

Le traitement des génotypages

Jusqu'à présent (année 2018), les génotypages sont réalisés uniquement sur les béliers collectifs, c'est-à-dire les béliers candidats à l'IA.

Utilisation de puces de moyenne densité 50k et basse densité 15k

Deux types de puces commercialisées par Illumina sont utilisés :

- la puce de moyenne densité 50k (54241 SNP);
- la puce LD 15k (15000 SNP).

Le matériel biologique est constitué d'échantillons de sang prélevés sur les agneaux (entre 1 et 6 mois selon la situation), puis sur les adultes passés à l'IA (à 1 an).

Deux laboratoires réalisent les génotypages : Labogéna et Aveyron Labo. Lorsque les génotypages sont effectués, les résultats sont envoyés dans le système d'information dédié. En 2016, les professionnels, dans le cadre du CNBL, ont décidé d'utiliser la plateforme Valogène comme système d'information des génotypages. Celle-ci avait l'avantage (i) d'être largement éprouvée par les bovins (les adaptations ovines ont été mineures) et (ii) d'être connectée au processus de traitement des génotypages développée par l'UMT eBIS à Jouy-en-Josas. Cette chaîne de traitement (qui inclut notamment l'imputation), conçue multi-espèce (avec un paramétrage en fonction de l'espèce), permet l'utilisation de plusieurs puces de densités et de versions différentes. Les laboratoires de génotypages déposent donc les résultats sur un espace FTP dédié, hébergé au CTIG de Jouy-en-Josas. Les différentes règles de gestion d'entrée en base de données sont alors appliquées puis les génotypages sont stockés dans la base Valogène.

Les génotypages bruts sont stockés avec comme identifiant le numéro d'échantillon (*sampleid*) qui contient le numéro figurant sur les tubes de sang (*idecha*).

Ex. *sampleid* : WG5884771-DNA_A10_COL602024

Ex. *idecha* : COL602024

La correspondance entre le numéro idecha et l'identifiant officiel de l'animal est fournie à la fois par le client et par le laboratoire.

Gestion des génotypages par la plateforme mutualisée Valogène

Avant une indexation, si de nouveaux génotypages sont entrés dans la base Valogène, l'UMT GPR demande un traitement des génotypages qui consiste à réaliser différentes étapes (vérification de la qualité des génotypages, vérification de la paternité, propositions d'assignation pour les animaux dont la paternité annoncée n'est pas compatible, imputation, phasage, mise en forme définitive et dépôt dans la BDIR). Ces différentes étapes sont décrites ci-après :

- **Contrôle des fichiers déclarés et apportés** : un échantillon génotypé doit avoir fait l'objet d'une commande de génotypages par le client, un échantillon ayant fait l'objet d'une information par le laboratoire d'analyse doit ensuite avoir un génotypage, les animaux génotypés doivent être présents dans la base de données génétique SIEOL.
- **Contrôle du CR** (call rate = % de SNP génotypés / nombre de SNP total) global et par chromosome. Le CR global est fourni par le laboratoire ; le typage est rejeté si le CR global est inférieur à 95%. Le CR par chromosome est recalculé en fonction de la puce sur les 38696 SNP ; si au moins 1 chromosome a un CR inférieur au seuil de 50%, la totalité du typage est rejetée.
- **Contrôle race et sexe biologique**. Le contrôle du sexe biologique n'est pas réalisé en ovin à ce jour.
- **Contrôle de Filiation**. Le test est réalisé sur 227 SNP du panel de contrôle de filiation. Le système compte le nombre d'erreurs mendéliennes entre un couple animal-typage et son père-typage ainsi qu'avec sa mère-typage en fonction des informations disponibles. Si on compte plus de 2% des 227 marqueurs incompatibles, la filiation est déclarée incompatible.
- **Assignation éventuelle** : en cas de contrôle de filiation incompatible, le système effectue une recherche des parents (en l'occurrence des pères) potentiels compatibles. La liste des ovins « assignés » est fournie aux Entreprises de Sélection qui peuvent ainsi corriger la paternité dans SIEOL.
- **Cohérence et consolidation** : il s'agit de ramener à un seul typage lorsque plusieurs typages existent pour un même animal.
- **Imputation/Phasage**. L'imputation consiste à mettre au même niveau les différentes densités. Il s'agit de combler les informations manquantes à partir de ce que l'on peut observer au sein d'une famille et/ou au sein d'une population. Les informations manquantes sont, soit les SNP d'une puce qui ne donnent pas de résultats au laboratoire, soit les SNP placés sur une puce mais pas sur une autre. En ovin lait, pour l'ensemble des races, la densité vers laquelle on impute est constituée de 38696 SNP, choisis parmi les 54241 de la puce 50k en tenant compte des critères suivants :
 - Call Freq (% d'animaux génotypés pour un SNP) supérieur à 0,97
 - MAF (fréquence de l'allèle mineur) supérieure à 0,01
 - en équilibre de Hardy Weinberg
 - non impliqués dans trop de problèmes de filiation

L'indexation ssGBLUP utilise un fichier de géotypages recodifiés en 0,1,2.

Ex. : FR20000399010098 38696 10121220010etcetcetc

- FR20000399010098 est le n° officiel de l'animal ;
- 38696 est le nombre de SNP
- 10121220010etcetcetc est le géotypage pour les 38696 SNP, chaque position représentant 1 SNP (0 = allèle absent ; 1 = hétérozygote ; 2 = allèle présent 2 fois)

Le phasage consiste à déterminer l'allèle issu du père et l'allèle issu de la mère. Les phases sont produites mais pas utilisées pour l'évaluation ovine laitière.

- Dépôts en BDIR. Les fichiers de géotypages et de phase recodifiés sont déposés dans la BDIR, de même qu'un fichier de suivi dans lequel les différentes étapes décrites ci-dessus sont formalisées pour chaque échantillon.

Environ 4000 béliers géotypés tous les ans

Le tableau suivant présente le nombre total de géotypages ovins présents dans la base Valogène, et pour le millésime de béliers 2017 : le nombre total de béliers, le % de d'échantillons avec qualité insuffisante (CR) et le % de paternité incompatible

Nb total géotypages	Nb géotypages réalisés en 2017	% problèmes qualité en 2017	% paternités incompatibles en 2017
27769	3776	1,0%	4,3%

Le délai moyen en 2017 entre la prise de sang et la diffusion des index des agneaux était de l'ordre de 45 jours.