WP7 : Biology-driven selection

T3 :Validating biology-driven genomic selection within and across small breeds

Pascal Croiseau, Didier Boichard

BovReg Final Conference - Brussels

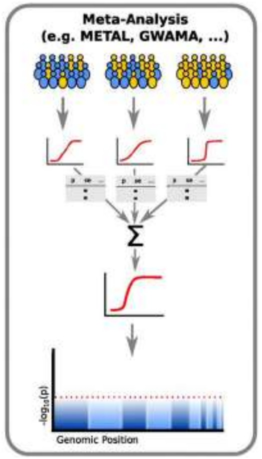
(14-15 February 2024)

Objectives

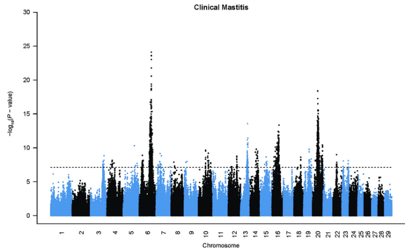
- Combining heterogeneous sources of information for biologically driven genomic selection

50K

Meta-Analyses
Conducted in BovReg

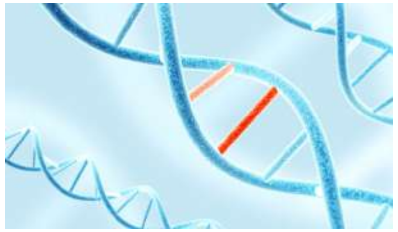


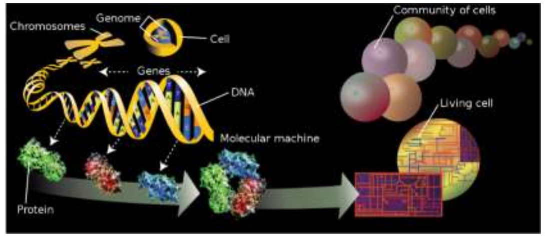
Nasirigerdeh et al., 2022



GWAS at the sequence level

Known causal
mutation





Genomic features
(BovReg & CGTeX)

- Adipose
- Liver
- Mammary
- Milk cell

Material

3323 Normand bulls

4261 Montbéliard bulls

8789 Charolaise cows

Material

		Training pop	Validation pop	total
Montbéliarde	5 production traits	3408	853	4261
Normande	5 production traits	2657	666	3323
Charolaise	weight at 18 months	6250	1564	7814
	thickness of bones, muscular and skeletal development	6790	1699	8789

Genomic features

- Milk production traits investigated
- 3 tissues used : Adipose / Liver / Mammary Gland
  - ✓ CGTEX
    - Step1 : Based on TWAS, selection of all significant genes (p-value <0.05)
    - Step2 : For each of these genes, selection of significant variants (p-value $<10^{-7}$) in the tissues of interest
  - ✓ BovReg
    - Use of the cis-eQTL top Variant results from BovReg WP4
    - 3 tissues for Montbéliarde and Normande breeds
      - ○ Liver
      - ○ Adipose
      - ○ Mammary gland
    - 1 tissue for the Charolaise breed
      - ○ Muscle

Genomic features

					Genomic Features			
		50K	GWAS	Meta-Analysis GWAS	CCTeX	BovReg Adipose	BovReg Liver	BovReg Mammary gland
Montbéliarde	Milk yield	43570	9899	12874	10834	2307	8807	2333
	Fat yield	43570	9899	6661	5017	2307	8807	2333
	Protein yield	43570	9899	6211	11133	2307	8807	2333
	Fat content	43570	9899	6211	11133	2307	8807	2333
	Protein content	43570	9899	30367	9430	2307	8807	2333
Normande	Milk yield	43570	9899	12827	10412	2307	8807	2333
	Fat yield	43570	9899	6655	5045	2307	8807	2333
	Protein yield	43570	9899	6573	11349	2307	8807	2333
	Fat content	43570	9899	18684	5666	2307	8807	2333
	Protein content	43570	9899	30089	9399	2307	8807	2333

