Sequencing, annotation, and pangenomics in wheat

Frédéric Choulet

To cite this version:

Frédéric Choulet. Sequencing, annotation, and pangenomics in wheat. Webinaire de l'AFBV, Association Française des Biotechnologies Végétales, May 2022, Clermont-Ferrand, France. hal-04090082

AFBV webinar, May 10th 2022

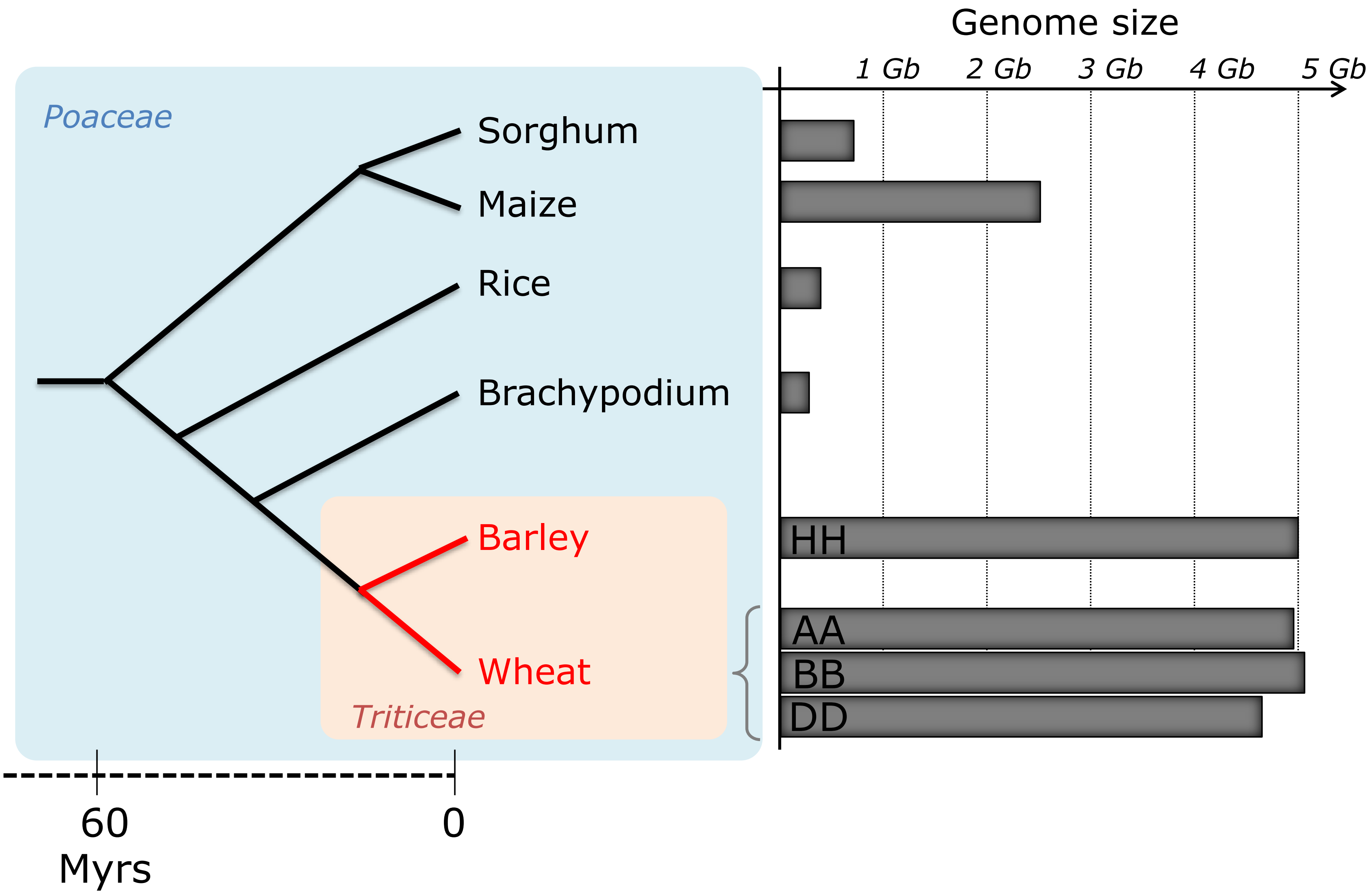
Sequencing, annotation, and pangenomics in wheat

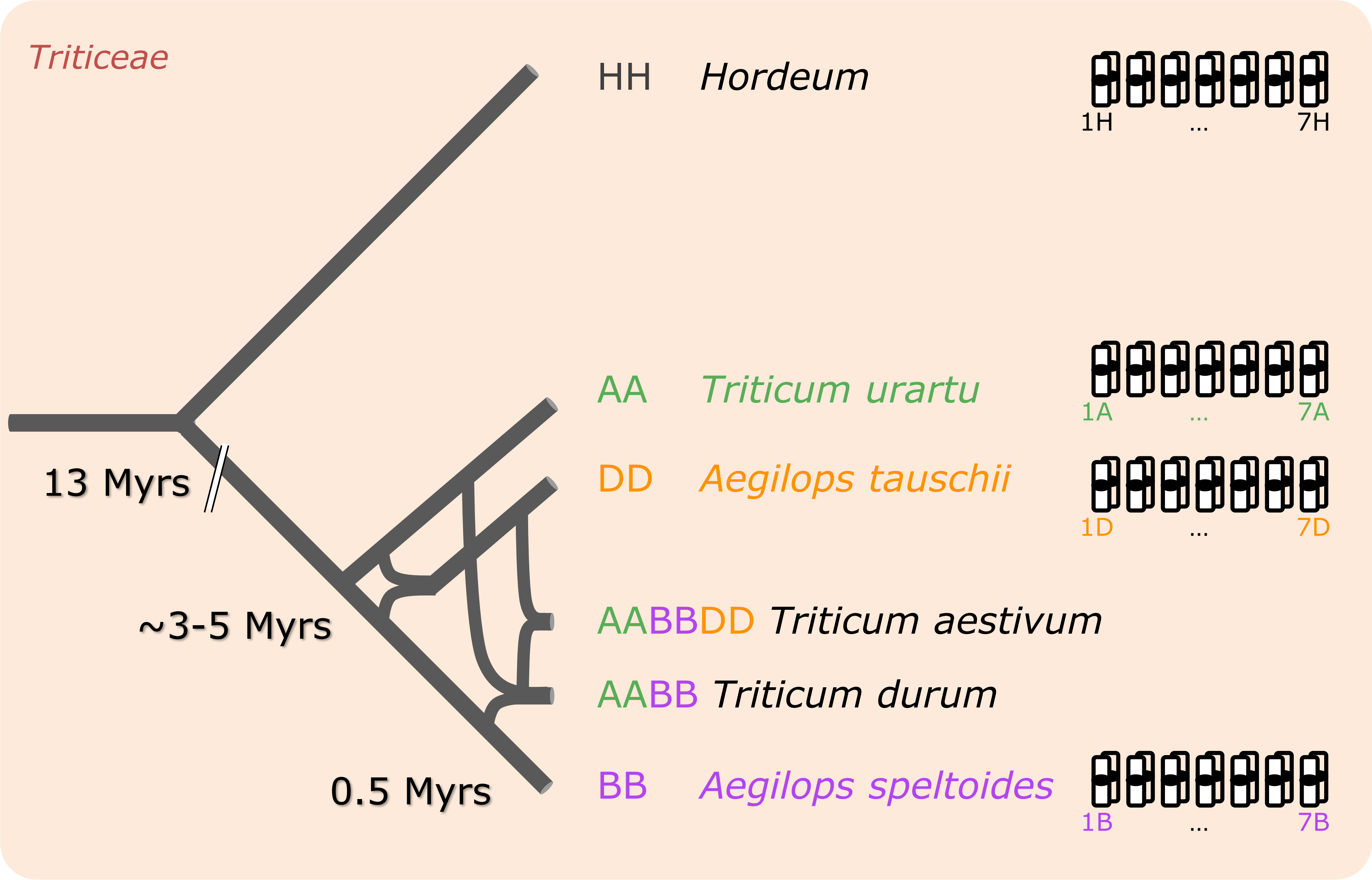
Frédéric Choulet

GDEC, Clermont-Ferrand, France

1. Introduction

- Wheat/*Triticeae*
- Pangenomics


Genome size
1 Gb
2 Gb
3 Gb
4 Gb
5 Gb
Poaceae
Sorghum
Maize
Rice
Brachypodium
Barley
Wheat
Triticeae
HH
AA
BB
DD
60
Myrs
0



Triticeae

HH *Hordeum*

1H … 7H

AA *Triticum urartu*

1A … 7A

DD *Aegilops tauschii*

1D … 7D

AABBDD *Triticum aestivum*

AABB *Triticum durum*

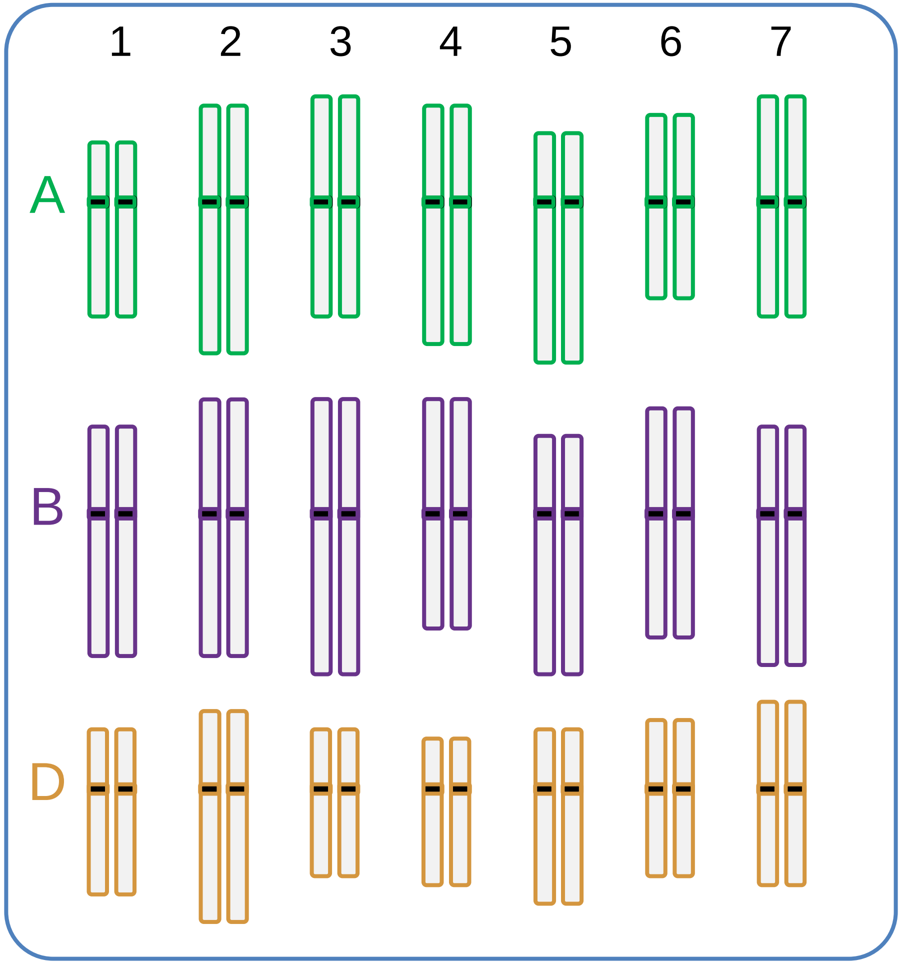
BB *Aegilops speltoides*

1B … 7B

13 Myrs

~3-5 Myrs

0.5 Myrs



Complex genome

- 15 Gb
- Hexaploid AA-BB-DD
- 85% TEs

- Ref cultivar: Chinese Spring

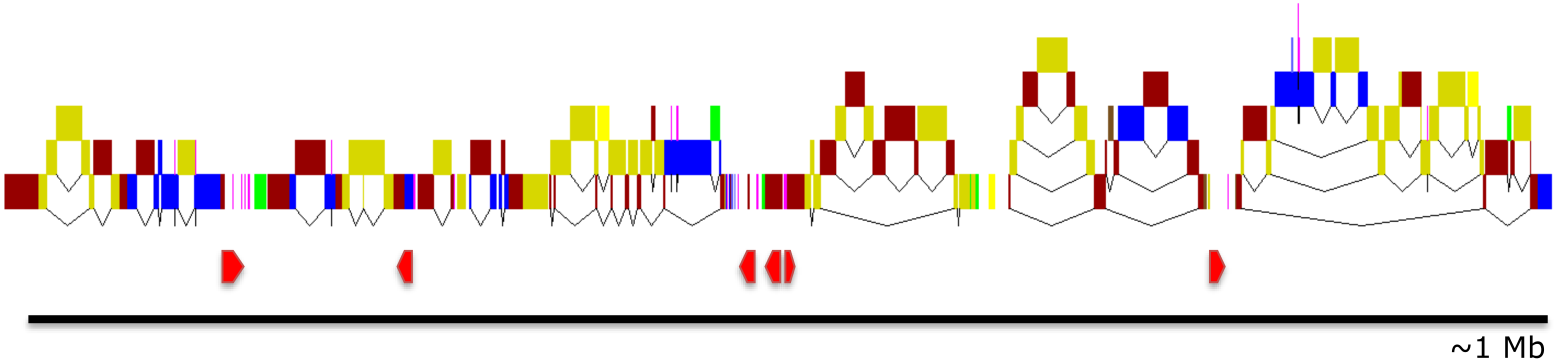
Assembly (v1)

- 21 chromosomes (14.5 Gb)

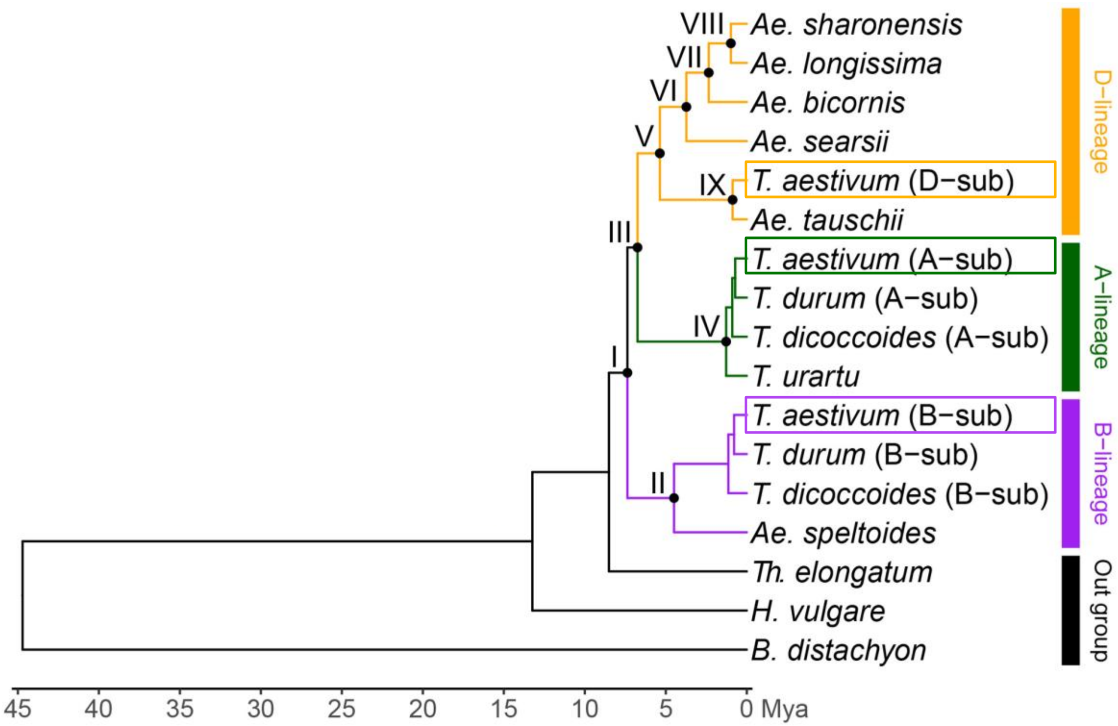
Annotation (v1.1)

