Updates for Goat SNP Chips

Guilherme Borba Neumann

Humboldt-Universität zu Berlin, Germany

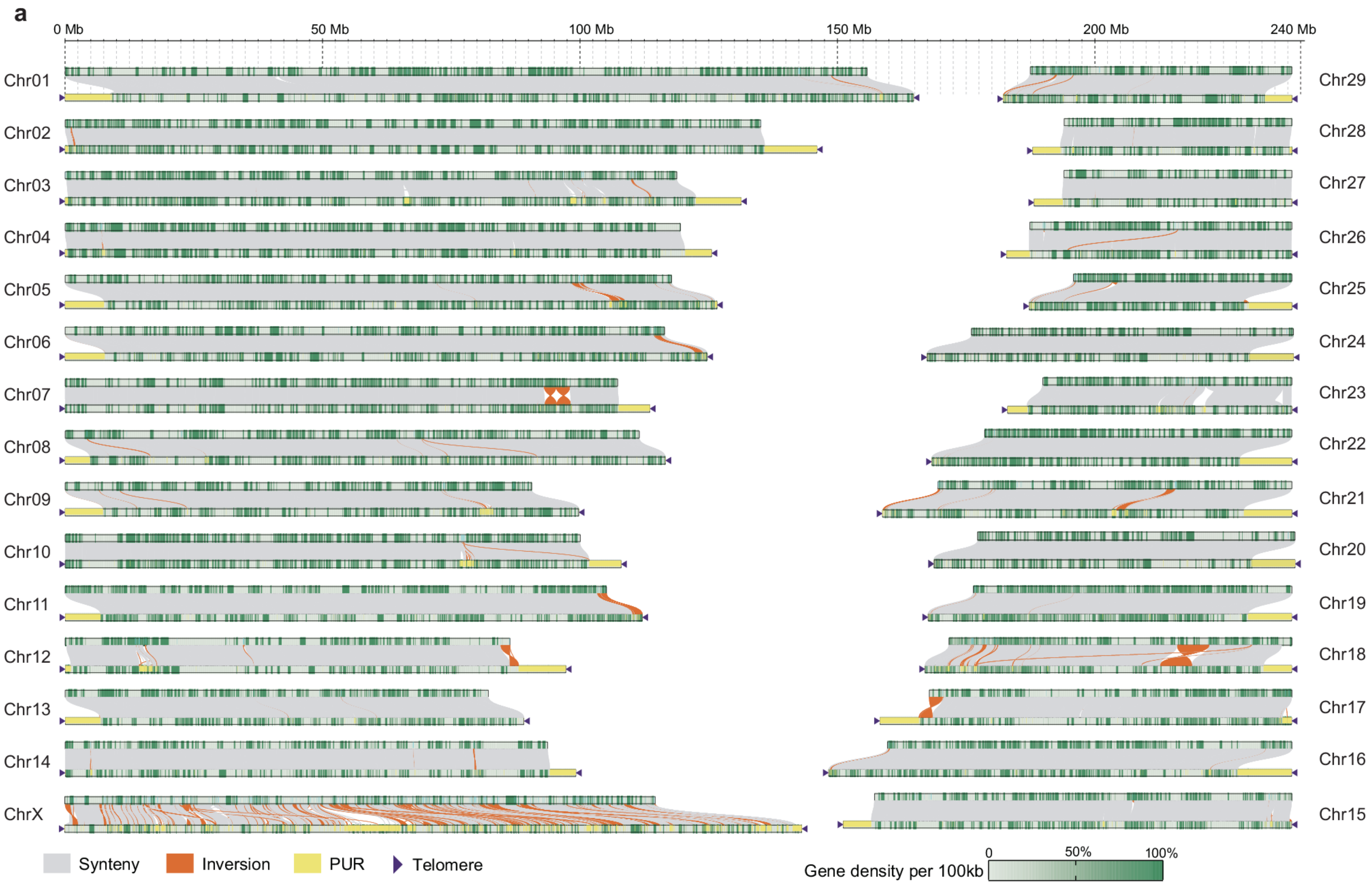
Current address: Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

On Behalf of the International Goat Genome Consortium

President of the IGGC, Gwenola Tosser (gwenola.tosser@inrae.f **)**

Goat genomics

- The first 2011 assembly (CHIR_1.0) has been upgraded to the currrent ARS1 reference (2016) - Genome annotation by Ensembl team (version 92)
- The assembly is still improving (21 published assemblies at NCBI)
- Latest version: November 2024- Telomere to telomere



Wu, H., Luo, LY., Zhang, YH. *et al.* Telomere-to-telomere genome assembly of a male goat reveals variants associated with cashmere traits. *Nat Commun* **15**, 2024.

