HAL
open science

Présentation et premiers résultats du programme 1000 génomes caprins : VarGoats

Philippe Bardou, Céline Bourdon, Laure Denoyelle, Thomas Faraut, Estelle Talouarn, Adriana A. Alberti, Mekki Boussaha, Stefan Engelen, Isabelle Palhière, Fabienne Le Provost, et al.

To cite this version:

Philippe Bardou, Céline Bourdon, Laure Denoyelle, Thomas Faraut, Estelle Talouarn, et al.. Présentation et premiers résultats du programme 1000 génomes caprins : VarGoats. Journées Scientifiques du Département de Génétique Animale, Oct 2018, Dienné, France. hal-03817570

HAL Id: hal-03817570

https://hal.inrae.fr/hal-03817570

Submitted on 17 Oct 2022

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.

Présentation et premiers résultats du programme 1000 génomes caprins : VarGoats

P Bardou[1], C Bourdon[2], L Denoyelle[3], T Faraut[1], E Talouarn[1], A Alberti[4], M Boussaha[2], S Engelen[4], I Palhière[1], F Le Provost[2], F Pompanon[3], C Robert-Granié[1], R Rupp[1], G Tosser-Klopp[1] et le Consortium VarGoats

[1] INRA, UMR 1388 GenPhySE, Castanet-Tolosan, [2] INRA, UMR 1313 GABI, Jouy-en-Josas, [3]UMR UGA-USMB-CNRS 5553, LECA, Grenoble, [4] Génoscope, Evry

CONTEXTE ET OBJECTIFS

Le projet **VarGoats** est un projet international de reséquençage de génomes caprins. Notre objectif est de caractériser la diversité génétique et de détecter des traces de sélection, à l'aide de l'étude de variants (SNP, CNV, variants de structure). Pour représenter la diversité génétique internationale, mais aussi caractériser finement les races françaises et les races africaines (projet **AGIN**), nous avons choisi 650 animaux en nous appuyant sur le projet **ADAPTmap** (données 50K). Nous avons incorporé en outre les chèvres marocaines et iraniennes (berceau de domestication) de **NextGen** (n≈180) et les données publiques (n≈200). Quelques races « inbred » et des chèvres sauvages ont été ajoutées. Les données sont analysées au sein de groupes de travail internationaux.

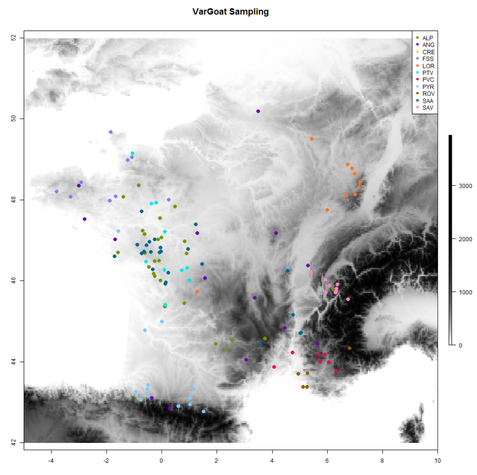
Nous présentons ici les résultats pilotes portant sur les chèvres françaises (n=200).

OBTENTION DES DONNEES

Echantillonnage, séquençage, ajout de données publiques

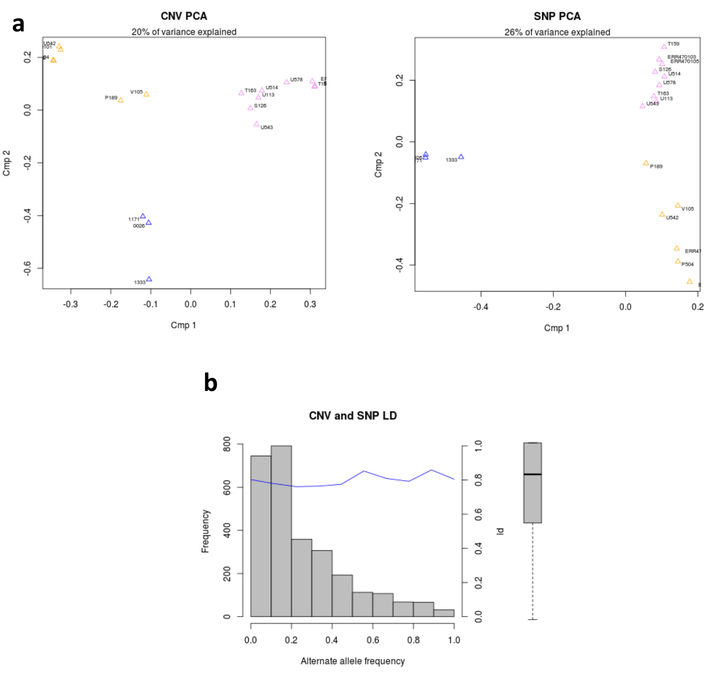
Alignement des fastq sur le génome ARS1 (BWA)
↓
fichiers bam
↓
Détection des variants (GATK)
↓
fichiers gvcf
↓
30 fichiers vcf
(un par chromosome)

