



HAL
open science

Utilisation de données transcriptomiques pour l'étude du déterminisme génétique de la qualité du jambon cuit

Frédéric Hérault

► **To cite this version:**

Frédéric Hérault. Utilisation de données transcriptomiques pour l'étude du déterminisme génétique de la qualité du jambon cuit. Sciences du Vivant [q-bio]. 2009. hal-02649986

HAL Id: hal-02649986

<https://hal.inrae.fr/hal-02649986>

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Centre de Rennes
65, rue de St-Brieuc
CS 84215
35042 RENNES Cedex



Campus de Beaulieu
Avenue du Général Leclerc
CS 74205
3504 RENNES Cedex



INRA - UMR SENAH
Domaine de la Prise
35590 Saint Gilles

Mémoire de Fin d'Études

MASTER Biologie, Agronomie, Santé (BAS)

Spécialité Biologie Appliquée aux Productions Animales et à la Santé (BAPSA)

Parcours de Recherche

UTILISATION DE DONNÉES TRANSCRIPTOMIQUES POUR L'ÉTUDE DU DÉTERMINISME GÉNÉTIQUE DE LA QUALITÉ DU JAMBON CUIT.

Par M. Hérault Frédéric

Devant le jury :

C. Labbé,
E. Moreau,
F. Lecerf,
F. Ferriere,
V. Rioux,

Soutenu le 23 Juin 2009.

Sous la présidence de :

Rapporteur :

F. Le Gac

Maître de stage :

Damon Marie (CR1 INRA)

Le Roy Pascale (DR2 INRA)

« Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celles d'AGROCAMPUS OUEST, de l'Université de Rennes1 et de l'UMR SENAH.»

Diffusion du mémoire

A remplir par l'auteur avec le maître de stage.

Aucune confidentialité ne sera prise en compte si la durée n'en est pas précisée.

Préciser les limites de la confidentialité ⁽¹⁾ :

Mémoire de fin d'études

Consultable sur place :	<input type="checkbox"/> oui	<input type="checkbox"/> non
Reproduction autorisée :	<input type="checkbox"/> oui	<input type="checkbox"/> non
Prêt autorisé :	<input type="checkbox"/> oui	<input type="checkbox"/> non
Confidentialité absolue : (<i>ni consultation, ni prêt</i>)	<input type="checkbox"/> oui	<input type="checkbox"/> non
<i>Si oui</i>	<input type="checkbox"/> 1 an	<input type="checkbox"/> 5 ans <input type="checkbox"/> 10 ans
Diffusion de la version numérique :	<input type="checkbox"/> oui	<input type="checkbox"/> non

Fiche de résumé du mémoire de fin d'études :

Résumé diffusable : oui non

Si oui, l'auteur complète l'autorisation suivante :

Je soussigné(e) _____, propriétaire des droits de reproduction dudit résumé, autorise toutes les sources bibliographiques à le signaler et le publier.

Date :

Signature :

Rennes, le

Le Maître de stage⁽²⁾,

L'auteur,

L'Enseignant responsable⁽²⁾,

(1) L'administration, les enseignants et les différents services de documentation du Pôle Agronomique de Rennes s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme.

Remerciements

Je remercie Jean Noblet, directeur de l'Unité Mixte de Recherches «Système d'Élevage Nutrition Animale et Humaine» (UMR SENAH) de St Gilles, ainsi que Bénédicte Lebret pour m'avoir accueilli au sein de l'équipe «Croissance et qualité».

Je tiens à exprimer ma profonde reconnaissance à Marie Damon et Pascale Le Roy (UMR Génétique Animale) pour m'avoir guidé et encadré tout au long de ce travail.

Mes remerciements s'adressent également à Annie Vincent, sans qui ce travail n'aurait pu avoir lieu, pour avoir réalisé l'ensemble des expérimentations préalables à ce travail ainsi que les contrôles expérimentaux nécessaire à mes travaux.

Enfin un très grand remerciement à l'ensemble des personnels de l'équipe « Croissance et qualité », les personnels de l'UMR SENAH, et plus généralement l'ensemble des personnels travaillant sur le site de St Gilles.

Sans oublier de remercier Christian Diot, directeur de l'UMR Génétique Animale, Didier Boichard chef du département de Génétique Animale et l'INRA pour m'avoir permis de suivre cette année de formation dans les meilleures conditions possibles.

TABLE DES MATIÈRES

Glossaire.....	1
Liste des tableaux.....	3
Liste des figures.....	4
Introduction.....	5
1.Contexte scientifique.....	6
1.1.La qualité des viandes de porc.....	6
1.2.Transformation du muscle en viande.....	7
1.3.Facteurs de variation de la qualité des viandes.....	8
1.4.Objectif de l'étude.....	9
2.Matériels et méthodes.....	9
2.1.Animaux et mesures phénotypiques.....	9
2.2.La puce musculaire porcine 15K.....	10
2.3.Mesure de l'expression génique.....	10
2.4.Traitement des données d'expression.....	11
2.5.Analyse des données.....	13
2.6.Analyse des catégories fonctionnelles.....	14
3.Résultats.....	15
3.1.Qualité des données d'expression.....	15
3.2.Détection de QTL dans le SM.....	16
3.3.Détection de eQTL dans le SM.....	16
3.4.Co-localisation des eQTL et QTL.....	17
3.5.Analyse de l'expression différentielle entre le SM et le LD.....	17
4.Discussion.....	18
Analyse critique de l'étude.....	19
Localisation des QTL.....	20
Caractérisation des eQTL: Cis ou Trans?.....	21
Spécificités musculaires.....	22
Les mécanismes mis en jeu.....	23
Focus sur le gène RYR1.....	25
Conclusion.....	25
Références bibliographiques.....	27
Annexes.....	32

Glossaire.

a* : indice de rouge

ACP: analyse en composante principale

ADN, ADNc : acide désoxyribonucléique et acide désoxyribonucléique complémentaire

ALDOA : aldolase A

ANOVA: analyse de variance

ARN, ARNc : acide ribonucléique et acide ribonucléique complémentaire

ATP : adenosine tri phosphate

b* : indice de jaune

B2M-F : beta 2 microglobuline

BP: processus biologique

CC: composant cellulaire

CEBPA : CCAAT/enhancer binding protein alpha

CIELAB: Commission internationale de l'éclairage L* a* b*

cM : centimorgan

CPT1M: carnitine palmitoyl transférase 1 musculaires

DAVID: Database for Annotation, Visualisation and Integrated Discovery

DFD : viandes sèches, fermes et sombres

EASE: Expression Analysis Systematic Explorer

eQTL : QTL d'expression

FDR : taux de faux positifs

GAPDH: glyceraldéhyde phosphate déshydrogénase

GO : ontologie génique

Hal: gène Halothane

IPR : terme ontologique associé à la base de données protéique InterPro (InterPro Records)

KEGG : Kyoto Encyclopedia of genes and Genomes

L* : clarté

LD : muscle Longissimus Dorsi

LIM : lipides intramusculaires

MF: fonction moléculaires

NA: donnée manquante (not available)

NCBI: National Center for Biotechnology Information

PCR : réaction de polymérisation en chaîne

pHu : pH ultime

PRE : pouvoir de rétention en eau

PSE : viande pâle, molle et exsudative

QTL : locus contrôlant un caractère quantitatif

RN : rendement napole

RT: transcription reverse

RYR1 : récepteur à la ryanodine

SIGENAE: système d'information des GENomes des Animaux d'élevage

SM : muscle semimembraneux

SYBR green: agent intercalant fluorescent

TNNC2 : troponine C2

Liste des tableaux

Tableau 1: Caractérisation des muscles Longissimus dorsi (LD) et Semimembranosus (SM)..	7
Tableau 2: Séquence des amorces utilisées pour la réaction de RT-PCR quantitative.....	11
Tableau 3: Critères de qualité des spots détectés par Agilent Feature Extraction.....	11
Tableau 4: Résultats des étapes de filtration sur les critères de qualités dans les dispositifs eQTL et comparaison SM/LD.....	14
Tableau 5: QTL détectés.....	15
Tableau 6: Analyse de l'enrichissement en catégories fonctionnelles au sein des 5 eQTL.....	16
Tableau 7: Comparaison des résultats d'expression différentielle par RT-PCR quantitative et puces.....	17
Tableau 8: Analyse de l'enrichissement en catégories fonctionnelles au sein des 6 clusters de gènes.....	17

Liste des figures

Figure 1: Représentation schématique des relations entre l'évolution post mortem du pH musculaire et la qualité de la viande chez le porc.....	6
Figure 2: Trie de la viande en fonction de la couleur.....	6
Figure 3: Classification des fibres musculaires, leurs caractéristiques contractile et métabolique.....	7
Figure 4: Dispositif familial pleins-frères et demi-frères.....	10
Figure 5: Origines des annotations des sondes oligonucléotidiques.	10
Figure 6: Effet de la transformation logarithmique de l'intensité du signal.	12
Figure 7: Filtration des données d'expression.....	12
Figure 8: Co-localisation de QTL.....	16
Figure 9: Répartition des eQTL sur l'ensemble des 18 autosomes du génome porcin.....	17
Figure 10: Double classification hiérarchique ascendante	18
Figure 11: Représentation de trois modèles possibles de relations simples entre deux caractères quantitatifs subissant l'effet d'un même locus.....	21
Figure 14: Illustration du mode de prélèvement à l'aide d'un trocart.....	22
Figure 15: Coupe transversale de la cuisse de porc.....	22

Introduction.

La qualité de la viande de porc est un caractère complexe encore mal maîtrisé. Pour les industriels, la qualité technologique d'une viande, qui se définit comme sa capacité à subir des transformations, est primordiale puisqu'elle influence le rendement de transformation de la matière première (rendement de cuisson, rendement de tranchage...). La qualité sensorielle, quant à elle, tient compte de la flaveur, la couleur, l'odeur et l'aspect du produit... ces caractéristiques étant déterminantes dans l'acte d'achat du consommateur. Ces deux composantes de la qualité ont un impact économique important. La maîtrise de la qualité de la viande de porc est donc devenue dans les années 80, une des préoccupations majeures de la filière porcine.

La viande provenant de la maturation du muscle, les caractéristiques (composition et structure) et le métabolisme énergétique musculaires influencent fortement les propriétés physico-chimiques de la viande : le pouvoir de rétention en eau (PRE), la teneur et la composition en lipides intramusculaires (LIM), la couleur... (Hocquette *et al.*, 2000). Ces caractéristiques du muscle sont sous la dépendance de nombreux facteurs environnementaux (techniques d'élevage, alimentation, conditions de transport et d'abattage...) et génétique (race, sexe, génotype aux gènes majeurs Halothane et RN)(Olsson & Pickova, 2005; Rosenvold & Andersen, 2003). Il est possible d'améliorer la qualité technologique des viandes en optimisant les conditions d'élevage et les conditions péri abattage. Cependant, les modes d'action et l'influence de ces facteurs environnementaux sur les critères de qualité demeurent assez mal compris. De plus, malgré de nombreuses études sur les facteurs de risque, la qualité de la viande est aujourd'hui souvent décrite comme hétérogène. L'héritabilité des critères de qualité de la viande est faible à moyenne (0,10 à 0,30, Sellier, 1998). L'obtention d'un progrès génétique par sélection est donc possible. Cependant, il n'existe pas à l'heure actuelle de prédicteur fiable de la qualité de la viande mesurable *in vivo*. Il s'ensuit l'obligation d'un contrôle sur performance effectué sur des collatéraux des candidats à la sélection.

La qualité de la viande de porc est donc un caractère sur lequel le gain de progrès génétique attendu par l'application d'une sélection assistée par marqueurs pourrait en théorie être substantiel (Le Roy 2002). De plus, les outils de post génomique offrent la possibilité d'accéder à une meilleure compréhension des mécanismes moléculaires déterminants pour la construction de la qualité. Ainsi, l'identification de nouveaux prédicteurs de la qualité des viandes devrait permettre d'améliorer notablement la réponse à la sélection. Enfin, la fixation de génotypes favorables à des gènes influençant fortement la qualité de la viande (gènes majeurs ou Quantitative Trait Loci) pourrait permettre d'homogénéiser les caractéristiques de

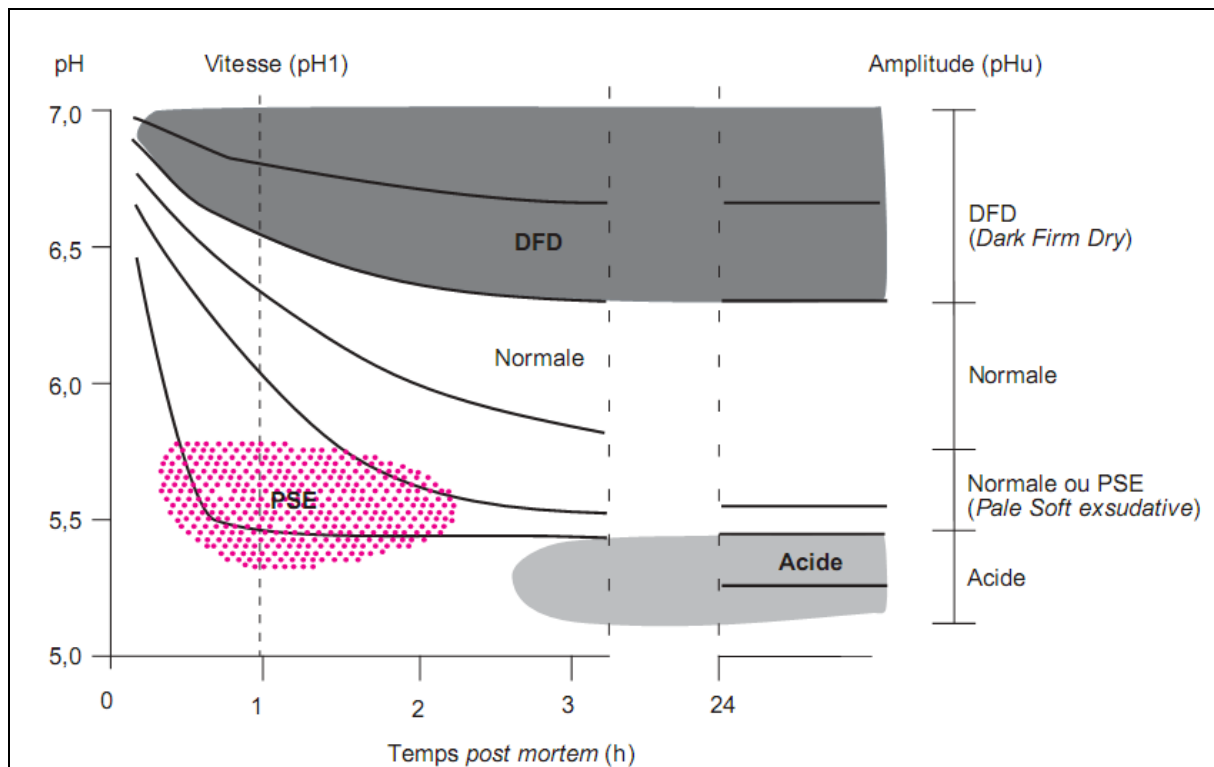


Figure 1: Représentation schématique des relations entre l'évolution post mortem du pH musculaire et la qualité de la viande chez le porc. Normalement le pH du muscle se situe aux alentours de 7,4. Le pH chute graduellement en 6-8 h jusqu'à 5,6 - 5,7 pour atteindre un pHu de 5,4 – 5,7. Pour une viande PSE la chute du pH est accéléré. Le pH descend en dessous de 6 durant la première heure post-mortem, tandis que le pHu se situe entre 5,4 et 5,7. Pour une viande Acide la vitesse de chute du pH est normale, mais le pHu est inférieur à 5,4. La viande DFD se caractérise par une chute du pH ralentie un pHu supérieur à 6,2 – 6,3 (Lebret 2004).

