



**HAL**  
open science

# Diversité génétique structurale et fonctionnelle du CMH chez le Poulet : Implication pour la résistance aux maladies

Olympe Chazara Trokiner

► **To cite this version:**

Olympe Chazara Trokiner. Diversité génétique structurale et fonctionnelle du CMH chez le Poulet : Implication pour la résistance aux maladies. Génétique animale. AgroParisTech, 2010. Français. NNT : 2010AGPT0025 . pastel-00601989

**HAL Id: pastel-00601989**

**<https://pastel.hal.science/pastel-00601989v1>**

Submitted on 21 Jun 2011

**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.

## THÈSE

pour obtenir le grade de

**Docteur**

de

**l'Institut des Sciences et Industries du Vivant et de l'Environnement  
(Agro Paris Tech)**

Spécialité : Génétique Animale

*présentée et soutenue publiquement  
par*

**Olympe CHAZARA**

**le 16 Mars 2010**

**Diversité génétique structurale et fonctionnelle du CMH chez le  
poulet : implication pour la résistance aux maladies**

*Directeur de thèse : Michèle TIXIER-BOICHARD*

*Codirecteur de thèse : Bertrand BED'HOM*

*INRA, UMR 1313 Génétique Animale et Biologie Intégrative, F-78352 Jouy-en-Josas*

Devant le jury :

|   |                   |
|---|-------------------|
| <b>M Pierre BOUDINOT, DR, INRA</b> .....                            | <b>Président</b>  |
| <b>M Olivier HANOTTE, Professor, University of Nottingham</b> ..... | <b>Rapporteur</b> |
| <b>M Pierre PONTAROTTI, DR, CNRS</b> .....                          | <b>Rapporteur</b> |
| <b>M Philippe LE BOUTEILLER, DR, INSERM</b> .....                   | <b>Examineur</b>  |
| <b>M Bertrand BED'HOM, IR, INRA</b> .....                           | <b>Examineur</b>  |
| <b>M Thomas HEAMS, MC, AgroParisTech</b> .....                      | <b>Examineur</b>  |



# REMERCIEMENTS

Je souhaiterais remercier en premier lieu mes chers directeurs de thèse, Michèle Tixier-Boichard et Bertrand Bed'hom pour m'avoir incité à accepter cette grande aventure. Et plus précisément Michèle, pour m'avoir fait partager sa passion des poules diverses et variées et Bertrand, pour avoir été disponible à tout moment pour discuter de mes derniers résultats et pour réussir à faire toujours preuve d'une grande rigueur scientifique et personnelle.

Merci également à Xavier Rognon et Rima Zoorob qui m'ont acceptée dans leur laboratoire, en stage, alors que mon cursus ne m'y conduisait pas et qui m'ont fait découvrir leur domaine d'expertise respectifs, la génétique des populations et la biologie moléculaire.

Merci à Marie-Hélène Pinard-van der Laan, Claire Rogel-Gaillard, Rima Zoorob, Thierry Robert et Thomas Heams d'avoir été les précieux membres de mon comité de thèse. Je remercie particulièrement Marie-Hélène pour ses conseils avisés et l'accès aux lignées Réponse Immune.

Merci aux différents membres des équipes dans lesquelles j'ai pu travailler, en particulier l'unité GDA, devenue équipe PSGen, ses jovaciens: Agathe, Marie-Hélène, Sandrine, Tatiana, Thu et Wendy, André et André, Francis mon Sensei, Fred, Mathieu et Thierry ; et ses parisiens: Etienne, Grégoire, Thomas et Xavier. Mais également à l'équipe GIS et ses différents membres, du bâtiment 440 ou de la ferme, pour nous avoir accueillis à bras ouverts. A special thought also for the Danish chicks' team led by Helle for welcoming me warmly in her lab in Foulum.

Un énorme merci au soutien technique infailible de Nicolas et Jean-Luc, mais aussi Hanne, Véro et Anne, ainsi qu'à leur patience...

Merci aux différents étudiants avec qui j'ai pu partager les joies de la vie de labo, Lucille, Claire, Marie-Laure, Roopika, Rui et Valérie, Chung-Ming, Grégoire, Jean-Claude, Thomas, et mes camarades de tous les instants, Giulietta, ma grande soeur de recherche et Chi-Sheng, my favorite Taiwanese !

Enfin, merci à ceux qui m'ont transmis leur passion pour la Biologie, c'est-à-dire mes parents, ma famille et mes chères professeures de Bio. Cependant, sans le soutien indéfectible de mes deux conseillers en tout, Aurélia et Benjamin, rien de tout ceci n'aurait abouti.



# SOMMAIRE

|  |                  |
|--|------------------|
| <b>RESUME</b>  | <b>5</b>         |
| <b>ABSTRACT</b>  | <b>6</b>         |
| <b>LISTE DES ILLUSTRATIONS</b>   | <b>7</b>         |
| <b>INTRODUCTION GENERALE</b>   | <b>9</b>         |
| <b><u>I. DIVERSITE GENETIQUE DU POULET ET AVICULTURE</u></b>                       | <b><u>9</u></b>  |
| <b><u>II. IMMUNOGENETIQUE AVIAIRE</u></b>  | <b><u>14</u></b> |
| <b><u>III LE COMPLEXE MAJEUR D'HISTOCOMPATIBILITE CHEZ LES OISEAUX</u></b>         | <b><u>19</u></b> |
| <b>A. Généralités</b>  | <b>19</b>        |
| <b>B. Diversité interspécifique et dynamique évolutive</b>                         | <b>22</b>        |
| <b>C. Le CMH du poulet</b>   | <b>31</b>        |
| 1. Historique du typage du CMH chez le poulet                                      | 31               |
| 2. Gènes de la région B du CMH du poulet   | 37               |
| 3. Autres gènes du CMH du poulet   | 45               |
| <b><u>IV. ROLE DU CMH DANS LA RESISTANCE AUX MALADIES CHEZ LE POULET</u></b>       | <b><u>46</u></b> |
| <b><u>V. PRESENTATION DE LA THESE</u></b>  | <b><u>50</u></b> |
| <br>   |                  |
| <b><u>MATERIEL &amp; METHODES</u></b>  | <b><u>51</u></b> |
| <b><u>I. MATERIEL ANIMAL</u></b>   | <b><u>51</u></b> |
| <b>A. Animaux utilisés tout au long du projet de thèse : les lignées R.I.</b>      | <b>51</b>        |
| <b>B. Animaux utilisés pour le reséquençage du CMH</b>                             | <b>53</b>        |
| 1. Echantillonnage AvianDiv  | 53               |
| <i>a) 7 lignées commerciales</i>   | 55               |
| <i>b) 8 lignées expérimentales</i>   | 55               |
| <i>c) 14 races provenant d'Europe ou du pourtour Méditerranéen</i>                 | 57               |
| 2. Echantillonnage complémentaire  | 60               |
| <b>C. Autres races utilisées pour les études de diversité du CMH dans la thèse</b> | <b>62</b>        |
| <b>D. Espèces sauvages</b>   | <b>63</b>        |
| <b><u>II. METHODES GENERALES UTILISEES</u></b>                                     | <b><u>64</u></b> |
| <b>A. Méthodes de biologie moléculaire</b>   | <b>64</b>        |
| <b>B. Méthodes de bioinformatique</b>  | <b>66</b>        |

|   |            |
|---|------------|
| <b>RESULTATS</b>  | <b>67</b>  |
| <hr/>   |            |
| <b><u>I PREMIERE PARTIE : Exploration de la diversité du CMH via le marqueur LEI0258 dans les collections maintenues au laboratoire et inférences sur le mode d'évolution de ce marqueur atypique</u></b>   | <b>67</b>  |
| A. Introduction   | 67         |
| B. Article de recherche   | 69         |
| C. Résultats complémentaires et Discussion  | 91         |
| <b><u>II DEUXIEME PARTIE : Etude de la corrélation entre génotypage et sérologie chez les lignées expérimentales White Leghorn R.I., à l'aide du marqueur LEI0258 et d'un jeu réduit de SNPs</u></b>  | <b>95</b>  |
| A. Introduction   | 95         |
| B. Article de recherche   | 97         |
| C. Discussion et Conclusion   | 105        |
| <b><u>III TROISIEME PARTIE : Identification et caractérisation de SNPs dans la région B du CMH du poulet par reséquencage de 11 gènes cibles</u></b>  | <b>107</b> |
| A. Introduction   | 107        |
| B. Article  | 108        |
| C. Résultats complémentaires et Discussion  | 118        |
| <b><u>IV QUATRIEME PARTIE : Caractérisation de la structure génomiques des trois gènes particuliers du CMH du poulet correspondant aux gènes <i>DM</i> des Mammifères et étude de leur polymorphismes dans trois espèces de <i>Gallus</i> sauvages et un large panel de races expérimentales, commerciales et locales</u></b> | <b>123</b> |
| A. Introduction   | 123        |
| B. Article de recherche   | 125        |
| C. Discussion et Conclusion   | 163        |
| <b><u>V. CINQUIEME PARTIE : Etude exploratoire des gènes de classe I classiques par clonage</u></b>   | <b>165</b> |
| <hr/>   |            |
| A. Introduction   | 165        |
| B. Rapport  | 166        |
| C. Discussion et Conclusion   | 177        |
| <b><u>VI SIXIEME PARTIE : Mise au point d'une puce de 96 SNPs dédiée pour la région B du CMH du poulet</u></b>  | <b>179</b> |
| A. Introduction   | 179        |
| B. Article en préparation   | 180        |
| C. Discussion   | 185        |

|   |            |
|---|------------|
| <b><u>DISCUSSION GENERALE &amp; PERSPECTIVES</u></b>  | <b>187</b> |
| <b><u>I. APPORT DES DIFFERENTES STRATEGIES CHOISIES POUR L'ETUDE DU CMH DU POULET</u></b>       | <b>187</b> |
| <b><u>II. L'ETUDE DU CMH COMME MOYEN D'ANALYSE DE LA DIVERSITE GENETIQUE CHEZ LE POULET</u></b> | <b>192</b> |
| <b><u>III. L'EVOLUTION DU CMH</u></b>   | <b>192</b> |
| <b><u>IV. RESISTANCE AUX MALADIES ET CMH CHEZ LE POULET</u></b>                                 | <b>198</b> |
| <b><u>V. PERSPECTIVES</u></b>   | <b>200</b> |
| <br>  |            |
| <b>ANNEXES</b>  | <b>217</b> |





## RESUME

Le complexe majeur d'histocompatibilité (CMH) est une région génomique complexe des Vertébrés, encore imparfaitement connue chez le poulet, qui présente à certains locus une très grande variabilité génétique et qui joue un rôle central dans l'organisation de la réponse immunitaire d'un animal aux pathologies infectieuses. Le CMH est également une région de choix pour étudier le déterminisme génétique de l'adaptation aux agents pathogènes dans un contexte évolutif. De plus une meilleure connaissance du déterminisme génétique de la réponse immunitaire contre les agents pathogènes est un atout important pour développer une stratégie globale de lutte contre les maladies infectieuses. Nous avons donc entrepris d'utiliser les récents outils de la génomique, notamment des marqueurs génétiques de type SNP (Single Nucleotide Polymorphism) afin de caractériser la région B du CMH du poulet. En effet, c'est cette région qui a été essentiellement impliquée dans des phénomènes de résistance/susceptibilité aux maladies infectieuses.

En premier lieu, la diversité génétique a été évaluée dans plus de 80 races ou populations en utilisant le marqueur LEI0258 dont la complexité a nécessité le séquençage des différents allèles identifiés ce qui a permis une caractérisation en détails des races ou populations étudiées mais également de mieux comprendre les mécanismes d'évolution de ce marqueur atypique. La corrélation entre le typage sérologique traditionnel du CMH et le typage moléculaire a ainsi pu être examinée dans les lignées expérimentales INRA Réponse Immune. Ensuite, afin de couvrir l'ensemble de la région à l'aide de marqueurs SNPs, nous avons choisi d'identifier ces polymorphismes par reséquençage de 11 gènes d'intérêt et par comparaison des séquences obtenues entre elles et avec la séquence du génome et les séquences de référence disponibles dans les bases de données. Les animaux utilisés pour le reséquençage étaient de diverses origines, afin de représenter la diversité des ressources existantes pour le poulet, il s'agissait notamment de races locales, pouvant présenter des caractères intéressants d'adaptation au milieu mais pour lesquelles le recueil de données expérimentales est difficile. En parallèle, ces travaux ont également permis d'améliorer la connaissance de certains gènes, notamment trois gènes de classe II non classiques de type DM. Enfin, un travail exploratoire a été mené afin d'évaluer les possibilités d'étude des gènes de classe I classiques, qui n'avaient pas pu être reséquencés directement en raison de leurs caractéristiques particulières. Une puce de 96 SNPs dédiée à la région B du CMH du poulet a été produite et devrait rapidement permettre d'exploiter des études d'infections expérimentales réalisées à l'INRA.

## ABSTRACT

The major histocompatibility complex (MHC) is a complex genomic region in Vertebrates, still imperfectly known in the chicken and which shows a great genetic variability. The MHC is also an interesting region for studying the genetic determinism of adaptation to pathogens in an evolutionary context. Moreover, the MHC plays a central role in the immune response of an animal to infectious diseases, while a better understanding of the genetic determinism of the immune response against pathogens is important for developing a comprehensive strategy to fight against infectious diseases.

We therefore used recent tools of genomics, including genetic markers such as SNPs (Single Nucleotide Polymorphism) to characterize the B region of the chicken MHC. This region has been mainly involved in observed resistance / susceptibility to infectious diseases.

First, genetic diversity has been evaluated in more than 80 breeds or populations with the LEI0258 marker, whose complexity required the sequencing of the different alleles identified, which made possible a detailed characterization of the breeds or populations studied but which also led to a better understanding of the evolution mechanisms of this unusual marker. The correlation between the traditional typing of the MHC by serology and its molecular typing was subsequently examined in the experimental INRA lines Immune Response. Then, to cover the entire region using SNP markers, we chose to identify these polymorphisms by resequencing 11 genes of interest and comparing the sequences obtained with the genome sequence and reference sequences available in databases. The animals used for resequencing were of various origins, to represent the diversity of resources available for chicken, including local breeds, which may have interesting adaptive traits but for which the record of experimental data is difficult. It also led to the improvement of the knowledge of a number of genes, including three DM-like non-classical class II genes. Finally, an exploratory work was conducted to assess the possibilities of studying classical class I genes, which were not resequenced directly because of their characteristics. A 96 SNPs chip, dedicated to the B region of the chicken MHC, has been produced and will soon provide the genotypes of a number of infectious challenge studies conducted at INRA.

# LISTE DES ILLUSTRATIONS

|  |            |
|--|------------|
| <b>INTRODUCTION GENERALE</b>   | <b>9</b>   |
| <b>Figure 1.</b> Les flux mondiaux de viande de volaille en 2005   | 11         |
| <b>Figure 2.</b> Mode de génération par conversion génique du répertoire d'anticorps chez le poulet (Janeway <i>et al.</i> , 2008)   | 16         |
| <b>Figure 3.</b> Carte du CMH de l'Homme (Horton <i>et al.</i> , 2004)   | 20         |
| <b>Figure 4.</b> Carte comparée du CMH de la caille japonaise (quail) et du poulet (fowl) (Shiina <i>et al.</i> , 2006)  | 24         |
| <b>Figure 5.</b> Modèle d'organisation du microchromosome GGA16 (Delany <i>et al.</i> , 2009)  | 30         |
| <b>Figure 6.</b> Carte du CMH du poulet présentant les différents gènes annotés (d'après Shiina <i>et al.</i> , 2007)  | 36         |
| <b>Figure 7.</b> Réponse vaccinale pour la Maladie de Marek pour 3 lignées congéniques de poulet différant par leur sérotype au CMH. Après vaccination, la taille de l'histogramme représente l'effet protecteur du vaccin contre la mort ou les lésions causées par le virus. Quelque soit le vaccin utilisé, certains sérotypes ont mieux répondu. ( Lee <i>et al.</i> , 2004) | 48         |
| <b>Tableau 1.</b> Comparaison de 13 sérotypes B (d'après Simonsen <i>et al.</i> , 1982)  | 34         |
| <b>Tableau 2.</b> Comparaison des gènes de la région B du CMH du poulet avec les gènes d'autres espèces (Shiina <i>et al.</i> , 2007)  | 44         |
| <b>MATERIEL &amp; METHODES</b>   | <b>51</b>  |
| <b>Figure 8.</b> Réponse à la sélection pour ND3 (réponse anticorps) de la lignée 1, pour PHA (réponse cellulaire) de la lignée 2 et pour CC (phagocytose) de la lignée 3, représentée par la moyenne des valeurs génétiques estimées (Pinard-van der Laan 2002)   | 52         |
| <b>Tableau 3.</b> Principales caractéristiques des populations étudiées  | 54         |
| <b>RESULTATS</b>   | <b>67</b>  |
| <b>Figure 9.</b> Représentation graphique des valeurs d'hétérozygotie observée, en abscisse, au niveau du marqueur LEI0258, en ordonnée, au niveau de marqueurs microsatellites répartis ailleurs dans le génome   | 92         |
| <b>Figure 10.</b> Comparaison des dendrogrammes obtenus en Neighbour Joining pour le gène <i>CD1A1</i> (à gauche) et le gène <i>TRIM27.1</i> situés aux extrémités de la zone centrale de la région B du CMH du poulet   | 121        |
| <b>Tableau 4.</b> Résultats de l'analyse de corrélation de Spearman (Rho) et du test des rangs signés de Wilcoxon (W) pour les données en totalité ou les différentes sources, incluant ou excluant les populations expérimentales et commerciales   | 93         |
| <b>Tableau 5.</b> Comparaison des caractéristiques, en terme de polymorphismes, des 24 gènes composant la zone centrale de la région B du CMH du poulet  | 119        |
| <b>DISCUSSION GENERALE &amp; PERSPECTIVES</b>  | <b>187</b> |
| <b>Tableau 6.</b> Récapitulatif des informations recueillies pour vingt principaux sérotypes ou haplotypes   | 190        |



# INTRODUCTION GENERALE

Le principal objectif de cette thèse est d'explorer la diversité génétique du complexe majeur d'histocompatibilité, CMH, chez le poulet. Cette exploration a pour but de mieux connaître l'organisation et les caractéristiques de la région mais également de mettre au point une méthode universelle de génotypage moléculaire du CMH chez le Poulet, et d'évaluer les possibilités d'implémentation de ces techniques dans des dispositifs expérimentaux.

Les connaissances actuelles sur la diversité génétique du poulet domestique seront d'abord rappelées afin de décrire le contexte des populations étudiées. Ensuite, le CMH sera présenté sous l'angle de sa dynamique évolutive chez les Oiseaux, avant d'aborder de façon détaillée l'état des connaissances sur la structure du CMH du poulet et son implication dans la résistance aux maladies.

Etant donné le rôle majeur mis en évidence pour la région B au sein du CMH, la suite du travail s'est focalisée sur la caractérisation de cette région. Les différentes étapes du travail sont introduites à l'issue de la revue bibliographique

## **I. DIVERSITE GENETIQUE DU POULET ET AVICULTURE**

Le poulet, *Gallus gallus*, est un oiseau domestique appartenant à l'ordre des Galliformes. Les milliards de poulets élevés par an au niveau mondial constituent une source très importante de protéines pour l'alimentation humaine, grâce à leur production d'œufs et de viande. En 2007, la production mondiale de viande de poulet était de 76 millions de tonnes (environ 28% de la production totale de viande dans le monde), et la production totale d'œufs au monde était de 59 millions de tonnes (<http://faostat.fao.org>). La majorité de ces productions est obtenue à partir d'un nombre limité de lignées commerciales sélectionnées de manière intensive par un très petit nombre d'entreprises d'élevage, dont la plupart sont internationales. Il existe également un large éventail de races de poules dans le monde, phénotypiquement très diversifiées, à la suite du long processus de domestication. La diversité génétique des lignées commerciales et d'autres populations a été évaluée avec des marqueurs microsatellites (Hillel *et al.*, 2003; Granevitze *et al.*, 2007), elle est faible pour les pondeuses d'œufs blancs (hétérozygotie attendue,  $H_e$ , de 0,35 à 0,40), plutôt faible pour les pondeuses d'œufs bruns

(He de 0,45 à 0,50), modérée à élevée pour les poulets de chair (He de 0,55 à 0,60) et toujours plus élevée (He de 0,60 à 0,65) pour les populations non sélectionnées et pour l'espèce sauvage Coq Bankiva (*Gallus gallus*). Une étude récente et plus détaillée basée sur un génotypage SNP indique que 50% de la diversité génétique ancestrale a été perdue dans les lignées commerciales (Muir *et al.*, 2008), alors que les races locales présentent des niveaux d'hétérozygotie variables, en fonction de leur histoire et leur taille de population. Un nombre croissant de races locales sont menacées d'extinction, le risque pour l'espèce étant la perte de leurs caractères spécifiques et de combinaisons alléliques originales (Blackburn, 2006). Les diverses races de poules domestiques actuelles ont pour origine largement admise les espèces sauvages de poules d'Asie du Sud-est, et principalement le Coq Bankiva. La contribution majeure de cette espèce a été démontrée au niveau mitochondrial (Fuhimito *et al.*, 1996). Cependant il a été récemment montré que le caractère phénotypique peau jaune des poulets domestiques, largement présent dans de nombreuses lignées commerciales, pondeuses ou de chair, provient du Coq de Sonnerat (*Gallus sonneratii*) (Eriksson *et al.*, 2008). En outre, les poules de jungle sauvages sont toujours présentes en Asie et des flux de gènes peuvent avoir lieu entre ces animaux sauvages et les poulets dans les villages de certaines régions (Berthouly *et al.*, 2009).