		50K	Meta-Analysis GWAS	BovReg Muscle
Charolaise	muscular devlopment	43570	4383	450
	skeletal development	43570	3484	450
	thickness of bones	43570	3484	450
	weight at 18 months	43570	1630	450

Genomic Evaluations approaches

- BayesRCO : software developed by F. Mollandin during her PhD funded by GenSwitch
  - ✓ Bayes C, Bayes R, Bayes RC, Bayes RCπ

Mollandin *et al. BMC Bioinformatics* (2022) 23:365
https://doi.org/10.1186/s12859-022-04914-5

BMC Bioinformatics

RESEARCH **Open Access**

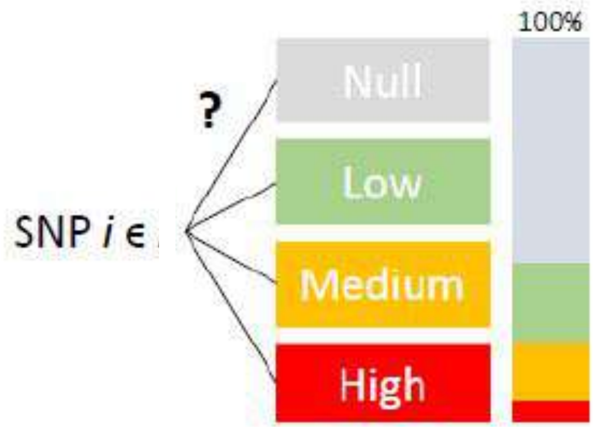
Accounting for overlapping annotations in genomic prediction models of complex traits

Fanny Mollandin[1*], Hélène Gilbert[2], Pascal Croiseau[1] and Andrea Rau[1,3]

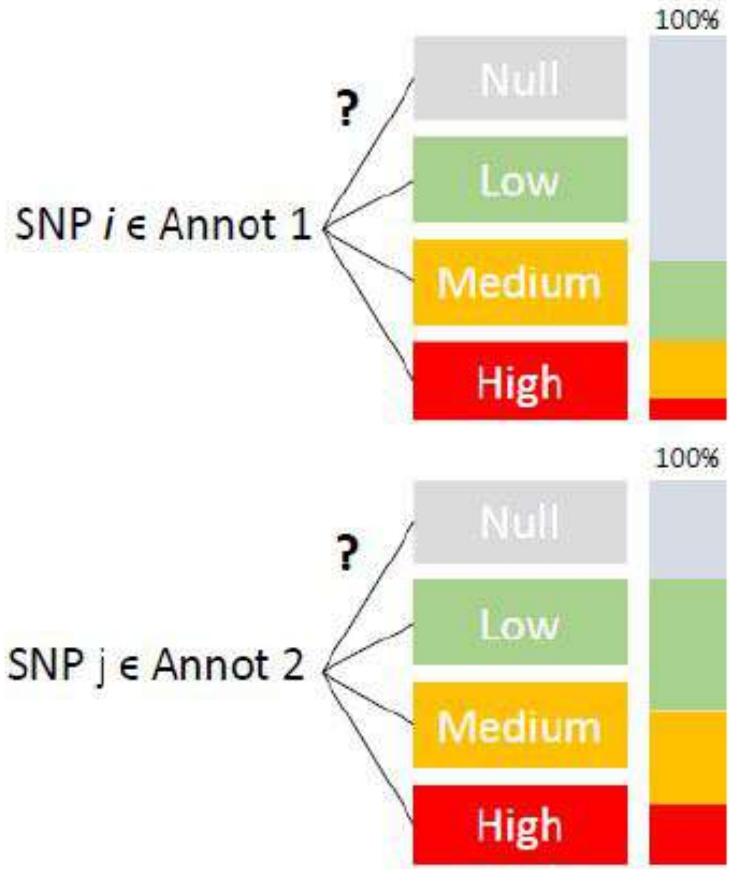
Funding

This work is part of the GENE-SWitCH project that has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement n°817998. This work also benefited from the clustering activities organized with the BovReg project, part of the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement n°815668. The financial support of the French National Agency of Research (ANR PigHeaT, ANR-12-ADAP-0015) is also gratefully acknowledged.