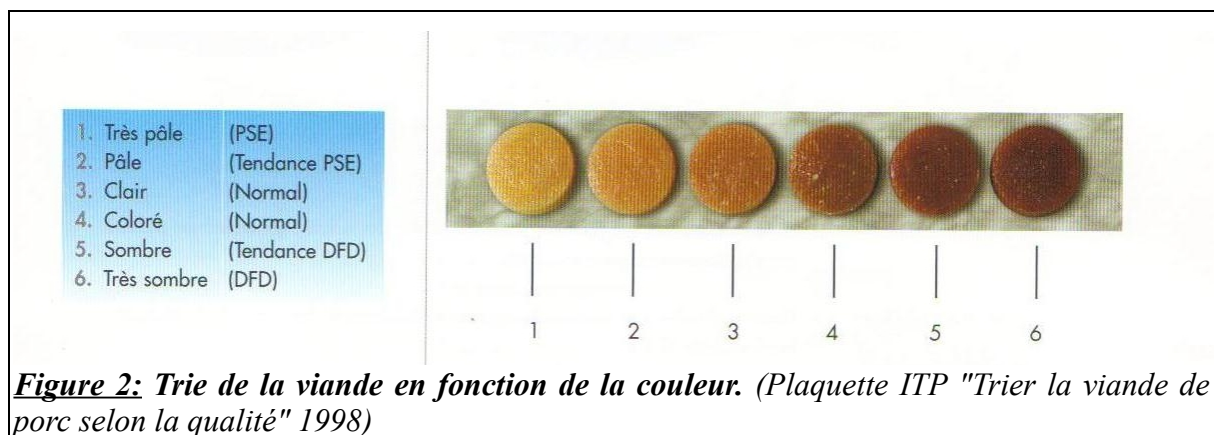


Figure 2: Trie de la viande en fonction de la couleur. (Plaquette ITP "Trier la viande de porc selon la qualité" 1998)

la matière première. Il serait alors plus facile d'optimiser les conditions environnementales afin d'obtenir au final un produit le plus homogène possible pour le transformateur.

1. Contexte scientifique.

1.1. La qualité des viandes de porc.

La viande de porc, que ce soit sous forme de viande fraîche ou de produits de charcuterie, est la première viande consommée en France. Les produits transformés représentent environ 70 % de cette consommation. En 2007, la fabrication de jambon cuit correspondait à 18 % des produits transformés soit 245 millions de tonnes (IFIP, 2009b). Après une forte croissance de la production de viande porcine au cours des années 60 à 80, la préoccupation de la filière porcine s'est tournée vers la maîtrise de la qualité de la viande.

L'industrie de transformation qui recherche une matière première de qualité technologique homogène s'intéresse aux critères ayant une influence sur le rendement de transformation de la viande : rendement de cuisson, qualité de tranchage du jambon, capacité de conservation...

Pour faire face à ces exigences de qualité, deux indicateurs de qualité sont actuellement utilisés dans les abattoirs industriels : l'appréciation de la couleur et la mesure du pH ultime (pHu). Le pHu est la valeur de pH mesuré 24h *post mortem* lorsque tout métabolisme énergétique est arrêté et que le pH est stabilisé. Le pHu et la couleur sont fortement corrélés au PRE et à la perte en eau à la cuisson (Huff-Lonergan *et al.*, 2002). Ils permettent d'effectuer un tri sur la matière première et d'éliminer ainsi les principaux défauts de qualité aujourd'hui identifiés.

Le pH *post mortem* permet de différencier les viandes : (Figure 1)

- **PSE** (Pale, Soft, Exsudative) : viandes pâles, molles et exsudatives,
- **DFD** (Dark, Firm, Dry) : viandes sombres, fermes et sèches,
- **acides** : faible valeur du pHu.

L'appréciation de la couleur de la viande permet d'identifier les viandes : (Figure 2)

- **bicolores** : différences de couleur très marquées,
- **déstructurées** : le muscle semble sans structure apparente et apparaît mou avec une teinte pâle et grise. Ce défaut touche principalement le muscle *Semimembranosus* (SM) du jambon et n'est visible que lorsque la pièce est désossée (IFIP, 2009a).

L'appréciation subjective de la couleur peut être instrumentalisée (spectrocolorimétrie). Actuellement, les paramètres les plus largement utilisés sont ceux du système CIELAB (1976) L* (clarté), a* (indice de rouge) et b* (indice de jaune). Le L* est fortement corrélé à la perte

Type de fibre musculaire:

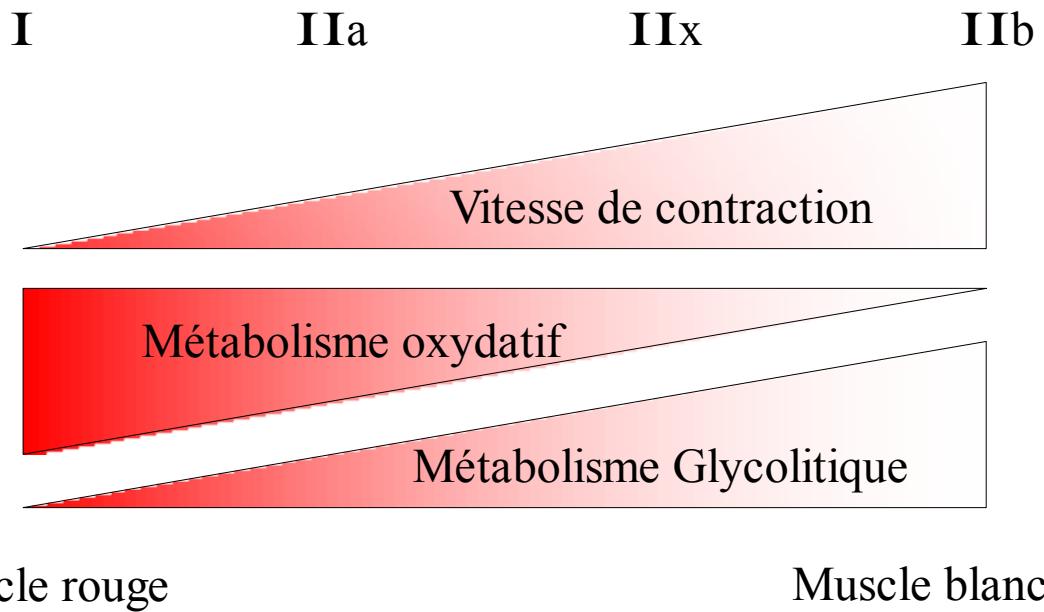


Figure 3: Classification des fibres musculaires, leurs caractéristiques contractile et métabolique.

	LD	SM
PH ultime	5,34	5,43
ATPase 10 ⁻³ μm KOH / mn / mg protéine	518	587
Citrate synthase 10 ⁻³ μm / mn / mg protéine	22,2	26,4
GAPDH μm / mn / mg protéine	7,1	1,6
Lactate déshydrogenase μm / mn / mg protéine	6,6	5,4
I % de fibre	6,8	6,5
IIa / IIx % de fibre	42,6	55,5
IIb % de fibre	50,6*	38

Tableau 1: Caractérisation des muscles *Longissimus dorsi* (LD) et *Semimembranosus* (SM). Caractéristiques biochimiques et type de fibres musculaires. (Adapté de Laborde, 1984; Melody et al., 2004).

en eau et à l'importance du niveau de déstructuration du SM (Franck *et al.*, 2000; Minvielle *et al.*, 2001).

Les qualités sensorielle et technologique sont déterminées par les caractéristiques du tissu musculaire au moment de la mort de l'animal. Elles dépendent en grande partie du type de fibres musculaires qui composent le muscle et du métabolisme énergétique de celui-ci (Hocquette *et al.*, 2000; Lebret, 2004).

1.2. Transformation du muscle en viande.

1.2.1. La structure musculaire.

Les muscles sont composés à 73 % d'eau, 20 % de protéines, 1 à 6 % de lipides, de glucides et sels minéraux. La plus grande partie de cette eau (85 %) est contenue dans les espaces entre les filaments d'actine et de myosine (Huff-Lonergan & Lonergan, 2005, 2007). Le tissu musculaire est très hétérogène, il contient des fibres musculaires ou myofibrilles, qui occupent 90 % du volume musculaire, du tissu conjonctif, du tissu adipeux, et des éléments du système nerveux et vasculaire (Lefaucheur, 1989). Les muscles sont composés de différents types de fibres musculaires. La distinction entre ces types de fibres est effectuée en fonction du type contractile de la fibre (fibre à vitesse de contraction lente ou rapide), et du type métabolique de la fibre (métabolisme oxydatif ou glycolytique). Cette classification permet de distinguer 4 types de fibres allant des fibres à contraction lente de métabolisme majoritairement oxydatif (type I et IIa), aux fibres à contraction rapide au métabolisme glycolytique (type IIx et IIb) (Figure 3)(Lefaucheur *et al.*, 1998). Les muscles dits « rouges » sont majoritairement composés de fibres lentes oxydatives, les muscles « blancs » de fibres rapides et glycolytiques. Les muscles SM et *Longissimus dorsi* (LD) sont deux muscles de type blanc. Cependant, ils se différencient par leurs niveaux de pHu, et d'activités enzymatiques, ATPases, LDH, GAPDH et citrate synthase, représentant la vitesse de contraction et les métabolismes glycolytique et oxydatif, ainsi que par leur proportion en différents types de fibres (Tableau 1)(Laborde, 1984; Melody *et al.*, 2004).

1.2.2. La transformation du muscle en viande.

L'évolution biochimique du muscle après l'abattage conduit à sa transformation en viande. Elle met en jeu des réactions d'hydrolyse qui dissipent les réserves énergétiques par voie anaérobie au cours de l'installation de la *rigor mortis* (rigidité cadavérique), puis affectent l'organisation des protéines myofibrillaires au cours de la phase de maturation. La cinétique et

l'intensité de ces réactions agissent sur les qualités technologique et sensorielle de la viande (Valin, 1988).

Durant l'installation de la *rigor mortis*, le muscle dégrade la majorité de ses réserves énergétiques. Ce catabolisme anaérobie produit de l'acide lactique ce qui acidifie le milieu intra-cellulaire. Le pH du muscle, aux alentours de 7, chute progressivement après la mort de l'animal pour atteindre un pHu de 5,4-5,7 (Monin, 2003; Scheffler & Gerrard, 2007). Le pHu dépend de l'importance des réserves musculaires en glycogène, au moment de la mort de l'animal. L'évolution *post mortem* du pH est directement dépendante de l'intensité du métabolisme musculaire *ante mortem* (niveau des réserves énergétiques au moment de la saignée) et *post mortem* (vitesse d'hydrolyse des molécules d'ATP). Les effets conjugués de la contraction musculaire provoquée lors de la *rigor mortis* et de la chute du pH contribuent au rétrécissement de la fibre musculaire. Il en résulte un déplacement de l'eau vers l'espace extra-cellulaire (Bertram *et al.*, 2004). Au cours de l'étape de maturation, la contraction des muscles disparaît, les protéines de structure musculaire (myofibrilles) et du cytosquelette sont dégradées sous l'action de protéases endogènes. L'eau jusqu'alors maintenue dans les espaces extra-cellulaires par la rigidité et la contraction des fibres, peut alors s'écouler hors de la pièce de viande. Ce phénomène contribue à la présence d'exsudat dans les barquettes de viande pré-emballée, à la perte de rendement à la fabrication de jambon cuit (efficacité du saumurage, perte à la cuisson), mais également à l'apparente pâleur de la viande due au reflet de la lumière sur la viande humide (Fischer, 2007; Hocquette *et al.*, 1998; Sellier *et al.*, 1992).

1.3. Facteurs de variation de la qualité des viandes.

De nombreuses études mettent en évidence l'effet des facteurs environnementaux et génétiques sur le PRE, la teneur et la composition en LIM, la couleur, le pHu. Ces études mettent clairement en évidence la complexité des interactions entre ces facteurs et des mécanismes mis en jeu (Olsson & Pickova, 2005; Rosenvold & Andersen, 2003). Parmi tous ces facteurs, il faut souligner la part importante de variabilité génétique inter et intra-race (10 à 30 %) ou le statut génotypique de l'animal au locus de gènes majeurs (RN ou Hal) (Davoli & Braglia, 2007).

L'amélioration de la qualité doit donc passer par une diminution de la variabilité due au génotype. Ainsi, lorsque la part de variabilité expliquée par des facteurs génétiques sera réduite, il sera alors possible d'optimiser les conditions environnementales afin d'obtenir au final un produit dont les qualités seront les plus homogènes possibles pour le transformateur.

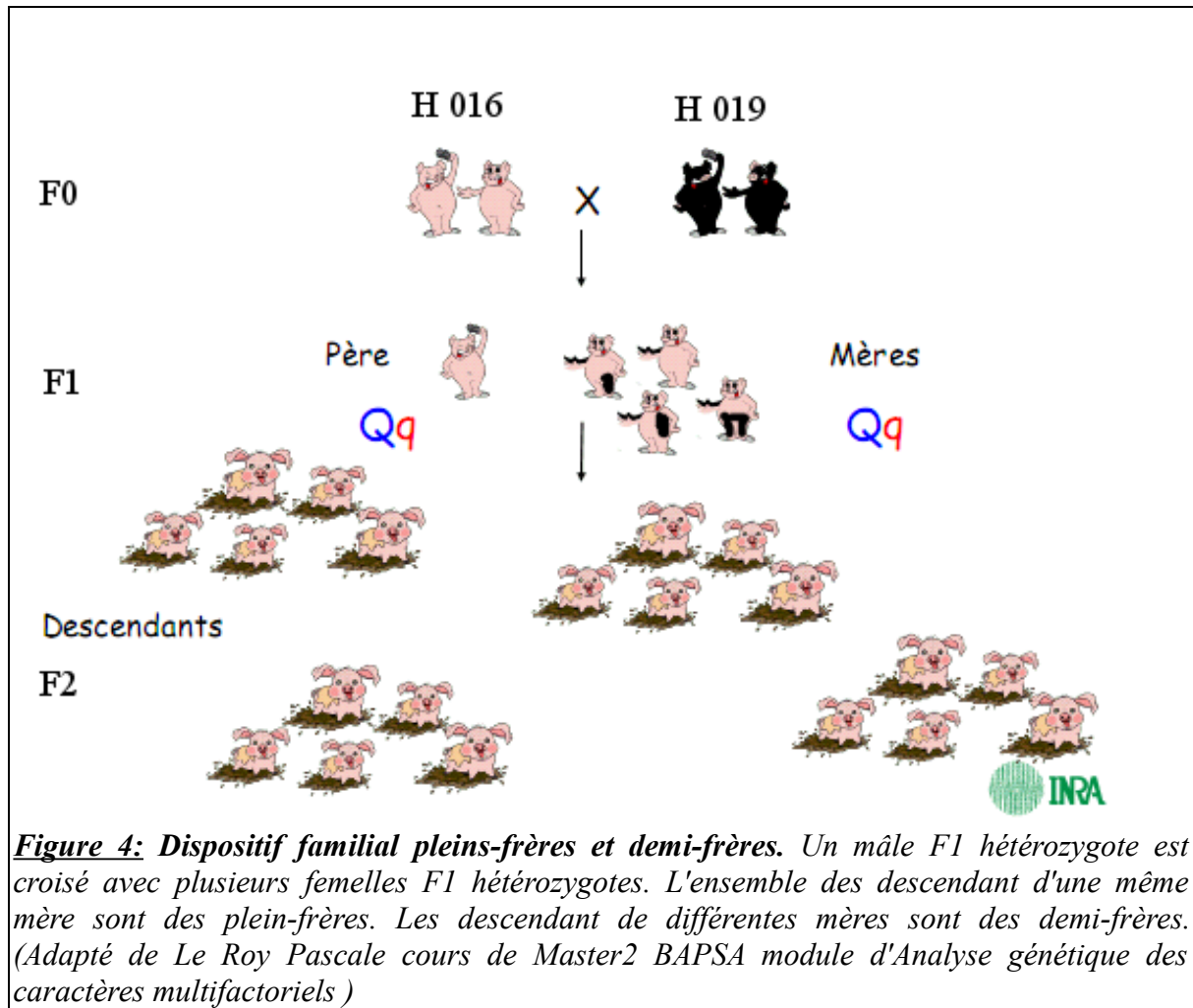


Figure 4: Dispositif familial pleins-frères et demi-frères. Un mâle F1 hétérozygote est croisé avec plusieurs femelles F1 hétérozygotes. L'ensemble des descendant d'une même mère sont des pleins-frères. Les descendant de différentes mères sont des demi-frères. (Adapté de Le Roy Pascale cours de Master2 BAPSA module d'Analyse génétique des caractères multifactoriels)

1.4. Objectif de l'étude.

L'étude de la variabilité génétique d'un caractère peut conduire à l'identification d'une région chromosomique (QTL : Quantitative Traits Loci) contrôlant ce caractère. Cependant, la détection du polymorphisme causal de l'effet observé reste un problème majeur du fait de la taille des QTL détectés qui comportent plusieurs centaines de gènes. L'utilisation conjointe de données phénotypiques et d'expression génique pourrait permettre de préciser le caractère et par là même d'affiner la position de la région d'intérêt : analyse eQTL (Rothschild *et al*, 2007).

