

A lncRNA gene-enriched atlas for GRCg7b chicken genome and its functional annotation across 47 tissues

Fabien Degalez, Mathieu Charles, Sylvain Foissac, Haijuan Zhou, Dailu Guan, Coralie Alain, Lingzhao Fang, Christophe Klopp, Laetitia Lagoutte, Bénédicte Lebez, et al.

▶ To cite this version:

Fabien Degalez, Mathieu Charles, Sylvain Foissac, Haijuan Zhou, Dailu Guan, et al.. A lncRNA gene-enriched atlas for GRCg7b chicken genome and its functional annotation across 47 tissues. 39. Congress of the international society for animal genetics (ISAG), Jul 2023, Cape Town, South Africa., pp.78, 2023, Abstracts. hal-04217582

HAL Id: hal-04217582 https://hal.inrae.fr/hal-04217582v1

Submitted on 14 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A lncRNA GENE-ENRICHED ATLAS FOR THE GRCg7b CHICKEN GENOME AND ITS FUNCTIONAL ANNOTATION ACROSS 47 TISSUES

Fabien Degalez, Sandrine Lagarrigue

▶ To cite this version:

Fabien Degalez, Sandrine Lagarrigue. A lncRNA GENE-ENRICHED ATLAS FOR THE GRCg7b CHICKEN GENOME AND ITS FUNCTIONAL ANNOTATION ACROSS 47 TISSUES. ISAG 2023 - 39th International Society for Animal Genetics Conference (2023), Jul 2023, CAPE TOWN, South Africa. hal-04342390

HAL Id: hal-04342390

https://hal.science/hal-04342390

Submitted on 13 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A Incrna Gene-Enriched atlas for the Grcg7b Chicken Genome and its functional annotation across 47 tissues

fabien.degalez@inrae.fr

F. Degalez¹ & S. Lagarrigue¹

sandrine.lagarrigue@institut-agro.fr

¹PEGASE, INRAE, Institut Agro, 35590 Saint Gilles, France

CONTEXT

While protein-coding genes (PCG) are relatively well known in livestock genomes, long non-coding RNA (IncRNA) gene loci, known to be regulatory genes, are more poorly described and differ greatly between reference genome annotations. Even if new expression data contribute to improve their identification, their low expression and high context-specificity remain a challenge.

For the chicken, in 2022, the new GRCg7b chicken genome assembly with its associated genome annotations have been released.

OBJECTIVES

- Provide an annotation of the chicken genome according to the GRCg7b assembly.
- Integrate the two reference annotations including "NCBI-RefSeq" (Ncbi) and "EMBL-EBI Ensembl/GENCODE" (Ensembl).
- Ease the joint use of gene models from Ncbi & Ensembl and the switch between the galgal5, GRCg6a and GRCg7b chicken assemblies.
- Increase the number of IncRNA identified using additional resources from multi-tissue projects or specialist databases.
- Provide a genomic and functional annotation for the community working on gene expression (IncRNAs and/or PCGs) to elucidate, for example, the molecular mechanisms linking non-coding variants and relevant phenotypes.

MATERIALS & METHODS Fig. 1: Origin and aggregation rules used to build the gene-enriched annotation Fig. 2: A diversity of 47 tissues used for gene expression Endocrine system (5/170) Successive order for gene models gathering Nervous system (7/184) Reference with non-overlapping transcripts – *i.e.,* no exon genome has base pairs in common on the same strand. Lymphatic system (3/104) annotations **C.** Ensembl Ensembl Ncbi Circulation (1/84) Ensembl Respiration (2/51) Fr-AgENCODE **FAANG** Fr-AgENCODE 3 Skin system (3/114) FAANG pilot **UC Davis** Muscles (2/50) projects **UC Davis** exon ovlp. **INRAe** Reproduction (5/107) No ovlp. Noncode Digestion (10/368) INRAC **INRAe** Embryo (2/40) **IncRNAs & PCGs** Additional Cell populations (7/128) gene-enriched resources **ENONCODE** Noncode atlas