Le poulet domestique s'est répandu en Europe par deux voies, la voie maritime par la Méditerranée d'une part, la voie terrestre par l'Asie centrale et la Russie (Coquerelle, 2000). Ainsi, les races traditionnelles comme la Gauloise dorée représentent le type méditerranéen des races européennes les plus anciennes. Au XIX<sup>ème</sup> siècle, d'autres races européennes ont été élaborées par des croisements entre les poules locales et des races asiatiques importées, par exemple de type combattant malais dans le cas des races de type chair. L'apparition de races à œufs bruns en Europe date également de cette période. Au XX<sup>ème</sup> siècle, la sélection est devenue le principal facteur d'évolution des races de poules en Europe. Ces différentes origines et pressions de sélection ont contribué à la biodiversité des races actuelles de poulet.. Mais les croisements ont également pu homogénéiser les différentes populations ou races et seules certaines de ces origines ont contribué au développement des lignées qui dominent actuellement l'industrie de la volaille dans le monde (Muir *et al.*, 2008). Alors que de nombreuses races polyvalentes sont menacées d'extinction, elles représentent les plus importantes ressources génétiques pour le futur. Malgré cela, la tendance actuelle favorise une sélection intensive des populations commerciales avec une poursuite d'objectifs de gain de production et de rentabilité. Certaines caractéristiques de l'espèce, notamment sa petite taille

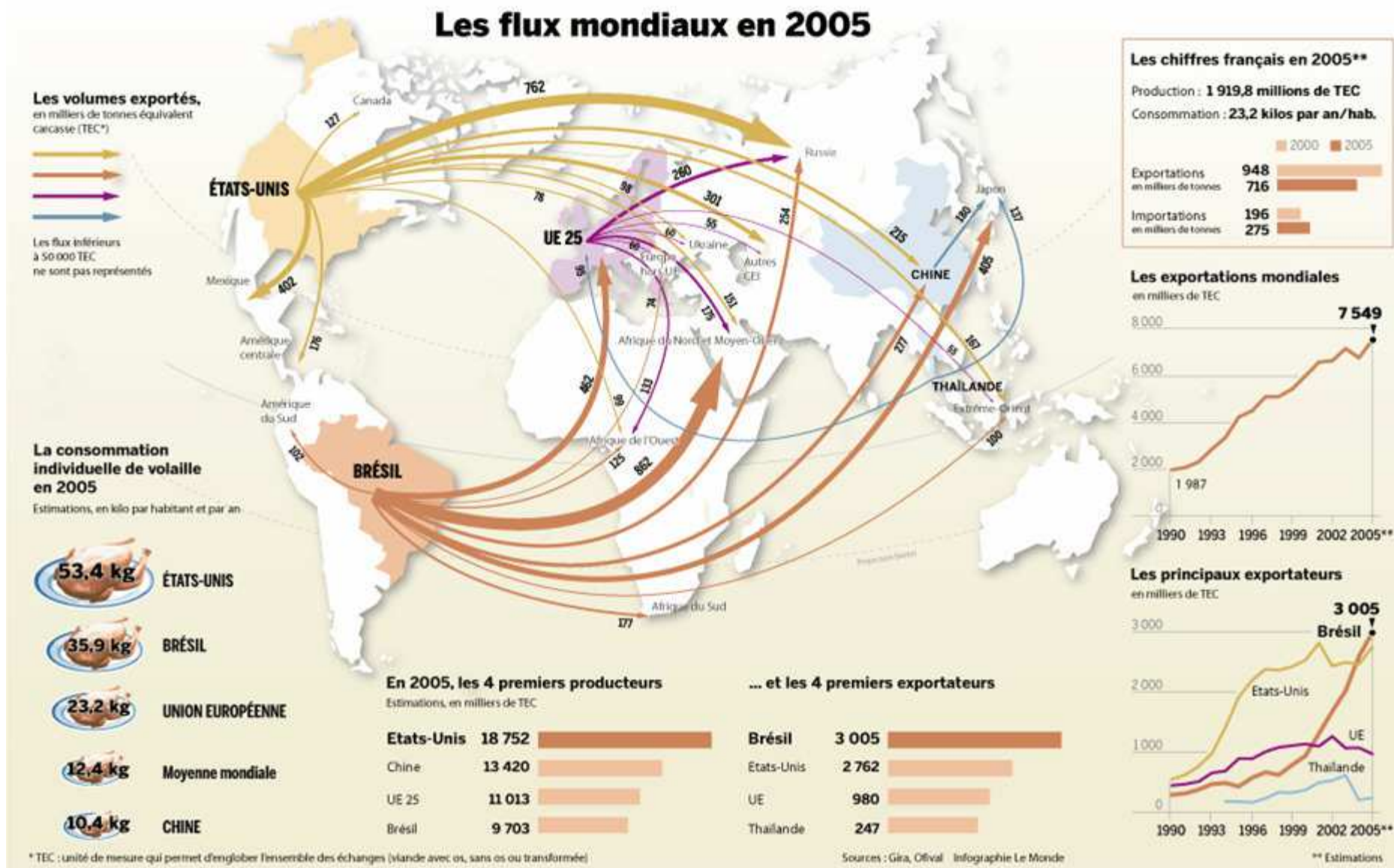


Figure 1. Les flux mondiaux de viande de volaille en 2005



et son intervalle de génération court se prêtent particulièrement bien à une sélection intensive, en particulier pour le niveau de production. Si la sélection avicole s'est montrée particulièrement efficace, elle s'est accompagnée d'une moindre adaptabilité des animaux à des systèmes plus variés. Depuis une dizaine d'année, la diminution du nombre d'entreprises de sélection conduit également à une diminution inquiétante du nombre de lignées commerciales, qui contribue également à la réduction de la variation génétique globale.

L'augmentation de la productivité des élevages avicoles s'est également traduite au cours du XXe siècle par l'exportation de la production excédante des pays du Nord vers les pays du Sud. Certains pays, comme le Brésil et la Chine, ont réagi en mettant en place leur propre système de sélection et de production (Figure 1). D'autres pays, notamment en Afrique, ont vu se développer des systèmes de production intensifs utilisant des animaux importés. Les modalités de ces introductions font craindre la réduction de la diversité des populations de poulets des pays importateurs, avec le remplacement des races locales ou le croisement entre races commerciales importées et locales, les résultats de ce type de croisement étant peu prévisibles.

Dans ce contexte, la valorisation des ressources génétiques de poulet, populations, races ou lignées est d'une grande importance mais nécessite la mise en évidence des particularités démontrant leur intérêt. Au-delà de leur valeur culturelle indéniable, ces ressources génétiques représentent un potentiel important sur le plan scientifique et peuvent apporter des alternatives intéressantes pour la production commerciale : leurs facultés d'adaptation à différentes conditions environnementales et leurs capacités de résistance aux maladies sont actuellement les motivations les plus pertinentes pour leur caractérisation et leur valorisation.

En effet, actuellement, les objectifs de sélection sont revisités pour répondre aux exigences du marché en termes de bien-être animal et de résistance aux maladies, en plus des objectifs habituels. En Europe, les nouveaux systèmes de production alternatifs pour les pondeuses, par exemple, au sol, nécessitent une meilleure compréhension de la variabilité génétique des caractères liés au comportement et à la résistance aux maladies, ceci s'accompagnant d'une législation réduisant l'utilisation d'antibiotiques et d'antiparasitaires pour toutes les filières.

Pour les poulets, la résistance aux maladies concerne d'une part les pathogènes affectant la santé humaine et la sécurité alimentaire, d'autre part les maladies non zoonotiques ayant des

conséquences économiques. En outre, la résistance aux maladies n'est pas un caractère simple à évaluer, elle peut signifier la résistance ciblée à certaines pathologies, mais également une meilleure réponse immunitaire générale. L'environnement pathogénique est fluctuant, même en milieux très contrôlés, et une approche durable que de plus en plus d'industriels prennent en considération serait de continuer la sélection tout en maintenant une certaine diversité génétique entre et au sein des lignées, afin de permettre de faire face à des pathogènes diversifiés. En effet, différentes théories s'accordent à rapprocher le maintien d'une forte diversité génétique, en particulier au niveau des gènes impliqués dans la présentation des antigènes du non-soi, à la capacité de développer une réponse immunitaire appropriée face au grand nombre de pathogènes qu'un organisme peut rencontrer au cours du temps, dans ses différents milieux de vie.

Actuellement, on considère que sélectionner pour la résistance aux maladies est possible (Stear *et al.*, 2001). Plusieurs programmes de recherche chez le poulet ont été couronnés de succès, autant pour l'identification ou la sélection d'animaux présentant une meilleure réponse immunitaire face à un pathogène donné (Calenge *et al.*, 2009 ; Kaufman, 2000), que pour la sélection d'animaux sur des caractères immunitaires généraux (Lamont *et al.*, 2003 ; Pinard-van der Laan, 2002). Cependant, le manque de connaissance sur les bases génétiques et les mécanismes sous-jacents aux caractères concernés est un facteur limitant.

## II. IMMUNOGENETIQUE AVIAIRE

Le système immunitaire du poulet présente de nombreuses similarités avec celui des Mammifères, mais également quelques particularités. Plusieurs types de tissus, organes, cellules et molécules interagissent dans un système qui peut être défini par deux aspects étroitement liés, la réponse innée et la réponse acquise. Quatre grandes catégories de gènes peuvent être considérées comme les principales composantes génétiques de la réponse immunitaire : ce sont les gènes codant pour les cytokines, ceux codant pour les anticorps ainsi que ceux codant pour les récepteur des lymphocytes T ou TCR (T Cell Receptor), et enfin les gènes du complexe majeur d'histocompatibilité (CMH).

### Réponse immune innée et cytokines

La réponse immunitaire innée regroupe les premières lignes de défense de l'organisme, des mécanismes de reconnaissance des pathogènes ainsi que des mécanismes de modulations du système immunitaire. La réponse immunitaire innée comprend les réactions inflammatoires, la production des protéines de phase aiguë, l'activation du système du complément, la libération de cytokines et la phagocytose.

Les cytokines sont des médiateurs cellulaires sécrétés pendant le développement et les activations immunitaires, pendant la réponse innée mais également acquise. Elles sont sécrétées par les cellules du système immunitaire ou ont une action sur ces cellules. En ce qui concerne les oiseaux d'élevage, des gènes codant pour des cytokines ont été identifiées chez le poulet mais également la caille japonaise (*Coturnix japonica*), la dinde (*Meleagris gallopavo*), les canards (de Pékin, *Anas platyrhynchos*, de Barbarie, *Cairina moschata*) et les oies cendrée et cygnoïde (*Anser anser*, *Anser cygnoides*) (Schultz et Magor, 2008). Le terme cytokine regroupe des interleukines, des interférons, des membres de la famille des TGF- $\beta$  (Transforming Growth Factor  $\beta$ ), des membres de la superfamille des TNF (Tumour Necrosis Factor), des facteurs CSF (Colony-Stimulating Factor) et des chimiokines. Brièvement, les interleukines ont des fonctions impliquant les lymphocytes alors que les interférons ont des effets anti-viraux. Les chimiokines ont pour première fonction de réguler les déplacements et recrutements des leucocytes.

La caractérisation des cytokines aviaires est relativement récente, étant donné le faible pourcentage d'identité (25-30%) des séquences protéiques aviaires avec celles des Mammifères. Pour le poulet, la disponibilité de la séquence du génome, depuis 2004, a permis un fort accroissement des connaissances et des travaux de recherche dans ce domaine. Globalement, la plupart des familles de cytokines connues chez les Mammifères existent chez le poulet et généralement, les membres des différentes familles de cytokines présentes chez le poulet sont en nombre plus réduit que chez les Mammifères (Kaiser et Staeheli, 2008).

### **Lymphocytes B et anticorps**

La bourse de Fabricius est un organe spécifique aux Oiseaux qui fait partie du tissu lymphoïde associé à l'intestin. C'est un organe actif chez les jeunes oiseaux qui régresse vers l'âge de 6 mois environ chez le poulet. Il est nécessaire à la maturation des cellules B. Ce sont des expériences de boursectomie chez le poulet dans les années 50 qui ont permis de démontrer que la bourse contrôlait l'immunité médiée par les anticorps, synthétisés et sécrétés par des lymphocytes alors appelés « B » pour Bourse (Davidson, 2003).

Les anticorps, qui sont donc sécrétés par les lymphocytes B, sont des glycoprotéines de la superfamille des immunoglobulines formées de 4 chaînes polypeptidiques: 2 chaînes lourdes et 2 chaînes légères. Les chaînes lourdes sont composées d'un fragment variable et de 3 ou 4 fragments constants, les chaînes légères, d'un domaine constant et d'un domaine variable. Les anticorps sont subdivisés en classes ou isotypes, selon l'organisation des domaines constants de leurs chaînes lourdes. Il a été observé chez les Oiseaux trois classes d'anticorps sécrétés : IgM, IgA et IgY. Les IgY sont les équivalents aviaires des IgG. La plupart des études portent sur le poulet et le canard, chez qui a été mis en évidence un IgY tronqué pour la région Fc (définie par clivage enzymatique et constituée de fragments constants), qui semble jouer un rôle important, notamment dans la neutralisation des antigènes (Schultz et Magor, 2008).

Chez les Oiseaux, le mode de génération du répertoire d'anticorps compétents est différent de celui connu chez les Mammifères. Chez ces derniers, le répertoire est généré par un processus de recombinaison somatique. En revanche, chez le poulet les gènes codant pour les segments variables et de jonction des chaînes légères et lourdes sont en simple copie. Mais des clusters de pseudogènes sont présents en amont des régions codant pour les chaînes lourdes et légères. Par un procédé appelé conversion génique, des fragments de gènes codant pour les régions variables sont remplacés par des fragments de pseudogènes (Ratcliffe, 2005, Figure 2). Une

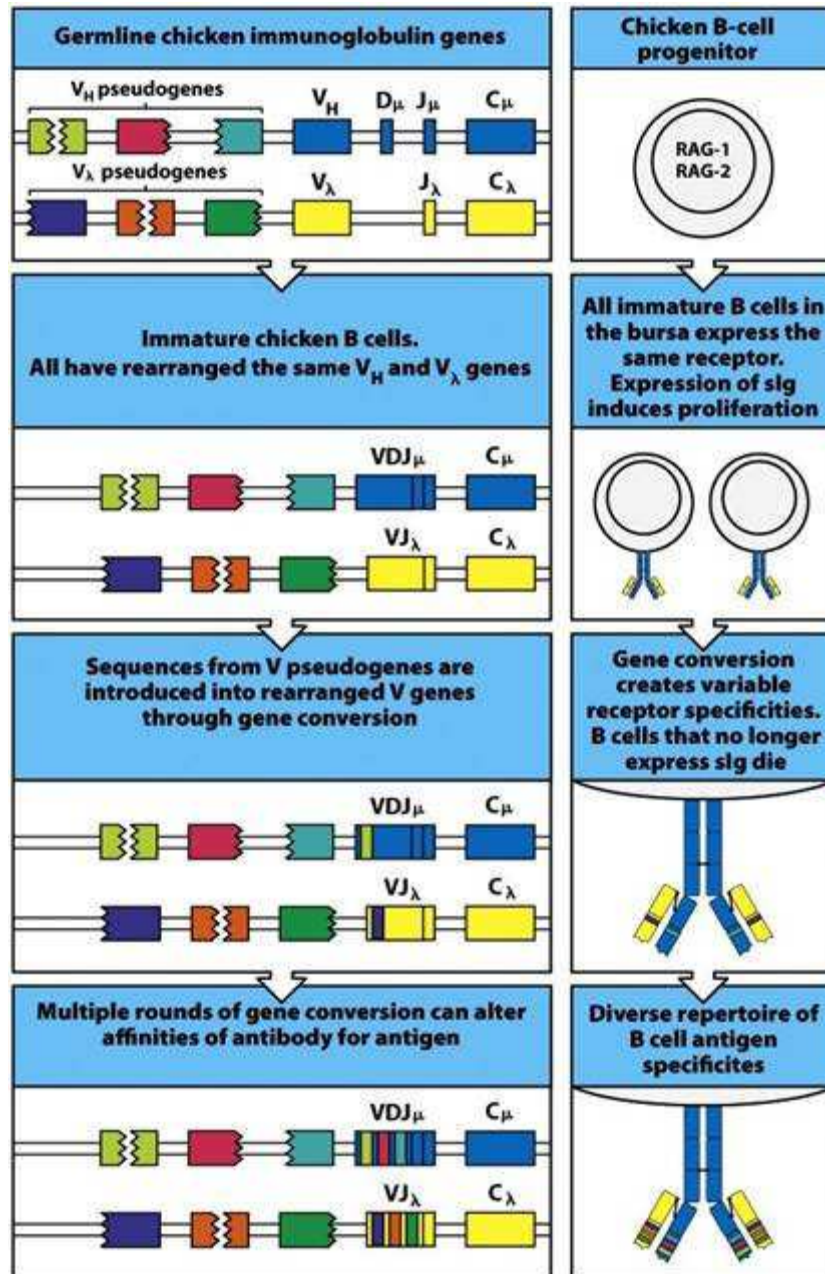


Figure 4-26 Immunobiology, 7ed. (© Garland Science 2008)

**Figure 2.** Mode de génération par conversion génique du répertoire d'anticorps chez le poulet (Janeway *et al.*, 2008)

enzyme, appelée AID (Activation-Induced Cytidine Deaminase) est nécessaire au processus (Arakawa *et al.*, 2002). Cette enzyme a été d'abord décrite pour le processus d'hypermutation somatique chez les Mammifères, comme déclencheur de la cascade d'événements aboutissant à l'accumulation de polymorphismes dans la séquence. Sa fonction première est de remplacer, par son activité désaminase, les cytidines par des uridines. Les mésappariements induits provoquent alors des cassures simples brins, qui s'ensuivent soit de réparations induisant des variations, soit de cassures double-brins activant alors une réparation par recombinaison qui peut s'avérer être non-homologues, notamment en présence de séquences similaires à proximité, et donc conduire à de la conversion génique. Chez le poulet, c'est cette conversion génique qui a lieu, mais uniquement pendant une partie de la maturation des lymphocytes B, dans la bourse de Fabricius. Si ce mode de génération du répertoire a été décrit chez d'autres animaux, notamment chez les bovins, les lapins, les ovins et les porcins, il semble qu'il constitue le seul mécanisme présent chez les Oiseaux.

### **Lymphocytes T et récepteurs TCR**

La plupart des propriétés des lymphocytes T des Mammifères se retrouve chez le poulet. Les lymphocytes T aviaires sont ainsi caractérisés par des fonctions différentes et sont classifiés en fonction des récepteurs (TCR) et des molécules CD4 et/ou CD8 présents à leur surface. Les études publiées chez d'autres Oiseaux démontrent également l'existence de lymphocytes T et sont actuellement limitées par le manque d'anticorps monoclonaux utilisables pour leur caractérisation précise. On trouve donc, chez le poulet, des lymphocytes TCR1 ( $\gamma\delta$ ), TCR2 et TCR3 ( $\alpha\beta$ ). L'analyse des gènes codant pour les TCR a révélé que malgré des divergences au niveau nucléotidique, les caractéristiques principales des polypeptides sont conservées entre le poulet et les Mammifères. De plus, contrairement aux anticorps, la série d'événements générant de la diversité chez les gènes codant pour les TCR chez le poulet est semblable à celle observée chez la souris et l'Homme. Ainsi, la majeure partie du répertoire des TCR de poulet est générée par réarrangement des membres des familles de gènes V, D et J. La seule différence notable est la relative simplicité des locus codant pour les TCR chez le poulet, comparés à ceux décrits chez les Mammifères. Fondièrément, seulement deux familles V $\beta$  ont été décrites chez le poulet, par rapport aux 20 à 30 familles décrites chez la souris et l'Homme (Lahti *et al.*, 1991).

La fonction et les processus de différenciation des lymphocytes T sont remarquablement similaires chez les Oiseaux et les Mammifères, avec notamment la distinction de lymphocytes  $T\alpha\beta$  et  $T\gamma\delta$ , indiquant ainsi une existence relativement ancienne de ces lignages cellulaires dans l'évolution des Vertébrés. Cependant, en comparaison avec la souris et l'Homme, chez qui ce pourcentage est d'environ 5%, le poulet fait partie des animaux qui présentent une proportion élevée de lymphocytes  $T\gamma\delta$ , jusqu'à 50 % des lymphocytes T périphériques, ce qui se retrouve également chez les bovins, lapins, ovins. Cette proportion élevée de lymphocytes  $T\gamma\delta$  a été associée au degré limité de diversité des gènes *TCR-V* et *Ig-V* du poulet (Plachy *et al.*, 2003). Chez le poulet, le rôle des lymphocytes  $T\gamma\delta$  n'a pas été clairement défini, leur capacité à se lier avec des antigènes indépendamment des molécules du CMH, comme chez les Mammifères, n'est pas connue.

### III LE COMPLEXE MAJEUR D'HISTOCOMPATIBILITE CHEZ LES OISEAUX

#### A. Généralités

L'étude de la région du complexe majeur d'histocompatibilité (CMH) chez les Vertébrés a démontré que cette zone était l'une des plus dynamiques du génome sur le plan évolutif (Kelley *et al.*, 2005). Il est défini historiquement comme le locus responsable du rejet de greffe : Gorer, en 1936, l'a découvert chez la souris et identifié comme l'ensemble de gènes qui contrôlent la compatibilité tissulaire. Le CMH de la souris, H2, et celui de l'Homme, HLA (Human Leucocyte Antigens, découvert par Dausset en 1958), ont été abondamment décrits tant du point de vue sérologique que sur le plan de leur organisation. Le CMH des Oiseaux, dont celui du poulet (*Gallus gallus*, complexe B), bien que découvert en 1948 par Briles, est moins bien connu. Le CMH, qui définit l'identité immunologique de l'individu, est un des composant génétique majeur du système immunitaire. Son organisation est très conservée entre les espèces, mais il comporte des gènes dont certains présentent un fort polymorphisme intraspécifique. En effet, le CMH comprend des gènes présentant le plus fort taux de polymorphismes des génomes, qui codent pour les glycoprotéines membranaires fixant les peptides antigéniques et les présentant en surface cellulaire aux lymphocytes T. L'origine des peptides, leur trajet intracellulaire et le type de lymphocytes T auxquels les peptides antigéniques seront présentés, permettent de classer ces molécules en deux classes : I et II. C'est ainsi que la région du CMH contenant les gènes codant pour ces molécules de classe I est appelée région de classe I et celle contenant les gènes codant pour ces molécules de classe II, région de classe II.

Les molécules de classe I, composées d'une chaîne lourde  $\alpha$  associée à la  $\beta$ -2 microglobuline, sont exprimées par l'ensemble des cellules de l'organisme. Ils présentent majoritairement aux lymphocytes T CD8<sup>+</sup> (cytotoxiques) des peptides antigéniques provenant de la dégradation par protéolyse de protéines endogènes synthétisées par la cellule et localisées dans le cytosol (protéines du soi, protéines virales ou tumorales). Le gène codant pour la  $\beta$ -2 microglobuline n'est pas situé dans le CMH. Les molécules de classe II, composées d'une chaîne lourde  $\alpha$  et d'une chaîne légère  $\beta$  sont exprimées principalement par les lymphocytes B et les cellules spécialisées dans la présentation d'antigènes (APC, Antigen-Presenting Cell). Ils présentent essentiellement aux lymphocytes CD4<sup>+</sup> (auxiliaires) des peptides antigéniques provenant de la dégradation endo-lysosomiale de protéines extracellulaires.