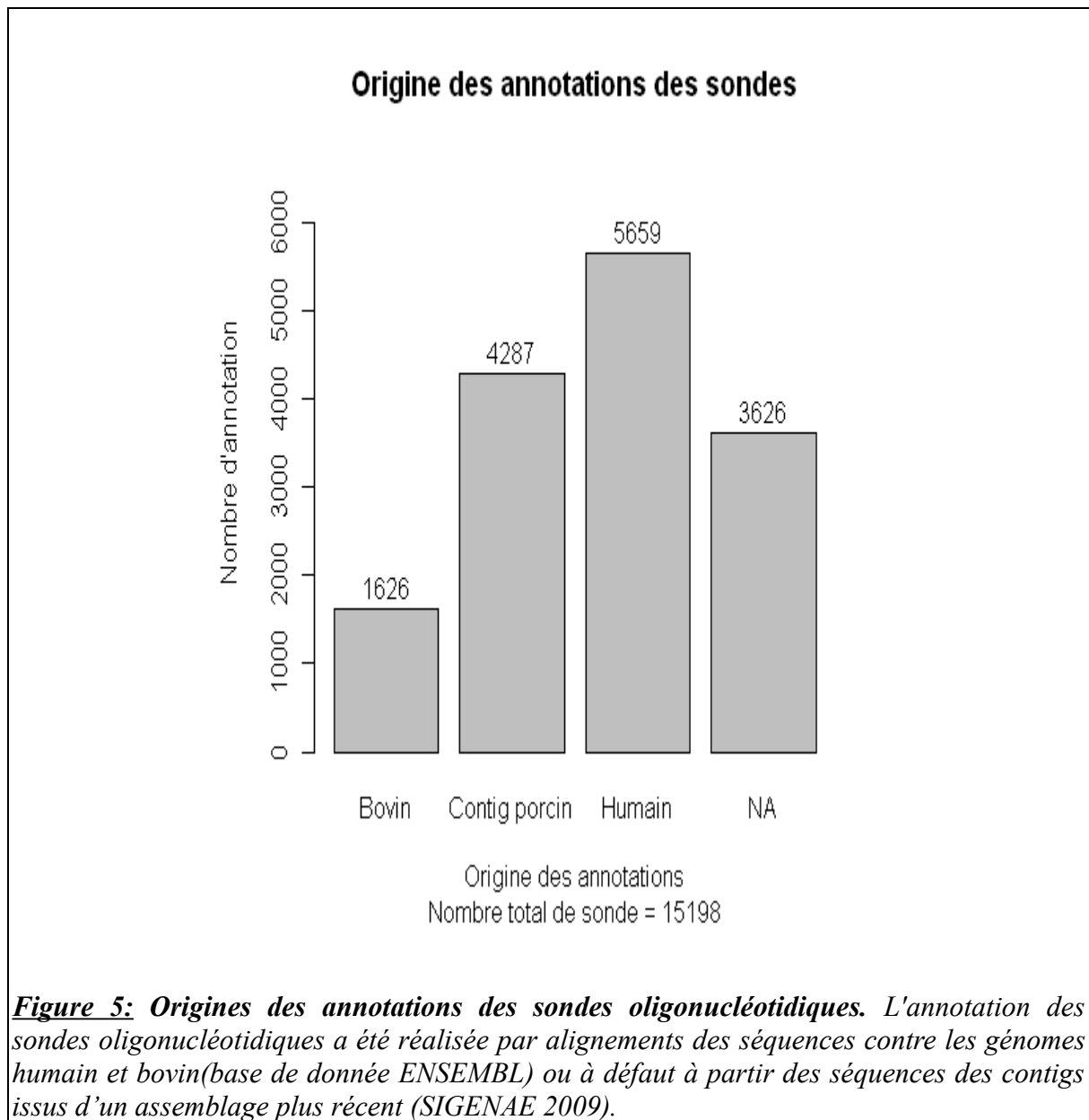
L'objectif de cette étude est de mieux comprendre le déterminisme de la qualité du jambon 1/ en recherchant de nouveaux gènes candidats ou « biomarqueurs » par une analyse génétique des caractères effectuée sur le muscle *Semimembranosus* (SM) et 2/ en déterminant les spécificités des profils d'expression de 2 muscles « modèles » pour la filière produit transformé (SM) et la filière viande fraîche (*Longissimus dorsi*, LD). L'intérêt de l'analyse génétique consiste en une approche combinée de détection de QTL phénotypiques (QTL) sur des mesures physico-chimiques associées à la qualité de viande (L^* , a^* , b^* et pHu) et de QTL d'expression (eQTL). Mon travail a porté sur le traitement et l'analyse de données d'expression génique préalablement acquises au laboratoire.

2. Matériels et méthodes.

2.1. Animaux et mesures phénotypiques.

Les animaux utilisés dans cette étude sont issus d'un protocole de croisement entre des animaux de races pures non porteurs des allèles RN^- et n défavorables pour la qualité de la viande. Les animaux sont issus d'une population commerciale de porcs élevés dans un élevage de sélection de la société France Hybrides situé à Sichamps (Nièvre). Le modèle expérimental complet est un dispositif F2 plein frère-demi-frère (Figure 4) créé à partir des lignées grand-parentales (F0) FH016 Piétrain et FH019 lignée synthétique (Hampshire, Duroc et Large White) de France Hybrides. Les animaux des trois générations du schéma expérimental (F0, F1, F2) ont été génotypés par France Hybrides à l'aide de 172 marqueurs microsatellites répartis sur les 18 autosomes du génome porcin.

Les animaux F2 ont été abattus dans un abattoir industriel (Orléans Viandes) à 105 kg. Des prélèvements de muscle LD et SM ont été effectués à l'aide d'un trocart sur la chaîne d'abattage après échaudage et immédiatement congelés dans l'azote liquide. Les mesures de couleur CIE « Lab » et de pHu ont été effectuées sur ces 2 muscles.



Dans le **dispositif QTL**, les mesures phénotypiques effectuées sur le SM de 735 animaux F2, issus de 9 familles de pères F1 ont été utilisées pour la détection de QTL dans ce muscle. Dans le **dispositif eQTL**, 2 familles de pères (familles 9 et 13) ont été sélectionnées sans privilégier de caractère phénotypique et les données d'expression génique de 112 descendants F2 (56 par famille) ont été utilisées. Dans le **dispositif de comparaison SM/LD**, les données d'expression génique acquises simultanément dans les 2 muscles (104 animaux F2 des familles 9 et 13) ont été utilisées pour effectuer l'analyse différentielle.

2.2. La puce musculaire porcine 15K.

Le génome du porc n'étant pas encore séquencé, le design de 43803 sondes a tout d'abord été réalisé à l'aide du logiciel eArray 4.5 (Agilent) en privilégiant la couverture du plus grand nombre de contigs musculaires possibles (29 811 contigs, SIGENAE 2006 v8). Après hybridation de cette puce, une sous sélection de 15198 sondes les plus informatives en terme de signaux d'hybridation validés a finalement été retenue. Ainsi, une puce 15K dédiée au muscle, a spécifiquement été dessinée pour ce programme et produite par la société Agilent. Les sondes sont des oligonucléotides (60 mers) synthétisés *in situ* sur la lame à l'aide d'une technologie « jet d'encre ». Chaque puce permet l'hybridation simultanée de 8 échantillons car elle comporte 8 cellules d'hybridation de 15198 sondes chacune (format 8*15K).

L'annotation de la puce a été réalisée par SIGENAE à partir de blasts entre les sondes et les génomes humain et bovin (base de donnée ENSEMBL) ou à défaut à partir des séquences des contigs issus d'un assemblage plus récent (SIGENAE 2009). Au final, 76% des sondes ont pu être annotées et elles correspondent à 9797 gènes (Figure 5).

2.3. Mesure de l'expression génique.

2.3.1. Pucés oligo

Brièvement, les ARN totaux ont été extraits au Trizol (Invitrogen) à partir de 100 mg de muscle. Ils sont rétro-transcrits en ADNc puis transcrits en ARN complémentaire (ARNc) marqué avec de la cyanine 3 (Agilent « Low RNA Input Amplification kit ») selon le protocole « One-color microarray-based gene expression analysis ». Le contrôle qualité des cibles est réalisé en vérifiant l'incorporation de la cyanine 3 et les ratios 260nm/280nm et 260nm/230nm avec le Nanodrop (Labtech). Les cibles marquées sont fragmentées puis hybridées (17h à 65°C). Après lavage, les lames sont immédiatement scannées sur « Agilent microarray scanner » (Agilent) avec une résolution de 5µm.

Numero d'accession	Amorces Sens	Amorces Antisens
Aldolase A (ALDOA)		
CB286787	CGC TGT CCC TGG GAT CAC	GCA CTT GTT GAT GGC GTT GA
Beta-2-microglobulin (B2M-F)		
DQ178123	AAA CGG AAA GCC AAA TTA CC	ATC CAC AGC GTT AGG AGT GA
Carnitine palmitoyltransferase (CPT-M)		
AF284832	CAC TGT CTG GGC AAA CCA AA	GCC ACC TGG TAG GAA CTC TCA AT
CCAAT/enhancer binding protein (C/EBP), alpha (CEBPA)		
AF103944.1	GTG GAC AAG AAC AGC AAC GA	CTC CAG CAC CTT CTG TTG AG
troponin C type 2 (fast) (TNNC2)		
AY575058	AGA TGA AAG AGG ACG CCA AG	GTC TGC ATT CCT GTC GAA GA

Tableau 2: Séquence des amorces utilisées pour la réaction de RT-PCR quantitative. Les amorces ont été dessinées à partir de séquences *Sus scroffa*. Les amorces pour l'amplification du 18S sont issu du Kit Pre-Developed TaqMan® Assay Reagents Human 18S rRNA (Applied Biosystems).

Signal de faible intensité	
« gIsFound »	<ul style="list-style-type: none"> signal - bruit de fond > 1,5 bruit de fond diametre > 0,3 diamètre de référence
« gIsPosAndSignif »	<ul style="list-style-type: none"> Moyenne du signal >> bruit de fond (test Student)
« gIsWellAboveBG »	<ul style="list-style-type: none"> « gIsPosAndSignif » Signal - bruit fond > 2,6 écart type du bruit de fond (SNR)
Signal non uniforme	
« gIsFeatNonUnifOL »	<ul style="list-style-type: none"> variance mesurée > variance estimée (prend en compte le bruit apporté par le scanner, le marquage, les lames...)
Signal saturé	
« gIsSaturated »	<ul style="list-style-type: none"> Plus de 50 % des pixels > au seuil de saturation

Tableau 3: Critères de qualité des spots détectés par Agilent Feature Extraction. Détails des critères d'assignation automatiques des « flags ».

2.3.2. Quantification d'expression par RT-PCR en temps réel.

Les amplifications sont réalisées à partir de 5 ng d'ARN transcrits en présence de 200 nM d'amorces spécifiques (Tableau 2) dans 12.5 µl de milieu réactionnel « Fast SYBR Green Master Mix » (Applied Biosystems) dans l'appareil StepOnePlus (Applied Biosystems). Les conditions de PCR sont : 2 min à 50°C, 20 sec à 95°C puis 40 cycles de dénaturation 3 sec à 95°C et hybridation-élongation 30 sec à 60°C. La spécificité des amplicons est vérifiée par l'analyse de leur courbe de fusion.

L'expression normalisée est calculée selon la formule suivante basée sur la méthode des Δ CT (Pfaffl, 2001) :

$$\text{expression normalisée} = E_{\text{cible}}^{-\Delta C_{\text{t cible}}} / \text{facteur de normalisation}$$


avec E, l'efficacité de la réaction de PCR; Δ CT= Ct échantillon - Ct échantillon calibrateur et facteur de normalisation = moyenne géométrique des quantités 18S et B2M-F calculée avec l'algorithme Genorm (Vandesompele *et al.*, 2002).

2.4. Traitement des données d'expression.

2.4.1. Acquisition des données d'expression.

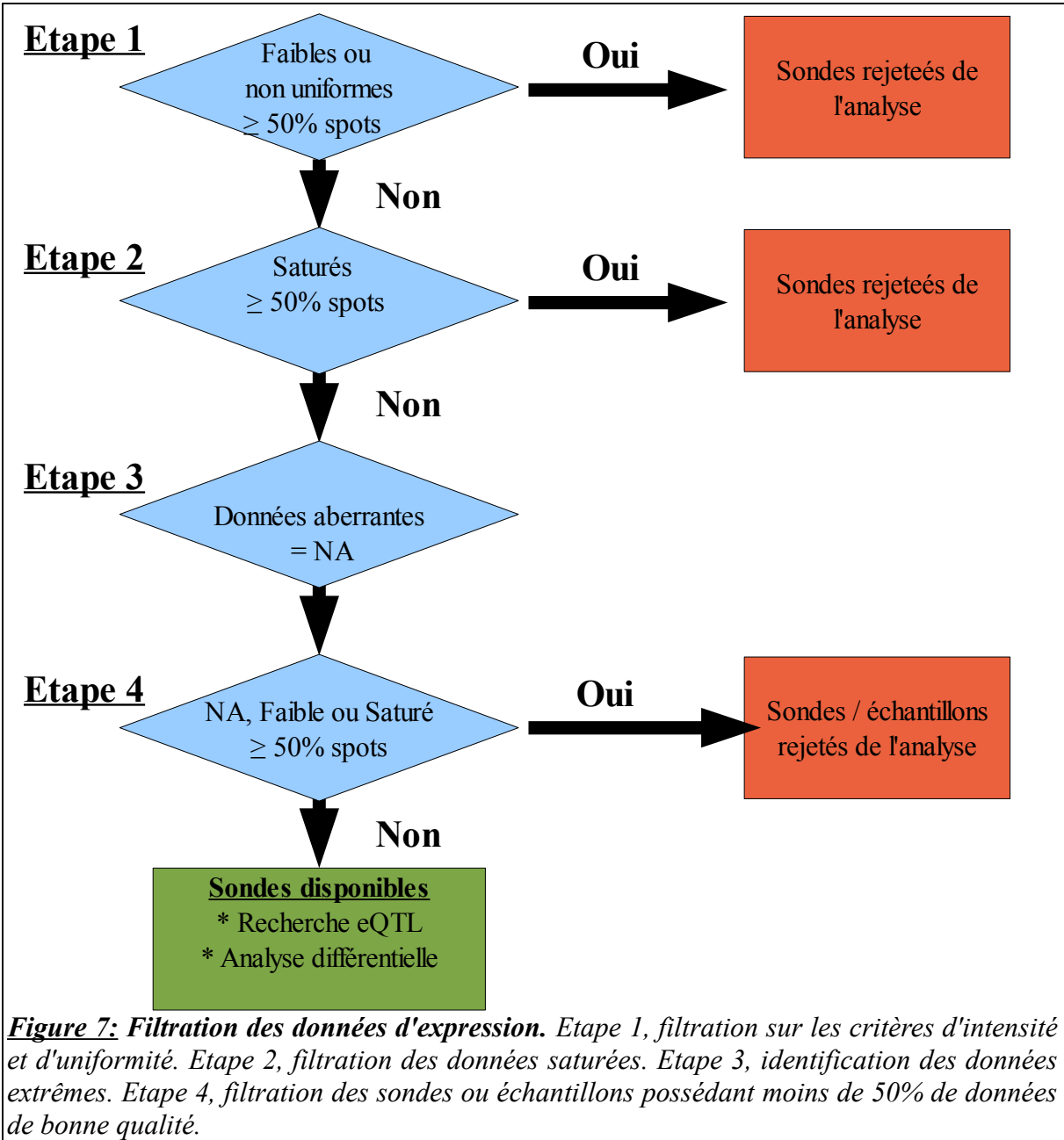
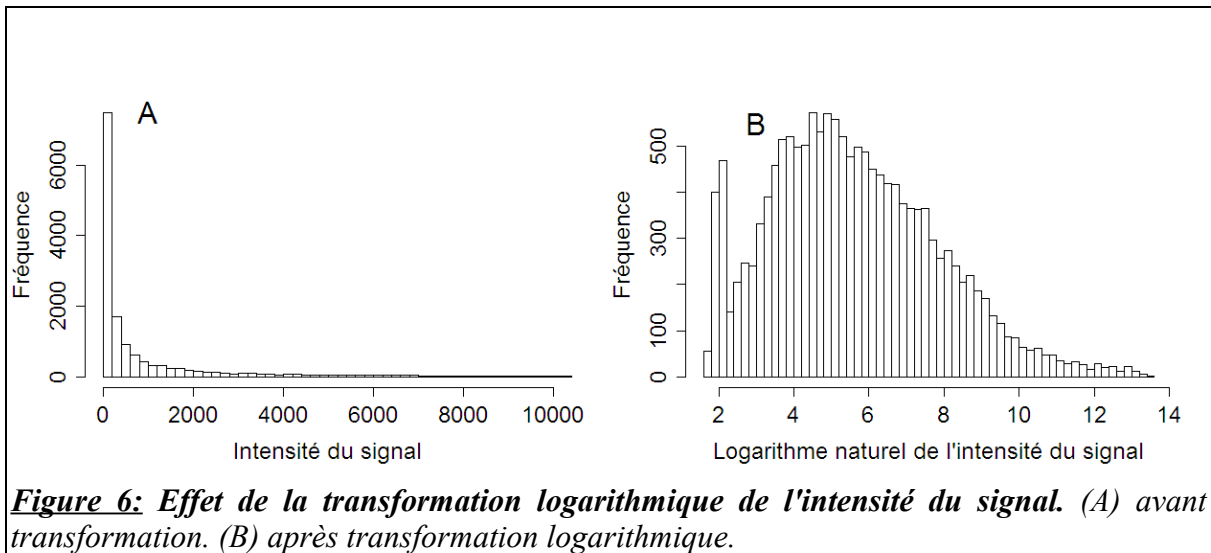
L'acquisition des données est effectuée à l'aide du logiciel Agilent Feature Extraction Software (V 9.5.1.1) en utilisant le programme pour hybridation mono-couleur GE1-V5_95_Feb07. Le logiciel détermine automatiquement 3 classes de critères de qualité pour les signaux de faible intensité, non uniformes et saturés (Tableau 3). Il effectue également une normalisation mono-couleur intra-lame du signal (« Spatial detrend »). Il s'agit d'une correction locale de l'impact du bruit de fond inclus dans le signal des spots de faible intensité, à partir d'une implémentation de l'algorithme Loess/Lowess.

