

Des puces à ADN pour décrypter le génome des bovins



Les progrès récents de la génomique doivent beaucoup au formidable essor des biotechnologies et notamment au développement des technologies de génotypage à haut débit qui permettent dorénavant d'analyser simultanément jusqu'à plusieurs millions de positions sur le génome. Lorsque les biotechnologies rencontrent les nanotechnologies, les possibilités d'analyses explosent et les coûts d'analyses baissent. En génomique bovine, plusieurs outils regroupés sous le terme de « biopuces » permettent d'explorer le génome et de repérer les meilleurs reproducteurs. Prêts pour un saut (de puce) dans le monde du tout petit ?

Plusieurs outils de génotypages en fonction du nombre de SNP à analyser

Le principe récurrent de ces puces de génotypage consiste à augmenter le nombre de sondes (c'est-à-dire le nombre de SNP analysables) tout en réduisant les volumes réactionnels de manière à réduire les coûts d'analyses. Plusieurs technologies de génotypage ont été développées dans ce sens ; l'outil adapté sera choisi en fonction du nombre de SNP (de 1 à 800.000 !) et du nombre d'échantillons que l'on veut analyser simultanément.

1 - De 1 à 3.072 SNP (3K), des outils dédiés, construits « à façon » : ces technologies permettent de choisir spécifiquement les marqueurs à analyser pour construire une biopuce dédiée à une application particulière.

Avantages : une puce flexible et souvent peu chère à développer.

Inconvénients : un nombre limité de marqueurs spécifiques pour une utilisation limitée. Seuls les SNP choisis sont analysés. Souvent ces outils sont rentables si un grand nombre d'échantillons est analysé.

2 - De 7.000 à 800.000 marqueurs, des puces ADN à moyenne et haute densité qui couvrent l'ensemble du génome : ces outils, initialement développés avec l'aide de consortium internatio-

naux, permettent d'analyser simultanément plusieurs milliers, voire plusieurs centaines de milliers de SNP répartis sur l'ensemble du génome d'un animal. La société américaine Illumina, leader dans le domaine, commercialise une gamme de 3 puces de génotypages dites « pangénomiques » permettant d'analyser 7.000 marqueurs (puce « LD » 7K), 54.000 marqueurs (puce 54K) et 777.000 marqueurs (puce « HD » 800K) voir tableau récapitulatif au dos.

Autre avancée qui devrait améliorer ces outils dans l'avenir : la possibilité de rajouter des panels de SNP supplémentaires sur les puces 7K et 54K de manière à créer des outils adaptés, flexibles et performants. Ainsi, grâce à ces technologies, il sera possible sur un même outil d'évaluer le potentiel génétique d'un animal, de contrôler les filiations ou de détecter des anomalies génétiques potentielles dès sa naissance !

La société américaine Affymetrix commercialise également une puce de génotypage à haut débit permettant l'analyse de 648.000 SNP ; cependant, l'utilisation de cette puce, principalement dédiée à des programmes de recherche, reste relativement marginale.

Pourquoi parle-t-on de « puce » à ADN ?

Ce terme de « puce » à ADN (appelé également « biopuce » ou « réseau ») provient d'une analogie avec les circuits intégrés ou (« chips » en anglais) qui miniaturisèrent les transistors dans les années 1970. A l'instar de ces « puces » électroniques on regroupe sous le terme de « puce » tout un tas d'outils miniaturisés visant à placer sur un support inerte de quelques mm (lame de verre, plastique, silicium...) des milliers, voire des millions de molécules bioactives ! Dans le cas des puces à ADN, ces molécules, appelées sondes sont de courts fragments d'ADN. Mais il existe également des « puces » à ARN, à protéines, à anticorps, ...

Imputation ? kézako ?

L'imputation est un processus mathématique qui permet de prédire statistiquement des informations manquantes d'un ou plusieurs marqueurs moléculaire en analysant les génotypes obtenus aux SNP voisins. En extrapolant ce principe, et à condition d'avoir une population d'animaux de référence suffisamment grande, il est possible d'imputer de manière précise les génotypes des 54.000 SNP de la puce 54K à partir du résultat des 7.000 SNP de la puce 7K !

Cette opération permet notamment d'appliquer des équations génomiques obtenues à partir d'une population d'animaux génotypés avec la puce 54K pour des animaux génotypés « seulement » avec la puce à faible densité. Le même travail est réalisable de la 54K à la 800K.



