

Newly annotated genomic features for biology-driven genome selection: the BovReg contribution



Gabriel Costa Monteiro Moreira, Lijing Tang, Sébastien Dupont, Meenu Bhati, Hubert Pausch, Doreen Becker, Mazdak Salavati, Richard Clark, Emily L. Clark, Graham Plastow, Christa Kühn and Carole Charlier on behalf of the BovReg consortium



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668

Disclaimer: the sole responsibility of this presentation lies with the authors. The Research Executive Agency is not responsible for any use that may be made of the information contained therein.

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

- BovReg consortium brings together scientists from twenty leading laboratories from different countries (EU, Switzerland, international organizations, Canada and Australia)
- Aims to improve our understanding of the cattle genome; also provide a comprehensive map of functionally active genomic features in cattle
- Critical for biological discovery and in linking genotype to phenotype; key knowledge for biology-driven genomic prediction needed by scientific and industry livestock communities.



www.bovreg.eu



@BovReg.H2020



@BovReg



BovReg



BovReg project

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

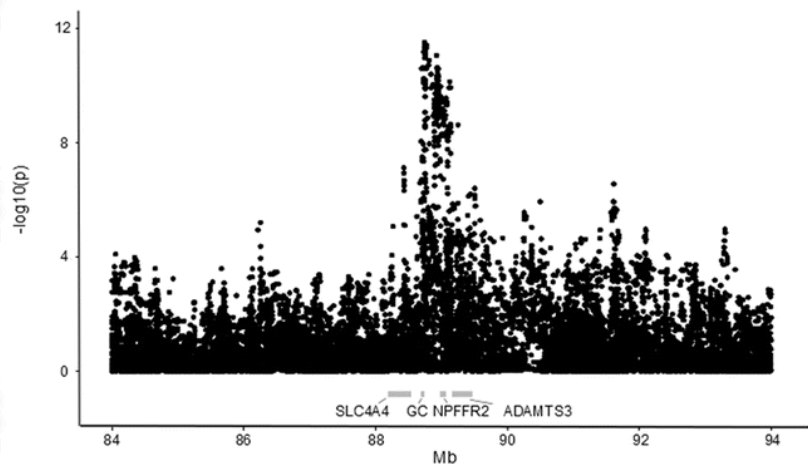
➤ **Epigenome maps enable functional follow-up of GWAS hits.¹**

PLOS GENETICS

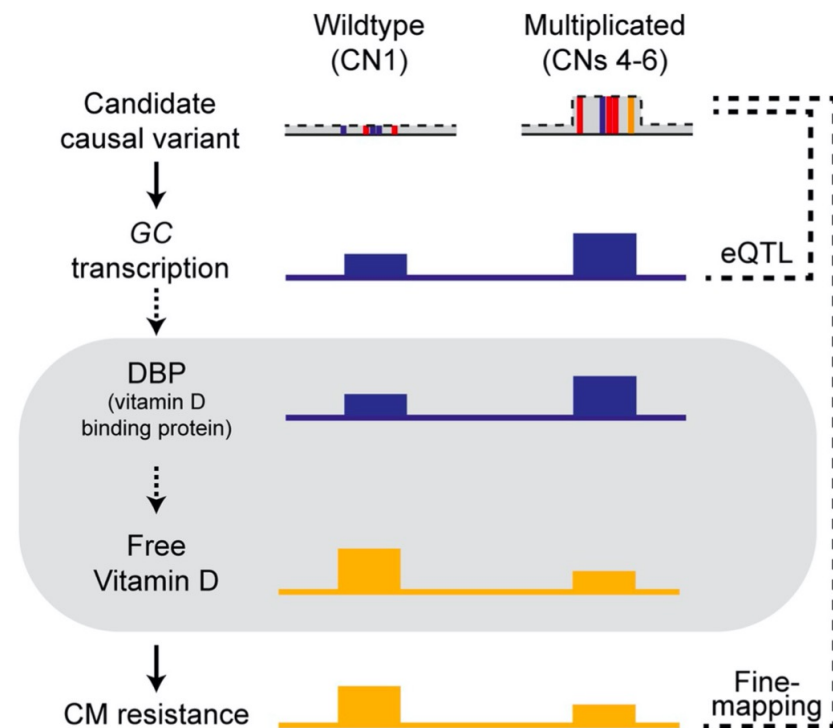
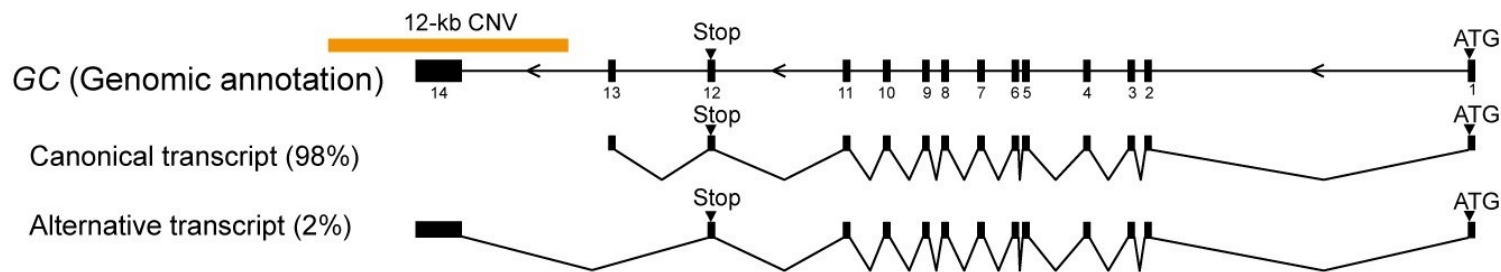
RESEARCH ARTICLE

A 12 kb multi-allelic copy number variation encompassing a GC gene enhancer is associated with mastitis resistance in dairy cattle

Young-Lim Lee^{1*}, Haruko Takeda², Gabriel Costa Monteiro Moreira², Latifa Karim³, Erik Mullaart⁴, Wouter Coppieters^{2,3}, The GplusE consortium⁵, Ruth Appeltant², Roel F. Veerkamp¹, Martien A. M. Groenen¹, Michel Georges², Mirte Bosse¹, Tom Druet², Aniek C. Bouwman¹, Carole Charlier²



Enhancer!
(ChIP-Seq data)



¹Georges, M., Charlier, C. & Hayes, B. Harnessing genomic information for livestock improvement. *Nat Rev Genet* **20**, 135–156 (2019). <https://doi.org/10.1038/s41576-018-0082-2>

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

➤ **Recent efforts from the FAANG consortium have contributed to the functional annotation of the bovine genome but, with a limited number of tissues, individuals and populations analyzed**

Chinese Simmental cattle



Zhang et al. *BMC Biology* (2022) 20:79
<https://doi.org/10.1186/s12915-022-01269-4> **BMC Biology**

RESEARCH ARTICLE **Open Access**

Transcriptional atlas analysis from multiple tissues reveals the expression specificity patterns in beef cattle

