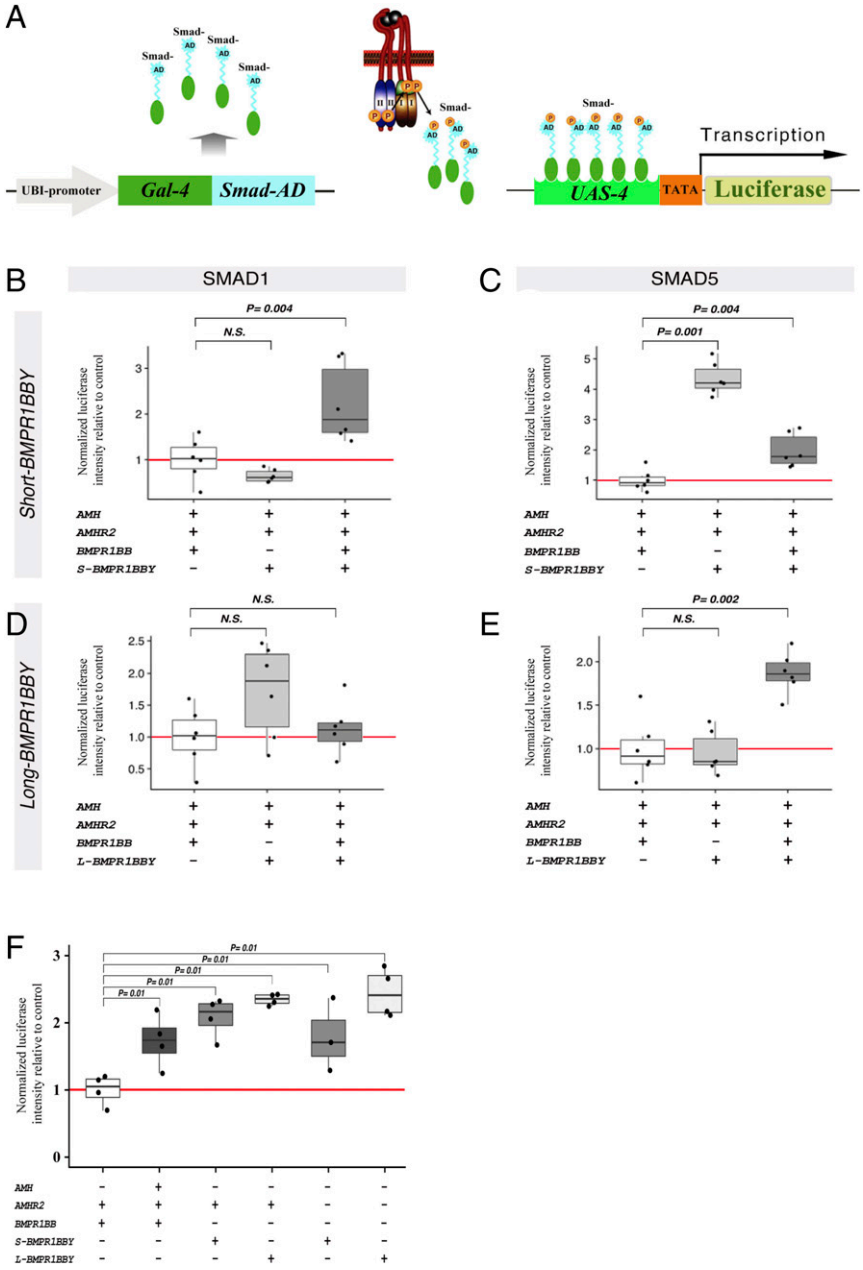


Fig. 3. Analysis of BMPR1BBY signaling. (*A*) Schematic of the UAS-Gal4-Smad-AD reporter assay for quantifying the degree of SMAD1 and SMAD5 phosphorylation upon application of different combinations of TGFβ ligands and receptors (see *Methods* for detailed information). In absence of AMHR2 or BMPR1BBY-induced stimulations, the Gal4-Smad-AD fusion protein remains in the cytoplasm and therefore does not activate transcription of the UAS-4 promoter driving expression of the luciferase. In this situation, only background luciferase expression levels are recorded and set a blank value. In presence of an active type I receptor, the Gal4-Smad-AD fusion proteins are selectively phosphorylated by the receptor through the Smad-AD domain and then translocate to the nucleus. Here they bind and transactivate the UAS-4 promoter through its Gal4 DNA binding domain, which leads to transcription of the luciferase gene. The resulting luciferase activity thus strictly correlates to the degree of activation (phosphorylation) of the Smads, which in turn is a measure of the specific upstream signals elicited by the various combinations of ligands and receptors. AD, specific transactivation domains of the different SMADS. (*B–E*) Normalized luciferase intensities relative to controls after transfection of different combinations of herring AMH, AMHR2, BMPR1BB, and S- and L-BMPR1BBY. (*F*) Normalized luciferase intensities relative to controls using the SMAD1 reporter after transfection of different combinations of herring AMH, AMHR2, BMPR1BB, and S- and L-BMPR1BBY.

are identified, the herring data strengthen an emerging picture of evolutionary significance: sperm preferentially adopt unique, sperm- and species-specific signaling molecules and pathways that are tuned to their respective habitat of reproduction (35).

In summary, we conclude that the herring Y-specific region contains genes encoding two proteins, CatSperG and CatSper3Y, expected to be essential for sperm function based on the functional conservation of the CatSper channel in both invertebrates and vertebrates.

Discussion

BMPR1BBY is a very strong candidate sex-determining factor in the Atlantic herring. It is clearly the best candidate out of the only three genes in the male-specific region because of the well-established role for TGF/BMPR signaling in sexual development (29). Its presence is strongly correlated with the male phenotype, it shows an accelerated protein evolution, and it can replace the AMH/BMPR1B/AMHR pathway for SMAD phosphorylation.