- 107k genes
- 4 million TEs

IWGSC 2018


~1 Mb

- RQAs: **Reference-Quality Assemblies** of *Triticeae* genomes

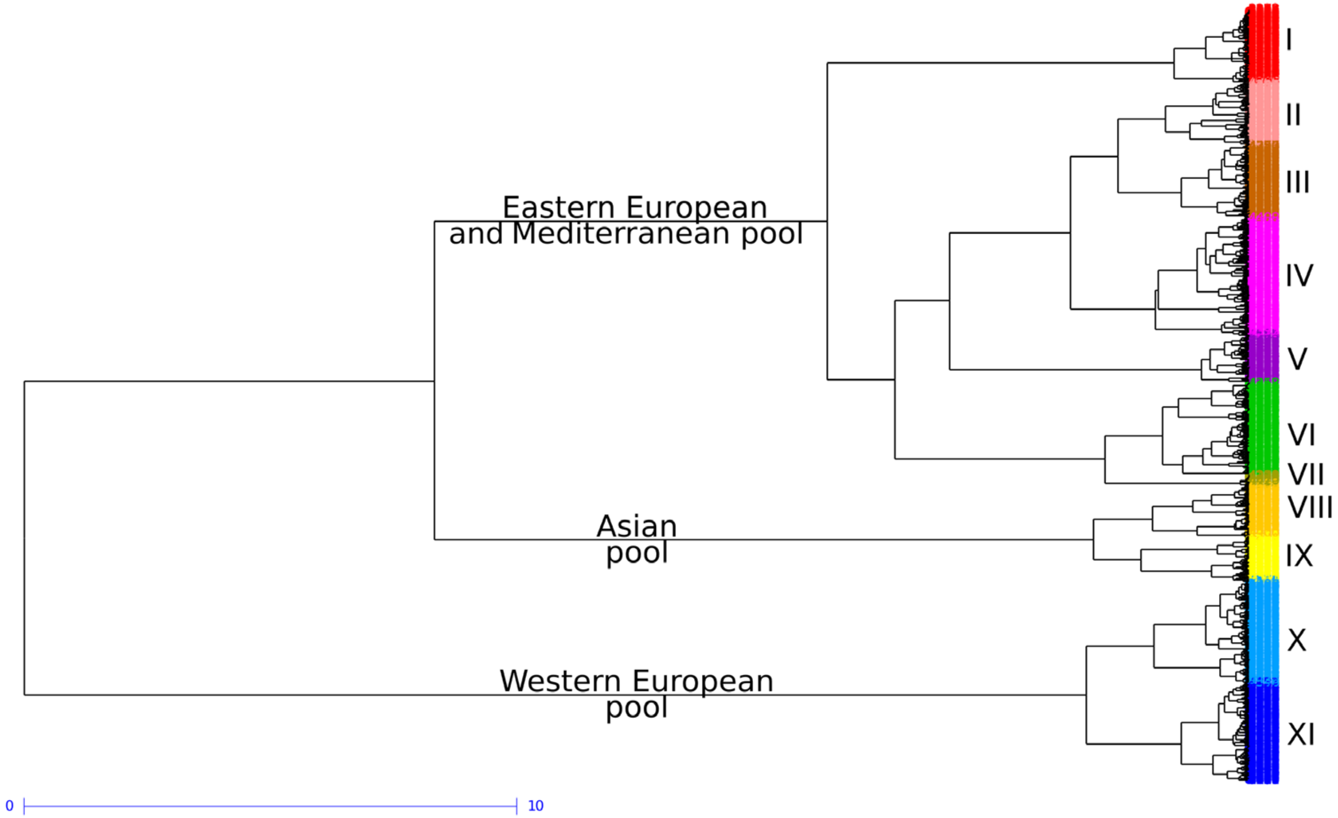


Li et al., Mol Plant 2022

- *T. aestivum:*

Chinese Spring + **17** accessions

- **SNP** genotyping of **4500** accessions representing world-wide diversity

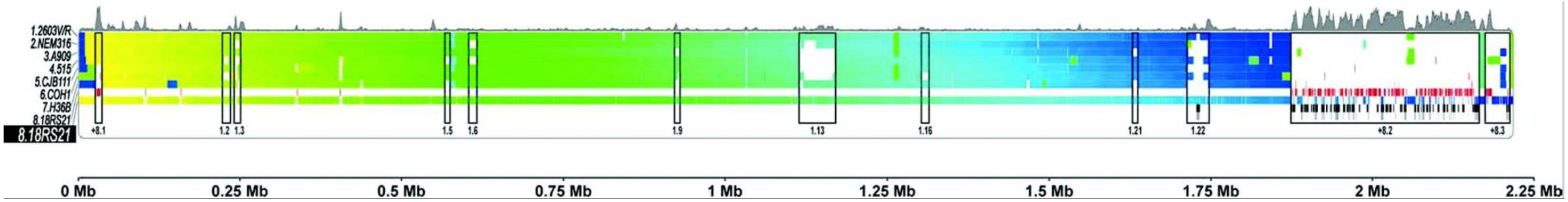


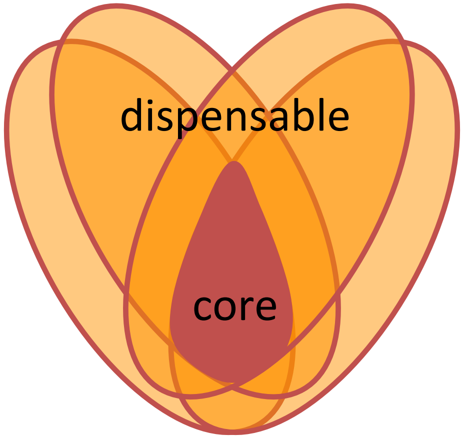
BreedWheat; Balfourier et al. 2018

○ Pangenomics

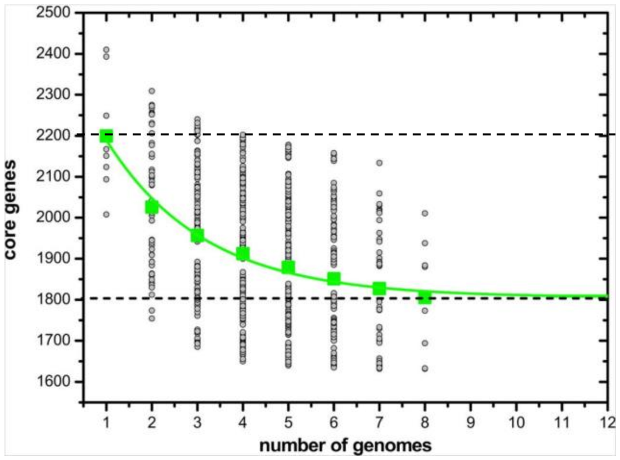
- **Tettelin et al. PNAS 2005**

8 genomes of *S. agalactiae*



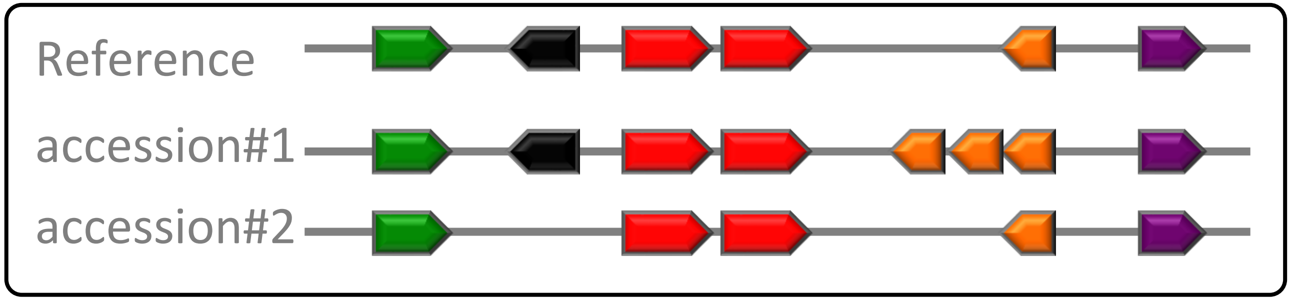


=pangenome



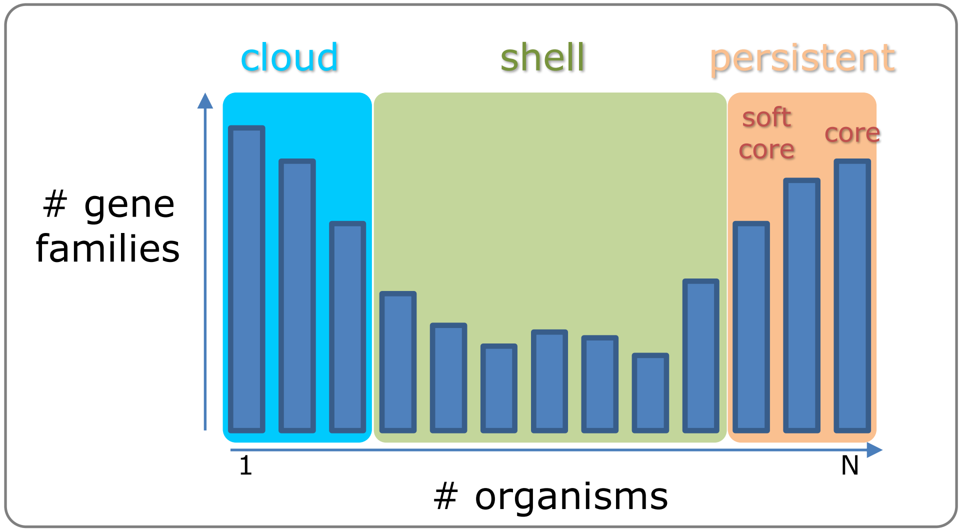
~80% (core)

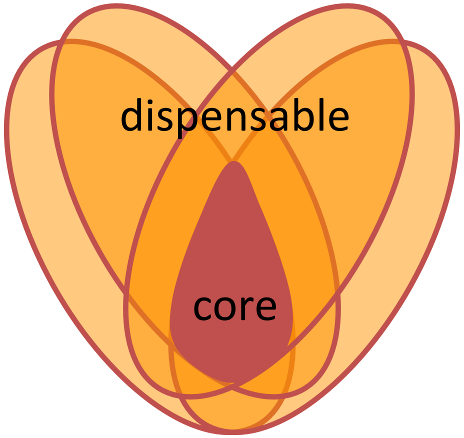
- Structural Variations (SVs)



PAV CNV

- **Koonin and Wolf NAR 2008**
bacteria, archea (~700 genomes)





=pangenome

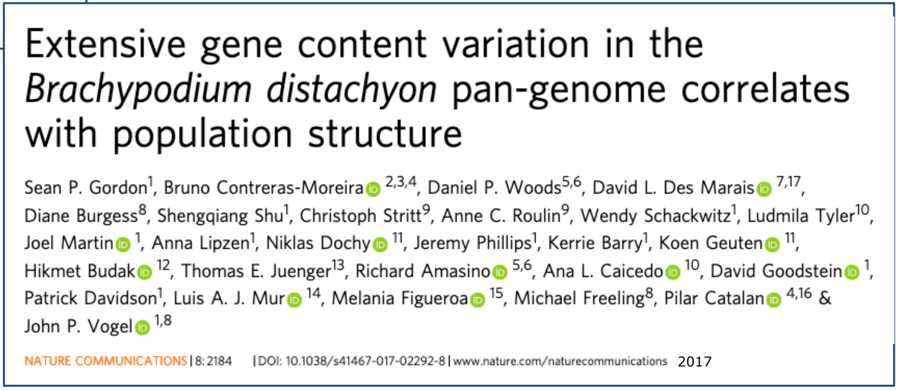
- Pangenomics in plants

De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits

Ying-hui Li[1–3,11], Guangyu Zhou[4,10,11], Jianxin Ma[5,11], Wenkai Jiang[4,11], Long-guo Jin[1–3], Zhouhao Zhang[4], Yong Guo[1–3], Jinbo Zhang[4], Yi Sui[1–3], Liangtao Zheng[4], Shan-shan Zhang[1–3], Qiyang Zuo[4], Xue-hui Shi[1–3], Yan-fei Li[1–3], Wan-ke Zhang[6], Yiyao Hu[4], Guanyi Kong[4], Hui-long Hong[1–3], Bing Tan[1–3], Jian Song[1–3], Zhang-xiong Liu[1–3], Yaoshen Wang[4], Hang Ruan[4], Carol K L Yeung[4], Jian Liu[4], Hailong Wang[4], Li-juan Zhang[1–3], Rong-xia Guan[1–3], Ke-jing Wang[1–3], Wen-bin Li[7], Shou-yi Chen[6], Ru-zhen Chang[1–3], Zhi Jiang[4], Scott A Jackson[8], Ruiqiang Li[4,9] & Li-juan Qiu[1–3]

NATURE BIOTECHNOLOGY VOLUME 32 NUMBER 10 OCTOBER 2014

7 genomes of *Glycine soja*
dispensable: **~20%** genes

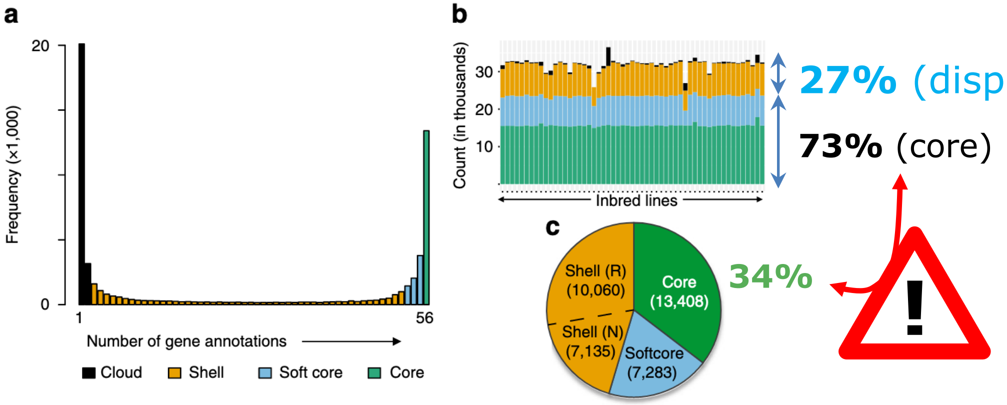


Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population structure

Sean P. Gordon[1], Bruno Contreras-Moreira[2,3,4], Daniel P. Woods[5,6], David L. Des Marais[7,17], Diane Burgess[8], Shengqiang Shu[1], Christoph Stritt[9], Anne C. Roulin[9], Wendy Schackwitz[1], Ludmila Tyler[10], Joel Martin[1], Anna Lipzen[1], Niklas Dochy[11], Jeremy Phillips[1], Kerrie Barry[1], Koen Geuten[11], Hikmet Budak[12], Thomas E. Juenger[13], Richard Amasino[5,6], Ana L. Caicedo[10], David Goodstein[1], Patrick Davidson[1], Luis A. J. Mur[14], Melania Figueroa[15], Michael Freeling[8], Pilar Catalan[4,16] & John P. Vogel[1,8]

NATURE COMMUNICATIONS | 8: 2184 | DOI: 10.1038/s41467-017-02292-8 | www.nature.com/naturecommunications 2017

54 genomes of *Brachypodium distachyon*



27% (dispensable)

73% (core)

34%

○ Pangenomics in plants

The pangenome of an agronomically important crop plant *Brassica oleracea*

Agnieszka A. Golicz[1], Philipp E. Bayer[2], Guy C. Barker[3], Patrick P. Edger[4], HyeRan Kim[5], Paula A. Martinez[1], Chon Kit Kenneth Chan[2], Anita Severn-Ellis[2], W. Richard McCombie[6], Isobel A.P. Parkin[7], Andrew H. Paterson[8], J. Chris Pires[9], Andrew G. Sharpe[10], Haibao Tang[11], Graham R. Teakle[3], Christopher D. Town[12], Jacqueline Batley[2] & David Edwards[2]

NATURE COMMUNICATIONS | 7:13390 | DOI: 10.1038/ncomms13390 | www.nature.com/naturecommunications 2016

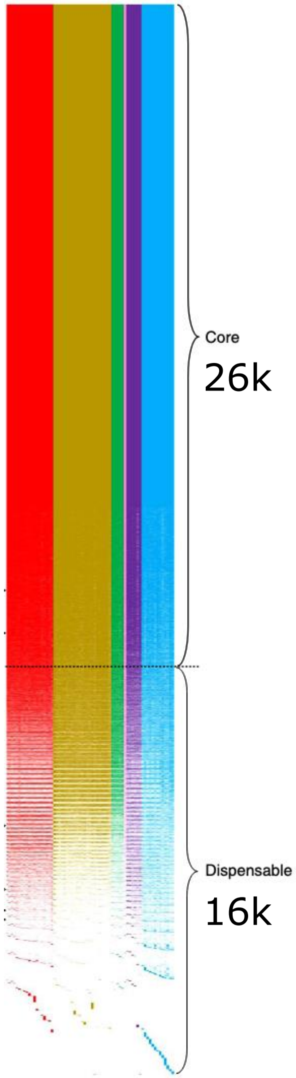
9 genomes of *Brassica oleracea*
dispensable: **~20%** genes

Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice

Qiang Zhao[1], Qi Feng[1], Hengyun Lu[1], Yan Li[1], Ahong Wang[1], Qilin Tian[1], Qilin Zhan[1], Yiqi Lu[1], Lei Zhang[1], Tao Huang[1], Yongchun Wang[1], Danlin Fan[1], Yan Zhao[1], Ziqun Wang[1], Congcong Zhou[1], Jiaying Chen[1], Chuanrang Zhu[1], Wenjun Li[1], Qijun Weng[1], Qun Xu[2], Zi-Xuan Wang[1], Xinghua Wei[2], Bin Han[1] and Xuehui Huang[1,3]*

NATURE GENETICS | VOL 50 | FEBRUARY 2018 | 278–284 | www.nature.com/naturegenetics

67 genomes of *Oryza sativa + rufipogon*
~20% genes



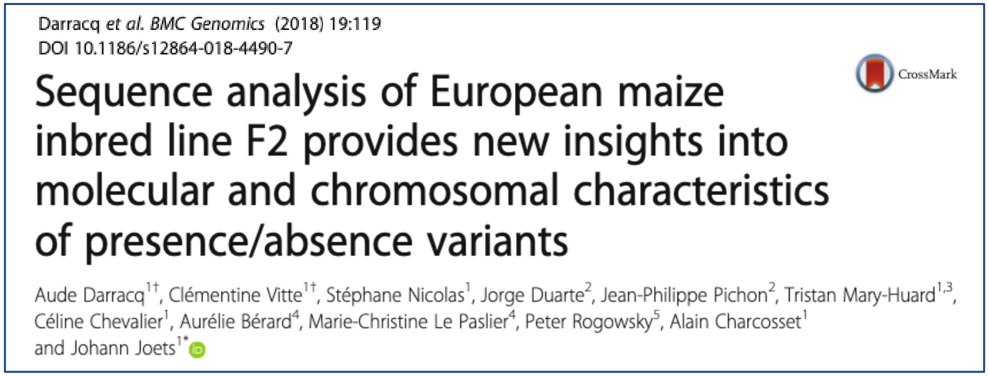
○ Pangenomics in plants

The Plant Cell, Vol. 26: 121–135, January 2014, www.plantcell.org © 2014.

Insights into the Maize Pan-Genome and Pan-Transcriptome[W][OPEN]

Candice N. Hirsch,[a,b,1] Jillian M. Foerster,[c,2] James M. Johnson,[c,3] Rajandeep S. Sekhon,[c,d] German Muttoni,[c,4] Brieanne Vaillancourt,[a,b] Francisco Peñagaricano,[e] Erika Lindquist,[f] Mary Ann Pedraza,[f] Kerrie Barry,[f] Natalia de Leon,[c,d] Shawn M. Kaeppler,[c,d] and C. Robin Buell[a,b,5]

503 maize accessions (RNASeq)
=> ~8700 novel transcripts



Darracq *et al. BMC Genomics* (2018) 19:119
DOI 10.1186/s12864-018-4490-7

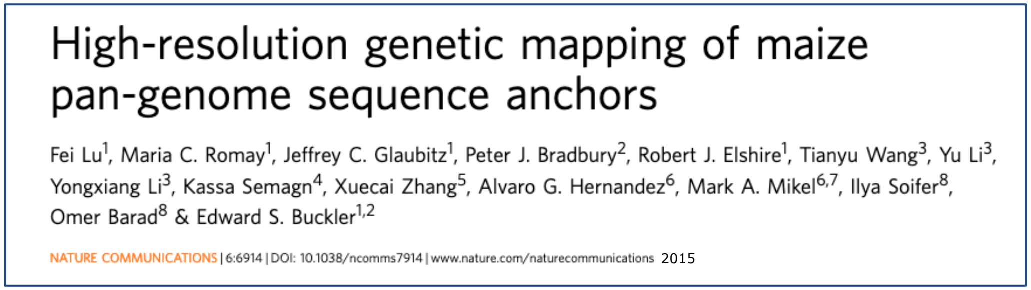
CrossMark

Sequence analysis of European maize inbred line F2 provides new insights into molecular and chromosomal characteristics of presence/absence variants

Aude Darracq[1†], Clémentine Vitte[1†], Stéphane Nicolas[1], Jorge Duarte[2], Jean-Philippe Pichon[2], Tristan Mary-Huard[1,3], Céline Chevalier[1], Aurélie Bérard[4], Marie-Christine Le Paslier[4], Peter Rogowsky[5], Alain Charcosset[1] and Johann Joets[1*]

Dispensable *vs* core genes:
- shorter, fewer introns
- less expressed, fewer conditions
- evolve more rapidly
- enriched in functions related to adaptation, response to env

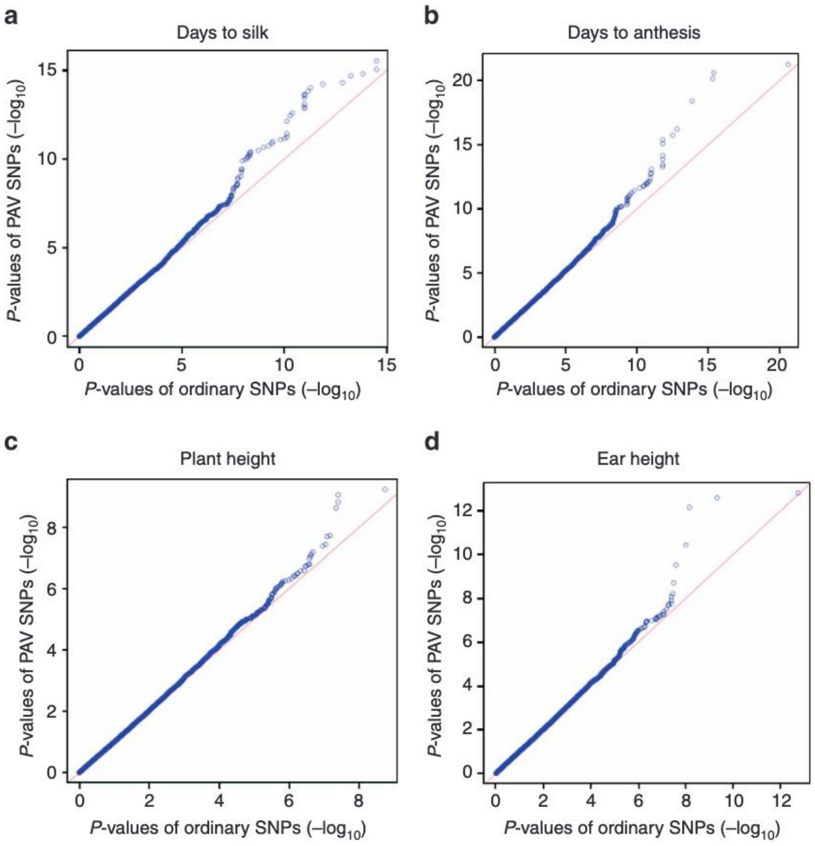
- Pangenomics in plants



High-resolution genetic mapping of maize pan-genome sequence anchors

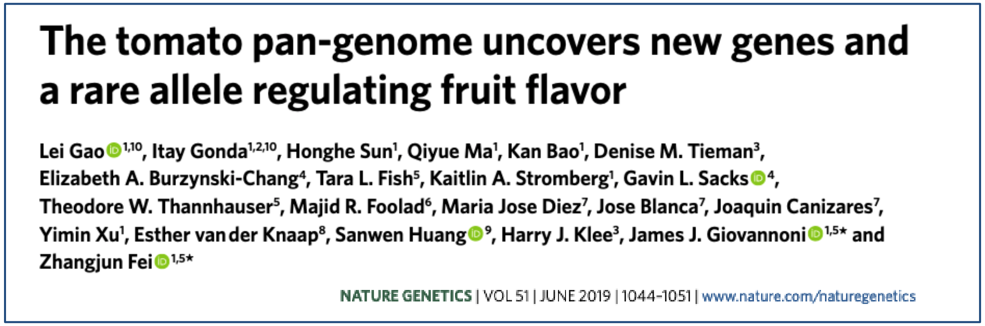
Fei Lu[1], Maria C. Romay[1], Jeffrey C. Glaubitz[1], Peter J. Bradbury[2], Robert J. Elshire[1], Tianyu Wang[3], Yu Li[3], Yongxiang Li[3], Kassa Semagn[4], Xuecai Zhang[5], Alvaro G. Hernandez[6], Mark A. Mikel[6,7], Ilya Soifer[8], Omer Barad[8] & Edward S. Buckler[1,2]

NATURE COMMUNICATIONS | 6:6914 | DOI: 10.1038/ncomms7914 | www.nature.com/naturecommunications 2015



=> "PAV SNPs exhibit enriched associations with traits"

○ Pangenomics in plants



The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor

Lei Gao[1,10], Itay Gonda[1,2,10], Honghe Sun[1], Qiyue Ma[1], Kan Bao[1], Denise M. Tieman[3], Elizabeth A. Burzynski-Chang[4], Tara L. Fish[5], Kaitlin A. Stromberg[1], Gavin L. Sacks[4], Theodore W. Thannhauser[5], Majid R. Foolad[6], Maria Jose Diez[7], Jose Blanca[7], Joaquin Canizares[7], Yimin Xu[1], Esther van der Knaap[8], Sanwen Huang[9], Harry J. Klee[3], James J. Giovannoni[1,5]* and Zhangjun Fei[1,5]*