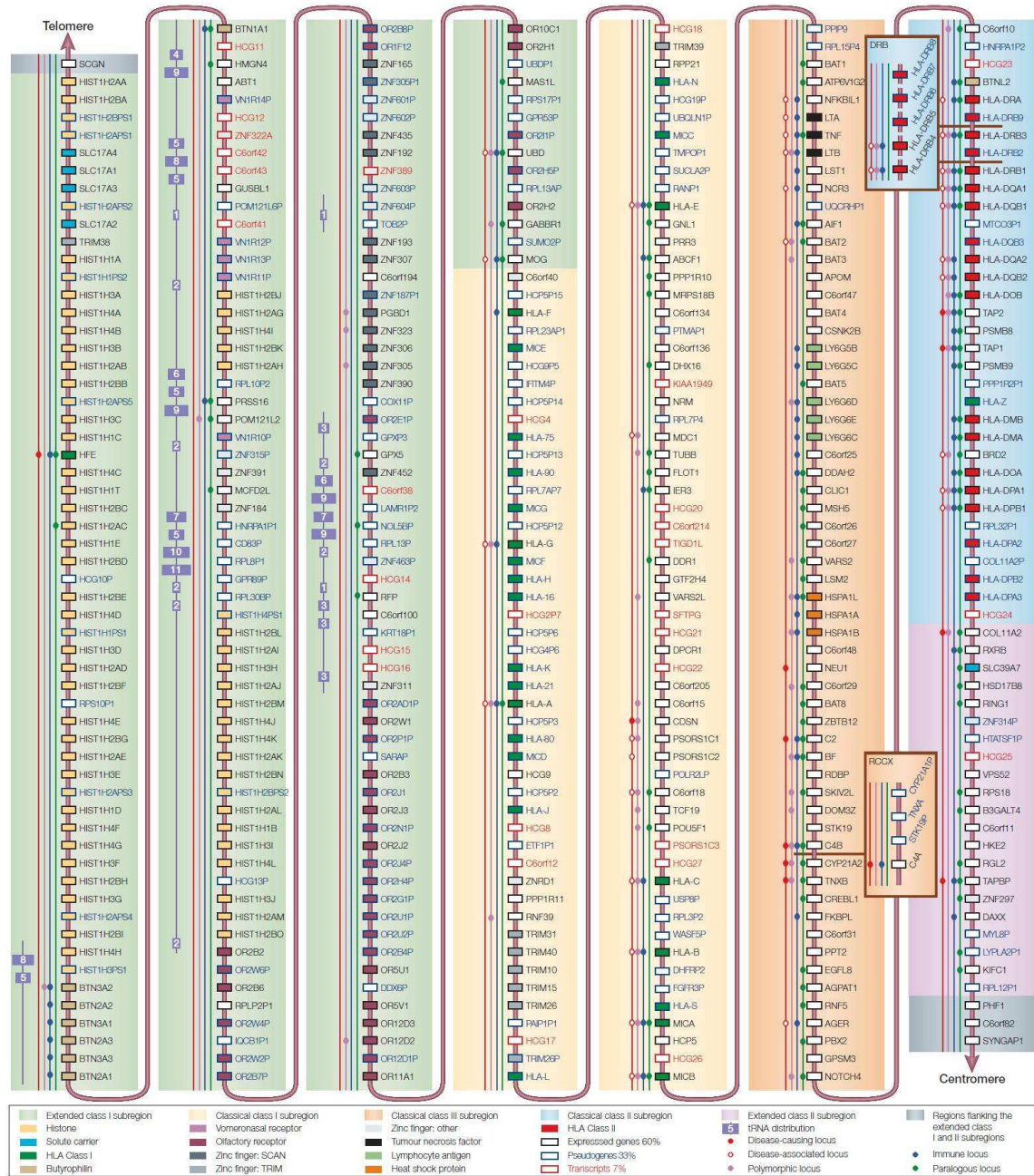


Figure 3. Carte du CMH étendu de l'Homme (Horton *et al.*, 2004)

La région du CMH chez les Mammifères est particulièrement étendue, en termes de distance physique et génétique, avec de nombreux gènes et pseudogènes mais également de nombreuses séquences répétées. On divise typiquement le CMH en trois sous-régions (Figure 3). La sous-région de classe I contenant de nombreux gènes de classe I classiques et non-classiques qui varient d'une espèce à l'autre, ainsi que d'autres gènes organisés selon un schéma commun à de nombreux Mammifères. La sous-région de classe II est beaucoup plus conservée, chez les Mammifères, et contient plusieurs couples de gènes de classes II classiques A et B, les gènes non-classiques *DM* et *DO*, impliqués dans le chargement des peptides par les molécules de classe II, quelques gènes codant pour des molécules impliquées dans le chargement des peptides par les molécules de classe I, notamment des gènes, inductibles par l'interféron, codant pour certaines sous-unités du protéasome (*LMP*, Low Molecular weight Protein), des gènes codant pour les transporteurs (*TAP*, Transporter of Antigenic Peptides ou Transporter Associated with antigen Processing) et pour la protéine associé au TAP, la tapasine, mais également pour une sérine-thréonine kinase (*RING3*, Really Interesting New Gene) ainsi que d'autres protéines diverses. L'appellation classique ou non-classique d'un gène de classe I ou II dépend essentiellement de sa fonction et son niveau de polymorphisme. Les gènes classiques sont ceux capables de présenter une grande diversité d'antigènes. Située entre les régions de classe I et de classe II chez l'Homme et la Souris, se trouve la sous-région particulièrement conservée de classe III, qui contient un grand nombre de gènes, certains codant pour des composants de la cascade du complément (*C2*, facteur B (*fB*) et *C4*), d'autres pour des gènes codant pour des cytokines de la superfamille des TNF (TNF et LT, LymphoToxines).

Par rapport à l'étendue des connaissances concernant le CMH des Mammifères, en particulier celui de l'Homme et de la Souris, il y a peu d'informations concernant le CMH des autres Vertébrés. Il semble que l'organisation du CMH soit apparue à la même période que le reste du système immunitaire acquis, chez les premiers Vertébrés (Kulski *et al.*, 2002). Les gènes codant pour les molécules de classe I et II, les TAP, la Tapasine, les TCR  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ , les anticorps et les molécules activant la recombinaison (RAG), ont été identifiés chez les Gnathostomes (Chondrichthyens et Actinoptérygiens) mais pas chez les Agnathes (*Hyperoartia* : lamproies et *Hyperotreti* : myxines) (Kaufman, 2008).

Le CMH du poulet est le mieux connu des CMH des Vertébrés non-mammaliens. Sa région B a été découverte en 1948 par Briles. Elle est considérée comme un CMH « minimal et

essentiel » (Kaufman, 2000). En effet, la région B du CMH du poulet ne contiendrait que deux gènes de classe I (*B-F*) et que deux gènes de classe II (*B-L*), avec dans chaque cas un seul gène majoritairement exprimé. Dans cette région, chez le poulet, ont également été décrits un autre type de gène codant pour des molécules présentes à la surface des cellules, les gènes *B-G*. Un autre cluster de gènes du CMH, génétiquement indépendant, a été décrit seulement dans les années 90, il se trouve sur le même microchromosome, GGA16, et constitue la région Y, elle est moins bien connue. On estime qu'il y a dans la région Y un grand nombre de gènes, certains de classe I dont l'expression a été démontrée, on peut donc exclure qu'il s'agisse seulement d'un groupe de pseudogènes (Afanassieff *et al.*, 2001). Si des gènes de classe I non classiques ont déjà été observés en dehors du CMH chez les Mammifères et les Amphibiens, la région Y est un nouveau type d'organisation génique avec une combinaison, dans un regroupement indépendant, de gènes codant pour des molécules impliquées dans la réponse immunitaire, dont des gènes de classe I.

## **B. Diversité interspécifique et dynamique évolutive**

Les oiseaux les plus étudiés, notamment en génétique, sont des Galliformes ou des Ansériformes, en effet ce sont les ordres correspondant à l'essentiel des oiseaux domestiques. Dans l'ordre phylogénétique des oiseaux actuels, ces ordres sont considérés comme relativement basaux, l'ordre des Passereaux étant considéré comme l'un des plus dérivés.

### **Faisan et Caille**

Ainsi, la plupart des études du CMH chez d'autres oiseaux concernent des espèces proches comme le faisan de Colchide (*Phasianus colchicus*) ou la caille japonaise (*Coturnix japonica*), appartenant à l'ordre des Galliformes. Pourtant le CMH de ces oiseaux ne semble pas présenter l'organisation minimale décrite chez le poulet.

Le faisan de Colchide est originaire d'Asie et a été introduit très largement en Europe et en Amérique du Nord. Il est proche phylogénétiquement du poulet, ces deux oiseaux appartiennent en effet à la même sous-famille (Ordre des Galliformes, famille des Phasianidae, sous-famille des Phasianinae). Le CMH du faisan, MHC-Phaco, a été notamment étudié chez des animaux sauvages (Wittzell *et al.*, 1999). Chez le faisan ont été identifiés des gènes de classe I, ainsi que des gènes de type B-G. Ont également été

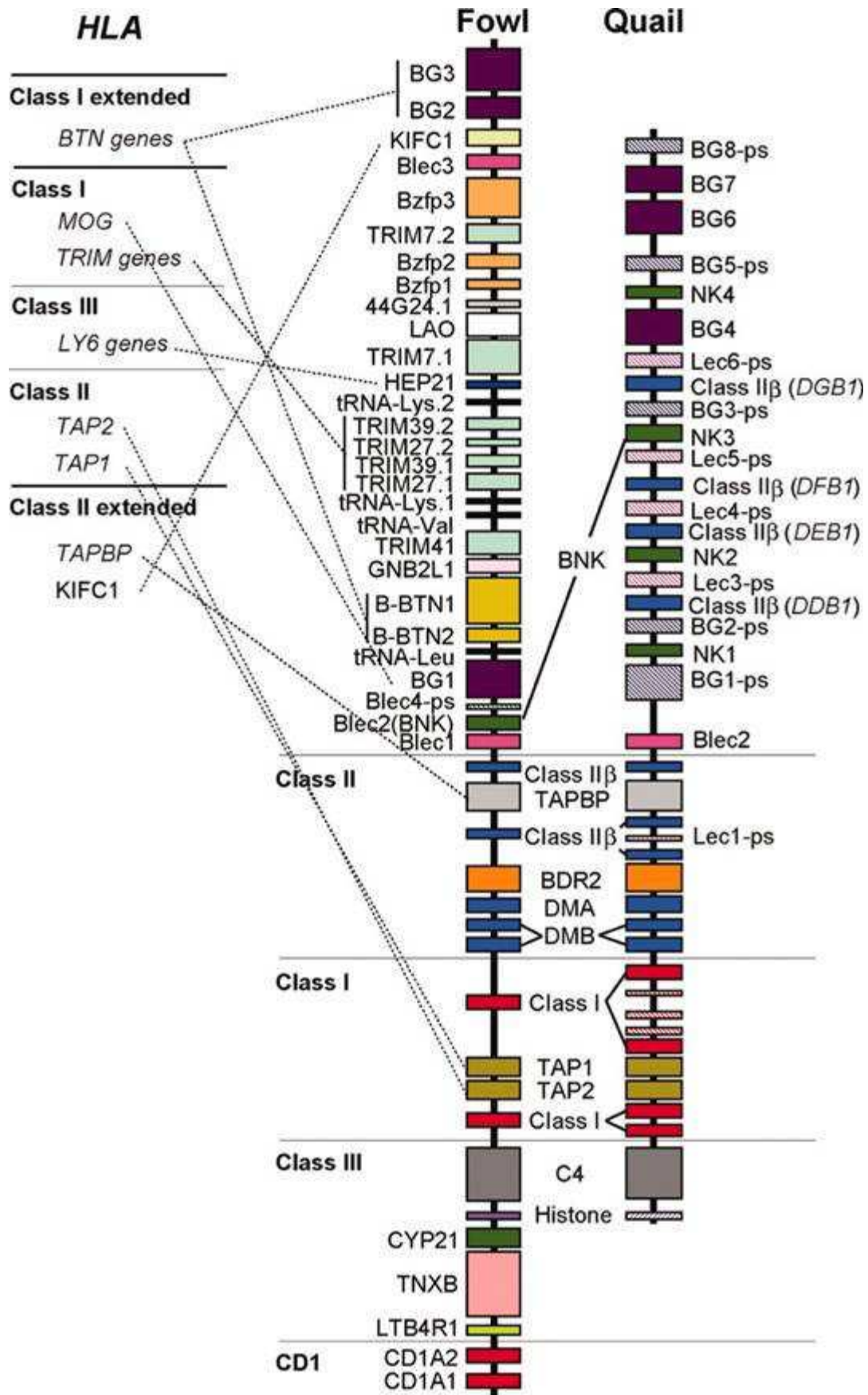
caractérisés, des gènes de classe IIB, certains analogues aux gènes de la région B du CMH du poulet, et d'autres apparentés aux gènes de la région Y, ce qui suggère que la région Y pourrait exister dans d'autres espèces d'oiseaux que le poulet (Jarvi *et al.*, 1996).

L'existence de deux gènes de classe IIB, *DAB1* et *DAB2*, homologues des gènes *B-LBI* et *B-LB2* du poulet, semble indiquer que deux gènes de classe IIB ancestraux existaient déjà avant la séparation des lignées évolutives qui ont donné naissance au poulet et au faisan, il y a 20 à 30 millions d'années, et pourraient provenir d'une duplication.

La caille japonaise (*Coturnix japonica*) appartient également à l'ordre des Galliformes, mais à une autre sous-famille (Famille des Phasianidae, sous-famille des Perdicinae). Elle a été domestiquée au Japon puis son élevage, pour les œufs ou la chair, s'est répandu dans le monde entier. C'est également un animal modèle pour lequel la région du CMH, MHC-Coja, a été partiellement séquencée et présente une organisation globalement similaire à celle de la région B du poulet (Numéro d'accèsion GenBank : AB078884, Figure 4). Cependant, chaque famille de gène semble y présenter un plus grand nombre de membres, avec au moins sept gènes de classe I, sept gènes de classe II, quatre gènes codant pour des récepteurs NK (Natural Killer), six gènes de type lectines et huit gène *B-G* (Shiina, 2004). Sur les sept gènes de classe I observés, il y aurait trois pseudogènes. Au final, la région de 180 kb séquencée contient 41 gènes décrits, là où la région correspondante du CMH du poulet en contient 19 pour 92 kb. Ce qui signifie que l'organisation reste globalement très similaire.

On remarque que chez la caille, comme chez le poulet, les gènes *TAP* sont dans la région de classe I. Ce type d'organisation est également retrouvé chez le poisson zèbre (*Danio rerio*, Takami *et al.*, 1997), laissant penser que l'organisation observée chez les Mammifères, avec les gènes *TAP* dans la région de classe II, est un état dérivé.

Un modèle de greffes embryonnaires a été mis au point entre le poulet et la caille, qui permet d'introduire des tissus étrangers avant la maturation du système immunitaire de l'hôte. Des animaux chimères naissent alors avec une aile appartenant au receveur et une aile appartenant au donneur. Quand le receveur est un poulet, la greffe est totalement rejetée, quelques jours après la naissance. Mais quand le receveur est une caille, l'aile provenant du poulet se développe jusqu'à dépasser de beaucoup la taille d'une aile normale de caille (Martin *et al.*, 1991). Les différences d'organisation et de polymorphismes des gènes de classe I et II, observées entre les CMH du poulet et de la caille donnent quelques pistes qui pourraient aboutir à la compréhension de ces phénomènes de rejet ou d'acceptation des xéno-greffes.



**Figure 4.** Carte comparée du CMH de la caille japonaise (quail) et du poulet (fowl) (Shiina *et al.*, 2006)

Ainsi, chez les Mammifères, les cellules T reconnaissant le soi sont éliminées dans le thymus ce qui permet de limiter les réactions auto-immunes. Plus les gènes du CMH sont nombreux, plus le répertoire de cellules T devrait être réduit. Si la caille obéit à ce schéma, vu son plus grand nombre de gènes de classe I et II, son répertoire de cellules T devrait être plus réduit que celui du poulet, ce qui pourrait expliquer l'immunotolérance observée lors des xéno-transplantations.

L'étude du CMH chez le faisan et la caille suggère que le petit nombre de gènes de classe I et II observé chez le poulet n'est pas représentatif de l'ensemble de l'ordre des Galliformes. Ces particularités pourraient être liées aux processus de domestication et de sélection. Dans cette optique, l'étude des espèces sauvages de poule de jungle, est particulièrement pertinente. Si des mécanismes dus à la domestication ont ainsi été mis en cause (Lambourne *et al.*, 2005), la séquence du génome provenant d'un individu Coq de Bankiva, semble similaire au niveau du CMH, à celle des poulets domestiques.

La plupart des autres données disponibles sur les CMH des Oiseaux concernent des oiseaux domestiques, notamment la dinde (*Meleagris gallopavo*) et le canard.

### **Dinde**

En ce qui concerne la dinde, les connaissances étaient assez limitées, par exemple, au niveau de polymorphisme d'un gène de classe II (Jones Dukes *et al.*, 2003), jusqu'à l'établissement des cartes génétique et physique du chromosome (Chaves *et al.*, 2007) et très récemment, la publication de la séquence génomique (Chaves *et al.*, 2009). Il en ressort que chez la dinde on retrouve deux locus correspondant au CMH, génétiquement indépendants situés sur le même microchromosome que la région NOR (Nucleolar Organizer Region). L'ensemble de la séquence disponible actuellement (Numéro d'accèsion GenBank : DQ993255) fait 197 kb et présente 34 gènes en grande synténie avec l'organisation du CMH du poulet avec cependant trois gènes *B-G* et trois gènes de classe IIB, là où le poulet n'en présente respectivement qu'un et deux.

## **Canard de Pékin ou Colvert**

Chez le canard de Pékin ou Colvert, un Anseriforme, la région de classe I contient cinq gènes de classe I et deux gènes *TAP*, avec l'expression majoritaire d'un seul gène de classe I, celui situé à proximité du gène *TAP2*. Ces données correspondent aux deux régions génomiques étendues décrites, l'une de 7 kb et l'autre de 37 kb, qui ne couvre que la région de classe I (Mesa *et al.*, 2004, Moon *et al.*, 2005). Aucune autre région du génome du canard ne semble contenir de gène de classe I.

Plusieurs phénomènes peuvent être liés à un nombre réduit de gènes du CMH exprimés, ce qui est donc observé chez le canard mais également chez plusieurs autres espèces d'oiseaux, et ce qui peut avoir des conséquences sur le répertoire des récepteurs des cellules T et des cellules NK, mais également, pour les gènes de classe I, à savoir la réduction du nombre de peptides antigéniques pouvant être présentés à la surface des cellules. Les hypothèses de coévolution des gènes de classe I et des gènes *TAP*, qui codent pour des molécules essentielles au processus de chargement des peptides par les molécules de classe I, et qui sont peu polymorphes, prédisent un intérêt évolutif de combinaisons privilégiées classe I / *TAP*.

## **Oiseaux sauvages**

Les autres travaux concernant les Oiseaux, ont été réalisés avec des optiques et dans des domaines différents. Depuis plusieurs années, des équipes d'écologistes et de généticiens ont entrepris l'étude de la coévolution hôte-pathogène chez des oiseaux. L'étude du CMH pourrait permettre de valider certaines hypothèses, en intégrant la génétique moléculaire et l'expression phénotypique de la résistance, et en prenant explicitement en compte les contraintes écologiques susceptibles d'affecter simultanément la virulence des pathogènes et la résistance des hôtes. Ces contraintes sont : la biologie de l'espèce hôte (aire de répartition, animal migrateur), le degré de socialité des hôtes (la socialité augmente les rencontres entre individus et, donc, la probabilité de transmission des pathogènes et les risques d'infestations multiples) et l'intensité de la sélection sexuelle (la plupart des caractères sexuels secondaires se développent sous l'action d'hormones stéroïdes qui interagissent au niveau moléculaire avec les produits (*sensu lato*) du système immunitaire). Ces phénomènes doivent être étudiés à des niveaux d'intégration différents allant des gènes aux individus d'une même population, aux individus d'une même espèce appartenant à des populations différentes dans des conditions écologiques différentes (latitudes différentes etc.), et enfin aux espèces

phylogénétiquement proches mais qui présentent des degrés de socialité et de sélection sexuelle différents. Une grande partie de ces études concernent des parasites, souvent considérés comme plus pertinents pour ces études chez les animaux sauvages.

Les forces évolutives dues à du parasitisme dans le milieu naturel conduisent en théorie, au niveau du CMH, à de la superdominance (hétérozygotie favorisée) et/ou à la présence d'allèles rares (sélection négativement corrélée à la fréquence : avantage du rare). L'étude d'oiseaux, chez qui on connaît de nombreux cas d'association entre haplotype au CMH et résistance aux maladies, peut permettre de rechercher les preuves de l'existence de ces forces au niveau du CMH. Certains modèles d'évolution trouveront peut-être alors des validations dans des populations naturelles.

En parallèle, ces études permettent d'évaluer la variabilité génétique au niveau du CMH dans des populations présentant des histoires démographiques très différentes, par exemple des goulots d'étranglement dans le passé, ayant conduit, à une époque, à des réductions drastiques de leur effectif efficace.

Les hypothèses de base de la génétique des populations, qui voudraient que ce type de phénomène affaiblisse les populations en réduisant la variabilité génétique et en conséquence réduise les réponses immunitaires possibles aux pathogènes, pourraient enfin être étudiées dans des populations naturelles.

Dans cette optique, la présence d'un seul gène de classe I exprimé chez le canard, qui est à l'état sauvage un oiseau migrateur, semble aller à l'encontre de certaines théories proposées chez d'autres espèces migratrices. En effet, du fait que les oiseaux migrateurs sont soumis à des pressions de sélection dues aux pathogènes plus diversifiées que les oiseaux sédentaires, certaines hypothèses prédisent que ces espèces migratrices doivent présenter un plus grand nombre de gènes ou un plus grand polymorphisme au niveau du CMH, que les oiseaux sédentaires. Mais il faut souligner que la plupart des cailles ou des canards étudiés proviennent de populations domestiques, voire expérimentales. Cette contradiction pourrait ainsi être expliquée. L'étude d'animaux provenant de populations sauvages migratrices serait alors déterminante.

Ainsi, les Oiseaux sauvages apparaissent comme des modèles très intéressants pour ces types d'études qui cherchent à élucider les liens entre la variabilité génétique au niveau du CMH,



l'expression phénotypique de la réponse immunitaire et la résistance aux parasites dans des populations naturelles.

Les Passereaux représentant plus de la moitié des espèces d'oiseaux, étant présents dans la plupart des écosystèmes et donc étant soumis à différents types de pressions parasitaires, sont un modèle privilégié de ce type d'études.

## **Passereaux**

L'ancêtre commun le plus récent entre les Passereaux et le poulet est approximativement aussi ancien que celui de la Souris et de l'Homme. L'étude du CMH chez les Passereaux et les comparaisons avec le poulet peuvent donc également servir à mieux comprendre si les forces évolutives suspectées chez les Mammifères sont généralisables à d'autres Vertébrés comme les Oiseaux.

Les particularités connues du CMH sont un polymorphisme très important au sein des populations au niveau des gènes de classe I et II, un excès de substitutions non synonymes par rapport aux substitutions synonymes au niveau des séquences codant le site de fixation de l'antigène (Antigen-Binding Site ou Peptide Binding Region : PBR), et la persistance à temps long de certains allèles. Chez les Mammifères, des modèles expliquant ces caractéristiques ont été proposés : l'existence d'une sélection équilibrante (Balancing Selection) qui expliquerait les ratios entre différents types de mutations au niveau de la PBR, et d'une sélection divergente entre locus qui se traduirait par le maintien des gènes dupliqués au cours de l'évolution (Hughes & Nei, 1990 ; Takahata *et al.*, 1992). La sélection équilibrante fait partie du modèle dans lequel les gènes possèdent de nombreux allèles présentant des fréquences intermédiaires, la plupart des individus sont hétérozygotes à plusieurs locus et tous les individus diffèrent. Le polymorphisme se maintient par sélection équilibrante : la valeur sélective du porteur d'un génotype dépend des autres allèles présents aux autres locus et du milieu.

Lorsque l'on considère des gènes de classe II, les polymorphismes sont en général étudiés au niveau second exon car il code pour la PBR. Le troisième exon est également étudié de façon à pouvoir comparer les séquences entre oiseaux et faire des hypothèses sur l'évolution à temps long au sein des Oiseaux des gènes de classe II.