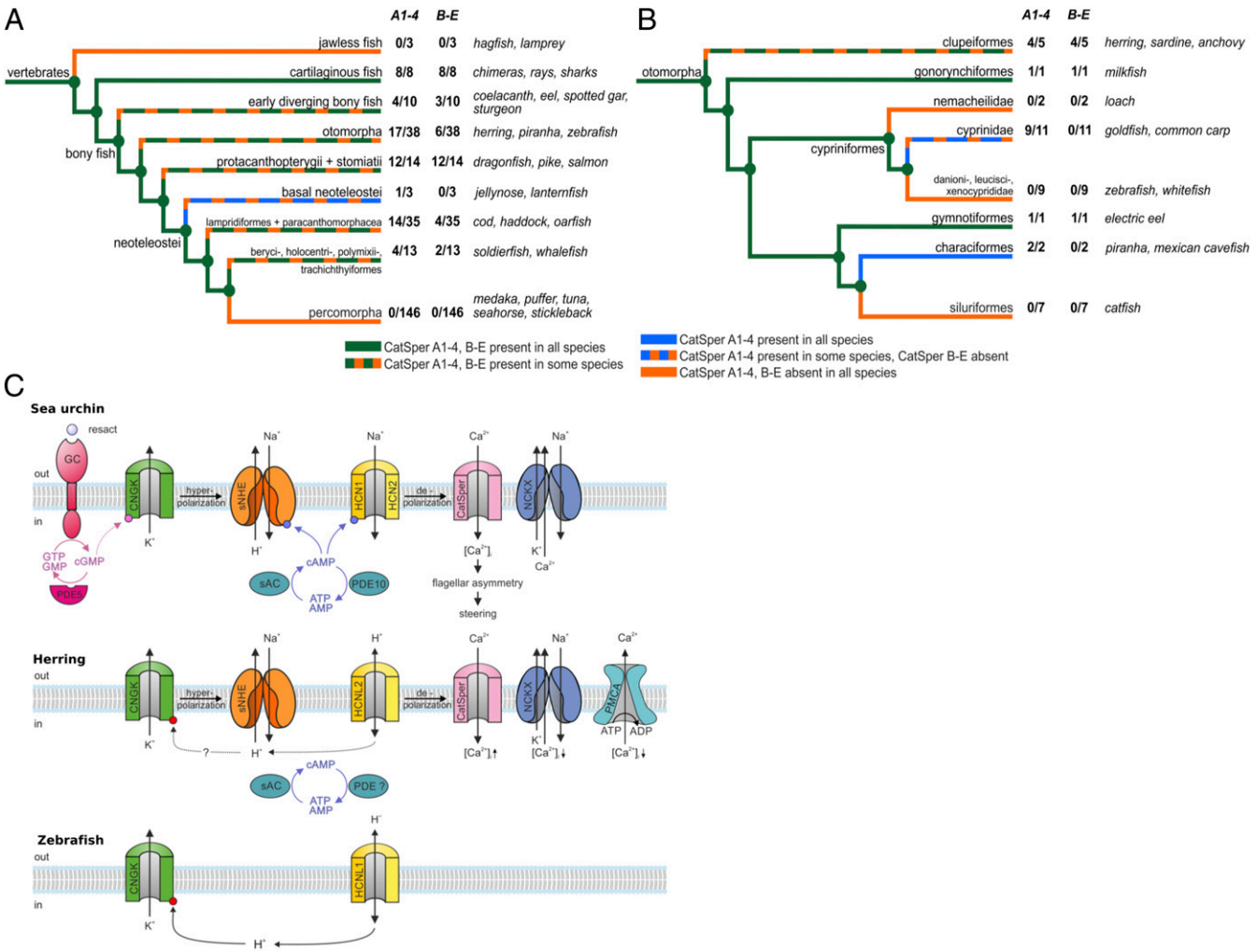


Fig. 4. Characterization of CatSper genes and the Ca^{2+} signaling pathway in sperm. (*A* and *B*) Multiple independent gene losses of CatSper1 to CatSper4 and CatSperB, CatSperG, CatSperD, and CatSperE in fish. Cladogram of major phylogenetic groups based on Betancur et al. (70). Colored branches indicate the presence of at least three of each the CatSper1 to CatSper4 and CatSperB, CatSperG, CatSperD, and CatSperE in all (green for both and blue for CatSper1 to CatSper4 only), none (orange), or some (green and orange for both, blue and orange for CatSper1 to CatSper4 only) species of the clade. Number of species with CatSper genes/all species investigated. Typical members of the respective clade are shown on the right. (*A*) Full tree including all fish species with searchable genomes. Clade groupings are simplified: early diverging bony fish comprise several groups—*Cladistia*, *Chondrostei*, coelacanths, eels, and *Holostei*. Note the near-complete retention of CatSper in cartilaginous fish and salmon-related fish (*Protacanthopterygii* and *Stomiatii*). (*B*) Presence and absence of CatSper in the *Otomorpha* clade. Note the almost complete retention of CatSper1 to CatSper4 in *Cyprinidae*, whereas all three sister clades have lost CatSper1 to CatSper4. Within *Clupeiformes*, four of five species have retained CatSper. (*C*) Comparison of signaling pathways in sperm of the sea urchin (*Arbacia punctulata*), Atlantic herring (*C. harengus*), and zebrafish (*D. rerio*). GC, chemoreceptor guanylate cyclase; CNGK, K^+-selective cyclic nucleotide-gated channel; sNHE, sperm specific Na^+/H^+ exchanger SLC9C1; HCN1 and HCN2, hyperpolarization activated and cyclic nucleotide-gated channels; CatSper, sperm-specific Ca^{2+} channel; NCKX, $Na^+/Ca^{2+}/K^+$ exchanger; PMCA, plasma membrane Ca^{2+}-ATPase; HCNL1, HCN-like 1 channel; HCNL2, HCN-like 2 channel.

The dN/dS ratio of 0.58 for *BMPR1BBY* is dramatically higher than the dN/dS = 0.01 to 0.02 for the autosomal paralogs (*SI Appendix*, Fig. S6). However, a dN/dS ratio of 0.58, well below 1.0, may be explained by relaxed purifying selection or a combination of positive selection at some sites and purifying selection at other sites. Our findings that the structural modeling predicts a conserved protein structure as well as its conserved kinase activity strongly suggest that the latter is the correct explanation. An important topic for future studies will be to further characterize the expression pattern of *BMPR1BBY* and in particular explore if it is expressed during early testis development at the sex determination stage. However, such experiments are challenging in a pelagic fish that does not reproduce in captivity.