Tianliu Zhang^{1†}, Tianzhen Wang^{1†}, Qunhao Niu¹, Lei Xu¹, Yan Chen¹, Xue Gao¹, Huijiang Gao¹, Lupei Zhang¹, George E. Liu², Junya Li^{1*} and Lingyang Xu^{1*}

- 3 animals; same age
- 51 tissue types
- Only PolyA+ data

Line 1 Herefords



nature COMMUNICATIONS

Functional annotations of three domestic animal genomes provide vital resources for comparative and agricultural research

Colin Kern¹, Ying Wang¹, Xiaoqin Xu¹, Zhangyuan Pan¹, Michelle Halstead¹, Ganrea Chanthavixay¹, Perot Saelao¹, Susan Waters¹, Ruidong Xiang^{2,3}, Amanda Chamberlain³, Ian Kort⁴, Mary E. Delany¹, Hans H. Cheng⁵, Juan F. Medrano¹, Alison L. Van Eenennaam¹, Chris K. Tuggle⁶, Catherine Ernst⁷, Paul Flicek⁸, Gerald Quon⁹, Pablo Ross¹⁰ & Huaijun Zhou¹⁰

- 8 tissues from 2 Herefords male cattle (14 months old)

Holstein



Foissac et al. *BMC Biology* (2019) 17:108
<https://doi.org/10.1186/s12915-019-0726-5> **BMC Biology**

RESEARCH ARTICLE **Open Access**

Multi-species annotation of transcriptome and chromatin structure in domesticated animals

Sylvain Foissac^{1†*}, Sarah Djebali^{1†}, Kylie Munyard², Nathalie Vialaneix³, Andrea Rau⁴, Kevin Muret⁵, Diane Esquerré^{1,6}, Matthias Zytnicki³, Thomas Derrien⁷, Philippe Bardou¹, Fany Blanc¹, Cédric Cabau¹, Elisa Crisci^{1,10}, Sophie Dhome-Pollet⁴, Françoise Drouet⁸, Thomas Faraut¹, Ignacio Gonzalez², Adeline Goubil⁴, Sonia Lacroix-Lamandé⁸, Fabrice Laurent⁸, Sylvain Marthey¹, Maria Marti-Marimon¹, Raphaëlle Momal-Leisenring⁴, Florence Mompant¹, Pascale Quérel⁸, David Robelin¹, Magali San Cristobal¹, Gwenola Tosser-Klopp¹, Silvia Vincent-Naulleau⁹, Stéphane Fabre¹, Marie-Hélène Pinard-Van der Laan⁴, Christophe Klopp³, Michèle Tixier-Boichard⁴, Hervé Aclouque^{1,4}, Sandrine Lagarrigue⁹ and Elisabetta Giuffra⁴

- Only Liver. CD4 and CD8 cells
- 4 adult animals (Holstein)
- PolyA+. ATAC and Hi-C

Line 1 Herefords



bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.05.510963>; this version posted October 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Functional genomics of cattle through integration of multi-omics data

Hamid Beiki¹, Brenda M. Murdoch², Carissa A. Park¹, Chandlar Kern³, Denise Kontechy²,

Gabrielle Becker², Gonzalo Rincon⁴, Honglin Jiang⁵, Huaijun Zhou⁶, Jacob Thorne², James E.

Koltes¹, Jennifer J. Michal⁷, Kimberly Davenport², Monique Rijnkels⁸, Pablo J. Ross⁶, Rui Hu⁵,

Sarah Corum⁴, Stephanie McKay⁹, Timothy P.L. Smith¹⁰, Wansheng Liu³, Wenzhi Ma³, Xiaohui

Zhang⁷, Xiaoqing Xu⁶, Xuelei Han⁷, Zhihua Jiang⁷, Zhi-Liang Hu¹, James M. Reedy¹

- 47 tissues from 4 Herefords (male and female) (14 months old)

nature genetics

Check for updates

comment

The missing diversity in human epigenomic studies

Charles E. Breeze, Stephan Beck, Sonja I. Berndt and Nora Franceschini

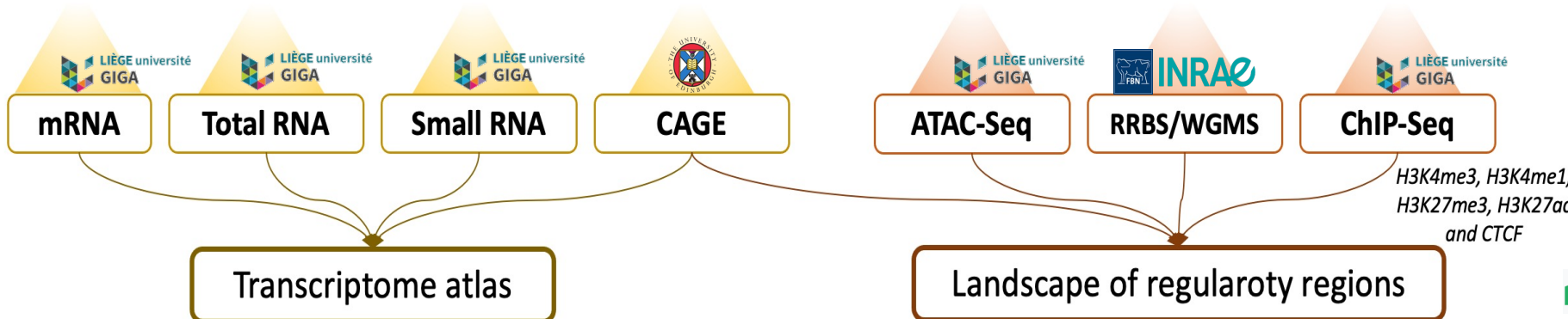
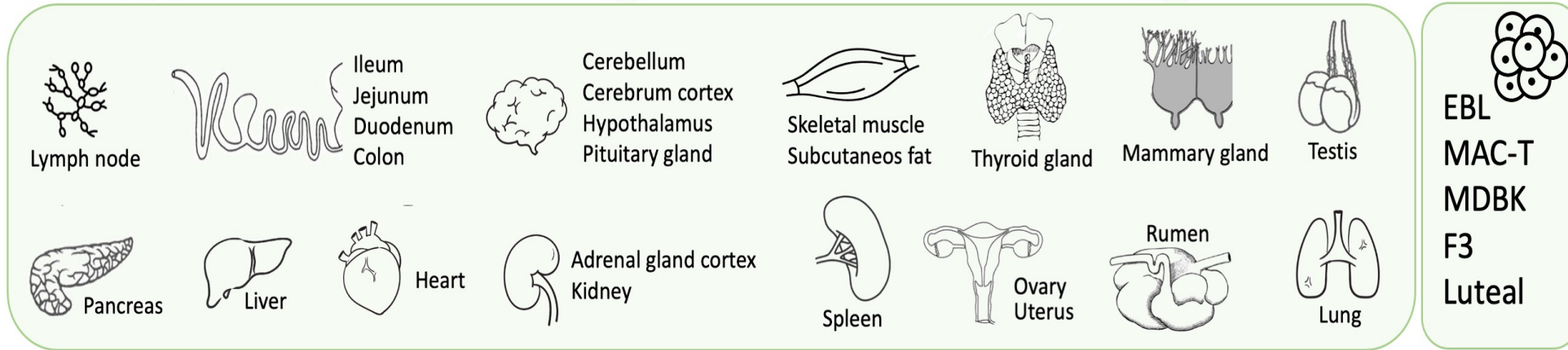
- Several assays performed in the same animal/tissue
- Different developmental stages
- Different environments
- Both sexes
- Different breeds



BovReg
 Understanding cattle genomes

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

➤ **BovReg** utilizes a diverse catalogue of tissues for functional annotation.



Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

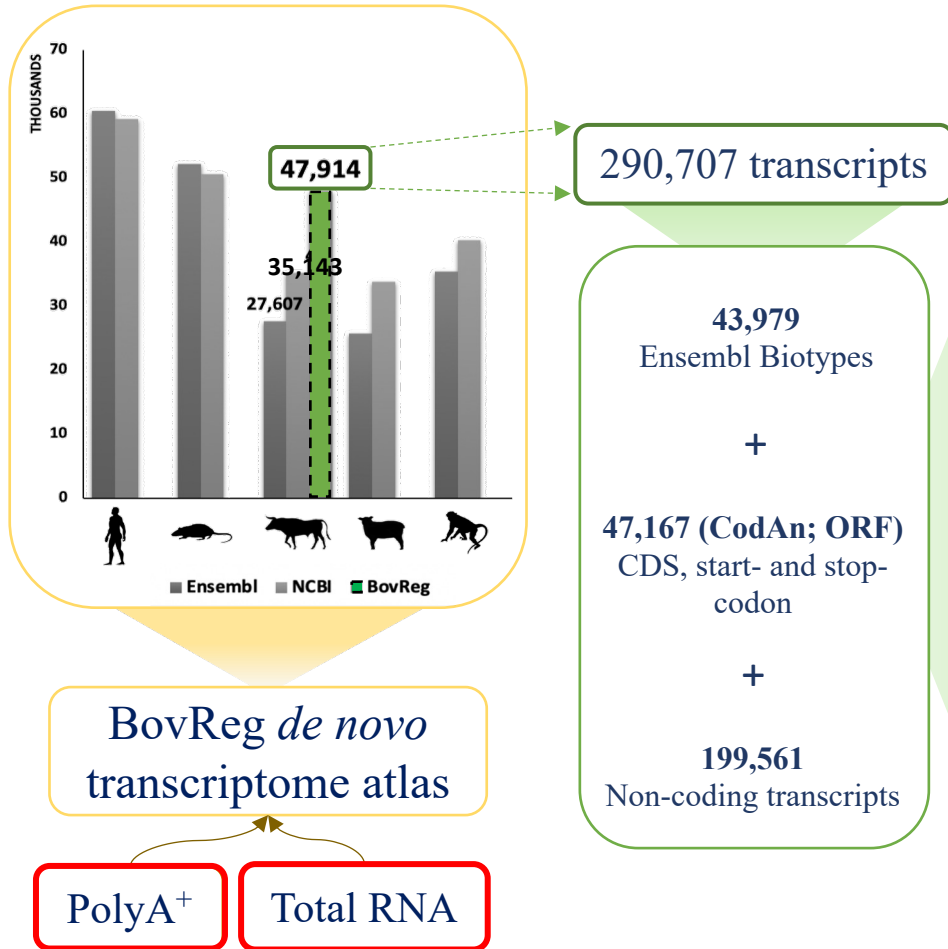


Transcriptome atlas



Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

- **BovReg expanded the catalog of bovine non-coding RNAs by including non-polyadenylated transcripts.**
- **48k genes models including $\geq 15k$ potentially novel transcripts!**



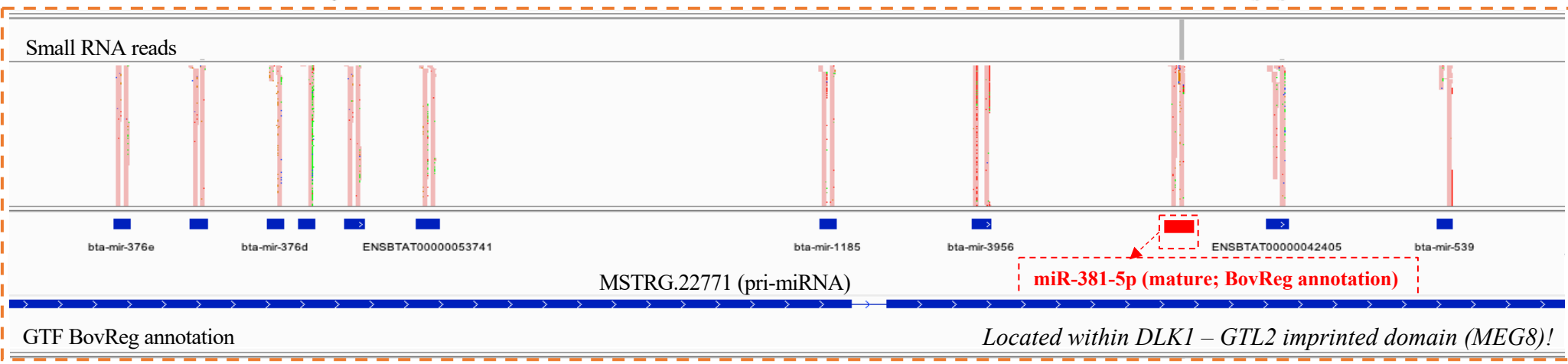
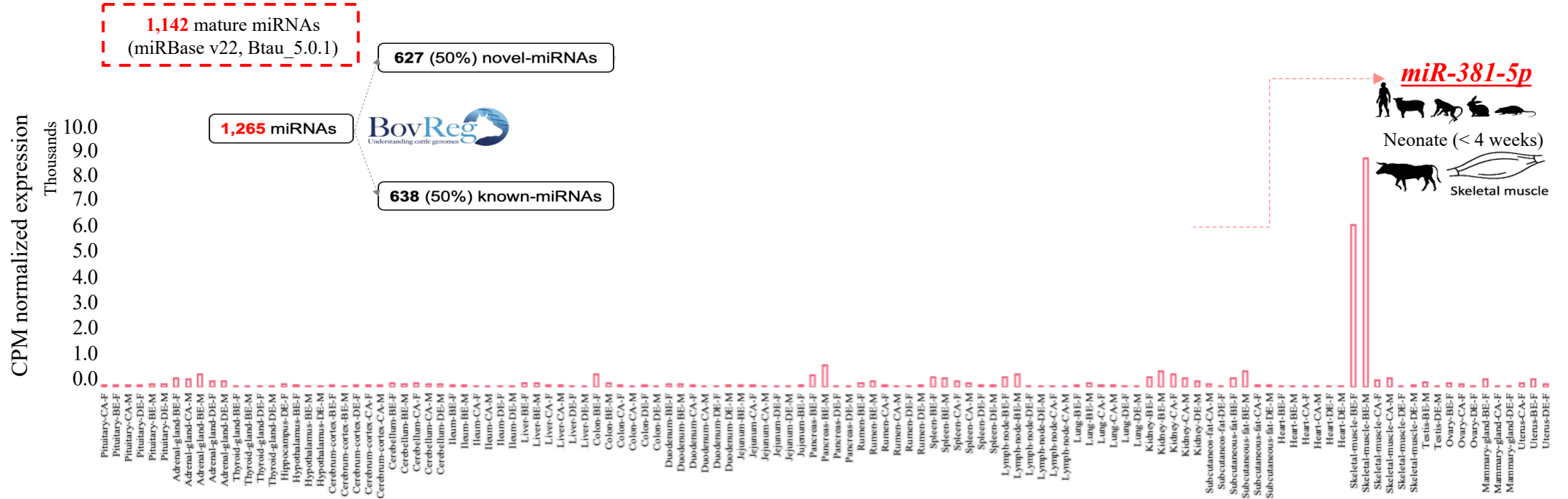
Class code ¹	Features ¹	Compared with bovine annotation	
		Ensembl r.105	NCBI r.106
=	complete transcripts with exact match of intron chain	42873 (14.75%)	36660 (12.61%)
c	transcripts contained in the reference (intron compatible)	4 (0.00%)	3082 (1.06%)
k	containment of reference (reverse containment)	25403 (8.74%)	24814 (8.54%)
m	transcripts with retained introns (all introns matched or retained)	4555 (1.57%)	6321 (2.17%)
n	transcripts with retained introns (not all introns matched or retained)	13812 (4.75%)	14760 (5.08%)
j	multi-exon transcripts with at least one junction match	142987 (49.19%)	150476 (51.76%)
e	single exon transcript partially covering an intron (possible pre-mRNA fragment)	0 (0.00%)	170 (0.06%)
o	other same strand transcripts overlap with reference exons	6422 (2.21%)	9458 (3.25%)
s	transcripts with intron match on the opposite strand (likely a mapping error)	63 (0.02%)	81 (0.03%)
x	exonic transcripts overlap on the opposite strand (like o or e but on the opposite strand)	10569 (3.64%)	8817 (3.03%)
i	transcripts fully contained within a reference intron	10961 (3.77%)	12419 (4.27%)
y	transcripts with a reference transcript within its intron	1871 (0.64%)	999 (0.34%)
p	possible polymerase run-on (no actual overlap)	357 (0.12%)	342 (0.12%)
r	repeat (at least 50% bases soft-masked)	0 (0.00%)	0 (0.00%)
u	none of the above (unknown, intergenic)	30830 (10.61%)	22308 (7.67%)
Total number of transcripts			290,707