NATURE GENETICS | VOL 51 | JUNE 2019 | 1044–1051 | www.nature.com/naturegenetics

725 genomes of *Solanum cultivated + 3 wild species*

core: **74%** genes

"The most striking feature of the tomato pan-genome was its ***high core*** *gene content (74.2%), as compared with those of Arabidopsis thaliana (70%), Brassica napus (62%), bread wheat (64%), rice (54%), wild soybean (49%) and Brachypodium distachyon (35%)"*

○ Pangenomics in plants

The pangenome of hexaploid bread wheat

Juan D. Montenegro[1,†], Agnieszka A. Golicz[1,2,†,‡], Philipp E. Bayer[2,†], Bhavna Hurgobin[1,2], HueyTyng Lee[1,2], Chon-Kit Kenneth Chan[2], Paul Visendi[1], Kaitao Lai[3], Jaroslav Doležel[4], Jacqueline Batley[1,2,5] and David Edwards[1,2,5,*]

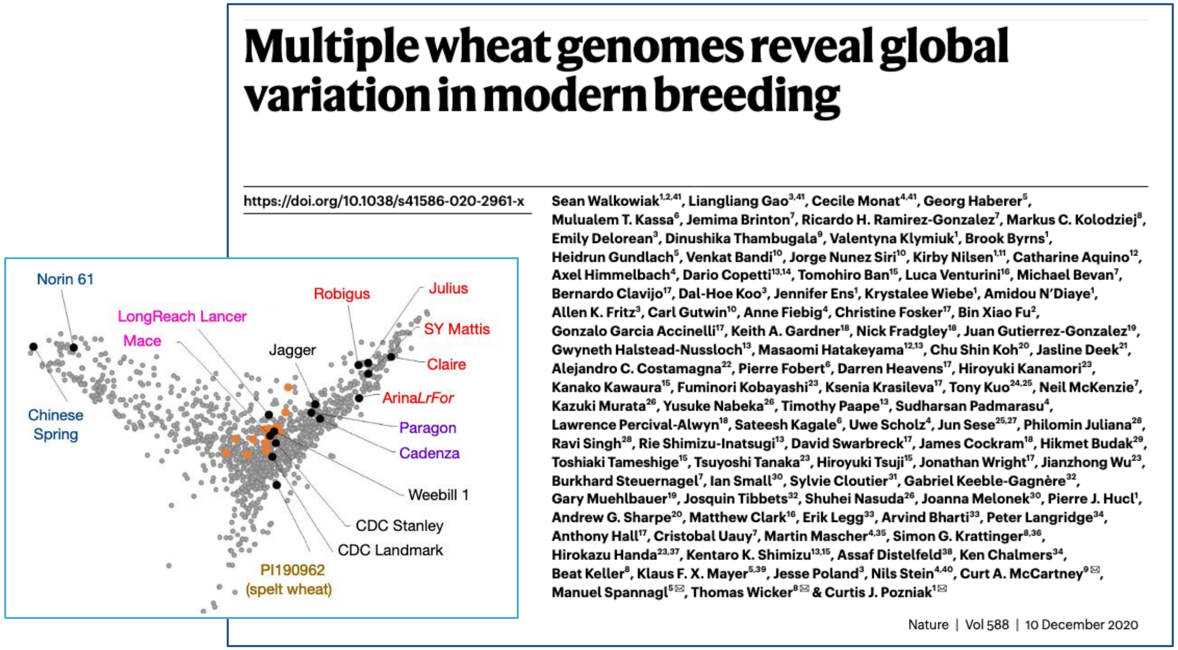
The Plant Journal (2017) **90**, 1007–1013

16 Australian wheat cultivars (Illumina reseq)

dispensable: **~35%** of genes

15 wheat genomes assembled

- *"12% of genes showed PAVs"*
- *"26% of the projected genes were found in tandem duplications, indicating that CNV is a strong contributor of genetic variation in wheat"*



Multiple wheat genomes reveal global variation in modern breeding

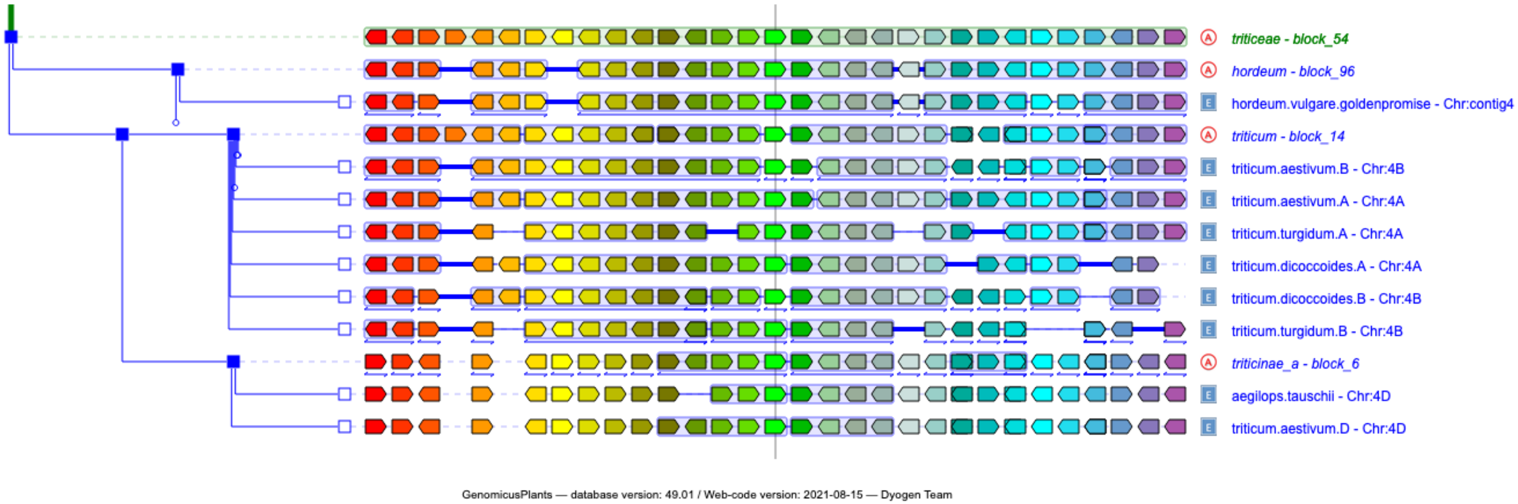
https://doi.org/10.1038/s41586-020-2961-x

Sean Walkowiak[1,2,41], Liangliang Gao[3,41], Cecile Monat[4,41], Georg Haberer[5], Mulualem T. Kassa[6], Jemima Brinton[7], Ricardo H. Ramirez-Gonzalez[7], Markus C. Kolodziej[8], Emily Delorean[3], Dinushika Thambugala[9], Valentyna Klymiuk[1], Brook Byrns[1], Heidrun Gundlach[5], Venkat Bandi[10], Jorge Nunez Siri[10], Kirby Nilsen[1,11], Catharine Aquino[12], Axel Himmelbach[4], Dario Copetti[13,14], Tomohiro Ban[15], Luca Venturini[16], Michael Bevan[7], Bernardo Clavijo[17], Dal-Hoe Koo[3], Jennifer Ens[1], Krystalee Wiebe[1], Amidou N'Diaye[1], Allen K. Fritz[3], Carl Gutwin[10], Anne Fiebig[4], Christine Fosker[17], Bin Xiao Fu[2], Gonzalo Garcia Accinelli[17], Keith A. Gardner[18], Nick Fradgley[18], Juan Gutierrez-Gonzalez[19], Gwyneth Halstead-Nussloch[13], Masaomi Hatakeyama[12,13], Chu Shin Koh[20], Jasline Deek[21], Alejandro C. Costamagna[22], Pierre Fobert[6], Darren Heavens[17], Hiroyuki Kanamori[23], Kanako Kawaura[15], Fuminori Kobayashi[23], Ksenia Krasileva[17], Tony Kuo[24,25], Neil McKenzie[7], Kazuki Murata[26], Yusuke Nabeka[26], Timothy Paape[13], Sudharsan Padmarasu[4], Lawrence Percival-Alwyn[18], Sateesh Kagale[6], Uwe Scholz[4], Jun Sese[25,27], Philomin Juliana[28], Ravi Singh[28], Rie Shimizu-Inatsugi[13], David Swarbreck[17], James Cockram[18], Hikmet Budak[29], Toshiaki Tameshige[15], Tsuyoshi Tanaka[23], Hiroyuki Tsuji[15], Jonathan Wright[17], Jianzhong Wu[23], Burkhard Steuernagel[7], Ian Small[30], Sylvie Cloutier[31], Gabriel Keeble-Gagnère[32], Gary Muehlbauer[19], Josquin Tibbets[32], Shuhei Nasuda[26], Joanna Melonek[30], Pierre J. Hucl[1], Andrew G. Sharpe[20], Matthew Clark[16], Erik Legg[33], Arvind Bharti[33], Peter Langridge[34], Anthony Hall[17], Cristobal Uauy[7], Martin Mascher[4,35], Simon G. Krattinger[8,36], Hirokazu Handa[23,37], Kentaro K. Shimizu[13,15], Assaf Distelfeld[38], Ken Chalmers[34], Beat Keller[8], Klaus F. X. Mayer[5,39], Jesse Poland[3], Nils Stein[4,40], Curt A. McCartney[9]✉, Manuel Spannagl[5]✉, Thomas Wicker[8]✉ & Curtis J. Pozniak[1]✉

Nature | Vol 588 | 10 December 2020

- Build and visualize pangenomes

Genomicus
Genomes in evolution



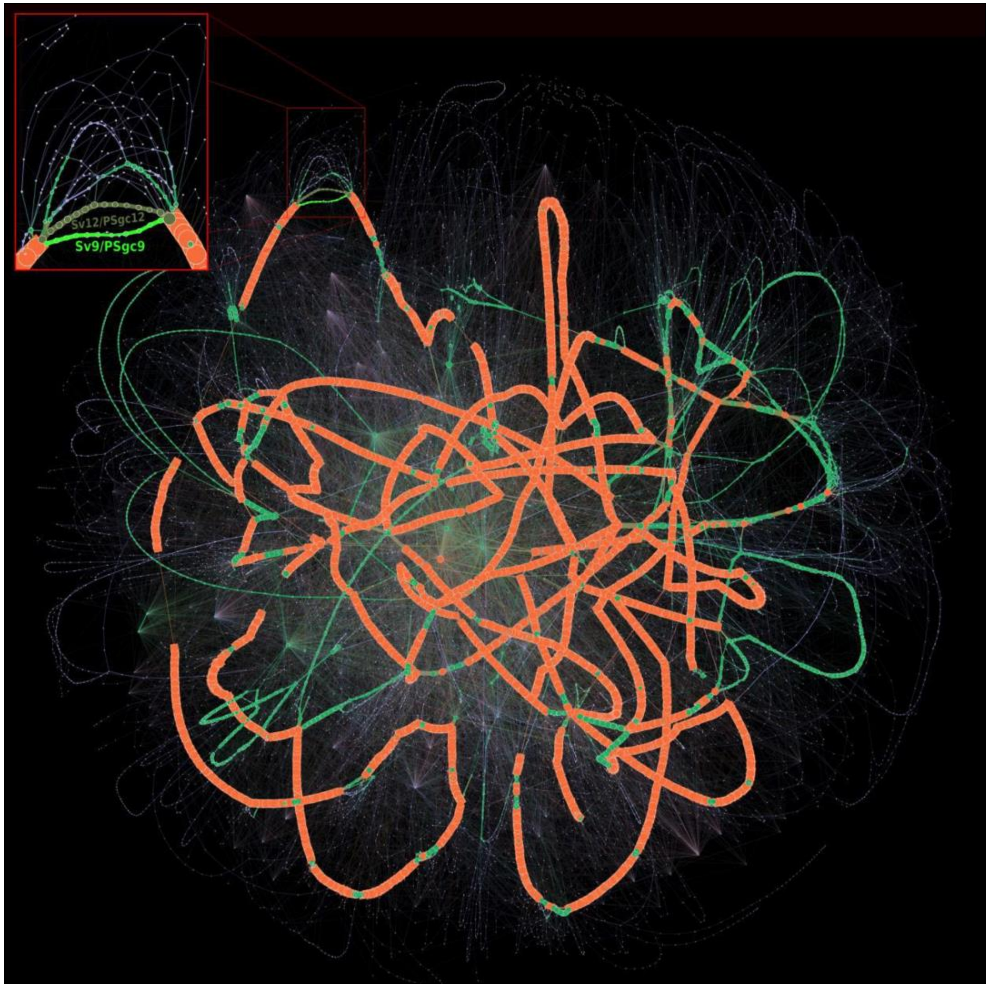
GenomicusPlants — database version: 49.01 / Web-code version: 2021-08-15 — Dyogen Team

Nguyen et al. NAR 2022

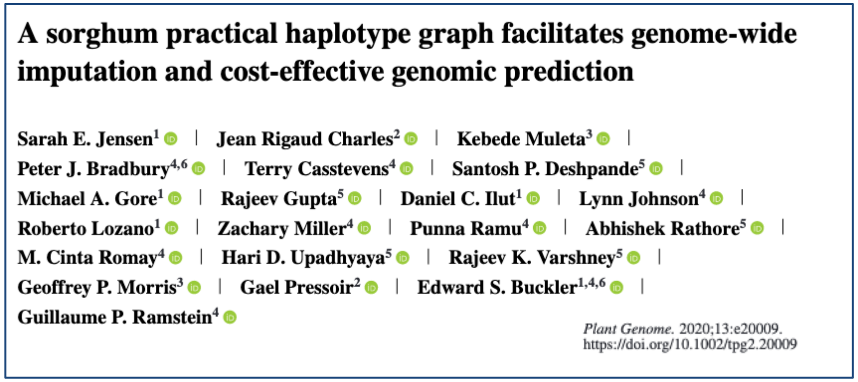
- **Gautereau et al. Plos Comput. Biol. 2020**

pangenome graph of ~3000 Acinetobacter

PPanGGOLiN



○ Practical Haplotype Graph

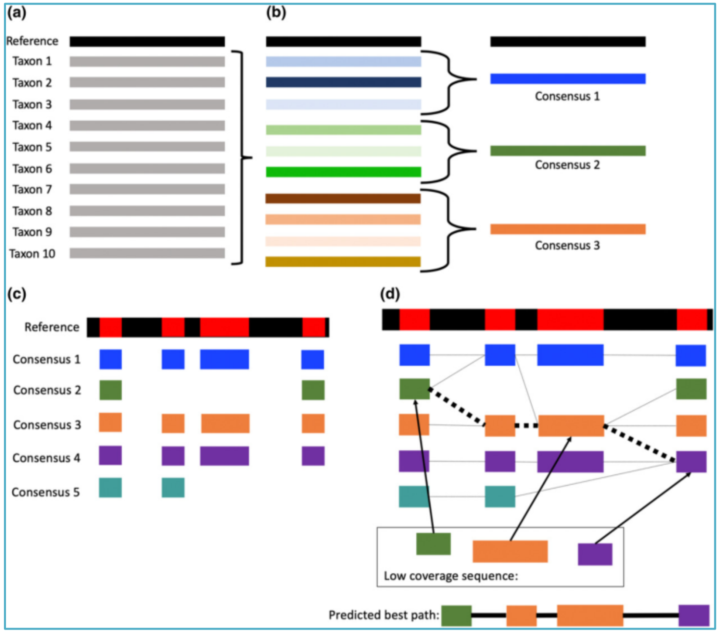


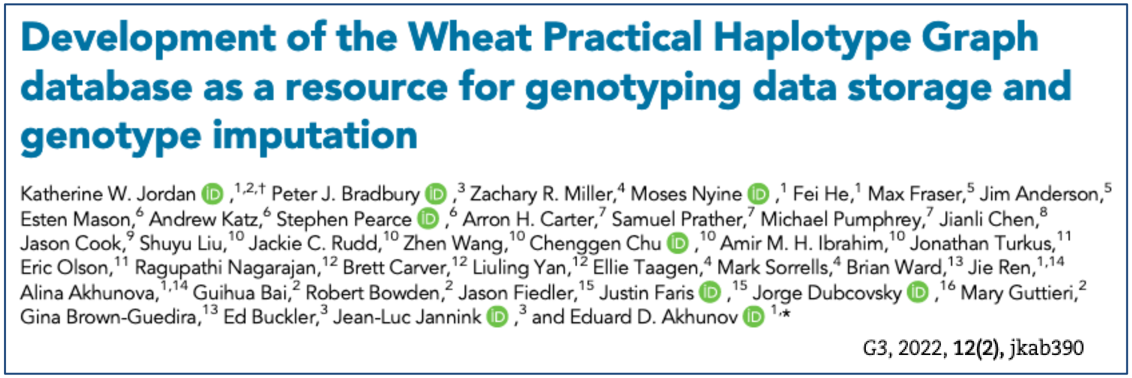
A sorghum practical haplotype graph facilitates genome-wide imputation and cost-effective genomic prediction

Sarah E. Jensen[1] | **Jean Rigaud Charles**[2] | **Kebede Muleta**[3] |
Peter J. Bradbury[4,6] | **Terry Casstevens**[4] | **Santosh P. Deshpande**[5] |
Michael A. Gore[1] | **Rajeev Gupta**[5] | **Daniel C. Ilut**[1] | **Lynn Johnson**[4] |
Roberto Lozano[1] | **Zachary Miller**[4] | **Punna Ramu**[4] | **Abhishek Rathore**[5] |
M. Cinta Romay[4] | **Hari D. Upadhyaya**[5] | **Rajeev K. Varshney**[5] |
Geoffrey P. Morris[3] | **Gael Pressoir**[2] | **Edward S. Buckler**[1,4,6] |
Guillaume P. Ramstein[4]

Plant Genome. 2020;13:e20009.
https://doi.org/10.1002/tpg2.20009

- for breeding programs
- based on a limited number of founder lines
- **0.01x** skim-seq ->call SNPs with 5% errors





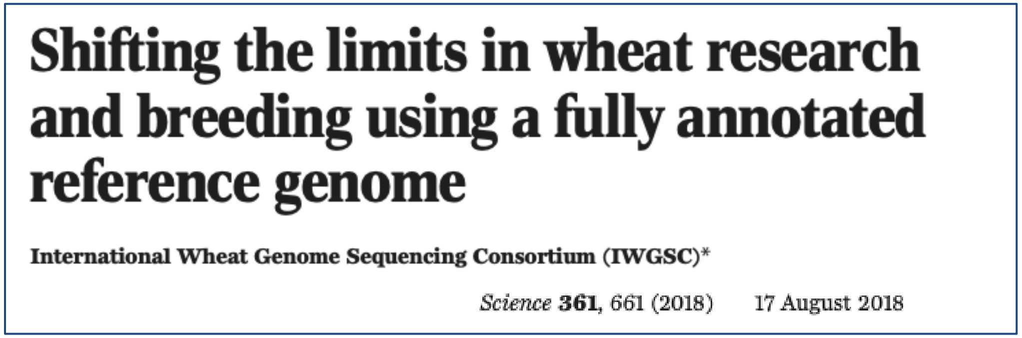
Development of the Wheat Practical Haplotype Graph database as a resource for genotyping data storage and genotype imputation

Katherine W. Jordan,[1,2,†] Peter J. Bradbury,[3] Zachary R. Miller,[4] Moses Nyine,[1] Fei He,[1] Max Fraser,[5] Jim Anderson,[5]
Esten Mason,[6] Andrew Katz,[6] Stephen Pearce,[6] Arron H. Carter,[7] Samuel Prather,[7] Michael Pumphrey,[7] Jianli Chen,[8]
Jason Cook,[9] Shuyu Liu,[10] Jackie C. Rudd,[10] Zhen Wang,[10] Chenggen Chu,[10] Amir M. H. Ibrahim,[10] Jonathan Turkus,[11]
Eric Olson,[11] Ragupathi Nagarajan,[12] Brett Carver,[12] Liuling Yan,[12] Ellie Taagen,[4] Mark Sorrells,[4] Brian Ward,[13] Jie Ren,[1,14]
Alina Akhunova,[1,14] Guihua Bai,[2] Robert Bowden,[2] Jason Fiedler,[15] Justin Faris,[15] Jorge Dubcovsky,[16] Mary Guttieri,[2]
Gina Brown-Guedira,[13] Ed Buckler,[3] Jean-Luc Jannink,[3] and Eduard D. Akhunov[1,*]