The herring Y chromosome provides a unique opportunity to reconstruct the birth of a sex chromosome due to the relatively small size of the male-specific region that is not overloaded with repetitive DNA and other junk material like most other sex chromosomes. The evolutionary process must have started with the emergence of a truncated copy of *BMPR1BB* that was translocated to chromosome 8 (Fig. 5). The truncation assigned a new protein function to the gene. At the same time, because of the loss of its 5′ end including the promoter and intron 1, *BMPR1BBY* must have lost the expression regulation of the autosomal gene and most likely acquired a new spatial and temporal expression pattern required to trigger male gonad development. Such transcriptional rewiring of emerging sex-determining genes, in particular if derived from gene duplication of a downstream gonadal factor, has been postulated to be a hallmark of sex chromosome evolution (4, 44).

The insertion of the duplicated fragment from chromosome 21 into chromosome 8 created a region of nonhomology that cannot recombine in meiosis. Thus, a proto Y was formed and maintained

by suppressed recombination ab initio. It is likely that *CATSPER3* and *CATSPERG* were already colocalized on the ancestral chromosome 8 when *BMPR1BBY* arrived on the scene because they are located on the same chromosome in many teleosts (*SI Appendix*, Table S3). In goldfish, the two genes are <1 Mb apart. Possibly the next step was duplication of *CATSPER3* and its translocation near *BMPR1BBY* during or after the duplication event (Fig. 5). The synteny in that region is highly conserved across vertebrates, including *CATSPER3* and the homeobox gene *PITX1*, although CatSper genes have been lost in birds and many teleosts (*SI Appendix*, Fig. S9, and Fig. 4). Notably, this region is more rearranged in the herring than in other vertebrates (*SI Appendix*, Fig. S9), suggesting that chromosomal rearrangements contributed to the evolution of the Y-specific region. A third event was the incorporation of *CATSPERG* into the Y-specific region and its loss from a pseudoautosomal location. Finally, *CATSPER3A* underwent the fate of most duplicated genes, namely, degeneration, because it was redundant and *CATSPER3Y* benefited from its tighter linkage to the male sex-determining gene. By this process, a minimal sex-determining region (SDR) has evolved in the Atlantic herring; the region does not recombine with chromosome X and contains only three functional genes, a candidate male-determining factor and two genes essential for sperm function.

It is possible that the SDR in herring as present today is much younger than the emergence of the BMPR1BBY gene about 20 Mya (*SI Appendix*, Fig. S6), as indicated by the modest sequence divergence in the sex-associated region which rather suggests a divergence time of a few million years (Fig. 1*A*). However, it is possible that there is ongoing genetic exchange (recombination or gene conversion) between X and Y in this region. In fact, one of the Pacific herring chromosomes included in Fig. 1*B* may represent a recombinant haplotype because it appears X-like over the *BMPR1BBY-CATSPER3Y* region but Y-like over *CATSPERG*. It will be of considerable interest to further explore the evolution of this SDR region in a closely related clupeid like the sprat (*Sprattus sprattus*).

The association of *BMPR1BBY* with *CATSPER3Y* and *CATSPERG* in the herring SDR is consistent with one of the most characteristic features of sex chromosomes, namely, suppressed recombination between a sex-determining factor and genes beneficial for that sex (11–15). However, the Atlantic herring Y differs from most previously described sex chromosomes as this suppressed recombination has not spread along the chromosome. The region showing strong genetic differentiation between males and females is remarkably small and extends only over ~300 kb despite the existence of the herring Y for at least a few million years. As a comparison, the Y chromosome in the threespine stickleback, which is less than 26 My old, has a nonrecombining region of 17.5 Mb that shows extensive degeneration (45). Unlike the stickleback Y chromosome, the herring Y does not lack any gene present on the X chromosome; it only carries three additional genes that are likely essential for male development and function. Any degeneration that may have occurred is expected to be restricted to the genes located in the sex-associated region showing sequence divergence between X and Y (*SI Appendix*, Table S1).

Methods

Whole-Genome Sequence Analyses. We analyzed whole-genome sequence (WGS) data of 16 individuals with known sex, collected from Atlantic and Baltic with on average ~12X sequence coverage (19). The reads were aligned to two assemblies of the Atlantic herring genome (17, 19), and high confidence SNPs were called as described (17, 19). A genome-wide association analysis for the detection of sex-associated regions was done by comparing allele frequency between males and females with a χ^2 test using read counts extracted from sites called in the older assembly (19). The depth of coverage across the chromosome-based genome assembly (17) was extracted using GATK:DepthOfCoverage (version 3.5.0) (46). The data were converted to normalized depth 1-kb windows, and we compared the depth between males and females. No Animal Care and Use Committee approval was required because the herring were not harvested specifically for research purposes.

Assembly of the Chromosome Y Region. The assembly of the Y region is based on two unassigned scaffolds (18 and 229) and parts from an earlier version of our current PacBio-assembly of the herring genome (17), which is based on a male Atlantic herring, but only the ChrX region was included in the previously reported assembly (17). The Y assembly presented here was validated by investigating uniformity of depth of coverage of individual PacBio reads across the assembled sequence.

We used eight males and eight females to estimate the genetic differentiation between males and females by Weir–Cockerham F_{ST} (47) using vcftools v0.1.16 (48) in 50-kb overlapping windows across chromosome Y.