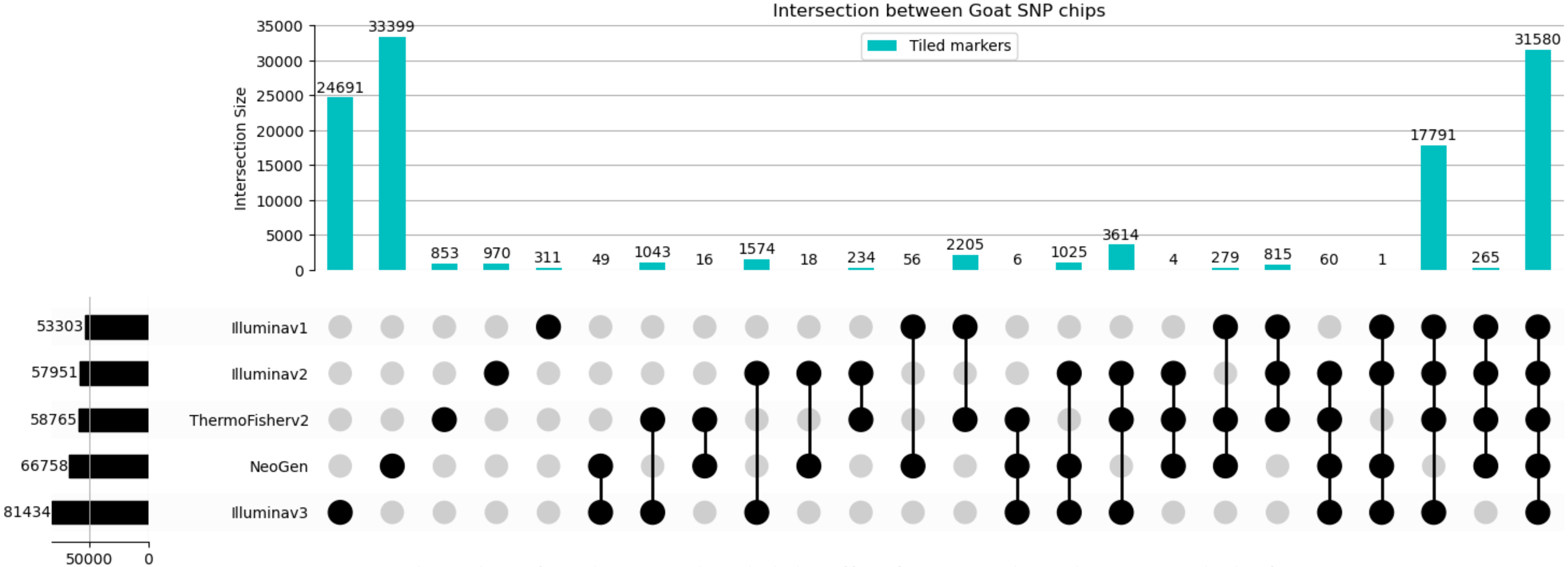
Goat SNP chips

- SNP chips were developped with two main technologies: Illumina & Thermofisher (formerly Affymetrix)
- Their evolution takes into account:
    - the assembly evolution
    - benefits from genome knowledge (inclusion of QTL regions, causal mutations, Structural Variants…)
    - Allows combination with alternative genotyping technologies (Genotyping by Sequencing)

- SNP chips allow the development of projects and collaborations:
    - 2013-2018: ADAPTMAP project (>4000 50K genotypes, 144 breeds worldwide), 6 publications
    - 2019- : VarGoats project (1,372 sequenced animals)

Introduction: several SNP chips for goats

- 2011: first goat SNP chip (GoatSNP50, Illumina, ~53K), CHIR_1.0 assembly
- **2020: 10K IMAGE chip (ThermoFisher) contains 8K from GoatSNP50**
- 2021: update of the Illumina chip: Goat_IGGC_65K_v2 (~60K), ARS1 assembly
- **2021: GGP Goat 70k chip (NeoGen/Geneseek, Illumina), ~40K from GoatSNP50**
- **2022: Axiom Goat v2 chip (ThermoFisher) contains Goat_IGGC_65K_v2 and CNV regions**
- **2024: Goat_IGGC_HTS_v3 (~80K), ARS1 assembly**