Par exemple, chez le Geai à gorge blanche (*Aphelocoma coerulescens*), le Carouge à épaulettes (*Agelaius phoeniceus*) et le Roselin familier (*Carpodacus mexicanus*), Edwards *et*

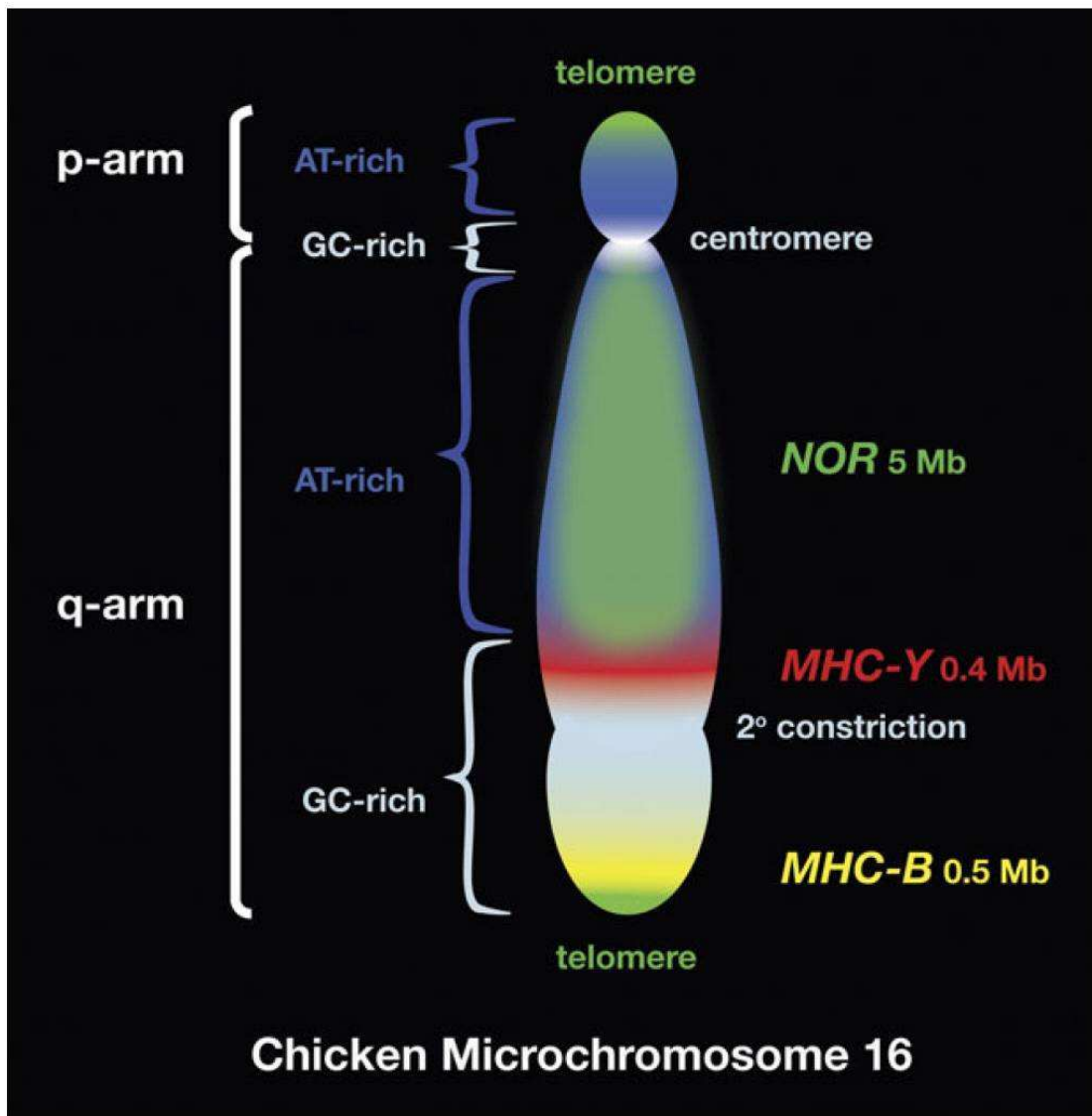
*al.*, en 1995, ont observé le même type de substitutions, au niveau de l'exon 2, que chez les Mammifères mais beaucoup moins de similitudes de séquences entre espèces. Le modèle d'évolution à temps court de sélection équilibrante serait alors valable chez ces oiseaux mais pas la dynamique à temps long proposée chez les Mammifères.

Chez les Oiseaux, le modèle serait plutôt une sélection convergente, les gènes paralogues du génome se ressemblant fortement au niveau de leurs régions codantes et non codantes. L'évolution se ferait d'une manière concertée, ou les duplications observées seraient assez récentes.

Il faut cependant noter que le nombre exact de gènes de classe II chez les espèces de Passereaux étudiés n'est pas connu, ni même l'organisation de leur CMH, en un ou deux systèmes. Cependant, les Passereaux ne semblent pas présenter de région CMH aussi compacte que celle du poulet (Edwards *et al.*, 2000, Westerdahl *et al.*, 2000, Bonneaud *et al.*, 2004)

Sans plus d'études sur le CMH dans d'autres clades chez les Oiseaux, il est assez difficile de conclure sur l'état ancestral ou dérivé du CMH compact observé chez le poulet. Actuellement, on n'explique pas l'origine de ce petit nombre de gènes, on observe simplement moins de classe I chez les Galliformes que chez les Passériformes, mais autant de classe II, à l'exclusion du poulet.

Il paraît alors évident que l'étude du CMH chez les Oiseaux ne doit pas se limiter à l'étude du CMH du poulet et le premier assemblage du génome du Diamant mandarin (*Taeniopygia guttata*) aurait pu permettre d'améliorer cette situation, cependant il ne couvre quasiment pas le microchromosome concerné. Actuellement, le poulet reste le modèle principal pour les Oiseaux, pour des raisons historiques et économiques. L'étendue des connaissances en général rend cette espèce incontournable et la particularité présentée par la compacité du CMH chez le poulet en font un modèle de choix pour l'immunogénétique.



**Figure 5.** Modèle d'organisation du microchromosome GGA16 (Delany *et al.*, 2009)

## C. Le CMH du poulet

Alors que la séquence du génome du poulet est disponible depuis 2004, et a été réassemblée en 2006, on ne connaît qu'environ 5% du microchromosome GGA16, sur lequel se situe le CMH. Cependant, deux autres séquences du CMH ont été publiées, la séquence référencée sous GenBank par le numéro d'accèsion AL023516, appelée GGB locus (Kaufman *et al.*, 1999) et celle ayant pour numéro AB268588 (Shiina *et al.*, 2007). Ces deux séquences ne concernent que la région B du CMH du poulet. Les difficultés rencontrées pour le séquençage de ce chromosome peuvent être mises en relation avec sa taille (c'est un microchromosome) mais également avec la présence de l'unique organisateur nucléolaire du génome du poulet (NOR), qui contient un très grand nombre de gènes ribosomiaux. L'ordre des régions B, Y et NOR a été récemment définitivement résolu, par hybridation fluorescente *in situ* (FISH) multi-couleur: sur le bras q du chromosome, c'est la région Y qui se trouve entre les deux autres, les régions Y et NOR étant proches, la région B plus distante, dans la partie terminale du bras q (Delany *et al.*, 2009). C'est une région riche en GC relativement grande physiquement qui sépare les régions B et Y, elle serait à la base de l'absence de liaison génétique entre les deux régions (Figure 5).

### 1. Historique du typage du CMH chez le poulet

Les premières études concernant les réactions immunes chez le poulet datent des années 30-40, notamment des études de variabilité individuelle concernant les alloantigènes érythrocytaires. Les premières études détaillées sur les groupes sanguins chez le poulet sont celles de Briles dans les années 50.

Dans un premier temps, ces études étaient donc basées sur l'injection d'érythrocytes, ensuite sur des greffes réciproques de peau: les greffes de peau échangées entre les poussins incompatibles étaient rejetées dans les 12 jours alors que les greffes entre les poussins compatibles se maintenaient à 40 jours post-opération. Si les greffes de peau se sont montrées utiles pour identifier les antigènes d'histocompatibilité, elles nécessitaient de longues périodes d'observation. Un autre indicateur d'histocompatibilité, découvert dans les années 50, concerne les réactions du greffon contre l'hôte, ou GVHR (Graft Versus Host Reaction). Payne et Jaffe ont ainsi démontré en 1962, en utilisant des lignées pures et leurs croisements, qu'un système majeur causait une splénomégalie (augmentation de volume de la rate) chez le

poulet. Ce système fut identifié comme le système B, par association avec les groupes sanguins B connus des animaux donneurs et receveurs. On remarquera que l'origine des animaux utilisés, par exemple la race, avait également été notée comme un facteur modulant l'histocompatibilité.

A la fin des années 70, grâce à la découverte d'un animal présentant un système B recombinant, issu de deux lignées consanguines, un modèle à trois locus fut proposé pour la région B (Longenecker et Mosmann, 1981). Ce modèle décrit une région génomique avec d'un côté les locus *B-F* et *B-L*, de l'autre le locus *B-G*, un crossing-over ayant permis une recombinaison entre le locus *B-G*, lié aux molécules présentes sur les érythrocytes et leurs précurseurs, et les locus *B-F* et *B-L*, liés aux molécules présentes sur les leucocytes.

Les allèles de la région B ont été rapidement numérotés chronologiquement B1 à Bn. Dans les années 70, une certaine confusion régnait, car des laboratoires utilisaient les mêmes numéros pour des allèles différents. De nombreux efforts furent déployés dans ce sens, dans les années 80, et la question fut partiellement résolue en ce qui concerne les lignées expérimentales, avec la définition en 1982 de 27 références (Briles *et al.*, 1982): B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, B17, B18, B19, B21, B22, B23, B24, B25, B26, B27, B28, B29. Une des raisons expliquant cette confusion est que contrairement à la souris, les études concernant le poulet ont concerné dès le départ des animaux divers et non pas des lignées consanguines. Une nomenclature désignant les recombinants fut également développée, utilisant la lettre « r » suivie d'un numéro, censé correspondre à un allèle particulier *B-F*.

Malgré l'identification de recombinants, les locus *B-F*, *B-L* et *B-G* sont considérés comme étant en liaison étroite au sein de la région B, ils sont ainsi transmis en bloc et l'association des allèles aux différents locus peut définir un haplotype. Ces haplotypes sont relativement conservés, les événements de recombinaison (crossing-over) entre deux haplotypes s'étant révélés particulièrement rares.

Un article publié en 1982 par Simonsen *et al.*, récapitule les résultats obtenus en hémagglutination, ce test consistant à additionner un sérum de référence au sang à typer. Les alloanticorps issus de l'anti-sérum allospécifique réagissent et provoquant l'agglutination avec les molécules présentes à la surface des érythrocytes. Ce typage est donc révélateur des allèles *B-G* qui sont exprimés sur les érythrocytes et leurs précurseurs, sauf lorsque le sérum a

été obtenu par immunisation avec des leucocytes purifiés, afin de typer les allèles *B-F*. Cet article présente également les résultats obtenus par des tests de compatibilité direct GVHR (mesure de splénomégalie), d'inhibition de la GVHR (GVH-I-R) qui est un typage spécifique des allèles *B-F*, et d'immunofluorescence indirecte (IIF) pour le typage spécifique des allèles *B-L*. Les principaux sérotypes B décrits sont les sérotypes B1, B2, B4, B5, B6, B7, B9, B12, B13, B14, B15, B19, B21 (Tableau 1).

Actuellement les sérotypes B du CMH sont toujours classés en fonction des tests de sérologie des groupes sanguins, des tests d'hémagglutination (voir Juul-Madsen *et al.*, annexe 1). Effectués *de novo*, ces typages concernent donc essentiellement les allèles *B-G*. La classification actuelle des sérotypes B du CMH est donc en grande partie basée sur les connaissances historiquement acquises, avec un premier récapitulatif publié en 1982 par Briles, un deuxième publié en 2004 par Miller *et al.*, et un troisième et dernier, publié en 2008 à la fin du livre Avian Immunology (Davidson *et al.*, 2008). La définition de nouveaux sérotypes est donc difficile, la vingtaine de sérotypes de référence étant les seuls bien définis, d'autres sérotypes restant confidentiels et les techniques de comparaison étant de moins en moins usitées. La plupart des typages réalisés aujourd'hui sont des tests de contrôle ou de comparaison simple, qui prennent en considération le très fort déséquilibre de liaison entre les locus *B-G*, *B-F* et *B-L*. On note également que ces techniques de typage ne tiennent compte que des molécules d'histocompatibilité portées par la région B alors que l'on connaît maintenant l'existence de la région Y.

Ces problèmes se posent plus particulièrement pour le typage de poulets provenant de races différant de la race expérimentale par excellence, la White Leghorn, une race pondeuse. Les réactifs de typage, préparés au sein de certaines populations, ne peuvent pas être facilement utilisés en raison de leur réactivité croisée lors du typage de poulets provenant de populations génétiquement différentes. De plus, lorsque le typage par sérologie reste confiné à la population, comme c'est le cas pour les poulets de chair ou pour les pondeuses d'œufs bruns, le système de nomenclature atteint ses limites, ne permettant pas de comparer les différents sérotypes B identifiés.

C'est pour faire face à ces difficultés que l'effort a été mis, dès la fin des années 80, sur un typage moléculaire utilisant les variations de l'ADN. Les premières techniques développées ont été des techniques de RFLP, basées sur les polymorphismes des gènes *B-G* (avec un brevet déposé par Miller en 1995) ou des gènes *B-F* et *B-L* (Lamont *et al.*, 1990) utilisant donc des sondes spécifiques à ces gènes. Un avantage de cette approche est que la

**Tableau 1.** Comparaison de 13 sérotypes B (d'après Simonsen *et al.*, 1982)

| Sérotipe B | Allèle L | Allèle F | Allèle G      |
|------------|----------|----------|---------------|
| 1          | 1        | 1 ≈ 7, 5 | 1 ≈ 4         |
| 2          | 2        | 2 ≈ 12   | 2 = 12        |
| 4          | 4 = 13   | 4 = 13   | 4 ≈ 1         |
| 5          | 5        | 5 ≈ 1, 7 | 5 ≈ 7, 19     |
| 6          | 6 ≈ 15   | 6        | 6             |
| 7          | 7        | 7 ≈ 1,5  | 7 ≈ 5, 19     |
| 9          | 9        | 9        | 9             |
| 12         | 12 = 19  | 12 ≈ 2   | 12 = 2        |
| 13         | 13 = 4   | 13 = 4   | 13            |
| 14         | 14       | 14       | 14 ≈ 19       |
| 15         | 15 ≈ 6*  | 15 ≈ 19  | 15            |
| 19         | 19 = 12  | 19 ≈ 15  | 19 ≈ 5, 7, 14 |
| 21         | 21       | 21       | 21            |

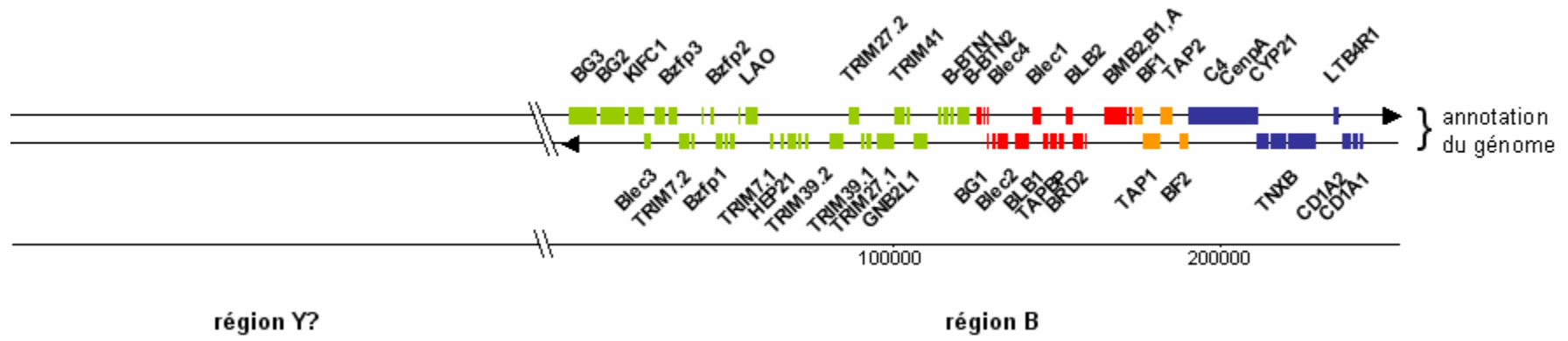
=, identité probable ; ≈, grande similarité mais des différences prouvées

connaissance préalable des séquences des gènes n'est pas nécessaire, car une autre méthode développée à la même époque utilisait comme sondes des oligonucléotides spécifiques de séquences correspondant à différents allèles des gènes *B-F* (Guillemot *et al.*, 1988).

Dans les années 90, des méthodes de PCR-SSCP (single strand conformation polymorphism), ciblant les gènes *B-G* puis d'autres gènes furent mises au point. C'est également à cette époque que l'existence de la région Y fut découverte. En conséquences, des techniques de typages furent mises au point, en RFLP ou en SSCP, permettant de typer à la fois la région B et la région Y (Goto *et al.*, 2002, Miller US patent 2004, annexe 1). En pratique, ces techniques sont assez lourdes à mettre en place et peu ou pas d'études ont été publiées permettant de comparer des sérotypes ou définir ainsi de nouveaux haplotypes. Leur utilisation chez les sélectionneurs privés est difficile à évaluer en raison des politiques de confidentialité mises en place. Cependant, ces techniques sont utilisées, en première approche, pour l'étude d'autres oiseaux, notamment sauvages, en RFLP, SSCP ou DGGE (denaturing gradient gel electrophoresis, Bonneaud *et al.*, 2004).

Le développement des marqueurs moléculaires, de type microsatellite, aurait du permettre de résoudre ces difficultés, cependant, aucun marqueur de type microsatellite permettant de typer le CMH chez le poulet n'a pu être identifié. Ceci est à mettre en relation avec la compacité de la région B, sa forte densité en gènes et donc le peu de séquences intergéniques. Les marqueurs mis au point furent donc des marqueurs atypiques et notamment les marqueurs COM dont COM154 (Wain *et al.*, 1998, numéro d'accèsion GenBank AF063648), marqueurs caractérisé par RFLP génomique et peu utilisés, le marqueur LEI0258 (Zoorob *et al.*, sous la dénomination r12.1, patent 1998, McConnell *et al.*, 1999, numéro d'accèsion GenBank Z83781), MCW0312 et MCW0371 (Crooijmans *et al.*, 1997, Groenen *et al.*, 2000). Ces marqueurs sont tous situés dans ou à proximité de la région B. En 2006, Fulton *et al.*, ont utilisé le marqueur LEI0258 et le marqueur MCW0371 pour typer plus de cinquante sérotypes dont la vingtaine de référence. Le marqueur MCW0371 est une répétition d'un mononucléotide, et dans cette étude 8 allèles ont été identifiés, d'une taille allant de 200 à 209 pb. Mais ce type de marqueur est difficile à typer. Le marqueur LEI0258 est un marqueur de type VNTR (variable number of tandem repeat) complexe composé principalement de deux motifs répétés consécutifs de 12 et 13 pb. Dans cette même étude, il s'est révélé particulièrement polymorphe, avec 26 allèles, allant de 182 à 552 bp. Mais la complexité de ce marqueur a été révélée lors du séquençage de ses allèles, avec de nombreuses





**Figure 6.** Carte du CMH du poulet présentant les différents gènes annotés (d'après Shiina *et al.*, 2007)

combinaisons possibles des nombres de répétition de chaque motif, ainsi que des insertions/délétions, allant jusqu'à 7 pb et de nombreux SNPs dans les régions flanquantes. Actuellement, c'est le marqueur LEI0258 qui sert de marqueur de génotypage moléculaire de référence du CMH chez le poulet (Annexe 1).

## **2. Gènes de la région B du CMH du poulet**

La notion d'haplotype du CMH chez le poulet utilisée actuellement correspond donc à l'association des allèles des gènes présents dans la région B. Ces haplotypes sont relativement conservés car les événements de recombinaison (*crossing-over*) entre deux haplotypes sont rares. Ce qui signifie qu'il existe un déséquilibre de liaison très fort entre les allèles de certains gènes. Toutes les combinaisons possibles ne sont pas représentées chez le poulet comme chez les Mammifères, certaines ont probablement disparu au cours de l'évolution et / ou de la domestication. Une des hypothèses avancée serait une pression de sélection favorisant le maintien de combinaisons procurant un avantage sélectif par la co-adaptation de gènes liés (Kaufman *et al.*, 1997, 2000).

En plus de la séquence du génome, deux autres séquences de la région B ont été publiées (Kaufman 1999, Shiina, 2007). Si ces trois séquences ne concernent que deux sérotypes, B12 et B14, des séquences partielles de la zone centrale des régions de classe I et II ont été publiées pour un beaucoup plus grande nombre d'animaux et sont entièrement disponibles pour 14 sérotypes (Hosomichi *et al.*, 2008, Figure 6). Toutes ces études indiquent une organisation similaire de la zone centrale de la région B, simple et compacte, avec une vingtaine de gènes pour 90 kb. On y retrouve de nombreux gènes caractéristiques du CMH chez les Mammifères, comme les gènes de classe I et IIB. L'organisation des différentes régions et gènes est cependant différente sur certains points, avec la région de classe III à l'extérieur des régions de classe I et II, les gènes *TAP* dans la région de classe I et le gène tapasine proche des gènes de classe II. On retrouve également dans cette région, des gènes *B-G*, pour lesquels il n'y a pas d'équivalent connu chez les Mammifères. Historiquement mais probablement à raison, la zone plus distante contenant les gènes *B-G* et la zone centrale de la région B, sont considérées comme soumises à la recombinaison.

Plusieurs articles décrivent les gènes de la région B du CMH du poulet, notamment Kaufman, 1999, Miller 2004, Shiina, 2006, Shiina 2007. Une nomenclature de référence a été publiée par Miller *et al.*, en 2004, mais elle n'est pas forcément encore bien respectée.

### **Gènes de classe I et voie de présentation des antigènes endogènes**

La voie de présentation des antigènes endogènes ou voie cytosolique est décrite chez les Mammifères comme la voie aboutissant à la présentation de peptides provenant du cytosol, par les molécules de classe I, aux lymphocytes T cytotoxiques CD8+. Les peptides, préalablement dégradés par le protéasome pénètrent dans le réticulum endoplasmique via les transporteurs TAP. Les molécules de classe I néosynthétisées s'associent aux molécules TAP par l'intermédiaire de la protéine tapasine, chargent un peptide, acquièrent leur conformation finale et sont exportées vers la surface.