G3, 2022, **12(2)**, jkab390

- 65 wheat lines, exome capture
- 0.01x -> imputations with 8% errors

2. Wheat genome sequencing, annotation, A-B-D comparative genomics



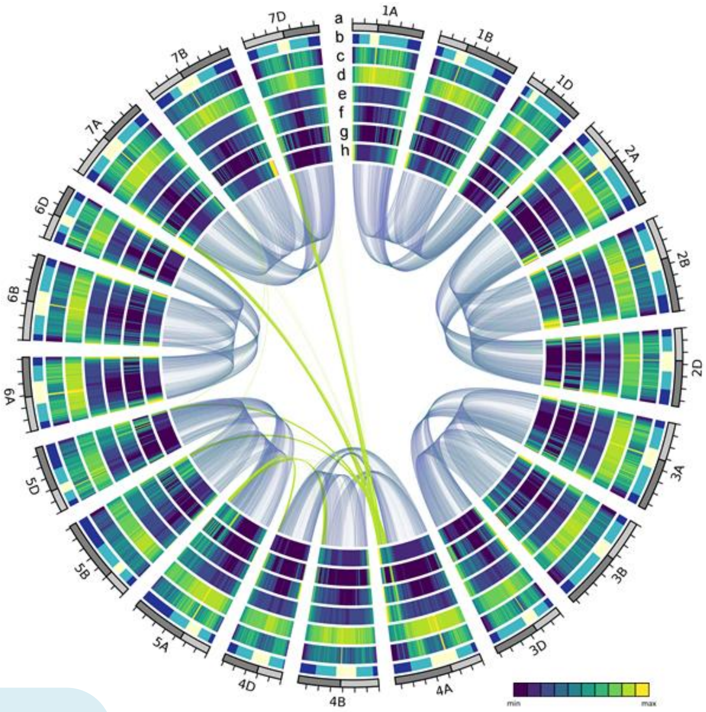
Shifting the limits in wheat research and breeding using a fully annotated reference genome

International Wheat Genome Sequencing Consortium (IWGSC)*

Science **361**, 661 (2018) 17 August 2018

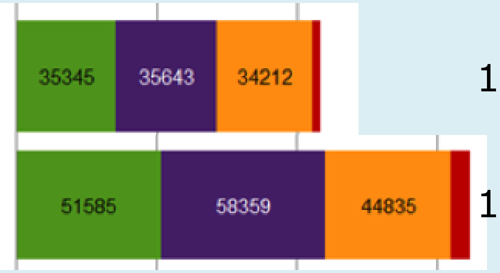
○ Predicting genes is still NOT routine

GDEC Clermont, France – *Rimbert Leroy Choulet et al.*
PGSB Munich, Germany – *Spannagl Twardziok et al.*
EI Norwich, UK – *Swarbreck Venturini et al.*

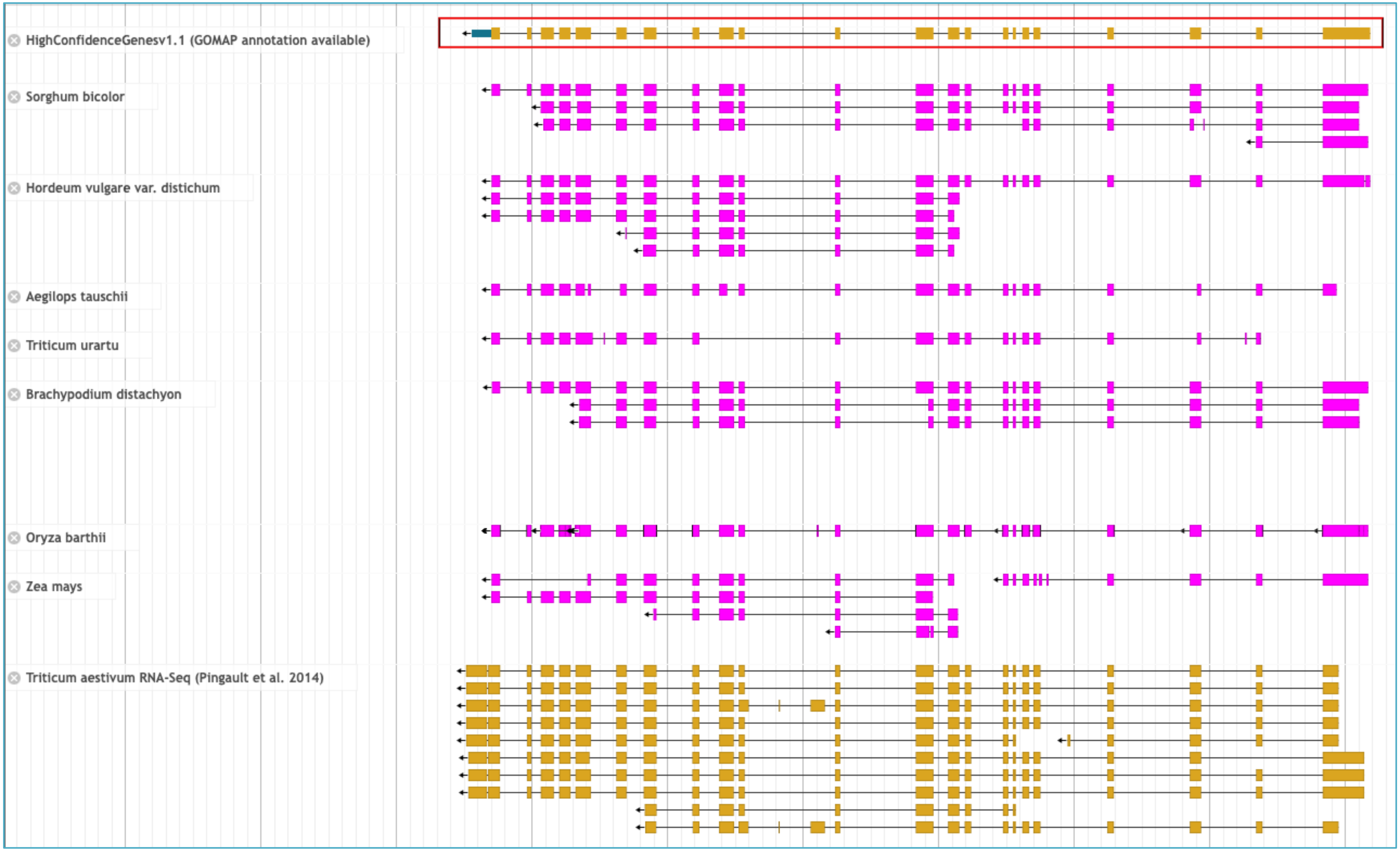


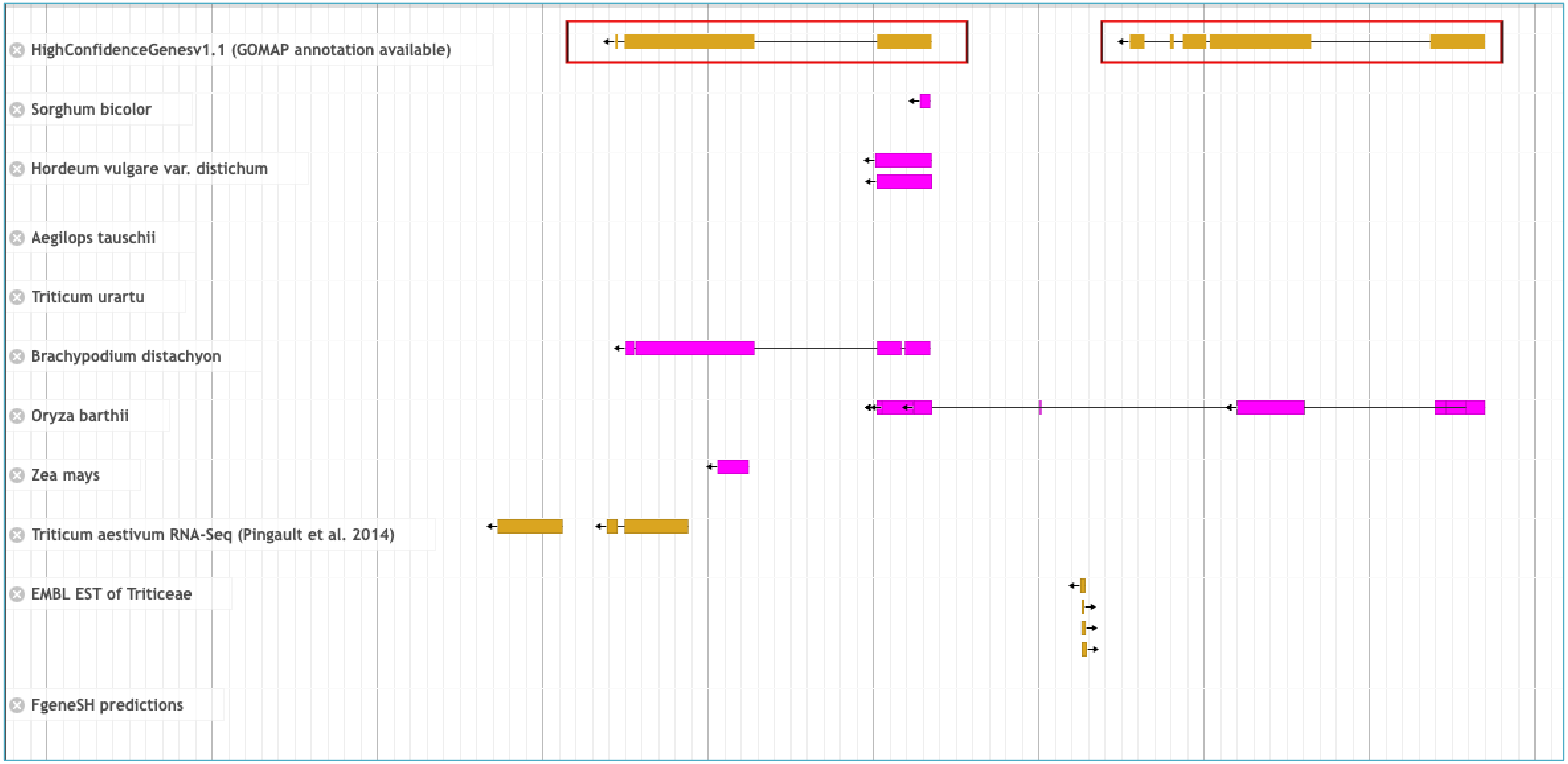
IWGSC RefSeq (Chinese Spring)

- 107,891 genes



HC genes	107k
LC genes	161k
pseudogenes	~300k

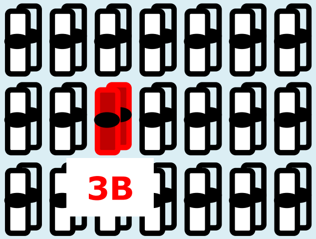

HighConfidenceGenesv1.1 (GOMAP annotation available)
Sorghum bicolor
Hordeum vulgare var. distichum
Aegilops tauschii
Triticum urartu
Brachypodium distachyon
Oryza barthii
Zea mays
Triticum aestivum RNA-Seq (Pingault et al. 2014)


HighConfidenceGenesv1.1 (GOMAP annotation available)
Sorghum bicolor
Hordeum vulgare var. distichum
Aegilops tauschii
Triticum urartu
Brachypodium distachyon
Oryza barthii
Zea mays
Triticum aestivum RNA-Seq (Pingault et al. 2014)
EMBL EST of Triticeae
FgeneSH predictions

Impact of annotation methods on knowledge

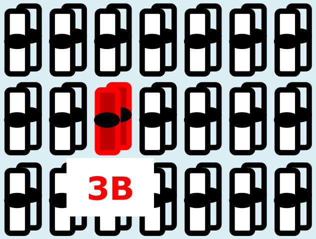
- in 2017: 3 versions of Chinese Spring chr3B

chr3B BAC-by-BAC



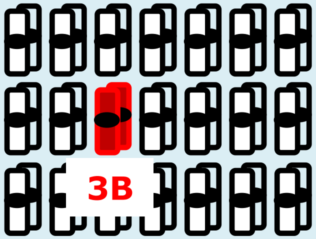
Choulet et al. Science 2014

TGAC_v1



Clavijo et al. Genom Biol 2017

Chr. Survey Seq



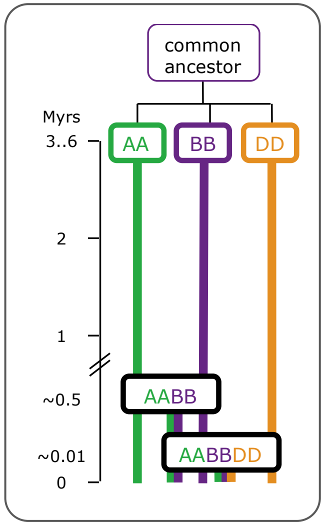
IWGSC Science 2014

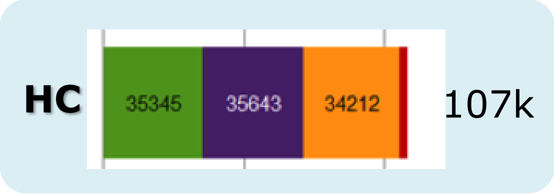
7264 genes

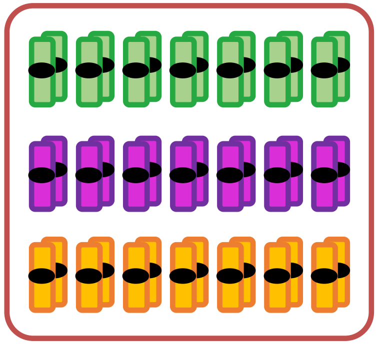
5728 genes
-> (**26**% identical)

5862 genes
(**12**% identical)

○ A-B-D Comparative Genomics





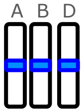


35k HC genes

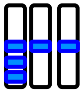
36k

34k

- 39k homeologous groups [including **29k** LC genes]

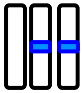


55% 1:1:1 triads



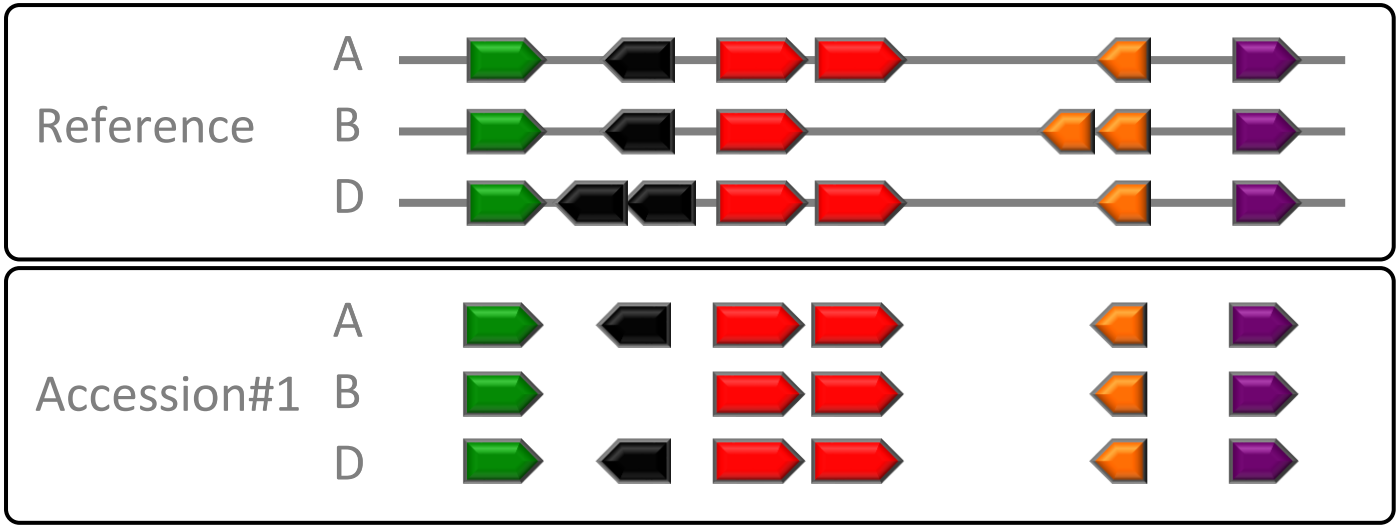
15% N:1:1 (duplications)

[27% of genes are duplicates]
ex: 207 PPRs

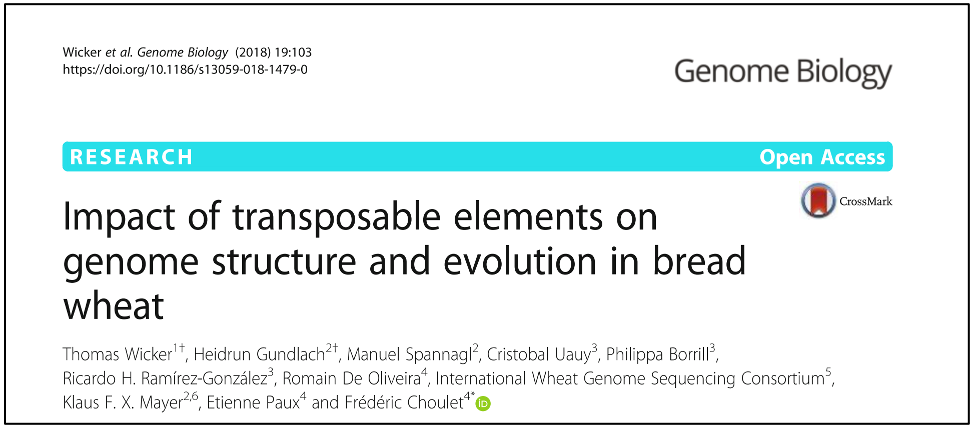


30%
0:1:1 11%
1:0:1 10%
1:1:0 9%

- Polyploidy + single gene duplications



3. A-B-D Comparative Analysis of **TEs**



Wicker *et al. Genome Biology* (2018) 19:103
https://doi.org/10.1186/s13059-018-1479-0

Genome Biology

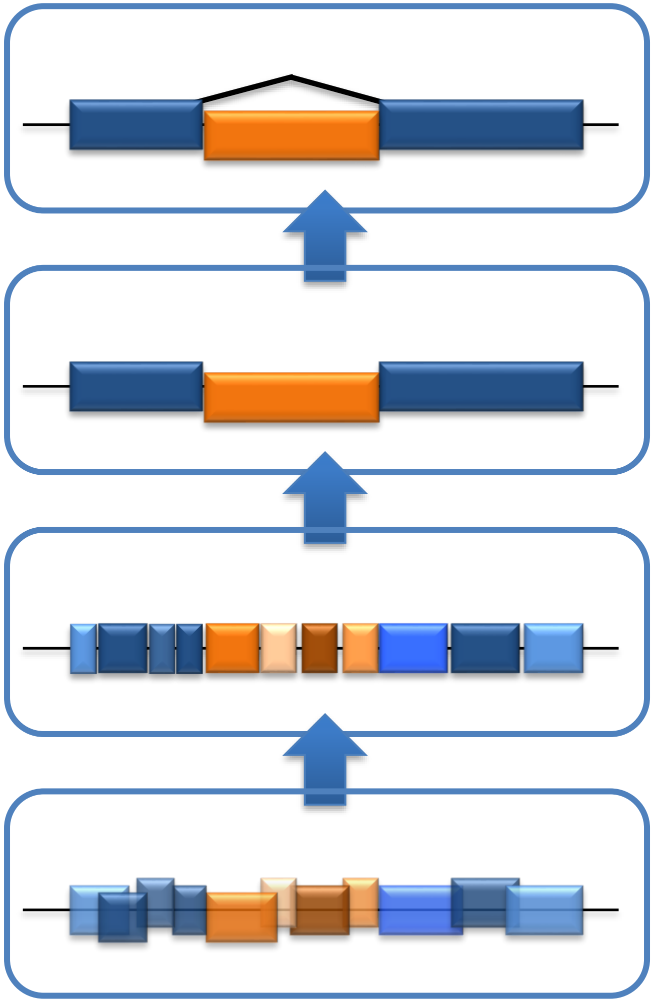
RESEARCH Open Access

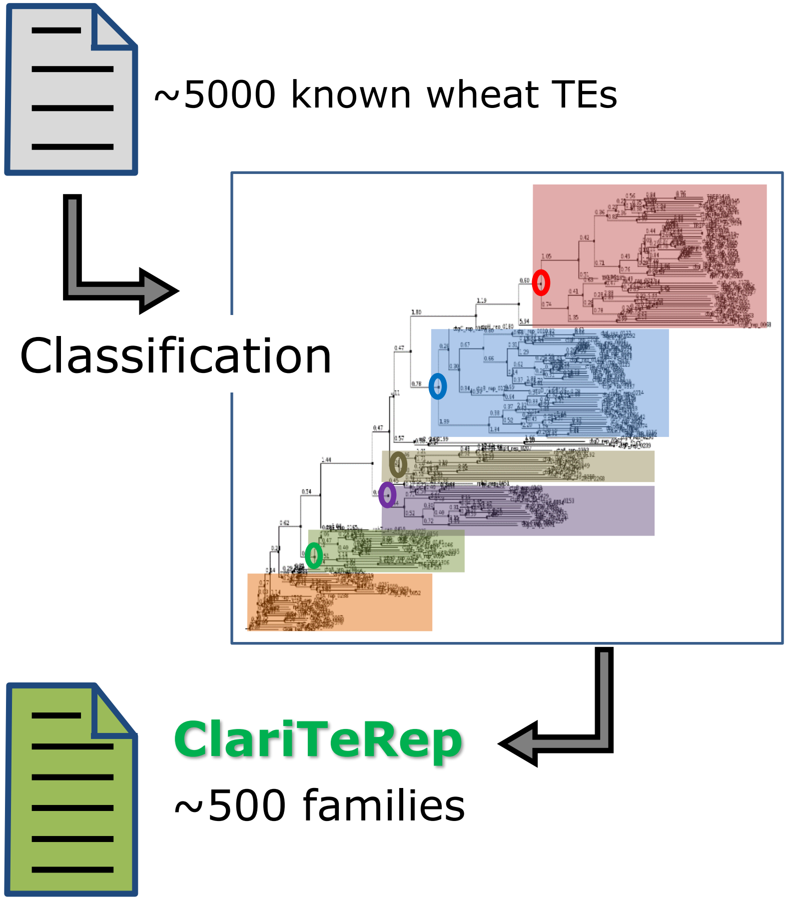
Impact of transposable elements on genome structure and evolution in bread wheat