- Total number of markers per chip slightly differs from actual numbers due to lack of some genomic coordinates used to calculate overlap between chips

Focus on Goat_IGGC_HTS_v3

83,951 SNPs in the design

- 2,703 monomorphic SNPs from V2

+ 21,928 new SNPs

+4,993 probes for 265 SVs + 20 probes for CSN1S1 SVs

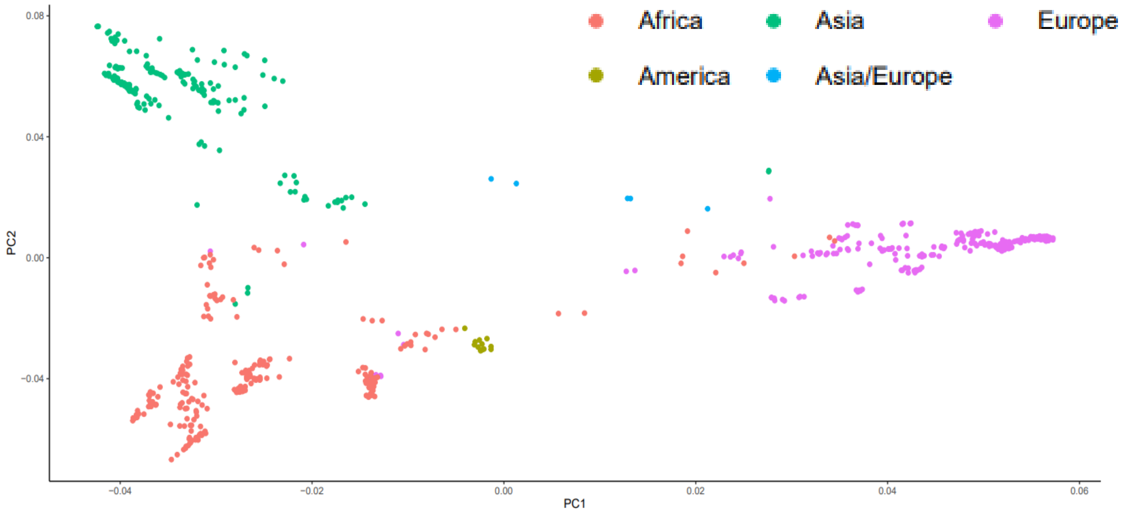
83,511 SNPs were synthetized

79,538 SNPs were clustered on 384 animals

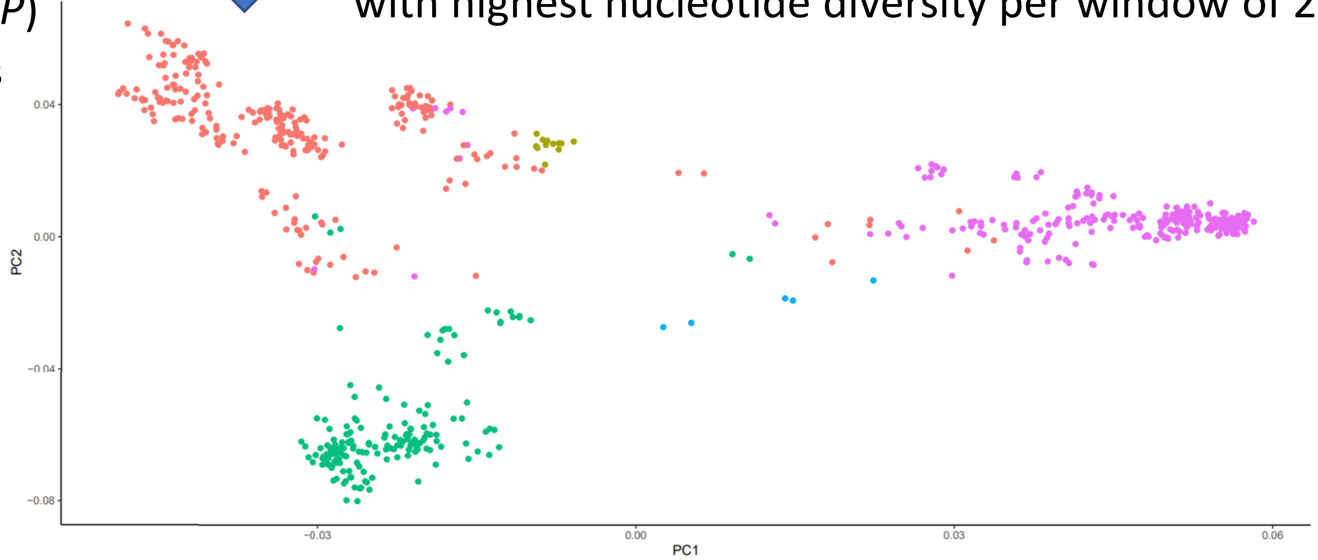
WORK IN PROGRESS

V3 improvements (22K new markers)

