THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Pathologie, Toxicologie, Génétique et Nutrition

Présentée et soutenue par :

Mme SAMIRA FATHALLAH

le mardi 29 septembre 2015

Titre :

VARIABILITES DE LA TAILLE DE PORTEE DES OVINS

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Génétique, Physiologie et Systèmes d'Elevage (GenPhySE)

Directeur(s) de Thèse :

M. LOYS BODIN

MME INGRID DAVID

Rapporteurs :

M. JUAN PABLO GUTIERREZ GARCIA, UNIVERSIDAD COMPLUTENSE DE MADRID

Mme ELISABETH LE BIHAN-DUVAL, INRA TOURS

Membre(s) du jury :

Mme FLORENCE PHOCAS, INRA JOUY EN JOSAS, Président

M. HERVE CHAPUIS, SYSAAF TOURS, Membre

M. LOYS BODIN, INRA TOULOUSE, Membre

Remerciements :

J'adresse mes remerciements au Coram (Collectif des Races des Massifs), et à l'INRA-GenPhySE pour le financement de cette thèse sur la canalisation de la taille de portée chez les ovins.

Je voudrais remercier à toutes les personnes qui, de prés, ou de loin, ont contribué à l'aboutissement de ce travail de thèse.

Tout d'abord, je tiens à remercier, Dr **Loys BODIN**, pour son soutien au long de mes travaux, sa présence. Il s'est montré toujours disponible et accueillant, il a su me faire évoluer grâce à ses remarques, et ses critiques constructives. J'ai eu beaucoup de chance de mener cette thèse sous sa direction, sans lui ce manuscrit ne serait jamais mené à terme. Un grand merci au Dr **Ingrid DAVID** d'avoir acceptée de co-encadrer cette thèse. Je vous suis particulièrement reconnaissante pour vos encouragements, vos conseils, et votre optimisme. Les échanges et les discussions que nous avons eues durant les réunions m'ont beaucoup appris.

Je tiens également à remercier les rapporteurs de cette thèse Mme Elisabeth LE BIHAN-DUVAL, Hervé CHAPUIS et Mr Juan Pablo GUTIERREZ GARCIA.

Merci aux membres de mon comité de thèse : **Laurianne CANARIO**, **Flavie TORTEREAU**, **Milagros GARCIA**, **Zulma VITEZIVA** et **Emeric JOUHET** d'avoir accepté d'y participer. Leurs remarques et idées ont permis d'approfondir certaines pistes de travail.

Je tiens à remercier l'ensemble des membres du département GenPhySE notamment l'ancien SAGA pour leur accueil.

Merci à mes collègues et mes amis (je ne citerai pas en risque d'en oublier) pour leurs soutiens, et les bonnes soirées que nous avons passé ensemble.

Merci aux personnes qui ont partagé mon bureau et particulièrement à Mme **Michel CLEMENT THAU,** ça fut un grand plaisir de t'avoir comme une collègue pendant un an et demi.

Merci à ma famille, et à mes amis pour leur soutien et encouragement. Enfin un merci tout particulier à mes parents pour leur soutien pour les petits et grand projets de vie.

Merci à nouveau à toutes les personnes que j'ai citées pour tout ce qu'elles m'ont appris durant ces trois années. MERCI

Table des matières

INTRODUCTION GENERALE 7

PARTIE 1 : Bibliographie 10

1 La prolificité – Les variabilités – La canalisation 12

1.1 La prolificité des brebis 12

1.1.1 Productivité numérique - prolificité : Contexte économique 12
1.1.2 Sélection de la prolificité 13
1.1.3 Facteurs de variation environnementaux 13
1.1.4 Facteurs de variation génétiques 14
1.1.5 Augmentation de la prolificité 14
1.1.6 Problèmes phénotypiques liés à l'augmentation de prolificité 14
1.1.7 Intérêt économique de la canalisation des tailles de portée des brebis 15

1.2 Variabilité de la distribution des tailles de portée ovine 17

1.2.1 Liaison moyenne variance 17
1.2.2 Variabilité entre races 17
1.2.3 Variabilité intra-race 18

1.3 La canalisation 18

1.3.1 Définitions 18
1.3.2 Pourquoi canaliser ? 20
1.3.3 Modèles génétiques 20
1.3.4 Mise en œuvre, réalisations 21

2 Modèles d'analyse 23

2.1 Rappel sur le modèle linéaire polygénique 23

2.2 Les modèles non linéaires polygéniques 24

2.2.1 Généralités 24
2.2.2 GLMM utilisés en génétique sur les données ordinales 25

2.3 Le modèle polygénique hétéroscédastique 26

PARTIE 2 : Modélisation des variabilités de la taille de portée des ovins comme une variable continue 28

1 Origine des données 30

1.1 Choix des races 30

1.2 Origine des données 30

1.3 Les échantillons complets 32

1.4 Les échantillons simplifiés 32

2 Modélisation de la "moyenne" : Modèle linaire mixte homoscédastique 34

2.1 Recherche des facteurs environnementaux du modèle. 34

2.2 Modèles génétiques 34

2.3 Résultats et discussion 35

Effet de l'âge de la brebis : 35

Effet de l'interaction âge à la première mise bas et la parité de mise bas : 35
Effet de l'intervalle entre agnelages : 35
Effet de la saison : 35
Effet de l'élevage : 36
Les facteurs de variation génétiques 36

3 Modélisation de la variance résiduelle : Modèle linéaire mixte hétéroscédastique 40

3.1 Résultats de la modélisation de la variance résiduelle 40

3.2 Conclusion 41

PARTIE 3 : Modélisation des variabilités de la taille de portée des ovins comme une variable ordinal 42

1 Modélisation de la "moyenne" : Modèle à seuils homoscédastique 44

1.1 Modèle à seuils homoscédastique 44

1.2 Résultats et discussion 44

2 Modélisation de la variance résiduelle : Modèle à seuils hétéroscédastique 45

2.2 Etude de la variabilité due à un gène majeur 47

2.2.1 Introduction 47
2.2.2 Article 1 : Relationship between the effects of the BMP15 gene and the polygenic effects on prolificacy in the Rasa Aragonesa sheep breed 48
2.2.3 Résultats majeurs obtenus sur l'article précédent 68

2.3 Etude de la variabilité due à plusieurs gènes 70

2.3.1 Introduction 70
2.3.2 Article 2 : Genetic heteroscedastic models for ordinal traits: Application to sheep litter size 70
2.3.3 Résultats complémentaires 99

2.4Discussion 101

RESUME 103

REFERENCES 105

Liste des tableaux

Tableau 1 : Effectif et prolificité des brebis adultes de 38 races contrôlées en France à la campagne 2013. (Données du Bilan de contrôle de performances ovin allaitant, campagne 2013 ; IE - INRA)... 16

Tableau 2 : Prolificité moyenne et variance des tailles de portée à la campagne 1982 ainsi que le pourcentage de chaque type de portée pour les 4 races les plus communes en Norvège. (d'après (Steine, 1985))..... 18

Tableau 3 : Différentes distributions..... 25

Tableau 4 : Effectif, prolificité naturelle, variance, coefficient de variation et pourcentages de chaque taille de portée pour les brebis adultes de 38 races contrôlées en France à la campagne 2013. (Données du Bilan de contrôle de performances ovin allaitant, campagne 2013 ; IE - INRA)..... 31

Tableau 5 : Caractéristiques descriptives des échantillons complets et réduits pour les trois races (Romane, Suffolk et Rouge de l'ouest)..... 33

Tableau 6 : Facteurs testés et retenus pour les analyses génétiques..... 34

Tableau 7 : Les estimations des paramètres génétiques et environnementaux de la fonction moyenne 37

Tableau 8 : Estimations des composantes de la variance du modèle sur la variabilité de la taille de portée et corrélations entre $\hat{u}$ et $\hat{v}$ 41

Tableau 9 : Estimations des composantes de la variance pour les différentes races par le modèle à seuils homoscédastique [TM]..... 44

Tableau 10 : Races ovines dans la quelle ségrége un gène majeur (extraites de la thèse (Mansanet, 2013)). 46

Tableau 11 : Corrélation calculée pour les diffèrent modèles ([TM], [HTM], [ITM]) entre la moyenne (variance) estimée et la moyenne (variance) observée..... 99

Tableau 12 : Variance génétique (calculée) sur la moyenne *u* et la variance de la sous-jacente *v* pour la Rouge de l'Ouest et pour la Suffolk. 99

Tableau 13 : Corrélation ρ entre les *u* et *v* estimées par le modèle à seuils homothétique [HTM]... 100

Table des figures

Figure 1 : Marge brute (hors prime) en fonction de la productivité numérique pour l'atelier Ovin allaitant des fermes de référence ovin viande, Productivité numérique en production ovine allaitante : composantes, intérêt économique, lien avec les facteurs de production et avec les aspects environnementaux ; Benoit, M. ; Laignel, G. ; Roulenc, M. Présentation orale : Assemblée Générale UPRA ROM Sélection, 2013. 12

Figure 2 : Evolution du taux de pertes au cours des premiers mois selon la taille de portée et la race (d'après Theriez, 1982) 15

Figure 3 : Distribution des variances des tailles de portée par race pour les données (n=2 399 063) contrôlées depuis la campagne 2000 de 36 races ovines françaises. (Les points en rouge correspondent à 3 races (Noire du Velay, Lacaune, Grivette) pour lesquelles un gène majeur d'ovulation est en ségrégation ; la courbe de tendance est estimée sans ces 3 races. (Données de la BDIR ovin allaitant, IE - INRA). 17

Figure 4 : Différents modèles génétiques expliquant la canalisation et les interactions GxE (tirée de (Bodin et al., 2010). 22

Figure 5 : Composantes de variance génétique, permanente et leur somme, ainsi que variance résiduelle et totale en fonction de la prolificité moyenne de la race. 36

Figure 6 : Influence de l'âge de la brebis sur la taille de portée des trois races (Suffolk, Rouge de l'Ouest, et Romane). 38

Figure 7 : Influence de l'âge à la première mise bas sur la taille de portée de la race Suffolk. 38

Figure 8 : Influence de l'âge à la première mise bas sur la taille de portée de la race Rouge de l'Ouest. 38

Figure 9 : Influence de l'âge à la première mise bas sur la taille de portée de la race Romane. 39

Figure 10 : Influence de l'intervalle entre agnelage sur la taille de portée des trois races (Suffolk, Rouge de l'Ouest, et Romane). 39

Figure 11 : Influence de la saison sur la taille de portée des deux races (Suffolk et Rouge de l'Ouest). 39

Figure 12 : Liaison moyenne – variance sur l'échelle observée pour différentes valeurs de la moyenne η de la variable sous-jacente et l'ensemble des seuils estimés pour chaque génotype R+ et ++. Quelle que soit la valeur de η (parmi les valeurs plausibles), la variance sur l'échelle observée des R+ est toujours supérieure à celle des ++ et notamment pour la même moyenne sur l'échelle observée. 68

Figure 13 : Liaison moyenne – variance sur l'échelle observée pour différentes valeurs de la moyenne η de la variable sous-jacente et l'ensemble des seuils estimés pour chaque génotype L+ et ++ (Martin et al., 2014). 69

Figure 14 : Relation moyenne-variance pour les trois modèles ([TM], [HTM], [ITM]) sur la race Rouge de l'Ouest (Prolif=1.89). 100

Figure 15 : Relation moyenne-variance pour les trois modèles ([TM], [HTM], [ITM]) sur la race Suffolk (Prolif=1.63). 100

Figure 16 : Relation moyenne-variance pour les trois races estimée par le modèle à seuils homothétiques ([HTM]). 102

INTRODUCTION GENERALE

La taille de portée des ovins a connu une augmentation remarquable ces dernières années due à l'amélioration de l'environnement au sens large ainsi qu'au progrès génétiquedégagé par la sélection. Cependant cette augmentation sans maitrise de la variabilité (sans« canalisation »)agénéré un certain nombre de défis associés aux agneaux nés multiples et principalement à ceux issus de portées supérieures à trois. Ces agneaux présentent en effet de plus faiblepoids à la naissance,de plus forte mortalitéet de moins bonne croissance; autant de facteurs qui se traduisent par des couts économiquessupplémentaires qui viennent contrebalancer l'intérêt des augmentations de prolificité.
Les recherches conduisant à modéliser la variabilité des caractères et qui ont connu un certain essor ces dernières années ne peuvent pas être appliqués directement au problème lié à la variabilité de la taille de portée des ovins à cause de la nature discrète du caractère.
Ce travail de thèse combine le domaine de la modélisation et de la génétique quantitative, il a pour ambition de mettre à la disposition des généticiens des modèles statistiques innovants répondant à la problématique liée à la nature discrète de la taille de portée chez les ovins.

La première partie de ce manuscrit consiste en une bibliographie succincte sur la prolificité des ovins et sur sa variabilitéainsi que sur l'intérêt de la canalisation; c'est-à-dire augmenter la moyenne vers un optimum, qui varie selon les systèmes de production,et diminuer la variabilité autour de cet optimum.
Nous présenterons aussi dans cette partie les différents modèles qu'il est possible d'utiliser sur ce type de données discrètes et notamment le modèle linéaire généralisé homoscédastique et le modèle linéaire généralisé hétéroscédastique.

L'homoscédasticité suppose que la variance des erreurs d'un modèle est la même quelle que soit l'observation considéréealors que l'hétéroscétasticité remet en cause cette hypothèse, et considère que cette variance d'erreur est affectée par des facteurs qui dans notre cas peuvent être environnementaux et génétiques. Notons qu'en général les modèles hétéroscédastiques ne modélisent pasdirectement la variance résiduelle de façon indépendante de son estimation par un modèle homoscédastique. Ce qui est modélisé n'est donc pas véritablement la variance résiduelle d'un premier modèle et l'on doit plutôt considérer que l'on a deux modèles emboités; l'un concerne le caractèrelui-même, et l'autre ce qui est **souvent appelée la variance environnementale**. Le fait qu'une partie d'un modèle hétéroscedastique concerne des termes de variance, fait qu'on parle du **modèle sur la variance** et **par abus de langage on parle aussi du modèle sur la moyenne** alors que ce n'est pas la moyenne des observations qui est traitée mais bien les observations individuelles.
Dans cette thèse, nous considérons d'une part la variabilité du caractère taille de portée et d'autre part la variabilité de sa variance environnementale. C'est pour cela que **nous parlons DES variabilités** de la taille de portée.

La partie 2 est consacrée à la modélisation des variabilités de la taille de portée considérée comme une variable continue. L'objectif de cette partie est d'identifier les facteurs de variation environnementaux influençant la taille de portée et sa variabilité puis d'estimer les paramètres génétiques du caractère et de sa variance résiduelle ou environnementale en utilisant un modèle linéaire mixte hétéroscédastique simplifié.

La partie 3 concernel'étude des variabilités de la taille de portée considérée comme une variable discrète dans le cadre théorique des modèles à seuils. Après avoir étudié rapidement les facteurs de variation qui affectent les tailles de portée avec un modèle homoscédastique, nous avons abordé les modèles hétéroscédastiques qui prennent en compte l'influence d'effets génétiques sur la variance environnemenatale du caractère. Dans un premier temps nous avons étudié l'effet d'un seul gène. Il s'agissait d'un gène connu pour avoir un effet majeur sur la moyenne du caractère et pour lequel nous avons recherché un effet possible sur la variance. Cette partie a conduit à un article accepté dans la revue ITEA, et à une communication dans un congrès. Puis nous avons considéré le cas polygénique dans un article soumis à GSE (non rejeté en 1ière lecture, en cours de révision). Dans cet article, nous avons d'abord présenté les modèles à seuils hétéroscédastiques utilisés dans la littérature pour analyser les réponses ordinales, puis nous avons montré comment ces modèles peuvent donner naissance à

d'autres modèles (modèle à seuils homothétiques, modèle à seuils indépendants) permettant de modéliser l'hétérogénéité de la variabilité des tailles de portée. Nous avons appliqué ces modèles sur différentes populations ovines (Romane, Rouge de l'Ouest et Suffolk).

PARTIE 1 : Bibliographie

Résumé :

Nous présentons une courte revue bibliographique sur la prolificité des brebis. Après avoir présenté rapidement ce caractère et son importance pour la filière allaitante, nous évoquons ses principaux facteurs de variation et l'intérêt de considérer la variabilité des tailles de portée. Nous présentons ensuite rapidement les concepts utilisés dans l'étude de la variabilité et de l'hétérogénéité des caractères. Les deux dernières parties sont consacrées à la présentation des modèles statistiques disponibles pour l'analyse génétiques des caractères continus et discrets dans le cadre homéoscédastique, puis les extensions qui en ont été faites afin de prendre en compte des facteurs d'hétérogénéité de variance.

Sommaire

1 La prolificité – Les variabilités – La canalisation 12

1.1 La prolificité des brebis 12

1.1.1 Productivité numérique - prolificité : Contexte économique 12
1.1.2 Sélection de la prolificité 13
1.1.3 Facteurs de variation environnementaux 13
1.1.4 Facteurs de variation génétiques 14
1.1.5 Augmentation de la prolificité 14
1.1.6 Problèmes phénotypiques liés à l'augmentation de prolificité 14
1.1.7 Intérêt économique de la canalisation des tailles de portée des brebis 15

1.2 Variabilité de la distribution des tailles de portée ovine 17

1.2.1 Liaison moyenne variance 17
1.2.2 Variabilité entre races 17
1.2.3 Variabilité intra-race 18

1.3 La canalisation 18

1.3.1 Définitions 18
1.3.2 Pourquoi canaliser ? 20
1.3.3 Modèles génétiques 20
1.3.4 Mise en œuvre, réalisations 21

2 Modèles d'analyse 23

2.1 Rappel sur le modèle linéaire polygénique 23

2.2 Les modèles non linéaires polygéniques 24

2.2.1 Généralités 24
2.2.2 GLMM utilisés en génétique sur les données ordinales 25

2.3 Le modèle polygénique hétéroscédastique 26

1 La prolificité – Les variabilités – La canalisation

1.1 La prolificité des brebis

1.1.1 Productivité numérique - prolificité : Contexte économique

En élevage ovin, la productivité numérique correspond au nombre d'agneaux produits sur une année par brebis de plus de 12 mois. C'est un paramètre particulièrement important du point de vue économique pour les élevages d'ovins allaitants; en effet, pour les producteurs d'agneaux de boucherie, les recettes proviennent quasi exclusivement de deux origines : les ventes d'agneaux et l'ensemble des aides, la part relative de ces deux facteurs variant selon les types d'élevages (Jousseins et al., 2012a). Or, lorsque le nombre d'agneaux augmente modérément, les charges s'accroissent mais elles restent inférieures à l'augmentation du revenu. Dans la plupart des situations françaises et étrangères, la marge brute augmente donc avec la productivité numérique (figure 1) (Benoit *et al*.).

La productivité numérique peut être vue comme le produit de trois composantes élémentaires : le taux de mises bas (nombre de mises bas annuel par brebis de plus de 12 mois), le taux de mortalité (nombre d'agneaux morts, sur le nombre d'agneaux nés) et la prolificité, définie par Desvignes et Darpoux(1964)comme le nombre d'agneaux nés pour cent mises bas. L'importance économique de la productivité numérique se retrouve donc en partie dans la prolificité. En race Blanche du Massif Central, à partir de deux modélisations des systèmes d'élevages majoritaires, l'effet d'une augmentation d'un point de prolificité a été évalué à une marge brute supplémentaire de 0.41 € par brebis (Dieny, 2011). Par ailleurs la prolificité contribue plus au poids d'agneaux sevrés par brebis que ne le fait la croissance individuelle des agneaux (Bradford, 1985).

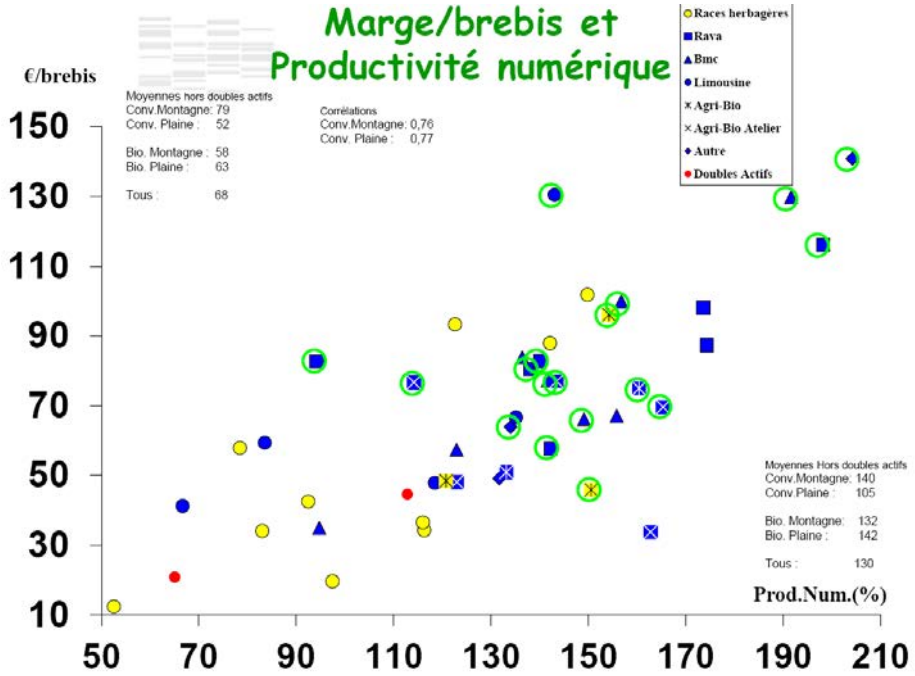


Figure 1 : Marge brute (hors prime) en fonction de la productivité numérique pour l'atelier Ovin allaitant des fermes de référence ovin viande, Productivité numérique en production ovine allaitante : composantes, intérêt économique, lien avec les facteurs de production et avec les aspects environnementaux ; Benoit, M. ; Laignel, G. ; Roulenc, M.Présentation orale : Assemblée Générale UPRA ROM Sélection,2013.

1.1.2 Sélection de la prolificité

Comme pour tout caractère quantitatif, l'amélioration de la prolificité des brebis peut être obtenue par l'amélioration des conditions d'élevage et par voie génétique. L'amélioration génétique d'une population pouvant se faire par croisement avec une race prolifique (Bradford, 1985; Fahmy, 1996), par sélection en race pure ou par l'introgression d'un gène majeur qui contrôle la prolificité (Sauvant and Perez, 2010). A chaque environnement (technique : ressources fourragères, type d'élevage, race ; économique : production d'agneaux de lait ou d'agneaux de 100 jours ; et social : disponibilité et coût de la main d'œuvre) correspond une stratégie optimale. En France, à l'heure actuelle, l'amélioration génétique de la prolificité se fait essentiellement en race pure même si quelques populations (Mérinos d'Arles, Lacaune Viande, Noire du Velay, Grivette) dans lesquelles il existe un gène majeur qui contrôle la prolificité ont choisi d'améliorer la prolificité à l'aide de ces gènes majeur(Bodin et al., 2011). Depuis leur origine dans les années 70, les schémas de sélection collectifs des ovins allaitants français ont considéré la prolificité comme un des objectifs prioritaires et des propositions d'optimisation des schémas ont été faites compte tenu des spécificités de ce caractère (Elsen et al., 1994).

1.1.3 Facteurs de variation environnementaux

De très nombreux facteurs de variation agissent sur la prolificité (Poivey et al., 1995) et doivent être pris en compte dans les analyses génétiques de ce caractère. En dehors des effets environnementaux (troupeau, année, saison) ou individuels (âge, numéro de performance) classiques qui jouent sur la prolificité (Tiphine et al., 2011) comme sur la plupart des caractères zootechniques, il existe des facteurs spécifiques liés au caractère prolificité et la physiologie de la brebis:

- Intervalle post-partum. L'intervalle entre une mise bas et la mise à la reproduction suivante affecte la fertilité que ce soit après insémination (David et al., 2008) ou monte naturelle (Janssens et al., 2004), et affecte aussi la prolificité bien que dans une moindre mesure. Ainsi la prolificité des brebis en rythme annuel de reproduction est plus élevée que dans un système de 3 agnelages en 2 ans (Rodriguez et al., 1998). Cet effet est probablement indirect et dû au fait que plus l'intervalle post-partum est long plus les brebis ont le temps de restaurer leurs réserves corporelles; or plusieurs études ont montré l'influence positive des bonnes conditions corporelles sur le taux d'ovulation et la prolificité (Atti et al., 2001; Gunn et al., 1984). Par ailleurs les brebis mis à la reproduction alors qu'elles sont encore en lactation ont généralement des intervalles post-partum plus courts que celui des brebis taries et l'on connait l'effet négatif de la lactation sur la taille de portée (Theriez and Molenat, 1975).

- Mode de reproduction. L'utilisation des traitements hormonaux consistant en une éponge vaginale imprégnée d'acétate de fluorogestone (FGA) posée pendant 12 ou 14 jours suivie d'une injection de "Pregnant Mare Serum gonadotrophin" (PMSG) permet d'obtenir des chaleurs et des ovulations lorsque les brebis sont en période d'anœstrus saisonnier (Thimonier et al.). Ces traitements sont aussi le seul moyen existant actuellement pour synchroniser efficacement les ovulations en vue de l'insémination (Baril et al., 1993) et représentent donc un outil pratiquement indispensable au bon fonctionnement des schémas de sélection (Elsen et al., 1984). Ces traitements modifient les performances de reproduction et augmentent la prolificité de l'ordre de 0.2 à 0.3 point dans la majorité des races françaises (Bodin and Elsen, 1989). Ils amènent à distinguer deux types de prolificité : la prolificité après œstrus naturel et la prolificité consécutive à un œstrus induit qui bien que fortement corrélées génétiquement ($r_g \sim 0.8$) sont considérées comme deux caractères différents (Baelden et al., 2005). Les héritabilités utilisées dans les programmes d'évaluation génétique sont respectivement de 0.10 pour la prolificité naturelle et 0.08 pour la prolificité sur œstrus induit (Tiphine et al., 2011).

1.1.4 Facteurs de variation génétiques

La taille de portée des ovins devrait être analysée par des méthodes propres aux variables discrètes en utilisant notamment des modèles à seuils issus du modèle de Wright (1934) et largement développés par Gianola (1982). Ces modèles considèrent l'existence d'une variable sous-jacente (communément appelée "liability") et d'un ensemble de *n*-1seuils qui transforment cette distribution en *n* catégories observables (voir ci-dessous). Ces modèles supposent que les effets environnementaux et génétiques affectent de manière (additive) la distribution de la variable sous-jacente et donc la distribution des observations de façon non linéaire. Ceci rend difficile l'interprétation des résultats et un grand nombre d'auteurs préfèrent utiliser les méthodes purement linéaires en traitant la taille de portée comme une variable continue.
Il existe une très grande différence de prolificité moyenne entre des brebis de races pures ou croisées élevées strictement dans le même environnement (Berrichones 1.26 – Romanov 3.00) (Ricordeau et al., 1976) ou dans des environnements différents (Bodin and Elsen, 1989; Bradford, 1985). Ces différences raciales sont importantes (l'écart-type entre races est du même ordre de grandeur que l'écart-type intra race) et sont pour une très large part d'origine génétique. L'absence de ségrégation du caractère au cours des générations successives de croisement à partir d'animaux de races prolifiques et peu prolifiques (Ricordeau et al., 1977) indique que ce caractère est très largement polygénique de plus, l'analyse des animaux croisés montre une absence d'hétérosis pour ce caractère dont le déterminisme génétique est donc essentiellement additif (Ricordeau et al., 1977).
Que ce soit en utilisant un modèle à seuil (Janssens et al., 2004; Martin et al., 2014; Matika et al., 2003; Matos et al., 1997) ou un modèle linéaire (Baelden et al., 2005), l'héritabilité de la prolificité est faible. La valeur moyenne de ce paramètre sur 30 publications analysées par Bradford (1985) est de 0.10, elle est de 0.082 (0.01 à 0.17) pour 14 publications analysées par Matika et al. (2003). Par ailleurs, le coefficient de variation de la prolificité est très grand : 35 à 40 % dans les conditions naturelles et 40 à 45 % après traitement hormonal (Bodin and Elsen, 1989) et la variabilité génétique additive est très importante. En effet, pour une prolificité moyenne de 1.50 on peut estimer à ± 0.4 agneau (± 2.5 écart type génétique additif) l'étendue de la variation génétique additive autour de la moyenne (Bodin et al., 1999). Ce qui autorise un progrès génétique de l'ordre de 2 % par an comme pour beaucoup d'autres caractères.

1.1.5 Augmentation de la prolificité

L'amélioration des conditions d'élevage (conduite d'élevage, reproduction, sanitaire (voir revue de Demers (2010)) et des valeurs génétiques permet d'augmenter relativement régulièrement la prolificité moyenne des brebis. Pour les 22 plus importantes races contrôlées en France, la prolificité est passée moyenne est passéé de 1.54 à 1.64 entre 1992 et 2012. Les 3 races pour lesquelles cette augmentation a été la plus forte sont: la Noire du Velay (+0.21), la Blanche du Massif Central (+0.15) et la Romane (+0.11) (données extraites des bilans du contrôles de performance ; IE – Races de France - INRA). Cette augmentation de la moyenne s'accompagne d'un accroissement du pourcentage de mises bas multiples et d'un accroissement supérieur du pourcentage d'agneaux nés multiples (pour une prolificité de 2.0, environ 1/3 des agneaux sont nés triples ou plus mais il y a seulement 20% des mises bas triple ou plus).

1.1.6 Problèmes phénotypiques liés à l'augmentation de prolificité

Bien qu'économiquement intéressante l'augmentation de prolificité induit des effets négatifs sur la rentabilité de l'élevage qui ont été largement passé en revu par Demers (2010). L'effet le plus direct est la réduction du poids des agneaux à la naissance. Cette réduction, peut atteindre 20 % en moyenne pour les jumeaux et 34 %, pour les triplés par rapport aux agneaux nés simples. Elle dépend de nombreux facteurs (alimentation, taille de la mère, …) et semble d'autant moins forte que la prolificité moyenne de la race augmente (Theriez, 1991). Cet effet entraine une augmentation importante de la mortalité des agneaux (figure 2) qui est aussi d'autant moins forte que la prolificité de la race est élevée (Everett-Hincks and Cullen, 2009 ; Gama et al., 1991 ; Theriez, 1982).

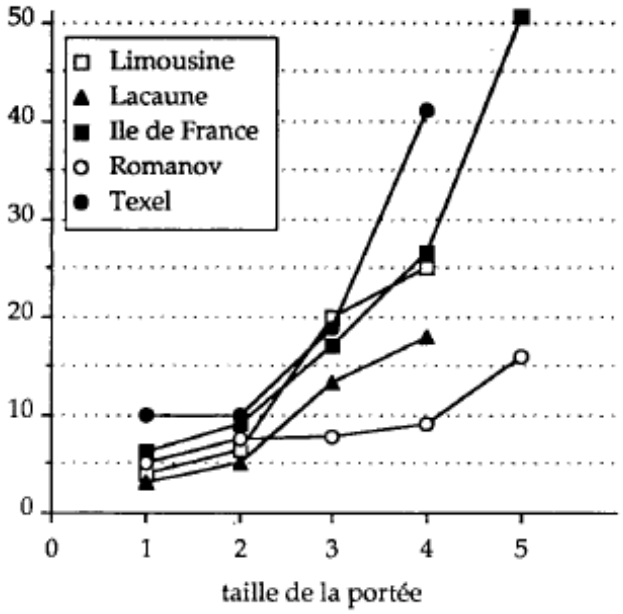


Figure 2 : Evolution du taux de pertes au cours des premiers mois selon la taille de portée et la race (d'après Theriez, 1982)

L'augmentation de la taille de portée à la naissance a aussi d'autres effets qui sont plus ou moins liés à la diminution du poids. Les agneaux nés multiples présentent une vigueur plus faible, mettent plus de temps à se lever et à aller téter (Dwyer et al., 2003)et vu leur nombre, ils reçoivent une quantité réduite de colostrum (Theriez, 1991) et bénéficient de moins de soins de la part de leur mère. De plus, une brebis n'a que deux tétines et pourra difficilement allaiter des triplés ou plus; généralement les agneaux supplémentaires sont mis à l'adoption à une autre brebis ou en allaitement artificiel, à l'origine d'un risque de mortalité et des coûts économiques supplémentaires.

1.1.7 Intérêt économique de la canalisation des tailles de portée des brebis

L'uniformité des tailles de portée à deux agneaux par portée est comme le signalait Bradford (1985) un objectif important pour un grand nombre de système de production ovin semi-extensif. En fait cette uniformité autour d'un optimum est à rechercher quel que soit le système de production, de très extensif, au plus intensif, mais l'environnement (disponibilité en ressources fourragères, main d'œuvre, climatologie, etc.) définit la valeur de l'optimum (1, 2 ou 3 agneaux) pour chaque système (Bodin and Elsen, 1989). Bradford (1995) a montré par simulation que pour une même prolificité de 1.95 un troupeau avec 75% de mises bas double (15% MB1, 75% MB2, 10% MB3) produit, selon les paramètres choisis, de 3.5 à 8.7% de plus de poids d'agneaux sevrés pour cent brebis qu'un troupeau avec seulement 45% de de mise bas de ce type (30% MB1, 45% MB2, 25% MB3). De façon analogue, une étude économique de Benoit mentionnée par Bodin et al.,(2000) rapporte une supériorité de 24% des brebis ayant consécutivement deux fois des doubles par rapport à une brebis ayant une fois des triples puis une fois un simple.
Il est clair que la canalisation des tailles de portée des populations ovines n'a d'intérêt que pour des populations dans lesquelles il y a au moins trois classes et donc que le pourcentage de mises bas triples n'est pas négligeable ce qui correspond à une prolificité moyenne supérieure à 1.60. En dessous, il suffit d'augmenter (ou de diminuer) la moyenne vers la valeur optimum. On constate qu'environ 60% des brebis contrôlées en France en 2013 appartiennent à des races dont la prolificité est supérieure à 1.60 (tableau 1) et donc pour lesquelles la canalisation devient un objectif important.

Tableau 1 : Effectif et prolificité des brebis adultes de 38 races contrôlées en France à la campagne 2013. (Données du Bilan de contrôle de performances ovin allaitant, campagne 2013; IE - INRA).