Chez le poulet, deux gènes de classe I ont été identifiés dans la région B du CMH : *B-F1* et *B-F2*. Ils codent pour la chaîne lourde des molécules d'histocompatibilité de classe I qui sont en effet des hétérodimères composés d'une chaîne lourde  $\alpha$  et d'une chaîne légère, la  $\beta$ 2-microglobuline. Le gène codant pour la  $\beta$ 2-microglobuline a été décrit chez le poulet, sur le chromosome 10 (Riegert *et al.*, 1996). Chez les Mammifères ce gène est également situé sur un chromosome différent de celui arborant le CMH (chromosome 15 chez l'Homme, 2 chez la souris). La chaîne lourde  $\alpha$  de classe I est formée de trois domaines extracellulaires  $\alpha$ 1,  $\alpha$ 2 et  $\alpha$ 3, les domaines  $\alpha$ 1 et  $\alpha$ 2 présentent en général le plus grand polymorphisme, en relation avec l'implication des acides aminés qu'ils contiennent dans les liaisons avec les différents peptides présentés, c'est la *peptide binding region*. L'organisation (nombre et taille des exons) de ces gènes est très conservée entre les Mammifères et les Oiseaux. Chaque exon correspond à un domaine fonctionnel de la protéine, l'exon 1 codant pour le signal peptide, les exons 2 à 4 pour les domaines  $\alpha$ 1,  $\alpha$ 2 et  $\alpha$ 3, l'exon 5 codant pour la région transmembranaire et les exons 6 à 8 codant pour le segment cytoplasmique. Les gènes du poulet se distinguent toutefois par la très faible taille de leurs introns (50-200 pb contre 1-2 kb chez les Mammifères). Par ailleurs les éléments de régulation qui se trouvent dans la région promotrice des gènes *B-F1* et *B-F2* sont semblables à ceux observés chez les Mammifères. La présence de motifs caractéristiques a été décrites, comme ENH-A, ISRE, W/S, site  $\alpha$  et ENH-B (Kroemer *et al.*, 1990). La transcription de ces gènes est inductible par interféron  $\gamma$  (Zoller *et al.*, 1992). Un seul de ces deux gènes serait majoritairement exprimé, le motif ENH-A du gène minoritaire a été décrit comme délété chez des poulets B12 et B19 et modifié chez des

poulets B2, B4 et B21 (Shaw *et al.*, 2007). Il est à noter que la nomenclature des deux gènes est variable et n'est donc pas un critère pour définir le niveau d'expression d'un gène *B-F*, bien que le gène minoritairement exprimé soit souvent noté *B-F1*. Il a été proposé (Livant *et al.*, 2004) que la molécule issue de *B-F1*, en relation avec un polymorphisme faible, soit plutôt un équivalent des molécules HLA-C et soit impliqué dans la régulation des cellules NK (Natural Killer).

Ces gènes *B-F* sont organisés en orientation transcriptionnelle opposée, juste à côté des deux gènes *TAP*, *TAP1* et *TAP2*, également en orientation transcriptionnelle opposée. Les gènes *TAP1* et *TAP2* codent pour deux protéines de la famille des transporteurs membranaires qui s'associent donc pour former un hétérodimère localisé à la surface du réticulum endoplasmique et effectuer la translocation, ATP-dépendante, des peptides cytosoliques à travers la membrane du réticulum endoplasmique. Les gènes *TAP* du poulet ont montré un polymorphisme important (Sironi *et al.*, 2008).

La proximité des gènes *B-F* et des gènes *TAP* chez le poulet, le fait qu'un seul gène *B-F* soit majoritairement exprimé, et les interactions fortes de leurs produits dans la voie cytosolique de présentation des antigènes a conduit à proposer un modèle de co-adaptations locales entre les allèles de ces gènes proches et par conséquent fortement liés, au sein d'un haplotype (Kaufman *et al.*, 1995).

Le gène codant pour la protéine tapasine, noté *TAPBP*, est situé dans la région de classe II chez le poulet, ce qui n'est pas le cas chez l'Homme. Il semblerait que ce gène soit soumis à épissage alternatif avec évitement possible de l'exon 4 (Frangoulis *et al.*, 1999).

Enfin, les gènes *LMP* (Low Molecular weight Protein) qui codent, chez les Mammifères pour les sous-unités du protéasome, le complexe d'enzymes protéolytiques assurant la dégradation des protéines cytosoliques, n'ont pas été identifiés chez le poulet. Cette absence a été mise en relation avec le fait que les molécules de classe I du poulet ont la capacité de charger une plus grande variété de peptides par rapport à ce qui a pu être observé chez les Mammifères où l'existence du protéasome permet une certaine homogénéisation de leur propriétés physico-chimiques (Kaufman *et al.*, 1995).

## Gènes de classe II et voie de présentation des antigènes exogènes

La voie de présentation des antigènes exogènes ou voie endocytaire est décrite chez les Mammifères comme la voie aboutissant à la présentation de peptides exogènes par les molécules de classe II, aux lymphocytes T auxiliaires CD4+. Les molécules de classe II s'assemblent dans le réticulum endoplasmique et y sont stabilisées par une troisième protéine, la chaîne invariante (notée Ii). Le gène codant pour la chaîne invariante a été identifié chez le poulet, sur le chromosome 13, et chez les Mammifères également ce gène est situé sur un chromosome différent de celui arborant le CMH (chromosome 5 chez l'Homme, 18 chez la souris). Lors d'une réaction catalysée par les molécules DM, cette molécule invariante est remplacée par des peptides provenant de la dégradation endo-lysosomiale de protéines extracellulaires. Les molécules de classes II chargées sont finalement transportées à la surface.

D'une façon similaire aux gènes de classe I, deux gènes de classe II ont été identifiés chez le poulet, dans la région B du CMH : *B-LB1* et *B-LB2*. Les molécules de classe II sont constituées d'une chaîne  $\alpha$  et d'une chaîne  $\beta$ , chacune correspondant à un gène distinct, A ou B. Chaque chaîne présente deux domaines extra-membranaires ( $\alpha 1$  et  $\alpha 2$ ,  $\beta 1$  et  $\beta 2$ ) associés par une courte séquence à une région transmembranaire et à un domaine intracytoplasmique. Les deux gènes *B-LB* de la région B ne codent donc que pour deux chaînes  $\beta$ . Leur organisation est très similaire à celle décrite chez les Mammifères, chaque exon codant pour un des domaines fonctionnels de la protéine : peptide signal, domaines extracellulaires  $\beta 1$  et  $\beta 2$ , région transmembranaire et domaine cytoplasmique. Le gène codant pour la chaîne  $\alpha$ , *B-LA*, se trouve à une distance génétique approximative de 5 cM des gènes *B-LB* (Salomonsen *et al.*, 2003), alors que chez les Mammifères, les gènes *A* et *B* codant pour les chaînes  $\alpha$  et  $\beta$  de chaque isotype sont clairement regroupés, avec au moins un gène *A* et un gène *B* disposés en tandem.

Toujours chez les Mammifères, ces gènes ne sont normalement exprimés que par l'épithélium thymique, les lymphocytes B, les macrophages et les cellules de Langerhans, ce qui semble également être le cas chez le poulet, où les motifs caractéristiques des promoteurs semblent conservés. Des motifs X, Y et GRE (un élément de réponse aux glucocorticoïdes) ont ainsi pu être identifiés (Chen *et al.*, 2000). A l'instar des gènes de classe I, un seul gène *B-LB* serait majoritairement exprimé chez le poulet (Kaufman, 2000). D'une manière similaire la

nomenclature des gènes *B-LB* est variable, le gène minoritairement exprimé étant plutôt noté *B-LB1*.

Des gènes *B-DMA* et *B-DMB* ont été prédits *in silico* chez le poulet, dans la région de classe II de la région B. Ils coderaient donc pour les chaînes  $\alpha$  et  $\beta$  des molécules DM. Ces molécules, dites molécules du CMH de classe II non-classiques, sont nécessaires, chez les Mammifères, au relargage de la molécule invariante et au chargement des peptides sur les molécules de classe II classiques, dans les endosomes tardifs.

### **Gènes B-G**

Trois gènes *B-G* ont été identifiés dans la région B chez le poulet. Ils codent pour des protéines appartenant à la superfamille des Immunoglobulines (IgSF), leur portion extracellulaire n'a pas de domaine de type PBR. Si elles ont principalement été décrites à la surface des érythrocytes et de leurs précurseurs, certaines sont exprimées par des lymphocytes ou des cellules du stroma de la bourse, du thymus et des tonsils caeaux. Elles pourraient avoir un rôle de régulation de la réponse immunitaire, plutôt comme inhibiteurs mais également un rôle dans la maturation des lymphocytes B dans la bourse de Fabricius (Plachy *et al.*, 2003). Ainsi le gène *B-G1* a récemment été mieux décrit comme codant pour une protéine appartenant à la superfamille des immunoglobulines de type 1, récepteur transmembranaire contenant un motif ITIM (Immunoreceptor Tyrosine-based Inhibition Motif) (Goto *et al.*, 2009). Cette protéine est donc similaire aux récepteurs inhibant l'activation des lymphocytes. Les gènes *B-G* sont polymorphes et soumis à épissage alternatif aboutissant à des transcrits de tailles différentes (Miller *et al.*, 1991).

### **Gènes CDI**

Les gènes *CDI* se situent chez les Mammifères en dehors du CMH (sur le chromosome 1 pour l'Homme, 3 pour la souris) : en effet ils font partie des nombreux gènes codant pour des protéines impliquées dans la réponse immunitaire qui se trouvent dans des régions paralogues du CMH chez les Mammifères. L'hypothèse la plus plausible propose qu'il y ait eu deux duplications complètes du génome comprenant la région primordiale du CMH chez les Vertébrés, aboutissant à 4 régions paralogues, dont les gènes auraient ensuite divergés ou bien été perdus. Les gènes *CDI* identifiés chez les Mammifères peuvent alors provenir de deux

histoires évolutives différentes, la première les plaçant dans le CMH primordial et ne les maintenant fonctionnels que dans une région paralogue au CMH principal, la seconde, les faisant évoluer à partir de gènes de classe I et II situés dans une région paralogue.

Le nombre de gènes *CD1* est variable chez les Mammifères : alors que la souris n'a que deux gènes *CD1*, l'Homme en présente cinq. Ces gènes codent pour des protéines aussi similaires aux molécules de classe I qu'aux molécules de classe II. De plus, alors que les molécules *CD1* s'organisent comme les molécules de classe I en 3 domaines extracellulaires qui s'associent avec la  $\beta$ 2-microglobuline, les voies qu'elles empruntent dans la cellule sont celles des molécules de classe II. En effet les molécules *CD1* chargent également des molécules dans des vésicules, mais des molécules hydrophobes, par exemple des glycolipides. Ce sont des lymphocytes T particuliers et des cellules NK qui interagissent avec les *CD1* à la surface des cellules.

L'identification de gènes *CD1* chez le poulet est une première chez un non-Mammifères. Le fait qu'ils se situent dans le CMH du poulet permet de privilégier la première hypothèse, celle de leur présence dans le CMH primordial des Vertébrés. Deux gènes *CD1* ont été caractérisés chez le poulet (Miller *et al.*, 2005, Salomonsen *et al.*, 2005), *CD1A1* et *CD1A2* leur expression est forte dans la bourse de Fabricius et dans la rate et l'un des deux gènes serait soumis à un épissage alternatif.

## **Gènes *TRIM* et domaines B30.2**

La famille des gènes *TRIM*, caractérisé par le motif TRIM (TRIPartite Motif) englobe un sous-ensemble de protéines, les TRIM-B30.2 qui contiennent les trois domaines RING, boîte B (B-Box), et domaine « coiled-coil » (hélice  $\alpha$ ) mais également un domaine B30.2. Le domaine B30.2 a été impliqué, chez les Mammifères, dans des interactions avec des immunoglobulines. Les gènes *BTN* des Mammifères, de type butyrophiline, contiennent également un domaine B30.2. On trouve chez l'Homme, dans la région de classe I étendue du CMH, des gènes *TRIM-B30.2* (*TRIM10*, 15, 26, 39) et dans une région plus télomérique, les gènes *TRIM-B30.2*, *TRIM27* et 38 ainsi que des gènes *BTN*. Cependant certains autres membres de la famille des gènes *TRIM-B30.2* ne sont pas situés dans le CMH chez l'Homme, notamment *TRIM41* et *TRIM7* qui sont sur le chromosome 5, adjacent à un gène *GNB2L1*. Chez les Mammifères, certains gènes *TRIM* ont une fonction impliquée dans la réponse immunitaire innée aux virus, en particulier *TRIM19* dans l'inhibition de la réplication ou l'expression de virus et *TRIM5 $\alpha$* , protégeant certains Primates de l'infection par le VIH-1. Des

gènes *TRIM* et *TRIM-B30.2*, d'autres gènes contenant des domaines B30.2 notés *B-BTN1*, *B-BTN2*, ainsi qu'un équivalent du gène *GNB2L1* (parfois noté *c12.3*), ont été identifiés dans la région B du CMH du poulet (Ruby *et al.*, 2005).

### **Gènes *B-lec***

Les protéines de la superfamille des lectines de type C, caractérisées par un domaine CTLD (C-Type Lectin-like Domain) comprennent, chez les Mammifères, de nombreuses protéines ainsi qu'un groupe de récepteurs des cellules NK, dont un grand nombre interagissent avec les molécules de classe I. Si un des gènes de ce type identifié chez le poulet est bien un récepteur NK, exprimé dans les cellules NK (noté *B-NK* ou *Blec 2*), les autres seraient plutôt des gènes codant pour des récepteurs impliqués dans la maturation et la régulation des cellules NK, gènes regroupés dans une région différente de celle du CMH ou de ses régions paralogues chez l'Homme, sur le chromosome 12 (Rogers *et al.*, 2005). *B-NK* (*Blec2*) coderait donc pour un récepteur inhibiteur exprimé à la surface des cellules NK, les autres *B-lec*, *Blec1* et *Blec3* coderaient plutôt pour d'autres types de récepteurs, alors que *Blec4* serait un pseudogène, avec un seul exon. Ces trois gènes (*Blec1*, *Blec2*, *Blec3*) codent pour des protéines glycosylées de type II qui forment des homodimères stabilisés par des ponts disulfures. Leurs fonctions et ligands précis restent indéterminés (Rogers et Kaufman 2008).

### **Gène *HEP21***

Parmi les particularités observées chez le poulet dans la région B du CMH, on note également le gène *HEP21*, exprimé dans l'oviducte, codant pour une protéine sécrétée dans le blanc d'œuf, similaire aux protéines de la famille LY6 des Mammifères (Nau *et al.*, 2003).

### **Autres gènes de la région B**

Les gènes actuellement annotés dans la région B du CMH du poulet sont présentés dans le tableau 2, d'après l'article de Shiina *et al.*, 2007.



**Tableau 2.** Comparaison des gènes de la région B du CMH du poulet avec les gènes d'autres espèces (Shiina *et al.*, 2007)

| Locus           | Species <sup>b</sup> | Chromosome Location in Humans <sup>c</sup> | Prominent Features of the Highest Amino Acid Similarity             | Amino Acid Acc. Num. | % Positive | E Value | Conserved Domain <sup>d</sup> |
|-----------------|----------------------|--|---|----------------------|------------|---------|-------------------------------|
| <i>BG3</i>      | Mouse                | <u>Extended class I</u>                    | MOG precursor   | Q61885               | 46         | 6e-9    | TraB_pillus                   |
| <i>BG2</i>      |                      | <b>None found</b>                          | <b>No significant homology</b>                                      |                      |            |         |                               |
| <i>KIFC1</i>    | African clawed frog  | <u>Extended class II</u>                   | C-terminal kinesin 2  | P79955               | 65         | 3e-140  | KISc_C-terminal               |
| <i>Blec3</i>    | Mouse                | 12p13-p12                                  | Early activation Ag CD69  | P37217               | 52         | 1e-19   | CLECT                         |
| <i>Bzfp3</i>    | Mouse                | 11q12                                      | Zinc finger protein 91 (Zfp-91)                                     | Q62511               | 68         | 4e-54   | SFP1                          |
| <i>TRIM72</i>   | Mouse                | 5q35.3                                     | TRIM protein 7  | Q923T7               | 58         | 6e-79   | RING BBOX<br>PRY SPRY         |
| <i>Bzfp2</i>    | Human                | <u>19p13.1-p12</u>                         | Zinc finger protein 43 (Zinc protein HTP6)                          | P17038               | 49         | 4e-67   | KRAB<br>Herpes_LMP2           |
| <i>Bzfp1</i>    | Human                | <u>1p36</u>                                | Zinc finger protein 436   | Q9C0F3               | 59         | 4e-89   | KRAB COG5048                  |
| <i>44G24.1</i>  |                      | <b>None found</b>                          | <b>No significant homology</b>                                      |                      |            |         |                               |
| <i>LAO</i>      | snake                | <u>19q13.3-q13.4</u>                       | L-amino acid oxidase precursor                                      | O93364               | 70         | 2e-149  | Amino_oxidase                 |
| <i>TRIM7.1</i>  | Human                | 5q35.3                                     | TRIM protein 7  | Q9C029               | 70         | 7e-139  | RING BBOX<br>BBC PRY<br>SPRY  |
| <i>HEP21</i>    |                      | <b>None found</b>                          | <b>No significant homology</b>                                      |                      |            |         |                               |
| <i>TRIM39.2</i> | Rat                  | <u>Classical class I</u>                   | TRIM protein 39   | Q6MFZ5               | 62         | 6e-86   | RING BBOX<br>PRY SPRY         |
| <i>TRIM27.2</i> | Human                | <u>Extended class I</u>                    | TRIM protein 27   | P14373               | 55         | 1e-75   | RING BBOX<br>PRY SPRY         |
| <i>TRIM39.1</i> | Rat                  | <u>Classical class I</u>                   | TRIM protein 39   | Q6MFZ5               | 60         | 2e-44   | PRY SPRY                      |
| <i>TRIM27.1</i> | Human                | <u>Extended class I</u>                    | TRIM protein 27   | P14373               | 57         | 3e-86   | RING BBOX<br>PRY SPRY         |
| <i>TRIM41</i>   | Human                | 5q35.3                                     | TRIM protein 41   | Q8WV44               | 69         | 8e-131  | RING BBOX<br>PRY SPRY         |
| <i>GNE2L1</i>   | Human                | 5q35.3                                     | Guanine nucleotide-binding protein $\beta$ subunit 2-like 1         | P63244               | 100        | 1e-179  | WD40                          |
| <i>B-BTN1</i>   | Rat                  | <u>Classical class I</u>                   | TRIM protein 39   | Q6MFZ5               | 57         | 2e-39   | PRY SPRY                      |
| <i>B-BTN2</i>   | Human                | <u>Extended class I</u>                    | TRIM protein 27   | P14373               | 58         | 6e-12   |                               |
| <i>BG1</i>      | Mouse                | <u>Extended class I</u>                    | MOG precursor   | Q61885               | 56         | 1e-22   | IG_like<br>COG3883            |
| <i>Blec2</i>    | Human                | 12p12.3-13.2                               | Killer cell lectin-like receptor subfamily F, member 1              | Q9NZS2               | 45         | 3e-09   | CLECT                         |
| <i>Blec1</i>    | Human                | 12p13-p12                                  | C-type lectin domain family 2 member B                              | Q92478               | 64         | 3e-26   | CLECT                         |
| <i>BLB1</i>     | Human                | <u>Classical class II</u>                  | HLA class II histocompatibility Ag, DRB1-1 $\beta$ -chain precursor | P04229               | 70         | 5e-57   | MHC_II_ $\beta$ IGc           |
| <i>TAPBPL</i>   | Human                | <u>Extended class II</u>                   | Tapasin precursor   | O15533               | 46         | 3e-25   | IG_like IGc                   |
| <i>BLB2</i>     | Human                | <u>Classical class II</u>                  | HLA class II histocompatibility Ag, DRB1-8 $\beta$ -chain precursor | Q30134               | 67         | 2e-56   | MHC_II_ $\beta$ IGc           |
| <i>RING3</i>    | Human                | <u>Classical class II</u>                  | Bromodomain-containing protein 2                                    | P25440               | 80         | 0.0     | 2 BROMO                       |
| <i>DMA</i>      | Mouse                | <u>Classical class II</u>                  | Class II histocompatibility Ag, M $\alpha$ -chain precursor         | P28078               | 58         | 1e-37   | MHC_II_ $\alpha$ IGc          |
| <i>DMB1</i>     | Mouse                | <u>Classical class II</u>                  | Class II histocompatibility Ag, M $\beta$ 1 chain precursor         | P35737               | 54         | 3e-33   | IGc                           |
| <i>DMB2</i>     | Human                | <u>Classical class II</u>                  | HLA class II histocompatibility Ag, DM $\beta$ -chain precursor     | P28068               | 56         | 2e-44   | IGc                           |
| <i>BF1</i>      | Mouse                | <u>Classical class I</u>                   | H-2 class I histocompatibility Ag, K-W28 $\alpha$ -chain precursor  | P03991               | 58         | 2e-62   | MHC_I IGc                     |
| <i>TAP1</i>     | Human                | <u>Classical class II</u>                  | Ag peptide transporter 1  | Q03518               | 67         | 6e-130  | ABC_ATPase                    |
| <i>TAP2</i>     | Human                | <u>Classical class II</u>                  | Ag peptide transporter 2  | Q03519               | 68         | 2e-141  | ABC_ATPase                    |
| <i>BF2</i>      | Mouse                | <u>Classical class I</u>                   | H-2 class I histocompatibility Ag, K-K $\alpha$ -chain precursor    | P04223               | 55         | 1e-57   | MHC_I IGc                     |
| <i>C4</i>       | Rat                  | <u>Class III</u>                           | Complement C4 precursor   | P08649               | 54         | 0.0     | A2M_N ANATO<br>A2M C345C      |
| <i>CenpA</i>    | Human                | 2p24-p21                                   | Centromere protein A  | P49450               | 85         | 1e-16   | H3                            |
| <i>CYP21</i>    | Human                | <u>Class III</u>                           | Cytochrome P450   | P08686               | 57         | 9e-74   | p450                          |
| <i>TNXB</i>     | Human                | <u>Class III</u>                           | Tenascin-X precursor  | P22105               | 58         | 0.0     | 14 FN3 FRcD                   |
| <i>LTB4R1</i>   | Human                | 14q11.2-q12                                | Leukotriens B4 receptor 1   | Q15722               | 56         | 8e-33   | 7tm_1                         |
| <i>CD1A2</i>    | Sheep                | <u>1q22-q23</u>                            | T cell surface glycoprotein CD1b-2 precursor                        | Q29422               | 53         | 2e-29   | Gc                            |
| <i>CD1A1</i>    | Human                | <u>1q22-q23</u>                            | T cell surface glycoprotein CD1b precursor                          | P29016               | 51         | 4e-34   | MHC_I IGc                     |

<sup>a</sup> Comparisons identifying homologies were made with GENETYX complemented by BLAST and SwissProt searches. Genes without significant match scores are shown in bold.

<sup>b</sup> Species in which highest scoring matches were found

<sup>c</sup> Location of homologous genes in human genome. Underlined values denote homologs present in the human MHC and its paralogous regions.

<sup>d</sup> Conserved domains revealed in *MHC-B* genes.

Le gène *RING3* est également appelé *BRD2* et le gène *CenpA*, *B-H3*.

### 3. Autres gènes du CMH du poulet

Parmi les gènes présents sur GGA16, identifiés mais dont la position sur la séquence n'a pas pu être définie, se trouvent les gènes situés dans la région Y, c'est à dire au moins un gène de classe I non-classique transcrit, deux gènes de classe II B non-classiques et des gènes de lectines de type C. Ces gènes, parfois appelés *Y-lec* coderaient pour des ligands des récepteurs des cellules NK (Shiina, 2006)

Il s'agit également du gène *fB*, codant pour le facteur B du complément (Kaufman *et al.*, 1999a) et du gène *G9 (BAT8)* (Spike et Lamont, 1995), tous deux présents dans la région de classe III du CMH chez l'Homme et du gène de class II classique *B-LA*, codant pour la seule chaîne  $\alpha$  connue de classe II classique du poulet, dont seul le cDNA a pu être séquencé (Salomonsen *et al.*, 2003).

