HAL open science

Mise en place d'évaluations génomiques dans les races bovines laitières à effectifs limités en France

Chris Hoze, Sebastien Fritz, Aurélia Baur, Florence Phocas, Didier Boichard, Vincent Ducrocq, Pascal Croiseau

To cite this version:

Chris Hoze, Sebastien Fritz, Aurélia Baur, Florence Phocas, Didier Boichard, et al.. Mise en place d'évaluations génomiques dans les races bovines laitières à effectifs limités en France. 21. Rencontres Recherches Ruminants, Dec 2014, Paris, France. hal-01194024

HAL Id: hal-01194024

https://hal.science/hal-01194024v1

Submitted on 2 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.

Mise en place d'évaluations génomiques dans les races bovines laitières à effectifs limités en France

HOZE C. (1), FRITZ S. (1), BAUR A. (1), PHOCAS F. (2), BOICHARD D. (2), DUCROCQ V. (2), CROISEAU P. (2)
(1) Union Nationale des Coopératives agricoles d'Elevage et d'Insémination Animale, Service Génétique, 149 rue de Bercy, 75595 Paris Cedex 12, France
(2) INRA, UMR 1313 Génétique Animale et Biologie Intégrative, 78352 Jouy-en-Josas, France

RESUME
Les évaluations génomiques sont aujourd'hui disponibles pour les principales races bovines laitières françaises. Cependant, construire une population de référence, ensemble des individus génotypés avec performances, de taille suffisante pour mettre en place des évaluations génomiques précises, s'avère un réel challenge pour les races laitières à effectifs limités. Pour contourner cette contrainte et développer des évaluations génomiques pour les races régionales, il a été initialement proposé de créer une population de référence commune entre races. Cette approche nécessite cependant, de disposer d'une densité en marqueurs suffisamment élevée pour que le lien QTL-marqueur soit conservé entre races. Les premiers travaux d'évaluations génomiques multiraciales ont démontré qu'une population de référence multiraciale génotypée sur une puce haute densité permettait d'augmenter la précision des valeurs génétiques estimées. Cependant, le gain observé est très dépendant de la proximité des races incluses dans la population de référence multiraciale. Ceci suggère que les évaluations multiraciales ne sont pas l'unique solution à la mise en place d'une sélection génomique dans les races à effectifs limités. D'autres stratégies basées à la fois sur l'augmentation de la population de référence intra-race (grâces aux échanges internationaux et au génotypage de femelles) et l'identification de mutations causales sont étudiées afin de mettre en place un prototype d'évaluations génomiques adapté à chacune des races régionales.

Implementing genomic evaluations in breeds with limited population size

HOZE C. (1), FRITZ S. (1), BAUR A. (1), PHOCAS F. (2), BOICHARD D. (2), DUCROCQ V. (2), CROISEAU P. (2)
(1) Union Nationale des Coopératives agricoles d'Elevage et d'Insémination Animale, Service Génétique, 149 rue de Bercy, 75595 Paris Cedex 12, France

SUMMARY
Within-breed genomic selection is now implemented in a number of large cattle breeds. However, building large enough reference populations (ie animals with both phenotypes and genotypes) remains a major challenge for smaller breeds. Combining reference populations and implementing a multi-breed genomic evaluation initially appeared to be an appealing alternative for small breeds. Such an approach requires a chip dense enough to maintain associations between QTL and markers. The first results of multi-breed genomic evaluations using high-density genotypes show that using a multi-breed reference population slightly increases the accuracy of genomic estimated breeding value. However the gain in accuracy was rather low suggesting that multi-breed reference population cannot be the unique solution to implement genomic selection in breeds with a limited population size. An alternative strategy is to increase within-breed reference population size through international exchange and female genotyping. The use of results from whole genome sequencing is also an appealing tool to improve the accuracy of genomic evaluation in breeds with limited reference population size.