Race	Nb Mises bas adultes	Prolificité 2013	Effectif RGA	Effectif contrôlé cumulé	% cumulé
Charmoise	3321	1.11	23923	3321	100.0%
Castillonnaise	1451	1.14	4004	4772	98.1%
Barégeoise	2111	1.15	4811	6883	97.3%
Lourdaise	307	1.15	1379	7190	96.2%
Tarasconnaise	8372	1.20	152100	15562	96.0%
Mérinos d'Arles	12193	1.22	283286	27755	91.3%
Mourerous	4218	1.24	36255	31973	84.5%
Boulonnaise	1358	1.30	5371	33331	82.1%
Limousine	6216	1.39	36369	39547	81.4%
Préalpes	4736	1.41	241126	44283	77.9%
Bizet	2566	1.43	9754	46849	75.2%
Hampshire	715	1.45	7659	47564	73.8%
Solognote	1571	1.46	3166	49135	73.4%
Dorset Down	103	1.51	2366	49238	72.5%
Blanc du Massif Central	18968	1.54	358954	68206	72.5%
Southdown	293	1.54	8069	68499	61.9%
Rava	6092	1.57	33622	74591	61.7%
Berrichon du cher	2220	1.61	37090	76811	58.3%
Berrichon de l'Indre	776	1.63	7223	77587	57.1%
Caussenard du Lot	24951	1.63	107695	102538	56.6%
Noire du Velay	4289	1.67	22266	106827	42.7%
Est à Laine Mérinos	2534	1.70	38580	109361	40.3%
Suffolk	**2681**	**1.70**	**165954**	**112042**	**38.9%**
Avranchin	117	1.74	3710	112159	37.4%
Ile de France	11248	1.75	235748	123407	37.3%
Cotentin	87	1.76	15976	123494	31.0%
Roussin	1338	1.76	21549	124832	31.0%
Mouton Charollais	6537	1.79	281753	131369	30.2%
Mouton Vendéen	5793	1.79	104745	137162	26.6%
Texel	4045	1.87	295897	141207	23.3%
Bleu du Maine	789	1.88	14318	141996	21.1%
Rouge de l'Ouest	**3680**	**1.90**	**111362**	**145676**	**20.6%**
Lacaune	10578	1.91	300000	156254	18.6%
Grivette	4571	1.98	20966	160825	12.7%
Romane (INRA 401)	**17488**	**2.14**	**41998**	**178313**	**10.1%**
Finnoise	186	2.71	186	178499	0.3%
Romanov	388	3.21	9143	178887	0.2%

1.2 Variabilité de la distribution des tailles de portée ovine

1.2.1 Liaison moyenne variance

L'augmentation du nombre de classes de taille de portée lorsque la prolificité moyenne augmente crée une liaison très forte entre la moyenne et la variance de ce caractère. Tant que le pourcentage de mises bas triples est inférieur à 1% dans la population la distribution des tailles de portée suit une loi binomiale pour laquelle la variance est directement liée à la moyenne. Pour les races françaises et des pourcentages de mises bas triples supérieurs, cette liaison reste très forte (figure 3) et la corrélation entre moyenne et variance est de 0.987. Compte tenu de cette liaison il est même possible de prédire les pourcentages de chaque type de mises bas sachant la prolificité moyenne de la population. Cependant, deux effets modifient cette relation générale de façon importante : les traitements hormonaux d'induction des ovulations (Bodin and Elsen, 1989) et la ségrégation de gènes majeurs d'ovulation dans la population (Martin et al., 2014).

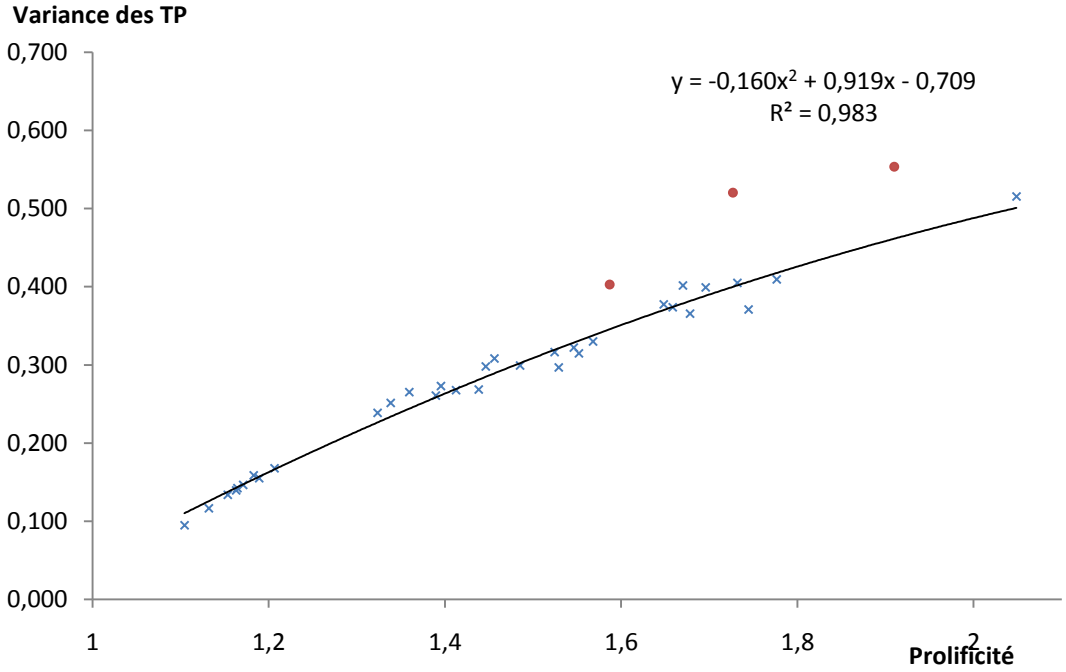


Figure 3 : Distribution des variances des tailles de portée par race pour les données (n=2 399 063) contrôlées depuis la campagne 2000 de 36 races ovines françaises. (Les points en rouge correspondent à 3 races (Noire du Velay, Lacaune, Grivette) pour lesquelles un gène majeur d'ovulation est en ségrégation ; la courbe de tendance est estimée sans ces 3 races. (Données de la BDIR ovin allaitant, IE - INRA).

1.2.2 Variabilité entre races

La prolificité étant une des composantes essentielles, sinon la plus importante, de la productivité de l'élevage ovin allaitant il existe une littérature très abondante sur ses facteurs de variation et la manière de les contrôler. Comme pour la plupart des caractères, il existe cependant peu de données sur les facteurs qui agissent sur la variabilité de cette prolificité. La liaison moyenne variance n'étant pas totale il existe pourtant pour une prolificité donnée une variabilité de la variance dont l'origine peut être environnementale et génétique. On sait depuis longtemps qu'il existe une certaine variabilité raciale et les races islandaises sont connues pour être prolifiques avec une très forte proportion d'ovulations doubles et un taux de survie embryonnaire élevé conduisant à une proportion de mises bas

doubles beaucoup plus importantes que ce qui serait attendu (Hanrahan, 1989). Plusieurs races norvégiennes (Steine, 1985) ainsi que la race Texel ont aussi été annoncées comme ayant un plus fort taux de mises bas doubles que les autres (Hanrahan, 1989). La Race Spaelsau parait tout à fait exceptionnelle avec plus de 73% de mises bas double; le maximum pour une race française est de 58% (race Texel; *données de la BDIR ovin allaitant, IE - INRA*). Au contraire, parmi les races prolifiques, la race finnoise présente une plus grande variabilité du taux d'ovulation et une plus forte mortalité embryonnaire que la Romanov. Ces deux facteurs contribuant à ce que cette race ait un plus faible taux de mises bas doubles que la Romanov pour une prolificité similaire (Hanrahan, 1986).

Tableau 2 : Prolificité moyenne et variance des tailles de portée à la campagne 1982 ainsi que le pourcentage de chaque type de portée pour les 4 races les plus communes en Norvège. (d'après Steine (1985))

Race	nb MB	Prolif.	Variance	% TP1	%TP2	%TP3	%TP3+
Dala	12812	1.93	0.35	20.4	66.4	12.8	0.4
Rygja	30463	1.93	0.35	20.7	66.2	12.6	0.5
Steigar	21282	1.90	0.34	21.5	67.0	11.1	0.4
Spælsau	51583	1.93	0.26	16.9	73.5	9.6	-

1.2.3 Variabilité intra-race

A ce jour il existe peu de données sur la variabilité intra race du pourcentage de chaque type de mise bas. Une estimation de l'héritabilité de la réalisation ou non d'une mise bas double a été faite en Lacaune (Elsen et al., 1994). Sa valeur était faible (h^2=0.13) et environ la moitié de celle de la taille de portée. Des estimations des paramètres génétiques de la variance des tailles de portée ont été faites sur cette même population en utilisant une méthode bayésienne (SanCristobal-Gaudy et al., 2001). Cette méthode, très lourde à mettre en œuvre, a montré l'existence d'une variabilité génétique de la variance des tailles de portée. Cependant la découverte d'un gène majeur d'ovulation dans cette population (Bodin et al., 2002) a pu fortement biaisé les résultats. En effet la descendance d'un mâle hétérozygote pour un gène qui agit sur la moyenne d'un caractère est en général plus variable que celle d'un mâle homozygote.
Une approche similaire à celle sur l'occurrence des mises bas double a été faite sur données de troupeaux Néo-Zémandais(Amer and Bodin, 2006). L'héritabilité de l'occurrence des mises bas de plus de 3 agneaux (h^2=0.036) était significativement différente de 0 et environ la moitié de l'héritabilité de la taille de portée (h^2=0.06). La corrélation génétique entre ces 2 caractères qui était quant à elle différente de 1 permettait d'envisager une sélection pour limiter les mises bas triples sans affecter la prolificité moyenne.

1.3 La canalisation

1.3.1 Définitions

Les objectifs de sélection s'adaptent en permanence aux évolutions de la production ainsi qu'aux demandes du marché et de la société. Les dernières décennies ont successivement vu émerger les objectifs de qualité de production, puis des objectifs d'adaptation des animaux à leur milieu de production au travers de l'amélioration des caractères dits fonctionnels (qualité des aplombs, morphologie mammaire, etc.), et finalement des objectifs non marchands pour répondre aux exigences environnementales, sociétales et de bien-être animal. Pour certaines espèces et certaines races, l'élargissement de leurs zones de dissémination, la diminution du nombre de sous-populations sélectionnées ainsi que l'augmentation des relations génétiques entre ces groupes imposent de considérer les interactions génétiques milieu dans les processus de sélection. Pour la grande majorité des populations en sélection, les relations des animaux à leur milieu de production deviennent donc

prépondérantes et les notions d'adaptation, de robustesse, de sensibilité environnementale voire de plasticité, sont en bonne place dans la liste des objectifs de sélection. Debat et David (2001) ont tenté de préciser ces concepts qui ont aussi fait l'objet d'un dossier spécial de la revue Production Animale (Sauvant and Perez, 2010). D'une manière synthétique on peut les définir de la façon suivante :

La robustesse : Il existe au moins deux écoles pour définir la robustesse en production animale, selon qu'on l'associe à un potentiel de production globale ou, à un potentiel de production spécifique. D'une façon générale, elle désigne la capacité d'animaux à maintenir un niveau élevé de performances dans des environnements variés, mais c'est aussi "l'aptitude à combiner un haut potentiel de production avec la capacité à endurer des stress, ce qui permet l'expression d'un haut potentiel de production dans une grande variété de conditions environnementales (Knap, 2005)".

La rusticité : C'est un concept très proche de la robustesse et ce vocable a été très largement utilisé pour caractériser des animaux bien adaptés aux aléas de l'élevage (alimentaires ou climatiques) et possédant une résistance plus ou moins forte à certaines affections (parasitisme en particulier). C'est ainsi la capacité à s'adapter, à croître dans un milieu défavorable, et à supporter des conditions dégradées. Mais il n'y a pas de notion de maintien d'un niveau de production dans le concept de rusticité contrairement à ce que sous-tend le concept de robustesse. En général les races ovines françaises dites rustiques sont des races ayant peu ou pas subi de croisements avec des races anglaises améliorées et dont le berceau d'élevage se situe aujourd'hui principalement dans la moitié sud de la France, en zones classées administrativement "zones de montagne" ou "zones de piémont" ou encore avec "handicap naturel".

La flexibilité :C'est un concept largement utilisé en sciences de gestion et sciences économiques et qui comme le rappelle Dedieu et Ingrand (2010) renvoie à l'image du "roseau qui plie mais ne rompt pas". Pour un système, la flexibilité est la capacité de s'adapter à des perturbations soit en modifiant sa structure, soit en modifiant sa finalité(Sauvant and Martin, 2010). Ce concept fait partie des outils utilisés pour l'étude des capacités d'adaptation des systèmes d'élevage à moyen et long terme à des perturbations d'ordre climatique (risque accru d'événements extrêmes), d'ordre économique (incertitude et instabilité des prix agricoles soumis à la spéculation), d'ordre sanitaire (émergence de crises sanitaires imprévisibles), etc. (Dedieu and Ingrand, 2010). Cependant, la recherche de systèmes d'élevage flexible ne met pas en jeu la flexibilité individuelle des animaux.

La plasticité : C'est la capacité d'un organisme à modifier son phénotype lorsque les conditions environnementales changent. Du point de vue évolutif, la plasticité peut être vue comme la capacité d'un génotype à produire des phénotypes qui sont le plus adaptés à chaque environnement (Gavrilets and Scheiner, 1993). D'une manière générale, la relation entre les phénotypes possibles produits par un génotype donné et la gamme de valeurs d'une variable environnementale à laquelle il est exposé constitue la norme de réaction. Ces normes de réactions qui peuvent être des fonctions linéaires, polynomiales voire quelconques constituent le principal outil pour étudier la plasticité (Callahan et al., 1997; Delpuech et al., 1995).

La canalisation :La relation entre les notions de robustesse et de canalisation vient de Waddington (1940) qui a créé le concept de canalisation pour décrire l'augmentation de la robustesse d'un phénotype vis-à-vis des perturbations génétiques ou environnementales; ce qui l'a amené à distinguer une canalisation génétique qui rassemble les mécanismes pour diminuer l'impact des effets des mutations et une canalisation environnementale qui regroupe les processus qui tamponnent les effets de l'environnement. Pour cet auteur, la robustesse et notamment la canalisation environnementale, est la capacité à maintenir l'expression d'un caractère à un niveau relativement stable dans des environnements variés. Ce concept est attachée ici à un caractère, contrairement à des notions beaucoup plus globales et multicritères de la robustesse telle que celle définie par Knap (2005), ou celle de Friggens (2010) qui fait intervenir les changements d'allocation de ressources entre fonctions biologiques en conditions limitantes de l'environnement. Malgré ces divergences de définition de la robustesse et bien que Waddington (1940) ait étroitement associé canalisation et robustesse, on peut à l'heure actuelle définir plus spécifiquement la canalisation comme l'action de sélectionner pour

diminuer la sensibilité aux variations environnementales et entraîner la moyenne du caractère vers une certaine valeur. Dans ce cadre, la diminution de la sensibilité aux conditions de milieu fait partie du concept élargi des interactions Génotype – Environnement (Bodin et al., 2010).

1.3.2 Pourquoi canaliser ?

L'intérêt de diminuer la sensibilité des animaux aux conditions environnementales qui agissent sur un caractère ou de diminuer la variance environnementale d'un caractère répond au moins à trois grandes préoccupations (Bodin et al., 2010).

- Certains caractères présentent une valeur optimum autour de laquelle il peut être intéressant de réduire la variabilité. C'est évidemment le cas pour la taille de portée des brebis, mais c'est aussi vrai pour celle des truies (Sorensen and Waagepetersen, 2003), pour le poids à la naissance des lapereaux (Bolet et al., 1996), ou des porcelets dont la variabilité intra-portée est largement lié à leur survie (Roehe and Kalm, 2000). L'intérêt de la canalisation a aussi été signalé pour des caractères aussi variés que le poids adulte des escargots (Ros et al., 2004), le Ph ultime du porc (Larzul et al., 2006), le rapport de la quantité de matière grasse à la quantité de matière protéique du lait de chèvre (SanCristobal-Gaudy et al., 1998a), le poids à la naissance chez les lapins (Garreau et al., 2008a)…

- L'agriculture durable mobilise des systèmes de production dans lesquels les facteurs de variation des caractères sont moins strictement contrôlés (animaux élevés en plein air, ou en cage collective). L'augmentation de variabilité qui en résulte doit être compensée par une plus grande stabilité des animaux aux contraintes environnementales. Par ailleurs, comme le souligne Phocas (2014) "la fluctuation croissante des conditions climatiques et la nécessité d'exploiter des environnements d'élevage moins favorables pour éviter la compétition avec les besoins humains obligent à comprendre les variations temporelles de la robustesse pour les inclure dans les stratégies de conduite et de sélection des animaux".
- Un des gages de la qualité des produits agricoles est l'homogénéité des productions (Soufflet, 1995) qui permet aux consommateurs d'identifier ou de différencier les produits dans une vaste gamme de proposition. Le tri et la classification des produits peuvent être faits par les industries agro-alimentaires mais il est plus intéressant que les productions soient naturellement homogènes.

1.3.3 Modèles génétiques

La nature et les bases génétiques de la canalisation ont été étudiés en détail par Rendel (1960) et plusieurs modèles génétiques rappelé parScheiner and Lyman(1991) ont été proposés (texte-figure 1). On distingue ainsi :

- Le modèle de "surdominance" proposé par Lerner(1954) et Gillespie and Turelli(1989) qui relie la variabilité phénotypique au taux de loci hétérozygotes.

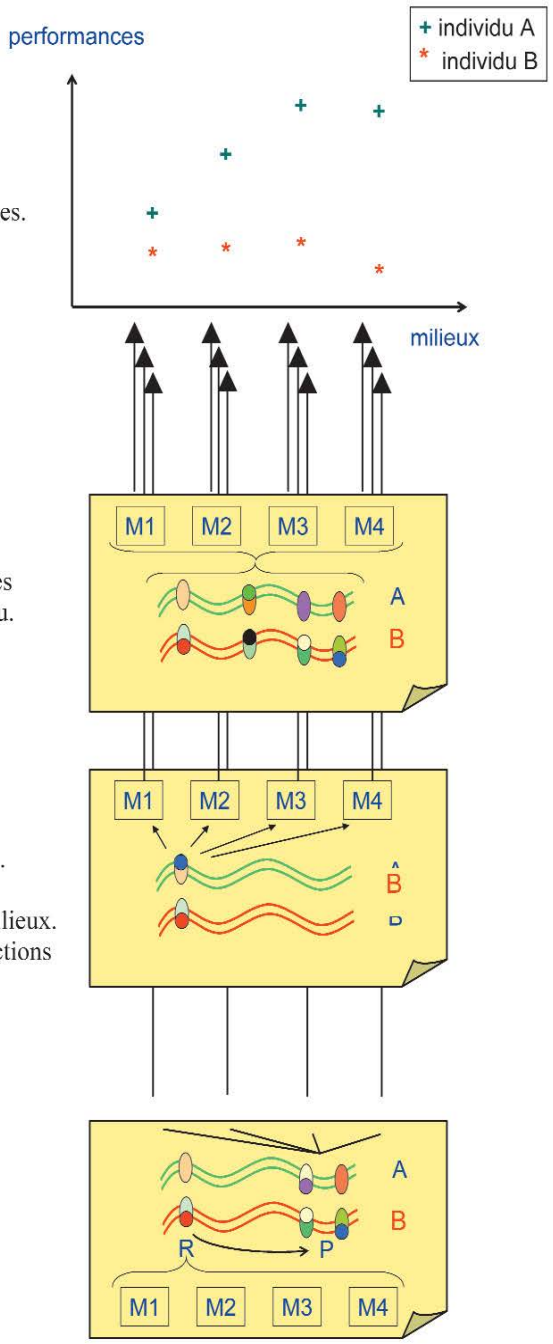
- le modèle "pléiotropique" énoncé par Falconer and Mackay (1996) et qui suppose que l'expression des gènes n'est pas constante et varie selon les environnements.

- Et le modèle "épistatique" (Lynch and Gabriel, 1987; Scheiner and Lyman, 1989) pour lequel les gènes qui déterminent l'expression d'un caractère sont partiellement différents de ceux qui contrôlent sa sensibilité environnementale.

1.3.4 Mise en œuvre, réalisations

De très nombreuses expériences ont montré qu'il était possible de réduire la variabilité environnementale d'un caractère ou sa sensibilité aux facteurs environnementaux par sélection. Elles concernent souvent des espèces de laboratoire comme la drosophile (Piñeiro, 1992; Waddington, 1940), le tribolium (Cardin and Minvielle, 1986), ou la souris (Cervantes et al., 2010; Ibáñez-Escriche et al., 2008a; Rendel, 1967 ; Pun et al., 2013), mais ont été aussi conduites dans des populations d'animaux de rente comme le lapin (Argente, M.J et al., 2011; Bolet et al., 2005; Garreau and Bolet, 2004; Ibáñez-Escriche et al., 2008b; Martínez-Álvaro et al., 2013)ou la truite (Dupont-Nivet et al., 2007, 2012). Les résultats indiquent le plus généralement que la sélection pour réduire la variance environnementale peut donner des résultats intéressants. Des estimations de paramètres génétiques pour la variance ont aussi été estimées dans un très grand nombre d'autres espèces et de situations: par exemple chez le porc (Canario et al., 2010; Damgaard et al., 2003; Huby et al., 2003; Ibáñez-Escriche et al., 2008c; Larzul et al., 2006; Sorensen and Waagepetersen, 2003), le poulet (Mulder et al., 2009; Wolc et al., 2009), les bovins laitiers (Rönnegård et al., 2013) ou à viande (Fina et al., 2013), voire même chez l'escargot (Ros et al., 2004) ou l'alpaca (Gutiérrez et al., 2011). Les études de paramètres génétiques de la variabilité environnementale de caractères discrets comme la taille de portée des brebis sont cependant peu fréquentes (Amer and Bodin, 2006; SanCristobal-Gaudy et al., 2001).

Encadré 1. ***Les trois modèles génétiques proposés pour expliquer les interactions génotype x milieu.***



On observe que les performances de l'individu **A** sont très variables dans 4 environnements différents, alors que celles de l'individu **B** sont beaucoup plus stables.

Les modèles proposés (Scheiner et Lyman 1991)

Le modèle de «surdominance»
Sous cette hypothèse, les individus les plus hétérozygotes sont les mieux armés pour tamponner les effets du milieu.

Le modèle «pléiotropique»
Les gènes présentent une expression différente dans les différents environnements et la variabilité de ces expressions induit la variabilité des performances. A l'extrême, certains gènes ne s'expriment pas du tout, et des gènes différents s'expriment dans les différents milieux. C'est l'hypothèse explicative la plus classique des interactions G x E.

Le modèle «épistatique»
Des gènes (R) sont sensibles aux variations du milieu et, selon le niveau du milieu, ils modulent l'expression des gènes de production (P). La variabilité des gènes de sensibilité à l'environnement détermine la variabilité résiduelle des productions.

Figure 4 : Différents modèles génétiques expliquant la canalisation et les interactions GxE (tirée de (Bodin et al., 2010).

2 Modèles d'analyse

Le modèle génétique « originel » repose explicitement sur les hypothèses d'homoscédasticité et de normalité des résidus. Cependant, dans la pratique, ces deux hypothèses ne sont pas toujours vérifiées. C'est notamment le cas pour la normalité lorsque l'on étudie la taille de portée qui est une variable discrète ordonnée. De surcroit, dans une volonté de réduire la variabilité de ce caractère, il est nécessaire de considérer une hétérogénéité des variances (modèles hétéroscédastiques). Après un bref rappel sur le modèle polygénique, nous présenterons les modèles linéaires mixtes généralisés utiles pour traiter les données non gaussiennes puis nous présenterons les modèles hétéroscédastiques.

2.1 Rappel sur le modèle linéaire polygénique

Les analyses de génétique quantitative et la pratique de l'amélioration génétique sont largement fondées sur la théorie du déterminisme polygénique selon laquelle un caractère est gouverné par l'expression d'une multitude de gènes à petits effets (Falconer D.S., 1974). Dans ce cadre le phénotype d'un caractère de distribution continue approximativement gaussienne, peut être décrit à l'aide d'un modèle linéaire mixte comme la somme d'effets génétiques additifs et environnementaux ; soit pour N observations sur n individus :

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{Wp} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{p} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{A}\sigma_u^2 & 0 & 0 \\ 0 & \boldsymbol{I}_n\sigma_p^2 & 0 \\ 0 & 0 & \boldsymbol{I}_N\sigma_\varepsilon^2 \end{bmatrix} \right)$$

Où $\boldsymbol{y}$ est le vecteur $(N \times 1)$ des observations ; $\boldsymbol{\beta}$ est le vecteur des coefficients des effets fixes ; $\boldsymbol{u}$ est le vecteur des effets génétiques additifs ; $\boldsymbol{p}$ est le vecteur des effets d'environnement permanent (inclus dans le modèle uniquement dans le cas de répétition des observations par sujet) ; $\boldsymbol{\varepsilon}$ est le vecteur des résidus normalement indépendamment et identiquement distribués. $\boldsymbol{X}$, $\boldsymbol{Z}$ et $\boldsymbol{W}$ sont les matrices d'incidence connues reliant les observations aux effets fixes et aléatoires qui les influencent. $\boldsymbol{A}$ est la matrice de parenté connue, et $\boldsymbol{I}_N$ et $\boldsymbol{I}_n$ sont les matrices identité de taille N et n(nombre de sujets avec observation), respectivement.

L'héritabilité du caractère est définie comme le rapport de la variance génétique sur la variance phénotypique, soit : $h^2 = \dfrac{\sigma_u^2}{\sigma_u^2 + \sigma_p^2 + \sigma_\varepsilon^2}$

L'estimation des effets fixes et la prédiction des effets aléatoires de ce modèle sont basées sur le BLUE (best linear unbiaised estimation) et le BLUP (best linear unbiased predictor)(Henderson, 1963). Cette méthode revient à trouver les solutions $\boldsymbol{u}$, $\boldsymbol{p}$ et $\boldsymbol{\beta}$ du système d'équations du modèle mixte(Henderson, 1973):

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} & \boldsymbol{X'W} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \dfrac{1-r}{h^2}\boldsymbol{A}^{-1} & \boldsymbol{Z'W} \\ \boldsymbol{W'X} & \boldsymbol{W'Z} & \boldsymbol{W'W} + \dfrac{1-r}{r-h^2}\boldsymbol{I}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{u} \\ \boldsymbol{p} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \\ \boldsymbol{W'y} \end{bmatrix}$$

Avec $\frac{1-r}{h^2} = \frac{\sigma_\varepsilon^2}{\sigma_u^2}$ et $\frac{1-r}{r-h^2} = \frac{\sigma_\varepsilon^2}{\sigma_p^2}$. Lorsque les composantes de variance sont inconnues, ces dernières sont remplacées par leur estimée dans le système d'équation du modèle mixte. Les estimations et prédictions sont alors appelés EBLUE et EBLUP pour empirical BLUE(P).

Il existe différentes méthodes générales de résolution de ce système d'équation linéaire. Les estimations des composantes de la variance peuvent être obtenues par différentes méthodes. La plus utilisée est la méthode REML (Maximum de vraisemblance restreint). Différents algorithmes permettent d'obtenir l'estimation REML : algorithme EM(Dempster et al., 1977), l'algorithme de Newton-Raphson, l'algorithme de Fisher, méthodes de Monte-Carlo par chaîne de Markov (MCMC).

Le modèle présenté ci-dessus est **homoscédastique** (variance résiduelle homogène). Cependant, l'hypothèse d'homogénéité peut ne pas être valide (Hill, 1984; Robert-Granié et al., 1997). Dans ce cas, il est nécessaire d'utiliser des modèles linéaires dits **hétéroscédastiques** (variance résiduelle hétérogène) afin d'éviter les biais dans l'estimation des paramètres et nuire à la sélection(SanCristobal et al., 2002). Différents types de facteurs peuvent influencer la variance résiduelle, des facteurs environnementaux (fixes ou aléatoires) (Ducrocq, 2000a)et/ou des facteurs génétiques (Foulley and Jaffrézic, 2010). Dans ce cas nous parlerons de modèles **hétéroscédastiques génétiques**.

2.2 Les modèles non linéaires polygéniques

2.2.1 Généralités

Afin de permettre l'analyse de données non gaussiennes, une extension en termes de loi du modèle linéaire classique a conduit au développement de la classe plus large des modèles linéaires généralisés. Ils ont été introduits par Nelder et Wedderburn (1972) et formalisés par McCullagh & Nelder (1989) au début des années 89. Cependant pour prendre en compte certains éléments lors de l'analyse (sur dispersion, structure de données…), ces modèles ont été agrémentés d'une ou plusieurs composantes aléatoires. Les modèles linéaires généralisées mixtes (GLMM ou GL2M) sont ainsi apparus à la fin des années 1980. Un GLMM se caractérise par les trois hypothèses suivantes :

Les observations y_i sont distribuées, conditionnellement aux effets aléatoires, selon une loi appartenant à la famille exponentielle $f_u(y_i, \theta_i) = \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right\}$ où θ_i est le paramètre canonique d'intérêt et ϕ un paramètre de dispersion. Les fonctions b et c sont spécifiques à chaque distribution (tableau 3).

On montre que $E(y_i|u) = b'(\theta_i) = \mu_u$ et $\text{var}(y_i|u) = a_i(\phi)\, b''(\theta_i)$.

On définit un prédicteur linéaire μ tel que $\boldsymbol{\mu}_u = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{Wp}$

Il existe une fonction de lien g, spécifique de chaque distribution, telle que $\mu_u = g\left(b'(\theta)\right)$.

Tableau 3 : Différentes distributions

Distribution y	Densité $f(y)$	$g(\mu)$	$a(\phi)$	$\mu(\theta)$	$b''(\theta)$
Normale $N(\mu,\phi^2)$	$\frac{1}{\phi\sqrt{2\pi}} e^{\left(-(y-\mu)^2/2\phi^2\right)}$	μ	ϕ^2	θ	1
Poisson $P(\mu)$	$\frac{\mu^y e^{-\mu}}{y!}$	$\ln(\mu)$	1	$\exp(\theta)$	μ
Bernoulli $B(\mu)$	$\mu^y(1-\mu)^{1-y}$	$\Phi^{-1}(\mu)$	1	$\frac{e^\theta}{1+e^\theta}$	$\mu(1-\mu)$
Négative binomiale $NB(\mu,k)$	$\frac{\Gamma(y+\mu)}{y!\Gamma(\mu)}\frac{\mu^y k}{(\mu+k)^{y+1}}$	$\ln\left(\frac{\mu}{k+\mu}\right)$	1	$\frac{ke^\theta}{1-e^\theta}$	$\mu+\frac{\mu^2}{k}$

Les GLMM peuvent être étendus au cas multivarié, c'est-à-dire considérant plusieurs variables expliquées et explicatives. Les « cumulative link models », notamment utilisés pour l'analyse de variables ordinales, font partie de cette classe de modèle(Fahrmeir and Tutz, 2001). A titre d'exemple, le modèle cumulatif de lien probit est caractérisé par une fonction de lien Φ^{-1} (Φ étant la fonction de répartition de la loi normale centrée réduite) telle que $P(y_i \le j) = \Phi(\mu_{ij})$ et un prédicteur linéaire $\mu_{ij} = \delta_j - x_i'\beta - u_i - p_i$.

Il existe différentes méthodes d'estimation des effets fixes et de composantes de la variance dans les GLMM. Certain nombre de travaux sont présentés par Trottier (1998). Anderson et Aitkin (1985) ont adopté une démarche classique, elle consiste en l'obtention de la fonction de vraisemblance marginale et sa maximisation moyennant des techniques d'intégration numériques. Les différentes intégrales sont ainsi approchées numériquement par quadrature gaussienne par exemple. La méthode « EK » proposée par Engel et Keen (1994), se décompose en deux étapes. La première étape consiste en une linéarisation conditionnelle et la deuxième étape procède ensuite à l'estimation dans ce modèle linéarisé par utilisation des équations d'Henderson pour les modèles linéaires mixtes.

Zeger et Karim (1991) ont, ainsi, proposé un algorithme de Gibbs Sampling pour l'estimation des paramètres. McCulloch (1997)propose également une méthode s'appuyant sur une étape de Metropolis-Hastings conduisant à la construction d'un algorithme de type MCEM (Monte Carlo EM). Gilmour, Anderson et Rae (1985) ont développé une méthode d'estimation dite méthode « GAR » dans un modèle à effet aléatoire pour des données binomiales avec un lien Probit. Foulley et Im (1993) l'ont étendue au cas d'une distribution de Poisson.

2.2.2 GLMM utilisés en génétique sur les données ordinales

Parmi les modèles linéaires généralisés (ou leur extension) utilisés en génétique pour étudier les données ordinales ; le modèle cumulatif de lien probit, ou plus précisément son interprétation en termes de variable latente, est le plus répandu. La théorie de ce modèle à variable latente a été introduite par Wright (1934)pour l'analyse du nombre de doigts des cobayes. Le principe consiste à considérer l'existence d'un phénotype sous-jacent l continu non observé qui suit une loi normale d'espérance μ_u, de variance 1 et $k-1$ seuils strictement croissants pour k réalisations possibles de y tels que $y_i = \begin{cases} 1 \text{ si } l_i \le t_1 \\ 2 \text{ si } t_1 < l_i \le t_2 \\ \cdots \\ k \text{ si } t_{k-1} < l_i \end{cases}$

(Figure 4). Ainsi la probabilité π_j que $y_i = j$ est $\Phi\left(t_k - l_i\right) - \Phi\left(t_{k-1} - l_i\right)$ en posant $t_0 = -\infty$ et $t_k = +\infty$.

De nombreux phénotypes ont été étudiés par ce modèle :la fertilité (Höschele et al., 1987), les difficultés de vêlage (Djemali et al., 1987; Meijering, 1984), la qualité de la laine (Thompson et al., 1985) et la taille de portée dans différentes espèces (Bodin and Elsen, 1989; Matos C.A.P, 1993; Mekkawy et al., 2010; Perez-Enciso et al., 1993; Petersson and Danell, 1985; SanCristobal-Gaudy et al., 2001).
Dans un modèle à seuil, l'héritabilité du caractère est facilement calculable sur l'échelle sous-jacente (Manfredi et al., 1991) $h^2_{l-scale} = \dfrac{\sigma_u^2}{1+\sigma_u^2}$. Du fait de la non linéarité, celle sur l'échelle observée l'est beaucoup moins. Gianola (1982)en a proposé l'approximation suivante :

$$h^2_{y-scale} = h^2_{l-scale} \frac{\sum_{j=1}^{k-1}\left[\phi\left(t_j\right)\left(a_{j+1} - a_j\right)\right]^2}{\sum_{j=1}^{k} a_j^2 \pi_j - \left(\sum_{j=1}^{k} a_j \pi_j\right)^2}$$

$\tau_1 \qquad \tau_2$

Où ϕ est la densité de la loi normale centrée réduite et a_j est le poids attribué à la classe j .
Le modèle de poisson a plus rarement été utilisé en génétique. Néanmoins, Perez-Enciso et al. (1993) ainsi que Matos et al. (1997), et Olensen et al. (1994) l'ont appliqué sur la taille de portée. Le principal problème de ce modèle est l'hypothèse forte d'égalité entre moyenne et variance qui ne s'avère pas toujours exacte. Une surdispersion est souvent observée (variance > moyenne) à cause de l'omission ou d'erreur de variables explicatives, d'une mauvaise spécification du modèle ou de non-homogénéité de l'environnement. Pour pallier ce problème il est possible de prendre en compte la surdispersion en considérant que le paramètre μ suit une loi gamma. La distribution de y devient alors binomiale négative. Ce modèle a été utilisé à titre d'exemple pour analyser le nombre d'inséminations avant conception chez les bovins laitiers (Tempelman and Gianola, 1999).
Remarquons néanmoins que, bien que théoriquement les GLMM soient plus appropriés pour l'analyse des variables ordinales que le modèle linéaire(Clogg and Shihadeh, 1994; Gianola, 1980; Thompson, 1979) ; de nombreux auteurs ont montré qu'il existait peu de différences au niveau de l'estimation des effets fixes et de la prédiction des effets aléatoires (donc le classement des animaux) entre les LMM et les GLMM(Olesen, 1995; Weller and Ron, 1992).