Annotation of the Male-Specific Region. We identified male-unique regions on unassigned scaffolds from the current PacBio-based herring assembly which was not annotated by the Ensembl annotation pipeline (https://www.ensembl.org) because the male-unique region was not included in this assembly (17). Thus, we annotated these regions by the annotation pipeline implemented by National Bioinformatics Infrastructure Sweden based on the Maker package (version 3.01.02) (49) (https://github.com/NBISweden/GAAS, version 2019-07-11). We first generated an evidence-built annotation based on protein similarity using Uniprot (50) (downloaded August 11, 2016). We also generated an ab initio built annotation by first training Augustus (version 2.7) (51) using 1,213 genes annotated by Ensembl (17). Then we combined the two annotations to generate the final annotation. We manually curated *BMPR1BBY* by comparing it with the autosomal copy on chromosome 21 and generated the final gene structure. In addition, we

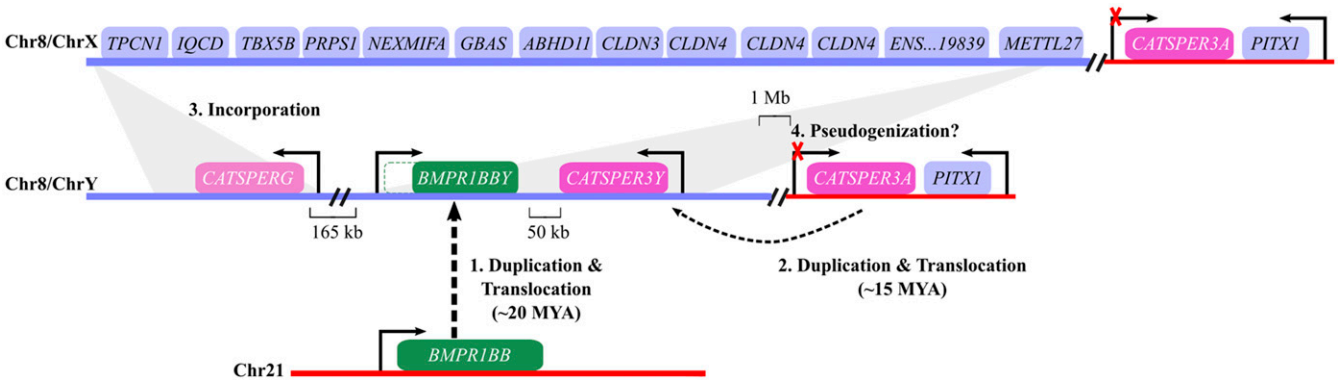


Fig. 5. Reconstruction of the evolution of sex chromosomes in Atlantic herring. The evolution of ChrY occurred in four steps, but the exact order of steps 2 to 4 is not yet known. 1) Duplication and translocation of *BMPR1BB* from Chr21 to Chr8. 2) Duplication and translocation of *CATSPER3* within Chr8/ChrY. 3) Incorporation of *CATSPERG* in ChrY and loss from ChrX. 4) *CATSPER3A* becomes pseudogenized or evolved a new function. MYA, million years ago.

used GenScan (52) to do another round of ab initio annotation which was merged with the Maker annotation.

To validate the structure of the gene model we conducted 3′ RACE and Nanopore whole-transcriptome and gene-specific cDNA sequencing. Total RNA was prepared from the testis of an adult Atlantic herring using the RNeasy Mini Kit (Qiagen). The 3′ RACE reaction was performed with 1 μg RNA using the FirstChoiceTM RLM-RACE Kit (Thermo Fisher Scientific). Nested RACE PCRs were performed in a 25-μL reaction containing KAPA HiFi Fidelity Buffer with $MgCl_2$, 0.32 mM dNTPs, 0.4 μM each of the forward and reverse primer, 0.8 U KAPA HiFi DNA Polymerase (Kapa Biosystems) and 1 μL of the cDNA or Outer RACE PCR product (1:5 dilution) as PCR template. Amplification was carried out following the program: 95 °C for 5 min, 35 cycles of 98 °C for 20 s, 58 °C for 30 s and 72 °C for 1 min, and a final extension of 10 min at 72 °C.

Targeted Resequencing Using Xdrop. To confirm the orientation of the male-specific region in relation to the flanking pseudoautosomal region we used the Xdrop method for targeted sequencing of long DNA molecules (20). Primer pairs for the amplification of fragments from the male-specific region in the near vicinity of the border to the pseudoautosomal region were designed. The following primer pairs were used to amplify a fragment in the vicinity of *CATSPER3Y*: for dPCR, Fw-TGCACTAGTGGGTGACTACCAA and R-TGTAAGACACCAGAATTTCCCCAC (product size 173 bp, Tm = 76.25 ± 0.08 °C), and for qPCR, Fw-TACTCAGATGGCCGACGAAATCTTTA and R-GCGTCAAAGACGTTATATCCGC (product size 111 bp, Tm = 77.88 ± 0.11 °C). For the amplicon from the scaffold18 region, the following primers were used: for dPCR, Fw-ACCGGCTTTATCCATCCGTC and R- AGTGACGAAAACGAAACC GTC (product length 165 bp, Tm = 76.97 ± 0.08 °C), and for qPCR, Fw:CTTCGGATCTGCACGATTCAC and R-TCAGAAGACCGGAGACTGGA (product length 127, Tm = 80.98 ± 0.11 °C). The reactions were run in a final volume of 10 μL, of which 5 μL were of dPCR mix, 0.5 μL of qPCR dye, 0.4 μL of 10 μM primer pair, and 0.1 μL of water. The enrichment and Nanopore sequencing of the genomic regions was performed by Samplix services (https://samplix.com/).

Genotyping Sex-Specific Markers. We genotyped 200 individuals for the presence/absence of *BMPR1BBY*. These samples had accurate sex records and were collected from two localities, Björköfjärden, Kattegat and Brofjorden, Skagerrak (19). The following primer pair was designed using primer3 with default settings, bby-fw CCAGACTGACTTGTGGTTCC and bby-R CTTCCCAGC AACAAGATGG, giving a 132-bp product with a melting temperature of 85.89 ± 0.07 °C (average and SD, respectively). The PCR reactions were composed of 5 μL SYBR Green PCR Master Mix (Applied Biosystems by Thermo Fisher Scientific), 1 μL of primer pair working solution at a concentration of 4 μM, and 4 μL of the DNA samples at a final concentration of 1 ng per reaction. Control reactions with no template were included at all points to spot any primer dimer formation. PCR was done using a QuantStudio 6 Flex instrument (Applied Biosystems by Thermo Fisher Scientific) with the thermal profile 2 min at 50 °C and 2 min at 95 °C, followed by 40 cycles of 15 s at 95 °C, 30 s at 57 °C, and 30 s at 72 °C. Melting curves were generated with incubation for 15 s at 95 °C, 1 min at 60 °C, and 15 s at 95 °C. The ramp rate between the two last steps was 0.05 °C/s. Primer sequences are also in *SI Appendix*, Table S4.