INTRODUCTION

Les évaluations génomiques reposent sur l'estimation des effets des régions du génome influençant un caractère quantitatif (QTL). Pour envisager la mise en place d'évaluations génomiques, il est nécessaire de constituer une population dite de référence (c'est-à-dire un ensemble suffisamment grand d'animaux génotypés avec performances) à partir de laquelle les effets des QTL sont estimés. Dans les races nationales (Holstein, Montbéliarde et Normande), le nombre important de taureaux à la fois testés sur descendance et génotypés a permis la mise en place d'évaluations génomiques officielles dès 2009 (Fritz et al., 2010). Cette approche est optimale car les index sur descendance, disponibles pour de nombreux caractères, se comportent comme des phénotypes à forte héritabilité.
Dans les races laitières régionales, le nombre plus limité de taureaux testés sur descendance rend difficile la constitution de cette population de référence et empêche le développement des évaluations génomiques suivant la même approche.
Une solution très attractive serait de mutualiser les populations de référence entre races. Cette solution permettrait aux races régionales de bénéficier de l'information apportée par les populations de référence de grande taille disponibles dans les races nationales. Elle devrait également constituer la stratégie optimale pour la sélection génomique dans toutes les populations en particulier pour des caractères difficiles à évaluer comme l'efficacité alimentaire.
Mais cette solution n'est pas unique et en parallèle, plusieurs dispositifs ont été mis en place pour augmenter la taille des populations de référence des races régionales. Il devient alors envisageable de mettre en place une évaluation génomique intra-race.
Nous décrirons ici les travaux de l'UMT3G (Gestion génétique et génomique des populations bovines, coordonnant les efforts de l'INRA, l'UNCEIA et Idele) pour développer des évaluations génomiques dans les races régionales.

1. UNE EVALUATION GENOMIQUE MULTIRACIALE

1.1 PRINCIPE

Mutualiser les populations de références entre races fait l'hypothèse que les QTL (ou tout au moins une fraction importante de ceux-ci) soient communs entre races. Cette hypothèse fait actuellement l'objet d'études qui ne seront pas développées ici, mais les résultats actuels sont assez favorables. Une seconde hypothèse est que l'association

observée dans une race entre un marqueur et un QTL soit conservée dans les autres races.
Une telle association entre QTL et SNP existant dans la population ancestrale est conservée dans les races actuelles si les recombinaisons accumulées au fil des générations entre ces deux loci ne l'ont pas détruite. Elle sera donc conservée entre races si celles-ci ont divergé récemment et/ou si la distance entre QTL et marqueur est suffisamment faible.
Dans le cadre du programme ANR GEMBAL (Génomique Multiraciale des Bovins Allaitants et Laitiers), il a été proposé de génotyper les animaux avec une puce haute-densité (HD) pour laquelle la distance moyenne entre deux marqueurs est de 4 kilobases (kb). Cette puce était supposée suffisamment résolutive pour mettre en évidence les segments chromosomiques conservés entre races, de l'ordre de 10 à 20 kb. D'un point de vue économique, il n'était pas envisageable de typer toutes les populations de référence avec la puce HD. Ce typage a été réalisé sur un échantillon de 99 à 773 individus par race, sélectionnés sur la base de leur contribution génétique. Dans chacune des races, les génotypes HD des animaux génotypés sur la puce 50k ont été imputés. Statistiquement il a été montré que cette imputation intra-race est très précise dès lors que 200 génotypes HD sont disponibles (Hozé et al, 2013).
Une comparaison de la précision des évaluations génomiques multiraciales et des évaluations intra-race a ensuite été réalisée en utilisant les génotypes imputés HD des populations de référence des principales races laitières.
Par ailleurs, Il a été démontré (Brondum et al., 2011) que le bénéfice des évaluations multiraciales était plus important dans le cas de races présentant une origine commune récente. Ainsi en race Simmental, les résultats ont été comparés en utilisant une population de référence Simmental uniquement, une population de référence Montbéliarde ou une population de référence conjointe.

1.2 CAS GENERAL

Initialement seules trois races nationales disposaient d'une population de référence de taille suffisante pour réaliser des études de la précision des évaluations génomiques. Les premiers travaux d'évaluation multiraciale ont donc été réalisés sur les populations de référence des trois grandes races laitières (Hozé et al., 2014a), pour simuler des races à effectifs plus réduits.