2.3 Le modèle polygénique hétéroscédastique

Dans le cas du modèle linéaire ; la prise en compte d'une hétérogénéité de variance est réalisée en considérant que la variance résiduelle est une fonction f d'un certain nombre d'effets environnementaux (et génétiques) $f(\boldsymbol{\eta}, \boldsymbol{v})$ où $\boldsymbol{\eta}$ **et** $\boldsymbol{v}$ correspondent aux vecteurs des effets fixes et génétiques additifsinfluençant la variance résiduelle (par soucis de simplicité nous n'avons pas considéré ici d'effets environnementaux aléatoires dans le modèle). Autrement dit,

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{Wp} + f(\boldsymbol{\eta}, \boldsymbol{v})\boldsymbol{\varepsilon}'$$

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{p} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{A}\sigma_u^2 & 0 & 0 \\ 0 & \boldsymbol{I}_n\sigma_p^2 & 0 \\ 0 & 0 & \boldsymbol{I}_N \end{bmatrix}\right).$$

ε' est le vecteur des résidus indépendamment et identiquement distribués suit une loi normale $N(0,1)$.

Plusieurs fonctions f positives ont été proposées dans la littérature: la fonction exponentielle (Neves et al., 2012a; Rönnegård et al., 2010; SanCristobal-Gaudy et al., 1998b) ou carré (Mulder et al., 2007). Néanmoins, récemment, Garcia et Ibáñez-Escriche(2009) ont montré qu'une fonction identité permettait également d'obtenir des estimations sans biais des paramètres.

Différentes méthodes d'estimation des paramètres des modèles linéaires hétéroscédastiques ont été proposées. Certains auteurs se limitent à analyser directement l'écart-type des observations. Ce qui revient à ignorer dans le calcul de cet écart-type tous les facteurs environnementaux et génétiques qui affectent la moyenne (Bodin et al., 2010). La méthode dite « en 2 étapes » décrite dans Garreau *et al*, (2004) et Guitterez *et al*. (2006) consiste à résoudre en deux étapes successives le modèle hétéroscédastique proposé par SanCristobal *et al*. (1998)

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{Wp} + \boldsymbol{e}^{\frac{1}{2}(\boldsymbol{X\tilde{\beta}}+\boldsymbol{Zv}+\boldsymbol{W\tilde{p}})}\,\varepsilon'$$

La première étape a pour but d'estimer les paramètres d'un premier modèle linéaire homoscédastique

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{Wp} + \varepsilon$$

Et transformer les résidus obtenus$\hat{\boldsymbol{e}} = \boldsymbol{y} - (\boldsymbol{X\hat{\beta}} + \boldsymbol{Z\hat{u}} + \boldsymbol{W\hat{p}})$ pour créer une nouvelle variable : $\tilde{\boldsymbol{y}} = \ln(\hat{\boldsymbol{e}}^2)$

La deuxième étape complète le modèle hétérocédastique en estimant les paramètres d'un second modèle linéaire : $\tilde{\boldsymbol{y}} = \boldsymbol{X\tilde{\beta}} + \boldsymbol{Zv} + \boldsymbol{W\tilde{p}} + \tilde{\varepsilon}$

Cette méthode simple permet une estimation très rapide mais approximative des paramètres. Plusieurs autres algorithmes ont donc été proposés pour estimer de manière itérative les facteurs agissant sur la moyenne et la variabilité du caractère : San Cristobal-Gaudy et al. (1998b) proposent l'algorithme EM, l'algorithme LINHE-ML qui fournit une estimation ML et utilise pour le calcul des moments les distributions conditionnelles (Trottier, 1998), l'algorithme LINHE-REML basé sur les travaux de Foulley et Quaas (1995), il fournit quant à lui une estimation REML utilisant les fonctions d'Henderson, et l'algorithme de MCMC basé sur la méthode de Gibbs Sampling (Garcia M and Ibáñez-Escriche; Sorensen and Waagepetersen, 2003).

Dans le cas des GLMM, il n'est pas toujours possible d'introduire une hétérogénéité de variance. En effet, pour certaines distributions définies par un seul paramètre (Poisson, Binomiale), on ne peut casser la relation fonctionnelle qui existe entre moyenne et variance.

Pour les distributions définies par deux paramètres (modèle à seuils), la modélisation d'une hétérogénéité de variance dans le modèle à seuil a été réalisée au niveau de la variable latente l de manière analogue à celle proposée dans le modèle linéaire hétéroscédastique $\mathbf{t} = (0,1,\tau_3,\ldots,\tau_{n-1})$; $l|u,v \sim N\left(\mu + u, f^2(\eta,v)\right)$.

Certains auteurs n'ont inclus dans cette hétérogénéité que des effets environnementaux fixes (Ducrocq, 2000b; Foulley and Gianola, 1996), d'autres aléatoires ; Foulley et Jaffrézic (2010) ont proposé un modèle à seuils heteroscédastique pour lequem la variance résiduelle dépend d'effets fixes et d'effets génétiques. Cette derniére modélisation a été utilisé pour traiter des données biomédicales avec le logiciel Winbugs/Openbugs. Les nombreuses données utilisées en génétique quantitative rendent cette implémentation irréalisable en raison de temps de calcul trop important.

Parmi les méthodes d'estimations utilisées on trouve la méthode itérative décrite par Trottier (1998). Chaque itération se présente en deux étapes :

La linéaration : est réalisée par l'introduction d'une variable dépendante comme dans le modèle GLM

L'estimation : s'appuie essentiellement sur les étapes suivies dans le cadre du modèle linéaire polygénique hétéroscédastique pour lesquelles nous pouvons utiliser les algorithmes décrits au-dessus.

PARTIE 2 : Modélisation des variabilités de la taille de portée des ovins comme une variable continue

Résumé:

Dans cette partie, nous avons modélisé les variabilités de la taille de portée, sans tenir compte de sa spécificité discrète. En effet afin, de mettre en œuvre les différentes approches d'analyse invoquées précédemment, il est nécessaire, de connaitreles principaux facteurs environnementaux qui influencent les variabilités de la taille de portée. Nous disposons pour cela de données de trois races dont la prolificité couvre l'ensemble des prolificités rencontrées en France.

Sommaire

1 Origine des données 30
1.1 Choix des races 30
1.2 Origine des données 30
1.3 Les échantillons complets 32
1.4 Les échantillons simplifiés 32
2 Modélisation de la "moyenne" : Modèle linaire mixte homoscédastique 34
2.1 Recherche des facteurs environnementaux du modèle. 34
2.2 Modèles génétiques 34
2.3 Résultats et discussion 35
Effet de l'âge de la brebis : 35
Effet de l'interaction âge à la première mise bas et la parité de mise bas : 35
Effet de l'intervalle entre agnelages : 35
Effet de la saison : 35
Effet de l'élevage : 36
Les facteurs de variation génétiques 36
3 Modélisation de la variance résiduelle : Modèle linéaire mixte hétéroscédastique .. 40
3.1 Résultats de la modélisation de la variance résiduelle 40
3.2 Conclusion 41

1 Origine des données

1.1 Choix des races

Les données françaises utilisées au cours de cette thèse concernent trois races différentes. Nous avons d'abord considéré la race Romane. C'est une race synthétique créée par l'INRA à partir de croisements entre les races Romanov (prolifique) et Berrichon du cher (race bouchère). C'est maintenant une race fixée qui a conservé les caractéristiques moyennes de ces races parentales: elle est en effet à la fois très productive et avec des très bonnes qualités maternelles et bouchère. Alors que les races prolifiques Romanov et Finnoise, d'effectifs très réduits, sont essentiellement exploitées comme races de mâles pour produire des femelles croisées prolifiques, la race Romane est exploitée pour produire directement des agneaux. C'est ainsi la race française de brebis à viande (tableau 4) qui est la plus prolifique (prolificité brute de 2.16 à la campagne 2013) et dont environ 43% et 7% des agneaux naissent respectivement triples ou plus et quadruples ou plus. Les fortes fréquences de taille de portée élevée qui permettent des analyses statistiques précises, ainsi que l'importance des effectifs contrôlés dont on connait la généalogie ont imposé le choix de cette race pour l'application des méthodes développées au cours de cette thèse. Pour tester si ces méthodes restent valables lorsque les fréquences de portées multiples diminuent nous avons aussi considéré deux autres races. La race Suffolk pour laquelle le pourcentage de mises bas multiples (triples et plus) est voisin de 5% (C'est sans doute la limite inférieure en deçà de laquelle il n'y a plus d'intérêt à réduire la variabilité des tailles de portée) et la Rouge de l'Ouest. La race Suffolk est la race la moins prolifique des 3 races choisies : sa prolificité moyenne brute est de 1.63 à la campagne 2013. C'est une race bouchère élevée principalement en plein air et dont la vitesse de croissance des agneaux est parmi les plus élevées de toutes les races. La Rouge de l'Ouest présente une prolificité élevée (1.89) et a été utilisée pour valider les méthodes et les tester sur une structure de population différente de la Romane. C'est aussi une race bouchère d'herbage de très bonne qualité maternelle.
Finalement ces trois races recouvrent le domaine de variation des prolificités des races françaises qui ont au moins 5 % de mises bas triples ou plus et pour lesquelles il serait possible de réduire la variabilité des tailles de portée.

1.2 Origine des données

Les données que nous avons utilisées proviennent de la BDIR ovinegérée par l'IDELE et le CTIG à l'INRA de Jouy en Josas. C'est une "Base de Données Indexation et Recherches" qui sert de support à l'évaluation génétique et aux équipes de recherche pour leurs études. Cette base rassemble une partie des données validées et traitées en provenance du contrôle de performances et qui sont stockées dans "OVALL", qui constitue le Système d'Information Génétique Ovins Allaitants.
Les informations proviennent d'une part d'un fichier rassemblant les mises bas de toutes les femelles depuis 35 ans avec les données qui y sont attachées (élevage, rang et date de mise bas, mode de reproduction, mortalité des agneaux nés, nb d'agneaux élevés sous la mère, etc.), et d'autre part d'un fichier de généalogie global qui regroupe les informations sur la naissance (élevage date et mode de naissance, père et mère, mode d'élevage et mode d'allaitement, etc.) de tous les individus nés depuis la même époque.

Tableau 4 : Effectif, prolificité naturelle, variance, coefficient de variation et pourcentages de chaque taille de portée pour les brebis adultes de 38 races contrôlées en France à la campagne 2013.
(Données du Bilan de contrôle de performances ovin allaitant, campagne 2013; IE - INRA).

Race	n	Prolif.	Var	CV	%TP1	%TP2	%TP3	%TP4+
Charmoise	2904	1.12	0.11	0.30	87.7	12.2	0.10	0.00
Castillonnaise	1154	1.14	0.12	0.31	85.7	14.3	0.00	0.00
Lourdaise	274	1.15	0.14	0.32	85.0	14.6	0.37	0.00
Barégeoise	1502	1.18	0.15	0.33	82.5	17.2	0.27	0.00
Aure et Campan	702	1.18	0.15	0.32	82.2	17.8	0.00	0.00
Tarasconnaise	6059	1.22	0.18	0.35	78.1	21.5	0.41	0.00
Mérinos d'Arles	10792	1.23	0.19	0.35	77.2	22.4	0.32	0.02
Mourerous	3421	1.27	0.20	0.36	73.5	26.2	0.26	0.03
Boulonnaise	1055	1.33	0.24	0.37	68.2	30.7	1.04	0.00
Limousine	5515	1.38	0.27	0.37	63.9	34.7	1.40	0.05
Préalpes	6491	1.40	0.28	0.38	62.2	36.0	1.65	0.09
Hampshire	571	1.46	0.26	0.35	54.8	44.8	0.35	0.00
Bizet	2521	1.46	0.27	0.36	55.6	43.1	1.27	0.00
Blanc du Massif Central	18390	1.51	0.32	0.37	52.5	44.5	2.92	0.15
Solognote	1254	1.51	0.28	0.35	50.2	48.4	1.28	0.08
Dorset Down	90	1.53	0.32	0.37	50.0	46.7	3.33	0.00
Southdown	241	1.54	0.32	0.37	49.8	46.9	3.32	0.00
Rava	7986	1.57	0.33	0.36	47.1	49.1	3.69	0.10
Berrichon du cher	954	1.60	0.33	0.36	44.7	51.2	4.19	0.00
Caussenard du Lot	23580	1.62	0.33	0.35	42.5	52.8	4.58	0.07
Suffolk	**2281**	**1.63**	**0.34**	**0.36**	**42.3**	**52.8**	**4.78**	**0.13**
Berrichon de l'Indre	686	1.64	0.34	0.36	41.4	53.5	4.81	0.29
Noire du Velay	5564	1.67	0.45	0.40	42.7	49.2	7.01	1.13
Est à Laine Mérinos	2289	1.70	0.33	0.34	35.9	58.5	5.55	0.13
Ile de France	13461	1.72	0.42	0.38	38.2	52.9	8.13	0.79
Mouton Vendéen	5520	1.72	0.41	0.37	37.3	53.6	8.44	0.63
Roussin	1289	1.76	0.31	0.31	29.7	64.2	5.97	0.08
Mouton Charollais	5957	1.77	0.38	0.35	32.4	58.3	8.91	0.42
Avranchin	112	1.78	0.36	0.34	31.3	59.8	8.93	0.00
Cotentin	77	1.78	0.44	0.37	35.1	51.9	12.99	0.00
Texel	3795	1.82	0.40	0.35	30.0	58.0	11.67	0.32
Lacaune Viande	7030	1.85	0.56	0.40	33.3	51.4	12.66	2.65
Rouge de l'Ouest	**2936**	**1.89**	**0.43**	**0.35**	**26.2**	**59.5**	**13.22**	**1.06**
Bleu du Maine	610	1.91	0.49	0.37	28.2	53.9	17.05	0.82
Grivette	4939	1.99	0.58	0.38	25.3	54.4	17.13	3.18
Romane	**15650**	**2.16**	**0.57**	**0.35**	**17.4**	**53.2**	**25.75**	**3.69**
Finnoise	186	2.72	1.21	0.41	15.1	28.0	32.80	24.19
Romanov	447	3.13	1.04	0.33	4.0	22.8	39.15	34.00

1.3 Les échantillons complets

Les données concernent les mises bas qui ont eu lieu entre 1990 et 2013. Pour des raisons de simplification des modèles utilisés par la suite, seules les mises bas consécutives à des montes naturelles (sans traitement hormonal préalable) ont été retenues. Les effectifs et les caractéristiques pour chaque race ainsi que les informations individuelles récoltées concernant chaque brebis, les données de chaque mise bas et les données déduites sont présentés dans le tableau 5.
Le pedigree des brebis en production a été extrait du fichier de généalogie global en remontant sur 10 générations. Il constituait un fichier de triplets (individu, père, mère) dont les effectifs sont aussi rapportés dans le tableau 5. Ces différentes informations ont été finalement recodifiées (Logiciel Recode) pour permettre l'utilisation des programmes d'analyses génétiques.

1.4 Les échantillons simplifiés

Pour permettre de tester et valider les différentes méthodes avec des temps de calcul plus court (quelques minutes au lieu des heures voire des jours pour certaines données et certains algorithmes), nous avons été amenés à travailler dans un premier temps sur des échantillons plus réduits mais pertinents. Ces échantillons réduits regroupent pour chacune des races (tableau 5) les performances enregistrées entre 1990 et 2013 pour les brebis qui ont eu au minimum 3 mises bas.

Tableau 5 : Caractéristiques des échantillons complets et réduits pour les trois races (Romane, Suffolk et Rouge de l'ouest)

Données sur chaque brebis			
Identification de la brebis Elevage naisseur Elevage détenteur Campagne de naissance Date de naissance Identification de la mère génétique Identification du père			
Données sur chaque mise bas	**Données déduites**		
Identification de la brebis Elevage détenteur Mode de fécondation Campagne d'agnelage Date de mise bas Taille de portée Parité ou rang de portée	Saison de mise bas Age à la mise bas Intervalle entre mise-bas Taille de portée MB précédente		
Race	**Suffolk**	**Rouge de l'ouest**	**Romane**
Echantillons complets			
Effectifs concernés			
Nombre de brebis	43 106	62 553	58 028
Nombre de portées	75 218	186 864	137 358
Nombre de pères connus	2 809	4 465	1 617
Nombre de mères connues	20 206	37 674	22 211
Nombre d'élevages de sélection	308	300	159
Données du pédigrée			
Individus dans le fichier pédigrée	43 106	75 247	58 028
Echantillon réduits			
Effectifs concernés			
Nombre de brebis avec plus de 3 MB	6 404	14 811	25 735
Nombre de portées	30 328	99 384	72 623
Nombre de pères connus	1 939	4 135	1 318
Nombre de mères connues	8 848	31 006	11 810
Nombre d'élevages de sélection	177	250	101
Données du pédigrée			
Individus dans le fichier pédigrée	15 391	62 553	25 735

2 Modélisation de la "moyenne" : Modèle linaire mixte homoscédastique

2.1 Recherche des facteurs environnementaux du modèle.

Une première analysesans effet génétique a été faite sur les données des échantillons réduits pour déterminer les facteurs de variations environnementaux qui avaient un effet significatif sur la taille de portée des brebis et qui devaient être conservés dans le modèle génétique homoscédastique. Le modèle utilisé était un modèle mixte (ASReml) dans lequel l'effet du troupeau et l'effet permanent ont été considérés comme des effets aléatoires à côté des effets environnementaux testés qui étaient considérés comme des effets fixes.
Le choix des facteurs du modèle a été fait en utilisant la méthode pas à pas (Stepwise). Cette méthode est une amélioration de la méthode d'introduction progressive. A chaque étape de la procédure, on examine si une nouvelle variable doit être ajoutée et si une des variables déjà incluses doit être éliminée. La procédure s'arrête lorsqu'aucune variable ne peut être rajoutée ou retirée du modèle selon les critères d'inclusion et d'exclusion choisis. Nous avons utilisé le critère d'Akaike (AIC)comme critère de comparaison des modèles. Les différents facteurs qui ont été testés puis retenus sont présentés dans le tableau 6.

Tableau 6 : Facteurs testés et retenus pour les analyses génétiques.

Facteurs testés	Facteurs retenus
Numéro cheptel détenteur de l'animal	√
Campagne d'agnelage	√
Age de la brebis	√
Mode d'élevage de la brebis	
Méthode d'allaitement	
Parité	√
Nombre d'agneaux allaités à la mise-bas précédente par la mère allaitante de l'agneau	√
Saison de mise bas	√
Age à la première mise bas	√
Intervalle entre mise-bas	√
Age à la première mise bas x âge	√
Age à la première mise bas x parité	√

L'analyse statistique a mis en évidence l'influence très significative des facteurs suivants :Le numéro cheptel détenteur de l'animal considéré comme une variable aléatoire,la campagne d'agnelage, l'âge de la brebis, le numéro de parité, la saison de la naissance, l'âge à la première mise bas,l'intervalle entre agnelage. Les combinaisons : âge x âge à la première mise bas et numéro de mise bas x âge sont les principales interactions qui ont une influence sur la taille de portée des ovins.

2.2 Modèles génétiques

La seconde étape d'analyse consistait à estimer les paramètres génétiques contrôlant la taille de portée, en utilisant un modèle linéaire mixte qui considérait les facteurs environnementaux sélectionnés dans l'étape précédente, un effet animal permanent ainsi qu'un effet animal génétique. Ces analyses ont été faites, avec le logiciel ASReml (Gilmour, Gogel, Cullis and Thompson , 2009).

2.3 Résultats et discussion

Effet de l'âge de la brebis:

La figure 6 montre l'effet de l'âge de la brebis à la mise bas sur le nombre d'agneaux nés. Nous avons observé une augmentation puis une diminution de la taille de portée avec l'âge de la brebis pour les trois races. Le nombre maximum d'agneaux nés était atteint à l'âge de 5-6 ans pour la Suffolk et la Rouge de l'Ouest, alors qu'il était atteint un peu plus tôt pour la Romane de 3-4 ans. On a pu observer que l'augmentation entre la prolificité de début de carrière et la prolificité maximum était plus élevée pour la Suffolk (+0.20 entre 2 et 6 ans), que pour la RW (+0.13 entre2 et 6 ans), et que pour la Romane (+0.10 entre 2 et 4 ans).
Le nombre d'agneaux nés augmente avec l'âge et la parité des femelles (Notter, 1980). Cette évolution résulte d'un accroissement du taux d'ovulation, d'une plus grande capacité de l'utérus, et d'une amélioration des autres caractères maternels reliés à la reproduction (Fahmy, 1990). Karam (Karam 1957) a aussi montré que la prolificité atteint son pic entre 5 et 7 ans chez la race Rahmani. Par la suite, à un âge voisin de 5ans, on observe généralementun déclin de la taille de portée (Dickerson and Glimp, 1975; Notter, 2000). Cette diminution avec l'âge est peut être reliée à la dégradation de la qualité des gamètes chez les femelles (García-Ispierto et al., 2007). La prolificité de la race Romane diminue bien avant 5 ans; un résultat similaire a été trouvé pour la race Galway (O'Ferrall, 1976) qui est aussi une race prolifique.

Effet de l'interaction âge à la première mise bas et la parité de mise bas :

L'âge à la première mise bas est un bon indicateur de la précocité sexuelle de la brebis pour un système de production donné. Nous avons trouvé que cet effet est significatif pour les 3 races : les femelles qui mettent bas plus tardivement ont des tailles de portée plus élevées (figure 7, 8, 9). De nombreux résultats allant dans ce sens ont été rapportés (Lajous, 1987; Waldron and Thomas, 1992). Cet effet est certainement liéau fait que les jeunes femelles qui mettent bas n'ont pas encore terminé leur croissance et atteint leur poids adulte.
La parité ou numéro de mise bas intervient de façon significative ($P<0.001$), sans que l'on puisse dégager une relation simple. De plus nous avons observé pour les trois races une interaction significative ($P<0.001$) entre l'âge à la première mise bas etle numéro de mise bas.

Effet de l'intervalle entre agnelages :

C'est l'intervalle de temps qui sépare deux mises bas et qui est directement relié à la durée de l'intervalle post-partum qui sépare une mise bas de la fécondation suivante. Nous avons observé quelaprolificitéaugmente avecl'intervalle postpartum précédent (+0.10 lorsque l'intervalle entre agnelages est supérieur à un an, Figure 10). Cependant, ce lien ne suppose aucune relation causale entre la prolificité et l'intervalle entre agnelage. On peut supposer qu'un intervalle post-partum long permet de restaurer les réserves corporelle de la femelle et que sa durée peut avoir un effet positif sur la prolificité comme il est observé pour la fertilité (David et al., 2008; TorrèS and Cognié, 1984). On peut également supposer que les femelles à forte prolificité ont tendance à avoir un grand intervalle d'agnelage (Komprej et al., 2011; Tosh, 2011).

Effet de la saison :

La saison n'a eu d'effet significatif ($P < 0.05$) que pour les races Suffolk et Rouge de l'Ouest, toutefois, il n'y a pas une grande différence entre les estimations de chaque niveau de cet effet. La taille de portée était légèrement plus élevée en saison 2 (entre Janvier et Mai) qu'en saison 1 (entre Juin et Décembre)pour la Suffolk(+0,01), alors que c'était l'inverse pour la Rouge de l'Ouest (-0,02) (Figure 11). En fait, les performances de prolificité sont meilleures lorsque les fécondations ont lieu en pleine saison sexuelle (ce qui correspond aux mises bas de la saison 1) et/ou que les ressources alimentaires sont en quantité importante et de bonne qualité (ce qui correspond aux mises bas de la

saison 2) (Faugere et al., 1990). Cet effet de la saison n'est pas toujours mis en évidence, par exemple (Notter et al., (2000) n'a pas trouvé d'effet significatif de la saison sur la taille de portée TP chez la Suffolk; ce qui pourrait venir de différences de méthode et de milieu d'élevage.

Effet de l'élevage :

Cette étude a montré un effet important de l'élevage sur la taille de la portée (P<0.001). Cet effet explique 7% à 10% de la variabilité de la taille de portée selon les races. Ce résultat classique, rapporté par d'autres auteurs, pour d'autres races, et dans d'autres régions du monde(Bunge et al., 1990),illustre la dépendance des performances reproductives à l'environnement.

Les facteurs de variation génétiques

Les héritabilités, les répétabilités et les variances génétiques associées sont représentées dans le tableau 7. Pour les 3 races, la somme des composantes de la variance $\Sigma^2 = \sigma_u^2 + \sigma_p^2 + \sigma_h^2$ ne représente que 15 à 17% de la variance globale des observations (tableau 4), alors que la composante résiduelle représente entre 73 et 80% de cette variabilité selon les races. La variance totale expliquée ainsi que la composante résiduelle sont proportionnelles aux prolificités moyennes des races (figure 5), ce qui n'est pas très étonnant étant donné la proportion que représente ces composantes, et la relation de liaison moyenne variance observée sur les données brutes (figure 2; page15). Par contre, on note que la somme des effets génétiques et permanents est aussi directement proportionnelle à la prolificité moyenne de la race.

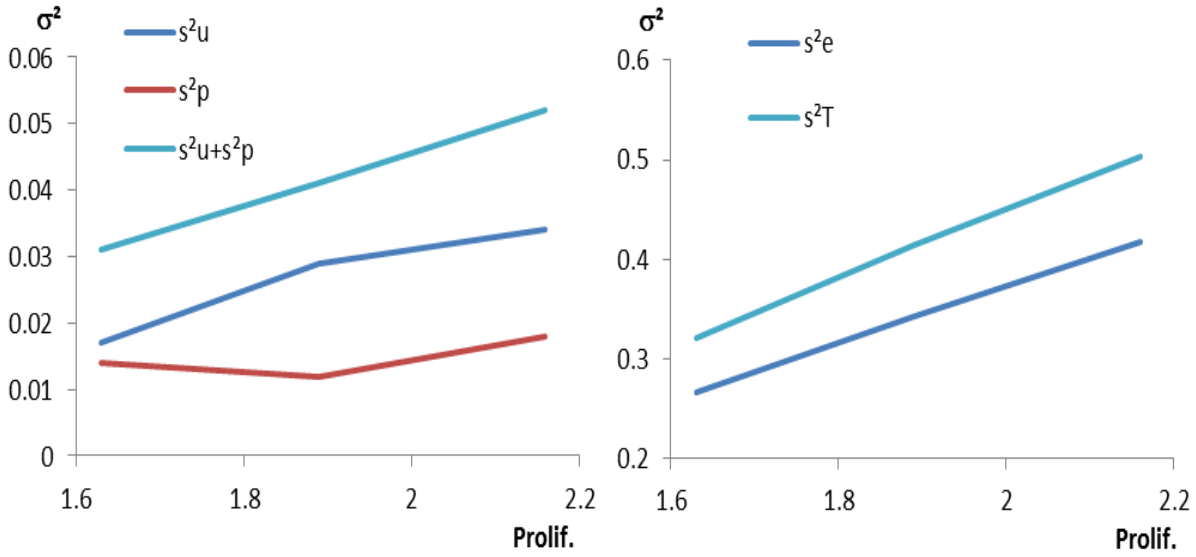


Figure 5 : Composantes de variance génétique, permanente et leur somme, ainsi que variance résiduelle et totale en fonction de la prolificité moyenne de la race.

Tableau 7 : Les estimations des paramètres génétiques et environnementaux de la fonction moyenne

	Races	σ_u^2	σ_p^2	σ_h^2	σ_e^2	h^2	r
Modèle linéaire homoscédastique	Suffolk	0.017	0.014	0.024	0.267	$0.052_{(0.005)}$	$0.096_{(0.005)}$
	Rouge de l'Ouest	0.029	0.012	0.031	0.343	$0.070_{(0.004)}$	$0.100_{(0.004)}$
	Romane	0.034	0.018	0.033	0.418	$0.068_{(0.003)}$	$0.078_{(0.003)}$

σ_u^2 :

On constate que la taille de portée est faiblement héritablepour les trois races. En effet, l'héritabilité varie de 0.052 à 0.070, mais elle est toujours significativement différente de zéro puisque sa précision varie de 0.003 à 0.005. Comme nous l'avons vu dans la partie bibliographique, ces valeurs sont dans la fourchette des valeurs rapportées pour ce paramètre dans l'espèce ovine.
La répétabilité varie entre les races de 0.078 à 0.100. Ces valeurs sont aussi dans la gamme habituelle mentionnée dans la littérature (voir par exemple Rao et Notter 2(000), Waldron et Thomas (1992)).

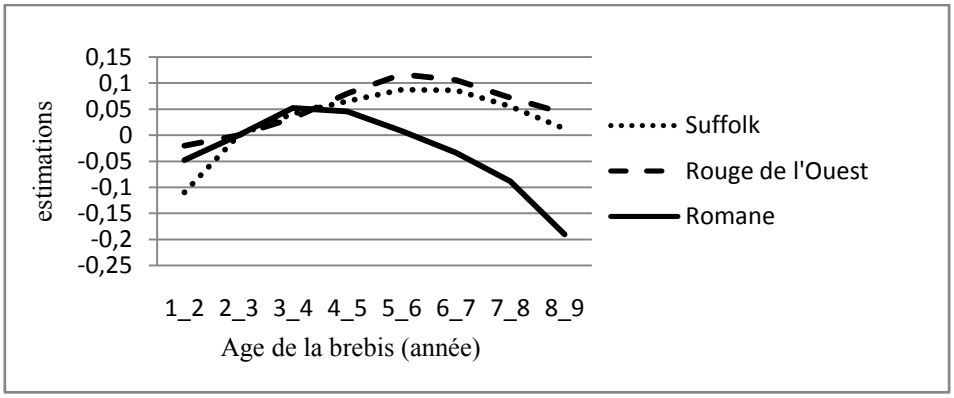


Figure 6 : Influence de l'âge de la brebis sur la taille de portée des trois races (Suffolk, Rouge de l'Ouest, et Romane).

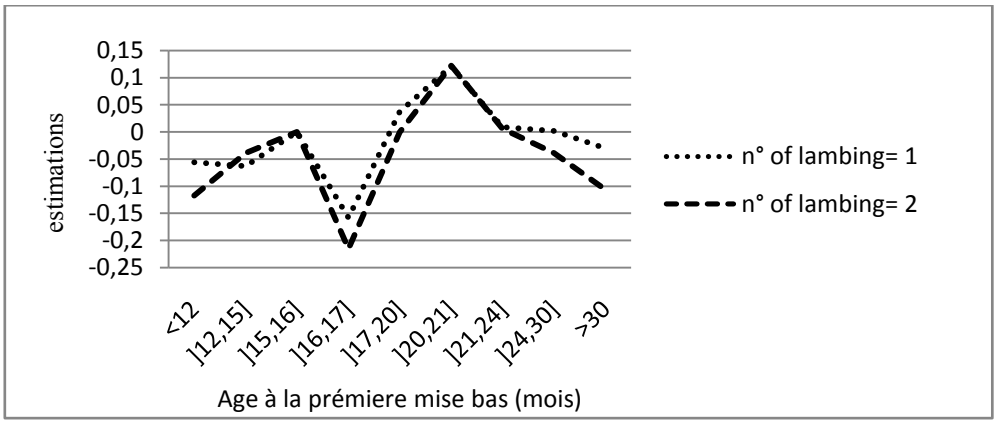


Figure 7 : Influence de l'âge à la première mise bas sur la taille de portéede la race Suffolk.

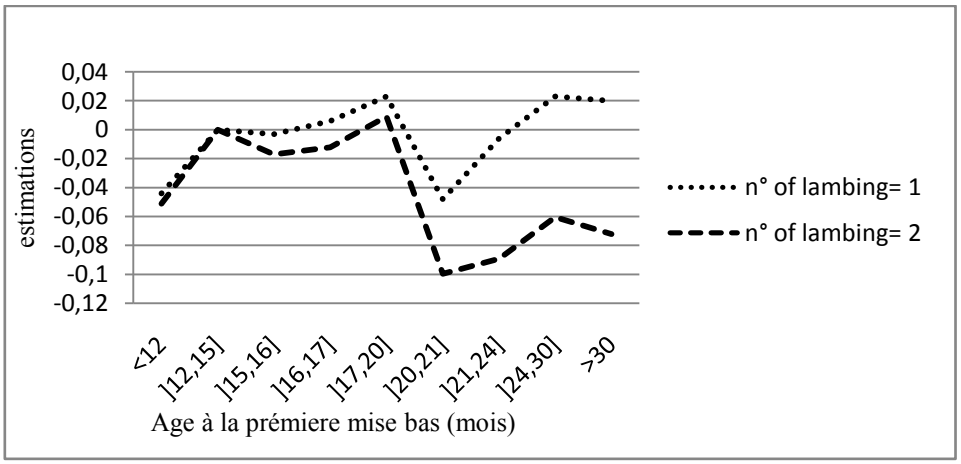


Figure 8 : Influence de l'âge à la première mise bas sur la taille de portéede la race Rouge de l'Ouest.

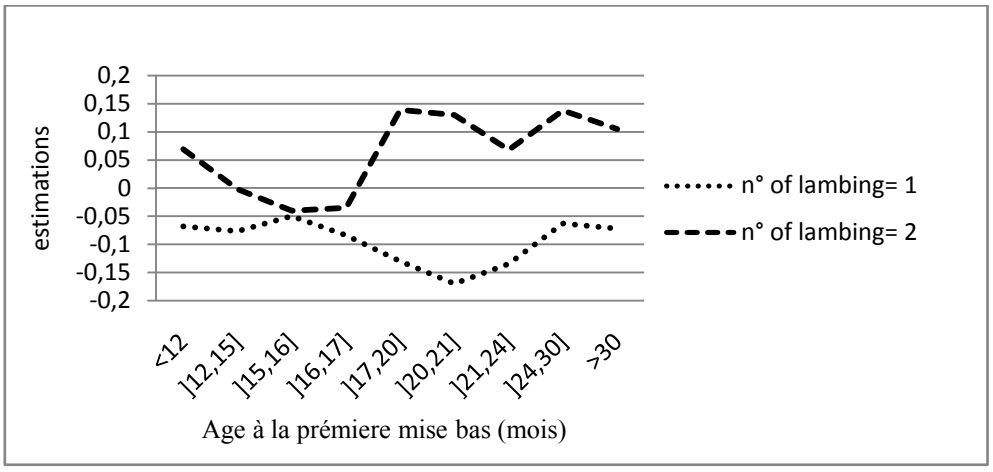


Figure 9 : Influence de l'âge à la première mise bas sur la taille de portéede la race Romane.

