

#### SEARCH FOR NEW MUTATIONS IN CATTLE BY SYSTEMATIC WHOLE GENOME RESEQUENCING

Mekki Boussaha, Camille Eché, Clémentine Escouflaire, Cécile Grohs, Carole Iampietro, Aurélien Capitan, Denis Milan, Christine Gaspin, Sébastien Fritz, Cécile Donnadieu, et al.

#### ▶ To cite this version:

Mekki Boussaha, Camille Eché, Clémentine Escouflaire, Cécile Grohs, Carole Iampietro, et al.. SEARCH FOR NEW MUTATIONS IN CATTLE BY SYSTEMATIC WHOLE GENOME RESE-QUENCING. 2ème colloque de l'infrastructure INRAE GENOMICS, Nov 2023, Bordeaux, France. . hal-04729114

#### HAL Id: hal-04729114 https://hal.inrae.fr/hal-04729114v1

Submitted on 9 Oct 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.

# RÉPUBLIQUE FRANÇAISE INRAØ SEARCH FOR NEW MUTATIONS IN CATTLE BY SYSTEMATIC WHOLE GENOME RESEQUENCING

Mekki Boussaha<sup>\*1</sup>, Camille Eché<sup>2</sup>, Clémentine Escouflaire<sup>3</sup>, Cécile Grohs<sup>1</sup>, Carole lampietro<sup>2</sup>, Aurélien Capitan<sup>1</sup>, Denis Milan<sup>2,4</sup>, Christine Gaspin<sup>5,6</sup>, Sébastien Fritz <sup>3</sup>, Cécile Donnadieu<sup>2</sup>, Didier Boichard<sup>1</sup>

<sup>1</sup>Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France, <sup>2</sup>INRAE, US 1426, GeT-PlaGe, Genotoul, France Genomique, Université Fédérale de Toulouse, Castanet-Tolosan, France, <sup>3</sup>Eliance, 75012 Paris, France, <sup>4</sup>GenPhySE, Université de Toulouse, INRAE, INPT, ENVT, Castanet-Tolosan Cedex, F-31326, France, <sup>5</sup>Université Fédérale de Toulouse, INRAE, BioinfOmics, GenoToul Bioinformatics facility, 31326, Castanet-Tolosan, France, <sup>6</sup>Université Fédérale de Toulouse, INRAE, BioinfOmics, GenoToul Bioinformatics facility, 31326, Castanet-Tolosan, France, <sup>6</sup>Université Fédérale de Toulouse, INRAE, BioinfOmics, GenoToul Bioinformatics facility, 31326, Castanet-Tolosan, France, <sup>6</sup>Université Fédérale de Toulouse, INRAE, MIAT, 31326, Castanet-Tolosan, France

\* Corresponding author

### Introduction

Systematic whole genome sequencing of artificial insemination bulls presents strong advantages:

- Discovery of all genetic variants, including the rarest ones
- Improvement of imputation up to the sequence level, especially in medium size populations
- Identification of candidate variants of QTL
- Early identification of new potential genetic defects and their

## Strategy to select potentially causal novel mutations



Variant calling with :

carriers before wide spreading in the population

As part of the SeqOccIn project, 571 AI bulls from 14 dairy and beef breeds were sequenced (Illumina NovaSeq technology). We present the strategy used to reduce false positives, with some of the results

#### GATK for SNVs Lumpy, delly & Pindel for SVs **Raw variants** Variant annotation with SnpEff/VEP **Annotated variants** ► Presence in .... **SNVs**: 152 millions 32 230<sup>\*</sup> SVs : New \* 13 288 SVs not yet published + 18 942 public SVs **Specific variants** Variants in known genes Link to phenotypes **OMIA** : Online Mendelian **OMIM** : Online Mendelian **Animal QTLdb** : Cattle QTLs Inheritance in Man Inheritance in Animals **MGI** : mammalian phenotypes

### **General results**

- \* Total # of SNVs : 34 252 085
- \* Novel candidate variants: 1 548
- \* Novel variants added into the EuroGMD chip : 1 342

#### \* Total # of SVs : 87 216

- \* Novel candidate variants: 1 219
- \* Novel variants added into the EuroGMD chip : 874

### Two examples of putative genetic defects

ACAN: Bulldog defect, already known in Dexter
ITGB4 : Junctional epidermolysis bullosa

Based on the severity and the certainty of these predictions, three AI bulls were culled

Several others are monitored with chip results (detection and phenotyping of homozygotes)

# Study of some candidate variants on milk production

Within-sire regression coefficients

Variant	Variant type	Breed	Milk	Fat	Protein	Fat%	Prot%	Haplotype frequency
PRLR	SV	BrownSwiss	ns	**	*	***	**	2.5%
DGAT1	SNP	Abondance	* * *	**	ns	* * *	* * *	4.3%
SLC25A21	SNP	Montholiordo	* * *	***	* * *	ns	ns	4.4%
BCO2	SNP		ns	**	ns	* * *	**	5.6%

**Candidate variants catalog** 

#### <u>Haplotypic tests</u> Identify haplotype in LD with the mutation Determine the ancestral origine of the mutation Identify homozygotes for the haplotype

#### Add variants to EuroGMD chip Large scale genotyping Backward imputation of the population with phenotypes GWAS

# Conclusion

- All AI bulls have a large impact in their population and are worth sequencing to fully characterize their genome
- New variants can be studied by genotyping with the chip widely used in genomic selection, by annotation, and by imputation in the population
- Here, we present a strategy to detect and characterize these new variants, while minimizing false positives
- 2 767 new variants with a strong annotation and an anticipated effect on phenotypes (abnormality or QTL)
- Large scale effect confirmation and identification of individuals of interest through SNP chip genotypes
- These variants will help to
  - > anticipate the emergence of genetic defects
  - To improve accuracy and persistence of genomic predictions

#### Acknowledgements: The SeqOccIn project was funded

by the Occitanie region, FEDER, and Apis-Gene.