Thomas Wicker[1†], Heidrun Gundlach[2†], Manuel Spannagl[2], Cristobal Uauy[3], Philippa Borrill[3], Ricardo H. Ramírez-González[3], Romain De Oliveira[4], International Wheat Genome Sequencing Consortium[5], Klaus F. X. Mayer[2,6], Etienne Paux[4] and Frédéric Choulet[4*]

TE modeling with CLARI-TE and ClariTeRep

https://github.com/jdaron/CLARI-TE





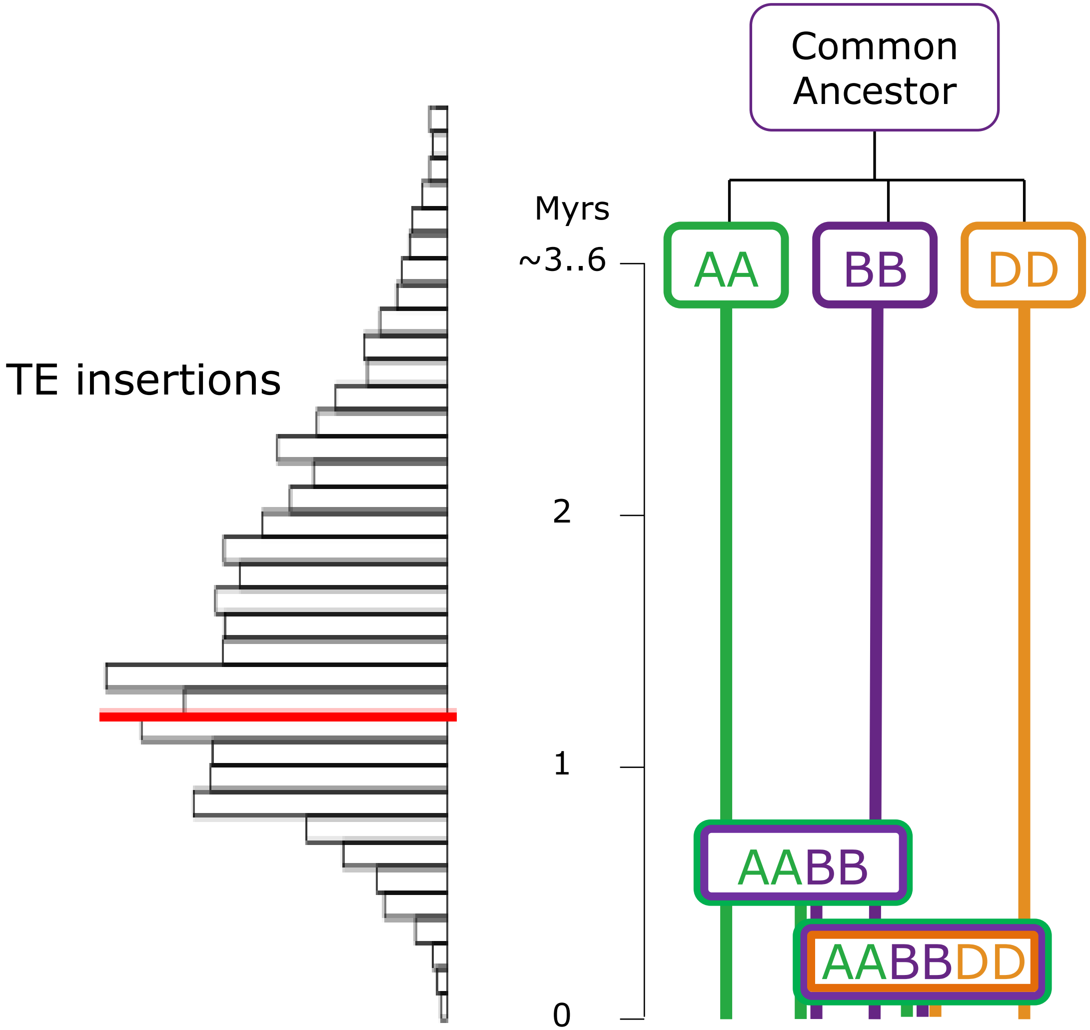
~5000 known wheat TEs

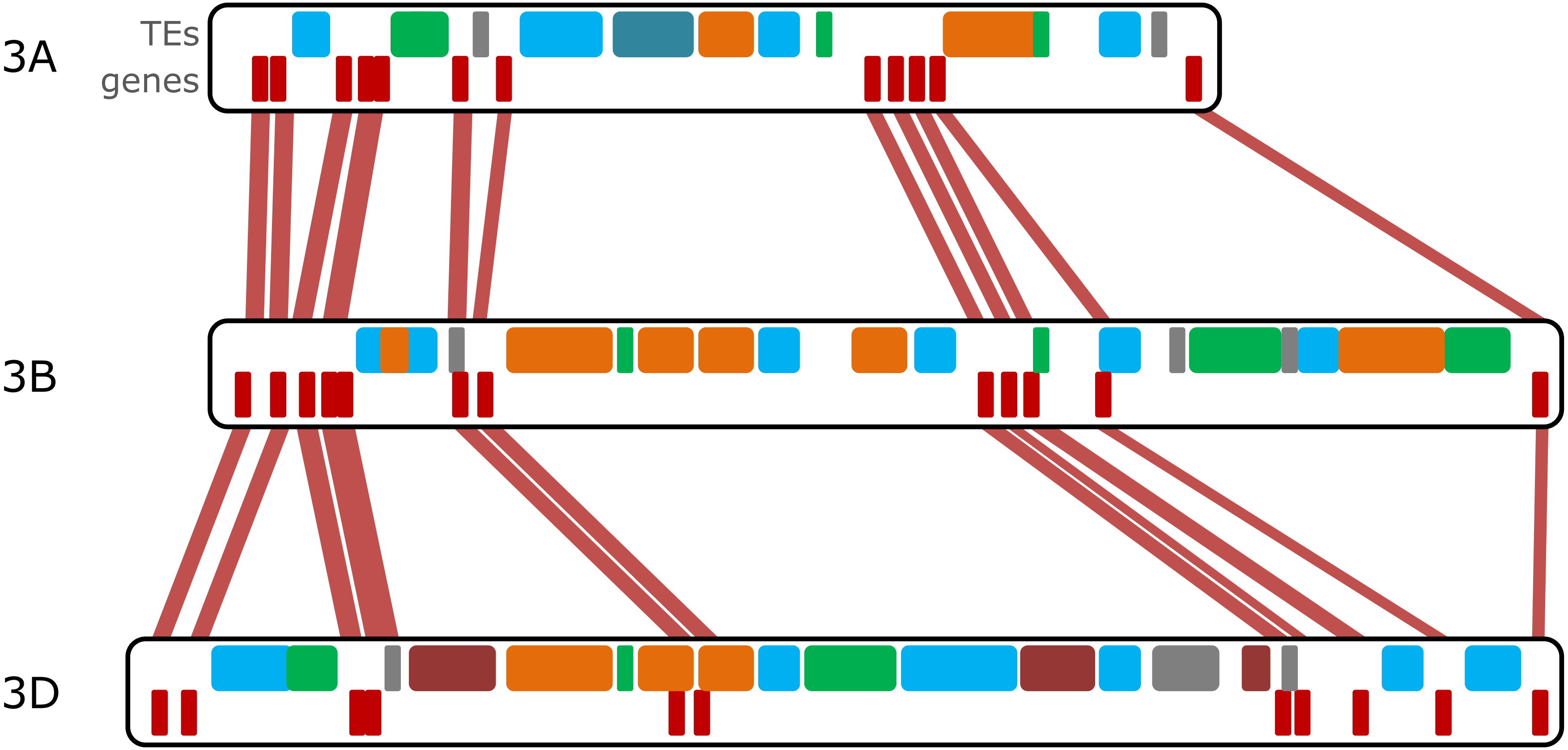
Classification

ClariTeRep

~500 families

Daron et al. Genom Biol 2015

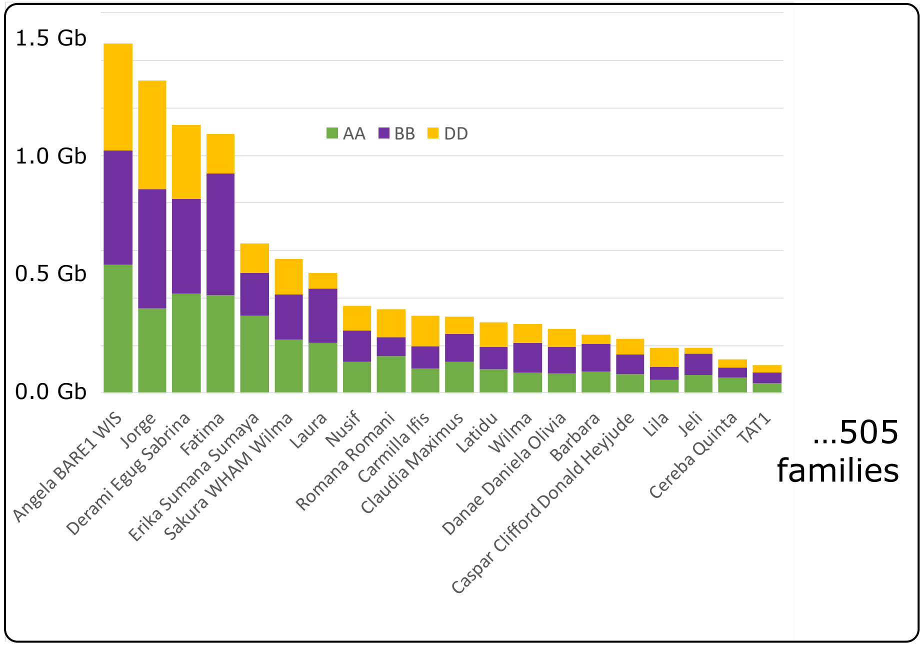

Common Ancestor
TE insertions
Myrs
~3..6
2
1
0
AA
BB
DD
AABB
AABBDD



- Near-complete **TE turnover** since A-B-D divergence

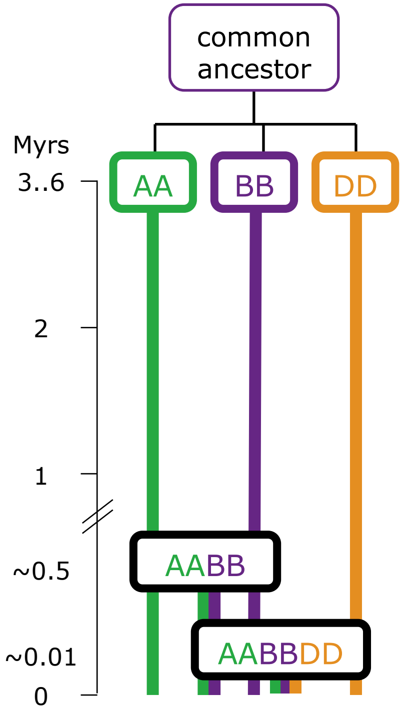
○ % TEs

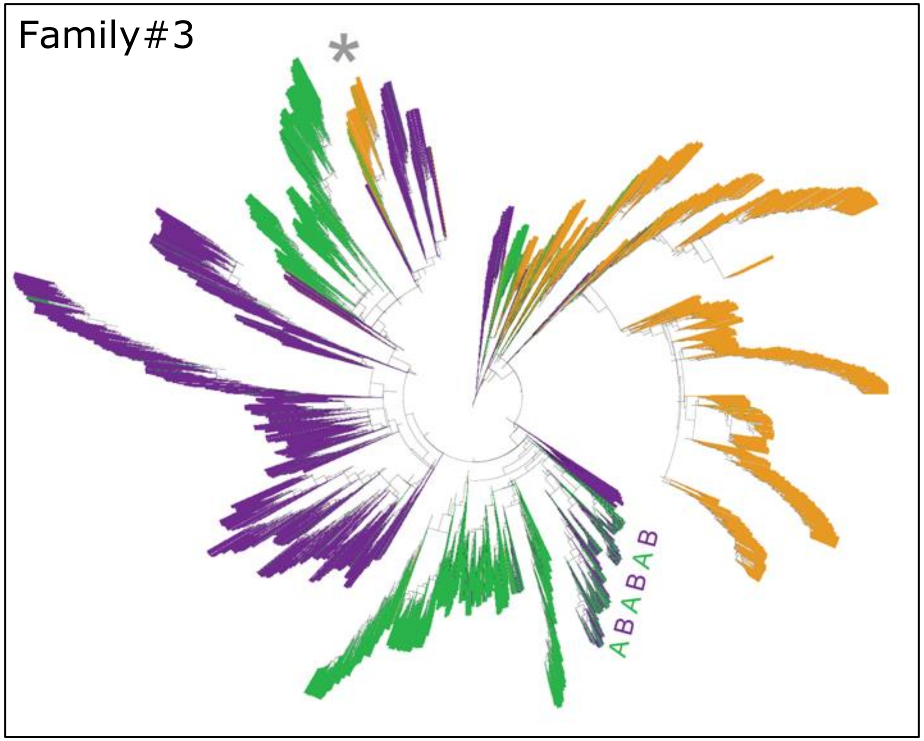
Families



- ➢ Abundant families are the same on A-B-D
- ➢ **0** family specific to 1 subgenome
- ➢ **76%** of the fam. account for similar % on A-B-D (<2 fold-change)

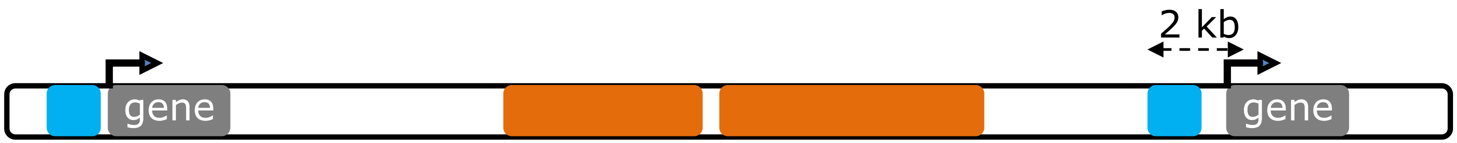
○ TE trees

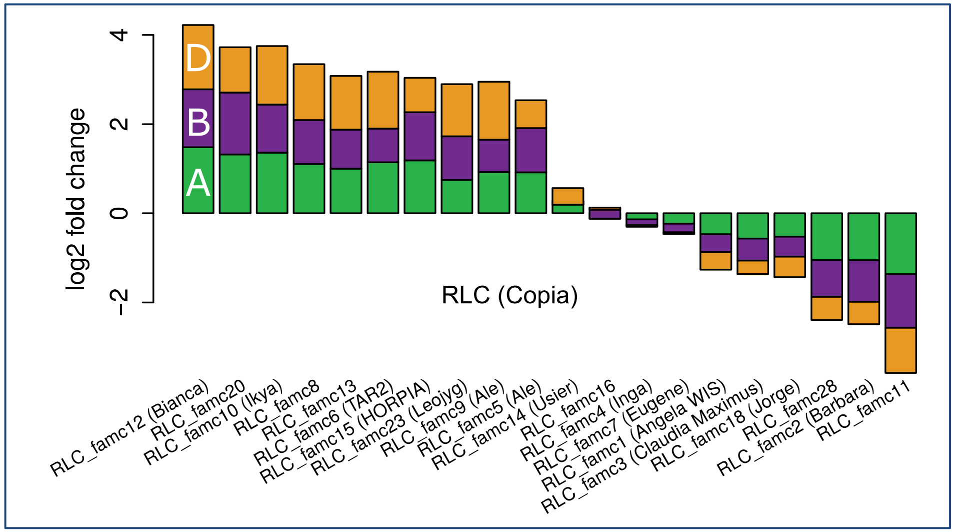


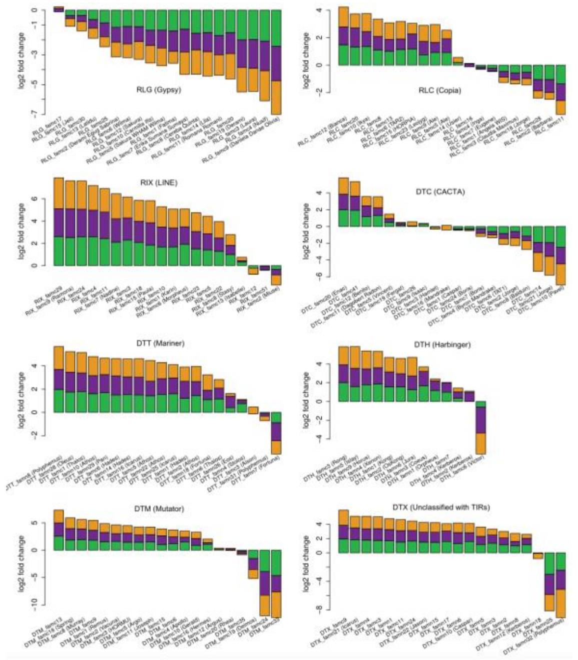


- ➢ Independent TE evolution in the diploids AA, BB, DD
- ➢ Some rare cases of TE amplif in the tetraploid AABB

○ TE content around genes

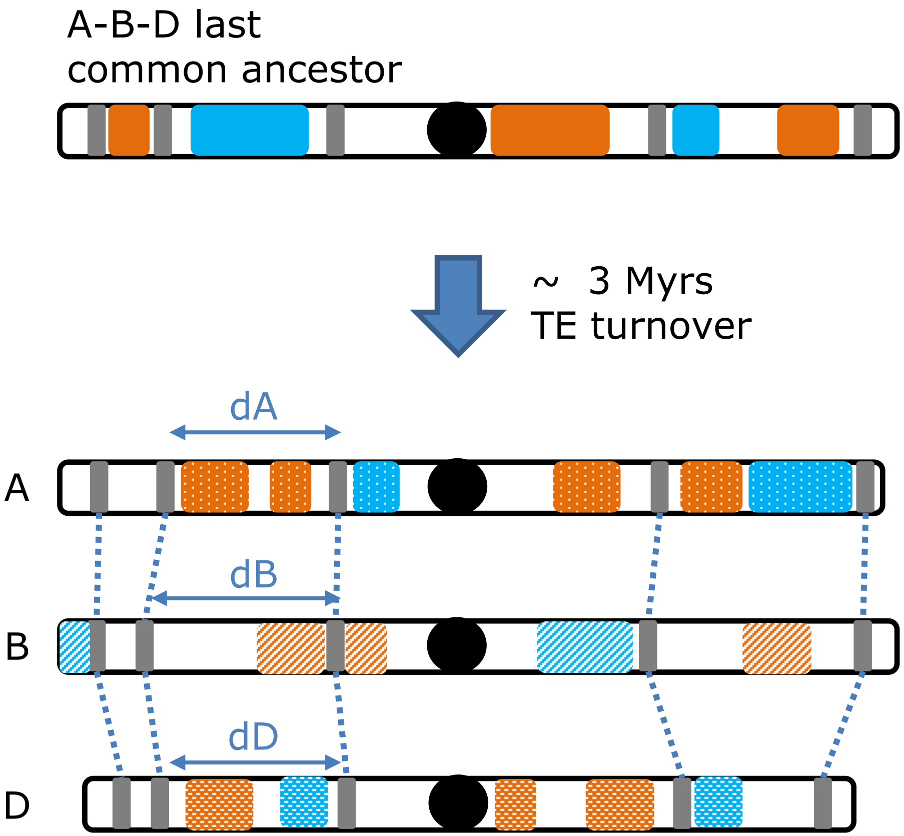






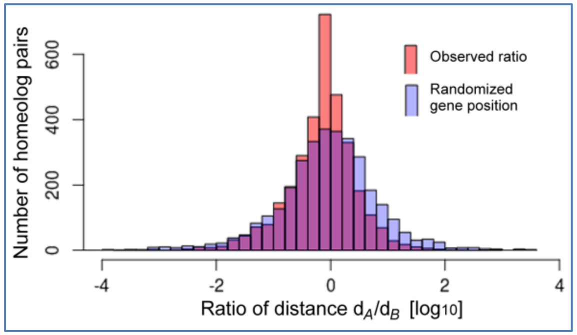
- ➢ Wide majority of TE families either enriched/underrepres. in gene promoters
- ➢ Enrichment/Underrepres. close to genes is conserved betw A-B-D

○ TE dynamics



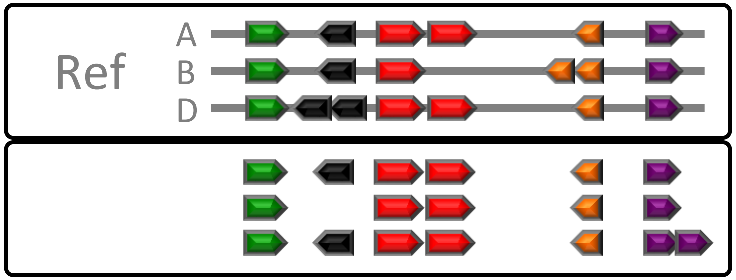
- Most families were active
- **Equilibrium**

-> Hypothesis of a structural role of TEs likely under selection pressure?