2.4.2. Pré-traitement des données d'expression.

Les pré-traitements des données d'expression ont été effectués avec le logiciel  version 2.8.1 (R Development Core Team, 2008a; R Development Core Team, 2008b).

2.4.2.1. Transformation des données et normalisation.

Le signal (« gProcessedSignal ») utilisé pour les analyses de données d'expression correspond au signal moyen normalisé intra-lame (« Spatial detrend »). Après transformation logarithmique (Ln), les données sont centrées par la médiane de l'échantillon. La transformation logarithmique des données permet d'obtenir une distribution quasi normale de



celles-ci, ce qui autorise d'une part l'utilisation de tests statistiques paramétriques, et d'autre part la détection des données extrêmes (Figure 6).

2.4.2.2. Filtration des données d'expression.

Le pré-traitement des données d'expression comporte 4 étapes de filtration permettant de sélectionner des données d'expression de bonne qualité (Figure 7). Dans le cadre du dispositif de comparaison SM/LD, les différentes étapes de filtration ont été effectuées pour chaque muscle.

- Etape 1 : les spots de faible intensité risquent d'être fortement parasités par le bruit de fond. D'autre part, les spots non uniformes risquent eux même d'apporter du bruit dans le jeu de données. Aussi, les sondes dont plus de 50 % des données ne satisfaisaient pas les critères de qualité, faible intensité et uniformité (« Flags ») ont été rejetées de l'analyse et la valeur « donnée manquante » (NA) est assignée aux spots non uniformes subsistant.
- Etape 2 : les sondes avec plus de 50 % des données saturées sont également éliminées.
- Etape 3 : Dans une distribution normale, 99,7 % des données se trouvent dans un intervalle de 3 écart-types autour de la moyenne. Les données non comprises dans cet intervalle sont considérées comme des données aberrantes et sont remplacées par la valeur « donnée manquante » (NA).
- Etape 4 : les sondes ou échantillons dont plus de 50 % des données sont manquantes (NA), de faible intensité ou saturées ont été supprimés de chaque dispositif.

2.4.2.3. Filtration des gènes de faible variance.


Les gènes restants sont regroupés en 3 catégories en fonction de leur variance à l'aide de l'algorithme K-means. Le groupe représentant les sondes de faible variance est considéré comme invariant et n'est pas conservé pour l'analyse. Cette filtration permet d'éliminer les gènes invariants, de diminuer le nombre de tests d'hypothèse effectués et concourt, avec le contrôle des faux positifs (FDR), à accroître la puissance des tests statistiques (Hackstadt & Hess, 2009).

2.5. Analyse des données.


2.5.1. Analyse de variance : ANOVA.

2.5.1.1. Correction des mesures phénotypiques et d'expression génique.

Les mesures phénotypiques (L*, a* et b* et pHu) ont été corrigées par une analyse de variance (ANOVA, procédure glm de SAS) pour les effets sexe, poids de carcasse, conditions *péri mortem* (date d'abattage) et l'interaction sexe*poids de carcasse. Les facteurs significatifs ont été sélectionnés après analyse du modèle complet en posant un seuil de risque d'erreur α de 20 %. Les résidus de l'ANOVA ont été utilisés pour la recherche de QTL.

De même, les données d'expression de chaque gène ont été corrigées pour les effets sexe, conditions d'abattage, conditions d'hybridation et poids de carcasse en utilisant la fonction glm du package « stats » de  et la fonction Anova du package « car » Version 1.2-12 (Fox, 2009). Les facteurs significatifs pour chaque gène ont été sélectionnés après analyse du modèle complet en posant un seuil de risque d'erreur α de 20 %. Les résidus de l'ANOVA ont été utilisés pour effectuer la recherche de eQTL.

2.5.1.2. Recherche de gènes différentiellement exprimés entre le SM et le LD.

Les gènes différentiellement exprimés entre les deux muscles ont été identifiés par ANOVA dans . Pour chaque gène, les facteurs significatifs ont été analysés par ANOVA sur la base d'un modèle complet. Les facteurs explicatifs pris en compte dans ce modèle sont l'effet muscle (SM/LD), l'effet sexe, l'effet conditions d'abattage (date d'abattage), l'effet famille niché dans l'effet conditions d'hybridation (date d'hybridation) et l'effet poids de la carcasse en covariable :

$$\text{Expr-gène} = \text{Muscle} + \text{Sexe} + \text{Date-Abat.} + \text{Famille} / \text{Date-Hyb} + \beta \text{ Poids-carcasse} + \varepsilon$$

avec Expr-gène : expression du gène, Muscle : effet du facteur muscle, Sexe : l'effet du facteur sexe, Date-Abat : l'effet des conditions *péri-mortem* de l'animal, Famille / Date-Hyb : l'effet famille par date d'hybridation (date, condition technique...), Poids-carcasse en covariable et ε l'erreur aléatoire.

Pour chaque gène, le modèle contenant les effets significatifs au seuil de 20% a été sélectionné pour réaliser l'analyse différentielle. Les valeurs de pValues ont été ajustées pour tenir compte des tests multiples en appliquant la correction de Benjamini et Hochberg (contrôle du FDR).

2.5.2. Détection de QTL.

La recherche de QTL consiste à détecter l'existence de régions du génome responsables de la variabilité d'un caractère. Cette recherche s'appuie, dans cette étude, sur un dispositif de croisement F2. Les mesures phénotypiques effectuées sur la génération F2 et le génotypage des générations F0, F1 et F2 permettent d'identifier la transmission d'allèles parentaux agissant sur le niveau du caractère étudié.

Les analyses statistiques de détection de QTL et de eQTL ont été réalisées avec le logiciel QTLMAP, basé sur la méthodologie du Maximum de Vraisemblance et la cartographie d'intervalle (interval mapping). L'analyse effectuée sur tout le génome avec un pas de 1 cM consiste à tester l'hypothèse H1, d'existence d'un QTL contre l'hypothèse H0, d'absence de QTL à cette position (Le Roy *et al.*, 1998). Pour cette analyse, les positions des marqueurs sont celles de la carte dite "consensus" internationale, la distance entre marqueurs ayant été ré-estimée sur les données de génotypage des animaux du programme D116 dont est issu notre dispositif (1200 porcs issus de 18 familles de père).

La significativité des régions QTL et eQTL a été testée par rapport à des seuils obtenus par simulation, respectivement 1000 et 5000, sous l'hypothèse nulle d'absence de QTL.

2.6. Analyse des catégories fonctionnelles.

Pour l'interprétation biologique des données d'expression génique, le logiciel DAVID (Dennis *et al.*, 2003; Huang, Sherman & Lempicki, 2008; NIAID - NIH, 2009) a été utilisé. Il permet notamment de définir des familles de catégories fonctionnelles dont l'abondance est associée spécifiquement à une liste de gènes d'intérêt (analyse d'enrichissement via l'étude de la co-occurrence des identifiants de plus de 40 catégories fonctionnelles). Par défaut, les algorithmes utilisent les termes GO (BP, biological process ; CC, cellular component et MF, molecular function), KEGG, InterPro records (IPR), SP_PIR_Keywords et Uniprot sequence feature. Dans l'analyse d'enrichissement (analyse EASE pour Expression Analysis Systematic Explorer), l'identification des catégories fonctionnelles les plus pertinentes est basée sur un test exact de Fisher modifié. La probabilité associée à ce test est appelée score EASE. La médiane du score EASE au sein du groupe d'annotations fonctionnelles est le critère permettant de sélectionner les catégories significatives les plus pertinentes (seuil de 10 % pour les analyses eQTL et seuil de 5 % pour l'analyse différentielle).

	EQTL	SM/LD	
	112 échantillons	SM 104 échantillons	LD 104 échantillons
Faibles intensités >50%	4346 sondes	4335 sondes	2662 sondes
Non uniformes >50%	11 sondes	11 sondes	16 sondes
Saturés >50%	2 sondes	2 sondes	1 sonde
Nombre de sondes rejetées	4362 sondes (28,7%)	4405 sondes (29%)	

Tableau 4: Résultats des étapes de filtration sur les critères de qualités dans les dispositifs eQTL et comparaison SM/LD.

Le génome de « référence » de la puce 15K sous DAVID est constitué de 6019 gènes sur les 9797 (redondance et/ou annotations obsolètes). Ce génome de référence comporte 4776 termes GO_BP, 4835 termes GO_CC, 5184 termes GO_MF, 1711 termes KEGG, 5525 termes SP_PIR_KEYWORDS et 5642 termes IPR soit environ 30 % des termes disponibles pour le génome humain sous DAVID.

3. Résultats.

3.1. Qualité des données d'expression.

3.1.1. Filtration des données sur les critères de « qualité ».

Les données d'expression de mauvaise qualité sont éliminées au cours de la filtration des données et seules les données de bonne qualité ont été normalisées. Les résultats de chacune des étapes sont ici brièvement présentés (Tableau 4).

- Intensité faible : Environ 28 % des sondes sont de faible intensité. Parmi les sondes rejetées dans le dispositif de comparaison SM/LD, 23 sont de faible intensité dans plus de 90 % des échantillons SM tandis qu'elles sont de faible intensité dans moins de 10 % des échantillons LD. Ces sondes ont été considérées comme différentiellement exprimées entre les deux muscles.
- Uniformité : 11 sondes sont concernées dans le SM, et 16 dans le LD.
- Saturation : 1 sonde est spécifiquement saturée dans le SM tandis qu'une sonde est saturée dans les 2 muscles. La sonde saturée dans le SM (72% des spots), n'est saturée que dans 22% des spots pour le LD. Cette sonde a été conservée dans la liste des gènes différentiellement exprimés entre les 2 muscles.
- Données extrêmes : les données extrêmes représentent 0,5 % des données.

Au final, dans chaque dispositif 71 % des sondes de la puce 15K et l'ensemble des échantillons disponibles ont été conservés soit : 10823 sondes et 112 échantillons SM pour le dispositif eQTL et 10785 sondes et 208 échantillons (104 SM et 104 LD) dans le dispositif SM/LD. Par ailleurs, un groupe de 24 gènes différentiellement exprimés entre les deux muscles a été sélectionné à ce stade de l'analyse.

3.1.2. Filtration des sondes de faible variance.

Cette ultime étape de filtration supprime de l'analyse les sondes considérées comme invariantes et élimine 38 % des sondes (4158 spots) du dispositif eQTL et 36 % des sondes (3921 spots) du dispositif différentiel SM/LD. Ainsi, 6665 et 6864 sondes sont respectivement

Chromosome	caractère	Position	Intervalle de confiance	Longueur	Seuil de détection
7	b*	21 cM	10 à 35 cM	25 cM	5 %
10	b*	47 cM	35 à 62 cM	27 cM	1 %
10	L*	49 cM	38 à 58 cM	20 cM	5 %
13	a*	81 cM	66 à 88 cM	22 cM	10 %
13	b*	153 cM	139 à 166 cM	27 cM	10 %
13	pHu	66 cM	55 à 75 cM	22 cM	10 %
16	a*	0 cM	0 à 1 cM	1 cM	5 %
16	b*	16 cM	8 à 27 cM	19 cM	10 %
17	a*	4 cM	0 à 14 cM	14 cM	5 %

Tableau 5: QTL détectés. Les familles 9 et 13 sont informatives pour les QTL 7b, 10b, 10L, 13pHu et 16b.

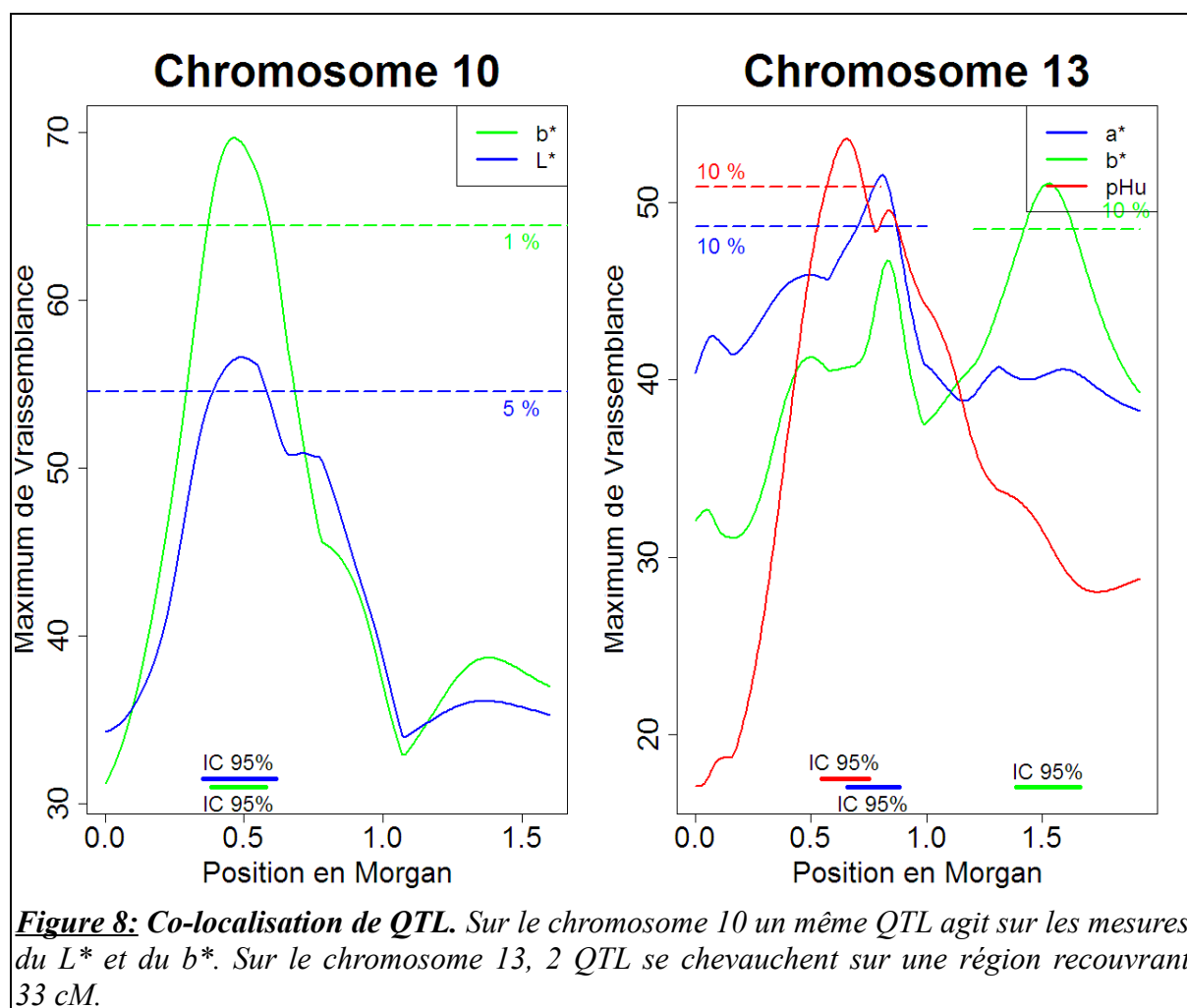


Figure 8: Co-localisation de QTL. Sur le chromosome 10 un même QTL agit sur les mesures du L* et du b*. Sur le chromosome 13, 2 QTL se chevauchent sur une région recouvrant 33 cM.

conservées dans les dispositifs eQTL et différentiel SM/LD. La variance des sondes conservées pour la suite de l'analyse est supérieure à 0.035.

