



HAL
open science

Whole genome reverse genetic screen for natural deleterious variants in 4000 domestic ruminants to gain insight into mammalian gene function

Maxime Ben Braiek, Mekki Boussaha, Philippe Bardou, Pierre Faux, Bertrand Servin, Cécile Grohs, Julien Sarry, Florian Besnard, Jeanlin Jourdain, Sébastien Taussat, et al.

► To cite this version:

Maxime Ben Braiek, Mekki Boussaha, Philippe Bardou, Pierre Faux, Bertrand Servin, et al.. Whole genome reverse genetic screen for natural deleterious variants in 4000 domestic ruminants to gain insight into mammalian gene function. JOBIM 2024 - Journées Ouvertes en Biologie, Informatique et Mathématiques, Jun 2024, TOULOUSE (FRANCE), France. hal-04698695

HAL Id: hal-04698695

<https://hal.inrae.fr/hal-04698695v1>

Submitted on 16 Sep 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.

Whole genome reverse genetic screen for natural deleterious variants in 4000 domestic ruminants to gain insight into mammalian gene function

Maxime BEN BRAIEK¹, Mekki BOUSSAHA¹, Philippe BARDOU², Pierre FAUX³, Bertrand SERVIN³, Cécile GROHS¹, Julien SARRY³, Florian BESNARD^{1,4}, Jeanlin JOURDAIN^{1,5}, Sébastien TAUSSAT^{1,5}, Chris HOZE^{1,5}, Clémentine ESCOUFLAIRE^{1,5}, Hélène LECLERC^{1,5}, Patrice DEHAIS², Florent WOLOSZYN³, Claire KUCHLY⁶, Camille ECHE⁶, Camille MARCUZZO⁶, Cécile DONNADIEU⁶, Marcel AMILLS⁷, Stéphane FABRE³, Didier BOICHARD¹, Gwenola TOSSER-KLOPP³ and Aurélien CAPITAN¹

(1) GABI, Université Paris-Saclay, INRAE, AgroParisTech, 78350, Jouy-en-Josas, France ; (2) Sigeneae, GenPhySE, Université de Toulouse, INRAE, ENVT, 31326, Castanet-Tolosan, France ; (3) GenPhySE, Université de Toulouse, INRAE, ENVT, 31326, Castanet-Tolosan, France ; (4) IDELE, 75012 Paris, France ; (5) Eliance, 75012, Paris, France ; (6) GeT-PlaGe, Genotoul, US 1426, INRAE, 31320, Castanet-Tolosan, France ; (7) CRAG, CSIC-IRTA-UAB-UB, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain

maxime.ben-braiek@inrae.fr

The study of laboratory mammalian models (such as knock-out rodents) has dramatically improved our knowledge of the molecular basis of many phenotypes, including genetic disorders (MGI, OMIM databases).

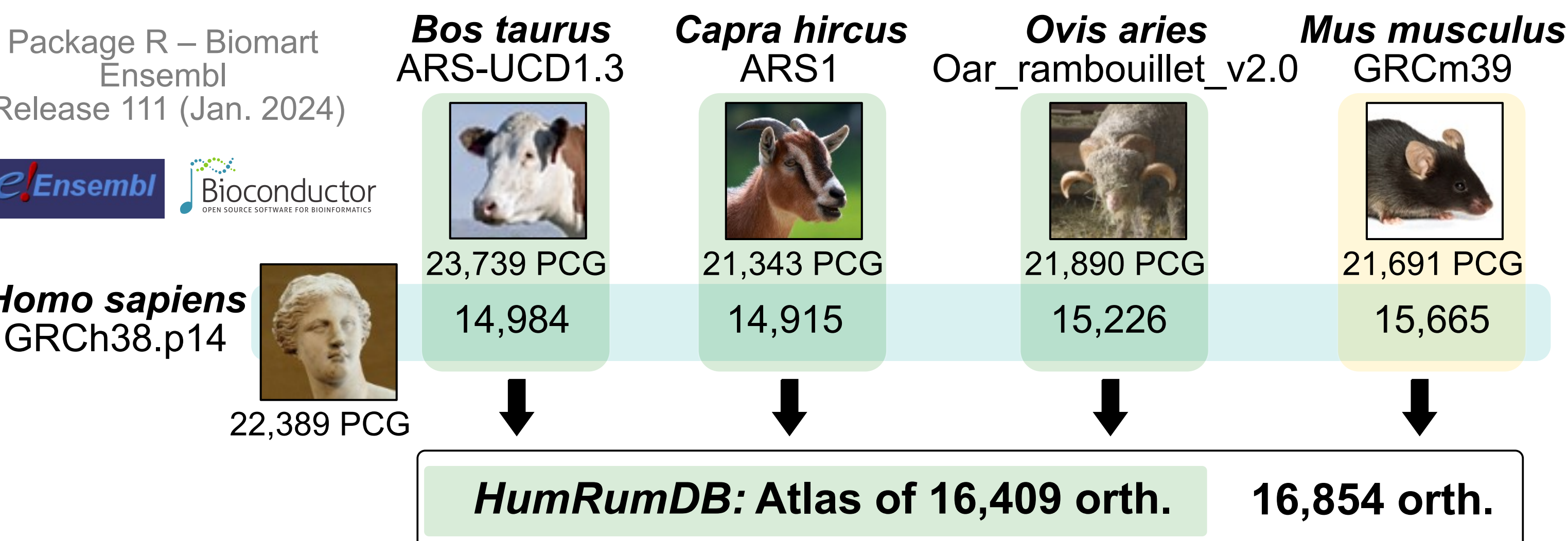
Advantages: high reproductive rate, short intergenerational interval and ease of laboratory breeding.