- **A representative set of ~10K of VarGoats SNPs (worldwide)**
- 3.5K GBS markers (allowing analysis with GBS data to be confirmed)
- 1.5K ancestral SNPs from IMAGE chip design
- 1.6K Loss of Function predicted mutations
- 4.3K gap filling markers
- 1K markers from candidate regions and genes (e.g. casein genes, *PRNP, SHBG, TMEM154, PAEP*)
    - Improved prediction of CSN1S1 genotypes (with null alleles O1 and O2)

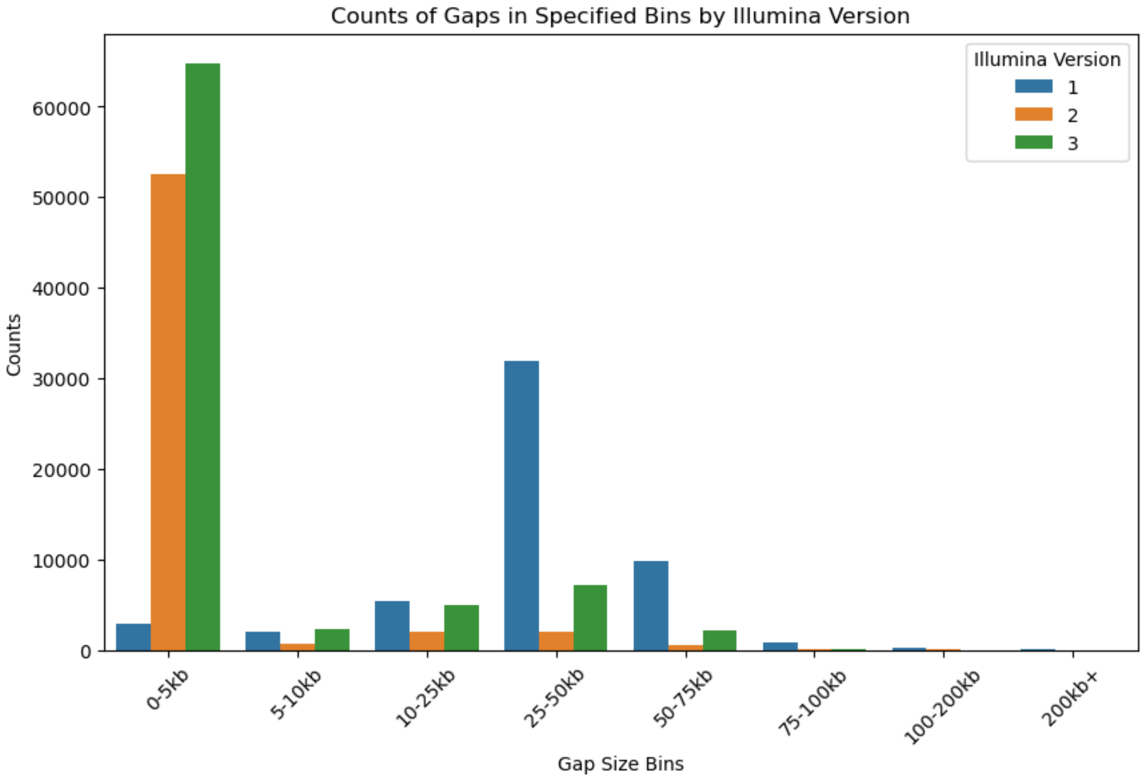


From 28 Mio to 10,010 autosomal SNPs, selecting SNPs with highest nucleotide diversity per window of 247 Kbp



V3 improvements (22K new markers)

- A representative set of ~10K of VarGoats SNPs (worldwide)
- 3.5K GBS markers (allowing analysis with GBS data to be confirmed)
- 1.5K ancestral SNPs from IMAGE chip design
- 1.6K Loss of Function predicted mutations
- **4.3K gap filling markers**
- 1K markers from candidate regions and genes (e.g. casein genes, *PRNP, SHBG, TMEM154, PAEP*)
    - Improved prediction of CSN1S1 genotypes (with null alleles O1 and O2)