¹ Transcript classification code from gffcompare v.0.12.2¹⁰.

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

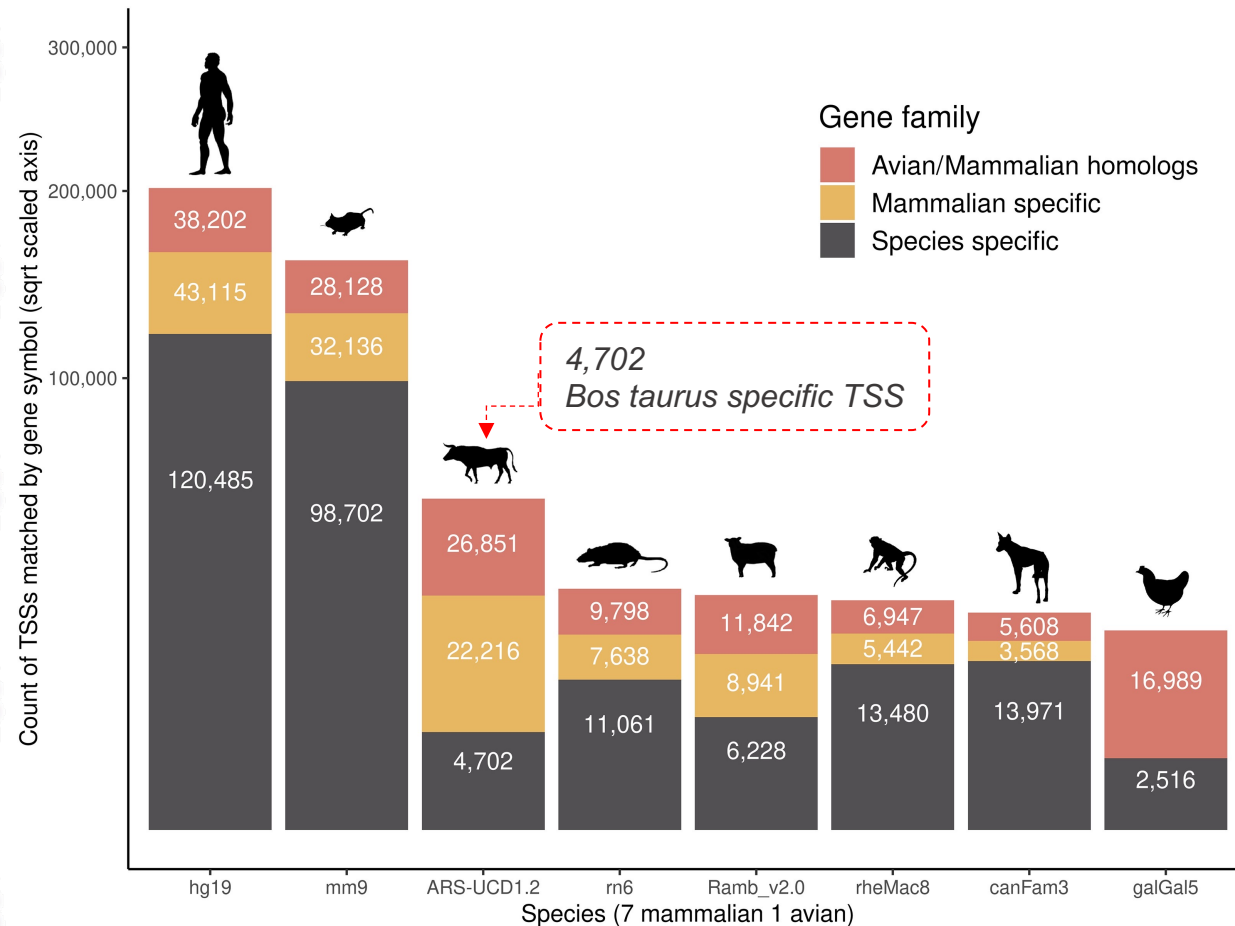
➤ **BovReg** expanded the repertoire of mature miRNAs annotated in bovine.