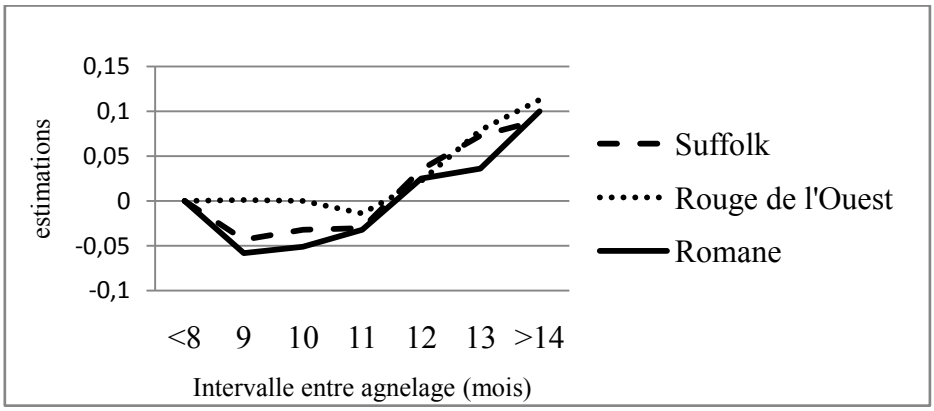


Figure 10 : Influence de l'intervalle entre agnelage sur la taille de portée des trois races (Suffolk, Rouge de l'Ouest, et Romane).

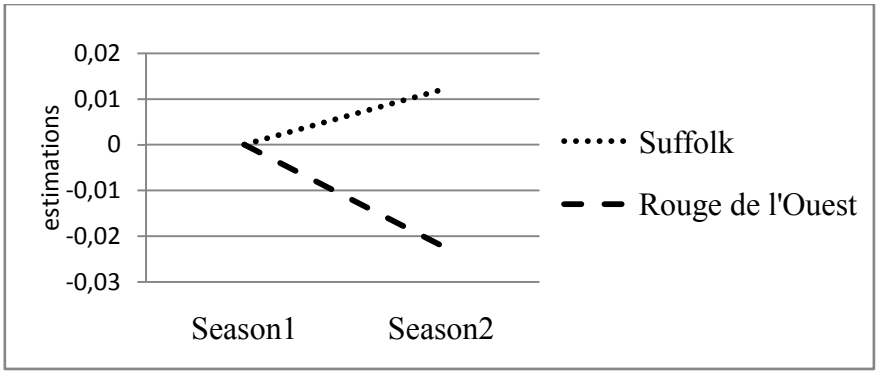


Figure 11 : Influence de la saison sur la taille de portée des deux races (Suffolk et Rouge de l'Ouest).

3 Modélisation de la variance résiduelle : Modèle linéaire mixte hétéroscédastique

Par mesure de simplification, nous avons dans un premier temps modélisé la variance environnementale de la taille de portée (considérée comme une variable continue) avec une procédure à deux étapes décrite dans la partie 1.
Dans notre cas, et pour les trois races concernées, les paramètres et les solutions du modèle de la première étape ont été estimés dans la partie précédente : "Modélisation de la fonction moyenne : modèle linaire mixte homoscédastique", et la significativité des facteurs de variation environnementaux à prendre en compte dans la deuxième étape a été testée par la procédure stepAIC sur le logiciel R (la sélection a été faite sur le critère BIC, le BIC aboutit à des modèles plus parcimonieux que le critère d'AIC dans le cas des grands échantillons). Le meilleur modèle retenu pour cette étape est celui impliquant les facteurs environnementaux troupeau, parité et intervalle post-partum.

3.1 Résultats de la modélisation de la variance résiduelle

Les estimations des paramètres génétiques pour les trois races, ainsi que la corrélation de Pearson ρ entre les valeurs $\hat{u}$ (EBV sur la moyenne), et $\hat{v}$ (EBV sur la variabilité environnementale) sont présentées dans le tableau 8.
L'espérance des résidus de la première étape $E(\varepsilon)$ étant proches de 0, et leur variance (σ^2_ε) étant faible, la variance des logarithmes des carrés des résidus $\sigma^2_{\ln(\varepsilon^2)}$ est très supérieure à la variance des résidus. ($\sigma^2_{\ln(\varepsilon^2)} \gg \sigma^2_\varepsilon$). Nous observons aussi, comme d'autres auteurs(Rowe et al., 2006; Neves et al., 2012) que les variances génétiques du modèle sur la variance sont supérieures à celles du modèle sur la fonction moyenne. Bien que la variance génétique de la variabilité des tailles de portée (σ^2_v) apparaisse fortement liée à la prolificité, on observe que pour les 3 races, la part de la composante génétique dans la variance globale de la variabilité est pratiquement constante (~6%) et indépendante de la prolificité de la race contrairement à la part de la composante génétique qui entre dans le contrôle de la moyenne du caractère (h^2_u).Les corrélations de Pearson $\rho_{\hat{u}_i,\hat{v}_i}$ entre les valeurs génétiques estimées dans la première étape pour la moyenne du caractère ($\hat{u}_i$) et dans la deuxième étape pour sa variabilité ($\hat{v}_i$)sont hétérogènes ; la corrélation $\rho_{\hat{u}_i,\hat{v}_i}$ est positive pour la Romane (+0.72), et négative pour la Rouge de l'Ouest (-0.58) et la Suffolk (-0.89). Selon cette procédure en deux étapes, la relation entre ces valeurs génétiques varietrès fortement selon les races ou selon le niveau de prolificité. Il n'y a pratiquement pas de données dans la littérature sur la corrélation entre valeurs génétiques pour la moyenne et la variance de caractères discrets. La seule information disponible concerne une corrélation de 0.19rapportée par SanCristobal-Gaudy et al., (2001)pour la taille de portée des ovins de race Lacaune; race dans laquelle un gène majeur contrôlant la prolificité a été trouvé. D'autres informations concernent des caractères continus et les valeurs rapportées sont très variables selon les caractères et les espèces : -0.92 pour la taille de portée chez les souris (Gutiérrez et al., 2006), 0.96 pour le poids chez les lapins (Garreau et al., 2008b), 0.40 pour le poids chez les bovins (Nellore) (Neves et al., 2012b).

Tableau 8 : Estimations des composantes de la variance du modèle sur la variabilité de la taille de portée et corrélations entre $\hat{u}$ et $\hat{v}$.

Races	Prolif.	σ^2_{Totale}	σ^2_v	$\sigma^2_{\tilde{p}}$	$\sigma^2_{\tilde{h}}$	$\sigma^2_{\tilde{e}}$	R	$\rho_{\hat{u}_i,\hat{v}_i}$
Suffolk	1.63	1.87	0.11	0.10	0.17	1.40	0.059	-0.89
Rouge de l'Ouest	1.89	6.00	0.35	0.22	0.48	5.00	0.058	-0.58
Romane	2.16	7.28	0.44	0.25	0.45	6.26	0.060	0.72

σ^2_{Total} : Variance totale, σ^2_v : Variance des effets génétiques additifs, $\sigma^2_{\tilde{p}}$: Variance des effets d'environnement permanent, $\sigma^2_{\tilde{h}}$: Variance des effets troupeaux, $\sigma^2_{\tilde{e}}$: Variance résiduelle, $R = \frac{\sigma^2_v}{\sigma^2_{Totale}}$,

$\rho_{\hat{u}_i,\hat{v}_i}$: corrélation entre $\hat{u}$ et $\hat{v}$.

3.2 Conclusion

Cette étude montre qu'en utilisant une méthode approximative sur une variable grossière (la taille de portée n'est pas une variable continue), il est possible de mettre en évidence des effets environnementaux et génétiques qui affectent la variabilité de la taille de portée des brebis. La composante génétique est sans doute faible mais elle existe et selon cette modélisation, son importance relative ne semble pas liée à la prolificité moyenne de la race. On sait par ailleurs que de prendre en compte le caractère discret de la variable taille de portée ne change généralement pas fondamentalement les résultats des analyses du caractère (son niveau moyen) ; il est vraisemblable qu'en raison de la liaison très forte entre moyenne et variance il n'en soit pas de même pour les analyses de sa variabilité. Dans cas, les analyses ne pourront pas être faites par la méthode "en deux étapes" qui est une méthode approchée et il faudra soit adopter une méthode fréquentiste itérative comme celle proposée par Ronnegard *et al*. (Rönnegård et al., 2010), soit une méthode bayésienne comme proposée par Sorensen (Sorensen and Waagepetersen, 2003). La question qui se pose maintenant est comment prendre en compte le caractère discret de la variable pour l'étude de sa variabilité.

PARTIE 3 : Modélisation des variabilités de la taille de portée des ovins comme une variableordinal

Résumé:

Comme évoqué lors de la partie une, la non linéarité des caractères est considéré comme l'une des problématiques des généticiens.Le modèle tel qu'il est décrit au paragraphe précédent ne permet de traiter que la variance résiduelle pour les variables gaussiennes. Parmi l'abondante littérature, de nombreux modèles ont été explorés, pour modéliser les caractères de nature discrète. La voix qui s'est révélé le plus fructueux, est celui du modèle à seuils, vu sa dépendance d'une sous-jacente. L'objectif principal, de point de vue mathématique, est de proposer un modèle statistique explicite qui tient en compte des effets environnementaux, génétiques sur la moyenne, et qui permet d'introduire l'information génétique sur la variance. Dans cette partie, nous avons aussi exploré d'autres modèles statistiques pertinents qui prennent en compte l'hétérogénéité de la variance résiduelle.

Sommaire

1 Modélisation de la "moyenne" : Modèle à seuils homoscédastique 44

1.1 Modèle à seuils homoscédastique 44

1.2 Résultats et discussion 44

2 Modélisation de la variance résiduelle : Modèle à seuils hétéroscédastique 45

2.2 Etude de la variabilité due à un gène majeur 47

2.2.1 Introduction 47

2.2.2 Article 1 : Relationship between the effects of the BMP15 gene and the polygenic effects on prolificacy in the Rasa Aragonesa sheep breed 48

2.2.3 Résultats majeurs obtenus sur l'article précédent 68

2.3 Etude de la variabilité due à plusieurs gènes 70

2.3.1 Introduction 70

2.3.2 Article 2 : Genetic heteroscedastic models for ordinal traits: Application to sheep litter size 70

2.3.3 Résultats complémentaires 99

2.4Discussion 101

1 Modélisation de la "moyenne" : Modèle à seuils homoscédastique

1.1 Modèle à seuils homoscédastique

L'analyse statistique de la "moyenne" des tailles de portée considérée comme une variable ordonnée a été faite en utilisant un modèle à seuil homoscédastique tel que décrit dans la partie bibliographique. Cette analyse a permis d'avoir des références pour la partie suivante dans laquelle sont modélisées conjointement la moyenne et la variance du caractère. Les données analysées dans cette partie sont issues de l'échantillon complet déjà utilisé dans les analyses qui considéraient la taille de portée nomme une donnée continue. Ces analyses ont donc porté sur un total de 75 218 observations pour la race Suffolk, 186 864 pour la Rouge de l'Ouest et 137 358 pour la race Romane. Nous avons considéré un modèle à seuils homoscédastique mixte dans lequel les effets environnementaux sont les mêmes que ceux du modèle linéaire mixte homoscédastique (partie précédente), à savoir : la saison de mise bas, l'âge de la brebis, son âge à la première mise bas, son millésime, l'intervalle post-partum, le numéro de parité et les combinaisons : âge*âge à la première mise bas et numéro de mise bas*âge. Les effets aléatoires sont l'effet génétique additif, l'effet permanent de la brebis et l'effet troupeau.
Cette analyse a été faite avec le logiciel ASReml (Gilmour, Gogel, Cullis and Thompson, 2009).

1.2 Résultats et discussion

Les variances génétiques, les héritabilités et les répétabilités associées sont présentées dans le tableau 9. Pour les trois races, nous constatons que les héritabilités et les répétabilités estimées sont faibles. Les valeurs génétiques estimées en considérant la variable comme continue (modèle linéaire mixte homoscédastique) ou discrète (modèle à seuils homoscédastique) sont hautement corrélées (0.95, $P < 0.001$). Ce qui est conforme à de nombreuses comparaisons publiées sur ces deux types de modèles. Il n'est pas dans notre but d'exposer ou de comparer ces analyses, citons toutefois les résultats de Olesen (Olesen, 1995)qui a comparé trois types de modèles (modèle linéaire, modèle de Poisson et modèle à seuils) pour l'analyse génétique de la taille de portée chez les ovins. Il trouve que les trois modèles donnent le même ajustement aux données et le même classement des individus selon leur valeur génétique.

Tableau 9 : Estimations des composantes de la variance pour les différentes races par le modèle à seuils homoscédastique [TM].

	Races	σ_u^2	σ_p^2	σ_h^2	σ_e^2	h^2	r
Modèle à seuils homoscédastique	Suffolk	0.093	0.041	0.126	1.000	$0.073_{(0.006)}$	$0.105_{(0.006)}$
	Rouge de l'Ouest	0.110	0.022	0.123	1.000	$0.088_{(0.005)}$	$0.114_{(0.005)}$
	Romane	0.098	0.039	0.099	1.000	$0.079_{(0.004)}$	$0.110_{(0.004)}$

σ_u^2 : Variance des effets génétiques additifs, σ_p^2 : Variance des effets d'environnement permanent , σ_h^2 : Variance des effets troupeaux, σ_e^2 : Variance résiduelle, h^2 **:** Héritabilité, r : Répétabilité

2 Modélisation de la variance résiduelle : Modèle à seuils hétéroscédastique

L'existence d'un contrôle génétique de la sensibilité environnementale de la taille de portée des ovins a été abordée à deux niveaux différents : monogénique et polygénique. Nous savons en effet depuis Piper et Bindon (1982), que la taille de portée des ovins est sous le contrôle de gènes majeurs dans certaines races (tableau 10) dont les effets s'ajoutent à ceux des nombreux autres gènes ayant des effets infinitésimaux (effet polygénique des polygènes). Les gènes majeurs comme les polygènes peuvent avoir des effets sur la variabilité du caractère. Même si dans le cas où un seul gène est en cause, les analyses sont forcément plus simples, les deux situations sont proches et des analyses similaires peuvent être faites.

Nous avons recherché des effets génétiques sur la variabilité de la taille de portée, en nous intéressant d'une part à l'effet du gène BMP15 qui a un effet majeur sur la moyenne et d'autre part à l'effet de l'ensemble des polygènes.

Pour l'étude monogénique nous avons utilisé des données de la race Rasa Aragonesa en collaboration avec les chercheurs espagnols en charge de la gestion et de l'évaluation génétique de cette population. Le cas polygénique a été traité sur les données des trois races françaises présentées dans la partie précédente. Ces deux approches ont donné lieu à publications.

Tableau 10 : Races ovines dans la quelle ségrége un gène majeur (extraites de la thèse(Mansanet, 2013)).

Race	Genotype, Allèle muté (autre appellation)	Gène, Coordonnée de la mutation (acides aminés)	Effets sur le taux d'ovulation (TO) et la taille de portée (TP)
Akkaraman			
Aragonaise	Rasa Aragonesa, *FecX*^{R}	BMP15, W154NfsX55	R+ : TP +1,3 RR : stérile
Belclare	Belclare, *FecX*^{B} Galway, *FecX*^{G} High Fertility, *FecG*^{H}	BMP15, S367I BMP15, Q239X GDF9, S395F	B+ : TO +1,0 / BB : sterile G+ : TO +0,7 / GG : sterile H+: TO +1,4 / HH: stérile
Belle-Ile			TO 1-8 ; TP 1-7
Cambridge	Galway, FecX^{G} *(FecC)* High Fertility, *FecG*^{H} *(FecC)*	BMP15, Q239X GDF9, S395F	G+ : TO +0,7 / GG : sterile H+: TO +1,4 / HH: stérile
Chios (Sakiz)			TO : 2,37±1,01 TP : 1,93±0,76
Coopworth	Woodlands, *FecX2*^{W}		W+ : TO +0,4 ; TP +O,25 WW : TO & TP ≥ W+
Florina			TO : 1,79±0,63 TP : 1,34±0,52
Garole	Booroola, *FecB*^{B}	BMPR1B, Q249R	B+ : TO +1,5 ; TP +1,0 BB : TO +3,0 ; TP +1,5
Grivette	Grivette, *FecX*^{Gr}	BMP15, T317I	Gr+ : TO +0,5 ; TP +0,1 GrGr : TO +2 ; TP +0,7
Han	Booroola, *FecB*^{B} Galway, *FecX*^{G}	BMPR1B, Q249R BMP15, Q239X	B+ : TO +1,5 ; TP +1,1 BB : TO +3,0 ; TP +1,4 G+ : TP +0,55 GG : sterile
Icelandic	Thoka, *FecG*^{T} (*FecI*)	GDF9, S427R	I+ : TO +1,2 ; TP +0,7 II : stérile
Javanese	Booroola, *FecB*^{B} (*FecJ*)	BMPR1B, Q249R	B+ : TO +1,5 ; TP +1,0 BB : TO +3,0 ; TP +1,5
Karya			
Lacaune	Lacaune, *FecX*^{L} Lacaune, *FecL*	BMP15, C321Y	L+ : TO +2 / LL : sterile L+: TO +1,5 / LL: TO +3
Merinos	Booroola, *FecB*^{B}	BMPR1B, Q249R	B+ : TO +1,5 ; TP +1,0 BB : TO +3,0 ; TP +1,5
Olkuska	Olkuska, *FecX*^{O}	BMP15, N337H	O+ : TO +1 ; TP +0,6 OO : TO +1,6 ; TP +1,2
Romney	Inverdale, *FecX*^{I} Hanna, *FecX*^{H} Wishart, *FecW*	BMP15, V299D BMP15, Q291X	I+ : TO +1,0 ; TP +0,6 II : sterile H+ : TO +1,0 ; TP +0,6 HH : sterile W+: TO +0,9 / WW: stérile
Sangsari	*FecG* ?	GDF9 ?	
Santa Ines	Embrapa, *FecG*^{E}	GDF9, F345C	E+ : TP +0,3 / EE : TO +1 ; TP +0,7
White Norwegian	WNS, *FecG*^{WN}	GDF9, V371M	WN+ : TP +0,2 / WNWN : TP +0,5

2.2 Etude de la variabilité due à un gène majeur

2.2.1 Introduction

Le but de cette partie est d'évaluer l'impact du gène BMP15 sur la variabilité de la taille de portée de la race Rasa Aragonesa. Une mutation (allèle ROA) du gène BPM15 porté par le chromosome X a été découverte en 2007 dans cette race(Jurado, 2007; Martinez-Royo et al., 2008). Son effet sur la moyenne est suffisamment intéressant pour que les organismes de sélection aient décidé de l'exploiter. Comme la plupart (6/8) des autres mutations connues de ce gène (tableau 10), la mutation ROA (*FecX*^{R}) entraine une hyper ovulation à l'état hétérozygote et une stérilité à l'état homozygote. La transmission particulière de cette mutation fait que les filles d'un mâle homozygote (porteur de la mutation sur son chromosome X) qui mettent bas sont toutes hétérozygotes. La connaissance du génotype des femelles dans cette population est donc grandement facilité ce qui nous a permis d'analyser l'effet du gène sur la variabilité à partir d'effectifs importants d'animaux de génotype identifié. Nous disposions en effet de 918 956 enregistrements dont plus de 10 000 provenaient de femelles hétérozygotes. L'effet du gène BMP15 a été analysé à partir de plusieurs modèles à seuils.
Bien que déjà présenté dans la partie bibliographique et redéveloppé pour notre cas particulier dans la partie suivante, rappelons que ce type de modèle suppose l'existence d'une variable sous-jacente qui suit une distribution normale standard: $l \sim N(0,1)$ et un vecteur de seuils $\boldsymbol{t} = (\tau_1, \tau_2, \tau_3)$ qui déterminent les 4 catégories de tailles de portée (TP=1, 2, 3 et ≥4). La paramétrisation adoptée par le logiciel utilisé (ASReml) permet de considérer par exemple que le vecteur des seuils est affecté par des effets environnementaux β et génétiques u :

$$\tau_1 = X\beta + Zu$$
$$\tau_2 = \tau_1 + \Delta_{2-1}$$
$$\tau_3 = \tau_2 + \Delta_{3-2}$$

où X et Z sont les matrices d'incidence connues et Δ_{2-1} et Δ_{3-2} sont les écarts entre les seuils communs pour l'ensemble de la population.
Nous avons d'abord analysé l'effet du gène sur la moyenne de la taille de portée et sa contribution à la variabilité génétique de la population. Puis comme l'organisme de sélection qui nous a fourni les données (Upra-Grupo Pastores) trouvait que les mâles porteurs de la mutation avaient systématiquement de plus faibles valeurs génétiques polygéniques que les non porteurs, nous avons étudié l'interaction entre la mutation et les effets polygéniques sur la moyenne. Finalement, en tenant compte des résultats précédents, nous avons évalué l'effet du gène sur la variabilité de la taille de portée. Cet effet a été analysé en considérant un vecteur de seuils spécifique à chaque génotype, c'est à dire en utilisant une facilité du programme ASReml qui permet d'introduire une interaction entre le vecteur des seuils et le génotype soit :

Genotype ++	Genotype R+
$\tau_1^{++} = X\beta + Zu^{++}$	$\tau_1^{R+} = X\beta + Zu^{R+}$
$\tau_2^{++} = \tau_1^{++} + \Delta_{2-1}^{++}$	$\tau_2^{R+} = \tau_1^{R+} + \Delta_{2-1}^{R+}$
$\tau_3^{++} = \tau_2^{++} + \Delta_{3-2}^{++}$	$\tau_3^{R+} = \tau_2^{R+} + \Delta_{3-2}^{R+}$

Les facteurs environnementaux et génétiques de ce modèle sont : l'âge, l'intervalle entre agnelage, l'interaction : troupeau x année x mois de naissance (R-A-M), l'effet génétique additif et l'effet permanent.

Cettepartie de l'étude a fait l'objet d'une publication dans la revue ITEA. Cette publication est présentée ci-après dans sa version finale.

2.2.2 Article 1 : Relationship between the effects of the BMP15 gene and the polygenic effects on prolificacy in the Rasa Aragonesa sheep breed

RELACIONES ENTRE LOS EFECTOS DEL GEN BMP15 Y LOS EFECTOS POLIGÉNICOS SOBRE LA PROLIFICIDAD EN LA RAZA OVINA RASA ARAGONESA

S. Fathallah[1*], J.L. Alabart[2], L. Bodin[1], M.A. Jiménez-Hernando[4], B Lahoz[2], E. Fantova[3], I. David[1], J.J. Jurado[4]

[1]INRA, UMR1388 Génétique, Physiologie et Systèmes d'Elevage, F-31326, Castanet-Tolosan, France.

[2]CITA de Aragón. Av. de Montañana 930. 50059 Zaragoza. España.

[3]UPRA-Grupo Pastores. Zaragoza. España.

[4]Departamento de Mejora Genetica Animal. INIA. Ctra. La Coruña, km 7,5. 28040-Madrid. España.

[*]autor para correspondencia: Samira.Fathallah@toulouse.inra.fr

RESUMEN

Los efectos del gen BMP15, localizado en el cromosoma X, sobre la media y la variabilidad de la prolificidad fueron estimados en la población Rasa Aragonesa, donde se analizaron 918.956 partos de ovejas de distintos genotipos. En primer lugar, se utilizaron modelos umbral que incluían o no el efecto del gen para estimar el efecto del genotipo *BMP15* y su contribución a la variabilidad genética de la prolificidad en esta población. También se utilizaron otros dos modelos para estimar las interacciones entre el genotipo y los poligenes, así como el efecto del genotipo sobre la variabilidad de la prolificidad. Todos los análisis se hicieron con el programa ASReml. El efecto del genotipo sobre la media fue de 0,32 corderos adicionales por parto cuando el alelo ROA® se encontraba presente. Debido a la selección por prolificidad durante muchos años desconociendo la presencia de esta mutación, los animales selectos portadores tienen un valor genético poligénico residual más bajo que

los no portadores. No se encontró interacción significativa entre el genotipo y los efectos poligénicos sobre el valor genético. Aunque la interacción entre el genotipo y el conjunto de umbrales fue significativa, la diferencia entre ambos genotipos en la varianza de la prolificidads sobre la escala observada a una prolificidad media dada es escasa y muy poco relevante.

Palabras clave: BMP15, prolificidad, efecto poligénico

RELATIONSHIP BETWEEN THE EFFECTS OF THE BMP15 GENE AND THE POLYGENIC EFFECTS ON PROLIFICACY IN THE RASA ARAGONESA SHEEP BREED

SUMMARY

The effects of the BMP15 gene, localized on the X chromosome, on litter size mean and variability were estimated in the Rasa Aragonesa population in which 918,956 lambing records from ewes of different genotypes were analysed. Threshold models including or not the gene effect were first run to determine the effect of the genotype and its importance in the total genetic variability of litter size. Two other models were also run to estimate the interaction between the *BMP15* genotype and the polygenic background as well as the effect of the genotype on the litter size variability. All models were run using the ASReml software. The effect of the allele ROA® of the BMP15 gene on the mean prolificacy was 0.32 lambs per lambing. Due to the selection during many years on prolificacy ignoring the presence of the major gene, animals carrying the mutation were found to have lower remaining polygenic estimated breeding values than non-carrier animals, and there was no interaction between the BMP15 genotype and the polygenic background. Although the interaction between the genotype and the set of thresholds was significant, the resulting between-genotypes difference of variance of the litter size on the observed scale at a similar mean litter size was very low and not relevant.

Keywords: BMP15, prolificacy, polygenic effect, breeding value

INTRODUCCIÓN

Desde que Piper y Bindon (1982) descubrieron el primer gen mayor que controla la prolificidad de los ovinos se han publicado hasta ahora 14 mutaciones localizadas en 4 genes distintos (8 mutaciones en *BMP15*; 4 en *GDF9*; 1 en *BMPR1b*, 1 en *B4GALNT2*) con un gran efecto sobre la tasa de ovulación y por lo tanto la prolificidad de las ovejas (Fabre *et al*. 2015). A pesar de esa multiplicidad, el número de poblaciones ovinas que realmente manejan alguno de esos genes como herramienta de mejora es escaso. Las situaciones de la raza Lacaune en Francia y de la raza Rasa Aragonesa en España son sumamente notables. Desde el descubrimiento del gen $FecL^L$(Bodin *et al.*, 2002) y de la mutación $FecX^R$(Jurado y Calvo, 2007; Martinez-Royo *et al.*, 2008), los organismos de selección de ambas poblaciones decidieron aumentar el número de hembras heterocigotas en los rebaños: $FecX^R/FecX^+$ (o sencillamente R+) en el caso de la Rasa Aragonesa y $FecL^L/FecL^+$ (o L+) en el caso de la Lacaune. Pero debido a la hyper-prolificidad de las homocigotas $FecL^L/FecL^L$ (o LL) y a la esterilidad de las homocigotas $FecX^R/FecX^R$ (o RR) estos organismos decidieron evitar la generación de tales animales (Folch *et al.*, 2010; Martin *et al.*, 2014). Si bien el efecto de esos genes sobre la tasa de ovulación y la prolificidad es bien conocido (Bodin *et al.*, 2002; Lahoz *et al.*, 2009), es preciso determinar las relaciones entre estos efectos y los efectos del fondo poligénico para una gestión más precisa y especialmente para mejorar la orientación de la selección y lograr una evaluación genética más eficaz. Estas relaciones no habían sido estudiadas por falta de datos suficientes. Sin embargo, debido a la transmisión particular del alelo ROA® del gen BMP15 en la población de producción (localización en el cromosoma X y esterilidad de las hembras homocigotas), disponemos ahora en Rasa Aragonesa de un gran número de hembras con genotipo conocido, obtenido bien por genotipado o por deducción a partir del genotipo de sus parientes. El objetivo de este trabajo es evaluar las relaciones entre el efecto del alelo ROA® del gen BMP15 y los efectos poligénicos sobre la prolificidad de la Rasa Aragonesa. La metodología de este estudio podrá ser utilizada posteriormente para la raza Lacaune cuando se disponga de una cantidad de datos suficiente así como para cualquier población donde un gen mayor que controla un carácter discreto está en segregación.

MATERIAL Y MÉTODOS

Los datos proceden de la base de datos utilizada para la evaluación genética de la población Rasa Aragonesa de UPRA-Grupo Pastores (Jurado, 2014) e incluyen: 1) los partos procedentes de monta natural de 269.503 ovejas, que se componen de 918.956 registros de parto con sus características recogidos durante 15 años en 309 rebaños; 2) el pedigrí tras 7 generaciones de dichas ovejas; 3) el genotipo de dichas ovejas y el de sus padres. Los efectos considerados en los análisis de esa muestra global fueron: la combinación rebaño-año-mes del parto (*RAM*: 15.704 niveles), la edad al parto medida en años (*Edad*: 9 niveles), el tiempo transcurrido entre el parto anterior y la fecha de fecundación (intervalo post-parto; *PostP*: 5 niveles : <90 días; 90-160 días; 160-1000 días; >1000; primer parto), el genotipo (*Genot*: 3 niveles: no portadora [++], n=1.023; heterocigota [R+], n=10.013; o desconocida [desc.] n=258.467), el efecto permanente ligado a cada oveja productiva (*Perm*: 269.503 niveles), y finalmente, el efecto genético aditivo que corresponde al número de animales en el pedigrí (*u*: 305.703 niveles). En este estudio, los partos de más de 3 corderos fueron agrupados en la categoría "4+". La distribución por genotipos en función del tipo de parto observado en esta muestra global aparece en la tabla 1.

Tabla 1. *Distribución (números y %) del tipo de parto observado para cada genotipo en la muestra global analizada.*

Table 1. Distribution (numbers and %) of the observed litter size for each genotype in the global sample.

Genotipo		TP=1	TP=2	TP=3	TP=4+	Partos totales	Prolificidad observada
++	n	1.453	588	22	3	2.066	1,31
1.023 ♀	%	70,3	28,5	1,1	0,15		
R+	n	12.501	14.200	2.526	210	29.437	1,68
10.013 ♀	%	42,5	48,2	8,6	0,71		
desc.	n	608.755	268.218	10.035	445	887.453	1,33
258.467 ♀	%	68,6	30,2	1,1	0,05		

El tipo de parto (TP) es una variable discreta que se analizó mediante el software ASReml (Gilmour *et al*. 2009), utilizando varios modelos umbral (Gianola, 1982). La parametrización utilizada para todos los modelos supone que la variable subyacente sigue una distribución normal estandarizada $l \sim \mathrm{N}(0,1)$ y que el vector de umbrales $\mathbf{t} = (\tau_1, \tau_2, \tau_3)$, que determina las cuatro categorías de parto, está afectado por efectos ambientales y genéticos. Todos los modelos incluyeron tres efectos aleatorios: la interacción rebaño-año-mes del parto ($RAM \sim \mathrm{N}(0, \sigma^2_{RAM})$), un efecto genético aditivo animal ($u \sim \mathrm{N}(0, \sigma^2_u \otimes A)$) asociado a la matriz de parentesco A entre los individuos, y un efecto ambiental permanente ligado a cada oveja ($Perm \sim \mathrm{N}(0, \sigma^2_{Perm})$).

El primer modelo (M1) consideró únicamente como efectos fijos la edad de la oveja al parto (*Edad*) y el intervalo post-parto (*PostP*):

$$t = \mu + RAM + Edad + PostP + u + Perm$$

El segundo modelo (M2) también incluyó el efecto fijo del genotipo (*Genot*), resultando el modelo:

$$\mathbf{t} = \mu + RAM + Edad + PostP + Genot + u + Perm$$

A continuación, se probó un tercer modelo (M3) que incluyó la interacción entre el efecto genético aditivo y el factor genotipo, para tener en cuenta la interacción entre los efectos poligénicos y el efecto del genotipo sobre la media del carácter: $\mathbf{t} = \mu + RAM + Edad + PostP + Genot + Genot.u + Perm$

Donde: $Genot.u$ es una variable normal $Genot.u \sim N(0, G \otimes A)$, siendo $G = \begin{pmatrix} \sigma^2_{u^{++}} & \sigma_{u^{++}u^{R+}} \\ \sigma_{u^{++}u^{R+}} & \sigma^2_{u^{R+}} \end{pmatrix}$ la matriz de varianza-covarianza entre los valores genéticos u^{++} y u^{R+} de cada individuo y A la matriz de parentesco entre los individuos. Debido a problemas de estructura de datos, no se pudieron estimar los parámetros de este modelo sobre la muestra global; se utilizó entonces para este modelo una submuestra que consideraba únicamente los animales de genotipo conocido con más de tres partos naturales (22.946 partos de 5.035 hembras de genotipo ++ o R+). Finalmente, para estimar un posible efecto del genotipo sobre la variabilidad del carácter, se probó sobre la muestra global un cuarto modelo (M4) con los mismos efectos que el modelo M1 y la interacción entre el factor genotipo y los

umbrales. Para ello se utilizó una posibilidad del programa ASReml que consistió en considerar un vector de umbrales para cada genotipo, en vez de considerar un solo vector de umbrales común para ambos genotipos. Al igual que en el modelo sin interacciones, el programa supone que estos vectores están sometidos a efectos ambientales y genéticos:

$$\left\{\mathbf{t}^{++},\mathbf{t}^{R+},\mathbf{t}^{desc}\right\} = \mu + RAM + Edad + PostP + u + Perm \ .$$

Las comparaciones entre los modelos se realizó a partir del criterio de información de Akaike (AIC). Las estimaciones de la prolificidad media y de su varianza sobre la escala observada se calcularon a partir de las estimaciones de cada tipo de parto proporcionadas con el modelo de umbrales: la probabilidad de repuesta en la *j-ésima* categoría del nivel *i* de un efecto(por ejemplo *i* =++, R+ o desc. para el efecto genotipo) fue estimada mediante $\hat{\Pi}^{i}_{j} = \Phi(\hat{\tau}^{i}_{j}) - \Phi(\hat{\tau}^{i}_{j-1})$ donde $\Phi(x)$ es la función de distribución de la distribución normal y el vector $\left\{\hat{\mathbf{T}}^{i}\right\}$ proviene de las soluciones para el nivel *i* del efecto en el modelo considerado. La media y la varianza sobre la escala observada se calcularon finalmente a partir de las fórmulas siguientes:

$$\mu^{i}_{obs} = \sum_{j=1}^{4} j\hat{\Pi}^{i}_{j} \quad y \quad \sigma^{2\,i}_{obs} = \sum_{j=1}^{4} j^{2}\hat{\Pi}^{i}_{j} - \hat{\mu}^{i\ 2}_{obs}$$

Aunque los valores genéticos se estimaron para todos los individuos del pedigrí, se presentan únicamente los de las ovejas que han tenido por lo menos un parto.