1.2.1 Matériel et méthodes

Dans chacune de ces races, plus de 500 individus ont été génotypés sur la puce HD. Le taux d'erreur allélique lié à l'imputation depuis la puce 50K est inférieur à 0,73% (Tableau 1). Les populations de référence génotypées sur la puce 50K ont été imputées vers la puce HD.

Tableau 1 : Constitution des populations de référence HD

Race	Nombre de génotypes HD	Population HD après imputation	Taux d'erreur d'imputation estimé
Montbéliarde	527	1788	0,51 %
Holstein	773	4989	0,73 %
Normande	535	1991	0,33 %

La précision des évaluations polygéniques, génomiques intra-race et génomiques multi-races a été comparée en utilisant les 394 plus jeunes animaux normands comme population de validation (V). Cet effectif est suffisant pour éviter des fluctuations liées à un échantillon trop réduit. La population restante (dite d'apprentissage) a été utilisée, dans sa totalité ou en partie, pour estimer les effets des marqueurs. Trois scénarios ont été testés pour simuler le cas d'une race disposant d'une population de référence de 1597 individus, de 404 individus ou 198 individus. En cas d'évaluation multi-raciale, les taureaux Holstein et Montbéliards sont également pris en compte en totalité.
Les évaluations génomiques ont été réalisées en utilisant une approche statistique bayesienne, dite BayesC, où l'on considère que seule une partie des marqueurs a un effet sur le caractère. Les phénotypes utilisés correspondaient aux déviations moyennes des filles d'un taureau (DYD) pour les caractères : quantité de lait, de matière grasse, de matière protéique, le taux butyreux, le taux protéique et le score cellulaire.

1.2.2 Résultats

Le tableau 2 présente les corrélations moyennes obtenues entre valeurs génétiques (ou génomiques) estimées (EBV) et les phénotypes (DYD) pour les 394 animaux de validation.
Dans le cas d'une population avec une population de référence de taille importante, une évaluation multiraciale ne permet pas d'augmenter la précision des valeurs génétiques estimées. En revanche, un gain est observé dans le cas d'une population de référence de taille réduite.
Il a également été observé, que ce gain de précision était dix fois plus élevé pour les animaux dont le père est absent de la population de référence (+0.141) que pour les animaux dont le père est présent (+0.016).

Tableau 2 : Corrélation moyenne sur les 6 caractères entre valeurs génétiques estimées (EBV) et performances moyennes des filles d'un taureau (DYD) pour les 394 taureaux V en fonction du scénario testé et du type d'évaluation génétique.

Taille de la population de référence (apprentissage)	Type d' évaluation	Corrélation (gain par rapport au polygénique)
1597 taureaux	Polygénique	0,357
	Génomique intra-race	0,550 (+ 0,193)
	Génomique multi-race	0,566 (+ 0,209)
404 taureaux	Polygénique	0,311
	Génomique intra-race	0,387 (+ 0,076)
	Génomique multi-race	0,416 (+0,105)
198 taureaux	Polygénique	0,311
	Génomique intra-race	0,368 (+ 0,057)
	Génomique multi-race	0,397 (+0,086)

1.3 CAS DE LA RACE SIMMENTAL

Les évaluations génomiques intra-race et multi-races ont également été comparées dans le cas de deux races proches : les races Simmental Française et Montbéliarde (Hozé et al., 2014b). Ces deux races sont évaluées conjointement pour la production laitière dans le cas de l'évaluation internationale Interbull Simmental, ce qui permet de disposer du même phénotype dans les deux races.

1.3.1 Matériel et méthodes

Parmi la population de référence Simmental initialement constituée de 229 taureaux (dont 123 dont le génotype a été échangé avec le LfL, Allemagne), les 46 animaux les plus jeunes ont été utilisés comme population de validation.
Quatre approches ont été comparées pour l'estimation des valeurs génétiques et génomiques (EBV) des animaux de validation : 1) Une évaluation polygénique, 2) Une évaluation génomique utilisant la population de référence Simmental seule, 3) Une évaluation génomique utilisant la population de référence Montbéliarde, 4) Une évaluation génomique utilisant la population de référence des deux races.