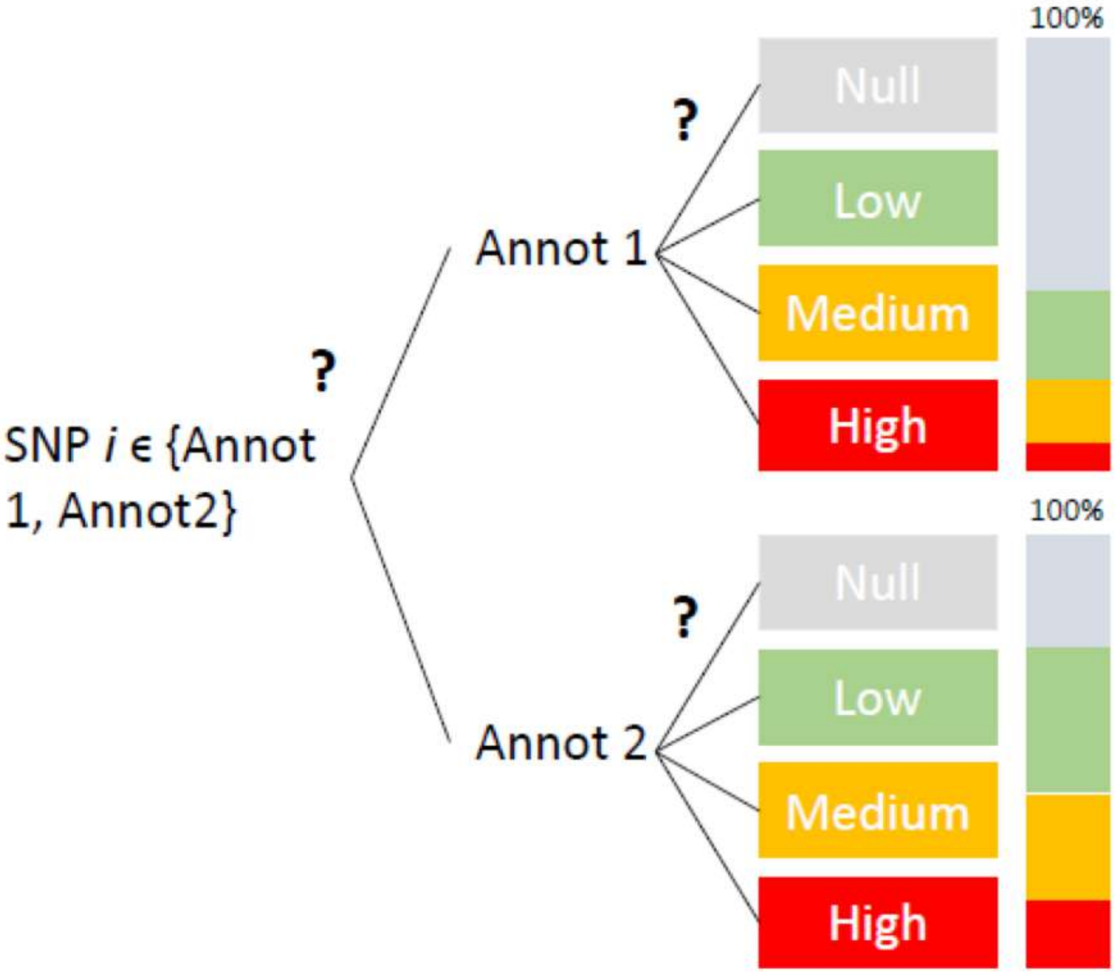
Bayes R



Bayes RC

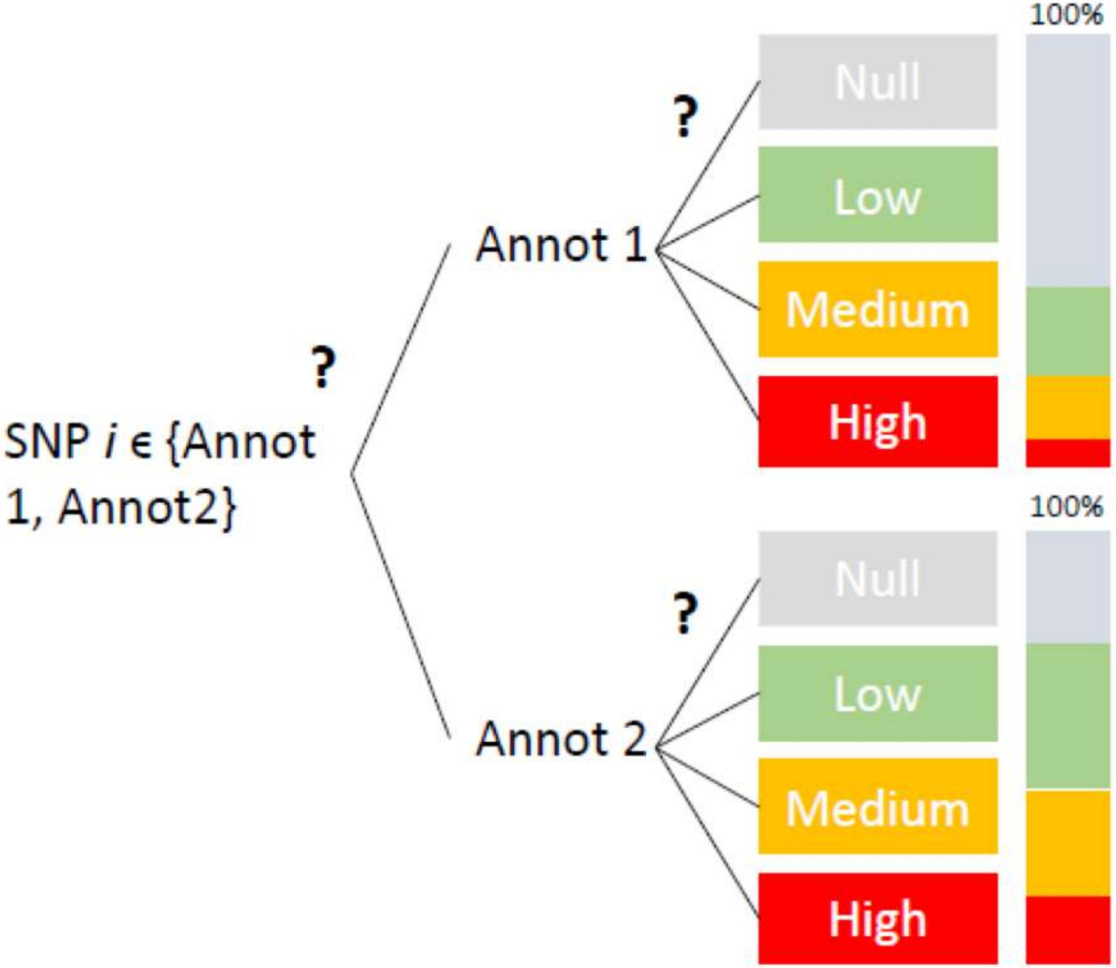

100%
?
Null
Low
SNP *i* ϵ Annot 1
Medium
High
100%
?
Null
Low
SNP j ϵ Annot 2
Medium
High

Bayes RCπ



SNP *i* ϵ {Annot 1, Annot2}
?
Annot 1
?
Null
Low
Medium
High
100%
Annot 2
?
Null
Low
Medium
High
100%

Bayes RCπ



In this study :