RESULTADOS Y DISCUSIÓN

Justificación del uso de un modelo umbral

Debido a la naturaleza discreta de la variable "tipo de parto", que además tiene pocas categorías (TP4+ <1%), existe una relación muy fuerte entre su media y su varianza. Por lo tanto, la metodología clásica de análisis de varianzas no está bien adaptada para estudiar este tipo de carácter. El modelo umbral, desarrollado por Gianola (1982), permite superar una parte de los problemas pero, al ser un

modelo no lineal, crea otras dificultades. Sin embargo, numerosos estudios demuestran que bajo ciertas condiciones (calidad del protocolo experimental, porcentaje de cada categoría, etc.) hay una escasa diferencia entre los resultados obtenidos con estas dos metodologías (David *et al.*, 2007). Además, Jurado y Espinosa (1996), mostraron precisamente en la Rasa Aragonesa que la evaluación genética de la prolificidad podía hacerse utilizando un modelo clásico lineal en vez del modelo umbral. En el presente estudio, se ha utilizado el modelo umbral porque, tal y como lo han mostrado Martin *et al.* (2014) y Bodin *et al.* (2015), la posibilidad de introducir una interacción entre los umbrales y cualquier otro factor del modelo (ambiental o genético), permite modelizar el efecto de ese factor sobre la variabilidad del carácter.

Modelos que ignoran la presencia del gen o suponen un simple efecto sobre la media

Los parámetros genéticos de la variable subyacente y la media de los valores genéticos estimados en los modelos que ignoran la presencia del genotipo ROA (M1) o que consideran que este tiene solamente un efecto sobre la media (M2), se presentan en la tabla 2. Como era de esperar, la varianza genética estimada es menor cuando el efecto del gen mayor ROA está incluido en los modelos. En estos casos, la varianza genética es puramente poligénica (debido a los genes de pequeño efecto), lo que permite estimar que el genotipo ROA representa alrededor del 20% de la variabilidad genética en esta población.

***Tabla 2**. Varianzas, parámetros genéticos de la prolificidad y medias de los valores genéticos estimados sobre la escala subyacente (errores estándar entre paréntesis).*
Table 2. Variance components of the models and genetic parameters estimated on the underlying scale (standard errors between parentheses).

Parámetro	Modelo M1 (sin el factor genotipo)	Modelo M2 (con el factor genotipo)
σ^2_{animal}	0,1202 *(0,0038)*	0,0967 *(0,0039)*
$\sigma^2_{permanente}$	0,0228 *(0,0037)*	0,0394 *(0,0039)*
Heredabilidad : h^2	0,105 *(0,001)*	0,085 *(0,003)*
Repetibilidad: r	0,125 *(0,003)*	0,120 *(0,001)*
Efecto del gen (escala observada)	/	0,32
- Valor genético[(1)]: media de ++	-0,1719 *(0,2273)*	-0,0147 *(0,1195)*
- Valor genético: media de R+	-0,5614 *(0,3419)*	-0,0049 *(0,1819)*
- Valor genético: media de desconocido	-0,0113 *(0,1520)*	-0,0104 *(0,1263)*

[(1)] *Valores genéticos calculados sobre la escala de la variable subyacente. Los valores más altos corresponden a la prolificidad más baja: cuanto más altos son los umbrales, mayor es la proporción de la variable subyacente que está por debajo (y viceversa).*

[(1)] *Breeding values are calculated on the underlying variable scale. The highest values correspond to a lower prolificacy: higher are the thresholds, higher proportion of the underlying variable is below (and vice-versa).*

El efecto medio del alelo *FecX*^{R}, estimado como la diferencia entre las medias del tipo de parto sobre la escala observada para los genotipos ++ y R+, es de 0,32 corderos por oveja y parto (1,62-1,29), lo que se corresponde con las estimaciones previamente obtenidas por Jurado *et al.* (2008) y Lahoz *et al.* (2011). Este efecto es de la misma magnitud que la de otros genes mayores, aunque ligeramente inferior al efecto de los genes Booroola o Lacaune (Martin *et al.*, 2014). Los valores genéticos estimados con el modelo M1, que ignora la presencia del gen mayor y que por lo tanto incluye de manera oculta el efecto de ese gen, son obviamente mejores para los individuos R+ que para los de genotipo desconocido y los ++ (tabla 2). Nótese que los valores genéticos están expresados en términos de la variable subyacente, de forma que los valores más altos corresponden a la prolificidad más baja: cuanto más altos son los umbrales, mayor es la proporción de la variable subyacente que está por debajo (y viceversa).

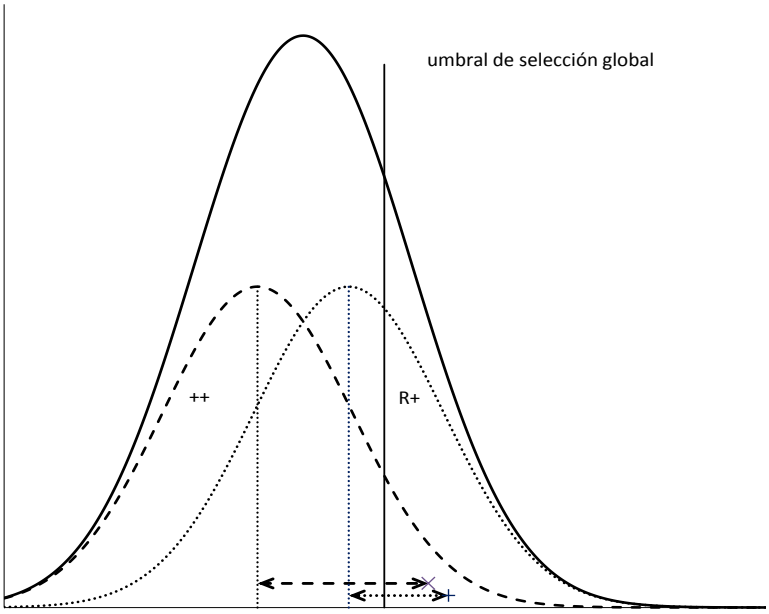


Figura 1. *Distribución teórica de los valores genéticos globales (línea continua) y media de cada genotipo (++: X y R+: +) sobre la escala observada, asi como la distribución de los valores poligénicos de cada genotipo con su media relativa (++:<---->y R+: <·····>).*

Figure 1. Theoretical distributions of the global breeding values (solid line) and mean for each genotype (++: X y R+: +) on the observed scale, as well as polygenic distribution for each genotype with its relative mean (++:<---->y R+: <·····>).

Los valores genéticos proporcionados por el modelo M2 representan los valores exclusivamente poligénicos y su media intra-genotipo es próxima a cero. Sin embargo nótese que estos valores genéticos poligénicos son significativamente mejores (P<0,001) para los ++ (μ=-0,0147) que para los R+ (μ=-0,0049). Este fenómeno es debido sencillamente a la selección de los individuos más prolíficos durante más de 35 años (Perez-Almero y Valls Ortiz, 1977) sobre un valor genético global que era la suma del valor poligénico y del valor del alelo del gen mayor. Para un valor global dado, los individuos ++ tienen necesariamente un mejor valor poligénico (figura 1). La tabla 3 presenta el porcentaje de las hembras reproductoras ++ y R+ de la muestra global que hay en distintos cuantiles de la distribución de los valores genéticos estimados con el modelo M1. Esta tabla indica también el valor poligénico (modelo M2) de estos animales y la diferencia de valores poligénicos entre los ++ y

R+ de cada grupo. Podemos averiguar que, dentro de las 25% mejores ovejas clasificadas sobre su valor genético global, están incluidas solo un 5% de las mejores ++, frente a un 44% de las R+. Como consecuencia, las ++ tienen un valor poligénico más alto que las R+. Este fenómeno clásico en un caso de selección de un carácter en presencia de un gen mayor, explica los valores poligénicos más bajos de los machos R+ comparando a los ++ en el catálogo de sementales de la cooperativa UPRA-Grupo Pastores (Jurado, 2014). Esta mera asociación entre genes, muy distinta de una interacción entre genes, seguramente empezó a evolucionar desde el año 2009 cuando la selección sobre el valor poligénico remplazó la selección sobre el valor genético global.

Tabla 3. *Porcentajes de las hembras ++ y R+ de la poblacion en diversas muestras de individuos elegidos sobre valores genéticos globales (modelo M1); valores genéticos poligénicos medios de estos individuos (modelo M2) y diferencia entre genotipos.*
Table 3. Percentages of the ++ and R+ ewes form the population in samples chosen on the global genetic value (Model M1); polygenic breeding values of these females (Model M2) and difference between genotypes.

Porcentajes de individuos (incluyendo los desconocidos)	Porcentajes de ++ y R+ incluidos en la muestra		Valor genético poligénico[(1)]		Diferencia de valor genético poligénico (R+) – (++)
	++	R+	++	R+	
10%	1%	18%	-0.437	-0.193	0,244
25%	5%	44%	-0.286	-0.092	0,194
50%	32%	75%	-0.125	-0.022	0.103
75%	69%	91%	-0.052	0.003	0,055
100%	100%	100%	-0.006	0.027	0,033

(1) *Valores genéticos poligénicos (modelo M2) calculados sobre la escala de la variable subyacente. Los valores más altos corresponden a la prolificidad más baja; diferencias positivas entre R+ y ++ indican que los ++ tienen mejores valores poligénicos (ver figura 1).*

La heredabilidad de los efectos poligénicos sobre la escala subyacente (tabla 2) es baja y consistente con los valores encontrados en la literatura, pero ligeramente superior a las estimaciones obtenidas por Jurado *et al.* (2008) en esta raza. Sin embargo, en el citado trabajo analizaron los partos como un carácter continuo y Gianola (1982) mostró que la heredabilidad sobre la escala subyacente es siempre superior a la heredabilidad de los datos observados. Por otra parte, se consideraron los partos procedentes de celo natural y de celo inducido por tratamiento hormonal como el mismo carácter

mientras que hoy sabemos que son dos caracteres distintos con una correlación genética de 0,75 (Baelden *et al.*, 2005).

Interacción entre el efecto del genotipo ROA y los efectos poligénicos sobre la media

Desafortunadamente, la estructura de la muestra global no permitió estimar la interacción entre el genotipo del animal y su efecto poligénico, cualesquiera que fueran los otros factores del modelo. La transmisión particular del gen y su modo de gestión en la población introducen confusiones entre el genotipo del animal, su valor poligénico y su valor de efecto permanente. De hecho, la mayor parte de la información que permite estimar los efectos poligénicos (separados de los efectos permanentes) viene de las relaciones entre medias hermanas de padre (sobre todo de los padres usados en inseminación con un gran número de hijas). Como intrafamiliarmente tienen casi todas el mismo genotipo (excepto algunas hijas de los machos ++), no se puede medir la interacción. El hecho de que los machos portadores no transmiten la mutación a sus hijos machos no permite evitar la confusión, puesto que estos hijos tienden a tener un valor poligénico más bajo que la media (ver párrafo anterior), y por lo tanto son muy poco utilizados y tienen poca descendencia. Sin embargo, considerar una submuestra limitada, construida únicamente con los partos de ovejas de genotipo conocido que tienen por lo menos 3 partos naturales, permite obtener parámetros satisfactorios con el modelo que incluye la interacción alelo ROA-poligenes (Tabla 4).

Tabla 4. *Varianzas y parámetros genéticos de la prolificidad sobre la escala subyacente en un modelo con interacción entre los efectos poligénicos y el genotipo ROA*; *(submuestra limitada a partos de individuos de genotipo conocido y con más de 3 partos).*
Table 4. Variance components and genetic parameters of litter size on the underlying scale in a model including the interaction between the polygenic effects and the ROA genotype (in a subsample limited to ewes of known genotype with at least 3 lambings recorded).

Parámetro	++	R+
$\sigma^2_{permanente}$	0,0306 *(0,028)*	
σ^2_{animal}	0,163 *(0,062)*	0,167 *(0,035)*
Heredabilidad : h^2	0,137 *(0,046)*	0,136 *(0,050)*
Repetibilidad : r	0,163 *(0,039)*	0,165 *(0,010)*
Correlación genética	0,983 *(0,293)*	

En esta submuestra, el efecto del gen sobre la escala observada es de 0,34 corderos por parto, lo que es muy próximo al efecto encontrado sobre la muestra global y confirma de una cierta forma la validez de ese muestreo. Según este modelo, es posible estimar efectos poligénicos para cada genotipo o en los dos ambientes genotípicos (++ y R+). Las varianzas genéticas y las heredabilidades (tabla 4) son ligeramente más altas que en la muestra global con el modelo sin interacción (tabla 2). Esto es debido al escaso porcentaje de individuos de genotipo desconocido de quienes se conoce por lo menos la madre (45%) en la muestra global, mientras que estos porcentajes son del 63% y 91% para los ++ y R+ en la submuestra. Sin embargo, estos parámetros son casi iguales en los dos ambientes genotípicos y la correlación genética es muy alta y no difiere de 1. Por lo tanto, se puede concluir que no hay interacción relevante entre el genotipo ROA y los otros genes. Este resultado deberá confirmarse más adelante cuando la gestión del gen y el uso de los valores genéticos poligénicos haya aumentado los vínculos entre individuos de genotipo diferente.

Efecto del genotipo ROA sobre la media y la variabilidad de la prolificidad

La introducción en el modelo de una interacción entre el genotipo del animal y los umbrales de la variable subyacente permite ajustar a la vez un efecto del genotipo sobre la media de los tipos de parto y sobre su variabilidad. El modelo umbral clásico supone la existencia de una variable subyacente normal estándar $N(\mu=0, \sigma^2=1)$ y unos umbrales que varían conjuntamente según los factores de variación. Estas variaciones (traslación del vector de umbrales) cambian las proporciones de las categorías delimitadas por los umbrales; y por lo tanto cambian, de manera totalmente ligada, la media y la varianza sobre la escala observada. La mayor parte de los cambios de varianza de los tipos de partos entre razas de ovino es una consecuencia directa de la variación de la prolificidad media (Bodin y Elsen, 1989). En nuestro estudio, la interacción entre el genotipo y los umbrales (figura 2 y tabla 4) es significativa ($P<0,001$) y produce dos efectos: un cambio de la media sobre la escala observada y un cambio adicional de la variabilidad. El efecto sobre la media es parecido al encontrado en el modelo sin interacción. La existencia de umbrales específicos a cada genotipo modifica más la distribución que si fuera solamente una traslación del conjunto de umbrales y conlleva un incremento de variabilidad adicional al meramente ligado al aumento de la media. Puede apreciarse (tabla 5) que para

unas medias de la variable subyacente que proporcionan la misma media en ++ y R+, la varianza sobre la escala observada es siempre superior para los animales R+. Sin embargo, esta diferencia debida a un solo gen es muy escasa. Por otra parte, la distancia entre los dos primeros umbrales es superior para las ++ (Δ^{++}_{2-1}=2,04) que para las R+ (Δ^{R+}_{2-1}=1,78) mientras que la distancia entre el segundo y tercero umbral es inferior para los ++ (Δ^{++}_{3-2}=0,88) que para los R+ (Δ^{R+}_{3-2}=1,34).

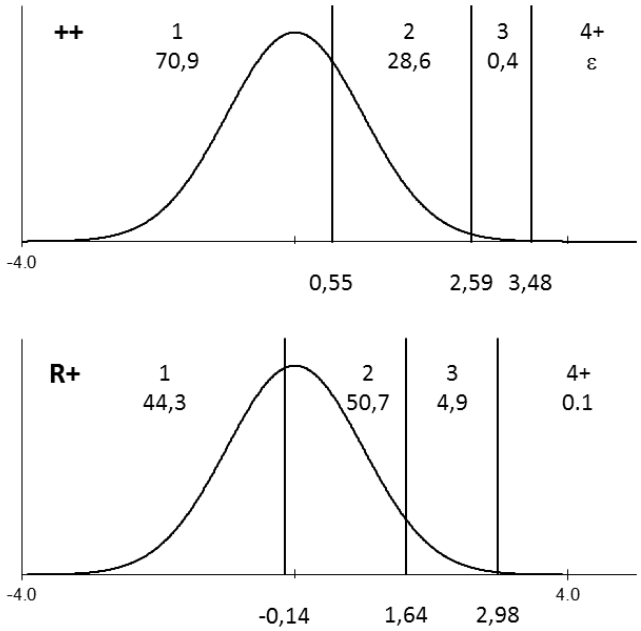


Figura 2. *Distribución de la variable subyacente Normal (0, 1) y valores de los umbrales de cada genotipo que delimitan las distintas categorías del tipo de parto.*
Figure 2. Distribution of the underlying variable N(0,1) and threshold values delimiting the litter size values for each genotype.

En consecuencia, no se puede establecer una transformación lineal entre estos dos conjuntos de umbrales, lo que no permite pasar de un conjunto al otro por un simple cambio de escala (scaling factor), lo que podría interpretarse como un cambio de varianza de la variable subyacente (Martin *et al.*, 2014). Aunque la interacción entre el genotipo y los umbrales es altamente significativa y los conjuntos de umbrales de las ++ y R+ son significativamente diferentes, como se ha visto en Bodin *et al.* (2015), el efecto adicional del gen sobre la varianza del tamaño de camada es muy escaso y parece carecer de relevancia. Por otra parte, el criterio AIC, que mide la calidad relativa de los modelos

estadísticos, es mejor para el modelo simple M2 (AIC=123.324) que para el modelo con la interacción entre el efecto del gen y los umbrales (AIC=123.400) lo que nos hace rechazar definitivamente el modelo más complejo.

Tabla 5. Umbrales estimados para cada genotipo y distribución de los tipos de partos según distintos parámetros de la variable subyacente.
Table 5. Estimates of the threshold values for each genotype and litter size distribution for different values of the parameters of the underlying variable.

Umbrales	Genotipo ++				Genotipo R+			
T_1	0,5517 *(0,0347)*				-0,1445 *(0,0164)*			
T_2	2,5955 *(0,0936)*				1,6358 *(0,0184)*			
T_3	3,4764 *(0,2366)*				2,9777 *(0,3340)*			
Parámetros de la variable subyacente								
μ	**0**	0,271	0,502	0,741	-0,710	-0,447	-0,225	**0**
σ^2	**1**	1	1	1	1	1	1	**1**
Distribución de los tipos de partos sobre la escala observada								
μ_{obs}	**1,296**	1,400	1,500	1,610	1,296	1,400	1,500	**1,610**
σ^2_{obs}	**0,219**	0,263	0,292	0,314	0,228	0,278	0,315	**0,346**
%TP1	**70,9**	61,1	52,0	42,5	71,4	61,9	53,2	**44,3**
%TP2	**28,6**	37,9	46,2	54,3	27,6	36,2	43,7	**50,7**
%TP3	**0,4**	0,9	1,7	2,9	0,9	1,8	3,1	**4,9**
%TP4+	**0,03**	0,07	0,15	0,31	0,01	0,03	0,07	**0,15**

Conclusión

La presencia del alelo ROA en la población Rasa Aragonesa aumenta la prolificidad de una manera muy favorable para los ganaderos puesto que no proporciona porcentajes muy elevados de corderos nacidos triples o de orden superior. La separación entre el efecto del gen y los efectos puramente poligénicos muestra la inferioridad poligénica de los animales portadores de la mutación, pero dicha inferioridad no es debida a una interacción o efectos epistáticos entre el alelo ROA y los poligenes. Esta asociación del alelo ROA a valores genéticos poligénicos más bajos se puede separar fácilmente mediante selección. A parte de esta ausencia de interacción, la ausencia de efecto relevante del gen sobre la variabilidad del tipo de parto es también particularmente interesante. Todos estos argumentos refuerzan aún más el interés de este gen para la gestión actual y futura de la prolificidad. Sin embargo se plantea ahora la cuestión de determinar cómo utilizar las informaciones genotípicas y poligénicas en el proceso de selección. Una cierta presión de selección tiene que estar orientada hacia las cualidades maternales de las ovejas, puesto que tienen que asegurar el crecimiento de los corderos suplementarios, pero no se puede eliminar sin embargo la prolificidad del objetivo de selección. De hecho, dejar de seleccionar poligénicamente la prolificidad o aplicar la misma presión de selección para los ++ que la presión actual sobre los R+ llevaría en un descenso en la prolificidad. Existen varias soluciones (Manfredi *et al*., 1998) que permiten no fijar la mutación sino mantenerla a una cierta frecuencia y gestionar los apareamientos conociendo los genotipos y los valores poligénicos. Si bien la solución óptima, descrita por Manfredi *et al*. (1998) es complicada de aplicar, la selección única utilizando la suma de los valores poligénicos y genotípicos es relativamente fácil. Una solución intermedia sería seleccionar los ++ y los R+ sobre su valor poligénico, pero utilizando reglas de gestión de los animales específicas para cada genotipo; por ejemplo, utilizando simplemente presiones de selección distintas.

AGRADECIMIENTOS

Ese trabajo se ha realizado en el marco de la tesis de Samira Fathallah financiada por el INRA (Francia) y el CORAM (Collectif des Races des Massifs - Francia). Los datos necesarios para la elaboración de este trabajo se obtuvieron gracias a la cofinanciación de proyectos INIA y CDTI.

REFERENCIAS BIBLIOGRÁFICAS

Baelden M, Tiphine L, Poivey JP, Bouix J, Bibe B, Robert-Granie C, Bodin L (2005). Estimation of genetic parameters for litter size after natural and hormone-induced oestrus in sheep. Livestock Production Science, 97: 275-281.

Bodin L, Alabart Alvarez JL, Fathallah S, Jimenez-Heranando M, Lahoz B, David I, Fantova E, Jurado JJ (2015). Efecto del gen BMP15 sobre la distribucion del tipo de parto en ovejas rasa aragonesa.En: XVI Jornadas sobre Producción Animal, Álvarez-Rodríguez J. *et al.*(Eds.), AIDA, Zaragoza, 19-20 de Mayo.

Bodin L, Elsen JM (1989). Variability of litter size of french sheep breeds following natural or induced ovulation. Animal Production, 48: 535-541.

Bodin L, SanCristobal M, Lecerf F, Mulsant P, Bibe B, Lajous D, Belloc JP, Eychenne F, Amigues Y, Elsen JM (2002). Segregation of a major gene influencing ovulation in progeny of Lacaune meat sheep. Genetics Selection Evolution, 34: 447-464.

David I, Bodin L, Lagriffoul G, Leymarie C, Manfredi E, Robert-Granie C (2007). Genetic analysis of male and female fertility after artificial insemination in sheep: Comparison of single-trait and joint models. Journal of Dairy Science, 90: 3917-3923.

Fabre S., Pierre A., Pisselet C., Mulsant P., Lecerf F., Pohl J., Monget P., Monniaux D., (2003). The Booroola mutation in sheep is associated with an alteration of the bone morphogenetic protein receptor-IB functionality. Journal of Endocrinology 177, 435-444.

Folch J, Jurado JJ, Calvo JH, Martínez Royo A, Alabart JL, Lahoz B, Fantova E (2010). Cómo mejorar la prolificidad de las ganaderías de ovino de carne mediante el programa de divulgación de la variante génica rasa oviaragón (ROA). Pequeños Rumiantes, 11: 17-20.

Gianola D (1982). Theory and analysis of threshold characters. Journal of Animal Science, 54: 1079-1096.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009). ASReml user guide release 3.0. VSN: Hemel Hempstead: VSN International Ltd.

Jurado JJ (2014). 21° Catálogo de reproductores de UPRA-Grupo Pastores.

Jurado JJ, Calvo JH (2007). ¿Un gen de gran efecto para prolificidad en raza rasa aragonesa? En: XII Jornadas sobre Produccion Animal, Zaragoza, Spain, 16-17 Mayo, 2007. Tomo I and II. pp. 504-506

Jurado JJ, Espinosa MJ (1996). Problemática del desarrollo de un programa de mejora genética en prolificidad en la raza rasa aragonesa. ITEA, 92A: 44-56.

Jurado JJ, Martinez-Royo A, Calvo JH (2008). Efecto fenotípico del alelo *BMP15*/*FecX*^{R} en la prolificidad de la poblacion de Carnes Oviaragon S.C.L. ITEA, 104: 149-154.

Lahoz B, Alabart JL, Echegoyen E, Sanchez P, Calvo JH, Martinez-Royo A, Jurado JJ, Fantova E, Folch J (2009). Efecto del alelo *FecX*^{R} del gen BMP15 sobre la tasa de ovulacion y la prolificidad en ovejas Rasa Aragonesa. En: XIII Jornadas sobre Producción Animal AIDA. M. Joy et al., (Eds.). AIDA, Zaragoza, pp. 678-680

Lahoz B, Alabart JL, Jurado JJ, Calvo JH, Martínez-Royo A, Fantova E, Folch J (2011). Effect of the *FecX*^{R} polymorphism in the bone morphogenetic protein 15 gene on natural or equine chorionic gonadotropin-induced ovulation rate and litter size in Rasa Aragonesa ewes and implications for on-farm application. J. Anim. Sci. 89(11): 3522-3530.

Manfredi E, Barbieri M, Fournet F, Elsen JM (1998). A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. Genetics Selection Evolution, 30: 127-148.

Martin P, Raoul J, Bodin L (2014). Effects of the *FecL* major gene in the Lacaune meat sheep population. Genet Sel Evol, 46: 48.

Martinez-Royo A, Jurado JJ, Smulders JP, Marti JI, Alabart JL, Roche A, Fantova E, Bodin L, Mulsant P, Serrano M, Folch J, Calvo JH (2008). A deletion in the bone morphogenetic

protein 15 gene causes sterility and increased prolificacy in Rasa Aragonesa sheep. Animal Genetics, 39: 294-297.

Perez-Almero J, Valls Ortiz M (1977). El control de producciones del Servicio de Mejora Ovina de la Excma. Diputacion Provincial de Zaragoza. Objetivos, metodología y primeros resultados. Diputacion Provincial de Zaragoza, 146 pp.

Piper LR, Bindon BM (1982). Genetic segregation for fecundity in booroola merino sheep. En Proceedings of the First World Congress On Sheep And Beef Cattle Breeding. November 1980. Palmerston North. pp. 395-400

2.2.3 Résultats majeurs obtenus sur l'article précédent

Conformément à ce qui a déjà été publié (Jurado, 2008; Lahoz et al., 2011), l'effet de l'allèle R sur la moyenne de la taille de portée (estimé par la différence sur l'échelle observée entre la prolificité des ++ et celle des R+) est de 0.32 agneau supplémentaire par mise bas. L'héritabilité polygénique est de h^2=0.08 et l'effet du gène contribue à environ 20 % de la variance génétique additive totale. On vérifie dans l'échantillon traité que les valeurs génétiques polygéniques des animaux porteurs sont systématiquement plus faibles que celles des non porteurs. Par contre nous n'avons pas mis en évidence d'interaction entre le génotype BMP15 et les polygènes (figure 12), de telle sorte que la relation trouvée jusqu'ici est une simple association due à la sélection sur une valeur génétique globale pendant plusieurs années en ignorant l'existence du gène.
Comme nous l'avons vu dans la partie bibliographique, il existe une liaison très forte entre la moyenne et la variance de la prolificité des ovins (figure 3 ; page 17) qui fait que sur l'échelle observée, une augmentation de la moyenne s'accompagne structurellement d'une augmentation de la variance. En fait on sait qu'avec le modèle à seuil, la modification d'un seul paramètre sur l'échelle sous-jacente (la moyenne de la variable sous-jacente) induit une modification des deux principaux paramètres de la distribution sur l'échelle observée :la moyenne et la variance. Nous avons repris la méthode utilisée par Martin *et al*. (2014) qui analyse l'effet qu'un gène majeur pour la moyenne peut avoir sur l'augmentation de variance de façon additionnelle à l'augmentation structurelle. Cette méthode consiste à mesurer l'évolution de la moyenne et de la variance sur l'échelle observée lorsque l'on fait varier la moyenne de la sous-jacente en gardant fixes les seuils estimés pour chacun des génotypes.

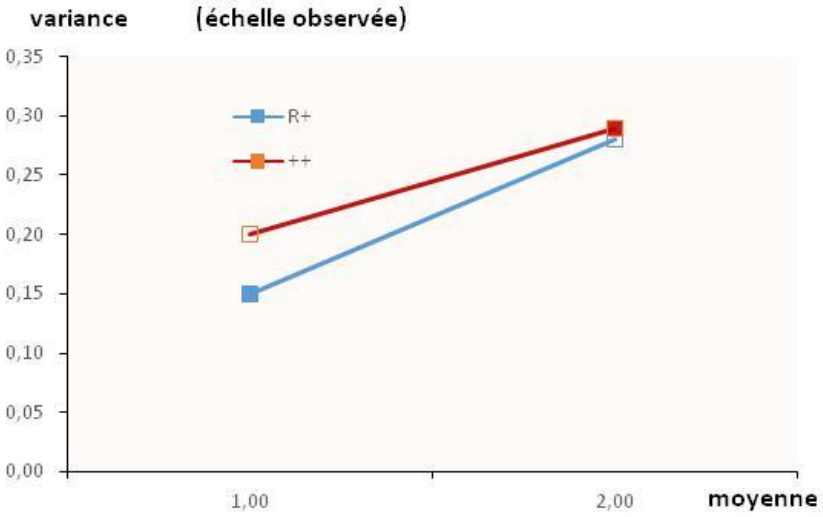


Figure 12 : Liaison moyenne – variance sur l'échelle observée pour différentes valeurs de la moyenne η de la variable sous-jacente et l'ensemble des seuils estimés pour chaque génotype R+ et ++. Quelle que soit la valeur de η (parmi les valeurs plausibles), la variance sur l'échelle observée des ++ est toujours supérieure à celle des R+ et notamment pour la même moyenne sur l'échelle observée.
Le modèle qui considère un ensemble de seuil spécifique à chaque génotype en ajustant une interaction seuils-x-génotype montre d'une part que cette interaction est significative, d'autre part que ce modèle est meilleur (LRT=76, ddl=3 et P-valeur << 0.001) que le modèle avec un seul ensemble de seuils. Cependant la traduction des pourcentages de chaque taille de portée montre que si la différence de moyenne est importante, l'augmentation de variance additionnelle est très faible et non pertinente (tableau 5 de la publication). Ce résultat est différent de ce qui a été trouvé par Martin *et al*.

(2014)(figure 13) qui montre que l'effet du gène Lacaune augmente de façon importante la variabilité des tailles de portée.

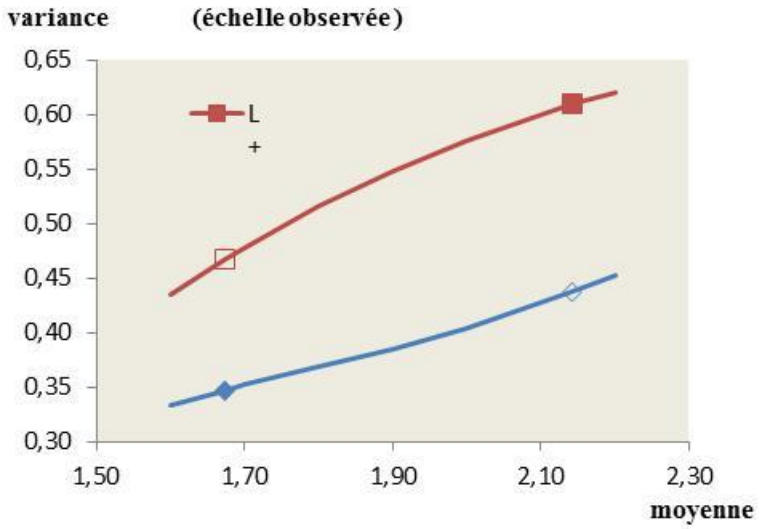


Figure 13 : Liaison moyenne – variance sur l'échelle observée pour différentes valeurs de la moyenne η de la variable sous-jacente et l'ensemble des seuils estimés pour chaque génotype L+ et ++(Martin et al., 2014).

2.3 Etude de la variabilité due à plusieurs gènes

2.3.1 Introduction

Nous avons vu dans lapartie précédente comment modéliser l'effet d'un gène sur la variabilité de la taille de portée. Nous proposons dans l'article 2 de modéliser, des effets polygéniques. Deux modèles, dérivés du modèle à seuil homothétique ont été développés :

Le modèle à seuils homothétiques Ce modèle fait l'hypothèse que, pour des seuils fixés, la variable sous-jacente l du modèle suit un modèle de canalisation tel que décrit en 2.3 (page 27):

$\mathbf{t} = (0, 1, \tau_3, \ldots, \tau_{n-1})$; $l \,|\, u, v \sim \mathrm{N}\left(\mu + u, f^2(\eta, v)\right)$.

Ce modèle a été appelé « homothétique » car, dans le cas de la paramétrisation alternative du modèle consistant à fixer les paramètres de la loi de la sous-jacente (i.e. $l \,|\, u, v \sim \mathrm{N}(0,1)$) et modéliser la matrice des seuils $\left(\tau_{\mathbf{1.}}^{'}, \tau_{\mathbf{2.}}^{'}, \tau_{\mathbf{3.}}^{'}, \ldots, \tau_{\mathbf{(n-1).}}^{'}\right)$; le passage du vecteur de seuils $\tau_{.\mathbf{i}}^{'}$ (individu i) à $\tau_{.\mathbf{i'}}^{'}$ (individu $i^{'}$) se fait par une translation et une homothétie. Ceci limite théoriquement le panel de variabilités modélisables. C'est pourquoi nous avons proposé un second modèle : le modèle à seuils indépendants.

Le modèle à seuils indépendants Ce modèle lève la contrainte de la translation-homothétie entre les seuils du modèle homothétique afin de permettre une plus grande flexibilité de modélisation. Ainsi, chaque seuil est fonction d'effets fixes et aléatoires qui lui sont spécifiques. Ce modèle peut également être paramétrisé de manière équivalente sous la forme d'un modèle multicaractères. En effet, la taille de portée ordinale à N classes peut être transformée en $N - 1$ variables à 2 classes comme indiqué ci-dessous :

Taille de portée	X_1	X_2	X_3
1	1	0	0
2	0	1	0
3	0	0	1
4	0	0	0

Le modèle à seuils indépendants correspond alors à un modèle multicaractères entre $X_1 \ldots X_{(N-1)}$, sous la contrainte que la somme des probabilités des variables X_j soit inférieure 1. Nous avons comparé les modèles à seuils homoscédastique, homothétiques et indépendants sur les données de taille de portée de Romane (Article 2). Les résultats obtenus ont montré que le modèle homothétique présentait le meilleur ajustement aux données.

2.3.2 Article 2: Genetic heteroscedastic models for ordinal traits: Application to sheep litter size

GENETIC HETEROSCEDASTIC MODELS FOR ORDINAL TRAITS: APPLICATION TO SHEEP LITTER SIZE

Samira Fathallah[a,b,c], Loys Bodin [a,b,c], Ingrid David [a,b,c]

[a] INRA, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), F-31326 Castanet-Tolosan, France

[b] Université de Toulouse, INP, ENSAT, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), F-31326 Castanet-Tolosan, France

[c] Université de Toulouse, INP, ENVT, GenPhySE (Génétique, Physiologie et Systèmes d'Elevage), F-31076 Toulouse, France

Corresponding author:

Samira Fathallah

Email addresses:
SF: samira.fathallah@toulouse.inra.fr

LB: loys.bodin@toulouse.inra.fr

ID: ingrid.david@toulouse.inra.fr

Abstract

Background: Classical genetic canalization models, that accommodate the mean and the variance of a trait separately, provide a flexible approach to take into account heteroscedasticity for continuous traits. However this model is not appropriate for discrete traits. The aim of this work is to propose heteroscedastic threshold models suitable for the genetic analysis of ordinal data.