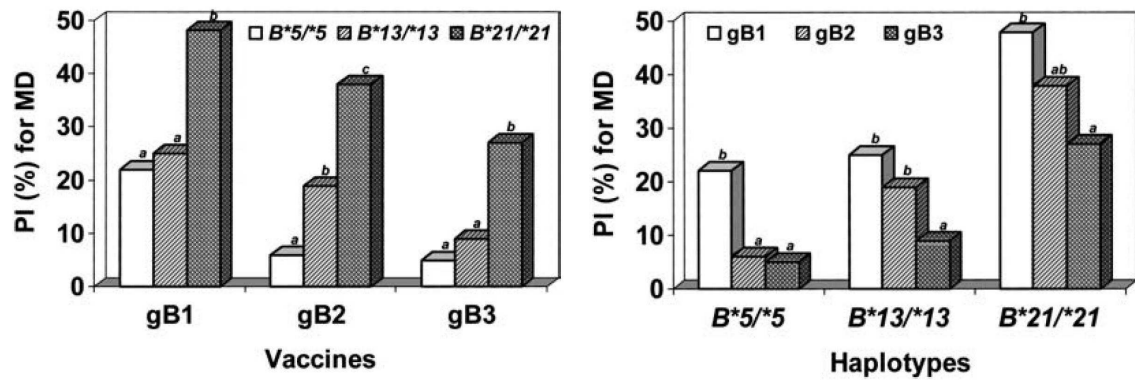
#### IV. ROLE DU CMH DANS LA RESISTANCE AUX MALADIES CHEZ LE POULET

Le premier exemple documenté d'association entre un CMH et la résistance à une maladie l'a été chez le poulet en 1977, quand Briles a décrit l'exceptionnelle résistance conférée par le sérotype B21 au virus de la maladie de Marek (MDV). D'autres associations de ce type ont été décrites par la suite, notamment chez la souris, avec des résistances au virus LCMV (Lymphocytic ChorioMenengitis Virus) et au virus de Theiler, associées à la région de classe I du CMH (Zinkernagel *et al.*, 1985, Monteyne *et al.*, 1997). Mais également, plus récemment avec des associations entre des locus du MHC et une résistance au virus ISAV (Infectious Salmon Anaemia Virus) mais également à la bactérie *Aeromonas salmonicida* chez le saumon atlantique (*Salmo salar*, Grimholt *et al.*, 2003). Par contre chez l'Homme, peu d'associations claires ont été démontrées entre des gènes du HLA et des résistances à des maladies infectieuses. En fait, la plupart des études ont démontré une plus grande susceptibilité plutôt qu'un effet protecteur de certains allèles avec une association concernant plutôt des gènes de classe II que des gènes de classe I (Hill, 2001) sauf dans le cas bien connu et documenté de la résistance conférée par certains allèles de classe I mais également de classe II, à la malaria (Hills, 1991). Bien que le taux de progression du virus HIV (human immunodeficiency virus) ou la susceptibilité à un virus de l'herpès (HSV2) aient également pu être associés à des allèles de classe I du HLA (Kroner *et al.*, 1995, Kaslow *et al.*, 1990, Lekstrom-Himes, 1999), ainsi que la présence de l'allèle HLA-A2 avec une charge réduite en virus HTLV-I (human T-lymphotropic virus I) et une protection contre la myélopathie associée à HTLV (Jeffery *et al.*, 1999). En comparaison avec la connaissance du complexe HLA et le nombre d'études réalisées sur le sujet, cette liste d'association est particulièrement courte. Ceci peut être mis en relation avec le fait que les études d'association ne sont pas faciles chez l'Homme, compte tenu de la diversité des allèles présents et de l'hétérogénéité des populations, et également de la complexité de la région HLA.

Le poulet devient dès lors, avec son CMH compact et simple, un bon modèle d'étude. Ce qui a été exacerbé par le fait que la zone centrale *B-F/B-L* de la région B du CMH a été associée à une importante résistance à plusieurs virus, comme le sarcome de Rous (RSV), la maladie de Marek (MDV), la leucose aviaire (lymphoid leukosis virus, Briles *et al.*, 1977, Bacon *et al.*, 1981, 2001, Lamont 1998) et la maladie de Gumboro (Infectious Bursal Disease Virus, Hudson *et al.*, 2002). On peut noter que plusieurs de ces virus ont pour cible les cellules du système immunitaire lui-même. Des associations ont également été observées entre des

haplotypes du CMH et la résistance à différentes bactéries, comme *Brucella abortus*, *Pasteurella multocida*, *Salmonella enteritidis* et *Staphylococcus aureus* (Lamont 1998, Zhou et Lamont, 2003). Ce sont en particulier les sérotypes B21, B130 et BW1 qui ont démontré une forte résistance à MDV, les sérotypes B2, B6, B14, une résistance modérée, les sérotypes B5, B15, B12 et B19 étant susceptibles (Plachy *et al.*, 2003). Actuellement, ce sont des différences de niveau et de type de transcription (Dalgaard *et al.*, 2003), voire de polymorphismes des gènes de classe I qui ont été identifiés comme différenciant entre les sérotypes B21 et B19. Le gène *Blec2* (Shiina *et al.*, 2006) a également été impliqué et le rôle des régions UTR des gènes *B-G* (Hosomichi *et al.*, 2008) a été confirmé récemment (Goto *et al.*, 2009) dans le cas des haplotypes recombinants BR2 et BR4 issus des haplotypes B2, B23 et B24. Par ailleurs, ce sont les gènes de classe I *B-F* qui ont été proposés pour la résistance au RSV (Hofmann *et al.*, 2003, Wallny *et al.*, 2006). De plus des sérotypes B ont également été associés à des maladies auto-immunes spontanées de type thyroïdite (Bacon *et al.*, 1977), vitiligo (Wang et Erf, 2004). La région Y du CMH du poulet a également été associée à une résistance au MDV (Lakshmanan et Lamont, 1998) et au RSV (Pinard van-der-Laan *et al.*, 2004).

Le contrôle ou la gestion des maladies infectieuses chez les animaux d'élevage font appel à un ensemble de pratiques, dont l'utilisation d'antibiotiques et d'antiparasitaires, la vaccination, les bonnes pratiques sanitaires, la traçabilité et les contrôles des animaux et des produits. La prise en compte de la génétique des pathogènes comme celles des hôtes et leur interaction sont maintenant reconnues comme nécessaires afin d'améliorer l'efficacité de ces pratiques, de les compléter ou parfois de les remplacer. Il est également important de pouvoir faire face à l'augmentation de la résistance des pathogènes, qu'ils soient parasites, bactéries ou virus, le virus de Marek étant particulièrement concerné, avec l'apparition d'un variant hypervirulent sélectionné à la suite des campagnes intensives de vaccination (Gandon *et al.*, 2001). Il s'agit également de s'adapter à de nouvelles réglementations ou bien d'utiliser les ressources génétiques propres aux races locales afin d'améliorer à bas coûts la situation dans les pays du Sud. Cependant les conséquences de l'introduction dans une filière d'animaux génétiquement résistants ou tolérants ne sont pas facile à prédire.



**Figure 7.** Réponse vaccinale pour la Maladie de Marek pour 3 lignées congéniques de poulet différant par leur sérotype au CMH. Après vaccination, la taille de l'histogramme représente l'effet protecteur du vaccin contre la mort ou les lésions causées par le virus. Quelque soit le vaccin utilisé, certains sérotypes ont mieux répondu. ( Lee *et al.*, 2004)

Une importante application de l'immunogénétique concernera l'évaluation de la part génétique de la variabilité des réponses individuelles aux infections et aux vaccinations. En effet dans ce type d'étude chez les animaux domestiques, il s'agit de déterminer l'importance relative des facteurs de type race et famille, des facteurs de type âge, cohorte, sexe et allèles de certains gènes dont le CMH. Un exemple peut être illustré chez le poulet avec les différences de protection apportées par la vaccination face au MDV en fonction du sérotype B du CMH concerné (Lee *et al.*, 2004, Figure 7).

## **V. PRESENTATION DE LA THESE**

La première partie du projet de thèse a consisté en l'exploration de la diversité du CMH via le marqueur LEI0258 dans les collections d'ADN maintenues au laboratoire, afin de proposer un mode d'évolution pour ce marqueur atypique.

La deuxième partie de la thèse repose sur l'étude de lignées expérimentales White Leghorn, disponibles au laboratoire, à l'aide du marqueur LEI0258 et d'un jeu réduit de SNPs, afin d'évaluer les possibilités offertes par ces marqueurs et leur corrélation avec la sérologie.

La troisième partie a étendu cette étude à un plus grand nombre de marqueurs et de races, en s'appuyant sur l'identification, par reséquençage, de nouveaux polymorphismes dans 11 gènes cibles de la région dans un large panel de races expérimentales, commerciales et locales.

La quatrième partie a plus particulièrement concerné la caractérisation de la structure génomique et de la transcription des trois gènes du CMH du poulet correspondant aux gènes *DM* des Mammifères, avec l'étude de leur polymorphisme dans trois espèces de *Gallus* sauvages et le même large panel de races.

Dans un cinquième temps, les polymorphismes identifiés et caractérisés au cours de la thèse ainsi que ceux disponibles dans les bases de données permettent la mise au point d'une puce 96 SNPs dédiée au CMH chez le poulet.

En complément, un résumé de l'étude exploratoire des gènes de classe I classiques par clonage sera présenté.

Pour conclure, une discussion plus générale sera proposée, dans un dernier chapitre.

# MATERIEL & METHODES

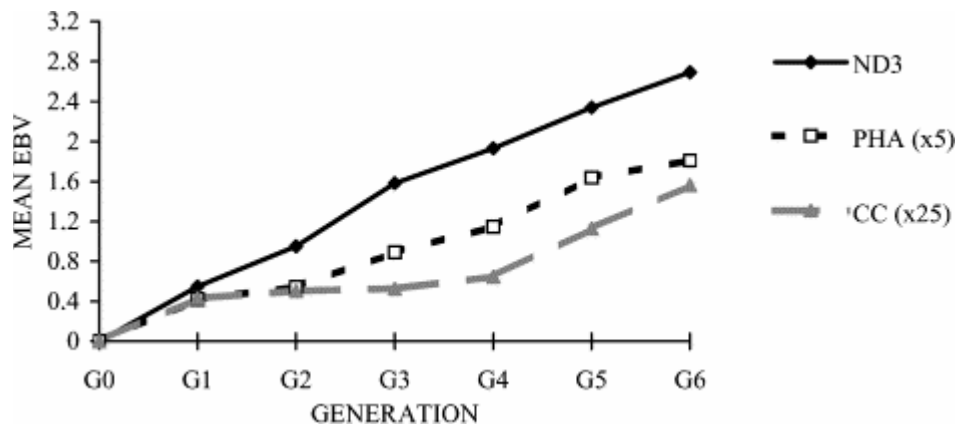
## I. MATERIEL ANIMAL

### A. Animaux utilisés tout au long du projet de thèse : les lignées R.I.

Trois lignées Réponse Immune (R.I.), de race White Leghorn, ont été sélectionnées à l'INRA depuis 15 générations sur différents critères relatifs à la réponse immunitaire par Marie-Hélène Pinard-van der Laan. Les animaux fondateurs étaient 9 coqs et 44 poules provenant d'un croisement entre une lignée White Leghorn expérimentale porteuse d'un allèle de nanisme lié au sexe (*dw*, pour *dwarf*) et une lignée commerciale Babcock White Leghorn. Les descendants de ce croisement ont produit 523 poulets non sélectionnés (génération  $G_0$ ), à partir desquels quatre lignées ont été constituées, trois lignées sélectionnées chacune sur un critère différent, et une lignée témoin non sélectionnée. Ces critères de sélection sont les suivants : lignée 1, réponse anticorps après vaccination contre le virus de la maladie de Newcastle (vaccin HB1) ; lignée 2, hypersensibilité retardée en réponse à la phytohémagglutinine (réponse cellulaire) ; lignée 3, clairance du carbone comme indicateur de la phagocytose. Chaque année 800 animaux, nés dans un seul lot, ont été testés pour les trois critères relatifs à la réponse immunitaire, avec 200 animaux par lignée, 100 mâles et 100 femelles (Figure 8).

Au sein de chaque lignée et pour chaque génération, 15 mâles et 30 femelles sont choisis comme reproducteurs sur les 100 candidats de chaque sexe. La sélection est basée sur la performance phénotypique individuelle. Le plan d'accouplement est réalisé de façon à minimiser l'augmentation de la consanguinité, afin d'éviter les problèmes à long terme. Les héritabilités, estimées à la sixième génération pour les trois critères de sélection ND3, PHA et CC, sont modérées, mais significatives, respectivement  $0,35 (\pm 0,02)$ ,  $0,13 (\pm 0,02)$  et  $0,15 (\pm 0,02)$ . Les corrélations phénotypiques et génétiques entre les critères ne sont pas significatives (Pinard-van der Laan, 2002). Des croisements entre les lignées et des backcross ont également été réalisés et étudiés pour les trois critères de sélection mais également pour d'autres critères relatifs à la réponse immunitaires (Minozzi *et al.*, 2008). D'autres mesures ont également été réalisées sur les animaux, comme le poids corporel et la température pour la quatrième génération. En outre, tous les animaux du début de l'expérience de sélection à la





**Figure 8.** Réponse à la sélection pour ND3 (réponse anticorps) de la lignée 1, pour PHA (réponse cellulaire) de la lignée 2 et pour CC (phagocytose) de la lignée 3, représentée par la moyenne des valeurs génétiques estimées (Pinard-van der Laan 2002)

génération neuf ont été typés en sérologie pour le CMH, puis à l'aide d'un typage moléculaire. À la sixième génération, les effets des sérotypes au CMH ont été estimés. Les fréquences des différents sérotypes varient avec le temps, par exemple en 2006, le sérotype B19 n'était plus présent dans la lignée 1 mais était le plus fréquent dans la lignée 2. L'effet global de chaque sérotype a été estimé (Loywyck *et al.*, 2006), des associations significatives avec le CMH ont été trouvées pour chaque critère, mais aucun des cinq sérotypes n'a été associé à une réponse plus élevée pour l'ensemble des trois critères.

## **B. Animaux utilisés pour le reséquençage du CMH**

Pour la partie concernant le reséquençage du CMH, un référentiel de 48 populations a été défini, en partie à partir d'un panel de populations établi en 1999 pour le projet européen AvianDiv. Il inclut 6 lignées commerciales autres que White Leghorn (4 lignées de poulets de chair et 2 lignées de pondeuses œufs bruns), 17 populations White Leghorn (issues de 15 lignées expérimentales dont les lignées R.I. et une lignée commerciale pondeuse) et 25 autres races dont 12 races européennes, 7 races locales asiatiques et 6 populations locales africaines. Chaque population est représentée par 2 animaux, un mâle et une femelle quand le sexe des animaux était décrit.

### **1. Echantillonnage AvianDiv (avec numéro de population de référence)**

L'INRA est gestionnaire de la banque d'ADN issue du projet européen AvianDiv, projet rassemblant 53 populations (<http://aviandiv.tzv.fal.de>). Cette banque a fourni 29 des 48 populations du référentiel : 7 lignées commerciales dont une lignée commerciale White Leghorn, 8 lignées expérimentales, dont 4 lignées homozygotes pour le CMH (C line), et 14 races provenant d'Europe ou du pourtour Méditerranéen.

Le tableau 3 récapitule les données concernant le mode de gestion, la taille, l'origine géographique et la période de création de ces populations qui sont identifiées par leur nom et leur numéro de référence AvianDiv.

**Tableau 3.** Principales caractéristiques des populations étudiées

| Nom et numéro AvianDiv                          |             | Mode de Gestion * | Pays d'échantillonnage | Population Fondatrice ** | Période de Creation *** | Taille de la population (mini-maxi) |
|---|-------------|-------------------|------------------------|--------------------------|-------------------------|-------------------------------------|
| <i>Gallus gallus spadiceus</i>                  | <b>101</b>  | <b>1</b>          | Thaïlande              | B                        | <b>1997</b>             | <b>100</b>                          |
| <i>Gallus gallus gallus</i>                     | <b>102</b>  | <b>1</b>          | Thaïlande              | B                        | <b>1997</b>             | <b>165</b>                          |
| Fayoumi   | <b>04</b>   | <b>3</b>          | France                 | B                        | <b>1978</b>             | <b>50-300</b>                       |
| Bedouin   | <b>05</b>   | <b>2</b>          | Israël                 | A                        | <b>1995</b>             | <b>30-50</b>                        |
| Westfaeliche Totleger                           | <b>06</b>   | <b>3</b>          | Allemagne              | B                        | <b>1904</b>             | <b>900</b>                          |
| Sundheimer                                      | <b>07</b>   | <b>3</b>          | Allemagne              | C                        | <b>1886</b>             | <b>100-500</b>                      |
| Light Brown Leghorn                             | <b>08</b>   | <b>3</b>          | Allemagne              | B                        | <b>1870</b>             | <b>&gt; 200</b>                     |
| Uilenbaarder                                    | <b>09</b>   | <b>3</b>          | Pays-Bas               | A                        | <b>1650</b>             | <b>200</b>                          |
| Friesian fowl                                   | <b>10</b>   | <b>3</b>          | Pays-Bas               | A                        | .                       | <b>50</b>                           |
| Bresse noire                                    | <b>11</b>   | <b>4</b>          | France                 | B                        | <b>1995</b>             | <b>400-2500</b>                     |
| Houdan  | <b>12</b>   | <b>3</b>          | France                 | C                        | <b>1994</b>             | <b>50-200</b>                       |
| Marans  | <b>13</b>   | <b>3</b>          | France                 | C                        | <b>1988</b>             | <b>200-350</b>                      |
| Dorking   | <b>14</b>   | <b>3</b>          | Royaume-Uni            | B                        | <b>1986</b>             | <b>85</b>                           |
| Cochin  | <b>15</b>   | <b>3</b>          | Royaume-Uni            | B                        | <b>1946</b>             | <b>130</b>                          |
| Icelandic landrace                              | <b>16</b>   | <b>2</b>          | Islande                | A                        | <b>900</b>              | <b>2000-4000</b>                    |
| Finnish Landrace                                | <b>17</b>   | <b>2</b>          | Finlande               | C                        | <b>1900</b>             | <b>600-1000</b>                     |
| Old Scand. Ref. Pop.                            | <b>18</b>   | <b>3</b>          | Danemark               | C                        | <b>1969</b>             | <b>200-700</b>                      |
| Jaerhoens                                       | <b>19</b>   | <b>3</b>          | Norvège                | A                        | <b>1916</b>             | <b>300-400</b>                      |
| Sicilian Buttercup                              | <b>20</b>   | <b>3</b>          | Italie                 | B                        | <b>1990</b>             | <b>150</b>                          |
| Padova  | <b>21</b>   | <b>3</b>          | Italie                 | B                        | <b>1986</b>             | <b>35-350</b>                       |
| Black Castellana                                | <b>22</b>   | <b>3</b>          | Espagne                | A                        | <b>1975</b>             | <b>200-300</b>                      |
| Red Villafranquina                              | <b>23</b>   | <b>3</b>          | Espagne                | A                        | <b>1980</b>             | <b>200-300</b>                      |
| Czech Golden Pencilled                          | <b>24</b>   | <b>3</b>          | République Tchèque     | C                        | <b>1995</b>             | <b>500-1000</b>                     |
| Oravka hen                                      | <b>25</b>   | <b>3</b>          | Slovaquie              | C                        | <b>1994</b>             | <b>50-100</b>                       |
| Transylvanian Naked Neck                        | <b>26</b>   | <b>3</b>          | Hongrie                | A                        | <b>1990</b>             | <b>70-220</b>                       |
| Green-legged Partridge                          | <b>27</b>   | <b>3</b>          | Pologne                | A                        | <b>1950</b>             | <b>600-1600</b>                     |
| Orlov   | <b>28</b>   | <b>3</b>          | Russie                 | C                        | <b>1960</b>             | <b>2000</b>                         |
| Yurlov crower Russia                            | <b>2901</b> | <b>3</b>          | Russie                 | C                        | <b>1976</b>             | <b>10000</b>                        |
| Ukrainian bearded                               | <b>30</b>   | <b>2</b>          | Ukraine                | B                        | <b>1850</b>             | <b>74-105</b>                       |
| Poltava clay                                    | <b>31</b>   | <b>3</b>          | Ukraine                | B                        | <b>1870</b>             | <b>2000-6000</b>                    |
| C line (White Leghorn)                          | <b>32</b>   | <b>5</b>          | République Tchèque     | A                        | <b>1932</b>             | <b>180-600</b>                      |
| 4 sous-populations                              |             |                   |                        |                          |                         |                                     |
| Gödöllő Nhx (New Hampshire x Rhode White)       | <b>33</b>   | <b>4</b>          | Hongrie                | C                        | <b>1996</b>             | <b>600-8000</b>                     |
| Line Sarcoma-Resistant (White Leghorn)          | <b>3401</b> | <b>4</b>          | Allemagne              | B                        | <b>1965</b>             | <b>200</b>                          |
| Line Sarcoma-Susceptible white-egg layer line A | <b>3402</b> | <b>4</b>          | Allemagne              | B                        | <b>1965</b>             | <b>200</b>                          |
| broiler dam line B                              | <b>37</b>   | <b>4</b>          | Allemagne              | B                        | <b>1959</b>             | <b>6000</b>                         |
| broiler dam line C                              | <b>41</b>   | <b>4</b>          | France                 | A                        | <b>1960</b>             | <b>5000-30000</b>                   |
| broiler dam line D                              | <b>44</b>   | <b>4</b>          | France                 | B                        | <b>1960</b>             | <b>5000</b>                         |
| broiler sire line C                             | <b>45</b>   | <b>4</b>          | Pays-Bas               | B                        | <b>1962</b>             | <b>1000</b>                         |
| broiler sire line D                             | <b>47</b>   | <b>4</b>          | Royaume-Uni            | A                        | <b>1974</b>             | .                                   |
| broiler dam line D                              | <b>49</b>   | <b>4</b>          | Royaume-Uni            | A                        | <b>1992</b>             | <b>8000</b>                         |
| High Immune Response line                       | <b>50</b>   | <b>4</b>          | Israël                 | B                        | <b>1970</b>             | <b>5000-20000</b>                   |
| Low Immune. Response. line                      | <b>51</b>   | <b>4</b>          | Pays-Bas               | C                        | <b>1980</b>             | <b>75-300</b>                       |
|   | <b>52</b>   | <b>4</b>          | Pays-Bas               | C                        | <b>1980</b>             | <b>75-300</b>                       |

\* Mode de gestion : 1 pour une population de type sauvage, 2 pour une population domestiquée non sélectionnée, 3 pour une population standardisée sur la morphologie, 4 pour une lignée sélectionnée sur des caractères quantitatifs, 5 pour une lignée consanguine.

\*\* A = un petit troupeau B = une race; C = un croisement entre races ou des introductions successives provenant d'autres races.

\*\*\* période de création du ou des troupeaux échantillonnés, à défaut date de création du standard de la race.

### ***a) 7 lignées commerciales***

Ce sont des lignées grand-parentales sélectionnées pour différents caractères qui sont ensuite utilisées en croisement pour produire les animaux commercialisés. Elles sont toutes maintenues en élevage protégé, la qualité sanitaire des reproducteurs grand-parentaux étant un objectif imposé par les sélectionneurs. Elles ont un phénotype fixé pour les caractères de coloration de peau, de plumage et de coquille.