Nanopore Sequencing. Total RNA (500 ng) from testis was converted into cDNA following the Oxford Nanopore Technologies (ONT) protocol for rapid cDNA sequencing with the SuperScript IV reverse transcriptase (Invitrogen). To improve yield of full-length cDNA the incubation times were extended to 50 min at 50 °C and 20 min at 42 °C. The cDNA was amplified with LongAmp polymerase (NEB) and click primers (cPRM, ONT) for 14 cycles with 6 min elongation. Sequencing was done as partial runs on R9.4 and R9.5 flow cells as 1D reads. *BMPR1BBY* and *CATSPERG* were amplified from testis cDNA with LongAmp and internal primers and in combination with the ONT forward and reverse adaptor primers in order to capture the full transcript. PCR products were sequenced using the LSK309 $1d^2$ ligation sequencing protocol (ONT).

Genomic Alignments. Similarity between *CATSPER3Y* and *BMPR1BBY* and their presumed homologous autosomal loci was assessed by performing local alignments of the genomic nucleotide sequence, including flanking regions, of the two Y-specific gene copies against genomic sequence fragments, extracted from assembly version Ch_v2.0.2 (GCA_900700415.1) that covered the potentially orthologous autosomal genes. The targets were identified from the Ensembl-provided annotation of the herring genome (Database version 99.202). The alignments were performed using Clustal Omega version 1.2.3 (53) using default parameters and analyzed in R (54) using the package "ape" version 5.3 (55).

We compared the gene content in the sex-associated region on X and Y using Clustal Omega version 1.2.3 (53) implemented in "ape" R package version 5.3 (55) as described above.

Phylogenetic Analysis by Maximum Likelihood (PAML). First, we collected a reference set of *CATSPER3* and *BMPR1BB* genes from the ENA database. After manual inspection of the possible transcripts, we retained the option for each species that most closely resembled the herring Y copy. Second, in-frame nucleotide alignments spanning the extent of the Y-specific copies were generated for both genes. For *BMPR1BB*, which is highly conserved, the nucleotide alignments were already in frame, as verified by the "trans" function from "ape" and subsequent comparison with the reference sequences. For *CATSPER3*, which is more variable, PAL2NAL (56) was used to translate the amino acid alignment to a nucleotide equivalent using the reference cDNA sequences. In both cases, ambiguous positions were discarded during the rate analysis. Next, we extracted the taxa represented in each alignment from the Fish Tree of Life (57) to generate unrooted guide trees for the PAML analysis. These trees were edited to split the herring entry into a Y-specific and autosomal copy, and for *CATSPER3*, the coelacanth was added as an additional unresolved branch. Finally, the alignments and guide trees were fed into the "codeml" module of PAML version 4.9 (58) and run under three different models: 1) All branches evolve at the same rate (NULL model). 2) Only the "herring Y-" (for *BMPR1BB*) or the "herring A-" (for *CATSPER3*) branch has a unique rate (two-rates model). 3) Both the "herring Y" and "herring A" branches have a unique rate (three-rates model).

The models were then compared using a likelihood ratio test statistic, which was evaluated against a χ^2-distribution with degrees of freedom equal to the difference in estimated parameters between the compared models. The PAML control files, guide trees, and alignments used are found at https://github.com/nimarafati/Atlantic_herring_SDR.

We used *dS* values from the PAML analysis to estimate divergence times by extracting fossil-calibrated LCA estimates from the Fish Tree of Life (57) and dividing the time interval proportionally to the lengths of the terminal and internal branches.

Structure Modeling. Structural prediction of the intracellular domain of BMPR1BB, full-length L-BMPR1BBY, and full-length S-BMPR1BBY was performed using the RaptorX server (59), which gave the best models using Protein Data Bank 3MDY (human BMPR1B) (60) as the template. Structure alignments and protein representations were generated using the PyMOL Molecular Graphics System (Version 1.3, Schrödinger, LLC).

Preparation and Mass Spectrometric Analysis of Flagella from Herring Sperm. The preparation of flagella from *C. harengus* sperm was as described (39) using Herring Ringer's (HR) solution instead of artificial seawater. Briefly, sperm were washed two times with HR (in mM: 206 NaCl, 7.2 KCl, 2.1 $CaCl_2$, 3.1 $MgCl_2$, 10 Hepes/NaOH, pH 7.6) (61). Hepes was used for buffering instead of $NaHCO_3$. Sperm suspension was sheared 20 times by centrifugation for 30 s at 75 × *g* and 4 °C through a 40-μm mesh of a cell strainer (BD Biosciences). Flagella were washed in HR and stored as pellet.

Flagella protein (20 μg) was subjected to GeLCMS as described (62) with the following modifications: excision of 10 gel slices, use of Proteome Discoverer Rev. 2.2, *C. harengus* protein sequences from National Center for Biotechnology (NCBI) release GCF_900700415.1 (date April 16, 2019) supplemented with additional sequences (Dataset S4), recalibration of precursor MS spectra, precursor mass error tolerance 8 ppm, MS/MS mass error tolerance 0.6 Da, oxidation of Met, phosphorylation at Ser, Thr, Tyr, and N-terminal pyro-Glu as variable, Cys-Propionamide as static modification.