4. Structural Variations in *Triticum* using resequencing data

PhD Romain De Oliveira



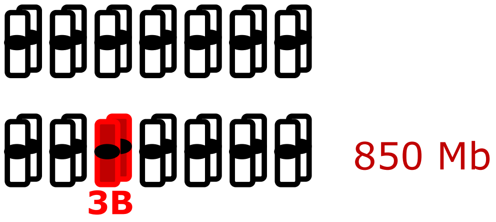

Frontiers in Genetics | www.frontiersin.org August 2020 | Volume 11 | Article 891

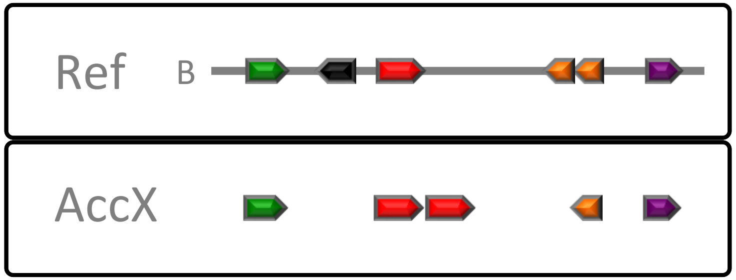
Structural Variations Affecting Genes and Transposable Elements of Chromosome 3B in Wheats

*Romain De Oliveira[1], Hélène Rimbert[1], François Balfourier[1], Jonathan Kitt[1], Emeric Dynomant[1], Jan Vrána[2], Jaroslav Doležel[2], Federica Cattonaro[3], Etienne Paux[1] and Frédéric Choulet[1]**

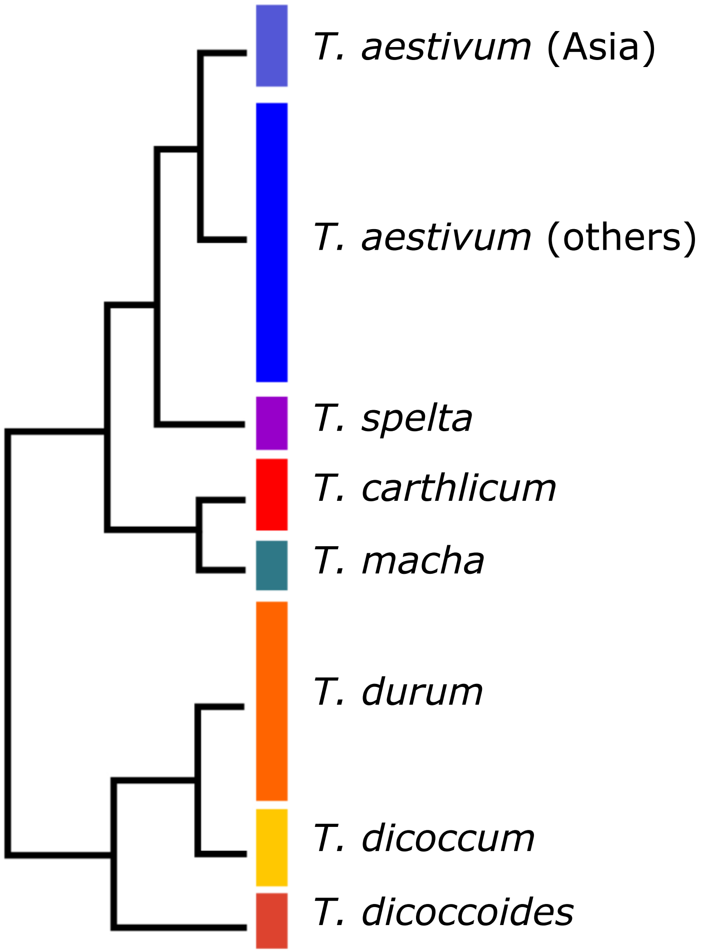
45 accessions / 7 species

-> sorted 3B DNA

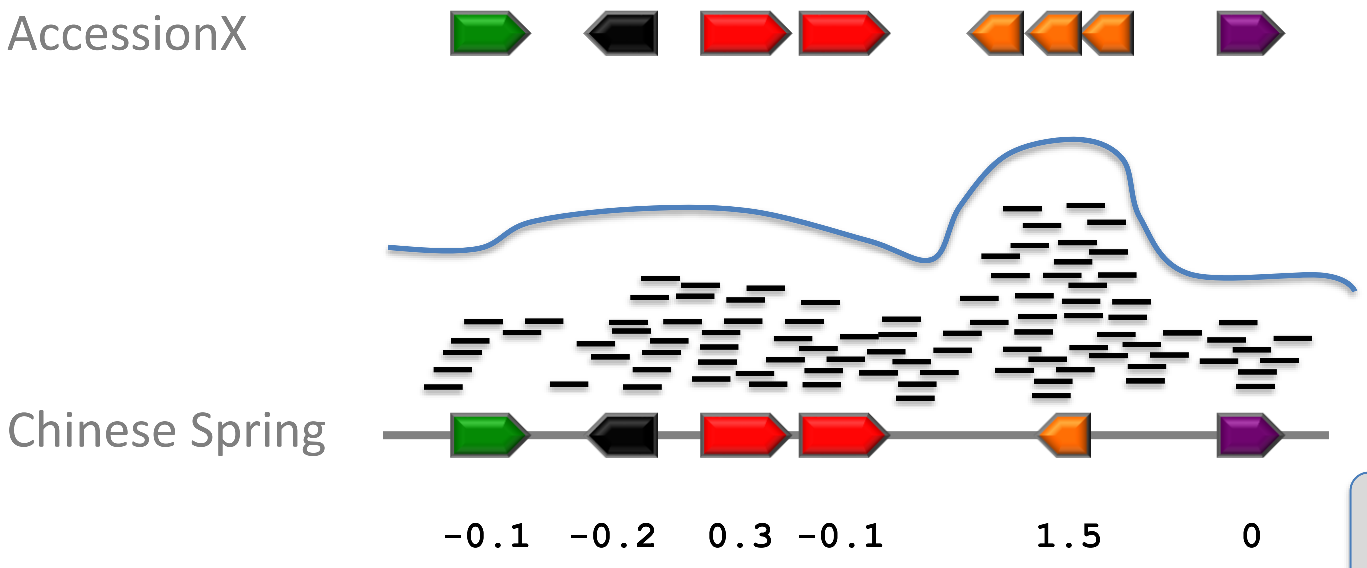




Illumina reseq ~30x

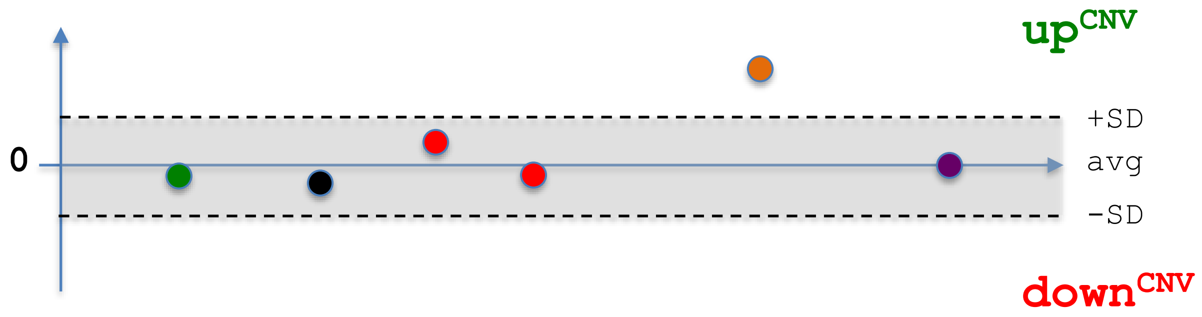


De Oliveira et al. Front. Genet. 2020

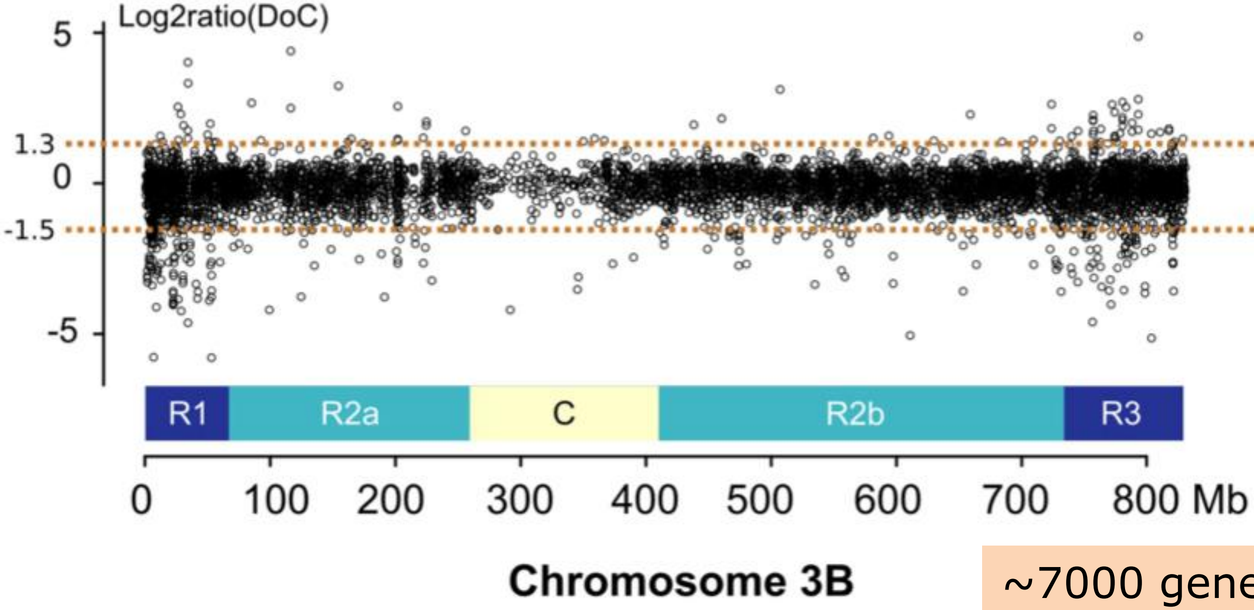


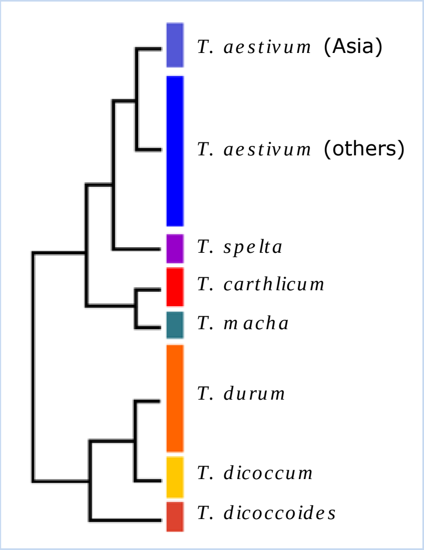
$$\log_2 \frac{cov^{accX}}{cov^{ChineseSpring}}$$

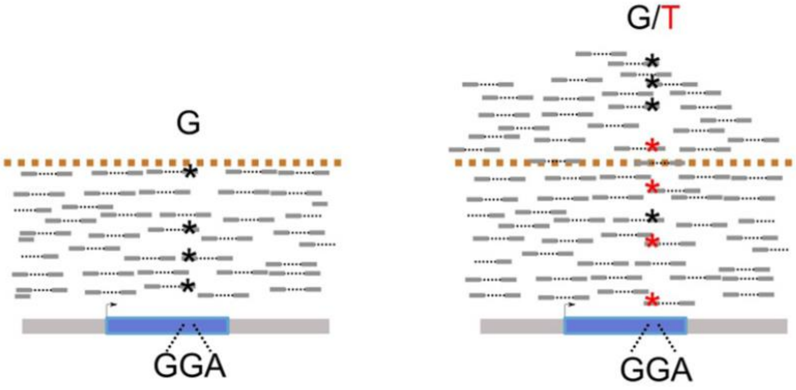
CNVs



De Oliveira et al. Front. Genet. 2020







~7000 genes

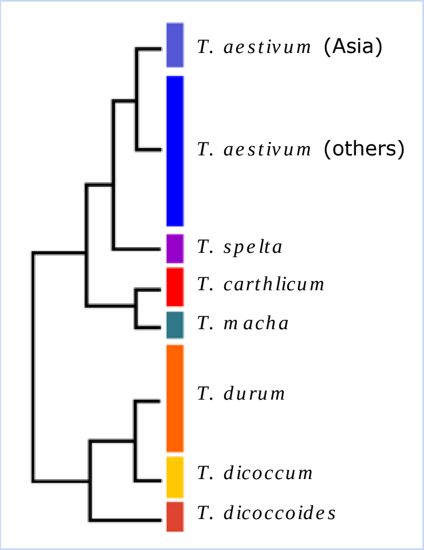
PAVs/CNVs: 20%

new genes: 330

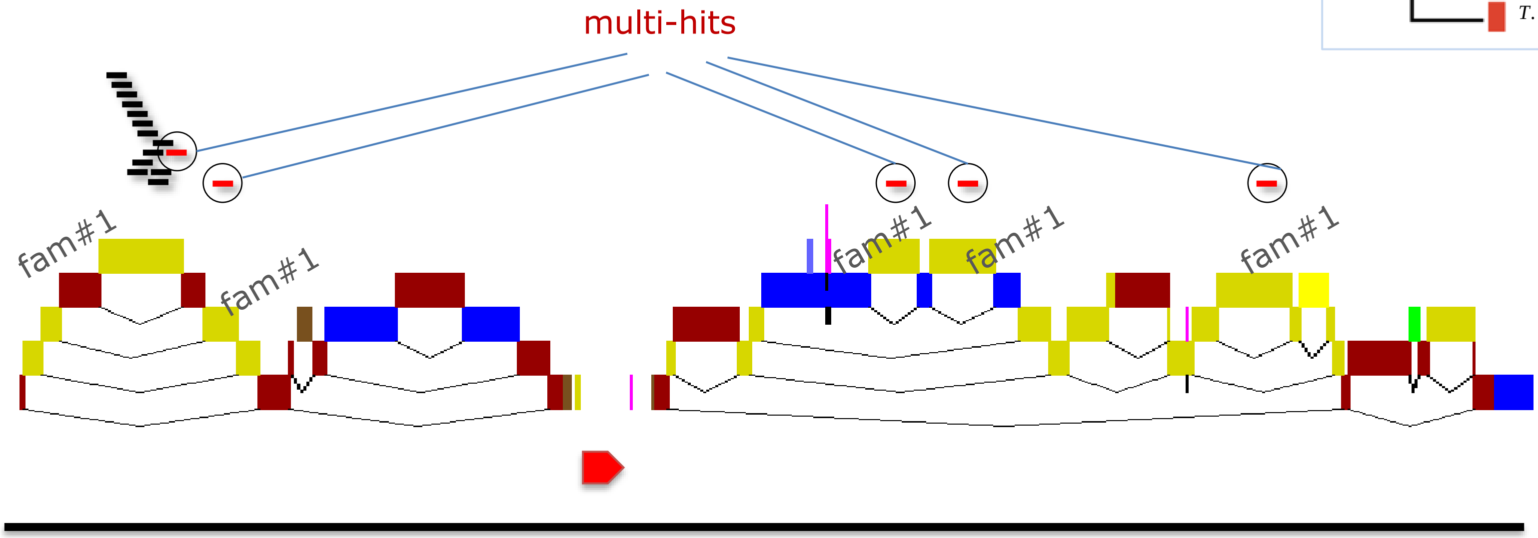
- functions: related to stress response
- expressed in more specific conditions