3.2. Détection de OTL dans le SM.

Deux modèles ont été retenus pour corriger les différentes mesures phénotypiques : le modèle complet pour L* et b* et un modèle comprenant la date d'abattage et le sexe pour a* et pHu.

Les mesures de 735 descendants F2, issus de 9 pères et 32 mères, ont permis de mettre en évidence 9 QTL (Tableau 5). Des QTL ont été détectés pour les caractères :

- a* sur les chromosomes 13, 16 et 17 au seuil α de 10 %, 5 % et 5 %,
- b* sur les chromosomes 7, 10, 13 et 16, à 5 %, 1 %, 10 % et 5 %,
- L* sur le chromosome 10, à 5 %,
- pHu sur le chromosome 13, à 5 %.

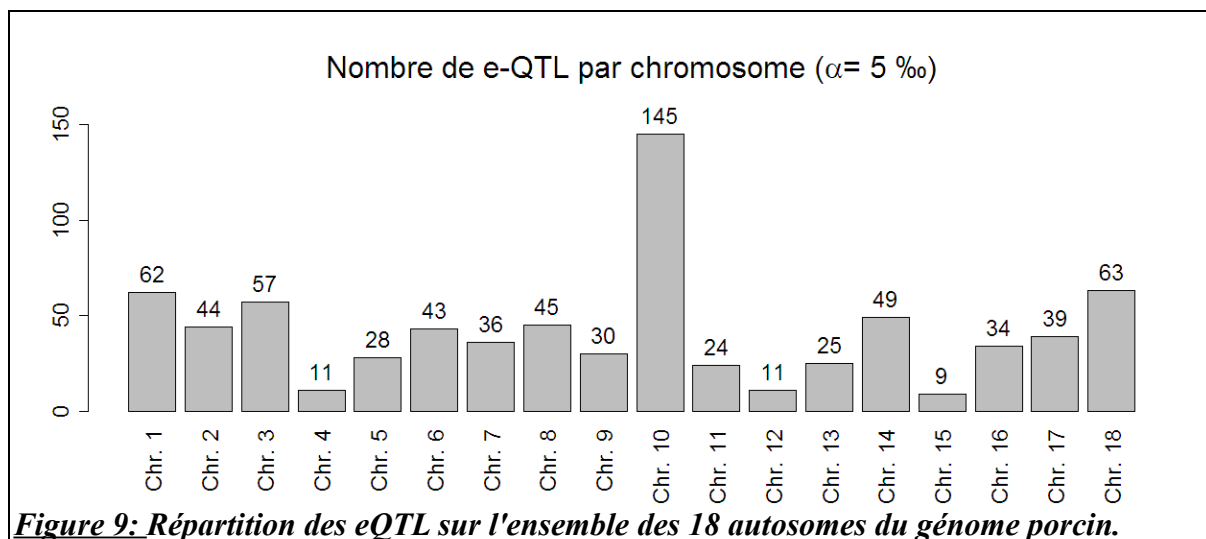
Parmi ceux-ci, le QTL 10L montre un fort effet (-20 sur une échelle de 1 à 100) de substitution allélique (Falconer, 1960) pour deux familles. Les QTL 10L et 10b sont localisés dans la même région du chromosome. Les QTL 13pHu et 13a, quant à eux, se chevauchent et recouvrent une région de 33 cM (Figure 8).

Cependant, les familles 9 et 13 (112 descendants F2) du dispositif eQTL ne sont informatives que sur 5 des 9 QTL détectés.

3.3. Détection de eQTL dans le SM.

L'analyse du modèle complet en ANOVA laisse apparaître que les conditions d'hybridation ont une influence sur une grande partie des gènes du dispositif (73 %), tandis que les facteurs sexe, condition d'abattage et poids de carcasse ne sont significatifs que pour 32 %, 23 % et 29 % des gènes.

Les facteurs significatifs ont été conservés pour définir un modèle de correction spécifique de chaque gène. Quinze combinaisons de facteurs significatifs ont été définies pour corriger l'ensemble des gènes du dispositif (Annexe I). Les résidus des ANOVA ont été utilisés pour effectuer une recherche de eQTL en utilisant le logiciel QTLMAP.



QTL 7b		sondes 90 gènes 59			
	Median Ease Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	5,66E-04	SP_PIR_KEYWORD	signal	9	15,25%
2	0,0620	INTERPRO	IPR002048:Calcium-binding EF-hand	4	6,78%
3	0,0648	GOTERM_CC_ALL	GO:0043292~contractile fiber	4	6,78%
4	0,0970	GOTERM_CC_ALL	GO:0031224~intrinsic to membrane	29	49,15%
QTL 10b		sondes 133 gènes 86			
	Median Ease Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	2,32E-10	GOTERM_CC_ALL	GO:0031012~extracellular matrix	16	18,60%
2	0,0011	SP_PIR_KEYWORD	collagen	16	18,60%
		KEGG_PATHWAY	hsa01430:Cell Communication		
3	0,0028	INTERPRO	IPR001611:Leucine-rich repeat	6	6,98%
4	0,0331	GOTERM_MF_ALL	GO:0030247~polysaccharide binding	4	4,65%
5	0,0783	GOTERM_BP_ALL	GO:0007166~cell surface receptor linked signal transduction	22	25,58%
6	0,0808	INTERPRO	IPR003961:Fibronectin, type III	4	4,65%
7	0,0942	INTERPRO	IPR013091:EGF calcium-binding	10	11,63%
QTL 16b		sondes 83 gènes 46		aucun cluster	
QTL 10L		sondes 218 gènes 126			
	Median Ease Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	1,42E-10	GOTERM_CC_ALL	GO:0031012~extracellular matrix	56	44,44%
2	0,0028	SP_PIR_KEYWORD	collagen	21	16,67%
		GOTERM_BP_ALL	GO:0015698~inorganic anion transport		
3	0,0059	SP_PIR_KEYWORD	leucine-rich repeat	7	5,56%
4	0,0216	GOTERM_BP_ALL	GO:0007166~cell surface receptor linked signal transduction	16	12,70%
5	0,0338	INTERPRO	IPR001881:EGF-like calcium-binding	7	5,56%
6	0,0374	GOTERM_BP_ALL	GO:0007166~cell surface receptor linked signal transduction	33	26,19%
7	0,0732	SP_PIR_KEYWORD	cell adhesion	11	8,73%
8	0,0888	GOTERM_MF_ALL	GO:0030247~polysaccharide binding	5	3,97%
QTL 13pHu		sondes 135 gènes 77			
	Median Ease Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	0,0182	INTERPRO	IPR003659:Plexin/semaphorin/integrin	6	7,79%
2	0,0620	INTERPRO	IPR013111:EGF, extracellular	4	5,19%
3	0,0663	GOTERM_MF_ALL	GO:0005509~calcium ion binding	8	10,39%
4	0,0723	GOTERM_MF_ALL	GO:0004871~signal transducer activity	36	46,75%

Tableau 6: Analyse de l'enrichissement en catégories fonctionnelles au sein des 5 eQTL. Seuil 10%.

Les mesures des 112 descendants F2, issus de 2 pères et 5 mères, ont permis de mettre en évidence 755 eQTL au seuil α de 5 %. Ces eQTL se répartissent sur les 18 autosomes du génome porcin (Figure 9).

3.4. Co-localisation des eQTL et QTL.

Pour les 5 régions QTL sélectionnées dans le dispositif eQTL, 11 eQTL co-localisent avec le QTL 7b, 43 eQTL co-localisent avec les QTL 10b et 10L, 9 eQTL co-localisent avec le QTL 13pHu et 11 eQTL avec le QTL 16b.

A partir de ces 5 listes de gènes, nous avons cherché à établir l'existence de Cis-eQTL en recherchant systématiquement la localisation de ces gènes sur le génome porcin sur la partie génome porcin du site du NCBI. Cette recherche s'est avérée infructueuse du fait du séquençage partiel et de l'annotation incomplète du génome porcin (NCBI, 2009).

L'analyse fonctionnelle, des 5 régions eQTL co-localisant avec un QTL, a été réalisée avec DAVID sur les gènes eQTL sélectionnés avec un risque α de 5 % afin de disposer d'un effectif de gènes suffisant pour l'analyse d'enrichissement. Cette analyse ne permet pas de mettre en évidence d'enrichissement particulier pour la région du eQTL 16b. Les 4 autres régions eQTL révèlent un enrichissement en gènes impliqués dans les catégories : communication cellulaire, matrice extra-cellulaire et récepteurs membranaires et plus particulièrement les récepteurs membranaires possédant un domaine de reconnaissance calcique (Tableau 6).

3.5. Analyse de l'expression différentielle entre le SM et le LD.

Le facteur muscle a un impact sur près de 92 % des sondes du dispositif SM/LD soit 6291 sondes. Les conditions d'hybridation associées au facteur famille de père ont également une influence sur un grand nombre de gènes (93 %). Les autres facteurs, sexe, poids de carcasse et condition d'abattage, ont un effet significatif sur respectivement 30 %, 35 % et 20 % des gènes du dispositif. Les facteurs significatifs ont été conservés pour définir le modèle ANOVA adapté à l'analyse des données d'expression de chaque gène. Seize modèles ont ainsi été retenus pour l'analyse de l'effet du muscle sur les 6864 sondes validées lors du pré-traitement (Annexe II).

Le facteur muscle a un effet sur 5960 sondes (p ajustée < 0.05). Seules 2338 sondes ont un ratio d'expression différentielle supérieur à 1.5, 1297 sondes sont plus exprimées dans le LD et 1041 le sont dans le SM. De plus, 70% des sondes ont un ratio d'expression différentielle compris entre 0.5 et 2.

GENES	RT-PCR		Puces	
	RATIO LD/SM	pValue	RATIO LD/SM	pValue
ALDOA	1,3	<0.001	3,2	<0.001
CPTM	1,58	0	3,5	<0.001
TNNC2	1,49	<0.001	Invalidé car trop faible dans le SM	
CEBPA	0,45	0,01	0,4	<0.001

Tableau 7: Comparaison des résultats d'expression différentielle par RT-PCR quantitative et puces. RT-PCR: n=8 par muscle; puces: n=104.

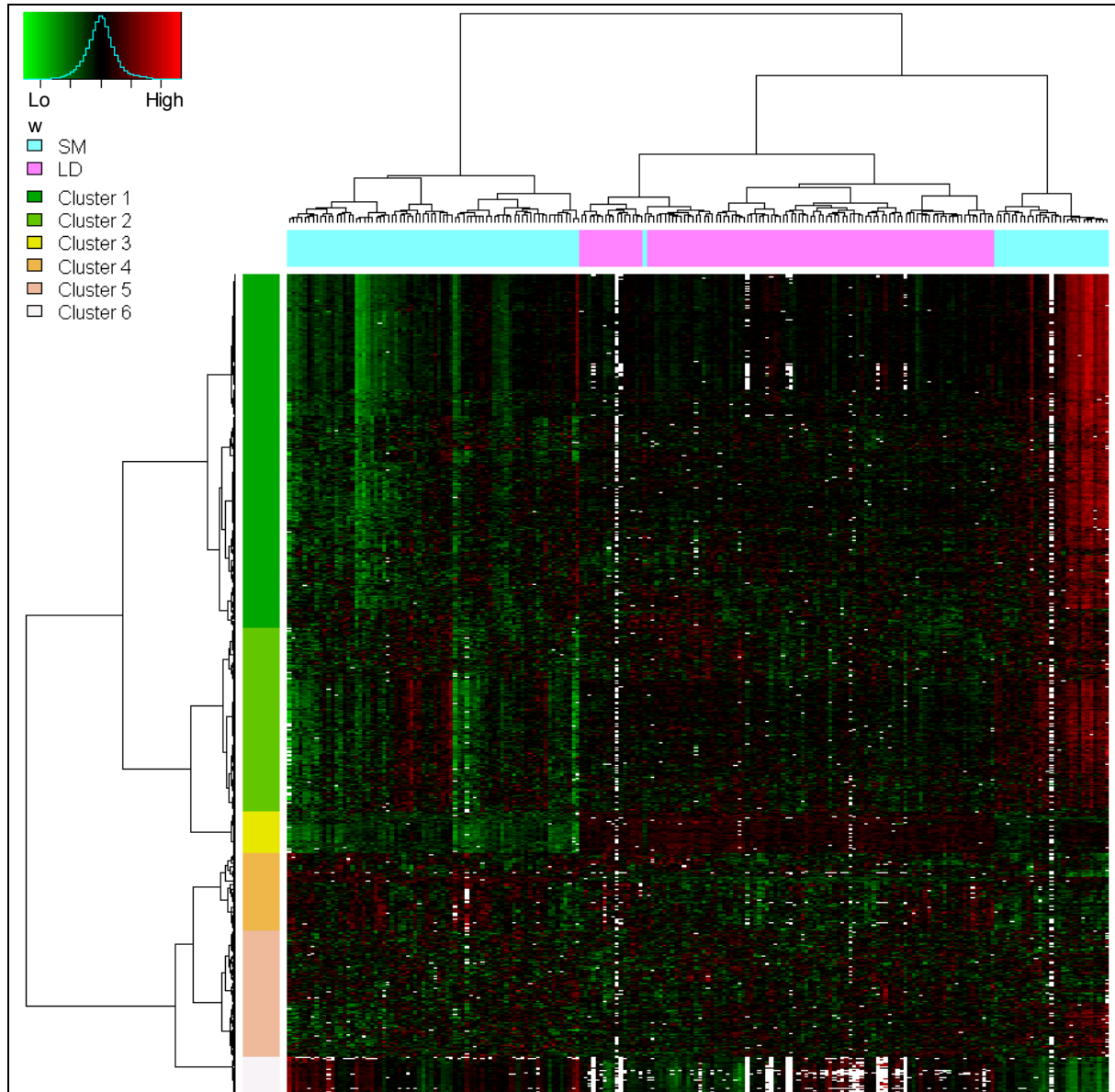


Figure 10: Double classification hiérarchique ascendante basée sur les corrélations, réalisées sur les sondes différemment exprimées entre le SM et le LD. Les sondes dont le différentiel d'expression est supérieur à 1,5 ont été utilisées.

Les gènes ALDOA, B2M-F, CPT-M, CEBPA et TNNC2 ont été analysés en RT-PCRQ dans les 2 muscles (n=8 par muscle). L'expression différentielle entre le SM et le LD est confirmée par cette autre méthode pour les gènes ALDOA, CPT-M, CEBPA et TNNC2 (Tableau 7). En revanche, pour ces 4 gènes les ratios d'expression entre les 2 muscles sont plus faibles que ceux obtenus avec les puces.

Une double classification ascendante hiérarchique utilisant les corrélations entre les niveaux d'expression des sondes, permet de regrouper les gènes en 6 clusters et les échantillons en 4 groupes différenciant nettement les muscles SM et le LD (Figure 10).

Les deux groupes SM (74 et 29 individus) présentent un fort contraste dans leur profil d'expression. Les clusters 1 et 2 présentent une forte variabilité d'expression au sein du SM alors qu'ils semblent moins variables pour le LD. Les clusters 1, 2 et 3 sont globalement plus fortement exprimés dans le LD que dans le SM. Les gènes des clusters 4 et 6 sont plus fortement exprimés dans le SM que dans le LD. Tandis que le cluster 5 ne présente pas de variation d'expression très forte entre les 2 muscles.

