



Towards a high density SNP International Goat Chip

Gwenola Tosser-Klopp, Abdullah Johari Jiken, Philippe Bardou, Cédric Cabau, Patrice Dehais, Thomas Faraut, Henri Heuven, Christophe C. Klopp, Patrice Martin, Carole Moreno-Romieux, et al.

► To cite this version:

Gwenola Tosser-Klopp, Abdullah Johari Jiken, Philippe Bardou, Cédric Cabau, Patrice Dehais, et al.. Towards a high density SNP International Goat Chip. Plant and Animal Genome, Jan 2011, San Diégo, United States. hal-01194170

HAL Id: hal-01194170

<https://hal.science/hal-01194170>

Submitted on 6 Jun 2020

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.

Towards a high density SNP International Goat Chip

Gwenola Tosser-Klopp
on behalf of the International Goat
Genome Consortium

Outline

- Introduction
- Data
- Data processing
- Expected outcome
- Questions to be discussed

High density SNP chips are useful

- Understanding species' evolution, domestication...
- Dissecting genetic mechanisms of complex agronomic traits
- Improving selection methods

SNP chips in many species

- SNP chips are available in several species
 - Humans
 - Model animals : mouse, rat
 - domestic species : chicken, dog, horse, cattle, sheep, pig
- Affymetrix and Illumina technologies are mainly used
- 20 K <density<1.1 M

Several SNP discovery projects for goats in the world

- University of Utrecht, Netherlands (H. Heuven / M.Groenen)
 - Reduced Representation Libraries of 17 dairy goats. 120 millions of 32bp paired-ends sequences = 3 Gb
- Malaysian Agricultural Research and Development Institute, Malaysia & DNA Landmarks, Canada (JJ Abdullah / TP Yu)
 - Whole genome sequencing of 3*20 Boer, Savanna and Kacang meat and indigenous goats = >210 Gb
- French National Institute for Agricultural Research (INRA), France (Ph. Mulsant / R. Rupp / C.Moreno):
 - RRL of 6 goats and whole genome sequencing of 13 Alpine, Saanen and Creole goats = >150 Gb

Goat genomic tools in progress

- Reference sequence : see W. Zhang presentation
- RH map in progress : see B. Sayre presentation
- 20K SNP chip: see poster P198
- International resequencing project of 25 goat breeds will also give more SNPs
- Several projects around expressed sequences (France, China)

We propose to federate these data

- Centralisation of the data at INRA
- Data processing with INRA pipeline
- www.sigenae.org & T.Faraut
- Data access to all partners

Steps of the Bioinformatic process

- alignment of the different sequences on *de novo* assembled goat contigs (BWA software)
- ordering of contigs on ovine or bovine genome reference sequence
- SNP discovery and annotation
- small set SNP test generation
- determination of the final SNP panel

Web interface

- Data visualisation with Ensembl



- Data extraction with Biomart : select SNP on different criteria = chromosome, specific contig, dataset



Based on Ensembl release 40 - Aug 2006

putative SNP:
snp_EE760096_333

snp_EE760096_333 - SHP info

CD285973.p.ov.8
106 - 306

- Graphical view
- Polymorphism view
- Export information about region
- Export sequence as FASTA
- Export EMBL file

Use Ensembl to...

- Search Contig Browser
- Data mining [BioMart]
- Export data
- Download data
- Digital Differential Display
- Venn Diagrams

Docs and downloads

- About Contig Browser
- Contig Browser data

Other links

- Home
- Sigenae
- Agenae



SNP Report

SNP	snp_EE760096_333 (putativeSNP1)
Synonyms	None currently in the database
Alleles	C/T (ambiguity code: Y)
Validation status	Unknown
Linkage disequilibrium data	No linkage data for this SNP
Sequence region	<pre>GGACGACCCCTGGAGATGATCGAACCGCCACGCTGCAATACAAAGCTACTGGAACTGTT XTGCTCTGGGCAAGGAGCGATTGCTGCTGCGGATTCGCGTCCGACTGAAAGGCTGCT G (SNP highlighted)</pre>

60b+SNP+60b

☐ Genotype frequencies per population

This SNP has not been genotyped in a population.

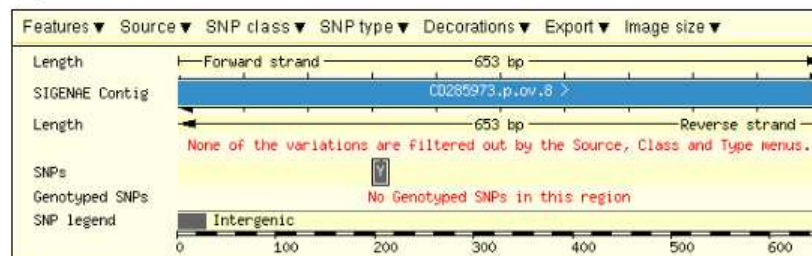
☐ Allele frequencies per population

This SNP has no allele frequencies per population.

☐ SNP snp_EE760096_333 is located in the following transcripts

There are no transcripts that contain this SNP.

☐ SNP Context - contig CD285973.p.ov.8 206



☐ Individual genotypes for SNP snp_EE760096_333

Criteria for SNP selection (1/2)

- Good sequence quality around the SNP
- Each allele of the SNP is included in at least 2 reads
- No repetitive elements surrounding the SNP (within 100bp)
- The SNP must be predicted by a minimum of 6 reads
- No overlap with complex regions

Ref : Matukumalli, 2009 & Ramos, 2009

Criteria for SNP selection (2/2)

- Representativity of the whole genome
- Also take SNPs in the regions that are not mapped on the genome
- Intra-species polymorphism in a maximum number of breeds
- Other criteria (to be defined)?

Expected outcomes

- Millions of SNPs
- A majority will be mapped on the genome
- Annotation will be facilitated by RNA seq data
- A first international chip order is planned at autumn 2011

Open questions

- Density of the chip
- Technology to be used
- The price depends on the interest of the goat community for this tool :
 - for example, Illumina's price is between \$130 and \$300 for a 50K SNP chip (20000 or 5000 chips)

PAG is an opportunity to join the project

- Contact to join IGGC and the HD SNP chip project :
gwenola.tosser@toulouse.inra.fr
- Website : www.goatgenome.org
- Illumina meeting at 12.30 today

Acknowledgements

- UNCEIA, Capgenes and Apis-gene (French breeding organizations)

I thank you for your attention