Bioinformatic Analysis of *CatSper* Genes. All fish genomes available from NCBI (https://www.ncbi.nlm.nih.gov/genome/) as of March 11, 2020, were analyzed for the presence of highly conserved β actin as positive control. The whole-genome shotgun databases (WGS) for these genomes were searched with TBlastN (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastn) for β actin, using the *Danio rerio* (zebrafish) sequence as query. Genomes which were unsearchable or yielded poor hits were excluded. The research resulted in 271 genomes, which were evaluated for the presence of *CatSper1* to *CatSper4* and *CatSperB*, *CatSperG*, *CatSperD*, and *CatSperE*. We then performed TBlastN searches in the WGS databases of fish genomes using previously annotated protein sequences for CatSper1 to CatSper4 and CatSperB, CatSperG, CatSperD, and CatSperE of *Lepisosteus occulatus* (spotted gar), *Latimeria chalumnae* (coelacanth), *Callorhinchus milii* (elephant shark), *Salmo salar* (Atlantic salmon), *Esox lucius* (northern pike), and *Carassius*

auratus (goldfish). The search was recursive, whenever more closely related sequences were identified. Candidates were evaluated according to e value until outgroup genes were found; a rigid cutoff by e value was ineffective. Candidate genes were predicted by genewise (https://www.ebi.ac.uk/Tools/psa/genewise/) (63) using closely related validated sequences as template. To validate candidates, we employed phylogenetic tree analysis. Trees were constructed from the candidate genes, the query sequences and other validated CatSper genes (e.g., *Mus musculus* [mouse], for full list, see Dataset S2). For CatSper1 to CatSper4, representative members from related ion channel families were used as outgroup (CACNA and SCN channels of mouse and zebrafish, TPCN and TRP channels of zebrafish). CatSperB, CatSperG, CatSperD, and CatSperE served as each other's respective outgroup. Sequences were aligned using the MAFFT (multiple alignment using fast Fourier transform) tool version 7 (https://mafft.cbrc.jp/alignment/server/) (64, 65) with the iterative refinement strategy called E-INS-I strategy and otherwise default parameters. Gap Strip Squeeze (https://hcv.lanl.gov/content/sequence/GAPSTREEZE/gap.html) was used to remove regions of the multiple sequence alignments which contained gaps in more than 90% of the sequences. Trees were subsequently calculated using the PhyML online tool, which is a phylogenetic tool that employs a maximum likelihood (ML) algorithm (http://www.atgc-montpellier.fr/phyml/) (66). Tree improvement was set to SPR (subtree pruning and regrafting), and branch support was set to χ^2-based aLRT (approximate likelihood ratio test). Otherwise the default settings were unchanged. Branch support for single CatSper1 to CatSper4 clades was always maximal; thus, CatSper1 to CatSper4 candidates that fell into these clades were considered validated. Branch support for outgroup clades (CACNA, SCN, and TRP) was maximal, and candidates which were sorted to the outgroup genes were discarded. For CatSperB, CatSperG, CatSperD, and CatSperE, branch support was always close to maximal; thus, candidates falling into these groups were considered validated.

RNA Preparation and Expression Analysis. Total RNA was extracted from the gonad, heart, spleen, kidney, gills, intestine, hypothalamus and saccus vasculosus (BSH), and brain without BSH (brain) of three adult male and three female Atlantic herring using the RNeasy Mini Kit (Qiagen). RNA was reverse transcribed into cDNA with a High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Real-Time PCR with SYBR Green chemistry was performed for *BMPR1BB*, *CATSPER3Y*, *CATSPERG*, *AMH*, and *AMHR2* in 10 μL reactions of SYBR Green PCR Master Mix (Thermo Fisher Scientific), 0.3 μM primers, and cDNA template, with a program of an initial denaturation of 10 min at 95 °C followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. TaqMan Gene Expression assay (Thermo Fisher Scientific) containing 0.30 μM primers and 0.25 μM *BMPR1BBY* TaqMan probe (Integrated DNA Technologies) was performed to compare the relative expression levels of *BMPR1BBY*. *ACTIN* was used as the housekeeping gene in all expression assays. All of the primers and probes are listed in *SI Appendix*, Table S4.

The expression of *CATSPER3A* in gonads and brain was investigated by RT-PCR. Briefly, cDNA was produced with the High Capacity cDNA Synthesis kit (Thermo Fisher Scientific), without RNase inhibitor. As this kit does not contain any oligo-dT primer, that was added to the reaction master mix to a final concentration of 5 mM [Oligo(dT)$_{18}$ primer; Thermo Fisher Scientific]. Primer pairs were designed with the NCBI tool (*SI Appendix*, Table S4). RT-PCRs were run with the AmpliTaq Gold DNA Polymerase with Buffer II and $MgCl_2$ (Thermo Fisher Scientific) together with dNTPs (dNTP Mix, 5 mM each) at a final concentration of 200 mM each. Primer pairs were added to a final concentration of 0.4 mM each, and the working concentration of DNA polymerase was 2 U/reaction. The total sample amount was 40 ng in a final volume of 10 μL. The reactions were incubated according to the following profile: 95 °C for 10 min, 35 cycles of 95 °C for 15 s, 57 °C for 30 s, 72 °C for 1 min, followed by 72 °C for 5 min. All reactions were checked by gel electrophoresis. No detectable amplification was observed for either primer pair.

Transfection Experiments. Full-length clones for *AMH*, *AMHR2*, and autosomal *BMPR1BB* as well as *L-BMPR1BBY* and *S-BMPR1BBY* clones were designed including a fish consensus Kozak sequence (67), synthesized and cloned (GenScript) into the expression vector pcDNA3.1 (-) (Thermo Fisher Scientific).

Medaka spermatogonial stem cells (Sg3) were cultured as previously described (68). For transfection, cells were grown to 80% confluency in six-well plates and subsequently transfected with 5 μg expression vectors using the FuGene (Roche) reagent as described by the manufacturer.