L'analyse des catégories fonctionnelles relative à ces 6 clusters de gènes a été réalisée avec le logiciel DAVID (Tableau 8). Les 2338 sondes montrant un profil d'expression différentielle entre les muscles représentent 1487 gènes, mais seuls 618 gènes se regroupent au sein de catégories fonctionnelles. Le cluster 6 ne présente pas d'enrichissement fonctionnel particulier. Les clusters 1, 2 et 3 montrent un enrichissement pour les catégories fonctionnelles représentant les récepteurs membranaires à domaine calcique et l'activité ATPase (cluster 1), les fonctions de régulation du cycle cellulaire et la transcription (cluster 2) et l'activité des phosphoprotéines (cluster 3). Le cluster 4 est enrichi dans les catégories fonctionnelles de réponse au stress, de transport ionique et de la matrice extra-cellulaire. Enfin le cluster 5 est enrichi pour les catégories fonctionnelles impliquées dans l'assemblage du cytosquelette et les modifications de la chromatine.

4. Discussion.

Les objectifs de cette étude étaient de mieux comprendre le déterminisme de la qualité du jambon par une approche combinée de détection de QTL sur des indicateurs de qualité et de eQTL et de déterminer les spécificités des profils d'expression de 2 muscles « modèles » pour la filière produit transformé (SM) et la filière viande fraîche (LD).

Nous avons mis en évidence 5 régions QTL dont une avec un fort effet de substitution allélique sur le L*. L'ensemble de ces 5 QTL montre une co-localisation des régions eQTL et QTL. Les gènes correspondant à ces eQTL sont impliqués dans les fonctions de

Cluster1					
		sondes 641 gènes 196			
Median Ease					
	Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	6,63E-04	INTERPRO	IPR000008:C2 calcium-dependent membrane targeting	9	4,59%
2	0,0026	INTERPRO	IPR015706:RNA-directed DNA polymerase (reverse transcriptase), related	3	1,53%
3	0,0127	GOTERM_MF_ALL	GO:0004386~helicase activity	13	6,63%
		GOTERM_MF_ALL	GO:0016887~ATPase activity		
4	0,0233	GOTERM_MF_ALL	GO:0016787~hydrolase activity	54	27,55%
5	0,0202	GOTERM_BP_ALL	GO:0060173~limb development	6	3,06%
6	0,0520	GOTERM_MF_ALL	GO:0004674~protein serine/threonine kinase activity	81	41,33%
		GOTERM_BP_ALL	GO:0043687~post-translational protein modification		
Cluster2					
		sondes 331 gènes 161			
Median Ease					
	Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	0,0053	GOTERM_BP_ALL	GO:0051726~regulation of cell cycle	23	14,29%
2	0,0150	GOTERM_BP_ALL	GO:0008380~RNA splicing	21	13,04%
3	0,0330	INTERPRO	IPR001623:Heat shock protein DnaJ, N-terminal	9	5,59%
		GOTERM_BP_ALL	GO:0006457~protein folding		
4	0,0173	GOTERM_MF_ALL	GO:0003676~nucleic acid binding	130	80,75%
Cluster3					
		sondes 76 gènes 39			
Median Ease					
	Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	0,0232	SP_PIR_KEYWORDS	phosphoprotein	31	40,79%
2	0,0226	GOTERM_CC_ALL	GO:0005622~intracellular	15	19,74%
Cluster4					
		sondes 141 gènes 91			
Median Ease					
	Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	1,04E-08	GOTERM_CC_ALL	GO:0031012~extracellular matrix	47	51,65%
2	0,0093	GOTERM_BP_ALL	GO:0006950~response to stress	24	26,37%
3	0,0060	GOTERM_BP_ALL	GO:0030198~extracellular matrix organization and biogenesis	13	14,29%
		SP_PIR_KEYWORDS	collagen		
		GOTERM_BP_ALL	GO:0006820~anion transport		
4	0,0244	SP_PIR_KEYWORDS	pyrrolidone carboxylic acid	7	7,69%
5	0,0179	GOTERM_MF_ALL	GO:0030414~protease inhibitor activity	8	8,79%
6	0,0370	GOTERM_BP_ALL	GO:0007155~cell adhesion	11	12,09%
Cluster5					
		sondes 229 gènes 97			
Median Ease					
	Score	Catégorie	Terme	nbre de gènes	% dans la catégorie
1	0,0125	INTERPRO	IPR001298:Filamin/ABP280 repeat	4	4,12%
2	0,0418	GOTERM_MF_ALL	GO:0018024~histone-lysine N-methyltransferase activity	5	5,15%
		GOTERM_BP_ALL	GO:0016568~chromatin modification		
Cluster6					
		sondes 69 gènes 34			

Tableau 8: Analyse de l'enrichissement en catégories fonctionnelles au sein des 6 clusters de gènes. Seuil 10%.

communication cellulaire, dans la matrice extra-cellulaire et des récepteurs membranaires à domaine calcique. Nous avons également montré des différences significatives de profil d'expression génique entre le SM et le LD. Six clusters de gènes différentiels représentent des groupes fonctionnels aussi variés que les récepteurs membranaires à domaine calcique, l'activité ATPase, la régulation du cycle cellulaire, la transcription, la réponse au stress, le transport ionique, la matrice extra-cellulaire, l'assemblage du cytosquelette et les modifications de la chromatine.

Notre dispositif expérimental ainsi que la méthodologie employée comportent des atouts et des inconvénients que je vais rappeler.

Analyse critique de l'étude.

Concernant la population expérimentale, souvent, la recherche de QTL pour un caractère donné s'appuie sur des lignées F0 possédant un phénotype divergent (Larzul *et al.*, 2008; Sanchez *et al.*, 2007; Wimmers *et al.*, 2006). Si le dispositif QTL utilisé lors de cette expérimentation n'est pas optimisé pour la recherche d'un QTL en particulier, il représente par contre l'intérêt de s'appuyer sur des animaux issus de lignées commerciales correspondant à une réalité en terme de sélection. En effet, la majorité des recherches de QTL a été réalisée avec des animaux expérimentaux (Davoli & Braglia, 2007).

Concernant la méthodologie employée, les données analysées dans cette étude correspondent à la quantification d'ARNm. Cependant, il est nécessaire de tenir compte du fait que les mécanismes moléculaires de la cellule ne sont pas linéaires. Le fait qu'un gène soit transcrit en ARNm n'implique pas que cet ARNm soit traduit en protéine ou encore que cette protéine soit fonctionnelle. Il faut tenir compte de l'épissage, de la stabilité moléculaire des ARNm mais également des différents processus de maturation post-traductionnelle qui confèrent aux protéines leurs propriétés (site de translocation, activité) et enfin des conditions environnementales physico-chimiques du milieu intra-cellulaire dans lequel se trouve la protéine et dont dépend son activité biochimique. Cette approche d'analyse du niveau d'expression des gènes n'est donc qu'une première étape dans la compréhension des mécanismes régulant la qualité des viandes, et devra être confirmée par des analyses protéiques et/ou biochimiques. De plus, la mesure de l'expression génique repose sur la spécificité de l'hybridation simultanée de transcrits avec un grand nombre de sondes. Cette spécificité est difficilement mesurable et repose également sur le soin apporté au choix des sondes. Dans cette étude, l'utilisation de puces commerciales incluant de nombreux contrôles « qualité » au cours de la procédure nous a permis sans doute de nous affranchir de difficultés

techniques. En effet, sur les 216 échantillons analysés sur puce aucun n'a été éliminé au cours de l'analyse.

Cependant, le séquençage et l'annotation du génome porcin sont encore à l'heure actuelle incomplets. Or, l'assignation des annotations géniques et fonctionnelles et la localisation physique de gènes candidats dépendent de ce séquençage. De ce fait, 25% des sondes ne sont pas annotées et une incertitude demeure sur l'annotation des sondes annotées en utilisant les contigs porcins SIGENAE.

Dans un premier temps, nous avons obtenu des résultats originaux grâce à la mise en place d'une analyse génétique combinant la détection de QTL et les mesures d'expression génique sur puce.

Localisation des QTL.

Nous avons mis en évidence 9 QTL influençant les indicateurs de qualité de la viande que sont la clarté (L*), la teinte rouge (a*), la teinte jaune (b*) et le pHu. Parmi ceux-ci, seul le QTL 13a se situe dans une région précédemment identifiée (animalgenome.org, 2009b; Harmegnies *et al.*, 2006; Hu *et al.*, 2005; van Wijk *et al.*, 2006). D'autre part sur les chromosomes 10 et 13 plusieurs QTL ont été localisés dans la même région.

L'analyse des QTL 10b et 10L est particulièrement intéressante car elle identifie une région chromosomique de 27 cM contrôlant l'expression de deux caractères en relation avec la qualité du jambon. De plus cette région possède un fort effet de substitution allélique pour le caractère L*. On peut donc penser y trouver un candidat positionnel contrôlant ces deux indicateurs de la qualité.

C'est également le cas sur le chromosome 13 où se situent des QTL pour le pHu, le a* et le b*. Les QTL 13pHu et 13a se chevauchent pour former une région QTL de 33 cM (55 à 88 cM). Le QTL 13b a quant à lui été identifié à 153 cM. Cependant, le profil des effets du QTL 13b suggère l'existence d'un deuxième QTL dans la région des QTL 13pHu et 13a. Il serait donc intéressant de tester l'hypothèse d'existence de deux QTL contre l'existence d'un seul QTL pour ce caractère et ainsi valider ou invalider l'hypothèse d'un QTL pleiotropique sur le chromosome 13 contrôlant a*, b* et pHu.

L'intérêt de tels QTL, influençant plusieurs indicateurs de la qualité, est d'identifier dans cette région un gène candidat placé particulièrement en amont dans la chaîne des événements conduisant à la mise en place de la qualité de la viande.

La taille des régions concernées par ces QTL demeure très large (22 cM en moyenne) ce qui représente environ 200 gènes. Afin de réduire cette taille et donc, le nombre de gènes

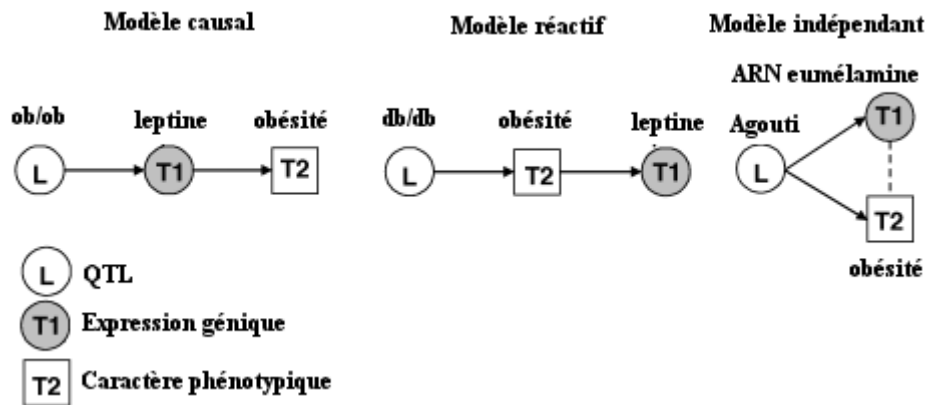


Figure 11: Représentation de trois modèles possibles de relations simples entre deux caractères quantitatifs subissant l'effet d'un même locus. Le modèle causal représente l'effet d'un QTL sur le phénotype T2 par l'intermédiaire du niveau d'expression du gène T1. Le modèle souris ob/ob est un exemple de ce type de relation où la diminution de la leptine provoque l'obésité. Le modèle réactif présente un effet du QTL sur le niveau d'expression du gène T1 en réponse à la variation du phénotype T2. Le modèle souris db/db en est un exemple. Dans ce modèle l'effet du QTL provoque une obésité ce qui par voie de conséquence induit une augmentation du niveau d'expression du gène de la leptine. Le modèle indépendant quant à lui représente l'effet d'un QTL agissant sur le caractère phénotypique et l'expression des gènes sans relation causale entre les deux. Le modèle souris agouti représente cette absence de relation causale entre le phénotype et l'expression d'un gène. Chez la souris agouti le QTL provoque une diminution de l'expression de l'eumélanine (modification de la couleur du pelage) et indépendamment une obésité de la souris (Adapté de Schadt, 2006).

candidats, il peut être envisagé de densifier la carte génétique et de génotyper les animaux pour de nouveaux marqueurs polymorphes situés dans cette région (microsatellites, haplotype SNP...). Si les animaux F1 sont informatifs pour ces marqueurs, il serait alors possible d'améliorer la résolution cartographique du QTL. Une autre solution consiste en l'obtention de nouvelles méioses informatives, en sélectionnant des reproducteurs hétérozygotes pour la région d'intérêt. Cette approche est cependant longue, coûteuse et d'une efficacité qui va en diminuant avec la taille de la région concernée, étant donné que 1 cM correspond à 1% de recombinaison au cours de la méiose.

C'est pourquoi nous avons choisi d'utiliser les données d'expression génique pour identifier des gènes candidats fonctionnels dans la région d'intérêt par la détection de eQTL co-localisés au QTL. Cette idée repose sur l'hypothèse que les gènes, dont le niveau d'expression est régulé par un eQTL localisé dans la même région qu'un QTL, participent à la régulation du phénotype ou subissent l'influence de ce mécanisme (Drake *et al.*, 2006). Nous avons ainsi identifié 755 eQTL. Par la suite nous nous sommes concentrés sur les eQTL qui co-localisaient avec les 5 régions QTL.

Caractérisation des eQTL: Cis ou Trans?

Au sein des eQTL il est possible de distinguer les cis-eQTL et les trans-eQTL. Les cis-eQTL sont des gènes physiquement localisés sur la région chromosomique correspondant au eQTL, ce qui implique que les variations d'expression sont dues à une régulation en cis du gène. Les trans-eQTL, qui co-localisent avec un QTL, sont des gènes localisés dans une autre région du génome. Dans ce cas, le gène responsable de la régulation en trans du e-QTL est localisé dans la région du eQTL. Ce gène causal n'est pas lui-même détecté en eQTL, soit parce qu'il s'agit d'une mutation ne modifiant pas son propre niveau d'expression, soit parce que l'effet du eQTL est trop faible, soit encore parce que le eQTL détecté n'est pas responsable de la variation du caractère étudié, mais est une conséquence de cette variation (Figure 11) (Drake *et al.*, 2006; Schadt *et al.*, 2005; Schadt, 2006).

Dans cette étude, nous n'avons pas pu mettre en évidence de cis-eQTL. Cependant, le séquençage et l'annotation du génome porcine sont, à l'heure actuelle, incomplets. Seuls sont disponibles des assemblages des chromosomes 1, 4, 5, 7, 11, 13, 14, 15, 17 et X (NCBI, 2009). Il pourrait donc être intéressant de rechercher la localisation de ces gènes sur d'autres génomes mieux annotés en utilisant la conservation de syntenie entre espèces (animalgenome.org, 2009a).

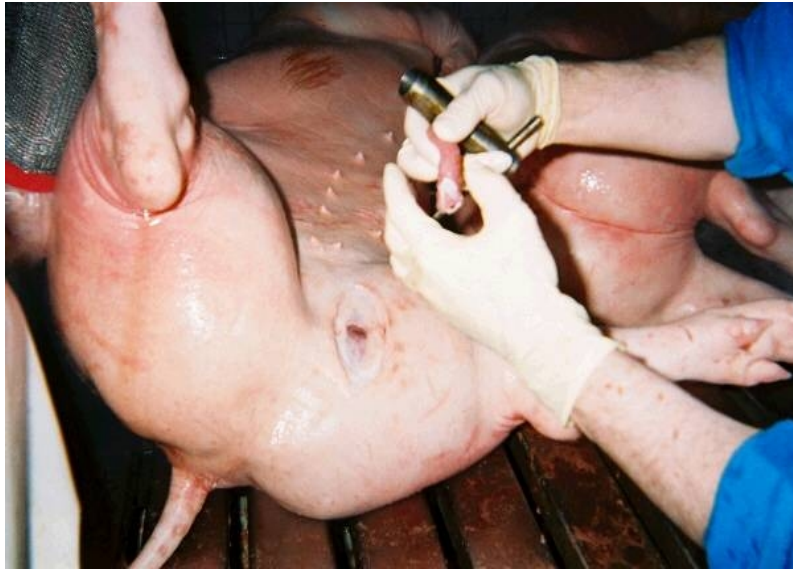


Figure 12 : Illustration du mode de prélèvement à l'aide d'un trocart.