Tissue-specific miRNA



Newly annotated genomic features for biology-driven genome selection: the BovReg contribution

- 51,295 transcription start sites (TSS)
- 2,328 TSS-Enhancer regions shared across the three populations

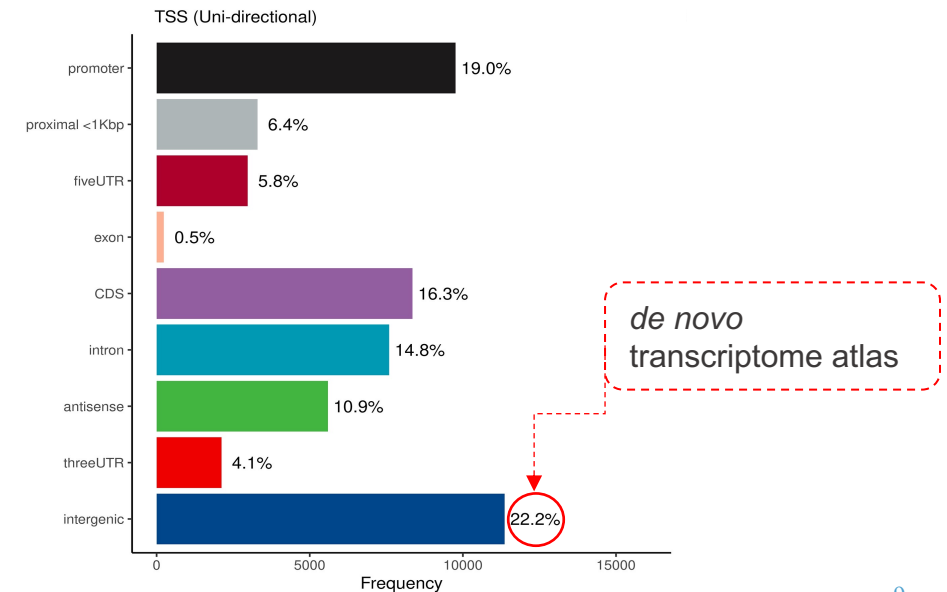


Improving the annotation of the cattle genome by annotating transcription start sites in a diverse set of tissues and populations using Cap Analysis Gene Expression sequencing

Mazdak Salavati^{1,*}, Richard Clark,² Doreen Becker,³ Christa Kühn,^{3,4} Graham Plastow,⁵ Sébastien Dupont,⁶ Gabriel Costa Monteiro Moreira,⁶ Carole Charlier,^{6,7} Emily Louise Clark,¹ on behalf of the BovReg consortium

¹The Roslin Institute, University of Edinburgh, Edinburgh EH25 9RG, UK
²Edinburgh Clinical Research Facility, Genetics Core, University of Edinburgh, Edinburgh EH4 2XU, UK
³Institute of Genome Biology, Research Institute for Farm Animal Biology (FBN), Dummerstorf 18196, Germany
⁴Faculty of Agricultural and Environmental Sciences, University Rostock, Rostock 18059, Germany
⁵Department of Agricultural, Food and Nutritional Science, Livestock Genetic, University of Alberta, Edmonton T6G 2H1, Canada
⁶Unit of Animal Genomics, GIGA Institute, University of Liège, Liège 4000, Belgium
⁷Faculty of Veterinary Medicine, University of Liège, Liège 4000, Belgium

*Corresponding author: The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh EH25 9RG, UK. Email: Mazdak.Salavati@roslin.ed.ac.uk
¹Present address: Dairy Research and Innovation Centre, Scotland's Rural College (SRUC), The Barony Campus, Dumfries DG1 3NE, UK.



Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

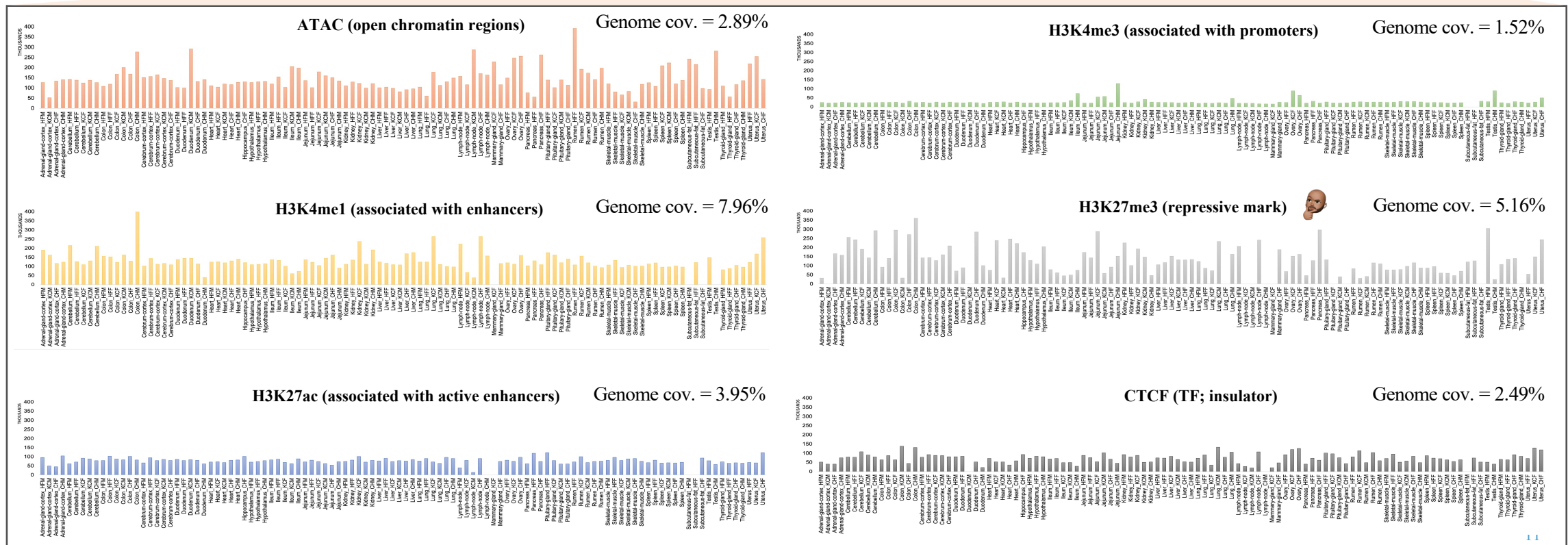
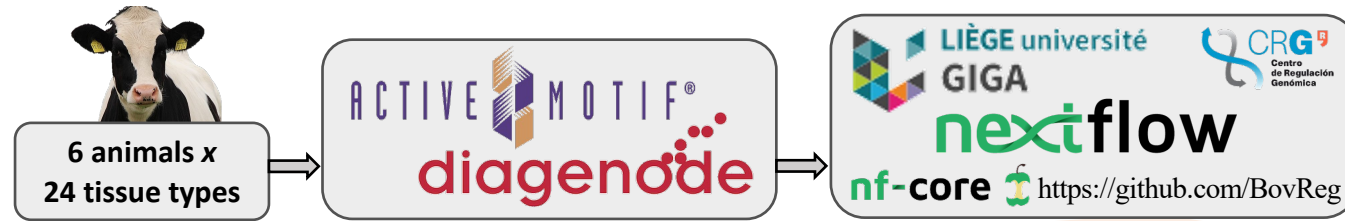