Clustering process

Out of 83,511 synthetized, 79,538 markers clustered

95% success rate

Validated for 384 animals from 11 breeds, including Alpine, Boer, Kiko, La Mancha, Nigerian Dwarf, Nubian, Saanen, San Clemente, Island, Savanna, Spanish, and Toggenburg

SNP origin	Clustered
v1	51494
v2	4519
V3-new	23525
Total	**79538**

SNP category	Synthetized	Clustered	Clustered Percentage	Clustered Duplicates
VarGoats	10055	8690	0.86	
Ancestral	1511	846	0.56	398
Candidate regions	800	726	0.91	10
CSN1S1_null -alleles	20	17	0.85	10
Gap filler	4247	3592	0.85	
GBS	3444	3126	0.91	347
LoF	1634	1507	0.92	
Casein cluster	46	43	0.93	
PRP new	8	8	1.00	6
SV	4970	4970	1.00	
Version2	56776	56013	0.99	707
Total	83511	79538	0.95	1478

Conclusion

- Four current low and medium density SNP chips available for goats with overlapping content

- New high density goat SNP chip available, with less gaps and reduced ascertainment bias (use of breeds worldwide)

- SNP submission available (sheep and goats) anytime for a next update https://image.sigenae.org/
  Send also by e-mail to gwenola.tosser@inrae.fr and Shannon.Clarke@agresearch.co.nz

- Imputation from chip to whole-genome sequence level and between chip versions are expected to be available soon

Acknowledgements

We thank the SNP providers and the scientists & engineers who analysed the genotyping data on a set of 384 relevant animals (kindly provided by AgResearch, LECA, UNIMI, VarGoats project, INRAE inhouse project and ??).
We thank Sigenae for bioinformatics analysis, data and information releases.
We thank Illumina for providing chips for genotyping 384 animals and the cluster file generation

Co-authors:

Hayley Baird, Rudiger Brauning and Shannon Clarke - AgResearch, Invermay Agricultural Centre, New Zeeland

Andrea Talenti - The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, United Kingdom

Antonia Noce, Marcel Amills and María Gracia Luigi-Sierra - Centre de Recerca Agrigenòmica (CRAG), Universitat Autònoma de Barcelona, Spain

Barbara Lazzari - Università Cattolica del Sacro Cuore, Piacenza, Institute of Agricultural Biology and Biotechnology (IBBA), National Research Council (CNR), Italy

Licia Colli - Dept. of Animal Science, Food and Nutrition - DIANA, Università Cattolica del S. Cuore, Italy

Marco Milanesi - Department for Innovation in Biological, Agro-food and Forest systems (DIBAF), University of Tuscia, Italy

Guilherme Borba Neumann and Siham Rahmatalla - Albrecht Daniel Thaer-Institute for Agricultural and Horticultural Sciences, Humboldt Universität zu Berlin, Germany

Casey Turk, Lee Ashley, André Eggen - Illumina

Rebecca Bellone and Anna S. Grulikowski – UCDavis, United States

Celine Vandecasteele and Joanna Lleda - INRAE, Get-PlaGe sequencing platform , France

Philippe Bardou , Mathieu Charles - INRAE, Sigenae platform, France

Mekki Boussaha and Marie-Pierre Sanchez - INRAE, GABI, France

Thomas Faraut, Julien Sarry, Hélène Larroque, Isabelle Palhière, Rachel Rupp, and Gwenola Tosser-Klopp - INRAE, GenPhySE, France

Thank you for your attention! Any questions?

Cluster

SNP category	Clustered in file v3_C1	duplicates
10133_VarGoats-SNP	8690	
Ancestral	846	398
candidate_region	726	10
CSN1S1_null-alleles	17	10
gap_filler	3592	
GBS	3126	347
LoF	1507	
new-variations-in-casein_cluster	43	
PRP_new	8	6
SV	4970	
v2	56013	707
Total	**79538**	**1478**

SNP origin	Clustered in file v3_C1
v1	51494
v2	4519
v3-submitted	23525
Total général	**79538**

Breed	n
Alpine	2
Boer	19
Crossbred	2
Experimental	1
Kiko	22
La Mancha	1
Nigerian Dwarf	7
Nubian	3
Saanen	1
San Clemente Island	1
Savanna	8
Spanish	20
Toggenburg	