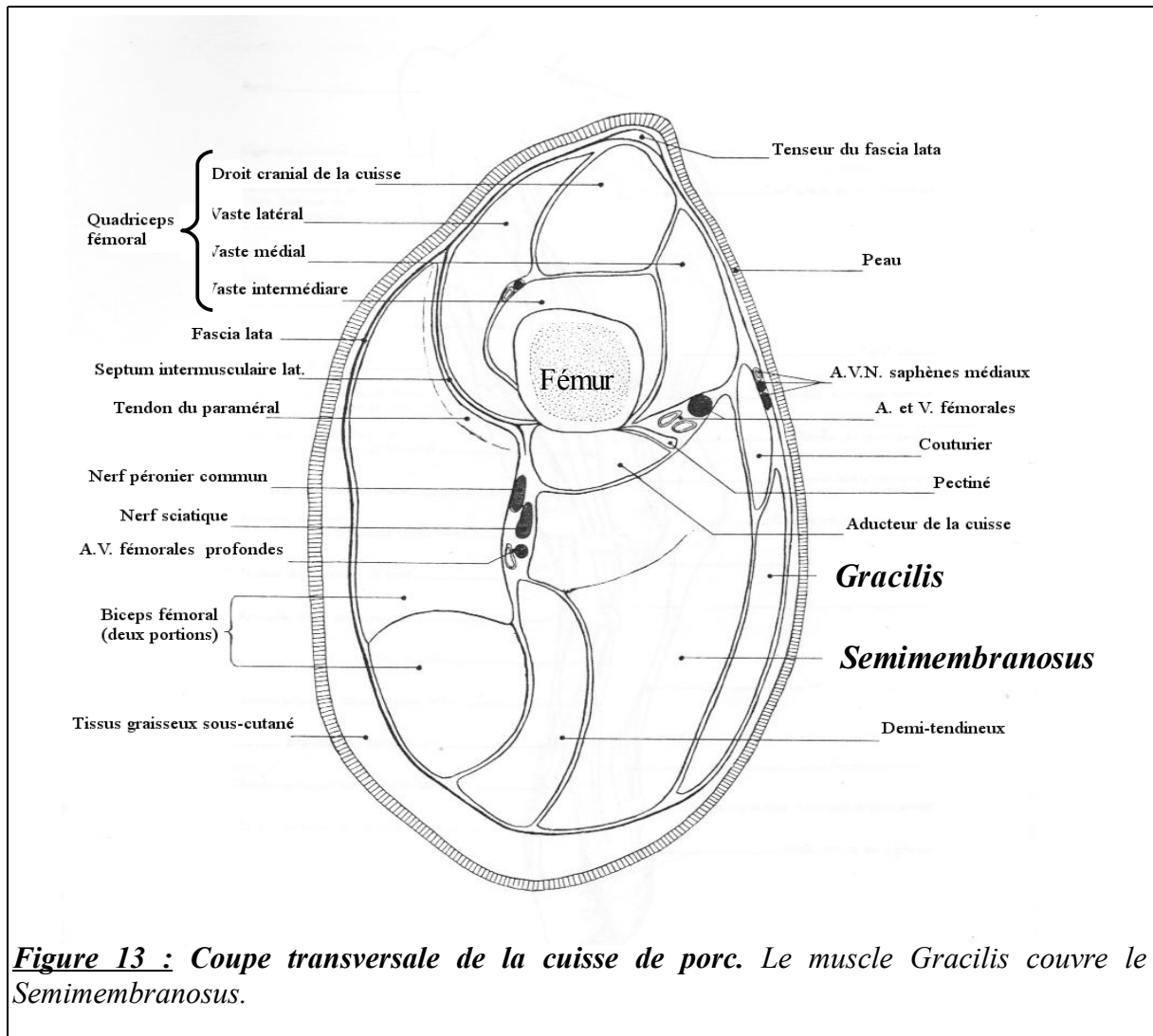


Figure 13 : Coupe transversale de la cuisse de porc. Le muscle Gracilis couvre le Semimembranosus.

Dans la deuxième partie de cette étude, nous avons clairement montré que les muscle SM et LD étaient différents.

Spécificités musculaires.

La comparaison des profils d'expression génique permet clairement de différencier les muscles LD et SM. Bien que ces muscles soient deux muscles blancs métaboliquement assez proches, leurs évolutions *post mortem* diffèrent (notamment au niveau du pHu). Des différences biochimiques entre ces 2 muscles ont été mises en évidence chez le porc (Purchas *et al.*, 2009 & Gil *et al.*, 2008). Ainsi, le pHu du LD est toujours inférieur au pHu du SM (-0.1 unité pH) (Chevillon *et al.*, 2006). En revanche, aucune analyse transcriptomique n'avait jusqu'ici été réalisée sur ces muscles car les études s'intéressaient à des muscles de types métabolique et contractile très distincts (Bai *et al.*, 2003 et Campbell *et al.*, 2001).

Au sein des échantillons de muscle SM une très forte distinction apparaît qui permet de séparer les échantillons SM en deux groupes dans des proportions 2/3 et 1/3. Cette différence semble plus forte que la différence observée entre les profils SM et LD. Nous émettons ici 2 hypothèses, qui pourraient expliquer cette différence.

La première repose sur la technique de prélèvement employée. Les prélèvements ont été effectués à l'aide d'un trocart directement sur la chaîne d'abattage (Figure 12), ce qui implique un prélèvement à un rythme assez élevé limitant la précision du site de prélèvement. L'anatomie de la cuisse de porc laisse clairement apparaître que le muscle *Gracilis* se situe au dessus du SM (Figure 13). Il est donc envisageable que certains prélèvements soient contaminé par du *Gracilis*.

L'autre hypothèse relève de la qualité même du muscle SM. Le SM est le principal muscle touché par le défaut de déstructuration des viandes (Laville *et al.*, 2005). Ce défaut se caractérise par une viande proche des viandes «PSE» avec un métabolisme glycolytique *post mortem* accru. La différence importante des profils d'expression entre les 2 muscles SM pourrait relever de la signature de ce type de défaut. En effet, les gènes ALDOA (impliqué dans la glycolyse) et TNNC2 (régulation par le calcium de la contraction musculaire) sont plus exprimés dans les muscles touchés par le défaut de déstructuration (Damon *et al.*, 2006) et semblent plus exprimés dans le SM ayant un profil d'expression proche du LD. Il serait donc intéressant de valider cette dernière hypothèse à l'aide d'une ACP pour voir si le pHu et le L* permettent d'identifier les individus suspectés d'être touchés par ce défaut.

Malgré l'utilisation d'une pValue corrigée, on peut s'attendre à l'existence de faux positifs. Nous avons donc choisi 5 gènes pour valider dans un premier temps les différentiels observés.

Ces 5 gènes ont été sélectionnés car les amorces étaient déjà disponibles au laboratoire. Avec cette technique, le gène de la bêta 2 microglobuline possède un niveau d'expression identique entre ces 2 muscles. Ce résultat concorde avec le fait que la beta 2 microglobuline est un gène de ménage susceptible d'être employé comme gène de référence (Erkens *et al.*, 2006; Nygard *et al.*, 2007). En revanche, la RT-PCR valide le résultat des puces pour les 4 autres gènes (ALDOA, TNNC2, CPT1M et CEBPA). La surexpression du gène ALDOA impliqué dans la glycolyse est en accord avec un métabolisme légèrement plus glycolytique du muscle LD comparé au muscle SM (Laborde, 1984). La troponine est une protéine du complexe actomyosine qui régule la contraction musculaire. La fixation du calcium sur la troponine C permet la levée de l'inhibition exercée par la troponine I sur l'activité ATPasique de la tête de myosine. La troponine C2 (TNNC2) est plus exprimée dans les muscles glycolytiques, tandis que la troponine C1 est prépondérante dans les muscles oxydatifs. L'expression plus importante de TNNC2 dans le LD est donc en accord avec le métabolisme glycolytique plus important de ce muscle et une disponibilité plus importante de calcium au travers d'une surexpression concomitante du gène RYR1 impliqué dans la libération de calcium dans le cytoplasme (cf plus bas). En revanche, la CPT1M est l'enzyme limitante de l'oxydation des acides gras dans la mitochondrie, sa surexpression dans le LD suggère une utilisation plus importante des lipides dans ce muscle. Le gène CEBPA est un facteur de transcription déterminant pour l'adipogenèse, sa surexpression dans le SM suggère qu'il y aurait plus d'adipocytes dans ce muscle. Le muscle LD contiendrait donc globalement moins de lipides que le SM car il posséderait moins d'adipocytes et serait plus apte à oxyder les lipides musculaires. Cependant, Purchas *et al.* (2009) ne montrent pas de différence en teneur en lipides entre ces 2 muscles.

Dans cette étude, nous avons effectué une analyse fonctionnelle en recherchant des catégories fonctionnelles associées à nos listes de gènes d'intérêt dans le but d'identifier les mécanismes moléculaires sous-jacents.

Les mécanismes mis en jeu.

Le phénomène de transformation du muscle en viande met en jeu des mécanismes biochimiques et physiques qui aboutissent à l'acidification de la viande lors de la synthèse d'ATP en condition anaérobie et à une modification de la couleur perçue suite aux effets de la perte en eau.

Les QTL 7b, 10b, 10L et 13pHu révèlent des groupes fonctionnels impliqués dans la matrice extra-cellulaire, la communication cellulaire, et les récepteurs membranaires possé-

nant un domaine de reconnaissance calcique. L'analyse des clusters de gènes différentiels met également en avant les voies de signalisation et les récepteurs membranaires à domaine calcique (sous exprimé dans le SM), ainsi que les mécanismes impliqués dans la réponse au stress et les structures de la matrice cellulaire (sur exprimé dans le SM) et les mécanismes d'assemblage du cytosquelette et de modification de la chromatine qui semblent peu variables entre le LD et le SM. Récemment, Ponsuksili *et al.* (2008) ont souligné l'importance de ces groupes fonctionnels dans le pouvoir de rétention en eau des viandes. L'importance de la matrice extra-cellulaire est corroborée par des études mettant en avant l'importance des protéines d'adhésion et de structure cellulaire dans la mise en place de la qualité de la viande. Lors de la phase de maturation de la viande, des protéases endogènes calcium-dépendantes (calpaïnes), dégradent les protéines de la plaque d'adhésion focale et plus globalement les protéines de structure et du cytosquelette des cellules musculaires (intégrines, desmines, vinculines, talines...) (Kristensen & Purslow, 2001; Lawson, 2004; Melody *et al.*, 2004). La plaque d'adhésion focale est un système qui assure la liaison entre la matrice extra-cellulaire et le cytosquelette. Les intégrines sont des protéines trans-membranaires de la plaque d'adhésion focale. Ces protéines assurent, d'une part la communication entre les milieux intra-cellulaire et extra-cellulaire, et d'autre part l'ancrage des cellules à la matrice extra-cellulaire (Moser *et al.*, 2009). La dégradation de ces protéines lors de la phase de maturation est impliquée dans certains critères de qualité de la viande (perte en eau, tendreté).

La voie de signalisation calcium possède une place particulièrement importante dans le contexte musculaire. L'ion calcium est l'un des éléments de régulation et de signalisation les plus importants dans le muscle. D'une part la concentration en calcium dans les fibres musculaires est l'un des facteurs limitant pour l'activité contractile des muscles, d'autre part le calcium est nécessaire à l'activité des protéases à calpaïnes. Le rôle important du calcium dans la mise en place de la qualité de la viande est particulièrement mis en avant par le défaut d'hypersensibilité au stress induit par le gène *Halothane*. Il s'agit d'une mutation du gène RYR1 (chromosome 6) qui code pour le récepteur à la ryanodine du muscle squelettique et dont le rôle est de contrôler les flux d'ions calcium sortant du réticulum sarcoplasmique. Sous l'effet du stress cette mutation provoque un relargage massif d'ions calcium, induisant une hyperactivité musculaire *post mortem* et une consommation accrue d'ATP aboutissant à une viande de type PSE.

L'ensemble de ces annotations, bien que soulignant l'importance de ces groupes fonctionnels dans les caractéristiques musculaires associées à la qualité des viandes, reste cependant très large et n'apporte malheureusement pas beaucoup d'informations

supplémentaires.

Au cours de cette étude, le gène RYR1 qui est un gène majeur de qualité de la viande responsable du défaut PSE (Hal), a particulièrement retenu notre attention.

Focus sur le gène RYR1

En effet, ce gène est nettement plus exprimé dans le LD que dans le SM (facteur 10). Dans le LD, une plus forte libération de calcium pourrait stimuler la contraction musculaire et conduire à un épuisement des réserves en ATP. Le métabolisme anaérobie glycolytique serait alors mis en jeu conduisant à un pHu plus bas du fait d'une acidification du milieu. Ce résultat est cohérent avec un pHu du LD plus faible que celui du SM.

De plus, ce gène appartient au cluster 1, cluster pour lequel les gènes sont très sur-exprimés dans 29 échantillons de SM. En admettant comme précédemment, que la sur-expression de ce gène augmente le nombre de récepteurs actifs et par la même le flux d'ions calcium, on peut supposer qu'en dehors de la mutation du gène Hal, un nombre important de récepteurs à la ryanodine provoquerait des lésions proches de celles des « viandes déstructurées » (hypercontractions des fibres et bas pHu) dans ces 29 échantillons SM en accord avec l'une de nos hypothèses sur ce sous-groupe SM.

CONCLUSION.

Ce travail avait pour objectif de mieux appréhender le déterminisme génétique de la qualité du jambon. L'analyse QTL nous a permis de mettre en évidence 8 nouveaux QTL, dont un possédant un fort effet de substitution allélique pour le L*. L'approche eQTL nous a permis d'identifier de nombreux eQTL co-localisant avec 5 de nos QTL précédemment, identifiés dont un grand nombre pour le QTL 10L.

Ces résultats permettent de mettre en évidence l'importance de cette région du génome porcin sur le caractère de clarté de la viande de porc, caractère fortement corrélé au pouvoir de rétention d'eau et à la perte en eau à la cuisson. Cependant il est nécessaire, d'une part de réduire la taille de cette région (27cM) et d'autre part d'analyser les gènes situés dans cette région, afin d'identifier des gènes candidats. Il faudra ensuite identifier dans les séquences de ces gènes candidats des mutations et analyser la co-ségrégation de ces mutations avec la mesure du L*.

Par ailleurs, nous avons analysé les profils d'expression génique des muscles LD et SM et ainsi mis en évidence une différence significative d'expression génique entre ces deux muscles. De plus, cette analyse nous a conduit à l'identification d'une sous-population au sein

du SM. Cependant, les catégories fonctionnelles différenciant ces 2 muscles sont très générales et peu informatives. Aussi le séquençage du génome porcin pourrait nous permettre d'avancer dans l'analyse fonctionnelle en améliorant notamment l'annotation de la puce.

Aussi, la recherche de bio-marqueurs de la qualité de la viande devra s'effectuer pour chacun de ces deux muscles. Cette conclusion rejoint celle de Huff-Lonergan *et al.* (2002) et de Chevillon & Vautier (2006) qui trouve une corrélation modérée entre les mesures de pHu des deux muscles et mettent en garde contre la valeur de prédiction modérée d'un marqueur unique de qualité pour ces deux muscles.

Références bibliographiques.