RESULTS

Fig. 3: Number of genes in the enriched atlas

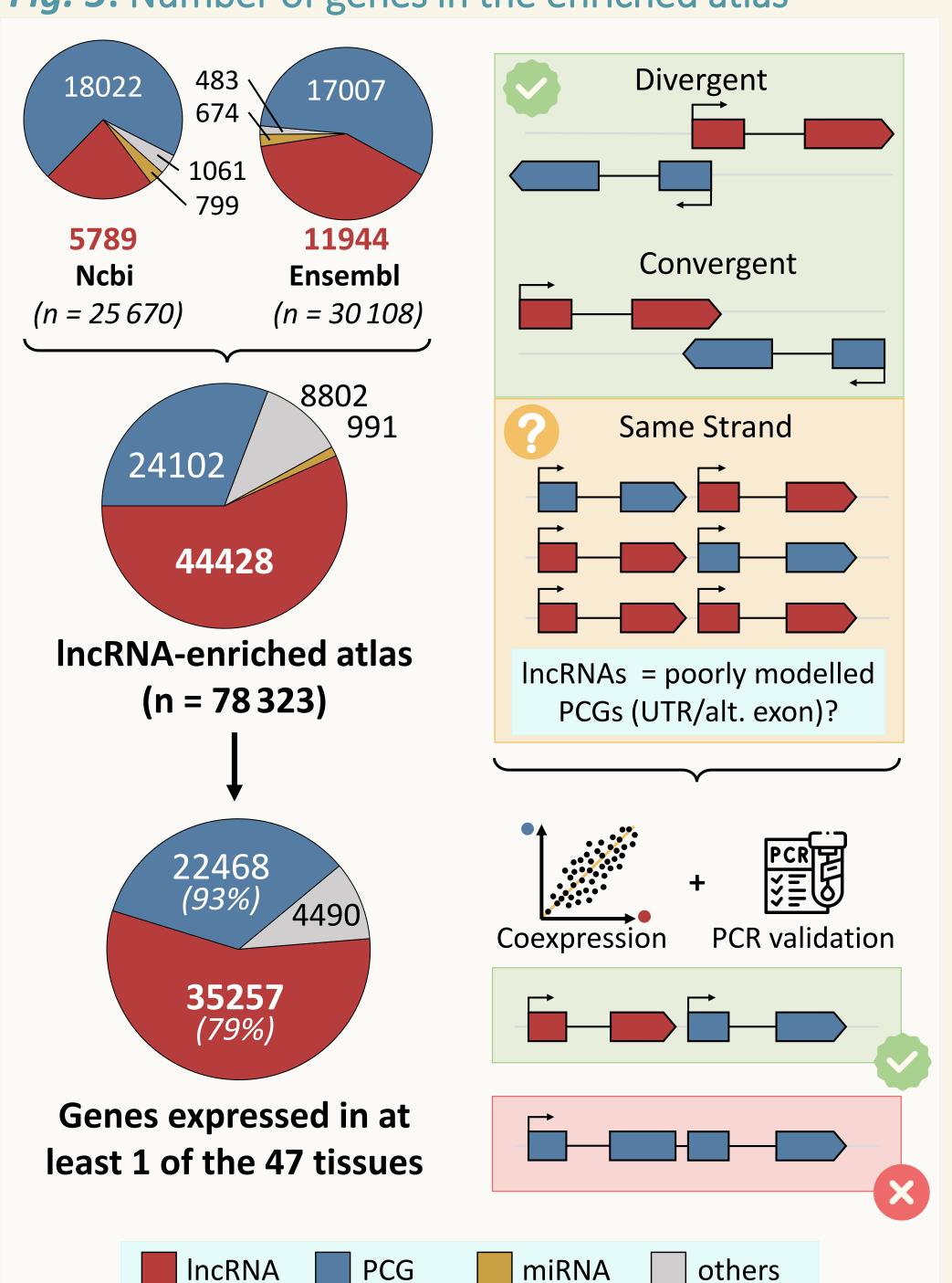
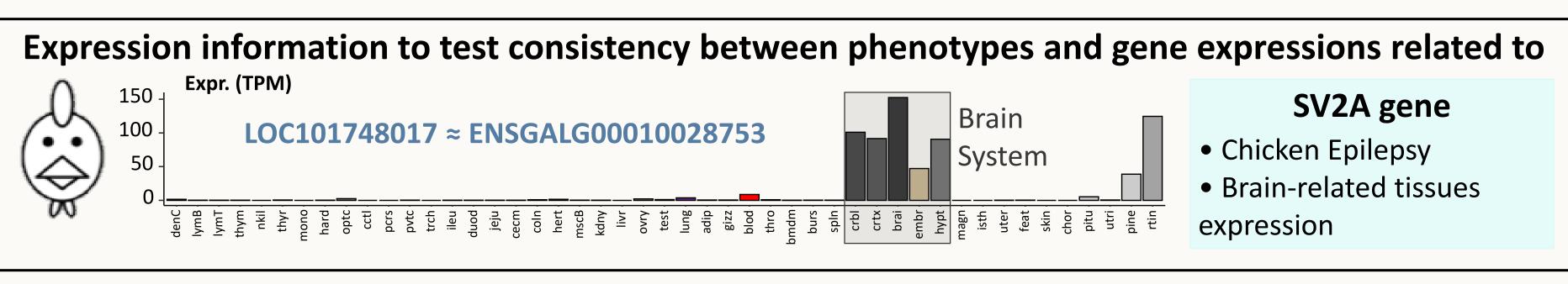
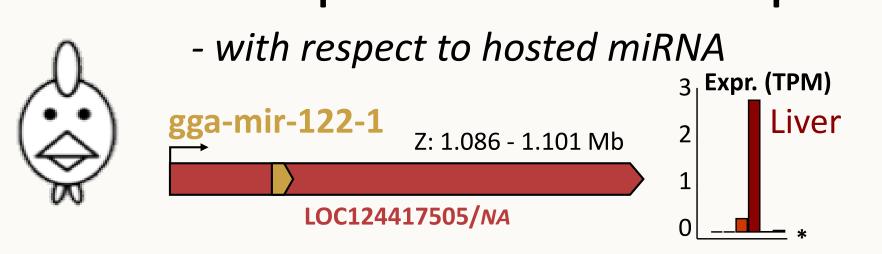