On peut distinguer 4 grandes catégories en fonction de l'objectif de sélection :

- les pondeuses à œufs blancs sélectionnées sur plusieurs caractères de ponte à partir de la race White Leghorn (37) ;
- les pondeuses à œufs bruns sélectionnées sur plusieurs caractères de ponte à partir de quelques races, dont la Rhode Island Red et la New Hampshire (44, 45) ;
- les poulets de chair dits de type 'femelle', c'est-à-dire que la lignée est utilisée comme grand-parent femelle dans le croisement commercial. Ce type de lignée est sélectionné sur la croissance et la reproduction. Les races fondatrices sont souvent la Plymouth Rock, mais également la New Hampshire (41, 50) ;
- les poulets de chair dits de type 'mâle' c'est-à-dire que la lignée est utilisée comme grand-parent ou parent mâle dans le croisement commercial. Ce type de lignée est sélectionné sur la croissance et la composition corporelle. Les races fondatrices sont de type 'combattant' d'origine asiatique (47, 49).

### ***b) 8 lignées expérimentales***

Par rapport aux lignées commerciales, ces lignées se caractérisent par un nombre plus faible de critères de sélection, un effectif plus limité et un élevage en conditions contrôlées mais généralement moins strictement protégées. Leur phénotype est généralement fixé, notamment lorsqu'il s'agit de lignées de race White Leghorn mais il peut également être variable.

#### **Les lignées « C line » (32)**

Cette lignée de poules White Leghorn a été créée en 1932 à partir d'une petite population provenant d'Angleterre, des laboratoires Reaseheath. Deux lignées CC et CB ont été dérivées en 1964, les poulets de la lignée CB sont homozygotes pour le sérotype B12 et les CC, homozygotes pour le sérotype B4. Ces deux lignées ont été répliquées en 1984 (CCR1 et CBR1). Les 4 lignées ainsi obtenues sont maintenues fermées et avec des accouplements

frère-sœur depuis l'origine. Le coefficient de consanguinité est de 99,9% et la sélection se fait sur l'état général. Ces lignées ont été utilisées pour étudier le déterminisme génétique de la résistance au virus du sarcome de Rous.

#### **Les lignées résistante ou susceptible au sarcome de Rous (3401 et 3402)**

Les lignées résistante ou susceptible au Sarcome sont des lignées expérimentales de White Leghorn, maintenues en Allemagne, mais importées des USA en 1965. Leur origine commune est la lignée K de Cornell et aucune introduction n'a eu lieu. La lignée résistante au sarcome, également appelée R11, a été sélectionnée jusqu'en 1974 pour la résistance aux virus A et B du sarcome de Rous, alors que la lignée sensible, également appelée R22, a été sélectionnée jusqu'en 1974 pour la sensibilité à cette infection. Le caractère résistant (ou sensible) semble se comporter comme un caractère autosomal récessif (ou dominant). Ces caractères ont été fixés dans les lignées respectives. Les deux lignées sont homozygotes pour le sérotype B15. Chaque lignée a été maintenue par accouplements non contrôlés entre 1974 et 1991 et suivant un plan de rotation depuis 1992.

#### **Les lignées High Immune Response line, Low Immune Response line (51, 52)**

Les lignées à réponse anticorps haute ou basse ont été créées en 1980 à partir d'une lignée commerciale pondeuse d'œufs bruns. Les 2 lignées ont les mêmes fondateurs et ont été sélectionnées aux Pays-bas sur la réponse immunitaire aux globules rouges de mouton (SRBC) à 6 semaines, avec le maintien d'une lignée témoin. Chaque lignée est maintenue par accouplements contrôlés de 25 mâles et 50 femelles. La coloration du plumage est variable, les plumages de la lignée à haute réponse anticorps sont surtout blancs, alors que ceux de la lignée à basse réponse anticorps sont à 50% doré à noir restreint (gold Columbian), 40% rouge homogène (solid red) et seulement à 10% blanc. La lignée témoin présente une répartition équilibrée des trois plumages. La peau est jaune, les oreillons sont bleutés et la coquille d'œuf est de couleur crème à brun.

### ***c) 14 races provenant d'Europe ou du pourtour Méditerranéen***

Il s'agit de races qui n'ont pas été soumises à une sélection intensive et qui dérivent de populations locales. Elles ont pu être échantillonnées d'autant plus facilement qu'elles font l'objet d'un programme d'étude ou de conservation. Leur phénotype est généralement standardisé, car le standard phénotypique constitue souvent le premier objectif de sélection.

#### **Fayoumi (4)**

Les poulets Fayoumi proviennent d'une d'une race égyptienne, dont un échantillon a été importé en France en 1978. Une lignée a été maintenue depuis en accouplement contrôlé et avec une intensité de sélection faible. La race Fayoumi est connue pour ses capacités de résistance à la chaleur et aux maladies, en particulier à la coccidiose et à la maladie de Marek. Le phénotype est homogène avec une barrure autosomale et des tarsi bleus. Les oreillons sont rouges avec fréquemment des taches noires. La coquille de l'oeuf est blanche et brillante.

#### **Bédouin (5)**

Les poulets Bédouin ont été échantillonnés auprès d'éleveurs dans le Sinaï. Une population d'environ 50 poulets est conservée en accouplement contrôlé dans une ferme universitaire en Israël. Le phénotype du plumage est hétérogène avec un mélange d'oiseaux noirs, blancs ou rouges, mais tous les oiseaux ont les oreillons rouges. La coquille de l'oeuf est blanche.

#### **Light Brown Leghorn ou Leghorn dorée (8)**

Il s'agit d'une race d'agrément représentée par plus de 20 troupeaux d'éleveurs amateurs. Elle présente un phénotype de type sauvage pour la coloration de plumage, avec des tarsi jaunes et des oreillons blancs. La coquille de l'oeuf est blanche.

#### **Friesian Fowl (10)**

La race Friesian Fowl est également une race d'agrément, typique des Pays-Bas, mais sans indication disponible sur son origine historique. Elle n'est représentée que par un seul troupeau fermé de 49 reproducteurs en accouplement non contrôlé. Ces oiseaux sont considérés comme étant vivaces, avec une bonne capacité de vol et de bonnes capacités d'adaptation. La coloration du plumage est assez hétérogène, en raison de la ségrégation des caractères argenté, doré, blanc dominant et barrure autosomale. Les tarsi sont bleus, les oreillons et la coquille de l'oeuf sont blancs.

### **Marans (13)**

La Marans est une race ancienne française de pondeuses, qui se caractérise par la coloration brun foncé de la coquille d'oeuf (presque chocolat). Cette race est connue pour avoir des origines ancestrales dans les races Cochin et Langshan. Le troupeau échantillonné est élevé en accouplement contrôlé par une société privée, dans un but de commercialiser cette race pour la production d'œufs de haute qualité. Les tarses sont emplumés et les oreillons rouges.

### **Icelandic Landrace (16)**

Cette population endémique à l'Islande provient des poulets importés par les premiers colons, vers 900. Ces poulets sont considérés comme présentant potentiellement une adaptation au climat froid. Ils ont été échantillonnés dans 14 élevages ne pratiquant pas de sélection. Cette population présente une très grande variabilité phénotypique, liée à la ségrégation de plusieurs gènes ayant des effets visibles. .

### **Jaerhoens (19)**

Cette race traditionnelle de Norvège a été sélectionnée dans un centre national d'élevage, pour les oeufs et la coloration du plumage, jusqu'en 1973. Elle est maintenue en conservation sous supervision de la Nordic Gene Bank depuis 1973. Son plumage barré permet l'auto-détermination du sexe par la couleur du duvet. Elle pond des oeufs blancs reconnus pour leur qualité de coquille.

### **Padovana (21)**

La poule de Padoue est une race très ancienne d'Italie. Le troupeau échantillonné a été constitué à partir d'un nombre limité de poulets choisis pour leur phénotype typique, dans une zone géographique restreinte autour de la ville de Padoue. Il est maintenu dans le but de produire une viande de grande qualité. La sélection prend en compte le poids corporel et le poids des oeufs et évite les anomalies squelettiques. Cette race présente certains caractères communs avec la race française Houdan (crête, favoris et barbe) mais ne présente pas de polydactylie, ni de plumage caillouté. Les oreillons et la coquille de l'oeuf sont blancs.

### **Red Villafrantina (23)**

La race Red Villafrantina est intégrée à un programme de conservation avec un plan d'accouplements contrôlé (rotation des mâles) depuis 1980, avec un seul troupeau fermé de

200-300 oiseaux. La coloration du plumage est homogène, de type perdrix doré, les tarses sont bleu-gris et les oreillons rouges. La coquille de l'oeuf est brun foncé.

#### **Czech Golden Pencilled (24)**

Cette race originaire de la République tchèque a été croisée avec des races importées pendant la seconde moitié du 19<sup>ème</sup> siècle. La population échantillonnée a été créée entre 1913 et 1930 et a toujours été fermée. Elle est maintenue en conservation. Des aptitudes pour la qualité de viande et de résistance au stress ont été évoquées. La coloration du plumage est homogène, de type perdrix maillé doré. Les tarses sont bleu-gris et les oreillons sont rouges. La coquille de l'oeuf est de couleur crème.

#### **Transylvanian Naked Neck (26)**

Cette race échantillonnée en Hongrie se retrouve également dans les pays limitrophes. Elle se caractérise par la présence de la mutation Cou Nu, qui diminue la surface des zones emplumées. Elle a été créée en 1950 à partir d'un petit nombre d'oiseaux. Le troupeau échantillonné est en conservation depuis 1991. Le comportement de ces animaux a été décrit comme très vivace. Leur plumage est entièrement noir, les tarses sont également noirs, les oreillons rouges. La coquille de l'oeuf est de couleur crème.

#### **Green-legged Partridge (27)**

Cette race originaire de Pologne a été créée en 1870 et pourrait avoir des origines provenant d'Italie, datant du 16<sup>ème</sup> siècle. Les animaux échantillonnés proviennent de deux troupeaux fermés, en conservation depuis plus de 30 générations. Parmi les caractéristiques rapportées pour cette race, on peut noter la résistance à la salmonelle, un faible niveau de cholestérol de l'oeuf et une qualité gustative de la viande. La coloration du plumage est de type sauvage avec oreillons rouges. Les tarses sont verts, la coquille de l'oeuf est blanche.

#### **Orlov (28)**

Cette race a été créée pendant la seconde moitié du 18<sup>ème</sup> siècle à partir de croisements entre des poulets de races locales et des poulets importés, de races telles que la Ghilany, la Old English Game, ainsi que des poulets Malay de type combattant. Des poulets Orlov provenant d'Allemagne et de Tchécoslovaquie ont également été introduits en Russie vers 1950. Le troupeau échantillonné est géré par une ferme d'Etat près de Moscou, qui distribue des poulets aux éleveurs amateurs. Une adaptation aux basses températures et des comportements de



couvaison sont évoqués. La coloration du plumage est de type sauvage avec une panachure de type ‘caillouté’, ce qui produit un plumage tricolore. Les caractéristiques typiques de la race sont l’emplument lent, la présence de favoris et la barbe, la crête en noix, la peau jaune et les oreillons rouges. La coquille de l'oeuf est blanche.

### **Yurlov crower (2901)**

Cette race a été créée à partir de croisements entre des poulets locaux et des races importées, telles que Brahma, Cochin, Langshan et des coqs de combat. La race a principalement été sélectionnée sur le timbre et la durée du chant du coq. Des poulets Yurlov apparaissent comme ancêtres de plusieurs races russes créées au 20ème siècle. Le troupeau échantillonné est maintenu dans une ferme d'Etat près de Moscou, avec environ 10000 poulets reproduits sans sélection. Une adaptation aux basses températures, des comportements de couvaison et de recherche de nourriture, ainsi que l’agressivité des coqs sont évoqués. La coloration du plumage est homogène de type ‘bouleau argenté’. Les caractéristiques typiques de la race sont l’emplument lent et la crête en rose, mais également la peau blanche, les tarses noirs et les oreillons rouges. La coquille de l'oeuf est brune.

## **2. Echantillonnage complémentaire**

### **Asie**

Le laboratoire est en collaborations régulière avec le département de sciences animales de l’université nationale de Chung-Hsing (NCHU) à Taiwan. Dans ce cadre, Chi-Sheng Chang réalise sa thèse en co-tutelle AgroParisTech/NCHU sur la diversité des races taiwanaises en relation avec la résistance aux maladies, ce qui m’a permis d’utiliser certains de ses échantillons. La NCHU maintient en conservation trois races taiwanaises Hsin-Yi, Ju-Chi et Hua-Tung et trois races importées Quemoy, Shek-Ki et Nagoya, avec un plan d’accouplement de 30 mâles reproducteurs et 60 femelles reproductrices.

Les races Hsin-Yi, Ju-Chi et Hua-Tung sont utilisées pour la production d’œufs et de viande. Ces trois races présentant des tarses noirs mais différentes colorations du plumage, la Hsin-Yi est de type brun, la Ju-Chi noire avec du brun et la Hua-Tung est noire. La Hua-Tung est une race de type combattant.

La race Quemoy est originaire de l'île de Quemoy, proche de la Chine continentale, en face de Taiwan. C'est également une race polyvalente au plumage de type sauvage et aux tarses noirs. La race Shek-ki (triple jaune) est originaire du sud de la Chine, elle est réputée pour sa viande, en relation avec sa peau, son plumage et ses tarses jaunes. Enfin la race Nagoya est originaire du Japon, c'est une race polyvalente au plumage également jaune mais aux tarses noirs.

C'est également dans le cadre de cette coopération que le laboratoire possède des échantillons de la race Silkie, provenant d'animaux au plumage soyeux blanc ou noir. Cette race présente plusieurs caractéristiques morphologiques très appréciées en Asie et notamment la coloration noire de sa peau mais également de sa chair et de ses os..

### **Afrique**

Des populations de poulets villageois ont été échantillonnées dans quatre pays d'Afrique. Pour trois d'entre eux, le Benin, le Ghana et la Côte d'Ivoire, les animaux ont été échantillonnés dans le cadre du projet « *Caractérisation morphologique, zootechnique et génétique des populations locales de volailles de l'espèce Gallus gallus dans les pays côtiers de l'Afrique occidentale* » financé par le programme DURAS (Promotion du Développement Durable dans les systèmes de Recherche Agricole du Sud) et coordonné par Xavier Rognon dans notre UMR (INRA, AgroParisTech). Un protocole similaire a été utilisé au Cameroun dans le cadre de la thèse de Jean-Claude Fotsa (Fotsa, 2008).

Les animaux ont été échantillonnés dans deux zones climatiques par pays : savane et forêt pour les pays du projet DURAS, forêt et haut-plateaux pour le Cameroun. La zone forestière est caractérisée par une pluviométrie bimodale. Les conditions villageoises sont associées à une absence totale de vaccination, une alimentation irrégulière et une forte pression des prédateurs et des aléas climatiques. Le taux de mortalité déclaré par les éleveurs est souvent élevé, proche de 50%, les maladies infectieuses représentant 50% des causes de cette mortalité (Fotsa 2008).

### **Lignées expérimentales histocompatibles de l'INRA**

Trois lignées de White Leghorn, congéniques pour le CMH et maintenues à Nouzilly ont également été utilisées. Ces lignées sont congéniques respectivement pour les sérotypes B12, B13 et B19 (Miller *et al.*, 2004).

## **C. Autres races utilisées pour les études de diversité du CMH dans la thèse**

### **Races européennes**

En plus des races précédemment décrites, d'autres races locales européennes du projet AvianDiv ont été utilisées dans le cadre de l'exploration de la diversité génétique du CMH. Il s'agit pour le Nord de l'Europe des races Westfaeliche Totleger et Sundheimer d'Allemagne (6 et 7), de la race Owl-bearded (Uilenbaarder) des Pays-Bas (9), de la Finnish Landrace de Finlande (17), d'une population témoin de White Leghorn, la Old Scandinavian Reference Population (18), ainsi que des races Dorking et Cochin du Royaume-Uni (14, 15). En provenance du Sud de l'Europe, il s'agit de la race Siciliana Buttercup d'Italie (20) et de la race Black Castellana d'Espagne (22). Enfin, pour l'est de l'Europe, il s'agit de la race Oravka de Slovaquie (25), des races Ukrainian bearded et Poltava clay d'Ukraine (31 et 30) et de la race Godollo Nhx de Hongrie (33).

### **Races anciennes françaises**

Quinze races anciennes françaises ont été étudiées au niveau du CMH, trois d'entre-elles font partie du projet AvianDiv, il s'agit de la Bresse Noire (11), la Houdan (12) et la Marans (13). Les douze autres sont des races dont l'ADN est conservé au laboratoire dans le cadre d'un projet de caractérisation et valorisation des races anciennes de poules. Ces races ont été échantillonnées soit dans un élevage pépinière au Centre de Sélection de Béchanne (Barbezieux, Bourbonnais, Bresse Blanche, Gauloise Grise, Gauloise Noire, Géline de Touraine) soit chez des éleveurs amateurs dans différentes régions françaises (Coucou de Rennes, Crève-cœur en Normandie, Gasconne, Gauloise Dorée, Limousin et Noire de Challans).

### **Maroc**

L'ADN d'animaux provenant d'une population locale est également disponible au laboratoire grâce à une collaboration de l'UMR avec le professeur Benabdeljellil de l'Institut Hassan II. La diversité de cette population a été estimée à l'aide des marqueurs microsatellites du projet AvianDiv par Xavier Rognon.

#### **D. Espèces sauvages**

Plusieurs espèces de poules sauvages ont pu être étudiées. Il s'agit de 25 animaux d'espèce Coq Bankiva (*Red Junglefowl*, *Gallus gallus*), 5 animaux d'espèce Coq de Sonnerat (*Grey Junglefowl*, *Gallus sonneratii*) et 4 animaux d'espèce Coq de Lafayette (*Sri Lanka Junglefowl*, *Gallus lafayetii*). Parmi les 25 Coqs Bankiva, 19 proviennent du projet AvianDiv et sont des animaux conservés en Thaïlande avec 2 sous-espèces identifiées : *Gallus gallus spadiceus* (101) et *Gallus gallus gallus* (102). Les autres animaux proviennent du parc zoologique de Clères, d'où proviennent également les Coqs de Sonnerat et de Lafayette.

## II. METHODES GENERALES UTILISEES

Les méthodes utilisées plus spécifiquement pour certaines parties de la thèse sont décrites dans les chapitres correspondants.

### A. Méthodes de biologie moléculaire

#### Extraction

L'ADN est extrait à partir du sang. Le sang est prélevé à la veine alaire des oiseaux, directement à l'aide d'une seringue et une aiguille, puis transféré dans des tubes sous vide contenant un anticoagulant (EDTA). Il est ensuite centrifugé de façon à éliminer le plasma, qui est remplacé par une solution de PBS/glycérol (20:80) afin de préserver les cellules lors de la congélation. Peu de sang suffit pour extraire de l'ADN génomique car les hématies sont nucléées chez les Oiseaux.

La méthode rapide d'extraction utilisée pour obtenir de l'ADN consiste en une étape de lyse des cellules, puis d'une lyse alcaline au NaOH suivie d'une neutralisation de la solution avec du Tris-HCl. Concrètement, environ 2  $\mu$ L de sang sont mis en solution dans 250  $\mu$ L de solution A (1% de Tris 1M à pH 8, 0,1% de MgCl<sub>2</sub> 1M, 5% de NaCl 3M, 1% de détergent Igepal, H<sub>2</sub>O milliQ). Après une première centrifugation (10 mn à 3000 rpm), le culot, blanc rosé, est remis en solution dans 250  $\mu$ L de solution A et une deuxième centrifugation identique est réalisée. Les culots, blancs, sont alors incubés 2 h à 65°C avec 250  $\mu$ L de solution de NaOH (200 mM) et une agitation horizontale, avant ajout, une fois à température ambiante de 250  $\mu$ L de Tris-HCL (Tris 100 mM et HCl, 200 mM). La neutralité de la solution finale est contrôlée sur papier pH.

Cette méthode a également été utilisée pour un grand nombre de populations du panel AvianDiv et d'origine asiatique, cependant d'autres ADN ont été extraits par une méthode d'extraction plus lente, fragmentant moins l'ADN, décrite par Tixier-Boichard *et al.*, en 1994. Les échantillons des populations africaines ont été extraits au Department of Animal Science, University of Ghana, avec un kit d'extraction QIAGEN. Un aliquot a été envoyé en France pour le génotypage des marqueurs microsatellites utilisés dans le projet AvianDiv.

## **PCR**

Les PCRs, sauf précision contraire, ont été réalisées dans 50  $\mu$ L en utilisant le kit HotStarTaq Master Mix de chez Qiagen, avec 100 ng d'ADN, 0,1  $\mu$ M de chaque amorce, 25  $\mu$ L de Master Mix et l'eau fournie dans le kit qsp 50  $\mu$ L. Le programme standard pour le thermocycleur est le suivant : une Hot-Start de 15 mn à 95°C suivi de 30 cycles composés de 45 s à 94°C, 1 mn 30 s à 60°C et 1 mn à 72°C, suivis d'une élongation finale de 15 mn à 72°C.

## **Electrophorèse**

Les gels d'électrophorèse utilisés en condition standard ont été des gels d'agarose à 1 ou 2 %. De 7 à 20  $\mu$ L de BET ont été rajoutés au gel selon son volume (de 50 mL à 200 mL). 10  $\mu$ L d'un mélange composé de 8  $\mu$ L de produit PCR et 2  $\mu$ L de tampon de charge 15X sont alors déposés pour la migration, ainsi que le marqueur de taille 100 bp DNA Ladder de chez New England BioLabs (7  $\mu$ L en qualitatif, 10  $\mu$ L en quantification).

## **Purification**

Les produits PCR ont été purifiés directement sur colonnes par centrifugation avec le QIAquick PCR purification Kit de chez Qiagen ou après leur découpage à partir d'un gel d'agarose avec le QIAquick Gel extraction Kit. Les temps et la vitesse de centrifugation ont systématiquement été minimisés à 30 s et 9000 g sauf pour la dernière étape d'élution réalisée en 60 s à 10000 g, ceci afin de maximiser les quantités de produit récupéré. De la même façon, les colonnes ont toujours été laissées à reposer au moins 2 mn avec le tampon d'élution avant la dernière centrifugation.

## **Séquençage**

Les séquençages, sauf en ce qui concerne le projet réalisé au CNS (Centre National de Séquençage), ont été réalisés par Eurofins MWG Operon, selon leur protocole de séquençage de produits PCR purifiés. Cependant, autant que possible, les produits PCR ont été envoyés directement à séquençer, dans des volumes minimaux de 15  $\mu$ L, accompagnés des amorces correspondantes, à la concentration de 2 pmol/ $\mu$ L. Un aliquot de 15  $\mu$ L de chaque amorce a été fourni par échantillon lors du séquençage d'un petit nombre de tubes ou pour quatre échantillons lors du séquençage par plaque de 96 échantillons. Lors de l'envoi de plaque, les produits à séquençer ont systématiquement été séchés pendant la nuit à 50°C.