Les évaluations génomiques ont été réalisées en utilisant l'approche BayesC. Par ailleurs, l'impact de la densité en marqueurs sur les évaluations a été étudié en utilisant les génotypes issus de la puce HD ou de la puce 50K.
Les phénotypes utilisés ici correspondent aux index dérégressés (DRP) issus des évaluations génétiques internationales dans l'échelle Simmental. Les caractères étudiés sont les quantités de lait, de matière grasse et de matière protéique, et le score cellulaire.

1.3.2 Résultats

Le tableau 3 présente les corrélations obtenues entre EBV et DRP pour les animaux de la population de validation. Les résultats présentés correspondent aux évaluations génomiques utilisant des génotypes 50K. Les corrélations observées dans le cas où les évaluations génomiques utilisent des génotypes HD étant très similaires, elles ne sont pas présentés ici.
En moyenne une évaluation génomique utilisant une population de référence Simmental et Montbéliarde permet un gain moyen de corrélation de 0.09. Ce gain est très supérieur à ce qui est observé dans le cas Montbéliarde, Normande, Holstein.
Ces résultats confirment que les évaluations multiraciales représentent une stratégie attractive dans le cas de deux races proches. Par ailleurs, la présence d'animaux Simmental (même en faible nombre) est indispensable à une bonne précision des évaluations génomiques.
Enfin on observe ici encore que le gain de précision permis par les évaluations génomiques multiraciales est beaucoup plus élevé pour les animaux dont le père est absent de la population de référence (+0.14) que pour ceux dont le père est présent (+0.05).

Tableau 3 : Corrélation moyenne sur les 4 caractères entre valeurs génétiques estimées (EBV) et performances moyennes des filles d'un taureau (DYD) pour les 46 taureaux V Simmental en fonction du type d'évaluation génétique.

Type d évaluation	Population de référence	Corrélation
Polygénique	Simmental	0,330
Génomique	Simmental	0,390
Génomique	Montbéliarde	0,280
Génomique multi-race	Montbéliarde + Simmental	0,480

1.4 BILAN SUR LES EVALUATIONS MULTIRACIALES

Dans les deux cas étudiés ici, la mise en place d'une évaluation génomique multiraciale permet d'augmenter la précision des évaluations génomiques. Cependant le gain observé est relativement limité en particulier pour les animaux fortement apparentés à la population de référence.
De plus, la précision des évaluations est à peine supérieure avec 198 taureaux Normands et 6777 taureaux Montbéliarde et Holstein qu'avec 404 taureaux Normands. Il est donc bénéfique, même dans le cas d'une évaluation génomique multiraciale de chercher à maximiser la taille des populations de référence intra-race.
Enfin, le gain de précision lié aux évaluations génomiques multiraciales est largement supérieur dans le cas Montbéliarde-Simmental, ce qui suggère que les évaluations multiraciales constituent une stratégie intéressante dans le cas de races génétiquement proches.

2. UNE EVALUATION GENOMIQUE INTRA-RACE

2.1 AUGMENTATION DE LA TAILLE DES POPULATIONS DE REFERENCE

Plusieurs approches sont envisageables pour augmenter la taille des populations de référence.

2.1.1 Génotypage complémentaire de mâles

L'approche la plus évidente consiste à réaliser des génotypages complémentaires de taureaux. Il est ainsi possible de génotyper les animaux à leur entrée en testage sur descendance afin qu'ils intègrent la population de référence à l'arrivée des performances de leur filles. Cependant le nombre réduit d'animaux testés chaque année dans les races régionales ne permettra pas d'augmenter fortement la taille des populations de référence.

2.1.2 Echanges internationaux

En race Holstein, le consortium européen EuroGenomics a été réalisé afin de mutualiser les populations de référence et améliorer la précision des évaluations génomiques. Les résultats ont été très favorables puisqu'il a été démontré que le passage de 4000 taureaux nationaux à 16000 taureaux européens permettait d'augmenter le CD des index génomiques de 0.10 (Lund et al, 2011).
Certaines races régionales françaises sont de grandes races à l'échelle internationale. Il est donc envisageable, de la même façon qu'en race Holstein, de réaliser des échanges internationaux afin d'augmenter la taille des populations et mettre en place une évaluation génomique précise en France.