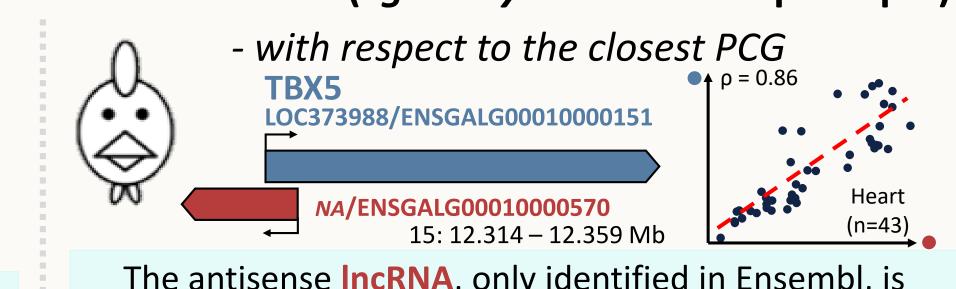
Fig. 4: Functional gene annotation, a resource for a variety of analyses



Genomic and expression information to predict IncRNA function ("guilt by association" principle)



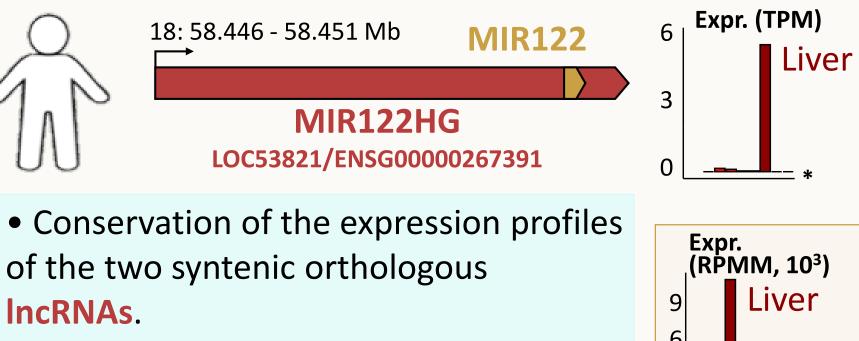
The IncRNA, only identified in Ncbi, is liver-specific and hosts mir-122 involved in liver function [1].



(47 tissues / 1400 samples) from 36 datasets

The antisense IncRNA, only identified in Ensembl, is co-expressed across tissues and especially in the heart with TBX5 associated to heart development [2].

Orthology information to extract genomic and expression information in other species



IncRNAs.

• The miRNA hosted in the IncRNA also shows liver-specific expression in human.

10³)
er
R122

TBX5 12: 114.354 – 114.413 Mb
LOC6910/ENSG00000089225

TBX5-AS1
LOC27402/ENSG00000255399

- Conservation of the expression profiles for both syntenic IncRNA and PCG.
- In human, TBX5-AS1 regulated TBX5 expression levels by forming an RNA duplex that protects the mRNA from RNase degradation [3].

*Tissues for which the gene is not/lowly expressed are not shown.

Occurrence of the gene loci in the 6 resources Gene model features Functional annotation (Go terms / MGI) Gene pair configuration (PCG / IncRNA / miRNA) Orthology (human / mouse) GEGA

WEBSITE

To help in performing studies such as QTL, eQTL or differentially expressed genes analyses, all these data are easily accessible via a user-friendly online tool called **GEGA (Gallus Enriched Gene Annotation)** enabling efficient data analysis thanks to a multitude of modules and an ergonomic design. All results and a multitude of graphical representation can be simply produced and downloaded, ranging from basic gene information to detailed visualization of expression and co-expression.

gega.sigenae.org

Acknowledgments: This project is funded by the European Union's Horizon 2020 research and innovation program under grant agreement N°101000236 and by ANR CE20 under 'EFFICACE' program.

References: [1] Bandiera S, et al. miR-122--a key factor and therapeutic target in liver disease. J Hepatol. 2015;62:448–57; [2] Steimle JD, Moskowitz IP. TBX5: A Key Regulator of Heart Development. Current Topics in Developmental Biology. Academic Press; 2017. p. 195–221. [3] Ma J, et al. Hypermethylation-mediated down-regulation of lncRNA TBX5-AS1:2 in Tetralogy of Fallot inhibits cell proliferation by reducing TBX5 expression. Journal of Cellular and Molecular Medicine. 2020;24:6472–84.