Methods: In order to first fit the mean and the variance of ordinal traitsseparately, we extended the classical threshold model $[TM]$ for discrete data by introducing non-genetic and genetic factors of heterogeneity on the variance of its underlying variable, leading to a homothetic threshold model $[HTM]$ and its alternative parameterization $[HTM']$ in which the thresholds of different individuals are linked by a homothety-translation. Relaxing the constraint between the thresholds led us to propose an independent threshold model $[ITM]$ that was more flexible than $[HTM']$ but required estimating more parameters. The $[TM]$,$[HTM]$ and$[ITM]$ models were applied to study 19,671 litter size records in Romane sheep.

Results: Both the $[HTM]$ and $[ITM]$ models were able to disentangle the link between the mean and the variance that holds in the classical homoscedastic threshold model. The results obtained for the litter size of Romane ewes showed that the data was best fitted with the$[HTM]$ model compared to the $[ITM]$ and $[TM]$ models. The correlations between the observed and predicted variances were 0.6 and 0.2 for $[HTM]$ and $[TM]$ models, respectively. These analyses showed the existence of a genetic component for the heterogeneity of litter size in sheep that was taken into account in the $[HTM]$ model.

Conclusion: The $[HTM]$ model is the most suitable model for studying the variability of litter size in sheep. It accommodates both the mean and variance separately while requiring only few parameters to be estimated.

Canalization / discrete data / model on each threshold / litter size / sheep

Background

Over the last decades the implementation of production systems with lower impacts on the environment has resulted in substantial changes in livestock production. The use of less constrained systems that are also more sustainable, results in a higher variability of production [1]. Due to this, and other reasons such as management, quality product, animal welfare, consumer demand, etc., robustness is becoming an important challenge in most livestock selection programs. According to Knap [2], robustness can be viewed as the capacity of an animal to "combine high production potential with resilience to external stressors, allowing for unproblematic expression of high production potential in a wide variety of environmental conditions". In this context, animal longevity or the duration of production can be criteria of robustness and can be analyzed and selected for with the same methodology as other production traits. However, robustness can also be viewed as the ability to maintain a stable phenotype whatever the environmental conditions. Contrary to the previous approach, in this case robustness is not itself a trait, but a feature of each trait [1]. In line with this second approach, several genetic models have been proposed to explain the greater or lesser stability of genotypes in different environments [3–5]. One of these models, presented by Scheiner and Lyman [6] for canalizing selection, considers that the expression of a trait is controlled by two sets of genes. One set controls the mean of the performance and the other the environmental variability of the trait. Various authors have developed heteroscedastic models (i.e. canalization models) to jointly estimate the influence of various factors on both the mean and the variability of a trait [7–10].

These models were implemented with frequentist [11, 12] or bayesian methods [13, 14], and applied to a large variety of continuous traits (birth weight of piglets, birth weightofrabbits, weight of chickens, etc.).

The methods described above mainly target continuous traits. Theoretically, these linear models do not analyze discrete traits adequately because of their non-normal distributions. However, some previous empirical studies have shown that linear and non-linear models show similar performance in terms of goodness-of-fit and predictive ability [15, 16] when analyzing the level of discrete livestock data. But

it is not known whether these properties are still true when dealing with the variability of the discrete data.

The main objective of this research was to propose new models that take into account the heterogeneous individual variability of ordinal traits using derived threshold models. The models are described in a consecutive manner in the Methods section. Heteroscedastic extensions of the threshold model are presented in a first subsection, and a numerical application of these models to litter size in Romane sheep is presented in the second subsection.

1) Methods

2-1) Heteroscedastic models for ordinal traits

2-1-1) Homothetic threshold models for ordinal traits

The new model we have developed is based first on heteroscedastic models for continuous variables described by Garcia et al. [17]

$$\boldsymbol{y} = \mu + \boldsymbol{u} + f(\eta, \boldsymbol{v})\,\boldsymbol{\varepsilon}$$

Where $\boldsymbol{y}$ is the vector of observations, μ the overall mean, $\boldsymbol{u}$ the vector of additive genetic effects affecting the mean, η and $\boldsymbol{v}$ the overall mean and the vector of additive genetic effects affecting the variability of the trait. $\boldsymbol{\varepsilon}$ is a random standardized normal variable: $\boldsymbol{\varepsilon} \sim \mathrm{N}(0,1)$. Several functions f have been proposed in the literature (exponential, linear or square root function [8]). As usual, the vectors of genetic values are assumed to be Gaussian:

$$\begin{pmatrix} u \\ v \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G} \otimes \boldsymbol{A}\right)$$

Where $\otimes$ denotes the Kronecker product, $\boldsymbol{A}$ is the known additive genetic relationship matrix and **G** the genetic matrix $\begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix}$ where σ_u^2 and σ_v^2 are the variances of the additive genetic effects affecting the mean and the variability of the trait, and ρ their coefficient of correlation. ,

The new model is based secondly on the homoscedastic threshold model $[\mathrm{TM}]$ for ordinal traits described by Gianola [18]

$$[\mathrm{TM}] \qquad \mathbf{t} = (0, 1, \tau_3, \ldots, \tau_{n-1}) \;\; ; \;\; l|u \sim \mathrm{N}\ (\mu + u, 1)$$

This model considers a hypothetical underlying normally distributed variable $l|u$ called liabilityand a set of thresholds $\mathbf{t}$ which transform this distribution into the observed discrete variable y in n categories.

Indeed, a simple way to analyze the genetic variability of an ordinal trait is to introduce heteroscedasticity on the liability of the threshold model. We call this new model the homothetic threshold model:

$$[\mathrm{HTM}] \qquad \mathbf{t} = (0, 1, \tau_3, \ldots, \tau_{n-1}) \;\; ; \;\; l|u, v \sim \mathrm{N}\ \left(\mu + u, f^2(\eta, v)\right)$$

It considers an underlying variable $l|u, v$ which varies in mean and variance according to environmental and genetic effects and a common set of thresholds $\mathbf{t}$ for all subjects. As before, the environmental effects affecting the mean and variance of the liability are (μ, η), and the parameters of the G (co)variance matrix are $(\sigma_u^2\ \sigma_v^2,\ \rho)$ of the genetic effects (u, v) that affect the mean and variance of the liability.

A fully equivalent parameterization of this $[\mathrm{HTM}]$ model can be written. This other parameterization (or $[\mathrm{HTM}']$ model) is reached by fixing the parameters of the liability and considering that the matrix of thresholds $\mathbf{t}'$ is affected by environmental and genetic effects.

$$\left[HTM^{'}\right] \quad \mathbf{t}^{'} = \begin{pmatrix} \tau^{'}_{11} & \tau^{'}_{21} & \cdots & \tau^{'}_{(n-1)1} \\ \tau^{'}_{1i} & \tau^{'}_{2i} & \cdots & \tau^{'}_{(n-1)i} \\ . & . & . & \\ \tau^{'}_{1N} & \tau^{'}_{2N} & \cdots & \tau^{'}_{(n-1)N} \end{pmatrix} = \begin{pmatrix} \mathbf{t}^{'}_{1} \\ \mathbf{t}^{'}_{i} \\ . \\ \mathbf{t}^{'}_{N} \end{pmatrix} = \left(\tau^{'}_{1.}, \tau^{'}_{2.}, \tau^{'}_{3.}, \cdots, \tau^{'}_{(n-1).}\right) \; ; \; l \sim \mathrm{N}\,(0,1)$$

Under this parameterization, the vector of thresholds $\mathbf{t}^{'}_{i}$ is specific to each individual i; with $\mathbf{t}^{'}_{i} = \boldsymbol{\mu}^{'} + u^{'}_{i}\boldsymbol{J} + \boldsymbol{v}^{'}_{i}$ where $\boldsymbol{\mu}^{'}$ is an overall set of $n-1$ thresholds, $u^{'}_{i}$ a subject-specific deviation from this overall set and $\boldsymbol{v}^{'}_{i}$ a vector of size $n-1$ which reflects the specific deviation of subject threshold from this overall set with:

$$\begin{bmatrix} \boldsymbol{u}^{'} \\ \boldsymbol{v}^{'}_{.1} \\ \boldsymbol{v}^{'}_{.2} \\ \boldsymbol{v}^{'}_{.3} \\ \\ \boldsymbol{v}^{'}_{.n-1} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma^{2}_{u^{'}} & 0 & \sigma_{u^{'},v^{'}}\left(\mu^{'}_{2}-\mu^{'}_{1}\right) & \sigma_{u^{'},v^{'}}\left(\mu^{'}_{3}-\mu^{'}_{1}\right) & \sigma_{u^{'},v^{'}}\left(\mu^{'}_{n-1}-\mu^{'}_{1}\right) \\ 0 & 0 & 0 & 0 & 0 \\ & & \left(\mu^{'}_{2}-\mu^{'}_{1}\right)^{2}\sigma^{2}_{v^{'}} & \left(\mu^{'}_{2}-\mu^{'}_{1}\right)\left(\mu^{'}_{3}-\mu^{'}_{1}\right)\sigma^{2}_{v^{'}} & \left(\mu^{'}_{2}-\mu^{'}_{1}\right)\left(\mu^{'}_{n-1}-\mu^{'}_{1}\right)\sigma^{2}_{v^{'}} \\ & & & \left(\mu^{'}_{3}-\mu^{'}_{1}\right)^{2}\sigma^{2}_{v^{'}} & \left(\mu^{'}_{3}-\mu^{'}_{1}\right)\left(\mu^{'}_{n-1}-\mu^{'}_{1}\right)\sigma^{2}_{v^{'}} \\ & & & & \left(\mu^{'}_{n-1}-\mu^{'}_{1}\right)^{2}\sigma^{2}_{v^{'}} \end{bmatrix} \otimes \boldsymbol{A}\right)$$

The transformation of $\mathbf{t}^{'}_{i}$ to $\mathbf{t}^{'}_{i^{'}}$ (from individuals i to $i^{'}$) is not a translation but a homothety-translation which means that, the ratio $\left(\tau^{'}_{k_2 i} - \tau^{'}_{k_1 i}\right) \Big/ \left(\tau^{'}_{k_4 i} - \tau^{'}_{k_3 i}\right)$ is constant for the whole population:

$$\left(\tau^{'}_{k_2 i} - \tau^{'}_{k_1 i}\right) \Big/ \left(\tau^{'}_{k_4 i} - \tau^{'}_{k_3 i}\right) = \left(\tau^{'}_{k_2 i^{'}} - \tau^{'}_{k_1 i^{'}}\right) \Big/ \left(\tau^{'}_{k_4 i^{'}} - \tau^{'}_{k_3 i^{'}}\right) = \left(\mu^{'}_{k_2} - \mu^{'}_{k_1}\right) \Big/ \left(\mu^{'}_{k_4} - \mu^{'}_{k_3}\right), \forall\, k_1, k_2, k_3, k_4$$

$\left[\mathrm{HTM}\right]$ and $\left[\mathrm{HTM}^{'}\right]$ are equivalent models with the same number of parameters (n+2) to estimate.

The relationship between $\mathbf{t}^{'}$ and $\mathbf{t}$ is: $\boldsymbol{t}^{'}_{i} = \dfrac{\boldsymbol{t} - (\mu + u_i)\boldsymbol{J}}{f\left(\eta, v_i\right)}$

The estimations obtained with the $\left[\mathrm{HTM}^{'}\right]$ model can be transformed into those that would have been obtained with the $\left[\mathrm{HTM}\right]$ model. In the case of a linear function f, the formula are (see Additional file 1: Appendix 1 for demonstration):

$$\sigma_v^2 = \frac{\sigma_{v'}^2}{\left(\mu_2' - \mu_1'\right)^2}$$

$$\sigma_u^2 = \frac{1}{\left(\mu_2' - \mu_1'\right)^2}\left(\left(\mu_1'\right)^2\sigma_{v'}^2 + \sigma_{u'}^2 + \sigma_v^2\sigma_u^2 + \left(\sigma_{v',u'}\right)^2 - 2\mu_1'\sigma_{v',u'}\right) \qquad \rho = \frac{1}{\left(\mu_2' - \mu_1'\right)^2}\left(\frac{\sigma_{v',u'} - \mu_1'\sigma_{v'}^2}{\sigma_u\sigma_v}\right)$$

The heritability of the trait at the level of the liability is estimated by the ratio of the animal variance acting on the mean to the total variance of the underlying variable $h_l^2 = \frac{\sigma_u^2}{\sigma_u^2 + f^2(\eta, \sigma_v)}$. For instance in the case of a linear function $f : h_l^2 = \frac{\sigma_u^2}{\sigma_u^2 + \eta^2 + \sigma_v^2}$.

2-1-2) Homothetic threshold model extension: Independent threshold model

A straightforward extension of the above model $\left[\mathrm{HTM}'\right]$ can also be envisaged. That consists in releasing the constraint of a homothetic-translation link between the thresholds of the individuals. The model therefore becomes:

$$[\mathrm{ITM}] \quad \mathbf{t}'' = \begin{pmatrix} \tau''_{11} & \tau''_{21} & \cdots & \tau''_{(n-1)1} \\ \tau''_{1i} & \tau''_{2i} & \cdots & \tau''_{(n-1)i} \\ \cdot & \cdot & \cdot & \\ \tau''_{1N} & \tau''_{2N} & \cdots & \tau''_{(n-1)N} \end{pmatrix} = \begin{pmatrix} \mathbf{t}''_1 \\ \mathbf{t}''_i \\ \cdot \\ \mathbf{t}''_N \end{pmatrix} = \left(\tau''_{\mathbf{1.}}, \tau''_{\mathbf{2.}}, \tau''_{\mathbf{3.}}, \cdots, \tau''_{\mathbf{(n-1).}}\right) \; ; \; l \sim \mathrm{N}(0,1)$$

with $\mathbf{t}''_i = \boldsymbol{\mu}'' + \boldsymbol{v}''_i$ where $\boldsymbol{\mu}''$ is an overall set of $n-1$ thresholds, and $\boldsymbol{v}''_i$ a vector of size $n-1$ which reflects the specific deviation of subject threshold from this overall set with:

$v'' \sim N(0, G'' \otimes A)$ where G'' is an unstructured (co)variances matrix, the parameters of which have to be estimated.

$$\begin{bmatrix} v''_{.1} \\ v''_{.i} \\ \\ v''_{.n-1} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma^2_{v''_{.1}} & \sigma_{v''_{.1}v''_{.2}} & \sigma_{v''_{.1}v''_{.n-1}} \\ & \sigma^2_{v''_{.i}} & \sigma_{v''_{.1}v''_{.i}} \\ & & \sigma^2_{v''_{.n-1}} \end{bmatrix} \otimes \boldsymbol{A}\right)$$

The total number of parameters in this model is $(n+2)(n-1)/2$. It should be noted that $[\mathrm{ITM}]$ cannot be parameterized like $[\mathrm{HTM}]$ using a fixed set of thresholds and a model on the liability $\boldsymbol{\iota}$. In this model fixing the thresholds to modify the parameterization would imply that the underlying variable would no longer be Gaussian.

Due to the lack of constraints between the thresholds (excepting the obvious relationship ($-\infty < \tau''_1 < \tau''_2 < \ldots < \tau''_{n-1} < \infty$)), $[\mathrm{ITM}]$ fits a much larger variability of situations than the homothetic threshold model which presumes a strong constraint either on a homothetic-translation relationship between thresholds ($[\mathrm{HTM}']$) or on the normality of the liability ($[\mathrm{HTM}]$).

Interestingly, we found that another parameterization of the $[\mathrm{ITM}]$ model is a constrained multiple-trait model $[\mathrm{MTM}]$ (see Additional file 1: Appendix 2). Indeed, each observation of the ordinal trait can be seen as the result of n exclusive binary traits:

$$1_k(y) := \begin{cases} 1 \text{ if } y = k \quad / \quad k = 1 \ldots n \\ 0 \text{ otherwise} \end{cases}$$

The observed variable y can thus be transformed into n indicator functions which can be analyzed using a multivariate binary threshold model. Only $n-1$ indicator functions have to be considered in this $[\mathrm{MTM}]$ since the value of one of the indicator functions is fixed conditionally to the others. Each sub-model k, $k \in [1, n-1]$ can be written:

$$[\mathrm{MTM}]_k \qquad \tau''^{*}_k \left| u''^{*}_k \right. ; \; l_k \sim \mathrm{N}(0,1)$$

According to this modelling, the probability that an individual might achieve the observation *k* or that the kth indicator function is 1 is:

$$\pi_k = P(1_k(y) = 1) = P\left(l_k \le \tau_k^{"*} \middle| u_k^{"*}\right)$$

and the *n*-1 probabilities π_k should satisfy the relation $\sum_{k=1}^{n-1} \pi_k < 1$. This constraint must be added to the multivariate binary threshold model.

As for $\left[\mathrm{HTM}^{'}\right]$, there are $(n+2)(n-1)/2$ parameters to estimate: $n-1$ thresholds and $n*(n-1)/2$ variance covariance parameters.

The relationship between the thresholds of $[\mathrm{MTM}]$ and $[\mathrm{ITM}]$ is $\tau_{ik}^{"*} = \Phi^{-1}\left(\Phi(\tau_{ik}^{"}) - \Phi(\tau_{i(k-1)}^{"})\right)$.

Under these last models, one cannot really consider that the expression of the ordinal trait is controlled by two sets of genes. The combination of elementary genetic values for each individual results in the heteroscedasticity of the global trait. In this model, the heritability can be calculated for each class as $h_k^2 = \dfrac{\sigma^2_{u_k^{"*}}}{\sigma^2_{u_k^{"*}} + 1}$.

The different parameterizations of the models are summarized in table 1. Models $\left[\mathrm{HTM}^{'}\right]$ and $[\mathrm{ITM}]$ can be fitted using the ASReml software.

Table 1: Main characteristics of the different models

Models		No. parameters	Population		For animal i	
			Thresholds	Underlying	Thresholds	Underlying
Homoscedastic	[TM]	n	$\mathbf{t}^{*} = \{\mathbf{t}_1^{*}, \mathbf{t}_2^{*}, \mathbf{t}_3^{*}, ..., \mathbf{t}_{n-1}^{*}\}$	$Y \sim N(0,1)$	$t_i^{*} = \{t_{1i}^{*}, t_{2i}^{*}, t_{3i}^{*}, ..., t_{(n-1)i}^{*}\}$ $t_{.i}^{*} = t_i + \mu_i + u_i$	$y_i = \varepsilon_i$
Heteroscedastic	[HTM]	n+2	$\mathbf{t} = \{0, 1, \mathbf{t}_3, ..., \mathbf{t}_{n-1}\}$	$Y \sim N(\mu + u, f^2(\eta, \mathbf{v}))$	$t_i = \{0, 1, t_{3i}, ..., t_{(n-1)i}\}$	$y_i = \mu_i + u_i + \varepsilon_i$
	[HTM']	n+2	$\mathbf{t}' = \{\mathbf{t}_1', \mathbf{t}_2', \mathbf{t}_3', ..., \mathbf{t}_{n-1}'\}$	$Y \sim N(0,1)$	$t_i' = \{t_{1i}', t_{2i}', t_{3i}', ..., t_{(n-1)i}'\}$ $t_{.i}' = t_i' + \mu_i' + u_i' + v_i'$	$y_i = \varepsilon_i$
	[ITM]	(n+2)(n-1)/2	$\mathbf{t}'' = \{\mathbf{t}_1'', \mathbf{t}_2'', \mathbf{t}_3'', ..., \mathbf{t}_{n-1}''\}$	$Y \sim N(0,1)$	$t_i'' = \{t_{1i}'', t_{2i}'', t_{3i}'', ..., t_{(n-1)i}''\}$ $t_{.i}'' = t_i' + \mu_i'' + u_i''$	$y_i = \varepsilon_i$

1. Numerical application

1.1 Data

The different models were applied to study litter size (LS) in Romane sheep. All data (litter size, environmental variation factors and pedigrees over ten generations) were extracted from the national database for genetic evaluation and research managed by the *Institut de l'Elevage* (French Livestock Production Institute) and the *Centre de Traitement de l'Information Génétique* (Genetic Information Processing Center, Jouy-en-Josas, France). Data consisted of the litter sizes of 11,073 multiparous ewes from 23 flocks with at least four lambings from 2000 to 2013. The mean prolificacy (μ = 2.12) in this sample was high and consequently there was a high global variability of litter size. The main features of the sample are presented in table 2. To avoid extreme case problems [16], litter sizes of more than four lambs were pooled into a single class for all analyses. This resulted in four classes of LS being considered: [1, 2, 3, 4+].

Table 2: Features of Romane sheep.

Item	
No. of records	19 671
No. of animals	11 073
No. of sires	1 096
No. of dams	6 550
Mean	2.12
Distribution of litter size (%)	
Single	17.70
Twin	55.01
Triplet	24.18
Quadruplet and +	3.11

The fixed effects included in the analysis were selected so as to be as similar as possible to those of the national genetic evaluation system [Institut de l'Elevage 2011]. This was done by comparing the likelihood-ratio of all the possible linear mixed models which included the following fixed effects: age of ewe (9 levels), time interval between lambing in months (7 levels), parity (7 levels), age of the ewe at first lambing in months (9 levels), season of lambing (2 levels), as well as the random effects: year-

flock and ewe. For this step of the analysis, the ewe effect was considered as random but the genetic relationships between individuals were ignored.

1.2 Models fitted - Estimated parameters – Comparison of methods

The three previously described threshold models $[\mathrm{TM}]$, $\left[\mathrm{HTM}^{'}\right]$ and $[\mathrm{ITM}]$ were applied to this data. The classical threshold model $[\mathrm{TM}]$, used as a base of comparison, included the previously selected fixed effects and the ewe as both a random genetic effect (u^{*}) and a random permanent environmental effect (p^{*}) (i.e. $\boldsymbol{t}_i^{*} = \boldsymbol{x}_i\boldsymbol{\beta} + u_i^{*}\boldsymbol{J} + p_i^{*}\boldsymbol{J}$). The homothetic threshold model $\left[\mathrm{HTM}^{'}\right]$ that fitted the data considered the fixed, genetic and permanent effects to be common for all thresholds and threshold-specific random genetic effects with constrained genetic (co)variance to ensure homothety-translation (i.e. $\boldsymbol{t}_i^{'} = \boldsymbol{x}_i\boldsymbol{\beta} + u_i^{'}\boldsymbol{J} + p_i^{'}\boldsymbol{J} + \boldsymbol{v}_i^{'}$). Finally, the further extension leading to the independent threshold model $[\mathrm{ITM}]$ was fitted. It also considered that the fixed and permanent effects were common for all thresholds and included a genetic effect specific to each threshold with no constrained genetic (co)variance (i.e. $\boldsymbol{t}_i^{''} = \boldsymbol{x}_i\boldsymbol{\beta} + p_i^{''}\boldsymbol{J} + \boldsymbol{v}_i^{''}$).

ASReml software [20] was used to implement all the models.

Models were compared using the Likelihood ratio test for nested models ($[\mathrm{TM}]$ vs. $\left[\mathrm{HTM}^{'}\right]$, LRT dof = 2 ; $[\mathrm{TM}]$ vs. $[\mathrm{ITM}]$, LRT dof = 5) and the Bayesian Information Criterion for non-nested models ($\mathrm{BIC} = \ln(N - p) * c - 2\ln L$, $\left[\mathrm{HTM}^{'}\right]$ vs. $[\mathrm{ITM}]$, where L is the restricted maximum likelihood of the model, N the number of observations, p and c the number of fixed effects and covariance parameters) [21]. The models were compared also based on their ability of predicting new records as follows: 75% of records were used to estimate parameters which were used to predict the remaining 25%. Five replicates of this design were sampled at random. To evaluate the predictive ability of the models, MSEP (mean square error of prediction) statistic was computed as:

$\mathrm{MSEP} = \frac{1}{\mathrm{M}} \sum_{j=1}^{\mathrm{M}} \left(\frac{1}{m} \sum_{i=1}^{m} \left(y_{ij} - y'_{ij} \right)^2 \right)$ where y , y' corresponds to the observed value and estimated mean, *m* is the number of data in the validation subset and M is the number of the subset.

For ewes with observations, we compared the raw means and variances of their litter sizes to the corresponding predicted values on the observable scale obtained with the different models. The raw mean and variance were calculated for each individual from the observed data. The predicted values were calculated as: $\hat{\bar{y}}_i = \sum_{k=1}^{4} k \hat{\pi}_{i,k}$ and $\hat{\sigma}_i^2 = \sum_{k=1}^{4} k^2 \hat{\pi}_{i,k} - \left(\sum_{k=1}^{4} k \hat{\pi}_{i,k} \right)^2$ where $\hat{\pi}_{i,k}$ was the probability that the litter size of a ewe *i* was *k*. These probabilities were calculated using $\hat{\pi}_{i,k} = \Phi(\hat{\tau}_{i,k}) - \Phi(\hat{\tau}_{i,k-1})$; with as before $\tau_0 = -\infty$ and $\tau_4 = +\infty$ ($\hat{\tau}_{i,k}$ corresponds to $\hat{\tau}^*_{i,k}, \hat{\tau}'^*_{i,k}$ or $\hat{\tau}''_{i,k}$ depending on the model).

To measure the ability of the models to break the link between $\hat{\bar{y}}_i = t$ and $\hat{\sigma}_i^2 = z$, we propose using the criterion $\mathrm{B} = \frac{E(Var(z \mid t))}{Var(z)}$ which varies from 0 to 1[1]. Because the density associated to the variance $Var(z \mid t)$ is unknown, data were ordered in ascending order of LS and were discretized into s subgroups, and the variance of y in each subgroup was calculated $\hat{\sigma}_m^2$ in order to calculate $E(Var(z \mid t)) = E(\hat{\sigma}_{y_i}^2 \mid \hat{\bar{y}}_i) = \sum_{m=1}^{s} \hat{\sigma}_m^2 \Big/ s$. The number of subgroup was defined in order to get enough data in each subgroup to obtain a good estimation of the subgroup variance.

The intensityof the dispersion of each model was also assessed by calculating the correlation, either for the mean or the variance, between the observed values and the predicted values on the observable scale.

1

$Var(z) = E\left(Var(z \mid t)\right) + Var\left(E(z \mid t)\right)$

with $E\left(Var(z \mid t)\right) = 0 \Rightarrow z = f(t)$ is one-to-one

and $Var\left(E(z \mid t)\right) = 0 \Rightarrow z$ is a constant

2. Results

3.1 Model comparisons

The values of the criteria (LRT, BIC, MSE and MSEP) used to compare the adjustment quality of the models are provided in tables 3 and 4.

Results showed that the homoscedastic model should be rejected and that the models that account for a genetic heterogeneity of litter size distribution should be favored. Among the heteroscedastic models, BIC values were found to be lower with the $[HTM']$ model than with the $[ITM]$ model. The MSE estimates of the mean were similar for the three models, and the MSE estimates of the variance were different, the MSE of the variance were higher in $[TM]$ model (0.13 in $[TM]$ model vs 0.10 in $[HTM']$ model). The differences between the MSEP values are small; this was explained by the fact that this parameter permits at estimating only prediction error for the level of the trait and not for its variability. For measuring the quality of model to predict the variance, we calculated the B criteria on the validation subset discretizing the whole data into 200 adjacent subgroups. The results are reported in the table 4. The $[HTM']$ and $[ITM]$ models have the better predictive ability of the relationship between the mean and the variance than that of the $[TM]$ model l (0.36 for $[HTM']$ model (0.37 for $[ITM]$ model) vs 0.005 (for the $[TM]$ model).

Table 3: Values obtained from Likelihood-ratio, BIC tests, and mean squared error (MSE) between observed and fitted values in the different models.

Models/tests	Likelihood-ratio (LRT)	BIC
Threshold model ([TM*])	[TM*] vs. [HTM*]: >> 11	54 005
Homothetic threshold model ([HTM*])	(Reject H0.P-value < 0.001)	49 517
Independent threshold model ([ITM])	[TM*] vs. [ITM]: >>16 (Reject H0.P-value < 0.001)	50 663

Table 4: Values obtained from mean squared error of prediction (MSEP) and B criteria on the validation subset in the different models.

Models/tests	MSEP	B criteria
Threshold model ($[TM]$)	0.473	0.005
Homothetic threshold model ($[HTM']$)	0.471	0.363
Independent threshold model ($[ITM]$)	0.481	0.371

3.2 Homoscedastic threshold model as reference ([TM])

With the $[TM]$ model, the genetic additive variance of the liability was $\sigma^2_{u.} = 0.079$ and the variance of the permanent environmental effect was $\sigma^2_{p.} = 0.034$, resulting in a heritability of $h^2 = 0.071$. The correlation between the predicted and the observed mean was very high (r = 0.80) (table 8), but low between the predicted and observed variances (r = 0.20). Despite the curvilinearity between the predicted means and variances, the correlation between these variables was high (r = 0.99) and not different from 1, and the B parameter was 0.013, evidencing the one–to-one relationship between these variables when this homoscedastic model was applied as shown in the scatterplot in figure 1A.

3.3 Homothetic threshold model ($[HTM']$)

The genetic additive variances obtained with $[HTM']$ are given in table 5. By design, the correlation between the genetic values associated to thresholds (v) was 1 in this model and the variance $\sigma^2_{v'.1}$ was 0. Moreover, and as expected due to the homothetic relationship, we observed an increase of $\sigma^2_{v'.3}$ compared to $\sigma^2_{v'.2}$. The variances were 0.04, 0.05 and 0.20 for $\sigma^2_{u'.1}$, $\sigma^2_{v'.2}$ and $\sigma^2_{v'.3}$ respectively. The variance of the permanent environmental effect was 0.015.

Table 5: Genetic variance and correlation between thresholds for Romane sheep using the [HTM'] model.

Thresholds /genetic variance	$\sigma^2_{u'}$	$\sigma^2_{v'.2}$	$\sigma^2_{v'.3}$
	0.04	0.13	0.13
		0.05	1
			0.20

By transforming the results of the $[HTM']$ parameterization to that of $[HTM]$, we could estimate the variance of the genetic effects affecting the mean of the liability as well as the correlation between these effects. The values found for the mean of the liability, the scaling function and the correlation between effects were $\sigma^2_u = 0.069$, $\sigma^2_v = 0.020$ and $r_{g(u,v)} = -0.40$, respectively. The variance of the permanent effect was $\sigma^2_p = 0.015$ while by design the mean of the residual variance was $\eta^2 = 1$. Since the model fitted for the environmental variance did not include fixed effects but only a general mean and a genetic effect, the heritability of the liability mean could be estimated independently of a given environment and was: $h^2_l = \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_p + \eta^2 + \sigma^2_v} = 0.062$.

The correlation between observed and predicted means was high (r = 0.79) and not different from the value obtained with the $[TM]$ model. This contrasts with the correlation between predicted and observed variances (r = 0.61). Although the prediction of the variances was still weaker than that of the means, it was much higher than with the $[TM]$ model. The $[HTM']$ model could therefore break the link between the predicted means and variances and the B parameter was 0.51 even if the correlation between these variables was still high (r = 0.88). The scatterplot in figure 1B shows this variability of the predicted variance for a given predicted mean.

In this homothetic-translation model, the gap Δ_{1-2} between the two first thresholds is an indirect measure of its variability for each individual. As an example, we estimated the distribution of the litter size of the daughters of two sires in particular (having more than 30 daughters) that had extreme Δ_{1-2}

values but the same average litter size(table 6). For these extreme sires, the percent difference of grand-offspring born as quadruplets or more can be very high (2% vs. 15%).

Table 6: Predicted litter size distribution for 2 sires with different prolificacy and different gaps between the first 2 thresholds:

Sire	Daughter prolificacy	Variance	Δ_{1-2}	%LS 1	%LS 2	%LS 3	%LS 4+
1	2.12	0.38	1.79	14%	61%	24%	1%
2	2.12	0.58	1.32	24%	46%	22%	8%

Δ_{1-2} : The gap between the first thresholds.

3.4 Independent threshold model ([ITM])

The genetic variances and correlations obtained with the $[ITM]$ model are provided in table 7. The genetic variances were low, but higher for the second threshold and the third threshold than for the first. The genetic correlations were moderate ($r_g > 0.4$), the highest correlation found existing between v_2'' and v_3''. Using the transformation formulae given in the appendix, it was possible to estimate the genetic variance associated with twin lambing ($\sigma^2_{ls=2} = 0.14$), the corresponding heritability ($h^2_{ls=2} = 0.128$) as well as the genetic correlation between the values for the first and the second classes of LS ($r_g = 0.44$). The correlations between the predicted and observed means (r = 0.78) and between the predicted and observed variances (r = 0.59) were close to the corresponding values obtained with the $[HTM]$ model (table 8). The B parameter ($B = 0.42$), the correlation between the predicted means and variances (r = 0.91) as well as the scatter plot (figure 1C) were also similar to those of the $[HTM']$ model. The Spearman rank correlations between the thresholds of the $[HTM']$ and $[ITM]$ models were high; the thresholds were close (table 9).

Table 7: Genetic variance and correlation between thresholds for Romane sheep using the [ITM] model

Thresholds / genetic variance	$\sigma^2_{v".1}$	$\sigma^2_{v".2}$	$\sigma^2_{v".3}$
	0.035	0.63	0.43
		0.138	0.92
			0.176

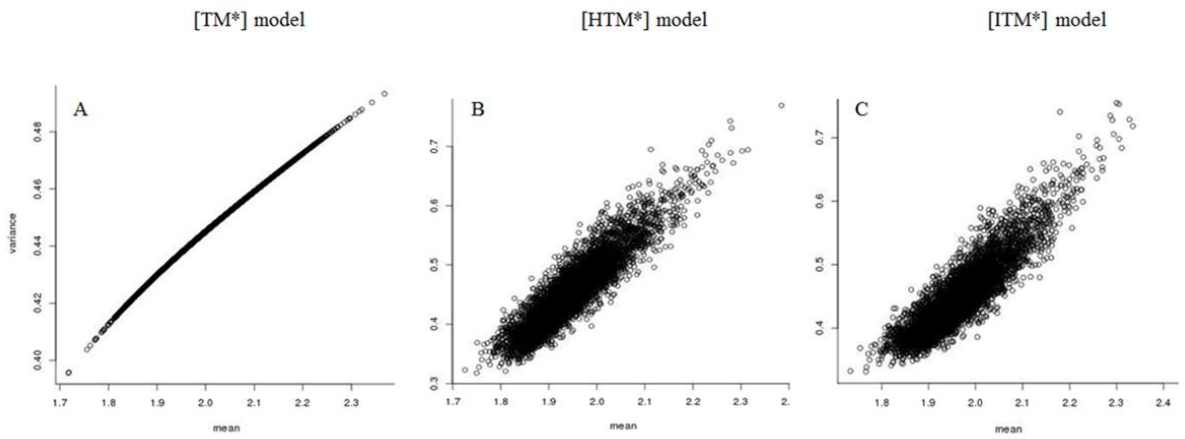


Figure 1: Scatterplots of estimated mean vs. estimated variance over the observed range for the three models.

Table 8: Correlations between the predicted and observed means as well as the predicted observed variances in the different models.