2.1.3 Inclusion de femelles dans les populations de référence

Lorsque le nombre de taureaux testés sur descendance est limité, il est possible de génotyper des femelles avec performances dont plusieurs milliers sont disponibles afin de compléter les populations de référence des races régionales. C'est une solution plus coûteuse mais théoriquement possible. La quantité d'information apportée par une femelle avec performances propres dépend de l'héritabilité du caractère. En utilisant la formule de Daetwyler et al. (2013) on peut estimer que pour apporter la même information qu'un taureau disposant d'un index sur descendance de CD 0,9, il faut entre 2 et 10 femelles, pour un caractère avec une héritabilité comprise entre 0,5 et 0,1. Dans les races régionales, l'inclusion de femelles dans les populations de référence peut donc permettre une forte augmentation de la taille de la population de référence et de la précision des évaluations génomiques.

3. DISCUSSION ET PERSPECTIVES DE DEVELOPPEMENT D'EVALUATIONS GENOMIQUES DANS LES RACES REGIONALES

3.1 UNE SOLUTION ADAPTEE A CHACUNE DES RACES ET NON UNE SOLUTION UNIQUE

Plusieurs stratégies sont donc possibles pour développer des évaluations génomiques dans les races régionales. En revanche, toutes les solutions ne sont pas disponibles dans l'ensemble des races. Il est donc probable que la stratégie finalement adoptée diffère en fonction des situations.
La solution la plus générique, consiste à augmenter la taille de population de référence via des génotypages complémentaires de mâles et de femelles.
Les situations de la race Brune et des races de montagnes (Abondance, Tarentaise, Montbéliarde, Simmental, Vosgienne) seront traités ici. Nous décrirons également une troisième perspective basée sur l'utilisation de données de séquençage complet.

3.2 CAS DE LA RACE BRUNE

En race Brune, la Fédération Européenne de la race a initié avec Interbull le projet visant à mettre en commun tous les génotypes et créer une population de référence internationale. La participation de 7 pays (Autriche, France, Allemagne, Italie, Slovénie, Suisse et Etats-Unis) à ce consortium a permis de constituer une population de référence de 5736 taureaux Bruns dont 90 français.

Cette population aujourd'hui accessible à chacun des pays membres, a permis de développer, à l'été 2014, une évaluation génomique française (Baur et al. 2014). Le gain de corrélation permis par les évaluations génomiques est compris entre 5,2 et 26,8% selon les caractères. Ce chiffre est très supérieur à ce qui aurait été atteint en utilisant uniquement les 90 taureaux français puisqu'il est proche des gains de précision observés en 2010 en race Montbéliarde avec 1500 taureaux,

3.3 CAS DES RACES ABONDANCE, TARENTAISE, VOSGIENNE ET SIMMENTAL

Les études sur les évaluations génomiques multiraciales ont abouti à deux grandes conclusions : 1) les évaluations multiraciales permettent un gain de précision limité pour les animaux apparentés à la population de référence, 2) pour ces mêmes individus la précision des évaluations génomiques intra-race est correcte malgré la taille limitée de la population de référence.

Sur la base de ces résultats, de nombreux génotypages de mâles et de femelles ont été réalisés afin d'augmenter la taille de la population de référence. Le nombre de génotypages complémentaires disponibles dans chaque race en mars 2014 est présenté dans le tableau 4. Des discussions sont également en cours pour réaliser des génotypages complémentaires en race Simmental.

Tableau 4 : Nombre de génotypages complémentaires disponibles dans les races régionales en mars 2014

Race	Mâle	Femelle
Abondance	150	475
Tarentaise	150	382
Vosgienne	0	996

Des tests d'évaluations génomiques intra-race intégrant ces génotypages et utilisant, une population de référence à la fois mâle et femelle sont en cours. Les performances utilisées correspondront alors aux déviations (*Yield Deviation* :YD) pour les femelles et aux déviations moyennes des filles (DYD) pour les mâles.