Regulatory regions



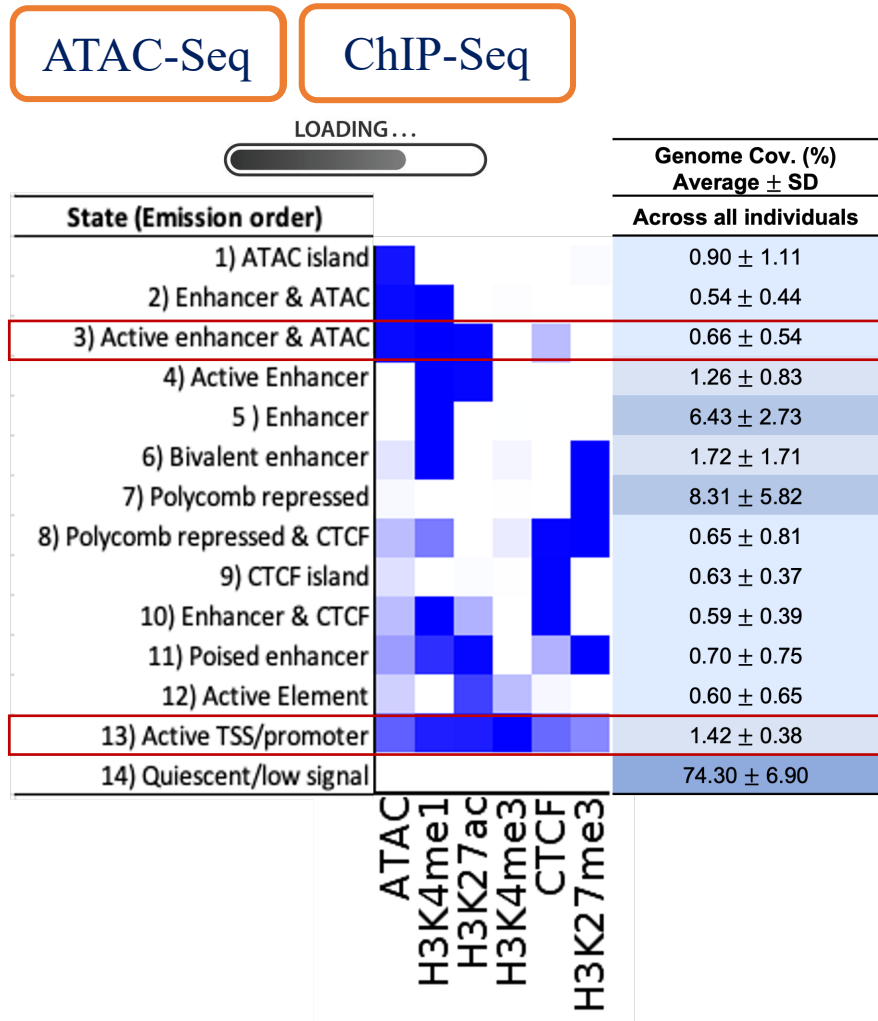
Newly annotated genomic features for biology-driven genome selection: the BovReg contribution

➤ We identified thousands of peaks for the different marks in our catalogue of tissues!

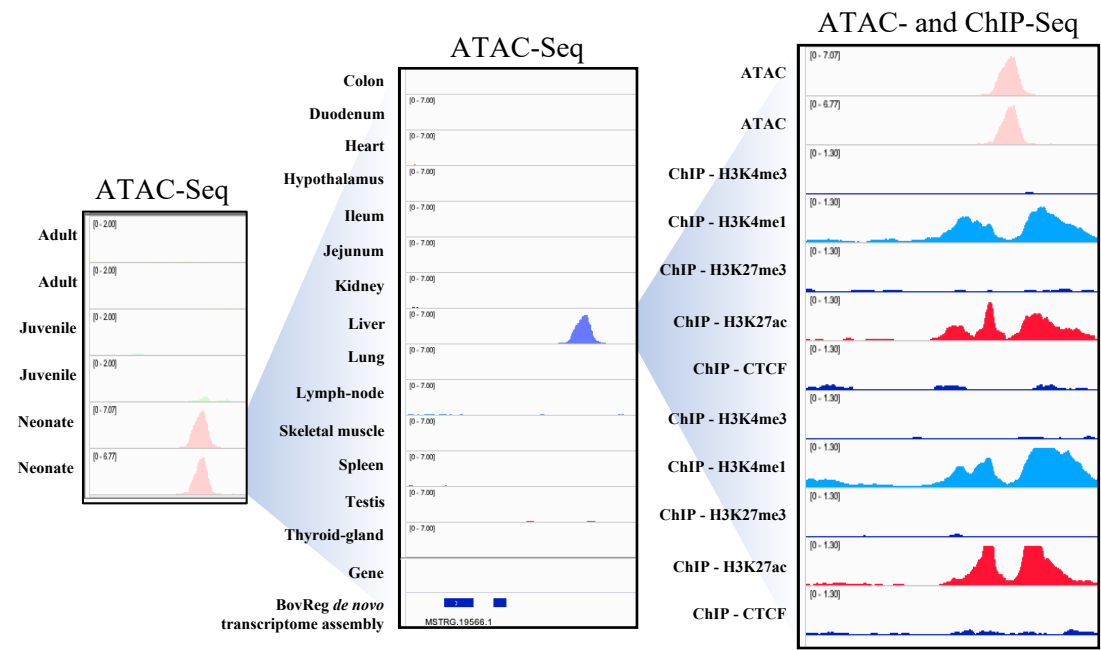


Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

➤ Around 25% of the genome was covered by different chromatin states (except quiescent)



Active enhancer (liver specific) only detected in neonate animals!!