Limitations: specific strain (difficult to reproduce), embryonic lethality (difficult to observe), pathologies with incomplete penetrance and variable expressivity (genetic predisposition, environment, etc.)

Alternatives: study of natural genetic variations in animal populations in order to gain new insights into the mammalian gene functions and the pathological consequences of their inactivation. Ruminants are good models due to their effective fertility (by artificial insemination) greater than mice, better genetic diversity, similarities with humans (physiology, size, longevity...), high-throughput data (reference genome, functional annotation, WGS data from "1,000 genomes" projects, phenotypes ...)

1. Human-Ruminant orthologous protein-coding genes (PCG) database

Identification of 1:1 pairwise protein-coding genes orthologous to human:

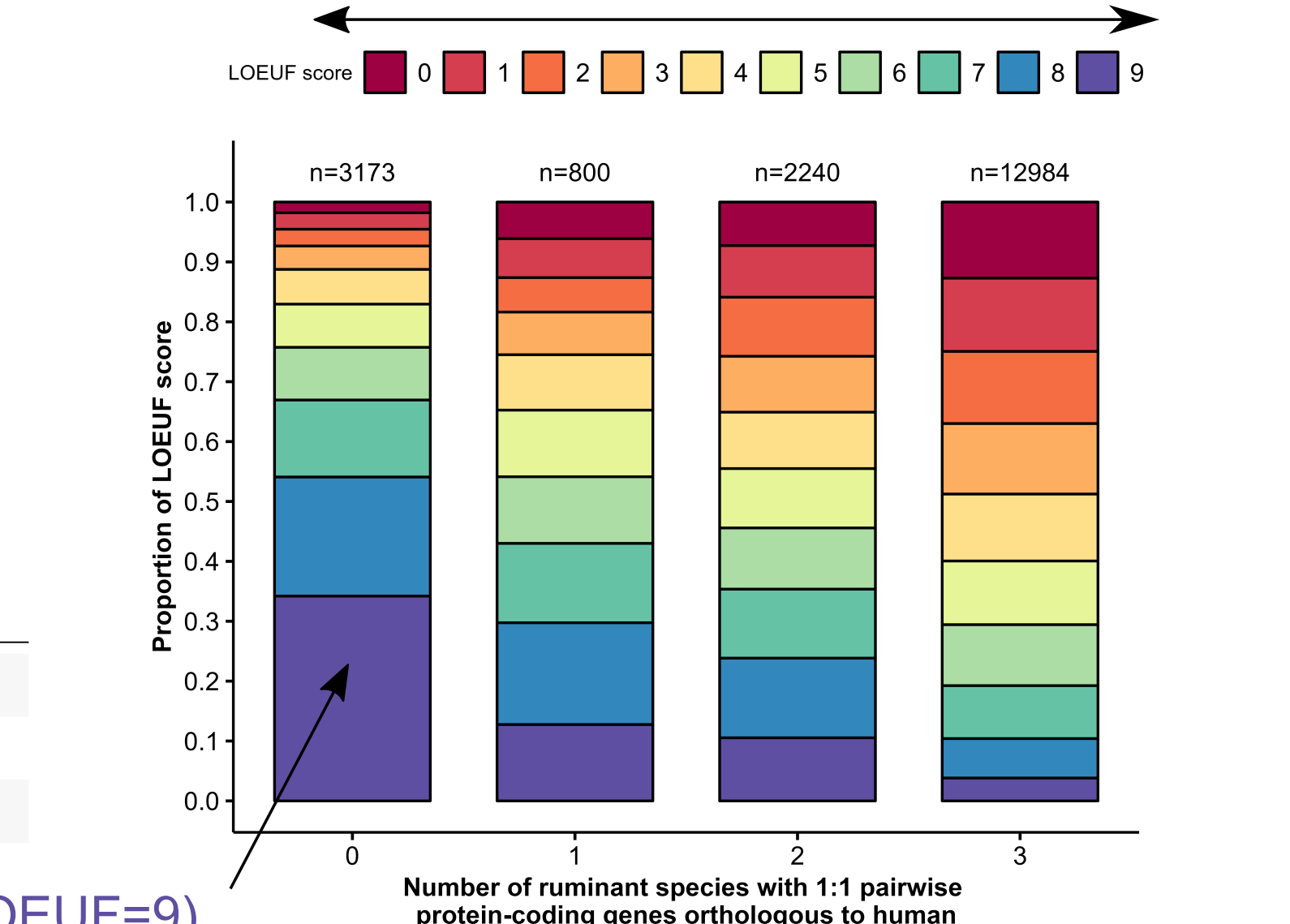


Phenotype information

Tolerance score (LOEUF* score)

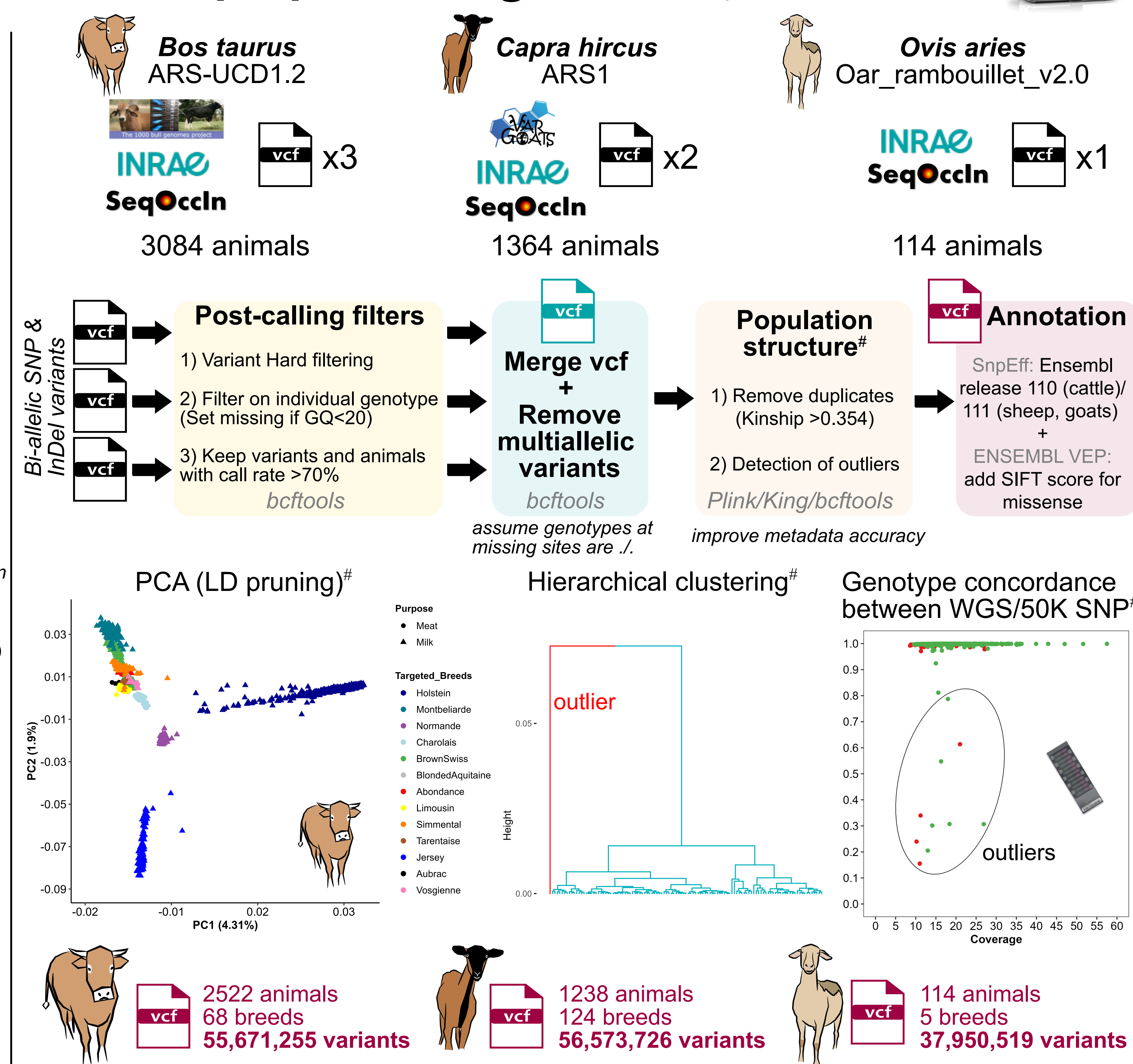
*LOEUF= loss-of-function observed/expected upper bound fraction
Karczewski et al. 2020