3.3 UTILISATION DES DONNEES DE SEQUENCAGE COMPLET ET INTEGRATION DE MUTATIONS CAUSALES

En dépit d'une résolution élevée, la puce HD ne permet pas de transférer l'information d'une race vers une autre avec la précision souhaitée. Les travaux d'I. Van Den Berg et al (2014) montrent que la densité requise est encore supérieure (inférieure à 1 kb) pour conserver les associations marqueur-QTL entre races. Plutôt que de concevoir des puces correspondantes à très haute densité (au moins 5 fois plus que la puce HD), la stratégie en cours vise plutôt à rechercher les principales mutations causales impliquées dans le déterminisme des caractères sélectionnés.
Des travaux très importants sont donc en cours dans les grandes races pour identifier les principales mutations causales, en combinant informations de séquence de génome complet (Daetwyler et al, 2014) et populations de référence. Une fois ces mutations identifiées (ou bien des mutations en étroit déséquilibre de liaison avec elles, ayant le même effet apparent), on peut vérifier qu'elles ségrégent également dans d'autres races, en particulier les races régionales. Cette stratégie est facilitée par le développement de puces personnalisables dites « *custom* » permettant l'ajout de marqueurs spécifiques. Il est ainsi possible d'ajouter progressivement les mutations identifiées sur cette puce afin de disposer d'un génotypage exact aux mutations à moindre coût. Si ces mutations et leurs effets sont confirmés, leur prise en compte en évaluation génomique est simple dans le modèle d'évaluation génétique français incluant l'information de QTL. Il est donc probable que les modèles d'évaluation génomique du futur prendront davantage en compte l'information biologique sur le déterminisme génétique des caractères et les mutations en cause.

CONCLUSION

Du fait du nombre limité d'individus génotypés avec performances dans les races régionales, les évaluations génomiques ne sont disponibles actuellement que dans les races nationales. Cependant, la mise en place d'évaluations génomiques est souhaitable dans toutes les races et nous proposons plusieurs stratégies pour réussir cette évolution.
La solution envisagée initialement était de regrouper les populations de référence entre races et de les analyser avec une densité de marqueurs supposée suffisante (4 kb). Cependant, les travaux réalisés montrent que le gain de précision est limité lorsque les races sont génétiquement éloignées.
Plusieurs autres pistes, nécessitant pour la plupart d'augmenter la taille des populations de référence intra race sont envisageables. A court terme, les échanges internationaux et l'inclusion de femelles dans la population de référence semblent être les meilleures stratégies pour améliorer la précision des évaluations génomiques.
A moyen terme, les travaux de recherche autour de l'utilisation des données de séquençage complet et l'intégration de mutations causales dans les modèles d'évaluation devraient eux aussi permettre d'améliorer la précision des évaluations en levant la contrainte de la conservation de l'association entre QTL et marqueurs.
Il faut également garder à l'esprit que la sélection génomique peut être un cercle vertueux. Une fois un premier système d'évaluation génomique en place, les candidats à la sélection et finalement retenus, mâles et surtout femelles, seront intégrés à la population de référence, permettant sa croissance et son renouvellement. Le maintien du contrôle de performances est donc plus que jamais nécessaire pour augmenter les tailles de population de référence mais également pour maintenir le lien, essentiel à la précision des évaluations génomiques, entre population de référence et candidats à la sélection.

*Le projet GEMBAL (2010-2014) est un programme financé conjointement par l'ANR, APIS-GENE, Races de France et l'INRA dans le cadre du GIS AGENAE. Les travaux ont été réalisés dans le cadre de l'UMT3G (*Gestion génétique et génomique des populations bovines). Les génotypes HD Simmental allemands ont été fournis par LfL.

Baur A. *et al*., 2014. 21èmes Renc. Rech. Rum..
Brondum R. *et al*., 2011. J. Dairy Sci. 94:4700-4707.
Daetwyler H. et al, 2013. Genetics, 193, 347-365
Daetwyler H. et al, 2014. Nat Genet, 46, 858-867
Fritz *et al.,* 2010. 17èmes Renc. Rech. Rum., 455-458.
Hozé C. *et al*., 2013. Genet. Sel. Evol. 45, 33.
Hozé C. *et al*., 2014a. J. Dairy. Sci. 97 : 3918-3929
Hozé C. *et al*., 2014b. 10th WCGALP, 459
Lund M. *et al*., 2011. Genet. Sel. Evol. 43:43
Van Den Berg I. *et al*., 2014. 10th WCGALP, 490