0

10^{-6}

10^{-3}

10^{-2}

Results : Montbéliarde

Correlations in the validation population

	50K	GWAS	Meta-Analysis GWAS	genomic features (CGTeX+BovReg)	BayesR	BayesRC	BayesRCpi
milk yield	0.365	0.354	0.321	0.268	**0.426**	0.419	0.419
fat yield	0.403	0.402	0.352	0.254	0.431	**0.433**	0.429
protein yield	**0.463**	0.371	0.324	0.243	0.437	0.436	0.435
fat content	**0.463**	0.182	0.250	0.297	0.452	0.448	0.447
protein content	0.544	0.432	0.466	0.158	**0.551**	0.546	0.548

- The correlations obtained using the selected groups of SNPs (Meta, GWAS and genomic features) range from 0.158 to 0.466.
- The combined use of these SNP groups and the 50K chip improves prediction accuracy for 3 of the 5 production traits (between 1 and 6 points of correlation).
- Models that take into account multiple annotations show no improvement

Results : Normande

Correlations in the validation population

	50K	**GWAS**	**Meta-Analysis GWAS**	**genomic features (CGTeX+BovReg)**	**BayesR**	**BayesRC**	**BayesRCpi**
milk yield	**0.399**	0.237	0.181	0.069	0.351	0.353	0.351
fat yield	**0.406**	0.254	0.179	0.093	0.402	0.396	0.388
protein yield	0.345	0.223	0.133	0.114	**0.350**	0.347	0.343
fat content	**0.537**	0.440	0.434	0.344	0.536	0.537	0.535
protein content	0.487	0.173	0.474	0.150	**0.524**	0.440	0.436

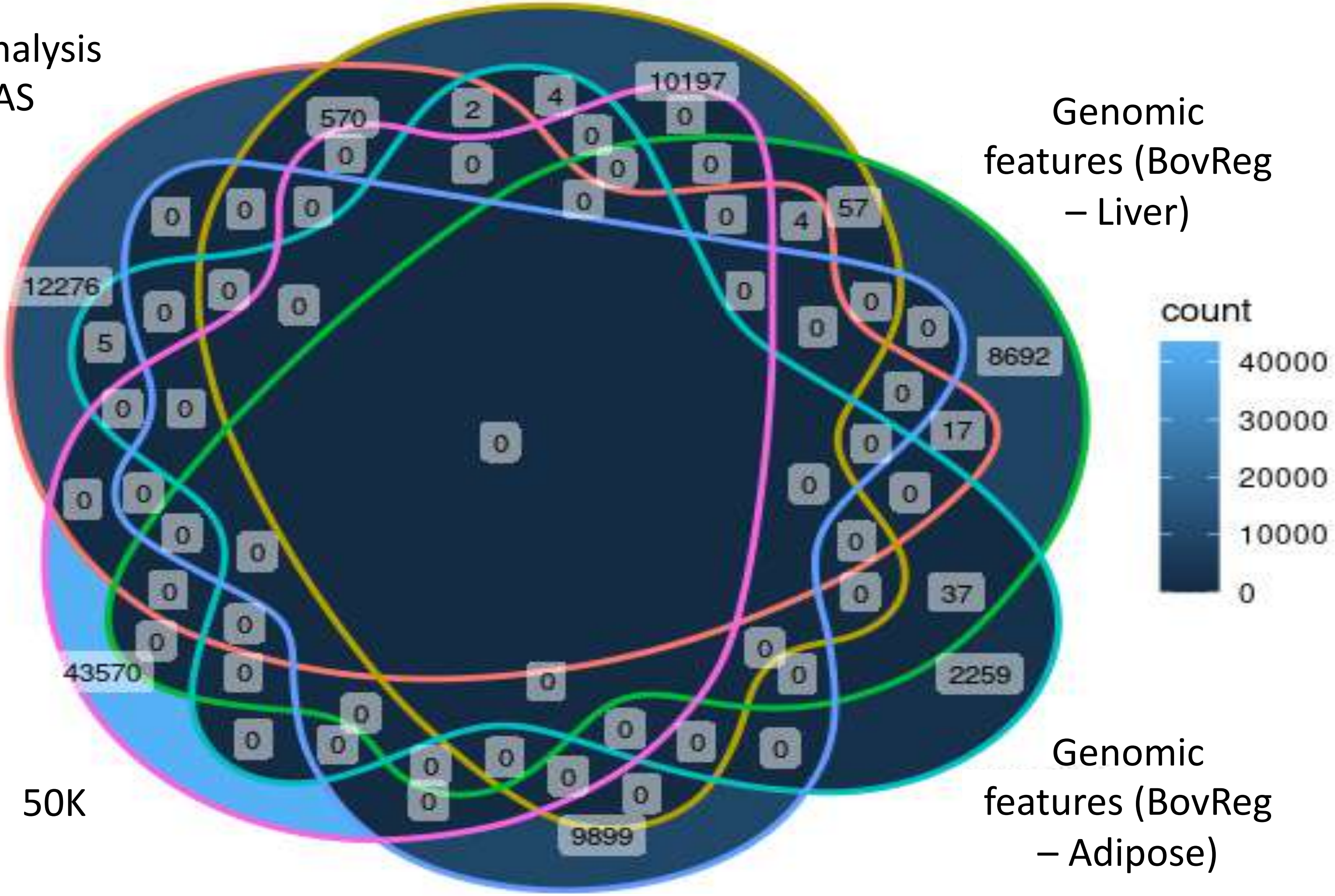
- Accounting for annotation classes does not improve the accuracy of genomic predictions, except for protein content.