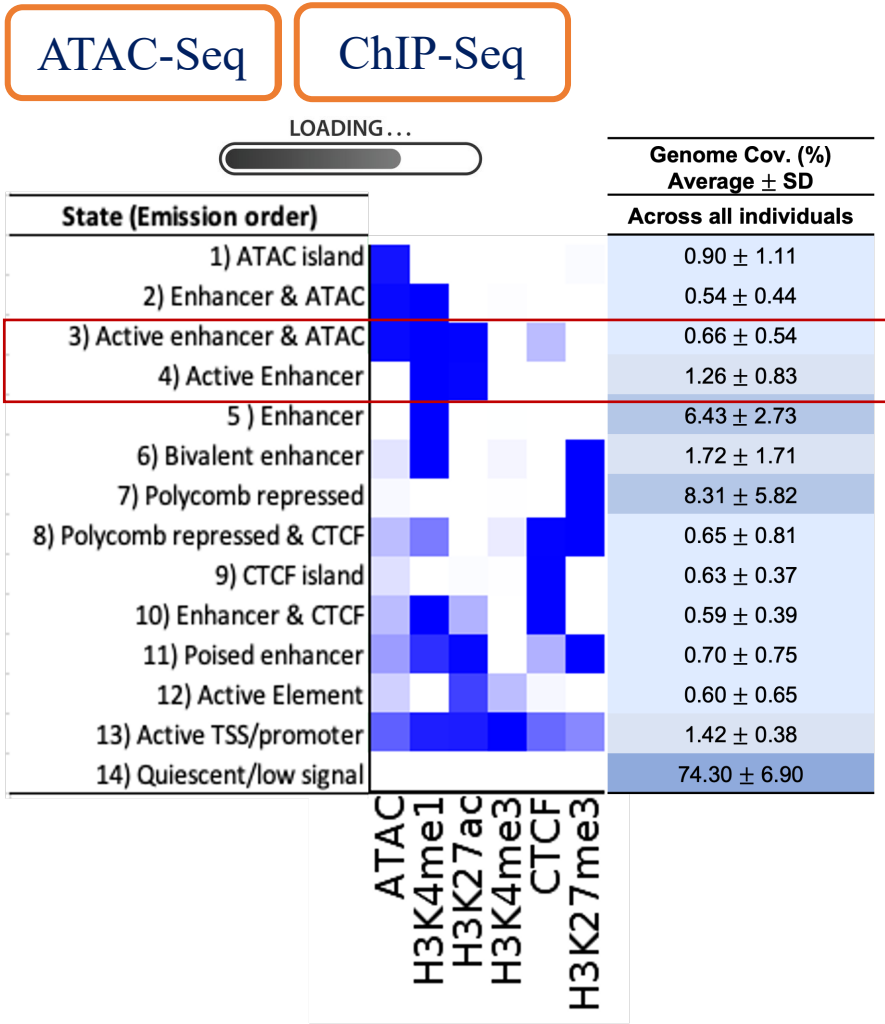
79% TSS annotated on RNA-Seq were validated by chromatin states

Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

➤ We recaptured known enhancers and annotated new ones!



Line 1 Herefords



BovReg active enhancers

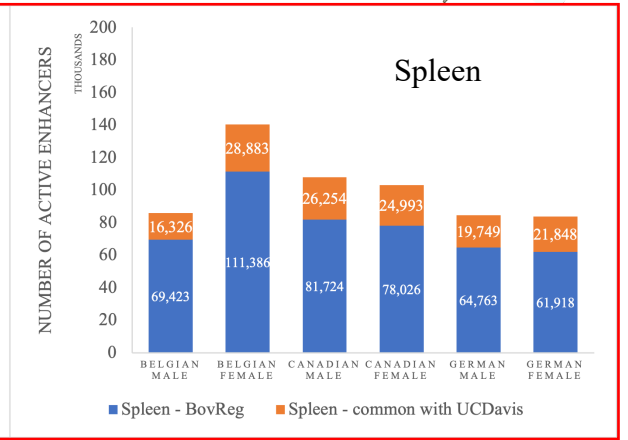
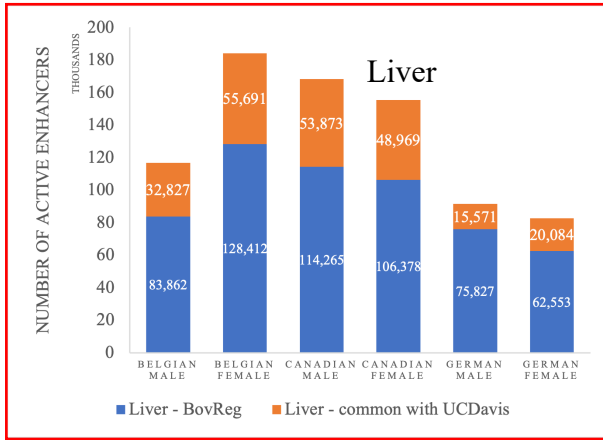
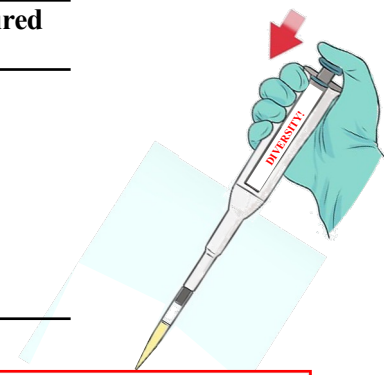
vs

Functional annotations of three domestic animal genomes provide vital resources for comparative and agricultural research

Colin Kern¹, Ying Wang¹, Xiaoqin Xu¹, Zhangyuan Pan¹, Michelle Halstead¹, Ganrea Chanthavixay¹, Perot Saelao¹, Susan Waters¹, Ruidong Xiang^{2,3}, Amanda Chamberlain³, Ian Kort⁴, Mary E. Delany¹, Hans H. Cheng⁵, Juan F. Medrano¹, Alison L. Van Eenennaam¹, Chris K. Tuggle⁶, Catherine Ernst⁷, Paul Flicek⁸, Gerald Quon⁹, Pablo Ross¹⁰ & Huaijun Zhou¹⁰

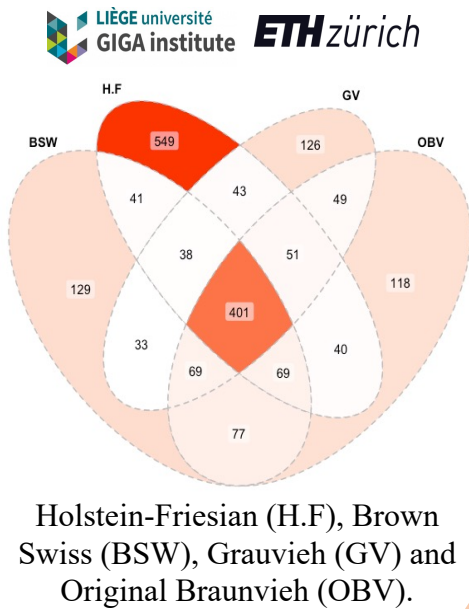
- 8 tissues from 2 Herefords male cattle (14 months old)

Tissues	Active enhancers (Kern et al., 2021)	Enhancers recaptured BovReg
Cerebellum	16,612	12,993 (78.21%)
Cortex	28,235	24,408 (86.45%)
Hypothalamus	27,796	17,277 (62.16%)
Liver	57,457	46,312 (80.60%)
Lung	49,861	38,537 (77.29%)
Muscle	28,023	19,442 (69.38%)
Spleen	29,707	24,520 (82.54%)