For monitoring differential activation of Smads upon selective AMH, AMHR2, BMPR1B, or S- or L-BMPR1BBY expression, Sg3 cells were seeded in six-well plates and cotransfected with a combination of four kinds of plasmids (Fig. 3*A*): 1) an expression plasmid which encodes a fusion protein (31, 69) (Smad1-AD-Gal4-DBD or Smad5-AD-Gal4-DBD; 300 ng per well) that will translocate to the nucleus upon phosphorylation by the different TGFßs and in return transactivate the UAS-4 promoter through its Gal-4 DNA binding domain; 2) a reporter plasmid, which codes for luciferase under the control of a minimal promoter, which contains UAS sequences (UAS-luc, *Firefly* luciferase; 300 ng per well); 3) plasmids coding for the different ligands and receptors to be tested for signaling activity [either pcDNA3.1(-)-AMH, pcDNA3.1(-)-AMHR2, pcDNA3.1(-)-BMPR1BB, pcDNA3.1(-)-S-BMPR1BY, pcDNA3.1(-)-L-BMPR1B or a control plasmid (empty pcDNA3.1(-)); 400 ng per well]; and 4) a *Gaussia* luciferase expression plasmid for normalization (pCMV-GLuc; 5 ng per well). After 24 h, cells were washed with 2× PBS solution and lysed with 75 μL of passive lysis buffer (Dual Luciferase Reporter Kit Assay; Promega) and then subjected to luciferase assay. *Firefly* luciferase activity (UAS-luc reporter constructs) was quantified using the Dual Luciferase Reporter Assay System (Promega) and normalized against cotransfected *Gaussia* luciferase expressing plasmid. Datasets are the results of at least four independent cell transfections and luciferase measurements. Statistical significance was assessed by means of the Wilcoxon–Mann–Whitney test ($n = 4$ or 8). Raw data are in Dataset S5.

Data Availability. The reference genome used in this study is deposited in European Nucleotide Archive (ENA) (PRJEB31270) and the sequence of the SDR together with its annotation as well as other sequences and codes used in this study are available on Github: https://github.com/nimarafati/Atlantic_herring_SDR. Variant calls used in the χ^2-analysis were from Ref. 19 deposited in Dryad (5r774). Also, Samplix and cDNA data generated for this study are deposited in ENA (PRJEB38031)

ACKNOWLEDGMENTS. We thank Carl-Henrik Heldin and Aristidis Moustakas for valuable comments and Chunheng Mo for technical assistance. The study was supported by the Knut and Alice Wallenberg Foundation (to L.A.), Vetenskapsrådet (to L.A.), Research Council of Norway project 254774, GENSINC (to A.F., F.B. and L.A.), and AquaCRISPR (ANR-16-COFA-0004-01) (to A.H.). The National Genomics Infrastructure (NGI)/Uppsala Genome Center and UPPMAX (Uppsala Multidisciplinary Center for Advanced Computational Science) provided service in massive parallel sequencing and computational infrastructure. Work performed at NGI/Uppsala Genome Center has been funded by RFI/VR (Council for Research Infrastructure/Vetenskapsrådet) and Science for Life Laboratory, Sweden.

1. D. Bachtrog *et al.*; Tree of Sex Consortium, Sex determination: Why so many ways of doing it? *PLoS Biol.* **12**, e1001899 (2014).
2. D. Charlesworth, J. E. Mank, The birds and the bees and the flowers and the trees: Lessons from genetic mapping of sex determination in plants and animals. *Genetics* **186**, 9–31 (2010).
3. G. S. van Doorn, Evolutionary transitions between sex-determining mechanisms: A review of theory. *Sex Dev.* **8**, 7–19 (2014).
4. A. Herpin, M. Schartl, Plasticity of gene-regulatory networks controlling sex determination: Of masters, slaves, usual suspects, newcomers, and usurpators. *EMBO Rep.* **16**, 1260–1274 (2015).
5. M. Schartl, D. Galiana-Arnoux, C. Schultheis, A. Böhne, J.-N. Volff, "A primer of sex determination in poeciliids" in *Ecology and Evolution of Poeciliid Fishes*, A. Pilastro, J. Evans, I. Schlupp, Eds. (The University of Chicago Press, 2010), pp. 264–275.
6. M. Matsuda, M. Sakaizumi, Evolution of the sex-determining gene in the teleostean genus Oryzias. *Gen. Comp. Endocrinol.* **239**, 80–88 (2016).
7. W. J. Gammerdinger, T. D. Kocher, Unusual diversity of sex chromosomes in African cichlid fishes. *Genes (Basel)* **9**, 480 (2018).
8. M. W. Pennell, J. E. Mank, C. L. Peichel, Transitions in sex determination and sex chromosomes across vertebrate species. *Mol. Ecol.* **27**, 3950–3963 (2018).
9. R. H. Devlin, Y. Nagahama, Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364 (2002).
10. M. Schartl, A. Herpin, "Sex determination in vertebrates" in *Encyclopedia of Reproduction*, M. K. Skinner, Ed., (Elsevier, 2018), pp. 159–167.
11. R. A. Fisher, The evolution of dominance. *Biol. Rev. Camb. Philos. Soc.* **6**, 345–368 (1931).
12. J. J. Bull, *Evolution of Sex Determining Mechanisms*, (Benjamin/Cummings, Menlo Park, CA, 1983).
13. W. R. Rice, The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex-chromosomes. *Evolution* **41**, 911–914 (1987).
14. D. Charlesworth, B. Charlesworth, G. Marais, Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118–128 (2005).
15. M. Kirkpatrick, The evolution of genome structure by natural and sexual selection. *J. Hered.* **108**, 3–11 (2017).