Results : Charolaise

	50K	GWAS	Meta-Analysis GWAS	genomic features (BovReg)	BayesR	BayesRC	BayesRCpi
muscular development	0.328	x	0.329	0.171	**0.398**	0.396	0.397
skeletal development	0.365	x	0.240	0.168	0.378	**0.383**	**0.383**
thickness of bones	0.312	x	0.238	0.171	0.325	**0.326**	**0.326**
weight at 18m	0.297	x	0.178	0.135	**0.308**	0.307	0.307

- The inclusion of annotation classes improves the accuracy of genomic predictions for all traits, but only marginally, except for muscle development (+7 points of correlation).
- As for the other breeds, BayesRC and BayesRCп do not improve the prediction accuracy.

Montbéliarde – Milk Yield



Genomic features (CGTeX)

Meta-Analysis GWAS

Genomic features (BovReg – Liver)

Genomic features (BovReg – Adipose)

50K

GWAS

Montbéliarde – Milk Yield

		effect			
		null	***small***	***medium***	***strong***
BayesR	CGTeX	0	10828	6	0
	BovReg Liver	0	8798	9	0
	BovReg Adipose	0	2301	6	0
	Meta-Analysis GWAS	0	12873	1	0
	GWAS	0	9897	2	0
	50K	0	43551	19	0

		effect			
		null	***small***	***medium***	***strong***
BayesRCπ	CGTeX	0	10826	8	0
	BovReg Liver	0	8791	16	0
	BovReg Adipose	0	2294	13	0
	Meta-Analysis GWAS	0	0	12873	1
	GWAS	0	9897	2	0
	50K	43132	411	27	0

Conclusion

- No significant improvement in the accuracy of genomic predictions using different sources of biological information.
- However, we can hope that the most important markers for a given trait will be better exploited by these new prediction models.
  - ✓ more stable effects over time
  - ✓ less sensitive to LD

BovReg *PARTNERS*

AARHUS UNIVERSITY

CRG Centre for Genomic Regulation

diagenode Innovating Epigenetics Solutions

edinethics

EMBL-EBI

ETH zürich

FBN

FACULDADE DE MEDICINA VETERINÁRIA UNIVERSIDADE DE LISBOA

INSTYTUT GENETYKI I HODOWLI ZWIERZĄT PAN

INRAE

LIÈGE université GIGA Medical Genomics

Université de Limoges

Luke LUONNONVARAKESKUS

MARTIN-LUTHER UNIVERSITÄT HALLE-WITTENBERG

ROSLIN

UNIVERSITY OF ALBERTA

THE UNIVERSITY OF EDINBURGH

Universiteit Utrecht

WAGENINGEN UNIVERSITY & RESEARCH

Thank you for your attention

www.bovregproject.eu

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