	Predicted mean vs. observed mean	Predicted variance vs. observed variance
Threshold model ([TM*])	0.80	0.21
Homothetic threshold model ([HTM*])	0.79	0.61
Independent threshold model ([ITM])	0.78	0.59

3. Discussion

A genetic heteroscedastic model has been proposed previously [11] to analyze categorical variables and break down the link between mean and variance that holds with the classical homoscedastic threshold model. This model originates from the combination of two base models: the homoscedastic threshold model for categorical data [22] and the structural heteroscedastic model for continuous variables [7]. In the present paper we have studied further extensions by analyzing the relationship that exists between the categories and in particular by releasing the main constraint.

In this paper we propose two kinds of heteroscedastic models, both with different parameterizations. The first type ($[HTM]/[HTM']$) was fully derived from the canalization model proposed for continuous traits [7] with which it was easy to draw comparisons. The second type consisted in relaxing the relationship that exists between thresholds (translation in the $[TM]$ or homothety-translation in the $[HTM]$) and led to a model which considers each threshold independently $[ITM]$. We demonstrated that this model was equivalent to a constrained multiple trait model $[MTM]$ on each class of the ordinal trait. This second type of model is much more flexible than the first and its application in selection would be of particular interest because allows for each class to be controlled through a multiple selection index [23] with specific environmental effects and economic weights for each class. Modeling each class of the trait to control its distribution or modeling only the most important class are trivial issues and have been already considered [24], however the correct multiple trait model must include a constraint on the sum of the probabilities of the n-1 binary variables. In spite of their potential interest, the use of this kind of model ($[ITM]$ or $[MTM]$) for genetic evaluation and selection would be limited since they require a larger number of parameters to be

estimated than ($[HTM]/[HTM']$) and would probably show convergence problems as soon as the number of classes is higher than 3. Furthermore, their biological interpretation is difficult; for instance we could not link the parameters of the $[ITM]$ or $[HTM]$ models to the biological model of Schneider and Lyman (2002) which assumes the existence of different genes that control the mean and the variance of a trait.

We compared the $[TM]$, $[HTM']$ and $[ITM]$ models through the analysis of LS data from both Romane and Rouge de l'Ouest sheep (results not shown). The low heritability of LS liability obtained with an homoscedastic threshold model ($[TM]$) was in accordance with previous studies which used similar threshold models [25–27]. However, LRT, BIC, and MSE parameters showed that this homoscedastic model was less successful than both of the heteroscedastic ones. Consequently, the LS distribution of sheep shows heteroscedasticity of genetic origin as revealed by the homothetic $[HTM']$ and independent threshold $[ITM]$ models.

Taking into account the genetic heteroscedasticity with the $[HTM']$ model and subsequent transformation of the parameters to those of the $[HTM]$ model did not substantially change the heritability of the liability compared to the $[TM]$ model. However with this heteroscedastic $[HTM]$ model, the additive genetic variance for the variability of the trait was lower than the corresponding variance for its mean. This is in line with the literature seeing that studies that have jointly estimated the additive genetic variances of the mean and the variability of a trait have generally found lower values for the variability than for the mean [11, 14, 28–33]. This seems to be a common rule whatever the methodology used for such estimations, the trait or the species. In contrast, the genetic correlation between the mean of a trait and its variability is highly variable and covers the whole range of its variation interval. For instance in a mouse population, this genetic correlation was -0.93 for litter size and 0.97 for individual birth weight [26]. The genetic correlation between the mean and the variability of sheep litter size estimated in the present study by the $[HTM']$ model was negative (rg = -0.40). As highlighted by Yang *et al.*[34] , this parameter might be affected by an artefact of the scale of

measurement or skewness of the data, although with the threshold model no such skewness affected this correlation. On the other hand, a not very different value (rg = -0.25) has been estimated for another French sheep breed (Rouge de l'Ouest) for which the mean prolificacy (prolificacy = 1.80; LS3+ = 12.5%) is lower than that of Romane sheep (results not shown).

Relaxing the homothetic constraint between the thresholds in $[HTM']$ model did not improve the fit of the model or the predictive ability of the data as one might have expected, but instead increased the BIC values due to the higher number of parameters to estimate in $[ITM]$. Moreover, although the constraint on the genetic correlation between thresholds was relaxed in the $[ITM]$ model, the correlation between the second and third thresholds remained close to 1 and the correlation between the first and the second thresholds was high, and also the Spearman rank correlations between the $[HTM']$ and $[ITM]$ models close to 1 (table 9). This suggests that the true nature of the link between the thresholds is almost a homothety-translation in sheep and that the distributions of LS classes are not independent but linked by a strong common law. It also means that the concept of sensitivity thresholds that transform individual Gaussian variables differing in mean and variance into the observed categories might be relevant.

Table 9: Spearman rank correlations between the [HTM'] and [ITM] models.

First threshold	0.988
Second threshold	0.989
Third threshold	0.988

Nonetheless in the present study, only genetic effects were considered as threshold-specific in the $[ITM]$ model in order to be able to compare the different models using LRT or BIC criteria. Of course, threshold-specific fixed effects and permanent effects could be included in the $[ITM]$ model. In the same manner, threshold-specific fixed effects and permanent effects could be included in the $[HTM']$ model. Unfortunately, the $[HTM]$ model with a threshold-specific permanent effect did not

converge and threshold-specific fixed factors with proportional effects cannot be implemented in ASReml.

Using this model, Martin, et al. [35] showed that the polymorphism of a single gene affected not only the mean prolificacy of Lacaune sheep, but also modified its variability by changing the gap between the thresholds in a homothetic way. A superiority of the homothetic model ($\left[\mathrm{HTM}^{'}\right]$), as well as a similar genetic variability of the LS liability have been found for Rouge de l'Ouest sheep (results not shown), thus covering the whole range of prolificacy in breeds where canalization might be of interest. By design, the gap between the first two thresholds estimated for each individual in the $\left[\mathrm{HTM}^{'}\right]$ model is inversely proportional to the individual residual variance of the $[\mathrm{HTM}]$ model. This criterion, which has been used to compare the variability of about 50 French sheep breeds [36], could be used to estimate the breeding value for the LS variability independently from that of the mean.

Conclusions and Prospects

In the application study using the $\left[\mathrm{HTM}^{'}\right]$ model, we did not include fixed effects acting on the variability seeing that there is no currently available software that takes into account this type of model properly (fixed effects + genetic effect). It may be possible to estimate fixed effects in the residual variance using an iterative process, testing the influence of each effect on the variance of the model. We did not test such algorithms on our data and it will probably show to be a slow and difficult process. Another solution would be to modify existing software for canalizing selection of a continuous trait in order to take into account the discrete characteristics of the data. For instance, data augmentation could be included in the Gibbs sampler in the GSEVM software [13] in order to fit a heteroscedastic threshold model. The availability of such software would mean that the environmental effects affecting LS variability could be studied in more detail and, more generally, that different combinations of criteria could be used to simulate the genetic progress that could be made on the mean and the variability.

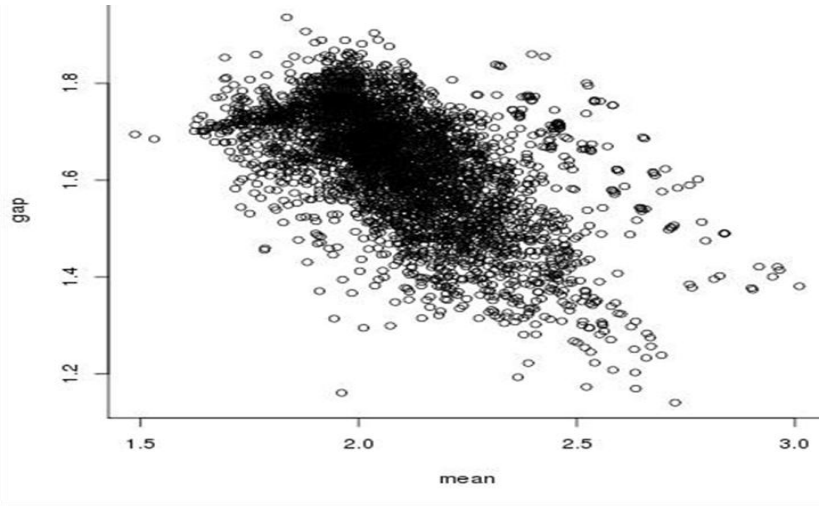


Figure 2: Scatterplot of estimates of the gap between the threshold and the mean predicted by the HTM model.

Appendices:

Appendix 1: $[\mathrm{HTM}]$ and $\left[\mathrm{HTM}^{'}\right]$

In this section we describe how we calculated the parameters σ_u^2, σ_v^2, and ρ_{uv} of a $[\mathrm{HTM}]$ model based on the $\left[\mathrm{HTM}^{'}\right]$ estimates. Let $t_k^{'}$ be the kth threshold of the $\left[\mathrm{HTM}^{'}\right]$ model with $\mathbf{t}_{.i}^{'} = \boldsymbol{\mu}^{'} + u_i^{'}\boldsymbol{J} + \boldsymbol{v}_i^{'}$, a simple variable transformation gives:

$$\begin{cases} t_2^{'} = \dfrac{t_2 - (\mu + u)}{f(\eta + v)} \\ t_1^{'} = \dfrac{t_1 - (\mu + u)}{f(\eta + v)} \end{cases}$$

Given that $t_1 = 0 \; and \; t_2 = 1$ in $[\mathrm{HTM}]$ and $t_{1.}^{'} = \mu_1^{'} + u^{'}$ and $t_{2.}^{'} = \mu_2^{'} + u^{'} + v_2^{'}$ in $\left[\mathrm{HTM}^{'}\right]$, we obtain:

$$\begin{cases} f(\eta, v) = \dfrac{1}{\mu_2^{'} - \mu_1^{'} + v_2^{'}} = \dfrac{1}{\mu_2^{'} - \mu_1^{'}}\left(\dfrac{1}{1 + v_2^{'}/\mu_2^{'} - \mu_1^{'}}\right) \\ u = -\mu - f(\eta, v)\left(\mu_1^{'} + u^{'}\right) \end{cases}$$

Considering that the ratio $v_2^{'}/\mu_2^{'} - \mu_1^{'}$ is small in comparison to 1, we apply Taylor's expansion to the previous equation:

$$\Rightarrow \begin{cases} f(\eta, v) \approx \dfrac{1}{\mu_2^{'} - \mu_1^{'}} - \dfrac{v_2^{'}}{\left(\mu_2^{'} - \mu_1^{'}\right)^2} \\ u \approx -\mu - \dfrac{1}{\mu_2^{'} - \mu_1^{'}}\left(1 - \dfrac{v_2^{'}}{\mu_2^{'} - \mu_1^{'}}\right)\left(\mu_1^{'} + u^{'}\right) \end{cases}$$

Thus, In the case of a linear function f and given that $v_2^{'} = \left(\mu_2^{'} - \mu_1^{'}\right)v^{'}$, we obtain:

$$\Rightarrow \begin{cases} v \approx C_1 - \dfrac{v^{'}}{\mu_2^{'} - \mu_1^{'}} \\ u \approx C_2 + v^{'}\left(\dfrac{\mu_1^{'}}{\mu_2^{'} - \mu_1^{'}}\right) - u^{'}\left(\dfrac{1}{\mu_2^{'} - \mu_1^{'}}\right) + u^{'}v^{'}\left(\dfrac{1}{\mu_2^{'} - \mu_1^{'}}\right) \end{cases}$$

where $C_1 = \frac{1}{\mu_2^{'} - \mu_1^{'}} - \eta$ and $C_2 = -\mu - \frac{\mu_1^{'}}{\mu_2^{'} - \mu_1^{'}}$ are constant terms. Therefore, given that the variance of the product of two centered random variables X and Y is $\sigma^2_{X*Y} = \sigma^2_X \sigma^2_Y + \left(\sigma_{X,Y}\right)^2$ and that $\sigma_{XY,Y} = 0$, we obtain:

$$\Rightarrow \begin{cases} \sigma_v^2 = \frac{\sigma_{v'}^2}{\left(\mu_2^{'} - \mu_1^{'}\right)^2} \\ \sigma_u^2 = \frac{1}{\left(\mu_2^{'} - \mu_1^{'}\right)^2}\left(\left(\mu_1^{'}\right)^2 \sigma_{v'}^2 + \sigma_{u'}^2 + \sigma_{v'}^2 \sigma_{u'}^2 + \left(\sigma_{v',u'}\right)^2 - 2\mu_1^{'}\sigma_{v',u'}\right) \end{cases}$$

The correlation between v and u is

$$\rho_{uv} = \frac{\sigma_{uv}}{\sigma_u \sigma_v}$$

with $\sigma_{v,u} = \operatorname{cov}\left(-\frac{v'}{\mu_2^{'} - \mu_1^{'}}, C_2 + v'\left(\frac{\mu_1^{'}}{\mu_2^{'} - \mu_1^{'}}\right) - u'\left(\frac{1}{\mu_2^{'} - \mu_1^{'}}\right) + u'v'\left(\frac{1}{\mu_2^{'} - \mu_1^{'}}\right)\right)$

$$= \frac{1}{\left(\mu_2^{'} - \mu_1^{'}\right)^2}\left(\sigma_{v',u'} - \mu_1^{'}\sigma_{v'}^2\right)$$

Appendix 2: [MTM] and [ITM]

In this section we describe how we calculated the genetic variance ($u_2^{"*}$) in the [MTM] model based on the genetic variances ($u_1^{"}, u_2^{"}$) of the [ITM] model.
The probability that the trait be of class 1 (or 2, 3, etc.) is the same for both models, thus:

$$\begin{cases} P(y=1)_{ITM} = P(y=1)_{MTM} \\ P(y=2)_{ITM} = P(y=2)_{MTM} \end{cases} \Rightarrow \begin{cases} L\left(t_1^{"}\right) = \mathrm{L}\left(t_1^{"*}\right) \\ L\left(t_2^{"}\right) - \mathrm{L}\left(t_1^{"}\right) = \mathrm{L}\left(t_2^{"*}\right) \end{cases} \Rightarrow \begin{cases} \dfrac{1}{1+e^{-t_1^{"}}} = \dfrac{1}{1+e^{-t_1^{"*}}} \\ \dfrac{1}{1+e^{-t_2^{"}}} - \dfrac{1}{1+e^{-t_1^{"}}} = \dfrac{1}{1+e^{-t_2^{"*}}} \end{cases}$$

where L is the Logit distribution function.

$$\Rightarrow \begin{cases} e^{-t_1^{"*}} = e^{-t_1^{"}} \\ e^{-t_2^{"*}} = \dfrac{1}{\dfrac{1+2e^{-t_2^{"}}+e^{-(t_2^{"}+t_1^{"})}}{e^{-t_2^{"}}-e^{-t_1^{"}}}} \end{cases} \Rightarrow \begin{cases} t_1^{"*} = t_1^{"} \\ t_2^{"*} = \log\left(\dfrac{1+2e^{-t_2^{"}}+e^{-(t_2^{"}+t_1^{"})}}{e^{-t_2^{"}}-e^{-t_1^{"}}}\right) \end{cases}$$

Thus, the variance $\sigma^2_{u_2^{"*}}$ is: $\sigma^2_{u_2^{"*}} = \sigma^2 \log\left(\dfrac{1+2e^{-t_2^{"}}+e^{-(t_2^{"}+t_1^{"})}}{e^{-t_2^{"}}-e^{-t_1^{"}}}\right)$

We well see that, from the parameters of the [ITM] model we can deduct the parameters of [MTM] model.

1. Bodin L, Bolet G, Garcia M, Garreau H, Larzul C, David I: **Robustesse et canalisation : vision de généticiens**. In Robustesse, rusticité, flexibilité, plasticité, résilience… les nouveaux critères de qualité des animaux et des systèmes d'élevage. Edited by Sauvant D, Perez JM: Dossier Inra Prod Anim 2010, **23**:11-22

2. Knap PW: **Breeding robust pigs**. Aust J Exp Agric 2005, **45**:763–773.

3. Lynch M, Gabriel W: **Environmental tolerance**. Am Nat 1987, **129**:283-303.

4. Gillespie JH, Turelli M: **Genotype-Environment Interactions and the Maintenance of Polygenic Variation**. Genetics 1989, **121**:129–138.

5. Falconer, D.S., Mackay, T.F.C: *Introduction to quantitative genetics*. Longmans Green, Harlow, Essex, UK. 1996.

6. Scheiner SM, Lyman RF: **The genetics of phenotypic plasticity .1. Heritability**. J Evol Biol 1989, **2**:95-107.

7. SanCristobal-Gaudy M, Elsen J-M, Bodin L, Chevalet C: **Prediction of the response to a selection for canalisation of a continuous trait in animal breeding**. *Genet Sel Evol GSE* 1998, **30**:423–451.

8. Garcia, M., David, I., Garreau, H., Ibañez-Escriche, N., Mallard, J., Masson, J.P., Pommeret, D., Robert-Granié, C., Bodin, L.: **Comparisons of three models for canalising selection or genetic robustness. 60th Annual Meeting of the European Association for Animal Production, Barcelona, Spain.** 2009.

9. Hill WG, Zhang X-S: **Effects on phenotypic variability of directional selection arising through genetic differences in residual variability**. *Genet Res* 2004, **83**:121–132.

10. Mulder HA, Bijma P, Hill WG: **Prediction of Breeding Values and Selection Responses With Genetic Heterogeneity of Environmental Variance**. *Genetics* 2007, **175**:1895–1910.

11. SanCristobal-Gaudy M, Bodin L, Elsen JM, Chevalet C: **Genetic components of litter size variability in sheep**. *Genet Sel Evol GSE* 2001, **33**:249–271.

12. Rönnegård L, Felleki M, Fikse F, Mulder HA, Strandberg E: **Genetic heterogeneity of residual variance - estimation of variance components using double hierarchical generalized linear models**. *Genet Sel Evol* 2010, **42**:8.

13. Sorensen DA, Andersen S, Gianola D, Korsgaard I: **Bayesian inference in threshold models using Gibbs sampling**. *Genet Sel Evol* 1995, **27**:229.

14. Ibáñez-Escriche N, Garcia M, Sorensen D: **GSEVM v.2: MCMC software to analyze genetically structured environmental variance models**. *J Anim Breed Genet Z Für Tierz Zücht* 2010, **127**:249–251.

15. Perez-Cabal M., De los Campos G, Vazquez A.: **Genetic evaluation of susceptibility to clinical mastitis in Spanish Holstein cows**. 2009.

16. Vazquez A., Gianola D, Bates D: **Assessment of Poisson, logit, and linear models for genetic analysis of clinical mastitis in Norwegian Red cows**. 2009.

17. Garcia M, Ibáñez-Escriche: **GSEVM v.3 : MCMC software to analyse genetically structred environmental variance models**. .

18. Gianola D: **Theory and Analysis of Threshold Characters**. *J Anim Sci* 1982, **54**:1079–1096.

19. Misztal I, Gianola D, Foulley JL: **Computing Aspects of a Nonlinear Method of Sire Evaluation for Categorical Data**. *J Dairy Sci* 1989, **72**:1557–1568.

20. Gilmour AR, Gogel BJ, Cullis BR and Thompson R: **ASReml User Guide Release 3.0. In VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.** 2009.

21. Schwarz G: **Estimating the Dimension of a Model**. *Ann Stat* 1978, **6**:461–464.

22. Foulley JL, Gianola D: **Statistical analysis of ordered categorical data via a structural heteroskedastic threshold model**. *Genet Sel Evol* 1996, **28**:1–25.

23. Hazel LN: **The Genetic Basis for Constructing Selection Indexes**. *Genetics* 1943, **28**:476–490.

24. Amer P.R, L Bodin B: **Quantitative genetic selection for twinning rate in ewes**. *Proc N Z Soc Anim Prod* 2006, **66**:429–433.

25. Matos CA, Thomas DL, Gianola D, Tempelman RJ, Young LD: **Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. Estimation of genetic parameters**. *J Anim Sci* 1997, **75**:76–87.

26. Janssens S, Vandepitte W, Bodin L: **Genetic parameters for litter size in sheep: natural versus hormone-induced oestrus**. *Genet Sel Evol GSE* 2004, **36**:543–562.

27. Matika O, van Wyk JB, Erasmus GJ, Baker RL: **Genetic parameter estimates in Sabi sheep**. *Livest Prod Sci* 2003, **79**:17–28.

28. Garreau H, Bolet G, Larzul C, Robert-Granié C, Saleil G, SanCristobal M, Bodin L: **Results of four generations of a canalising selection for rabbit birth weight**. *Livest Sci* 2008, **119**:55–62.

29. Gutiérrez JP, Nieto B, Piqueras P, Ibáñez N, Salgado C: **Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice**. *Genet Sel Evol* 2006, **38**:445.

30. Ros M, Sorensen D, Waagepetersen R, Dupont-Nivet M, SanCristobal M, Bonnet J-C, Mallard J: **Evidence for genetic control of adult weight plasticity in the snail Helix aspersa**. *Genetics* 2004, **168**:2089–2097.

31. Rowe SJ, White IM, Avendaño S, Hill WG: **Genetic heterogeneity of residual variance in broiler chickens**. *Genet Sel Evol GSE* 2006, **38**:617–635.

32. Sorensen D, Waagepetersen R: **Normal linear models with genetically structured residual variance heterogeneity: a case study**. *Genet Res* 2003, **82**:207–222.

33. Ibáñez-Escriche N, Moreno A, Nieto B, Piqueras P, Salgado C, Gutiérrez JP: **Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response**. *Genet Sel Evol GSE* 2008, **40**:279–293.

34. Yang Y, Christensen OF, Sorensen D: **Analysis of a genetically structured variance heterogeneity model using the Box-Cox transformation**. *Genet Res* 2011, **93**:33–46.

35. Martin P, Raoul J, Bodin L: **Effects of the FecL major gene in the Lacaune meat sheep population**. *Genet Sel Evol* 2014, **46**:48.

36. Bodin L, Elsen JM: **Variability of litter size of french sheep breeds following natural or induced ovulation**. *Anim Sci* 1989, **48**:535–541.

2.3.3 Résultats complémentaires

L'article du paragraphe 2.2.2 ne présentait que les résultats obtenus en race Romane. Une étude similaire a été réalisée chez les Suffolk et les Rouge de l'Ouest. Il n'a pas été possible de faire converger le modèle à seuils indépendants en race Suffolk. Les résultats obtenus pour ces deux races sont très proches de ceux de la Romane. Le modèle à seuil homothétique est celui qui estimait aussi bien la moyenne mais beaucoup mieux la variabilité des observations que le modèle homoscédastique (tableau 11). Nous remarquons aussi qu'il y n'a pas une grande différence entre les estimées par le modèle à seuils homothétiques et le modèle à seuils indépendants pour la race Rouge de l'Ouest. Ceci reste en cohérence avec les résultats trouvés sur la race Romane.

Tableau 11 : Corrélation calculée pour les diffèrent modèles ([TM], [HTM], [ITM]) entre la moyenne (variance) estimée et la moyenne (variance) observée.

Modèle / corrélation	Corrélation moyenne estimée - moyenne observée	Corrélation variance estimée-variance observée
Modèle de référence: modèle à seuils homoscédastique ([TM])	**0.83*** (Suffolk) **0.78***(Rouge de l'Ouest)	**0.24***(Suffolk) **0.16***(Rouge de l'ouest)
Modèle à seuils homothétiques ([HTM])	**0.84*** (Suffolk) **0.77***(Rouge de l'Ouest)	**0.38***(Suffolk) **0.50*** (Rouge de l'Ouest)
Modèle à seuils indépendants ([ITM])	Non convergence (Suffolk) **0.76 ***(Rouge de l'Ouest)	Non convergence (Suffolk) **0.46*** (Rouge de l'Ouest)

* la valeur de P-valeur est strictement inférieure à 0.001.

Le tableau 12 présente les variances génétiques sur la moyenne et la variance sur l'échelle sous-jacente calculée avec le modèle à seuil homothétique. Nous constatons que les variances génétiques sur la moyenne de la sous-jacente restent en cohérence avec les variances génétiques estimées par le modèle à seuils homoscédastique. Les variances génétiques calculées sur la variance résiduelle de la sous-jacente sont faibles.

Tableau 12 : Variance génétique (calculée) sur la moyenne*u* et la variance de la sous-jacente *v* pour la Rouge de l'Ouest et pour la Suffolk.

Races	σ_u^2	σ_v^2
Rouge de l'ouest	0.10	0.022
Suffolk	0.079	0.008

Le tableau 13 présente la corrélation génétique entre *u* et *v* estimée par le modèle à seuils homothétiques. Les corrélations calculées sont homogènes, elles sont négatives pour l'ensemble des races étudiées.

Tableau 13 : Corrélation ρ entre les *u* et *v* estimées par le modèle à seuils homothétique [HTM].

	Rouge de l'Ouest	Suffolk
ρ	-0.25	-0.47

Les figures 14 et 15 présentent les relations entre les valeurs génétiques pour la moyenne et la variance pour les trois modèles sur les deux races (Rouge de l'Ouest, Suffolk) ; nous remarquons que la liaison obtenue par les modèles [HTM] et [ITM] est moins forte que celle obtenue par le modèle [TM]. La dispersion de la variance pour chaque moyenne est bien plus faible en race Suffolk qu'en Rouge de l'Ouest.

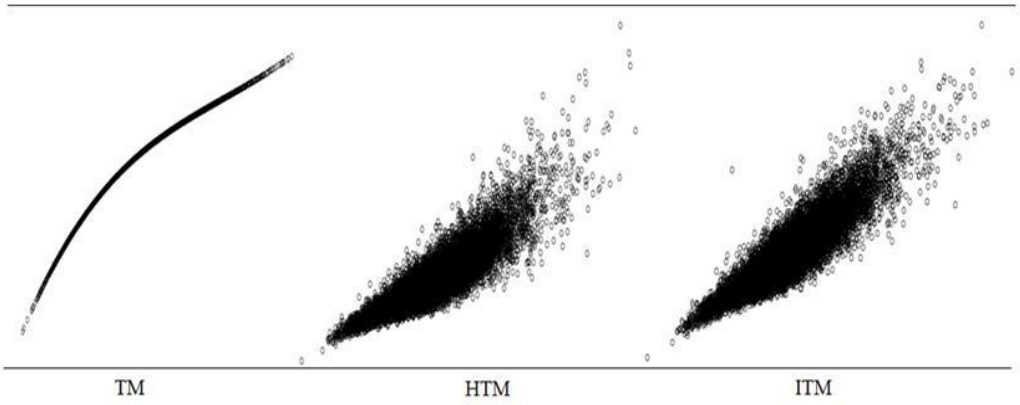


Figure 14 : Relation moyenne-variance pour les trois modèles ([TM], [HTM], [ITM]) sur la race Rouge de l'Ouest (Prolif=1.89).

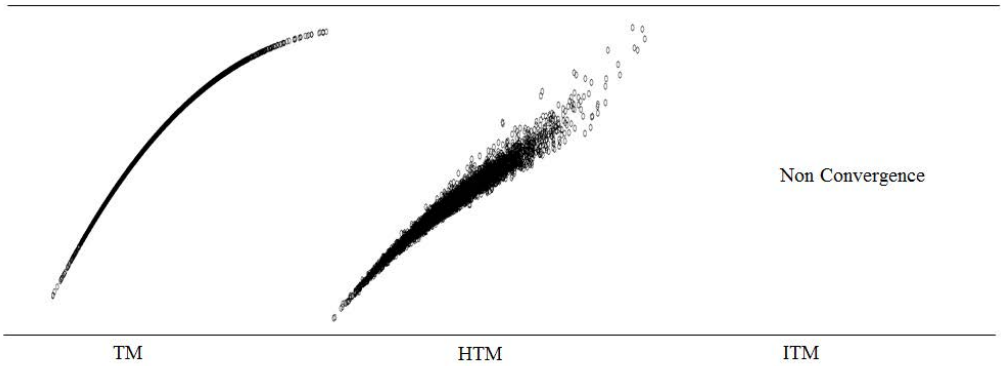


Figure 15 : Relation moyenne-variance pour les trois modèles ([TM], [HTM], [ITM]) sur la race Suffolk (Prolif=1.63).

2.4 Discussion

Nous avons proposé dans cette étude deux modèles qui permettent de modéliser la variabilité des tailles de portée e en s'appuyant sur le modèle à seuils : le modèle à seuils homothétiques qui permet de modéliser la moyenne et la variance de la sous-jacente séparément de la même manière que le modèle de canalisation « classique » et le modèle à seuils indépendants qui est l'équivalent d'un modèle multicaractère et qui modélise chaque seuil ou chaque classe de taille de portée. Les principaux avantages du modèle à seuils indépendants sont qu'il permet d'associer des pondérations économiques à chacune des classes pour construire un index de sélection et qu'il peut s'adapter à n'importe quelle distribution de tailles de portée sans imposer de contraintes particulières. En contrepartie il est difficile à justifier du point de vue biologique : comment expliquer en effet l'existence de facteurs de variation génétiques et environnementaux différents pour chaque taille de portée. Par ailleurs, lorsqu'il n'y a que 3 classes (2 seuils), tous ces modèles sont équivalents et ce modèle n'aurait donc d'intérêt que pour des populations pour lesquelles il y a suffisamment de données pour bien estimer les facteurs de variation de la classe 4 (ou du seuil 3). Le modèle homothétique à l'avantage d'avoir moins de paramètres à estimer que le modèle à seuil indépendant (6 contre 9 pour 4 classes) ; sa convergence est de ce fait plus rapide à obtenir et surtout nécessite moins de données et est moins sensible à leur structure (Rappelons que nous n'avons pas réussi à faire converger le modèle à seuils indépendants sur les données Suffolk). L'autre avantage de ce modèle est qu'il correspondbien au modèle de canalisation sur l'échelle sous-jacente en séparant des effets qui affectent le niveau et la variabilité du caractère. Néanmoins, l'interprétation des paramètres sur l'échelle observée n'est pas si aisée.
Nous avons présenté la théorie de ces deux modèles dans l'article 2. Néanmoins lors de leur mise en application ; nous nous sommes heurtés à un problème :

- Nous n'avons pas introduit les effets fixes sur la fonction f de la sous-jacente du modèle à seuil homothétique sous ASReml. Pour pallier ce problème, nous proposons l'algorithme itératif suivant :

  1) Faire passer le modèle à seuils homothétiques (le modèle [HTM*]) en introduisant des interactions entre les facteurs environnementaux qui doivent jouer sur la variance résiduelle et les seuils.
2) Récupérer les écarts entre les seuils
3) Créer les contrastes entre effets fixes et effets génétique pour maintenir l'écart
4) Faire repasser le modèle avec les contraintes.

Nous n'avons pas eu le temps au cours de ces 3 années de thèse de programmer et tester cet algorithme. Malgré cette difficulté qui reste à surmonter nous pensons que le modèle homothétique est le plus adapté à l'analyse de la variabilité de la taille de portée des ovins.

Cette étude a permis de montrer l'existence d'une hétérogénéité d'origine génétique de la variabilité des tailles de portée dans trois races dont les prolificités couvrent le domaine de variation où la canalisation serait intéressante. Cette hétérogénéité a pu être mise en évidence en considérant la taille de portée comme une variable continue, mais dans ce cas, la relation entre les valeurs génétiques pour la moyenne et pour la variance sont très variables et dépendent de la race ou de sa prolificité. Cette procédure ne permet pas de séparer la liaison structurelle qui existe entre moyenne et variance de la liaison génétique qui nous intéresse. La prise en compte du caractère discret de la taille de portée en utilisant des modèles à seuils permet de s'affranchir de cette "liaison structurelle" dans ce cas la corrélation génétique entre les valeurs pour la moyenne et pour la variance sont relativement homogènes pour les 3 races. Elles sont négatives ce qui montre qu'intra race, un gain de prolificité peut s'accompagner d'une diminution de la variabilité sur l'échelle sous-jacente. Lors de la transformation des valeurs sur l'échelle observée, la corrélation négative sur l'échelle sous jacente ne

contre-balance pas la très forte liaison structurelle existante, de sorte que sur cette échelle la relation entre moyenne et variance est positive.
Si la corrélation génétique est du même ordre de grandeur pour les 3 races, la dispersion de la variance pour chaque moyenne est bien plus faible en race Suffolk et en race Rouge de l'Ouest qu'en race Romane (figure 16). Cette différence de dispersion est évidemment liée à la moyenne pour de faible moyenne. En effet plus la moyenne est faible plus le pourcentage de portées triples est faible et plus la dispersion est faible pour une même moyenne. Nous montrons toutefois que pour des prolificités supérieures à 1.5 il est possible de séparer efficacement les deux composantes moyenne et variance et qu'il sera donc possible de canaliser les tailles de portées.

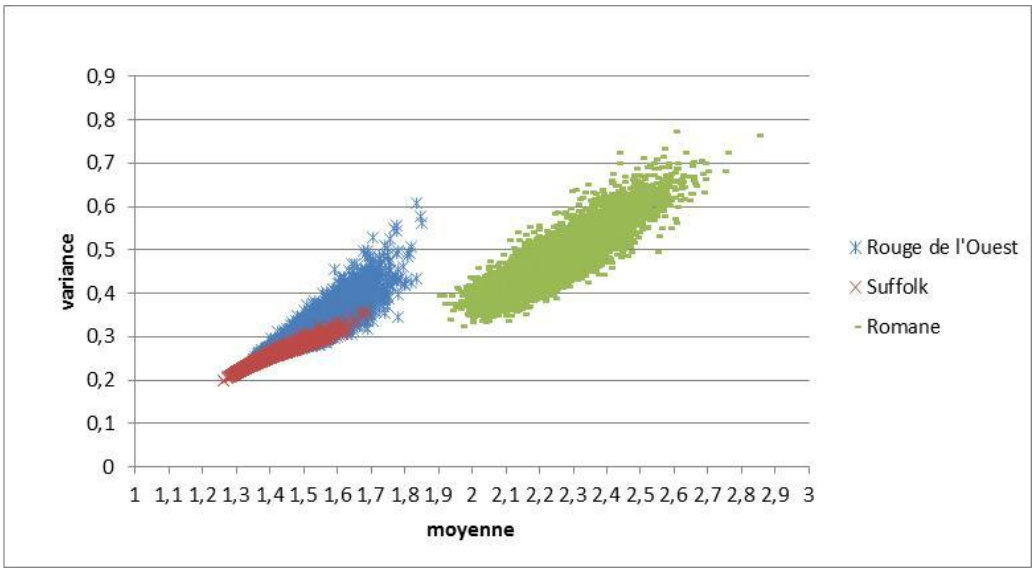


Figure 16 : Relation moyenne-variance pour les trois races estimée par le modèle à seuils homothétiques ([HTM]).

RESUME

Ce travail de recherche a été mené dans l'équipe MODGEN du département **GenPhySE**, 1388 UMR de l'INRA. Il avait pour perspectives de modéliser les variabilités de la taille de portée des ovins pour une meilleure évaluation génétique et de proposer des modèles d'analyse adoptés à ce type de caractère.Les objectifs de l'étude étaient d'identifier, et d'analyser les facteurs environnementaux et génétiques des variabilités des tailles de portée des brebis et de proposer une méthode adaptée pour canaliser les caractères discretsqui prenne en compte conjointement des objectifs sur la moyenne et sur la variabilité du caractère. La thèse s'appuyait sur des données issues de la chaîne nationale de contrôle de performances en relation avec des organismes de sélection ovin qui sont particulièrement intéressés par les résultats de ces recherches (CORAM).