core=80%

De Oliveira et al. Front. Genet. 2020



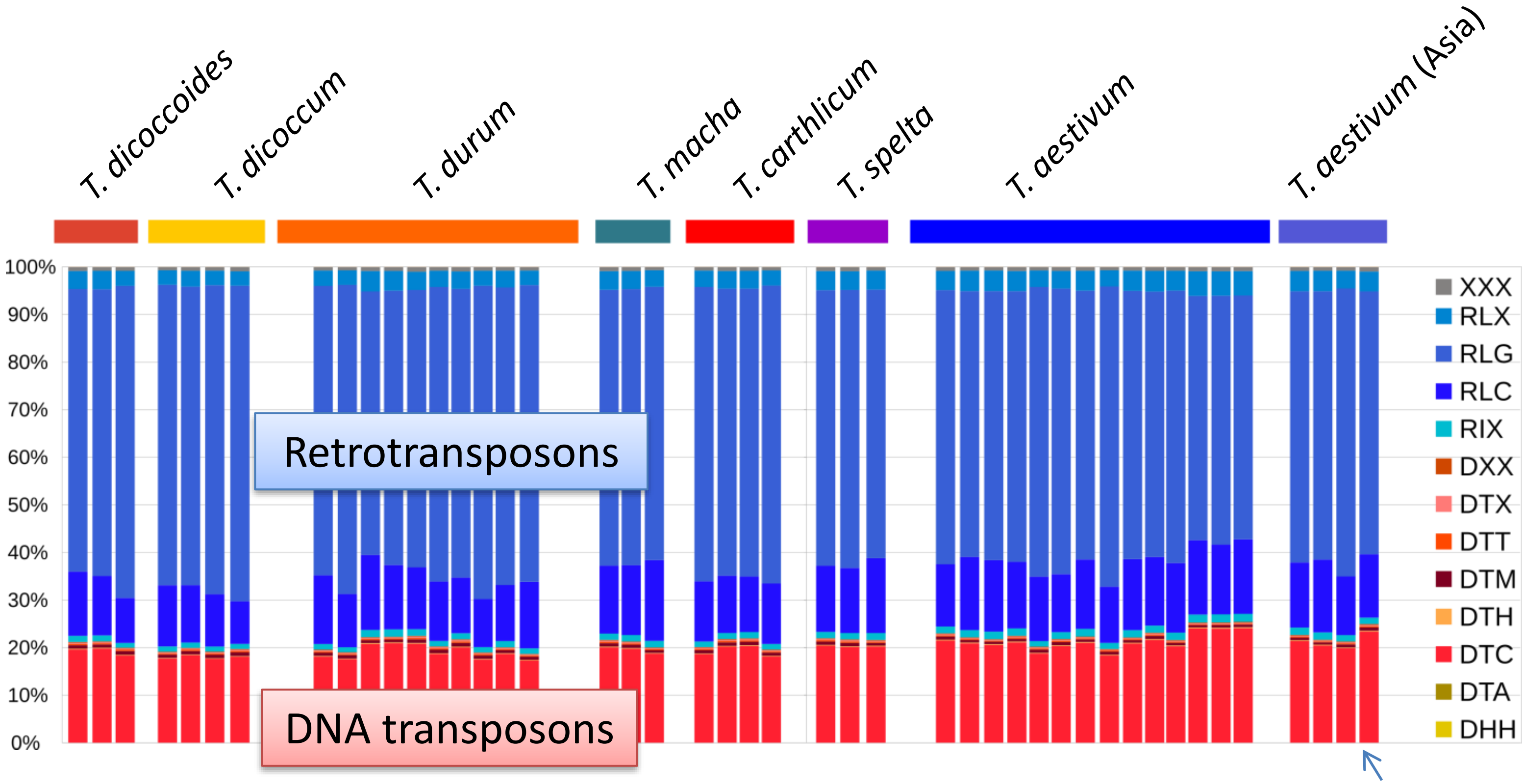
- ○ **TE CNVs**



Chinese Spring chr3B
~500 TE families

○ TE CNVs

TE% = **84%** in all 45 accessions



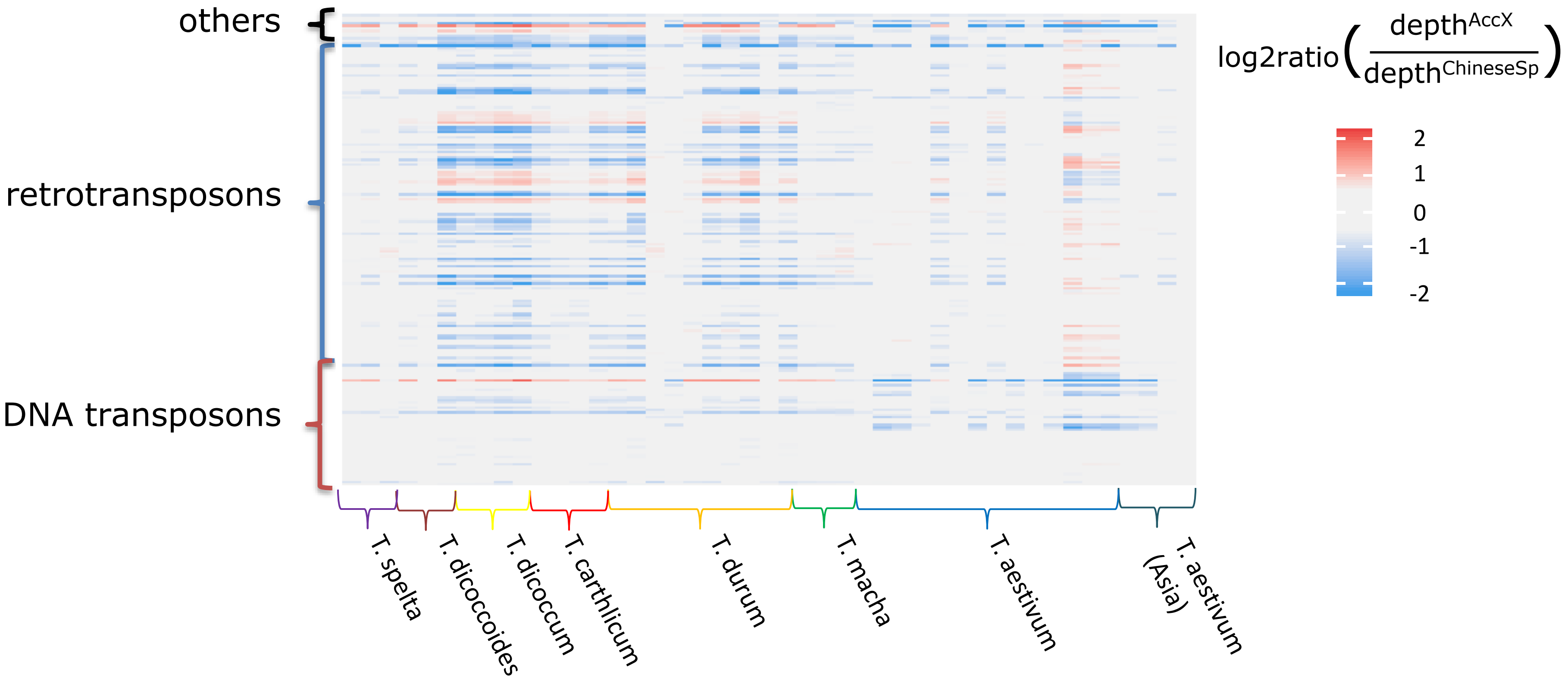
Retrotransposons

DNA transposons

Chinese spring

- ➢ No major difference of the number of TE copies for each superfamily

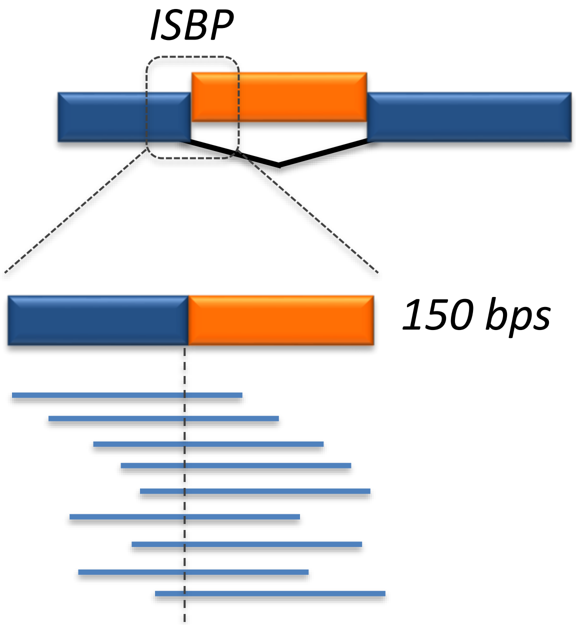
○ TE CNVs

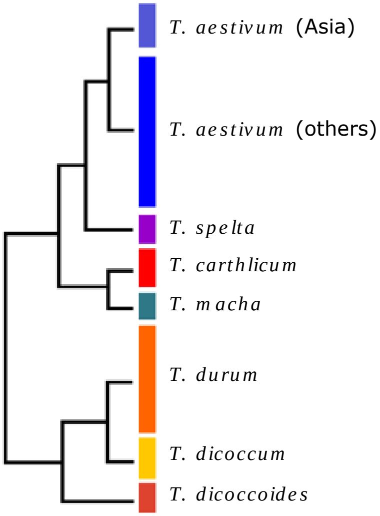


0 fam with log2ratio > 2

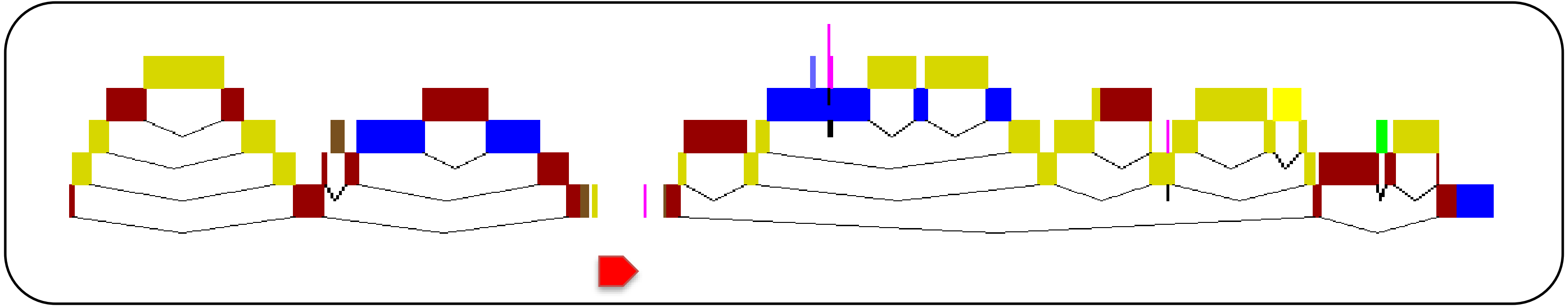
➢ At the family level: globally extremely conserved proportions

○ TE PAVs





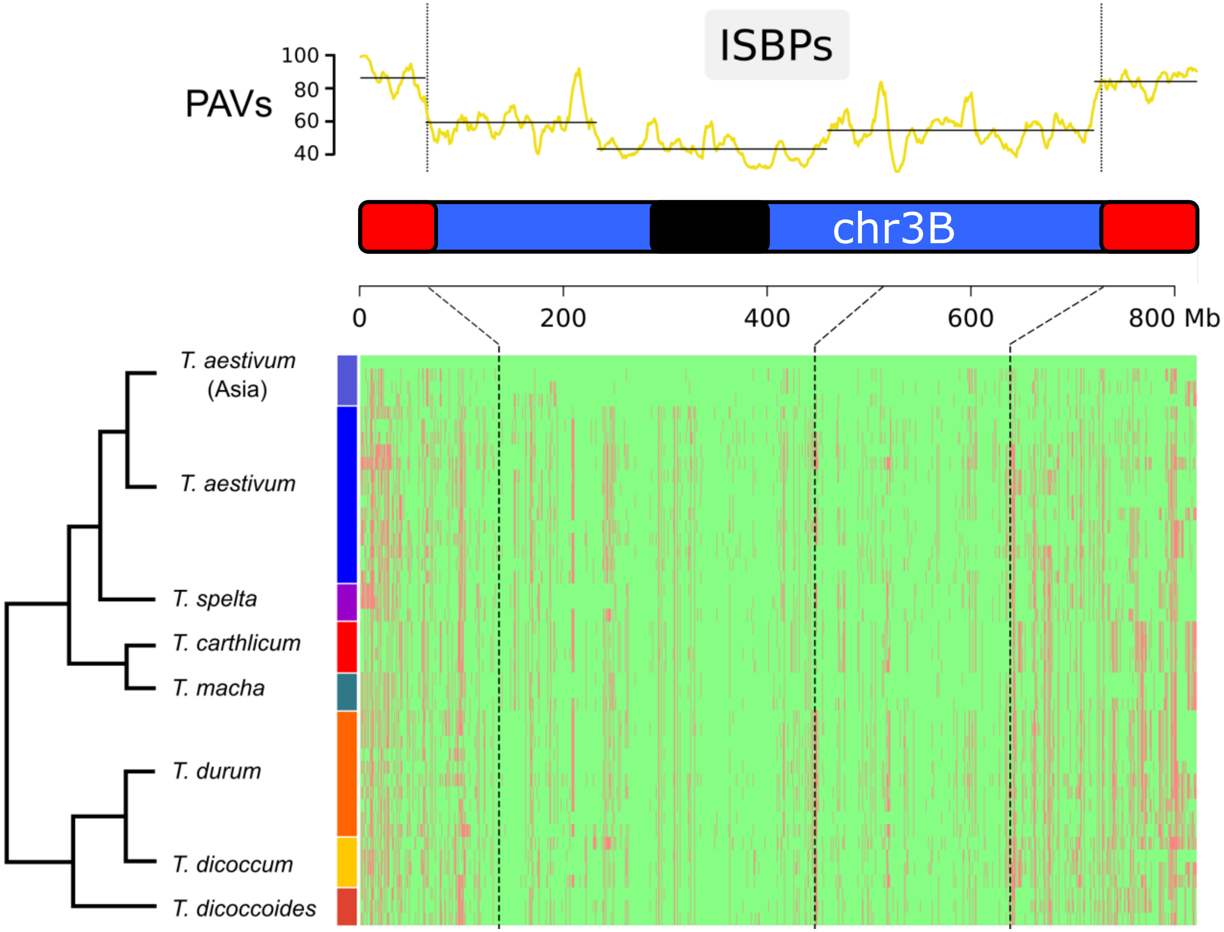
1 ISBP every 3 kb



RESULTS for 45 *Triticum* accessions:

- from **7%** to **19%** TE-PAVs
- cumulated variable: **60%** ISBPs

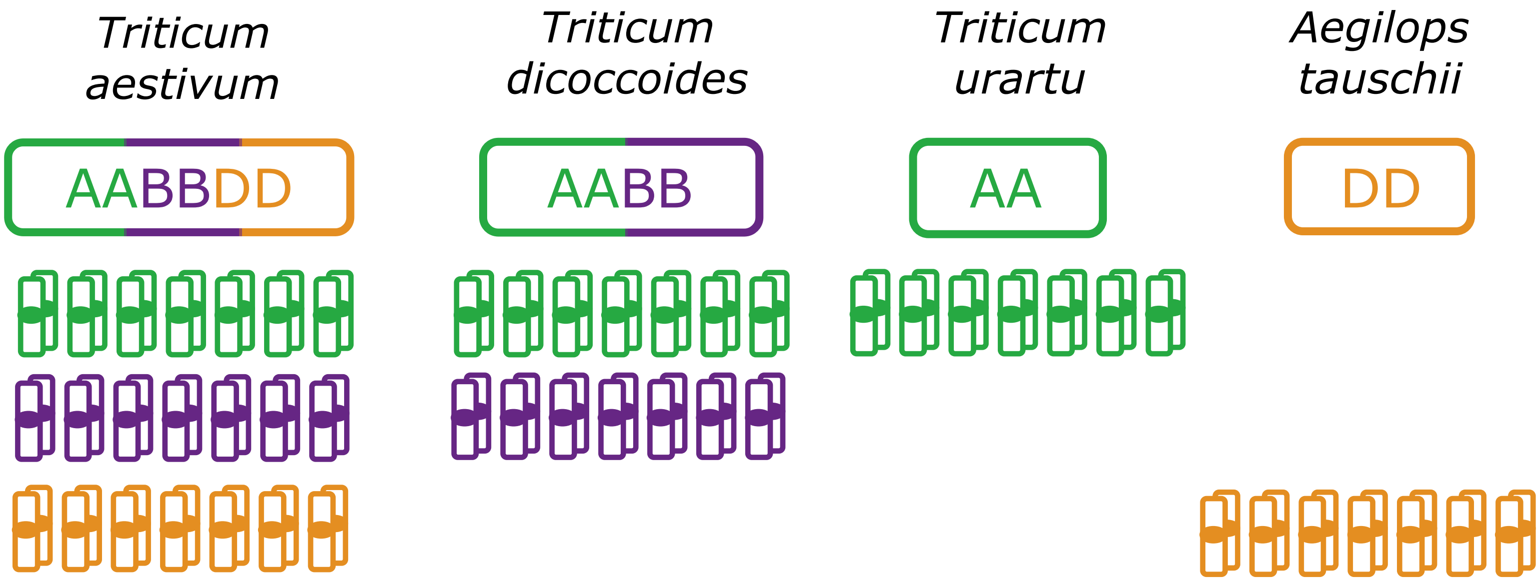
○ TE PAVs – distribution of polymorphisms



5. Genomic variability affecting **TEs** in wheat **using assembled genomes**

PhD Nathan Papon (2019-2022)

- *Methodology for comparative genomics of complex genomes with Gbps of repeats*
- *Extent of genomic variability?*
- *Active families? Amount of transposition events?*
- *Did polyploidization trigger a TE burst?*



Triticum aestivum

AABBDD

Triticum dicoccoides

AABB

Triticum urartu

AA

Aegilops tauschii

DD

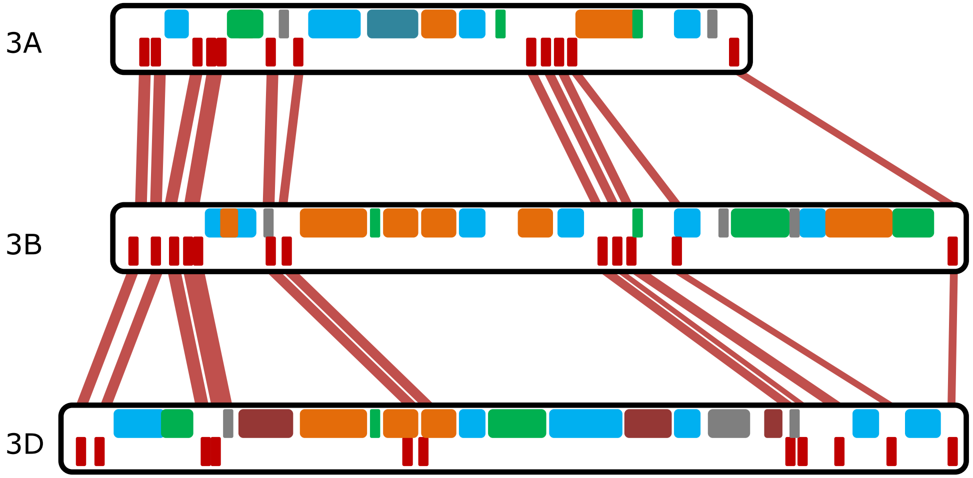
- Multiple RefSeq now available

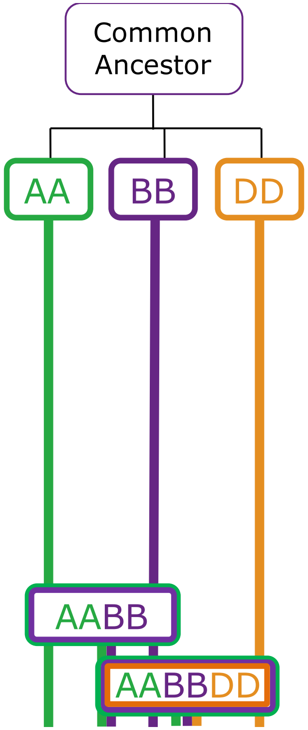
IWGSC 2018

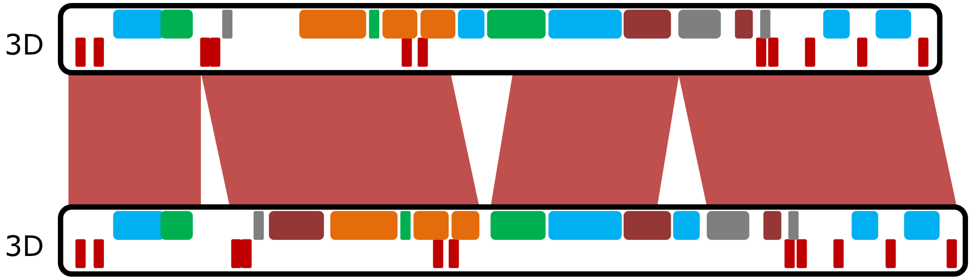
Avni et al. 2017

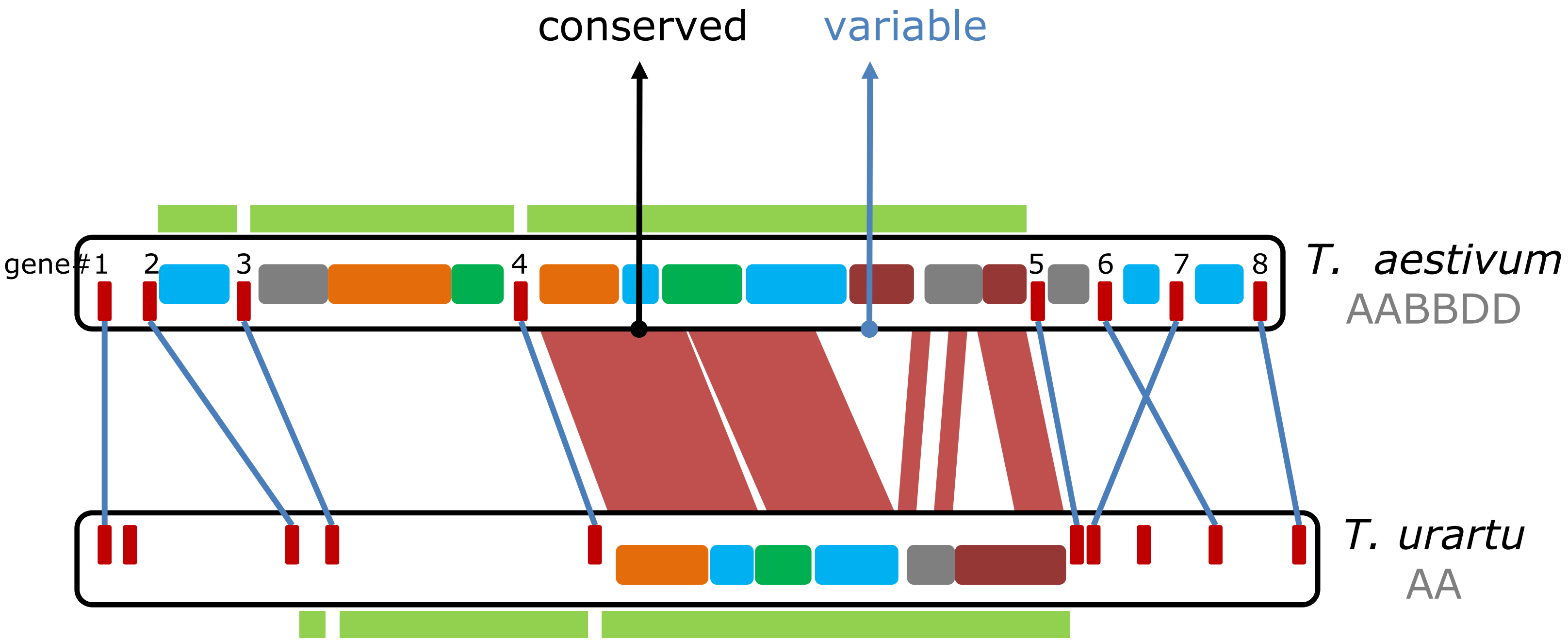
Ling et al. 2018

Luo et al. 2017


3A
3B
3D

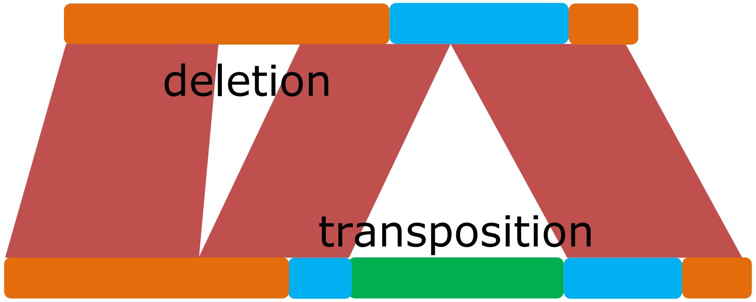

Common Ancestor
AA
BB
DD
AABB
AABBDD


3D
3D



Results:

- ➢ Variable regions represent (pairwise):
  - inter-specific: 8-34% of the sequences
  - intra-specific: 4-6% of the sequences

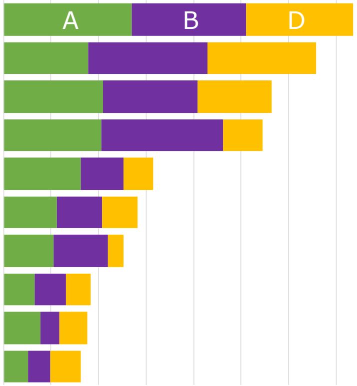


Extent of variability due to recent insertions?

- ➢ ~5k–10k transposition events / subgenome (<1% of the genome)
- ➢ 238 active families

Impact of polyploidy?