16. H. Ida, N. Oka, K. Hayashigaki, Karyotypes and cellular DNA contents of three species of the subfamily Clupeinae. *J. Ichthyol.* **38**, 289–294 (1991).
17. M. E. Pettersson *et al*., A chromosome-level assembly of the Atlantic herring genome-detection of a supergene and other signals of selection. *Genome Res.* **29**, 1919–1928 (2019).
18. S. Lamichhaney *et al*., Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3452–E3461 (2017).
19. A. Martinez Barrio *et al*., The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *eLife* **5**, e12081 (2016).
20. E. B. Madsen, T. Kvist, I. Höijer, A. Ameur, M. J. Mikkelsen, Xdrop: Targeted sequencing of long DNA molecules from low input samples using droplet sorting *bioRxiv*: 10.1101/409086 (6 September 2018).
21. D. Ren *et al*., A sperm ion channel required for sperm motility and male fertility. *Nature* **413**, 603–609 (2001).
22. R. Seifert *et al*., The CatSper channel controls chemosensation in sea urchin sperm. *EMBO J.* **34**, 379–392 (2015).
23. C. Brenker *et al*., Action of steroids and plant triterpenoids on CatSper Ca^{2+} channels in human sperm. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E344–E346 (2018).
24. J. F. Smith *et al*., Disruption of the principal, progesterone-activated sperm Ca^{2+} channel in a CatSper2-deficient infertile patient. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6823–6828 (2013).
25. M. S. Hildebrand *et al*., Genetic male infertility and mutation of CATSPER ion channels. *Eur. J. Hum. Genet.* **18**, 1178–1184 (2010).
26. Q. Pan *et al*., Identification of the master sex determining gene in Northern pike (*Esox lucius*) reveals restricted sex chromosome differentiation. *PLoS Genet.* **15**, e1008013 (2019).
27. A. Herpin, C. Lelong, P. Favrel, Transforming growth factor-β-related proteins: An ancestral and widespread superfamily of cytokines in metazoans. *Dev. Comp. Immunol.* **28**, 461–485 (2004).
28. S. Mawaribuchi, S. Yoshimoto, S. Ohashi, N. Takamatsu, M. Ito, Molecular evolution of vertebrate sex-determining genes. *Chromosome Res.* **20**, 139–151 (2012).
29. M. C. Adolfi, R. T. Nakajima, R. H. Nóbrega, M. Schartl, Intersex, hermaphroditism, and gonadal plasticity in vertebrates: Evolution of the Müllerian duct and Amh/Amhr2 signaling. *Annu. Rev. Anim. Biosci.* **7**, 149–172 (2019).
30. J. A. Visser, AMH signaling: From receptor to target gene. *Mol. Cell. Endocrinol.* **211**, 65–73 (2003).
31. L. Sèdes *et al*., Anti-Müllerian hormone recruits BMPR-IA in immature granulosa cells. *PLoS One* **8**, e81551 (2013).
32. X. Cai, D. E. Clapham, Evolutionary genomics reveals lineage-specific gene loss and rapid evolution of a sperm-specific ion channel complex: CatSpers and CatSperbeta. *PLoS One* **3**, e3569 (2008).
33. J.-J. Chung *et al*., CatSperζ regulates the structural continuity of sperm Ca^{2+} signaling domains and is required for normal fertility. *eLife* **6**, e23082 (2017).
34. D. Wachten, J. F. Jikeli, U. B. Kaupp, Sperm sensory signaling. *Cold Spring Harb. Perspect. Biol.* **9**, a028225 (2017).
35. U. B. Kaupp, T. Strünker, Signaling in sperm: More different than similar. *Trends Cell Biol.* **27**, 101–109 (2017).
36. P. V. Lishko *et al*., The control of male fertility by spermatozoan ion channels. *Annu. Rev. Physiol.* **74**, 453–475 (2012).
37. W. Bönigk *et al*., An atypical CNG channel activated by a single cGMP molecule controls sperm chemotaxis. *Sci. Signal.* **2**, ra68 (2009).
38. T. Strünker *et al*., A K^+-selective cGMP-gated ion channel controls chemosensation of sperm. *Nat. Cell Biol.* **8**, 1149–1154 (2006).
39. C. Trötschel *et al*., Absolute proteomic quantification reveals design principles of sperm flagellar chemosensation. *EMBO J.* **39**, e102723 (2020).
40. R. Gauss, R. Seifert, U. B. Kaupp, Molecular identification of a hyperpolarization-activated channel in sea urchin sperm. *Nature* **393**, 583–587 (1998).
41. F. Windler *et al*., The solute carrier SLC9C1 is a Na^+/H^+-exchanger gated by an S4-type voltage-sensor and cyclic-nucleotide binding. *Nat. Commun.* **9**, 2809 (2018).
42. L. Wobig *et al*., A family of hyperpolarization-activated channels selective for protons. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 13783–13791 (2020).
43. S. Fechner *et al*., A $K^{(+)}$-selective CNG channel orchestrates $Ca^{(2+)}$ signalling in zebrafish sperm. *eLife* **4**, e07624 (2015).
44. A. Herpin *et al*., Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. *PLoS Genet.* **6**, e1000844 (2010).
45. C. L. Peichel *et al*., Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biol.* **21**, 177 (2020).
46. A. McKenna *et al*., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
47. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
48. P. Danecek *et al*.; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
49. B. L. Cantarel *et al*., MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
50. M. Magrane; UniProt Consortium, UniProt knowledgebase: A hub of integrated protein data. *Database (Oxford)* **2011**, bar009 (2011).
51. M. Stanke, M. Diekhans, R. Baertsch, D. Haussler, Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637–644 (2008).
52. C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
53. F. Sievers *et al*., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
54. Team RC, *R: A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, Vienna, Austria, 2019).
55. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
56. M. Suyama, D. Torrents, P. Bork, PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–W612 (2006).
57. D. L. Rabosky *et al*., An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).
58. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
59. M. Källberg *et al*., Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–1522 (2012).
60. A. Chaikuad *et al*, Crystal structure of the cytoplasmic domain of the bone morphogenetic protein receptor type-1B (BMPR1B) in complex with FKBP12 and LDN-193189. Protein Data Bank. https://www.rcsb.org/structure/3MDY (accessed 3 February 2020).
61. C. A. Vines *et al*., Motility initiation in herring sperm is regulated by reverse sodium-calcium exchange. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2026–2031 (2002).
62. D. Raju *et al*., Accumulation of glucosylceramide in the absence of the beta-glucosidase GBA2 alters cytoskeletal dynamics. *PLoS Genet.* **11**, e1005063 (2015).
63. E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
64. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
65. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
66. S. Guindon *et al*., New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
67. S. J. Grzegorski, E. F. Chiari, A. Robbins, P. E. Kish, A. Kahana, Natural variability of Kozak sequences correlates with function in a zebrafish model. *PLoS One* **9**, e108475 (2014).
68. Y. Hong *et al*., Establishment of a normal medakafish spermatogonial cell line capable of sperm production *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8011–8016 (2004).
69. F. Liu *et al*., A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**, 620–623 (1996).
70. R. Betancur-R *et al*., *The Tree of Life and a New Classification of Bony Fishes*, (PLOS Currents Tree of Life, ed. 1, 2013).