10 millions de dollars de récompense !

Le décryptage du génome reste un enjeu majeur quelque soit l'espèce concernée. Une fondation américaine a doté un concours d'un prix de 10 millions de dollars pour la première société capable de réséquer 100 génomes humains en moins de 30 jours, et pour moins de 1.000\$ chacun ! De quoi susciter la motivation des 8 sociétés actuellement en course...



	Puce « LD » 7K	Puce 54K	Puce « HD » 800K
Date de sortie	2012	2008	2011
Description	Initialement déclinée sous la forme d'une puce 3K (2011) dédiée à l'industrie laitière, elle se décline maintenant sous la forme d'une puce LD (pour « Low density ») à 7.000 SNP.	Le best-seller d'Illumina, cette puce 54K est devenue l'outil de référence pour la quasi-totalité des équipes de génomique bovine dans le monde.	Un outil à très haut-débit qui permet d'explorer le génome plus finement. Puce HD pour « high density ».
Utilisation	Permet de génotyper un plus grand nombre d'animaux tout en réduisant le coût de l'analyse. Outil récent qui devrait s'imposer comme outil de routine.	Outil polyvalent : à la fois outil de recherche et outil d'évaluation génomique. Il est notamment utilisé pour l'outil EvaLiM® d'IngenomiX.	Outil dédié à la recherche principalement : recherche fine de QTL, recherche de mutations causales, étude des liens entre différentes races.
Avantages	Prix compétitif Densité adaptée pour réaliser une imputation efficace de la 7K → 54K.	Rapport densité SNP/prix Choix et répartition des SNP sur le génome Possibilité de faire une imputation 54K → 800K.	Le nombre de marqueurs analysés : support le plus dense actuellement disponible pour le génome bovin.
Inconvénients	Le nombre de marqueurs analysés.	Le prix qui reste élevé malgré une baisse significative ces derniers mois.	Le prix qui reste très élevé pour une utilisation autre que de la recherche.
Nb SNP	6.909 SNP	54.609 SNP	777.962 SNP
Nb échantillons	24 échantillons	24 échantillons	8 échantillons
			

Le reséquençage à haut-débit, l'outil ultime de génotypage ?

En 2007, le séquençage des 2.87 milliards de bases (Gb) constituant le génome de Dominette (la jeune vache de race Hereford dont le génome a été entièrement décrypté) a coûté 42 millions d'euros et 6 ans de travail avec la mobilisation de plus 300 chercheurs dans 25 pays. Ces travaux de reséquençage, ainsi que la recherche de polymorphisme associée sont à la base du développement de la puce ADN 54K d'Illumina.

Quelques années plus tard, le décryptage des 3 milliards de bases qui constituent le génome d'une vache nécessite une seule machine et l'expérience prend tout au plus deux semaines pour un coût total de moins de 200.000 €. La vitesse de progression dans le domaine est vertigineuse et le seuil psychologique de 1.000 dollars (environ 800€) pour le séquençage d'un génome complet devrait être atteint dès 2012 avec l'arrivée des séquenceurs de 3^e génération.

Compte tenu de ce coût, la question se pose déjà de savoir s'il n'est pas préférable d'investir quelques euros supplémentaires pour disposer de la séquence complète du génome (c'est-à-dire les 3 milliards de bases) au lieu des quelques 800.000 génotypes actuellement accessibles avec les outils de génotypages les plus performants. En effet, l'accès à la séquence complète d'un animal permettra d'accéder à l'exhaustivité de son information génomique, et ainsi d'identifier plus finement les gènes et les mutations causales impliquées dans la régulation des caractères intérêt.

En 2012, les séquenceurs de 3^e génération qui vont arriver sur le marché avec des annonces de coût de 1.000\$ le génome et des analyses qui auront lieu à l'échelle d'une seule base d'ADN. De fait, cette progression s'accompagne d'une compétition effrénée

entre les sociétés de biotechnologies qui ont toutes développées leur gamme de robot avec chacun sa particularité en termes de type de lecture, de type d'utilisation (génome complet, fragment de génome, génome bactérien et autres) et avec des technologies différentes. Et même si le séquençage systématique de tous les reproducteurs n'est pas (encore) à l'ordre du jour, de plus en plus de projets ambitieux de reséquençage sont actuellement en cours dans le monde. C'est notamment le cas du projet 1.000 génomes bovins dont l'objectif est de re-séquencer un grand nombre d'animaux de différentes races dans le monde.