- animalgenome.org (page consultée le 06/06/09). **Alignment of Pig Linkage Map, RH Map, and Human Maps.**
http://www.animalgenome.org/cgi-bin/QTLdb/SS/draw_rh2human
- animalgenome.org (page consultée le 04/06/09). **QTL for CIE-A* in the Pig Genome.**
http://www.animalgenome.org/cgi-bin/QTLdb/SS/draw_traitmap?trait_ID=422
- Bertram, H. C., Schäfer, A., Rosenvold, K. & Andersen, H.J. (2004). Physical changes of significance for early post mortem water distribution in porcine *M. longissimus*. *Meat science*, 66, pp. 915-924.
- Chevillon, P. & Vautier, A. (2006). Peut-on prédire le pH ultime des longes et échine à partir de la mesure sur le jambon ?. *TechniPorc*, 29.
- Damon, M., Vincent, A., Cherel, P., Liaubet, L. & Le Roy, P. (2006). **Transcriptomic analysis of destructured ham.** In: *1st European conference on pig genomics*, Lodi, Italie
- Davoli, R. & Braglia, S. (2007). Molecular approaches in pig breeding to improve meat quality. *Briefings in functional genomics & proteomics*, 6, pp. 313-321.
- Dennis, G., Sherman, B., Hosack, D., Yang, J., Gao, W., Lane, H. C. & Lempicki, R. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, 4, p. P3.
- Drake, T. A., Schadt, E. E. & Lusis, A.J. (2006). Integrating genetic and gene expression data: application to cardiovascular and metabolic traits in mice. *Mammalian Genome*, 17, pp. 466-479.
- Erkens, T., Van Poucke, M., Vandesompele, J., Goossens, K., Van Zeveren, A. & Peelman, L. (2006). Development of a new set of reference genes for normalization of real-time RT-PCR data of porcine backfat and longissimus dorsi muscle, and evaluation with PPARGC1A. *BMC Biotechnology*, 6.
- Falconer, D. S. (1960). *Introduction to Quantitative Genetics*. 1ed., Oliver & Boyd, Edinburgh.
- Fischer, K. (2007). Drip loss in pork: influencing factors and relation to further meat quality traits. *Journal of animal breeding and genetics*, 124, pp. 12-18.
- Fox, J. (page consultée le 20/01/2009). **car: Companion to Applied Regression.**
<http://cran.cict.fr/web/packages/car/index.html>

- Franck, M., Monin, G. & Legault, C. (2000). **Observations complémentaires sur le jambon déstructuré : caractérisation du phénomène par le pH et la couleur du muscle semi-membraneux**. In: *32 èmes Journées de la Recherche Porcine*, Paris, 1-3 Février 2000, pp. 345-349.
- Hackstadt, A. & Hess, A. (2009). Filtering for increased power for microarray data analysis. *BMC bioinformatics*, 10.
- Harmegnies, N., Davin, F., De Smet, S., Buys, N., Georges, M. & Coppeters, W. (2006). Results of a whole-genome quantitative trait locus scan for growth, carcass composition and meat quality in a porcine four-way cross. *Animal Genetics*, 37, pp. 543-553.
- Hocquette, J. F., Ortigues-Marty, I., Damon, M., Herpin, P. & Geay, Y. (2000). Métabolisme énergétique des muscles squelettiques chez les animaux producteurs de viande. *INRA Productions Animales*, 13, pp. 185-200.
- Hocquette, J. F., Ortigues-Marty, I., Pethick, D., Herpin, P. & Fernandez, X. (1998). Nutritional and hormonal regulation of energy metabolism in skeletal muscles of meat-producing animals. *Livestock Production Science*, 56, pp. 115-143.
- Hu, Z., Dracheva, S., Jang, W., Maglott, D., Bastiaansen, J., Rothschild, M. F. & Reecy, J.M. (2005). A QTL resource and comparison tool for pigs: PigQTLDB. *Mammalian genome : official journal of the International Mammalian Genome Society*, 16, pp. 792-800.
- Huang, D. W., Sherman, B. T. & Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, pp. 44-57.
- Huff-Lonergan, E. & Lonergan, S.M. (2005). Mechanisms of water-holding capacity of meat: The role of postmortem biochemical and structural changes. *Meat science*, 71, pp. 194-204.
- Huff-Lonergan, E. & Lonergan, S.M. (2007). New frontiers in understanding drip loss in pork: recent insights on the role of postmortem muscle biochemistry. *Journal of animal breeding and genetics = Zeitschrift für Tierzucht und Zuchtungsbiologie*, 124, pp. 19-26.
- Huff-Lonergan, E., Baas, T. J., Malek, M., Dekkers, J. C., Prusa, K. & Rothschild, M.F. (2002). Correlations among selected pork quality traits. *Journal of animal science*, 80, pp. 617-627.
- IFIP (page consultée le 26/05/2009). **Qualité de la viande de porc**.
<http://www.itp.asso.fr/qual/esstech.htm>

IFIP (page consultée le 26/05/2009). **La charcuterie-salaison.**

<http://www.itp.asso.fr/qual/esscharsalai.htm>

- Kristensen, L. & Purslow, P.P. (2001). The effect of ageing on the water-holding capacity of pork: role of cytoskeletal proteins. *Meat science*, 58, pp. 17-23.
- Laborde, D. (1984). **Etude de facteur de variation du pouvoir de rétention d'eau de la viande de porc.** Docteur Ingenieur, Université de Clermont II U.E.R Sciences Exactes et Naturelles, Clermont - Ferrand, 99 p.
- Larzul, C., Bidanel, J. P., Iannuccelli, N., Gruand, J., Mourot, J. & Milan, D. (2008). **Détection de QTL pour la teneur en lipides intramusculaires et les activités enzymatiques lipogéniques dans un croisement entre porcs Meishan et Large White.** In: *40 èmes Journées de la Recherche Porcine*, Paris, 5 - 6 février 2008, pp. 127-130.
- Laville, E., Sayd, T., Santé-Lhoutellier, V., Morzel, M., Labas, R., Franck, M., Chambon, C. & Monin, G. (2005). Characterisation of PSE zones in semimembranosus pig muscle. *Meat Science*, 70, pp. 167-172.
- Lawson, M. A. (2004). The role of integrin degradation in post-mortem drip loss in pork. *Meat science*, 68, pp. 559-566.
- Le Roy, P., Elsen, J. M., Boichard, D., Mangin, M., Bidanel, J. P. & Goffinet, B. (1998). **An algorithm for QTL detection in mixture of full and half sib families.** In: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Australia, 11-16 janvier 1998, pp. 257-260.
- Lebret, B. (2004). Conséquences de la rationalisation de la production porcine sur les qualités des viandes. *INRA Productions Animales*, 17, pp. 79-91.
- Lefaucheur, L. (1989). Les différents types de fibres musculaires chez le porc. *INRA Productions Animales*, 2, pp. 205-213.
- Lefaucheur, L., Hoffman, R. K., Gerrard, D. E., Okamura, C. S., Rubinstein, N. & Kelly, A. (1998). Evidence for three adult fast myosin heavy chain isoforms in type II skeletal muscle fibers in pigs. *Journal of animal science*, 76, pp. 1584-1593.
- Melody, J. L., Lonergan, S. M., Rowe, L. J., Huiatt, T. W., Mayes, M. S. & Huff-Lonergan, E. (2004). Early postmortem biochemical factors influence tenderness and water-holding capacity of three porcine muscles. *Journal of animal science*, 82, pp. 1195-1205.

- Minvielle, B., Le Strat, P., Lebret, B. & Houix, Y. (2001). **Viandes déstructurées : situation dans cinq abattoirs de l'Ouest de la France ; facteurs de risque et proposition d'un modèle; caractérisation colorimétrique, biochimique et histologique.** In: *33 èmes Journées de la Recherche Porcine*, Paris, 30 Janvier - 1 Février 2001, pp. 95-101.
- Monin, G. (2003). Abattage des porcs et qualités des carcasses et des viandes. *INRA Productions Animales*, 16, pp. 251-262.
- Moser, M., Legate, K. R., Zent, R. & Fassler, R. (2009). The Tail of Integrins, Talin, and Kindlins. *Science*, 324, pp. 895-899.
- NCBI (page consultée le 02-06-09). **Sus scrofa (pig) genome view.**
http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=9823
- NIAID - NIH (page consultée le 06/06/09). **DAVID Bioinformatics Resources 2008.**
<http://david.abcc.ncifcrf.gov/home.jsp>
- Nygaard, A. B., Jørgensen, C. B., Cicera, S. & Fredholm, M. (2007). Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Molecular Biology*, 8.
- Olsson, V. & Pickova, J. (2005). The influence of production systems on meat quality, with emphasis on pork. *Ambio*, 34, pp. 338-343.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29, pp. 2002-2007.
- Ponsuksili, S., Jonas, E., Murani, E., Phatsara, C., Srikanchai, T., Walz, C., Schwerin, M., Schellander, K. & Wimmers, K. (2008). Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. *BMC genomics*, 9, pp. 367-381.
- Purchas, R., Morel, P., Janz, J. & Wilkinson, B. (2009). Chemical composition characteristics of the longissimus and semimembranosus muscles for pigs from New Zealand and Singapore. *Meat Science*, 81, pp. 540-548.
- R Development Core Team (2008a). **R: A Language and Environment for Statistical Computing.** R Foundation for Statistical Computing, Vienna, Austria.
<http://www.R-project.org>.
- R Development Core Team (page consultée le 22/12/2008). **R-2.8.1 for Windows.**
<http://cran.cict.fr/bin/windows/base/old/2.8.1/>
- Rosenvold, K. & Andersen, H.J. (2003). Factors of significance for pork quality--a review. *Meat science*, 64, pp. 219-237.

- Rothschild, M. F., Hu, Z. & Jiang, Z. (2007). Advances in QTL mapping in pigs. *International journal of biological sciences*, 3, pp. 192-197.
- Sanchez, M., Iannuccelli, N., Basso, B., Bidanel, J., Billon, Y., Gandemer, G., Gilbert, H., Larzul, C., Legault, C., Riquet, J., Milan, D. & Le Roy, P. (2007). Identification of QTL with effects on intramuscular fat content and fatty acid composition in a Duroc x Large White cross. *BMC Genetics*, 8, pp. 55-63.
- Schadt, E. E., Lamb, J., Yang, X., Zhu, J., Edwards, S., GuhaThakurta, D., Sieberts, S. K., Monks, S., Reitman, M., Zhang, C., Lum, P. Y., Leonardson, A., Thieringer, R., Metzger, J. M., Yang, L., Castle, J., Zhu, H., Kash, S. F., Drake, T. A., Sachs, A. & Lusis, A.J. (2005). An integrative genomics approach to infer causal associations between gene expression and disease. *Nature genetics*, 37, pp. 710-717.
- Schadt, E. E. (2006). Novel integrative genomics strategies to identify genes for complex traits. *Animal Genetics*, 37, pp. 18-23.
- Scheffler, T. L. & Gerrard, D.E. (2007). Mechanisms controlling pork quality development: The biochemistry controlling postmortem energy metabolism. *Meat science*, 77, pp. 7 16.
- Sellier, P., Bouix, J., Renand, G. & Molénat, M. (1992). Les objectifs et les critères de sélection : Les aptitudes bouchères : croissance, efficacité alimentaire et qualité de la carcasse. *INRA Productions Animales, Hors série 1992 - Eléments de génétique quantitative et application aux populations animales*, pp. 147-159.
- Sellier, P. (1998). **Genetics of meat and carcass traits**. IN: *The genetics of the pig*. CAB INTERNATIONAL; Wallingford; UK ,pp. 463-510.
- Valin, C. (1988). Différenciation du tissu musculaire. Conséquences technologiques pour la filière viande. *Reproduction, nutrition, development*, 28, pp. 845-856.
- van Wijk, H. J., Dibbits, B., Baron, E. E., Brings, A. D., Harlizius, B., Groenen, M. A. M., Knol, E. F. & Bovenhuis, H. (2006). Identification of quantitative trait loci for carcass composition and pork quality traits in a commercial finishing cross. *Journal of animal science*, 84, pp. 789-799.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3, pp. 1-11.
- Wimmers, K., Fiedler, I., Hardge, T., Murani, E., Schellander, K. & Ponsuksili, S. (2006). QTL for microstructural and biophysical muscle properties and body composition in pigs. *BMC genetics*, 7, pp. 15-28.

Annexes.

Annexe I

Pas de correction	Conditions d'hybridation
743 gènes	1663 gènes (25% des gènes)
Poids de carcasse	Poids de carcasse + conditions d'hybridation
310 gènes	700 gènes
Conditions d'abattage	Conditions d'abattage + conditions d'hybridation
187 gènes	551 gènes
Conditions d'abattage + poids de carcasse	Conditions d'abattage + poids de carcasse + conditions d'hybridation
118 gènes	234 gènes
Sexe	Sexe + conditions d'hybridation
211 gènes	1032 gènes (15% des gènes)
Sexe + poids de carcasse	Sexe + poids de carcasse + conditions d'hybridation
97 gènes	342 gènes
Sexe + conditions d'abattage	Sexe + conditions d'abattage + conditions d'hybridation
71 gènes	253 gènes
Sexe + conditions d'abattage + poids de carcasse	Sexe + conditions d'abattage + poids de carcasse + conditions d'hybridation
55 gènes	98 gènes

Corrections effectuées sur les données d'expression génique: 15 modèles de correction et les effectifs de gènes concernés.

Annexe II

Muscle	Muscles + Famille-hybridation par muscles
214 gènes	2355 gènes (34% des gènes)
Muscles + poids de carcasse	Muscles + poids de carcasse + Famille-hybridation par muscles
78 gènes	1171 gènes (17% des gènes)
Muscles + conditions d'abattage	Muscles + conditions d'abattage + Famille-hybridation par muscles
48 gènes	527 gènes
Muscles + conditions d'abattage + poids de carcasse	Muscles + conditions d'abattage + poids de carcasse + Famille-hybridation par muscles
19 gènes	350 gènes
Muscle + sexe	Muscle + sexe + Famille-hybridation par muscles
56 gènes	949 gènes
Muscle + sexe + poids de carcasse	Muscle + sexe + poids de carcasse + Famille-hybridation par muscles
26 gènes	681 gènes
Muscle + sexe + conditions d'abattage	Muscle + sexe + conditions d'abattage + Famille-hybridation par muscles
22 gènes	248 gènes
Muscle + sexe + conditions d'abattage + poids de carcasse	Muscle + sexe + conditions d'abattage + poids de carcasse + Famille-hybridation par muscles
8 gènes	112 gènes

Facteurs pris en compte dans le modèle d'analyse de variance : 16 modèles et les effectifs de gènes concernés.



Département : Sciences de l'animal.

Spécialisation

option

Enseignant responsable :

Tuteur scientifique :

Auteur(s) :
HÉRAULT Frédéric

Organisme d'accueil : INRA UMR SENAH

Adresse :

Nb pages : 36 Annexe(s) : 2

Domaine de la Prise

Année de soutenance : 2009

35590 Saint Gilles

Maître de stage : DAMON Marie (CR1 INRA)

LE ROY Pascale (DR2 INRA)

Titre :

« UTILISATION DE DONNÉES TRANSCRIPTOMIQUES POUR L'ÉTUDE DU DÉTERMINISME GÉNÉTIQUE DE LA QUALITÉ DU JAMBON CUIT. »

Résumé :

La qualité de la viande est un critère de choix important pour le consommateur et l'industriel. Actuellement, il n'existe pas de critère unique, fiable permettant d'estimer la qualité de la viande. Celle-ci est influencée par des facteurs d'ordre environnementaux ou génétiques. On estime que 10 à 30 % de la variation de la qualité de la viande est sous l'influence de facteurs génétiques.

L'objectif de cette étude est d'une part de mieux comprendre le déterminisme génétique de la qualité du jambon et sa relation avec les caractéristiques musculaires et d'autre part de déterminer les spécificités des profils d'expression de deux muscles le *Longissimus dorsi* et le *Semimembranosus*.

Dans un premier temps une approche combinée de détection de QTL et de eQTL a été utilisée pour identifier des loci responsables de la variation de la qualité de la viande de porc. Nous avons ainsi identifié huit QTL impliqués dans la variabilité du pHu et de la couleur. Le QTL détecté sur le chromosome 10 pour le caractère L* a un fort effet de substitution allélique.

De plus, nous avons identifié de nombreux eQTL co-localisés à ces 5 QTL. La cartographie fine des régions d'intérêt est maintenant nécessaire pour localiser les gènes candidats. Ces approches de cartographie fine bénéficieront des avancées technologiques faites en terme de génotypage et de la finalisation du séquençage du génome porcin.

Dans la seconde partie de cette étude, nous avons montré l'existence de profils d'expression génique divergeant entre les muscles *Longissimus dorsi* et *Semimembranosus*. Cette différence d'expression génique met en évidence la valeur de prédiction modérée que l'on peut attendre d'un marqueur unique de qualité pour ces deux muscles.

Abstract :

Meat Quality is of critical importance for both consumers and processing industry. Currently, it is really difficult to measure this complex trait in a unique manner. The quality of raw pig meat is influenced by a large number of genetic and non-genetic factors including rearing conditions, transport, slaughtering and meat processing. Genetic effects play a crucial role in determining pig meat quality since 10% to 30% of the variation in meat quality traits is determined by the genetic background of the animal.

This study aims to identify genetically determined variations in ham quality and their genetic relationships with muscle characteristics, and to characterized differentially expressed genes between two types of muscles *Longissimus dorsi* and *Semimembranosus*.

First, both positional (QTL research) and expressional (eQTL) genomics were intended to identify loci involved in the quality of pork meat. Thus, 8 QTLs involved in the variation of ultimate pH and colour have been identified. One of them is involved in variation of « L » and shows an important effect of allele substitution.

Moreover we have identified many eQTLs co-localised with QTL. Fine mapping will be now required to identify more accurately these potential regulatory genes. This future research will benefit from the progress made in genotyping technologies (with the access to high-density SNP markers) and from the Porcine Genome Sequencing Project.

In the second part of this study we have identified divergent gene expression pattern between *Longissimus dorsi* and *Semimembranosus*. This divergent pattern clearly demonstrate moderate reliability in predicting the quality of one muscle based on the gene expression profiling of another one.

Mots-clés :

QTL, eQTL, qualité de la viande, type musculaire, transcriptomique.