- ➢ none



Top10 families	Recent TE insertions *T. aestivum* *AABBDD*	 *T. durum* *AABB*
RLC_famc1	3032	3239
DTC_famc2	177	191
RLG_famc2	118	133
RLG_famc1	259	228
RLG_famc7	521	537
RLG_famc5	67	66
RLG_famc3	2312	2515
RLG_famc4	36	27
RLG_famc11	157	168
RLG_famc10	104	106

- ➢ Equilibrium

6. Sequencing the Renan genome

OXFORD (GIGA)^n SCIENCE

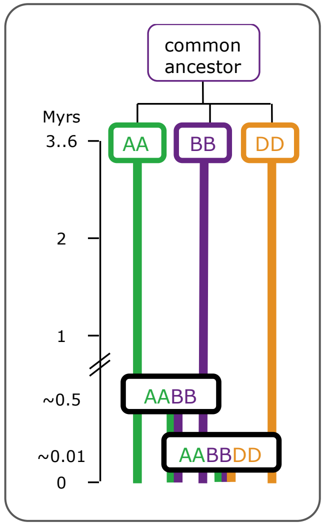
GigaScience, 2022, **11**, 1–18

DOI: 10.1093/gigascience/giac034

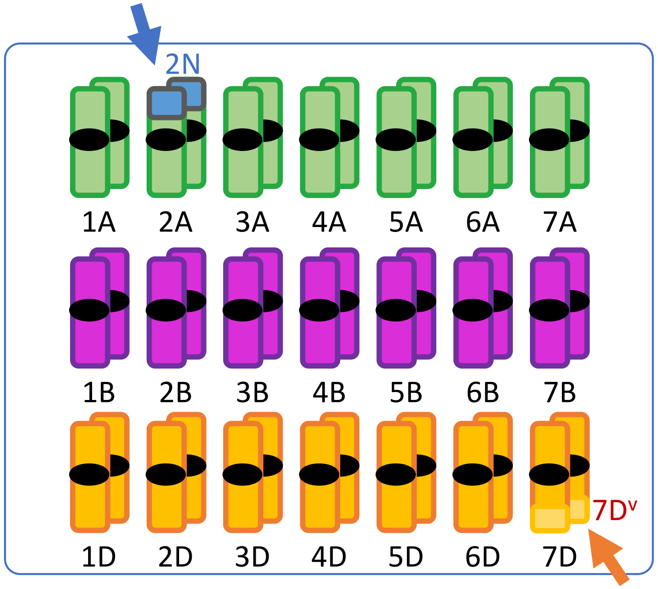
RESEARCH

Long-read and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding

Jean-Marc Aury[1,*], Stefan Engelen[1], Benjamin Istace[1], Cécile Monat[2], Pauline Lasserre-Zuber[2], Caroline Belser[1], Corinne Cruaud[3], Hélène Rimbert[2], Philippe Leroy[2], Sandrine Arribat[4], Isabelle Dufau[4], Arnaud Bellec[4], David Grimbichler[5], Nathan Papon[2], Etienne Paux[2], Marion Ranoux[2], Adriana Alberti[1,6], Patrick Wincker[1] and Frédéric Choulet[2,*]

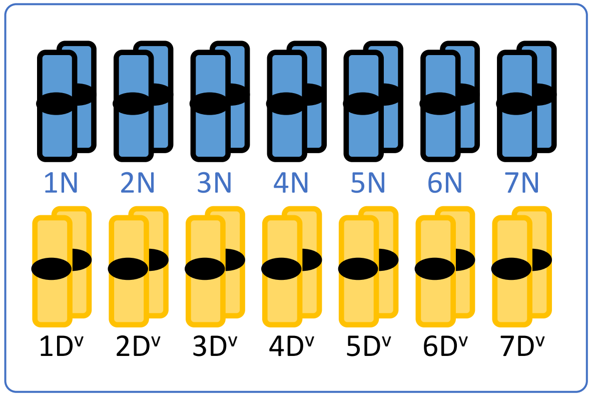


Yr17 Lr37
Sr38 Cre5



Pch1

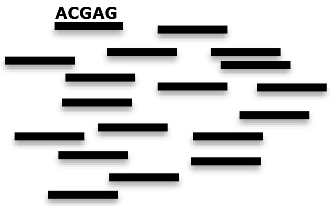
T. aestivum
Renan



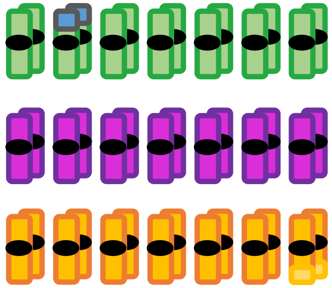
Ae. ventricosa

G. Doussinault, J. Jahier, M. Bernard et al. (1980's 90's 00's)

○ Genome Assembly



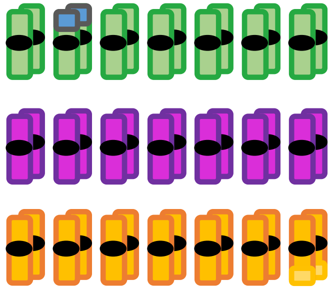
- ❖ Nanopore
- ❖ Illumina
- ❖ Bionano
- ❖ HiC



21 chromosomes

○ Annotation of genes

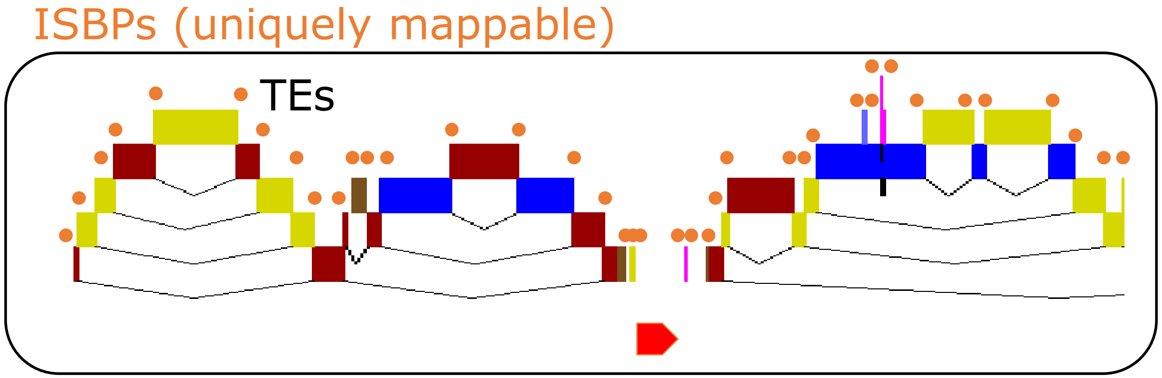
107k HC genes
Chinese Spring

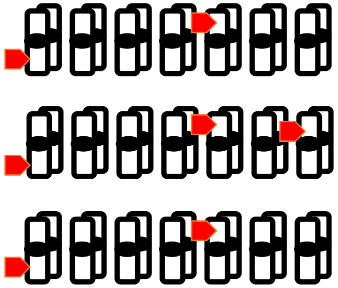


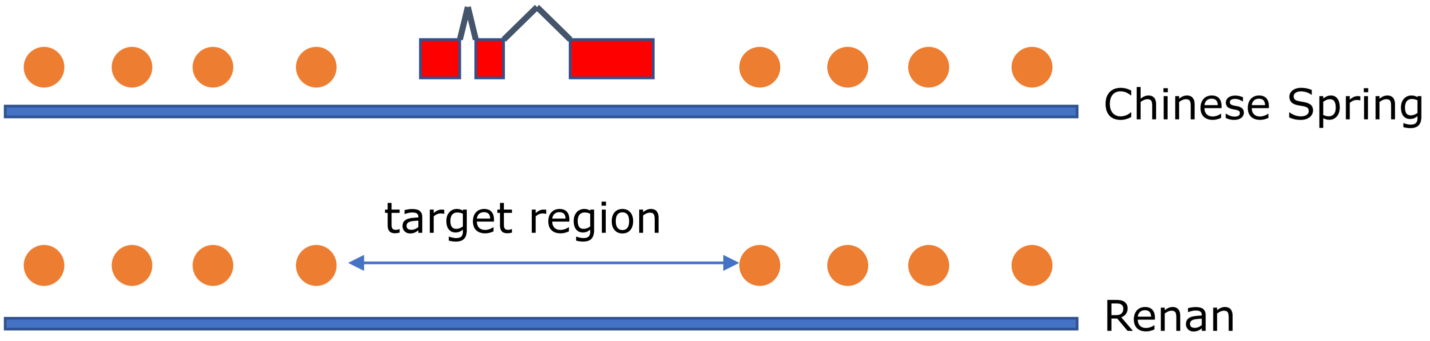
~12% ambiguous mapping

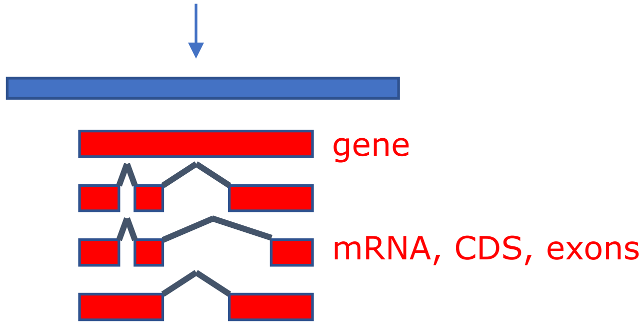
- *denovo* gene annotation would generate to many differences
- gene mapping from a reference is NOT satisfying
  - ➢ Develop **MAGATT**

❑ **MAGATT** (**M**arker-**A**ssisted **G**ene **A**nnotation **T**ransfer for ***T**riticeae*)



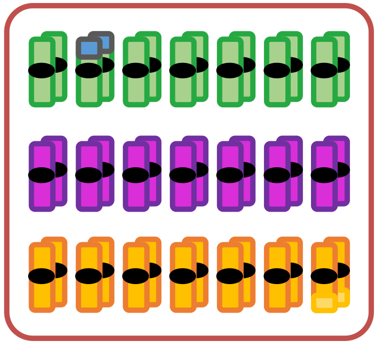


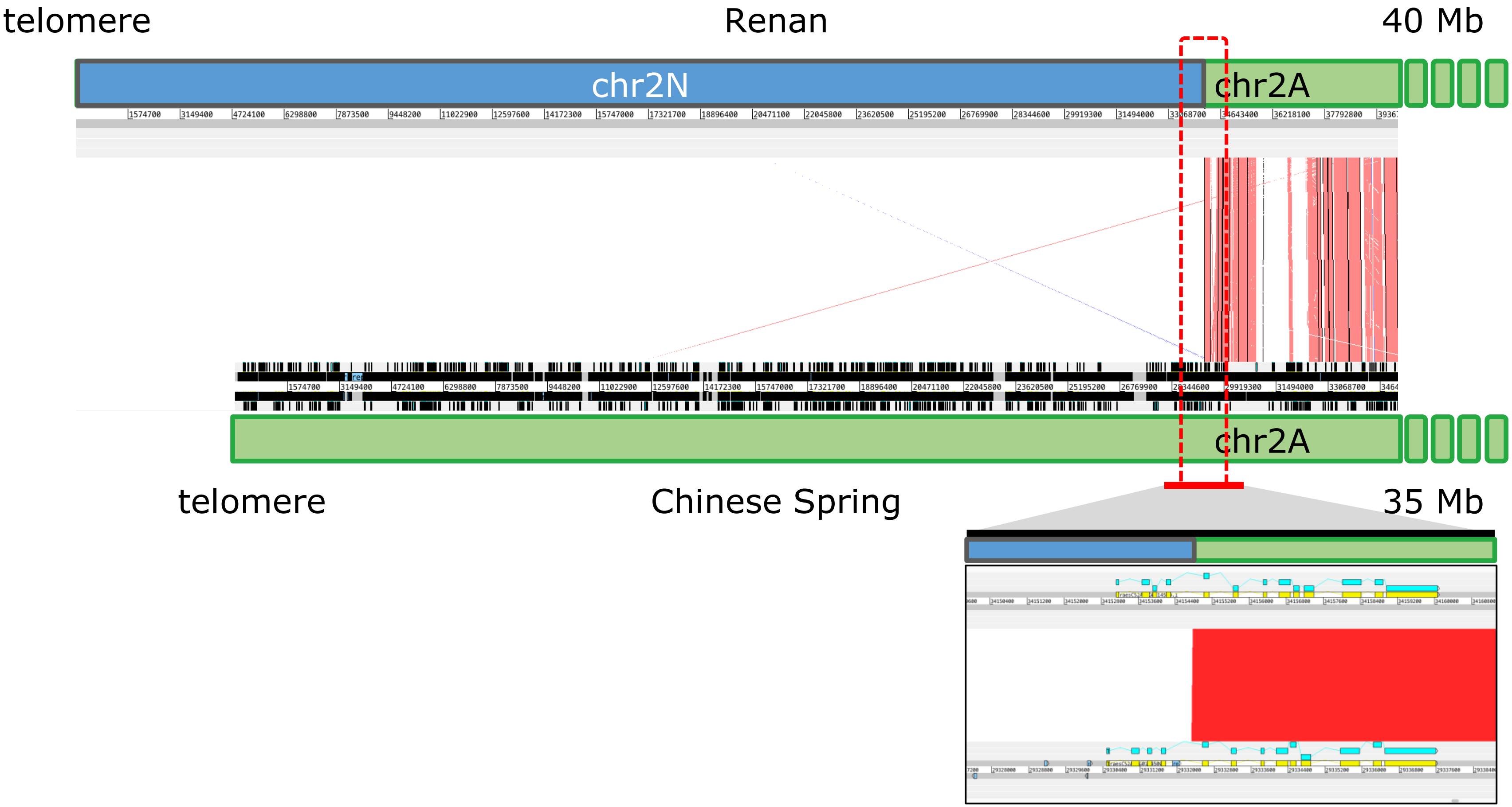


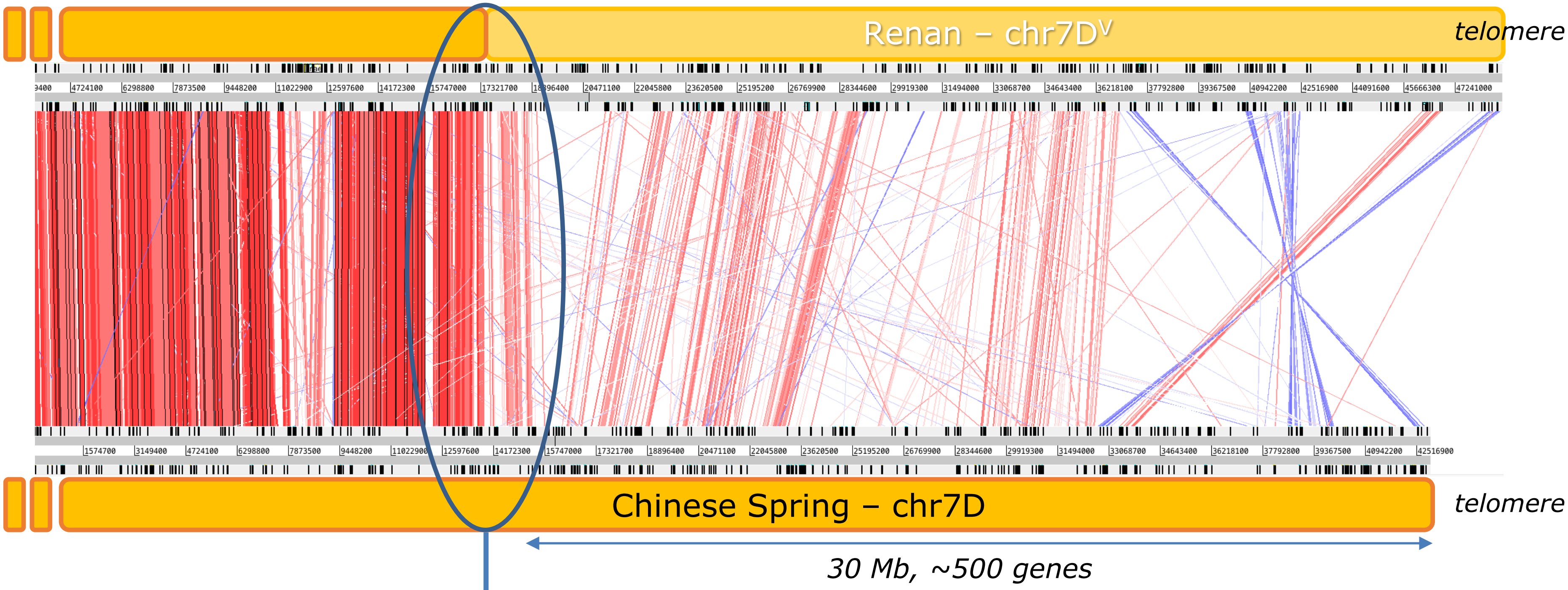


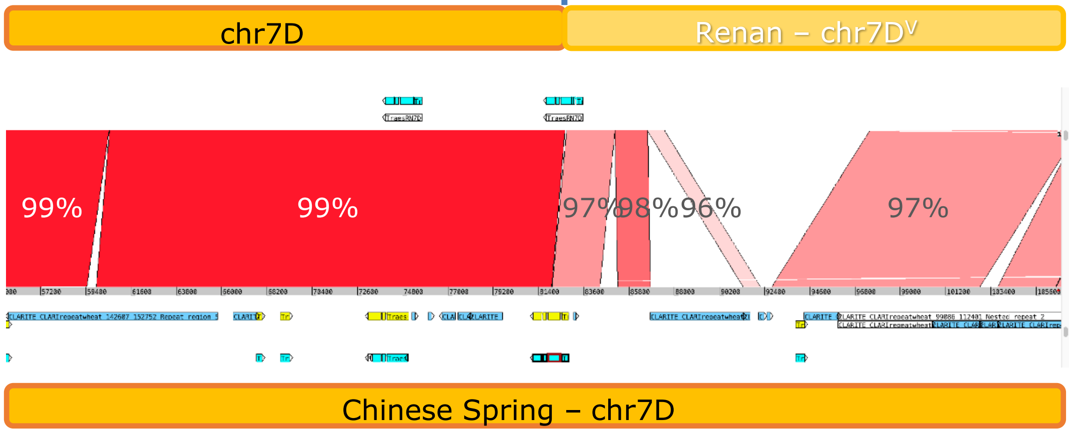
- ➢ 98% of genes mapped accurately
- ➢ + 4400 Renan specific genes

❑ *Introgressions*









- Introgressions: source of diversity in modern breeding lines

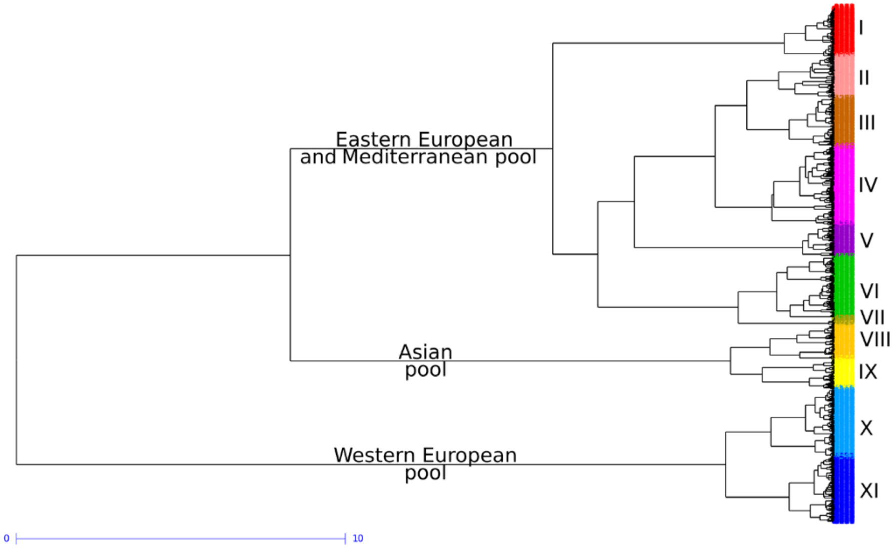
analyses in progress…

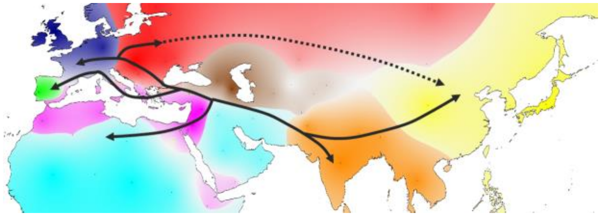
○ Conclusions

- Wheat has reached the pangenomic area
- Annotation matters

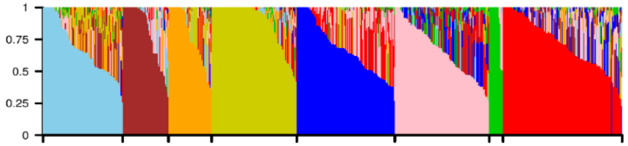
○ Perspectives

- Build a wheat pangenome based on 8-12 founder lines
    - IWGSC/NSF proposal under review
    - GDEC, IPS2





STRUCTURE



Balfourier et al. 2018

❑ Acknowledgments

- *GDEC, INRAE, UCA, Clermont-Fd*

Nathan Papon
Romain De Oliveira
Cécile Monat
Caroline Juery
Josquin Daron
Emeric Dynomant
Hélène Rimbert
Philippe Leroy
Pauline Lasserre-Zuber
Jonathan Kitt
Marion Ranoux
Florence Exbrayat
François Balfourier
Etienne Paux
Pierre Sourdille
Sophie Bouchet

- *Collaborators*

Jean-Marc Aury's team (Genoscope)
Arnaud Bellec (INRAE CNRGV)
Michael Alaux (INRAE URGI)
Thomas Wicker (U Zurich)
Heidrun Gundlach (PGSB Munich)
Manuel Spannagl (PGSB Munich)
Jan Vrana (IEB Olomouc)
Jaroslav Dolezel (IEB Olomouc)
Federica Cattonaro (IGA, Udine)

- *IWGSC*

K. Eversole, et al.