Essential genes (Strong selection against LoF) vs Non-essential genes (Tolerance to LoF accumulation)



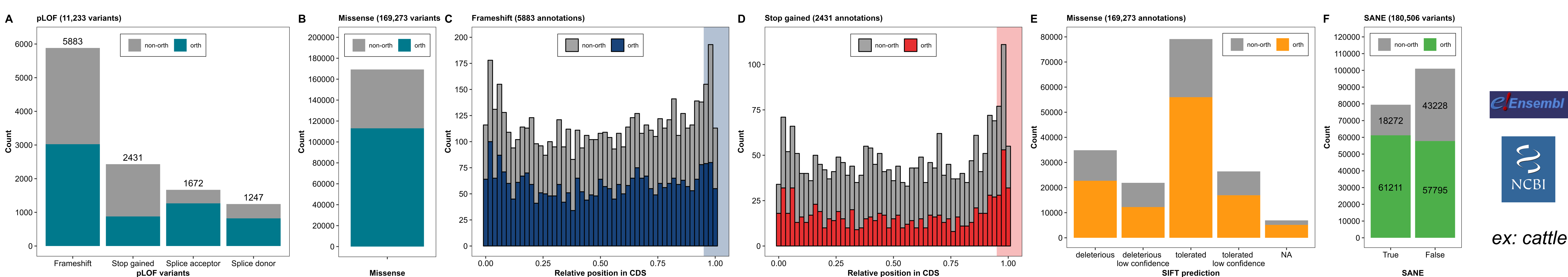
Enrichment of olfactory receptor activity (LOEUF=9)

2. Data preprocessing and analysis of WGS



3. Genomic landscape of putative Loss-Of-Function (pLOF) and missense variants in ruminants

Based on the annotation, we kept pLOF frameshift (Fig. A,C), stop-gain (removed if located in the last 5% of the CDS, Fig. A,D), essential splice (Fig. A) and missense (SIFT score: deleterious, Fig. A,E) variants located in human-ruminants orthologous genes and having the same annotation between NCBI and Ensembl (SANE, Fig. F). After filtering, we retained 15,673 variants (among 180,506) in cattle, 20,244 variants (among 221,292) in goats and 5742 variants (among 90,742) in sheep.



Conclusion and prospects

- Creation of a human-ruminant orthologous gene database with an inventory of deleterious variants in cattle, goats and sheep
- Specific focus on variants segregating in French breeds and located on nuclear genes encoding mitochondrial proteins ~ 1000 genes (Rath et al., 2021)
- Development of haplotype tests to predict genotype status for all genotyped animals in each breed
- Prediction of putative consequences on life trajectories (fertility, embryonic/juvenile/adult mortality) and production traits
- Functional validation (in vivo and in vitro)

The study of natural variants in livestock, especially those located in genes poorly characterized in mammals, could provide valuable information on the molecular mechanisms that determine phenotypic variation

References

Karczewski et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434-43.
Rath et al. MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res*. 2021;49:D1541-7.

Fundings

M.B.B. was supported by a postdoctoral fellowship from the Animal Genetics division of INRAE