Choix des races Nous avons considéré trois races dont les fréquences de la taille de portée sont plus ou moins élevées. La race Romane connue par sa forte prolificité ainsi des effectifs contrôlés importants a permis des analyses précises. Les deux autres races choisies sont la race Suffolk et la race Rouge de l'Ouest qui ont servi à tester les modèles sur des populations comportant de moins en moins de portées multiples.

La taille de portée des ovins considérée comme une variable continue Au cours de l'étude, nous avons identifié et hiérarchisé les facteurs de variation environnementaux du niveau de prolificité et de de sa variance pour trois races ovines (Suffolk, Rouge de l'Ouest et Romane). Si les facteurs de variation environnementaux du niveau de prolificitéétaient déjà bien connus, peu de données concernaient les facteurs de variation environnementaux de la variance.
Nous avons également estimé les paramètres génétiques du niveau de prolificité et de sa variance, ils sont cohérents entre eux. Les héritabilités estimées sur la moyenne sont faibles, en cohérences entre les trois races et avec des études antérieures. La non homogénéité des corrélations estimées par le modèle à deux étapes peut s'expliquer par la relation structurelle entre la moyenne et la variance sur l'échelle observée qui n'est pas prise en compte dans ce type d'analyse. A l'échelle d'une race, cette relation dépend de l'écart spécifique entre les seuils; elle est différente de la relation moyenne – variance observée pour l'ensemble des races et qui est calculée sur des données brutes (figure 3 ; page 17).

La taille de portée des ovins considérée comme une variable discrète Nous avons montré l'existence d'une variabilité d'origine génétique dans les trois races étudiées et nous avons montré que le modèle à seuils homothétique [HTM] s'ajustait mieux aux données et qu'il permettait de calculer conjointement l'effet génétique sur la moyenne et sur la variance. Nous avons aussi montré les limites de ce modèle et nous avons proposé une solution afin de donner un cadre plus général à son utilisation et l'utiliser dans les évaluations génétiques. Nous avons également identifié un critère de sélection pertinent qui permet de sélectionner les individus de même moyenne ayant des variances différentes.
Nous avons proposé le modèle à seuils indépendants [ITM] ou le modèle multi caractères avec contrainte, ce modèle considère que chaque réalisation est un caractère. Nous avons montré la faisabilité de cemodèle sur les trois races. Ceci nous a permis de mettre en évidence les problèmes liés à la convergence du programme utilisé, et nous avons proposé d'implémenter ce modèle en modifiant le programme TM.

REFERENCES

Amer, P.R., Bodin, L., (2006). Quantitative genetic selection for twinning rate in ewes. Proceedings of the New Zealand Society of Animal Production, Napier, NZ, pp. 429-433.

Anderson, D.A., and Aitkin, M. (1985). Variance Component Models with Binary Response: Interviewer Variability. J. R. Stat. Soc. Ser. B Methodol. ***47***, 203–210.

Argente, M.J, Garcia, M.L, and Muelas, R (2011). Resultados de las cuatro primeras generaciones de seleccion divergente por variabilidad ambiental del tamano de camada en conejo.

Atti, N., Theriez, M., Abdennebi, L., (2001). Relationship between ewe body condition at mating and reproductive performance in the fat-tailed Barbarine breed. Anim. Res. **50**, 135-144.

Baelden, M., Tiphine, L., Poivey, J.P., Bouix, J., Bibé, B., Robert-Granié, C., and Bodin, L. (2005). Estimation of genetic parameters for litter size after natural and hormone-induced oestrus in sheep. Livest. Prod. Sci. ***97***, 275–281.

Baril, G., Chemineau, P., Cognie, Y., Guerin, Y., Leboeuf, B., Orgeur, P., Vallet, J.C. (Eds.), (1993). Manuel de formation pour l'insémination artificielle chez les ovins et les caprins. FAO, Rome, Italy.

Bodin, L., and Elsen, J.M. (1989). Variability of litter size of french sheep breeds following natural or induced ovulation. Anim. Sci. ***48***, 535–541.

Bodin, L., Elsen, J.M., Hanocq, E., Francois, D., Lajous, D., Manfredi, E., Mialon, M.M., Boichard, D., Foulley, J.L., SanCristobal-Gaudy, M., Teyssier, J., Thimonier, J., Chemineau, P., (1999). Genetique de la reproduction chez les ruminants - —Genetics of reproduction in ruminants". Productions Animales **12**, 87-100.

Bodin, L., Benoit, M., SanCristobal, M., Chevalet, C., Elsen, J.M., 2000. Selection for prolificacy: new prospects for an ever-interesting objective. Options Méditérranéennes. Série A, Séminaires Méditérranéens, pp. 117-120.

Bodin, L., SanCristobal, M., Lecerf, F., Mulsant, P., Bibé, B., Lajous, D., Belloc, J.-P., Eychenne, F., Amigues, Y., and Elsen, J.-M. (2002). Segregation of a major gene influencing ovulation in progeny of Lacaune meat sheep. Genet. Sel. Evol. GSE ***34***, 447–464.

Bodin, L., Bolet, G., Garcia, M., Garreau, H., Larzul, C., David, I., (2010). Robustesse et canalisation : vision de généticiens. In: Sauvant, D., Perez, J.M. (Eds.), Robustesse, rusticité, flexibilité, plasticité, résilience… les nouveaux critères de qualité des animaux et des systèmes d'élevage. Dossier Inra Prod. Anim. **23**, pp. 11-22.

Bodin, L., RAOUL, J., DROUILHE, L., MULSAN, P., and SARRY, J. (2011). Etat des lieux et gestion prat ique des gènes d'ovulation détectés dans les races ovines françaises.

Bolet, G., Esparbié, J., Falières, J., (1996). Relations entre le nombre de foetus par corne utérine, la taille de portée à la naissance et la croissance pondérale des lapereaux. Annales de Zootechnie **45**, 185-200.

Bolet, G., Garreau, H., Joly, T., Theau-Clement, M., Hurtaud, J., Bodin, L., (2005). Genetic homogenization of birth weight in rabbits. Evolution of the characteristics of the genital tract after two generations of selection. 56th Annual Meeting EAAP, Uppsala, Norway.

Bradford, G.E., (1985). Selection for litter size. In: Land, R.B., Robinson, D.W. (Eds.), Genetics of Reproduction In Sheep. Butterworths, London, pp. 3-18.

Bunge, R., Thomas, D.L., and Stookey, J.M. (1990). Factors affecting productivity of Rambouillet ewes mated to ram lambs. J. Anim. Sci.***68***, 2253–2262.

Callahan, H.S., Pigliucci, M., Schlichting, C.D., (1997). Developmental phenotypic plasticity: where ecology and evolution meet molecular biology. Bioessays **19**, 519-525.

Canario, L., Lundgren, H., Haandlykken, M., and Rydhmer, L. (2010). Génétique de la croissance et de l'hétérogénéité des poids de porcelets dans les portées de Landrace norvégien. Journ. Rech. Porc. ***42***, 181–182.

Cardin, S., and Minvielle, F. (1986). Selection on phenotypic variation of pupa weight in Tribolium castaneum. Can. J. Genet. Cytol. ***28***, 856–861.

Cervantes, I., Nieto, B., Salgado, C., Pérez-Cabal, M.A., Ibanez-Escriche, N., Gutierrez, J.P., (2010). Resultados de las dos primeras generaciones de selección divergente para variabilidad del peso al nacimiento en ratones., XV Reunión Nacional de Mejora Genética Animal, Vigo, Spain.

Clogg, C.C., and Shihadeh, E.S. (1994). Statistical models for ordinal variables (Sage Publications).

Damgaard, L.H., Rydhmer, L., Løvendahl, P., and Grandinson, K. (2003). Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. J. Anim. Sci. ***81***, 604–610.

David, I., Robert-Granié, C., Manfredi, E., Lagriffoul, G., and Bodin, L. (2008). Environmental and genetic variation factors of artificial insemination success in French dairy sheep. Anim. Int. J. Anim. Biosci. ***2***, 979–986.

David, I., Bodin, L., Gianola, D., Legarra, A., Manfredi, E., and Robert-Granié, C. (2009). Product versus additive threshold models for analysis of reproduction outcomes in animal genetics. J. Anim. Sci. ***87***, 2510–2518.

Debat, V., and David, P. (2001). Mapping phenotypes: canalization, plasticity and developmental stability. Trends Ecol. Evol. ***16***, 555–561.

Dedieu, B., and Ingrand, S. (2010). Incertitude et adaptation: cadres theoriques et application a l'analyse de la dynamique des systemes d'elevage. INRA Prod Anim ***23***, 81–90.

Delpuech, J.-M., Moreteau, B., Chiche, J., Pla, E., Vouidibio, J., and David, J.R. (1995). Phenotypic Plasticity and Reaction Norms in Temperate and Tropical Populations of Drosophila melanogaster: Ovarian Size and Developmental Temperature. Evolution ***49***, 670–675.

Demers, V., (2010). Impact de la prolificité sur la rentabilité de l'entreprise ovine québécoise : approche par modélisation. Mémoire de Msc. Dépatement des sciences animales. Université de Laval, Laval, Quebec, p. 125.

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. J. R. Stat. Soc. Ser. B Methodol. ***39***, 1–38.

Desvignes, A., Darpoux, R., (1964). Valeurs d'élevages des brebis utilisées pour le traitement industriel avec différentes races de béliers. , 195, 29-. Bull. Techn. Inf. **195**, 893-301.

Dickerson, G.E., and Glimp, H.A. (1975). Breed and Age Effects on Lamb Production of Ewes. J. Anim. Sci. ***40***, 397–408.

Dieny, M. (2011). Elaboration d'indices de synthèse en réponse à un objectif de sélection sur bases économiques en race ovin Blanche du Massif Central. Mémoire de fin d'études pour l'obtention du diplôme d'ingénieur agronome (VetAgro Sup, Lempdes).

Djemali, M., Berger, P.J., and Freeman, A.E. (1987). Ordered categorical sire evaluation for dystocia in Holsteins. J. Dairy Sci. ***70***, 2374–2384.

Ducrocq, V. 2000. Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects. Interbull Bulletin 25,123- 130.

Dupont-Nivet, M., Robert-Granie, C., Tiquet, F., Le Guillou, S., and Quillet, E. (2007). What clones tell us about genetic determinism of phenotypic plasticity in rainbow trout (Oncorhynchus mykiss). Aquaculture *272*, S253.

Dupont-Nivet, M., Robert-Granié, C., Le Guillou, S., Tiquet, F., and Quillet, E. (2012). Comparison of isogenic lines provides evidence that phenotypic plasticity is under genetic control in rainbow trout Oncorhynchus mykiss. J. Fish Biol. ***81***, 1754–1762.

Dwyer, C.M., Lawrence, A.B., Bishop, S.C., and Lewis, M. (2003). Ewe–lamb bonding behaviours at birth are affected by maternal undernutrition in pregnancy. Br. J. Nutr. ***89***, 123–136.

Elsen, J.M., Bodin, L., Bibe, B., Poivey, J.P., Tchamitchian, L., (1984). Utilisation des techniques modernes de reproduction pour l'amelioration genetique. 9emes Journees de la Recherche Ovine et Caprine, pp. 361-398.

Elsen, J.M., Bodin, L., Francois, D., Poivey, J.P., Teyssier, J., (1994). Genetic improvement of litter size in sheep. Proceedings, 5th World Congress on Genetics Applied to Livestock Production, University of Guelph, Guelph, Ontario, Canada, 7-12 August 1994. Volume 19. Selection and quantitative genetics; growth; reproduction; lactation; fish; fiber; meat., pp. 237-244.

Engel, B., and Keen, A. (1994). A simple approach for the analysis of generalizea linear mixed models. Stat. Neerlandica ***48***, 1–22.

Everett-Hincks, J.M., and Cullen, N.G. (2009). Genetic parameters for ewe rearing performance. J. Anim. Sci. ***87***, 2753–2758.

Fahmy, M.H., (1996). Prolific sheep. Cab International, Wallingford, UK.

Fahmy, M.H. (1990). Development of dls breed of sheep: genetic and phenotypic parameters of date of lambing and litter size. Can. J. Anim. Sci. ***70***, 771–778.

Fahrmeir, L., and Tutz, G. (2001). Multivariate Statistical Modelling Based on Generalized Linear Models (New York, NY: Springer New York).

Falconer, D.S., and Mackay, T.F.C. (1996). Introduction to Quantitative Genetics (Essex, England: Benjamin Cummings).

Falconer D.S. (1974). Introduction à la génétique quantitative (Paris: Masson).

Faugere, O., DOCKES, A.., and PERROT, C. (1990). L'élevage traditionnel des petits ruminants au Sénégal. 1. Pratiques de conduite et d'exploitation des animaux chez les éleveurs de la région de Kolda. 249–259.

Fina, M., Ibáñez-Escriche, N., Piedrafita, J., and Casellas, J. (2013). Canalization analysis of birth weight in Bruna dels Pirineus beef cattle. J. Anim. Sci. ***91***, 3070–3078.

Foulley, J., and Im, S. (1993). A marginal quasi-likelihood approach to the analysis of Poisson variables with generalized linear mixed models. Genet. Sel. Evol. GSE ***25***, 101–107.

Foulley, J.L., and Gianola, D. (1996). Statistical analysis of ordered categorical data via a structural heteroskedastic threshold model. Genet. Sel. Evol. ***28***, 1–25.

Foulley, J.-L., and Jaffrézic, F. (2010). Modelling and estimating heterogeneous variances in threshold models for ordinal discrete data via Winbugs/Openbugs. Comput. Methods Programs Biomed. ***97***, 19–27.

Foulley, J.L., and Quaas, R.L. (1995). Heterogeneous variances in Gaussian linear mixed models. Genet. Sel. Evol. ***27***, 1–18.

Friggens, N.C., Sauvant, D., Martin, O., (2010). Vers des définitions opérationnelles de la robustesse s'appuyant sur des faits biologiques: l'exemple de la nutrition. In: Sauvant, D., Perez, J.M. (Eds.), Robustesse, rusticité, flexibilité, plasticité, résilience… les nouveaux critères de qualité des animaux et des systèmes d'élevage. Dossier Inra Prod. Anim. **23**, pp. 43-52.

Gama, L.T., Dickerson, G.E., Young, L.D., and Leymaster, K.A. (1991). Effects of breed, heterosis, age of dam, litter size, and birth weight on lamb mortality. J. Anim. Sci. ***69***, 2727–2743.

García-Ispierto, I., López-Gatius, F., Santolaria, P., Yániz, J.L., Nogareda, C., and López-Béjar, M. (2007). Factors affecting the fertility of high producing dairy herds in northeastern Spain. Theriogenology ***67***, 632–638.

Garcia M, and Ibáñez-Escriche GSEVM v.3 : MCMC software to analyse genetically structred environmental variance models.

Garreau, H., Bolet, G., Hurtaud, J., Larzul, C., Robert-Granié, C., Ros, M., Saleil, G., SanCristobal, M., Bodin, L., (2004). Homogeneizacion genetica de un caractrer. Resultados preliminares de una seleccion canalizante sobre el peso al nacimiento de los gazapos. 12ème Réunion Nacional Mejora Genetica Animal, Arucas, Las Palmas, 7-9 juillet 2004,.

Garreau, H., Bolet, G., Larzul, C., Robert-Granié, C., Saleil, G., SanCristobal, M., and Bodin, L. (2008). Results of four generations of a canalising selection for rabbit birth weight. Livest. Sci. ***119***, 55–62.

Garreau, H., Bolet, G., Larzul, C., Robert-Granié, C., Saleil, G., SanCristobal, M., and Bodin, L. (2008b). Results of four generations of a canalising selection for rabbit birth weight. Livest. Sci. *119*, 55–62.

Gavrilets, S., and Scheiner, S.M. (1993). The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. J. Evol. Biol. ***6***, 31–48.

Gianola, D. (1980). A method of sire evaluation for dichotomies. J. Anim. Sci. ***51***, 1266–1271.

Gianola, D. (1982). Theory and Analysis of Threshold Characters. J. Anim. Sci. ***54***, 1079–1096.

Gillespie, J.H., and Turelli, M. (1989). Genotype-Environment Interactions and the Maintenance of Polygenic Variation. Genetics ***121***, 129–138.

Gilmour, A.R., Anderson, R.D., and Rae, A.L. (1985). The Analysis of Binomial Data by a Generalized Linear Mixed Model. Biometrika ***72***, 593–599.

Gilmour AR, Gogel BJ, Cullis BR and Thompson R (2009). ASReml User Guide Release 3.0. In VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Gunn, R.G., Doney, J.M., and Smith, W.F. (1984). The effect of level of pre-mating nutrition on ovulation rate in Scottish Blackface ewes in different body conditions at mating. Anim. Sci. ***39***, 235–239.

Gutiérrez, J.P., Nieto, B., Piqueras, P., Ibáñez, N., and Salgado, C. (2006). Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. Genet. Sel. Evol. *38*, 445.

Gutiérrez, J.P., Varona, L., Pun, A., Morante, R., Burgos, A., Cervantes, I., and Pérez-Cabal, M.A. (2011). Genetic parameters for growth of fiber diameter in alpacas. J. Anim. Sci. ***89***, 2310–2315.

Hanrahan, J.P., (1986). Reproductive efficiency in sheep. Conference on: Exploiting new technologies in animal breeding. Genetic developments. Oxford University Press; Oxford; UK, pp. 59-70.

Hanrahan, J.. (1989). Altering reproductive rate in sheep: some genetic and non-genetic options. Reproduction, growth and nutrition in sheep. 45–55.

Henderson, C.R. (1963). Selection index and expected genetic advance. In Statistical Genetics and Plant Breeding. **982** 141–163.

Henderson, C.R. Sire evaluation and genetic trends.

Hill, W.G. (1984). On selection among groups with heterogeneous variance. Anim. Sci. ***39***, 473–477.

Höschele, I., Gianola, D., and Foulley, J.L. (1987). Estimation of variance components with quasi-continuous data using Bayesian methods. J. Anim. Breed. Genet. ***104***, 334–349.

Huby, M., Gogué, J., Maignel, L., and Bidanel, J.. (2003). Corrélations génétiques entre les caractéristiques numériques et pondérales de la portée, la variabilité du poids des porcelets et leur survie entre la naissance et le sevrage. Journ. Rech. Porc. ***35***, 293–300.

Ibáñez-Escriche, N., Moreno, A., Nieto, B., Piqueras, P., Salgado, C., and Gutiérrez, J.P. (2008a). Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response. Genet. Sel. Evol. GSE ***40***, 279–293.

Ibáñez-Escriche, N., Varona, L., Sorensen, D., and Noguera, J.L. (2008b). A study of heterogeneity of environmental variance for slaughter weight in pigs. Anim. Int. J. Anim. Biosci. ***2***, 19–26.

I. Olesen, M.S. (1995). Application of a multiple-trait animal model for genetic evaluation of maternal and lamb traits in Norwegian sheep. Anim. Sci. ***60***, 457–469.

Janssens, S., Vandepitte, W., and Bodin, L. (2004). Genetic parameters for litter size in sheep: natural versus hormone-induced oestrus. Genet. Sel. Evol. GSE ***36***, 543–562.

Jousseins, C., Trocmé, E., Delmas, D., Ducourtieux, C., Mouret, J.., Dugat, J.., Defrance, M.., Tyssandier, P., Joliet, V., Lanne, P., et al. (2012a). Les coûts de production en élevage ovin allaitant dans le Sud-ouest – Campagne 2010. Résea ux d'élevage pour le conseil et la prospective, collection référence. Inst. L'Élevage Paris, 16p.

Jurado, J.. (2007). Un gen de gran efecto para prolificidad en raza Rasa Aragonesa? In: Calcodo, J.H., Jaime, R.C.M., Torrens, M.J., Pueyo, I.C., Gorriz, A.L., Pascua, A.S. (Eds.), XXXVIII Jornadas de Estudio.

Jurado, J.J. (2008). Efecto fenotipico del alelo BMP15/Fecx R en la prolificidad de la población de CarnesOviaragon S.C.L.

Knap, P.W. (2005). Breeding robust pigs. Aust. J. Exp. Agric. ***45***, 763–773.

Komprej, A., Gorjanc, G., and Kompan, D. (2011). Lambing Interval in Jezersko-Solčava and Improved Jezersko-Solčava Breed. Agric. Conspec. Sci. ***76***, 353–356.

Lahoz, B., Alabart, J.L., Jurado, J.J., Calvo, J.H., Martínez-Royo, A., Fantova, E., and Folch, J. (2011). Effect of the FecX(R) polymorphism in the bone morphogenetic protein 15 gene on natural or equine chorionic gonadotropin-induced ovulation rate and litter size in Rasa Aragonesa ewes and implications for on-farm application. J. Anim. Sci. ***89***, 3522–3530.

Lajous, D. (1987). Mesure du taux d'ovulation et de la mortalité embryonnaire chez la brebis Romanov. Utilisation et intérêt de la cœlioscopie (Institut national polytechnique, Toulouse,France).

Larzul, C., Roy, P.l., Tribout, T., Gogue, J., SanCristobal, M., (2006). Canalizing selection on ultimate pH in pigs: consequences on meat quality. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte.

Lerner, I.M., (1954). Genetic Homeostasis. Oliver and Boyd, London.

Lynch, M., and Gabriel, W. (1987). Environmental Tolerance. Am. Nat. ***129***, 283–303.

Manfredi, E., Foulley, J.L., Cristobal, M.S., and Gillard, P. (1991). Genetic parameters for twinning in the Maine-Anjou breed. Genet. Sel. Evol. ***23***, 421.

Mansanet, C. (2013). thèse : Contrôle génétique et physiologique de la pro lificité en race ovine Lacaune : caractérisation de la mutation causale et rôle fonctionnel du gène FecL (Tours).

Martin, P., Raoul, J., and Bodin, L. (2014). Effects of the FecL major gene in the Lacaune meat sheep population. Genet. Sel. Evol. ***46***, 48.

Martínez-Álvaro, M., Ibáñez-Escriche, N., Argente, M.J., García, M.L., and Blasco, A. (2013). Responses of divergent selection for environmental variability in the litter size in rabbit. (Asociación Interprofesional para el Desarrollo Agrario), pp. 460–462.

Martinez-Royo, A., Jurado, J.J., Smulders, J.P., Martí, J.I., Alabart, J.L., Roche, A., Fantova, E., Bodin, L., Mulsant, P., Serrano, M., et al. (2008). A deletion in the bone morphogenetic protein 15 gene causes sterility and increased prolificacy in Rasa Aragonesa sheep. Anim. Genet. ***39***, 294–297.

Matika, O., van Wyk, J.B., Erasmus, G.J., and Baker, R.L. (2003). Genetic parameter estimates in Sabi sheep. Livest. Prod. Sci. ***79***, 17–28.

Matos, C.A., Thomas, D.L., Gianola, D., Tempelman, R.J., and Young, L.D. (1997). Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. Estimation of genetic parameters. J. Anim. Sci. ***75***, 76–87.

Matos C.A.P (1993). Genetic analysis of reproductive discrete traits in sheep using linear and nonlinear models.

McCullagh, P., and Nelder, J.A. (1989). Generalized Linear Models, Second Edition (Boca Raton: Chapman and Hall/CRC).

McCulloch, C.E. (1997). Maximum Likelihood Algorithms for Generalized Linear Mixed Models. J. Am. Stat. Assoc. ***92***, 162–170.

Meijering, A. (1984). Dystocia and stillbirth in cattle — A review of causes, relations and implications. Livest. Prod. Sci. ***11***, 143–177.

Mekkawy, W., Roehe, R., Lewis, R.M., Davies, M.H., Bünger, L., Simm, G., and Haresign, W. (2010). Comparison of repeatability and multiple trait threshold models for litter size in sheep using observed and simulated data in Bayesian analyses. J. Anim. Breed. Genet. Z. Für Tierz. Zücht. ***127***, 261–271.

Mulder, H.A., Bijma, P., and Hill, W.G. (2007). Prediction of Breeding Values and Selection Responses With Genetic Heterogeneity of Environmental Variance. Genetics ***175***, 1895–1910.

Mulder, H.A., Hill, W.G., Vereijken, A., and Veerkamp, R.F. (2009). Estimation of genetic variation in residual variance in female and male broiler chickens. Anim. Int. J. Anim. Biosci. ***3***, 1673–1680.

Nelder, J.A., and Wedderburn, R.W.M. (1972). Generalized Linear Models. J. R. Stat. Soc. Ser. Gen. ***135***, 370–384.

Neves, H.H., Carvalheiro, R., and Queiroz, S.A. (2012a). Genetic and environmental heterogeneity of residual variance of weight traits in Nellore beef cattle. Genet. Sel. Evol. *44*, 19.

Neves, H.H., Carvalheiro, R., and Queiroz, S.A. (2012b). Genetic and environmental heterogeneity of residual variance of weight traits in Nellore beef cattle. Genet. Sel. Evol. *44*, 19.

Notter, D.. (2000). Effects of ewe age and season of lambing on prolificacy in US Targhee, Suffolk, and Polypay sheep. Small Rumin. Res. ***38***, 1–7.

O'Ferrall, G.J.M. (1976). Phenotypic and genetic parameters of productivity in Galway ewes. Anim. Sci. ***23***, 295–304.

Olesen, I., Perez-Enciso, M., Gianola, D., and Thomas, D.L. (1994). A comparison of normal and nonnormal mixed models for number of lambs born in Norwegian sheep. J. Anim. Sci. ***72***, 1166–1173.

Perez-Enciso, M., Tempelman, R.J., and Gianola, D. (1993). A comparison between linear and poisson mixed models for litter size in iberian pigs. Livest. Prod. Sci. ***35***, 303–316.

Petersson, C.J., and Danell, Ö. (1985). Factors Influencing Lamb Survival in Four Swedish Sheep Breeds. Acta Agric. Scand. ***35***, 217–232.

Piñeiro, R. (1992). Selection for canalization at extra dorsocentral and scutellar bristles in Drosophila melanogaster. J. Hered. ***83***, 445–448.

Poivey, J.., Tiphine, L., Berny, F., and Jullien, E. (1995). Indexation blup modele animal chez les ovins allaitants. 2ème Rencontres Recherches Ruminants. 2ème Rencontres Rech. Rumin. Paris Fr. 453–456.

Pun, A., Cervantes, I., Nieto, B., Salgado, C., Pérez-Cabal, M.A., Ibáñez-Escriche, N., and Gutiérrez, J.P. (2013). Genetic parameters for birthweight environmental variability in mice. J. Anim. Breed. Genet. Z. Für Tierz. Zücht. ***130***, 404–414.

Rao, S., and Notter, D.R. (2000). Genetic analysis of litter size in Targhee, Suffolk, and Polypay sheep. J. Anim. Sci. ***78***, 2113–2120.

Rendel, J.M., (1979). Canalisation and selection. In: Thompson, J.N., Thoday, J.M. (Eds.), Quantitative genetic variation. Academic Press, New York.

Rendel, J., and Sheldon, B. (1960). Selection for Canalization of the Scute Phenotype in Drosophila Melanogaster. Aust. J. Biol. Sci. ***13***, 36–47.

Ricordeau, G., Desvignes, A., Tchamitchian, L., Rastogi, R., and Lefevre, C. (1976). Amélioration de la productivité des brebis Berrichonnes du Cher (BC) par croisement. I. Productivité numérique des brebis BC, Cotentin, Border Leicester, Romanov et de 3 types de F1. Ann. Génétique Sélection Anim. ***8***, 367–389.

Ricordeau, G., Tchamitchian, L., Lefevre, C., Brunel, J.., and Desvignes, A. (1977). Amelioration de la productivite des brebis Berrichonnes du Cher (BC) par croisement. III. Performances de reproduction des trois premieres generations de brebis croisees entre les races Berrichonne du Cher et Romanov. Ann. Genet. Sel. Anim.***8***, 405–419.

D. R. Notter, J.S.C. (1980). Performance of finnish Landrace crossbred ewes under accelerated lambing. I. Fartility, Prolificity and ewe productivity.

Robert-Granié, C., Ducrocq, V., and Foulley, J.L. (1997). Heterogeneity of variance for type traits in the Montbeliarde cattle breed. Genet. Sel. Evol.

Rodriguez, R.O.L., Heredia, A.M., Quintal, F.J., and Velazquez, M.A. (1998). Productivity of Pelibuey and Blackbelly ewes mated at yearly and 8-monthly intervals over six years. Small Rumin. Res. J. Int. Goat Assoc.

Roehe, R., and Kalm, E. (2000). Estimation of genetic and environmental risk factors associated with pre-weaning mortality in piglets using generalized linear mixed models. Anim. Sci. ***70***, 227–240.

Rönnegård, L., Felleki, M., Fikse, F., Mulder, H.A., and Strandberg, E. (2010). Genetic heterogeneity of residual variance - estimation of variance components using double hierarchical generalized linear models. Genet. Sel. Evol. *42*, 8.

Rönnegård, L., Felleki, M., Fikse, W.F., Mulder, H.A., and Strandberg, E. (2013). Variance component and breeding value estimation for genetic heterogeneity of residual variance in Swedish Holstein dairy cattle. J. Dairy Sci. ***96***, 2627–2636.

Ros, M., Sorensen, D., Waagepetersen, R., Dupont-Nivet, M., SanCristobal, M., Bonnet, J.-C., and Mallard, J. (2004). Evidence for Genetic Control of Adult Weight Plasticity in the Snail Helix aspersa. Genetics ***168***, 2089–2097.

Rowe, S.J., White, I.M., Avendaño, S., and Hill, W.G. (2006). Genetic heterogeneity of residual variance in broiler chickens. Genet. Sel. Evol. GSE *38*, 617–635.

SanCristobal-Gaudy, M.S., Robert-Granié, C., and Foulley, J.-L. (2002). Héteroscédasticité et modèles linéaires mixtes : théorie et applications en génétique quantitative. J. Société Fr. Stat. ***143***, 155–165.

SanCristobal-Gaudy, M., Elsen, J.-M., Bodin, L., and Chevalet, C. (1998a). Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. Genet. Sel. Evol. GSE ***30***, 423–451.

SanCristobal-Gaudy, M., Elsen, J.-M., Bodin, L., and Chevalet, C. (1998b). Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. Genet. Sel. Evol. GSE ***30***, 423–451.

SanCristobal-Gaudy, M., Bodin, L., Elsen, J.M., and Chevalet, C. (2001). Genetic components of litter size variability in sheep. Genet. Sel. Evol. GSE ***33***, 249–271.

Sauvant, D., Martin, O., (2010). Robustesse, rusticite, flexibilite, plasticite... les nouveaux criteres de qualite des animaux et des systemes d'elevage: definitions systemique et biologique des differents concepts. In: Sauvant, D., Perez, J.M. (Eds.), Dossier INRA Productions Animales, pp. 5-9.

Sauvant, D., Perez, J.M., (2010). Robustesse, rusticité, flexibilité, plasticité, résilience... Les nouveaux critères de qualité des animaux et des systèmes d'élevage. Dossier INRA Productions Animales, 23(1). INRA.

Scheiner, S.M., and Lyman, R.F. (1989). The genetics of phenotypic plasticity I. Heritability. J. Evol. Biol.***2***, 95–107.

Scheiner, S.M., and Lyman, R.F. (1991). The genetics of phenotypic plasticity. II. Response to selection. J. Evol. Biol.***4***, 23–50.

Sorensen, D., and Waagepetersen, R. (2003). Normal linear models with genetically structured residual variance heterogeneity: a case study. Genet. Res. ***82***, 207–222.

Soufflet, F. (1995). Filière viande bovine: labels ou démarche de qualité? In: Nicolas, F., Valeschini, E. (Eds.), Agro-alimentaire: Une économie de la qualité. NRA Ed. 281–292.

Steine, T. (1985). Genetic studies of reproduction in Norwegian sheep. Genetics of reproduction in sheep. Butterworths Lond. UK 47–54.

Tempelman, R.J., and Gianola, D. (1999). Genetic Analysis of Fertility in Dairy Cattle Using Negative Binomial Mixed Models. J. Dairy Sci. ***82***, 1834–1847.

Theriez, M., (1982). La mortalité des agneaux : point de vue du zootechnicien. 7ième Journées de la Recherche Ovine et Caprine. Paris INRA-ITOVIC Ed., Paris, pp. 1-17.

Theriez, M. (1991). Effects of increase in prolificacy on the rearing of lambs and on meat production. Prod. Anim. ***4***, 161–168.

Theriez, M., Molenat, G., (1975). Conduite intensive des troupeaux ovins. Effets du tarissement des la mise bas sur la fecondite de brebis inseminees tous les six mois. Annales de Zootechnie **24**, 729-742.

Thimonier, J., Mauleon, P., and Cognie, Y. Declenchement de l‘oestrus et obtention de la gestation pendant l‘anoestrus post-partum chez les brebis a l‘aide d‘eponges vaginales impregnees d‘acetate de fluorogestone. Ann. Zootech. ***17***, 257–273.

Thompson, R. (1979). Sire Evaluation. Biometrics ***35***, 339.

Thompson, R., McGuirk, B.J., and Gilmour, A.R. (1985). Estimating the heritability of all-or-none and categorical traits by offspring-parent regression. J. Anim. Breed. Genet. ***102***, 342–354.

Tiphine, L., David, I., Raoul, J., Guerrier, J., Praud, J., Bodin, L., François, D., Jullien, E., Poivey, J.P., (2011). Estimation of breeding values for meat sheep in France. 62nd EAAP meeting, Stavanger, Norway.

TorrèS, S., and Cognié, Y. (1984). Superovulation and egg transfer in the ewe. Reprod. Nutr. Dév. ***24***, 623–631.

tosh J (2011). Analysis of Lambing Interval.Centre for Genetic Improvement of Livestock. Animal and Poultry Science,University of Guelph. http://www.uoguelph.ca/~jtosh/lambint2.

Trottier, C. (1998). Estimation dans les modèles linéaires généralisés à effets aléatoires. Institut National Polytechnique de Grenoble.

Waddington, C.H. (1940). The genetic control of wing development inDrosophila. J. Genet. ***41***, 75–113.

Waldron, D.F., and Thomas, D.L. (1992). Increased litter size in Rambouillet sheep: I. Estimation of genetic parameters. J. Anim. Sci. ***70***, 3333–3344.

Weller, J.I., and Ron, M. (1992). Genetic analysis of fertility traits in Israeli Holsteins by linear and threshold models. J. Dairy Sci. ***75***, 2541–2548.

Wolc, A., White, I.M.S., Avendano, S., and Hill, W.G. (2009). Genetic variability in residual variation of body weight and conformation scores in broiler chickens. Poult. Sci. ***88***, 1156–1161.

Wright, S. (1934). An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs. Genetics ***19***, 506–536.

Zeger, S.L., and Karim, M.R. (1991). Generalized Linear Models With Random Effects; A Gibbs Sampling Approach. J. Am. Stat. Assoc. ***86***, 79–86.