Traitement bio-informatique

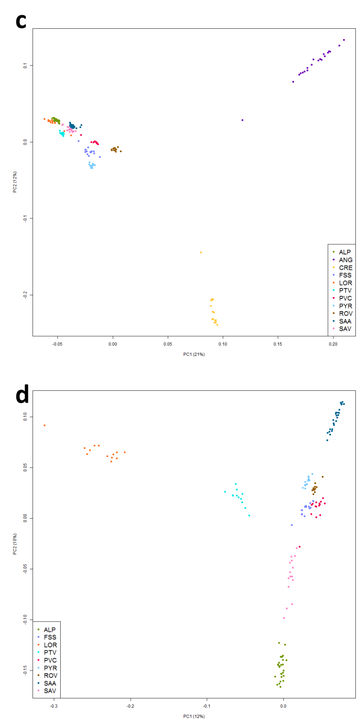


Analyse des chèvres françaises

PREMIERS RESULTATS



- Détection de CNV sur un jeu de données pilote
- La diversité génétique observée avec les CNV se superpose à l'organisation en races (**a**)
- Les CNV sont en fort déséquilibre de liaison avec les SNP flanquants (**b**)



- Créole et Angora sont très différentes des autres (**c**)
- Alpine, Savoie, Lorraine et Poitevine se distinguent (**d**)
- Fossés, Provençale, Pyrénnéenne et Rove sont groupées(**d**)

e

Génotypes 50k

2200 Alpine françaises 1358 Saanen françaises

Données de séquence filtrées

Alpine	Saanen
30	32

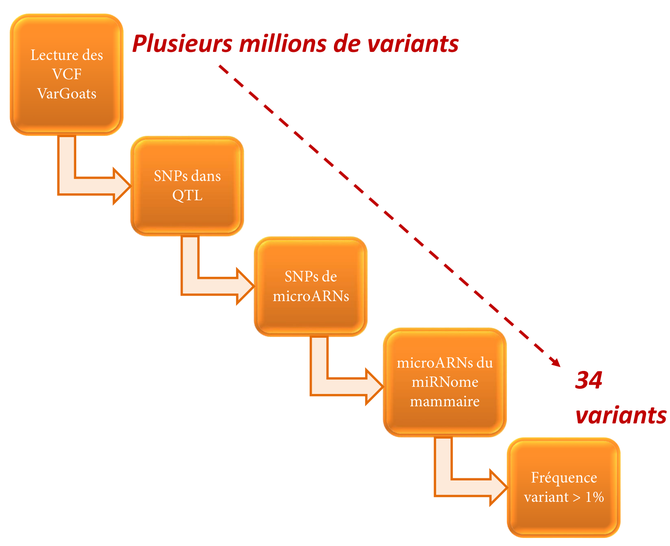
FImpute
Imputation intra-race (avec pedigree)

7 414 112 variants imputés

Alpine		Saanen	
CR	R	CR	R
88,48 %	0,81	87,82 %	0,80

- Imputation des données de génotypage 50K vers la séquence
- Plus de 7 millions de variants imputés avec de bons taux de concordance (CR) et corrélations (R) entre génotypes vrais et estimés (**e**)

f



- Recherche des variants génétiques de microARNs situés dans des régions de QTL laitiers et exprimés dans la glande mammaire des ruminants laitiers (programme miRQTLait)
- Chez les caprins, 34 variants (**f**) ont été trouvés et seront étudiés plus avant (rôle régulateur?).

CONCLUSIONS

Le développement de premiers pipelines d'analyse sur un nombre limité d'animaux français a permis de montrer la capacité de notre Consortium à analyser les données de séquence caprine. L'analyse des 1000 génomes (594 disponibles à ce jour) devrait nous permettre de générer des résultats originaux à l'échelle internationale.

Références : http://www.goatgenome.org/vargoats.html, http://www.goatadaptmap.org/, http://projects.ensembl.org/nextgen/, https://www.ars.usda.gov/office-of-international-research-programs/ftf-livestock-improvement/

Centre Occitanie-Toulouse
Chemin de Borde Rouge
CS 52627 • 31326 Castanet-Tolosan cedex
Tél : +33(0) 05 62 28 50 20

Alimentation Agriculture Environnement