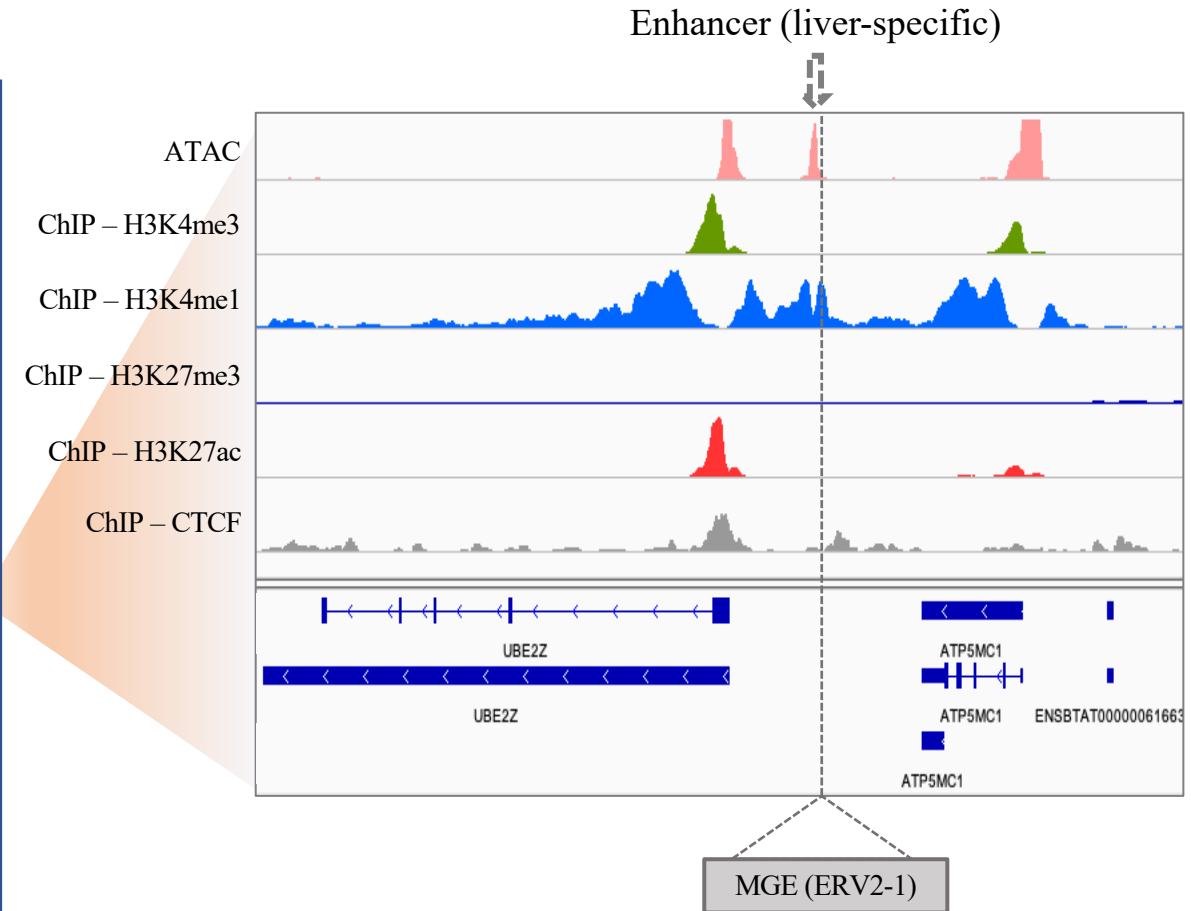
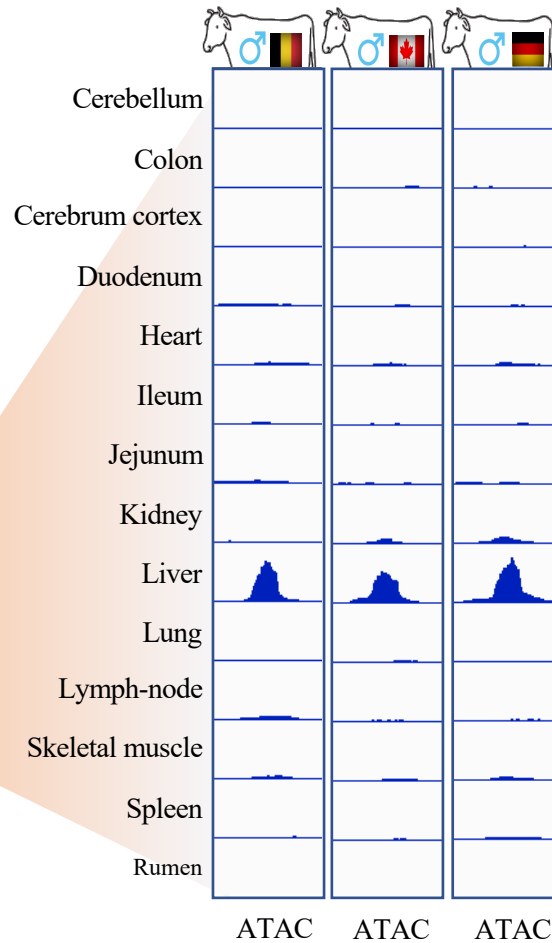


Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

➤ Mobile genetic element co-localized with liver-specific regulatory elements!



> 700 probes in genotyping array (public)



“Given the SV size distribution spike around the size of LTR elements, it is likely such transposable elements will increasingly be identified as a driving force behind bovine phenotypic diversity” - *Pangenome genotyped structural variation improves molecular phenotype mapping in cattle*; Leonard, Mapel & Pausch, bioRxiv 2023.



Perspectives & take-home message



Newly annotated genomic features for biology-driven genome selection: the **BovReg** contribution

Genomic features



eQTL in different bovine tissues
(liver, blood, jejunum, rumen, milk cells, muscle, adipose tissue and mammary gland)

Transcriptome map

- 48k genes models including $\geq 15k$ potentially novel transcripts
- 1,265 (638 known and 627 novel) miRNAs
- 51,295 transcription start sites (TSS)
- 2,328 TSS-Enhancer regions shared across the three populations

Regulatory regions

- 925,498 ATAC-Seq consensus peaks
- 327,236 H3K4me3 consensus peaks
- 396,828 H3K4me1 consensus peaks
- 316,262 H3K27me3 consensus peaks
- 454,530 H3K27ac consensus peaks
- 643,198 CTCF consensus peaks

Variations

- Mobile genetic elements (*727 probes, EuroGenomics array*)

LOADING...

WP7/Task7.2: Validating biology-driven genomic selection within and across small breeds (French, Nordic, Swiss, Canadian beef and dairy breeds)

BovReg
Understanding cattle genomes

BovReg Final Conference

University Foundation
Room "Felicien Cattier"
Bruxelles
(virtual participation available)

📅 14th (afternoon)
📅 15th (full day)
February 2024

Watch out our website www.bovreg.eu for further details to be announced soon!

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668.

Thank you for your attention!



www.bovreg.eu

@BovReg BovReg project

@BovReg.H2020 BovReg



WUR, the Netherlands. BovReg General Assembly, October 2022.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668

Disclaimer: the sole responsibility of this presentation lies with the authors. The Research Executive Agency is not responsible for any use that may be made of the information contained therein.