## **B. Méthodes de bioinformatique**

### **Conception des amorces**

Les amorces ont été choisies en utilisant principalement la séquence du CMH publiée en 2007 par Shiina *et al.*, 2007 (numéro d'accèsion GenBank AB268588) et le programme disponible en ligne Primer3 (Rozen & Shaletsky 2000) puis, à partir de sa mise à disposition, l'outil Primer-BLAST disponible sur le site du NCBI (<http://www.ncbi.nlm.nih.gov>). Un tableau récapitulant les amorces utilisées se trouve en Annexe 2.

### **Navigateurs de génome**

Le navigateur de génome de l'UCSC a été utilisé en priorité pour accéder au dernier assemblage du génome du poulet (mai 2006), ainsi que les fonctions ePCR pour la vérification des amorces et Blat pour la recherche de séquences dans le génome.

### **Analyse de séquences**

Les séquences d'acides nucléiques ont toujours été manipulées en première approche avec BioEdit (Hall 1999) et les séquences d'acides aminés avec MEGA 4 (Tamura *et al.*, 2007). Les outils de BLAST (Altschul *et al.*, 1990) proposés par le NCBI ont été utilisés de nombreuses fois pour de nombreuses tâches.

# RESULTATS

## I PREMIERE PARTIE : EXPLORATION DE LA DIVERSITE DU CMH VIA LE MARQUEUR LEI0258 DANS LES COLLECTIONS MAINTENUES AU LABORATOIRE ET INFERENCES SUR LE MODE D'EVOLUTION DE CE MARQUEUR ATYPIQUE

### A. Introduction

Les ressources génétiques concernant le poulet sont particulièrement diversifiées pour un animal d'élevage, en relation avec des foyers multiples de domestication et la présence, dans de nombreuses régions du monde d'un grand nombre de races (FAO, The State of the World's Animal Genetic Resources for Food and Agriculture), de populations sauvages, de populations villageoises, de races traditionnelles et races d'agrément, de lignées commerciales et expérimentales. Des études récentes de la diversité de poulet ont utilisé des marqueurs neutres comme des microsatellites (Rosenberg *et al.*, 2001, Berthouly *et al.*, 2008) ou des SNPs (Twito *et al.*, 2007) et des séquences mitochondriales (Liu *et al.*, 2006). Peu d'études ont été menées, en génétique des populations sur des locus fonctionnels ou des gènes, notamment des gènes associées à la réponse immunitaire (Downing *et al.*, 2009a et 2009b).

Le CMH du poulet est reconnu comme contenant les gènes les plus variables du génome. En effet, il comprend les gènes codant pour des protéines impliquées dans la présentation des peptides aux cellules T (principalement des molécules de classe I et de classe II). Le CMH du poulet est divisé en deux régions génétiquement indépendantes, B et Y, situées sur le microchromosome 16 du poulet (Delany *et al.*, 2009). Le CMH du poulet, notamment la région B, a été associé à des phénotypes de résistance, de susceptibilité et de réponse différentielle aux vaccins. Historiquement, la région B a été définie en utilisant des techniques de sérologie, conduisant à la description d'un certain nombre de sérologies B (ou encore sérotype), une vingtaine étant bien connus et répandus, principalement chez des animaux de race White Leghorn. Seules les études concernant des lignées expérimentales indiquent le sérotype B des animaux. D'un autre côté, de plus en plus d'études concernent des études de résistance aux maladies chez des poulets commerciaux ou locaux, et dans ce contexte, la sérologie est rarement connue. Depuis la fin des années 90, au moins un marqueur du CMH du poulet a été identifié, en 1999 par McConnell *et al.*, et décrit en 2006 par Fulton *et al.*, il s'agit du marqueur LEI0258 Ce marqueur a apporté un progrès important car il peut être relativement facilement et économiquement génotypé en première approche, c'est-à-dire en



termes de tailles approximatives d'allèles, par électrophorèse sur gel. En outre, ce marqueur présente une corrélation étonnamment forte avec le type B en sérologie dans les différentes lignées expérimentales et commerciales testées comme l'ont prouvé Fulton *et al.*, Cependant les correspondances ne sont pas parfaites, un allèle LEI0258 pouvant correspondre à plusieurs sérotypes, ou plusieurs allèles à un sérotype mais dans ce cas les sérotypes ont pu être être reclassifiés en sous-types, par exemple B15, B15.1, B15.2.

Ce marqueur est un marqueur atypique de type « nombre variable de répétitions en tandem » (VNTR), très polymorphe, avec 26 tailles d'allèles rapportées (Fulton *et al.*, 2006). Il est caractérisé par la répétition de deux courtes séquences conservées de 12 et 13 pb, et par un nombre important de polymorphismes de séquence dans les régions flanquantes, des indels et des SNPs. Ces polymorphismes et le nombre et type de répétitions définissent la taille des allèles. Le marqueur LEI0258 est localisé dans la zone centrale de la région B du CMH, dans un intron du gène *B-BTN1*, à proximité des gènes *B-BTN2* et *B-G1*. Le gène *B-G1* a été récemment identifié comme gène candidat pour expliquer l'association notable entre des sérologies B et la maladie de Marek chez le poulet (Goto *et al.*, 2009).

Ce travail de recherche a consisté en l'exploration de la diversité génétique du marqueur atypique LEI0258 chez plus de 1200 animaux provenant de 80 populations différentes de poulet en provenance d'Afrique, d'Asie et d'Europe, y compris les espèces de Galliformes sauvages, en utilisant pour le génotypage les techniques d'électrophorèse sur gel et de séquençage classique. Afin de pleinement caractériser les allèles génotypés, nous avons séquencé plus de 500 allèles, aboutissant à l'identification de 77 allèles différents. Ces différents allèles, ainsi que 5 allèles précédemment décrits, ont été comparés, conduisant à des résultats sur la ressemblance entre allèles mais également sur l'évolution du marqueur LEI0258.

**Cet article doit être soumis à Immunogenetics.**

## B. Article de recherche

### Diversity and evolution of the atypical marker LEI0258 located within the chicken *Mhc*

Olympe Chazara, Chi-Sheng Chang, Nicolas Bruneau, Khalid Benabdeljelil, Jean-Claude Fotsa, Boniface B. Kayang, Valentine Yapi-Gnaore, Issaka A.K. Youssao, Chih-Feng Chen, Michèle Tixier-Boichard, Bertrand Bed'Hom

O. Chazara, N. Bruneau, M. Tixier-Boichard, B. Bed'Hom  
INRA, UMR 1313, Génétique Animale et Biologie Intégrative,  
Domaine de Vilvert, 78352 Jouy en Josas, France  
e-mail: [olympе.chazara@jouy.inra.fr](mailto:olympе.chazara@jouy.inra.fr)  
Tel.: +33-1-34652720  
Fax.: +33-1-34652933

C-S. Chang, C-F. Chen  
Department of Animal Science,  
National Chung-Hsing University, Taichung, Taiwan

K. Benabdeljelil  
IAV Hassan II DPBA BP 6202  
Rabat Institute, 10101 Rabat, Morocco

B. B. Kayang  
Department of Animal Science  
College of Agriculture & Consumer Sciences  
University of Ghana, Legon, Ghana

I. A. K. Youssao  
Université d'Abomey-Calavi  
Ecole Polytechnique d'Abomey-Calavi  
Département de Production et Santé Animales  
01 BP 2009, Cotonou, Benin

V. Yapi-Gnaore  
Centre Nationale de Recherche Agronomique  
01 BP 1740, Abidjan 01, Côte d'Ivoire

J-C. Fotsa  
Institut de Recherche Agricole pour le Développement  
BP 30189, Yaoundé, Cameroon

## **Abstract**

The chicken major histocompatibility complex (*Mhc*) is located on the microchromosome 16 and is recognised as the most variable region in the genome. The primary aim of the study was to provide a global picture of the diversity of the chicken *Mhc B* region and to use the different alleles identified to assess relationship between *Mhc* types. In order to do so, we genotyped more than 1600 animals from 80 different populations or lines of chickens from Africa, Asia and Europe, including wild fowl species, using the marker LEI0258, located within the core area of the *B* region of the chicken *Mhc*. In order to fully characterize the alleles genotyped, we sequenced more than 500 alleles, defining 77 different alleles, and we studied their relationship, including alleles previously reported, resulting in the definition of relationships between 81 alleles. Such a definition is all the more useful as the marker LEI0258 exhibits a very strong correlation with chicken *Mhc B* serotypes associated with resistance, susceptibility, and differential response to vaccine. The comparison of the alleles also provided indications of this atypical and highly polymorphic marker evolutionary dynamics. This study constitutes the largest variation report of the chicken *Mhc* and provides a framework for future diversity or association studies involving the *Mhc* in chickens.

**Keywords** Major histocompatibility complex ; LEI0258 ; *Gallus gallus* ; diversity ; evolution

## Introduction

Chicken genetic resources comprise a wide range of breeds and populations including wild fowl species, fancy breeds, local breeds, commercial stocks and experimental lines. Recent studies of chicken diversity have used neutral markers such as microsatellites (Rosenberg *et al.* 2001, Berthouly *et al.* 2008) or SNPs (Twito *et al.* 2007), and mitochondrial sequences (Liu *et al.* 2006). But only few population genetics studies have been conducted on functional loci or genes, especially disease-associated genes (Downing *et al.* 2009).

The major histocompatibility complex (MHC) is a genomic region comprising genes coding for proteins involved in peptide presentation to T cells (mainly Class-I and Class-II molecules). The chicken MHC is divided in two genetically independent regions, B and Y, located on the microchromosome 16 of the chicken (Delany *et al.* 2009). The chicken MHC, especially the B region, has been associated with resistance, susceptibility, and differential response to vaccine (Pinard-van der Laan *et al.* 1998, Lee *et al.* 2004), and historically the B region has been defined using serology, leading to the description of a number of B serotypes, about twenty being well known, mainly in White Leghorns. Concerning the MHC, a number of papers have been published on specific MHC genes and specific chicken origins. Studies involving experimental lines of chickens use the serology to define B serotypes of the animals (Rogers and Kaufman, 2008), whereas those concerning other origins of chicken do not use serology and provide only some comparative indications of the MHC genotype (O'Neill *et al.* 2009). Thus, comparisons between studies are difficult, but could be facilitated by the use of at least one reference marker for the chicken MHC, identified in 1999 by McConnell *et al.* and described in 2006 by Fulton *et al.*: the LEI0258 marker, which can be easily genotyped, in terms of allele size, by gel electrophoresis, as shown in the latest article. Furthermore, this marker as been shown to correlate surprisingly well with B serotypes in the different lines tested (Fulton *et al.* 2006).

The chicken marker LEI0258 is an atypical variable number of tandem repeat (VNTR), highly polymorphic, with 26 allele sizes already reported (Fulton *et al.* 2006). It is characterized by the repetition of two conserved short sequences, 12 and 13 bp long respectively, and an important number of sequence polymorphisms in the flanking regions, including indels and SNPs, both types of polymorphisms determine the allele size. This marker is localised in the core area of the B region of the MHC, close to the gene BG1. This gene has just been reported as involved in the exceptional strong association between B types and Marek's disease in chicken (Goto *et al.* 2009).

One of the aims of this study is to identify the maximum number of alleles of the LEI0258 marker by examining its genetic diversity in 80 different populations or lines of chicken from Africa, Asia and Europe, using gel electrophoresis and sequencing for genotyping. Then, for the alleles sequenced, we compare the characteristics of these alleles, in order to infer the relationship between alleles and the dynamic evolution of the LEI0258 marker.

## **Materials and methods**

### **Genomic DNA origin**

Animals were sampled from 80 different populations or lines of chickens (*Gallus gallus*). This included 51 populations of various origins and selection history, sampled in the frame of the European AVIANDIV project (Hillel *et al.* 2003), twelve traditional French breeds (Berthouly *et al.* 2008), nine different populations of Asian chickens from National Chung-Hsing University (Taichung, Taiwan), four Immune Response lines and three lines fixed for a B serotype from INRA (Nouzilly, France, Minozzi *et al.* 2008, Miller *et al.* 2004), samples of village chickens from four African countries, three wild *Gallus sp.* kept in a zoological park (Clères, France), and samples of Moroccan chickens. Most of these populations had previously been included in at least one diversity study using microsatellite markers (Table 1). DNA was already available from the Aviandiv DNA bank kept at INRA, Jouy-en-Josas. Additional extractions were performed for the African chickens at the Department of Animal Science of Legon University in Ghana, using the DNeasy Blood & Tissue Kit (Qiagen). Additional extractions were performed for Asian breeds and French local breeds at INRA using a rapid extraction procedure described in Chang *et al.* 2006.

The *Mhc* haplotype information was only available, from serology typing, for the seven INRA experimental lines, reported as B12, B124, B13, B15, B19, B21, B34, for the B4 and B12 C line, and the B15 Sarcoma-resistant or susceptible lines, from the AVIANDIV project.

### **Genotyping: PCR and electrophoresis**

In most cases, 10 to 50 animals were randomly picked from the different populations for genotyping with the marker LEI0258, except for INRA experimental lines and for the zoo animals (Table 1).

Amplifications were done with PCR primers LEI0258-F: CACGCAGCAG-AACTTGGTAAGG forward and LEI0258-R: AGCTGTGCTCAGTCCTCAGTGC reverse (McConnell *et al.* 1999), in 10 µL with 20 ng of DNA, 1X of QIAGEN Multiplex PCR Kit and primer concentrations of 0.1 µM and the following PCR cycle setup: 94°C for 15 min, then 35 cycles of 94°C for 45 s, 60°C for 1 min 30 s, 72°C for 1 min, with a final extension of 72°C for 15 min. The total volume of PCR product was then loaded on 4% agarose gel containing ethidium bromide (BET). The gel was exposed to UV light to reveal the amplified fragments and their size, by comparison with the 100 bp DNA Ladder (New England BioLabs).

### **Genotyping: sequencing**

In order to get the sequence for all the alleles of the LEI0258 marker, each allele identified from a population or an origin was reamplified for sequencing in at least two individuals. Homozygous carriers were selected in priority but heterozygous carriers had to be used for rare alleles. Approximately 410 animals were used for sequencing, half were homozygotes, half were heterozygotes, those were often sequenced for both alleles. Amplifications were then done in 50  $\mu$ L with 40 ng of DNA with the same primers, in the conditions described above. It was sent directly to sequencing after PCR amplification for homozygotes or after separation of the two alleles on a 4% agarose gel, bands cutting and purification with QIAquick PCR Purification Kit from Qiagen. Sequencing was carried out in the forward direction, by Eurofins MWG Operon, using their standard protocol for purified PCR products.

### **Sequence and data analysis**

Sequence quality and variation analyses were carried out with BioEdit (Hall 1999). In order to estimate the relationship between alleles, we considered all the sequenced alleles. Polymorphisms in the flanking sequences, SNPs and Indels, were taken into account as binary information, with the presence of the rare allele for SNPs or the deletion been encoded 1, their absence, 0. Then, median-joining haplotype networks were constructed using Splitstree version 4.10 (Huson and Bryant 2006, Bandelt *et al.* 1999). In VNTR evolution, tandem expansion and contraction are considered to be much more frequent than point mutations. Therefore we did not use the repeat number characters for the network construction. Furthermore, for this marker, the difference of length between alleles was a poor estimate of allele divergence, as two very different alleles could have similar lengths. The information about the number of repeats, along with the size of the allele, was reported on Figure 2 for analysis purposes.

### **Population analysis**

All animals were genotyped on gels but not all of them were sequenced. For non-sequenced alleles, sizes on gels were extrapolated from the sequenced alleles in the same population. In the rare cases where the corresponding allele was not sequenced, the most plausible size is given. For each population or origin, allele frequencies, number of alleles, number of allele combinations, expected and observed heterozygosity values were calculated using GENETIX 4.4 (Belkhir *et al.* 2000).

## Results

### Polymorphism of the marker LEI0258

As described by Fulton *et al.* in 2006, the basis of the LEI0258 marker polymorphism are the number of repetitions of the two repeats, named R13 and R12 after the number of base repeated (R13: ATGTCTTCTTTCT, R12: TTCCTTCTTTCT), plus the sequence of the flanking regions, 62 bp upstream and 51 bp downstream (excluding primers). Whereas those authors identified nine polymorphisms in the flanking sequences, two deletions and two SNPs upstream and two deletions and three SNPs downstream, we could not find one of the upstream deletion (the A insertion at position -61 bp) but we found two additional SNPs, one upstream, at position +1 bp and one downstream, at +15 bp, numbered in accordance with Fulton *et al.* 2006. In addition, we propose to redefine the biggest deletion, found in the downstream sequence, as a deletion concerning 8 bp instead of 7 bp: ATTTTGAG. This has been discussed and we agreed on that point with the authors of the 2006 publication and therefore the deletion now concerns positions +23 to +30.

### Global allelic variability of the marker LEI0258

One of the aims of the study was to evaluate the variability of the LEI0258 marker in order to gain information on the variability of the MHC in the chicken. To achieve this, 534 alleles were sequenced, corresponding to 80 different origins of chicken. Out of this high number of sequenced alleles, 77 different alleles were identified, by the number of repetitions, the total size, and the combination of the indels and SNPs identified in the flanking regions. Among those alleles, we identified 52 new alleles, by comparison with Fulton *et al.* 2006, but we did not find 4 alleles they described. Thus, the number of alleles known for the LEI0258 marker is now 81 (Table 2), although this number is probably lower than the real number of existing alleles.

Then, we determined the number of different allele sizes in our samples: we found 51 different allele sizes, from 181 bp to 552 bp, this is almost included within the range already described (182 bp to 522 bp), therefore, the 8 new allele sizes we describe are mostly intermediate allele sizes.

In addition, as the size of the alleles reflects only the combination of repeats, and as those are probably of greater evolutionary meaning, the number of each repeat should be fully taken into consideration. With this study, we were able to describe wider ranges for the number of repeat, 14 states for the R13 repeat : 1, 2, 10, 15, 16, 17, 20, 22, 23, 24, 25, 26, 27, 28, and 23 states for the R12 repeat : 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,



23, 24 (Fig. 1). We found some of the “missing” combinations predicted by Fulton *et al.*: 1-9 and 1-10, as we identified 16 new combinations. But in comparison to the total number of mathematically possible combinations (322), we could only identify 37 combinations, including 21 of the 22 already described. Furthermore, 23 out of the 38 known combinations for the LEI0258 marker are composed of only 1 repeat for R13

### **Population analysis**

Another aim of this study was to evaluate variability at the Mhc level of 80 various populations, breeds or lines of chicken. The highest value for observed heterozygosity was obtained for the Ukrainian bearded breed, a traditional breed from East of Europe and also for a smaller sample of Red Junglefowl from the zoological park ( $H_o = 1$ ). Low values were obtained as expected for experimental lines fixed for *Mhc B*-types ( $H_o = 0$ ), for five commercial lines which could also be under selection for *Mhc B*-types ( $H_o = 0$  to  $H_o = 0.27$ ) but also for one experimental control line (P53,  $H_o = 0.29$ ) and one traditional French breed (Gauloise dorée,  $H_o = 0.18$ ). Besides those extreme values, observed heterozygosity ranged from 0.4 (the Ju-Chi breed from Taiwan) to 0.96 (the Bedouin breed from Middle-East) (Table 1).

### **Network analysis and allele relationships**

The network constructed from the 82 alleles discriminated 13 states, each state clustering one to 18 alleles (Fig. 2). The two single alleles are from a Grey Junglefowl and a Taiwanese local breed. Their removal from the analysis did not change the network. The rest of the clusters were supported by eight to 183 obtained sequences, including observations from Fulton *et al.* 2006. Whereas the number of repeats and the size of the alleles was not taken into consideration for the network construction, a certain pattern can be observed from the clustering obtained, with the small size alleles (R13 = 1, R12 = 2 to 7), smaller or equal to 241 bp, grouped on one side and the bigger alleles, bigger or equal to 237 bp, or with a high number of R13 repeats (R13 = 1, 10 to 28), grouped in one large cluster (Supplementary Fig.1).

## **Discussion**

### **Population**

For 73 out of 80 origins of chicken, we were able to find evaluation of observed heterozygosity from neutral microsatellite markers (Berthouly *et al.* 2008, Berthouly *et al.* 2009, Fotsa 2008, Granevitz *et al.* 2007, Hillel *et al.* 2003, Loukou *et al.* 2009, Loywyck *et al.* 2008, Osei-Amposah *et al.* 2009, Rosenberg *et al.* 2001, Youssao *et al.* 2009). In 58 cases out of 73 (79.5%), observed heterozygosity was higher with the LEI0258 marker than with the average of neutral marker. For the 53 traditional breeds or wild species only, this percentage was even higher: 92.5%. These results point out that the heterozygosity found in commercial egg production lines is generally very low, and in one case null, compared to other commercial lines or unselected chicken populations. Similar results for commercial stocks have also been reported with neutral markers by Muir *et al.* in 2008. But the very low level of heterozygosity observed in our study for some commercial lines and the comparison with the heterozygosity observed with neutral markers in the same lines, point toward a strong selection done on the *Mhc*, since any effect of population size, bottleneck or founder effect, should affect also the diversity estimates obtained with microsatellite markers. The benefit of this type of strong selection, on a region implicated in disease resistance and characterized by naturally balanced selection, should probably be evaluated carefully. If it is probably true that the nucleus populations of the different selection schemes have lost the ability to respond to a variety of pathogens, but the terminal crosses can still combine several alleles if the different grand-parental lines used in the cross are selected on different *Mhc* alleles.

### **LEI0258 and serology**

Serology information was only available for the following eight B types: B4, B12 from two origins, B124 (INRA I.R. lines, Pinard-van der Laan 2002), B13, B15 from two origins, B19, B21, and B34. All those B types originated from White Leghorns. For the B4, B12, B13, B15, B19 and B21, a corresponding LEI0258 allele had already been reported by Fulton *et al.* 2006 and our finding perfectly matched their description. The LEI0258 allele corresponding to B124 is identical to the allele corresponding to B12, a 487 bp long allele. B124 is not common and has been defined in the INRA I.R. lines, as serologically identical to the B124 described in ISA Warren experimental lines, selected for high and low antibody response to sheep red blood cells, and in 1993, this B124 was defined as distinct from B12 (Pinard and Hepkema, 1993). B124 from the the INRA I.R. lines and references B12 have never been compared with serology, but could provide valuable information. B34 has been described in

the same lines and is now associated with the same LEI0258 allele than B62, an allele 420 bp long. B34 has never been compared to B62 which is also found in White Leghorn but in a commercial line (Fulton *et al.* 2006).

### **LEI0258 evolution**

The evolution of the number of repeats of the LEI0258 marker is probably primarily based on a step-wise model, one step being a duplication or a loss of one repeat. This has been observed in the I.R. lines with the observation of the emergence of a new allele size, 552 bp, absent in the founder animals, explained by the duplication of a R13 repeat from the 539 bp allele, common in these lines. But the combination of the different numbers of repeats seems far from being random, as we could only identify 37 combinations, including 21 of the 22 already described, furthermore 23 out of the 38 combinations of repeats known for the LEI0258 marker are composed of only 1 repeat for R13. The model of evolution leading to this limited number of combinations is not obvious and could be explained by a limitation of the total allele size. This can make sense in the Mhc of the chicken, which is characterized by a compact organization, with small intergenic sequences (Kaufman *et al.* 1999).

Whatever the mechanism involved in the evolution of the number of repeats, the repeat region is associated with a number of SNPs and indels in the flanking region. The rate of polymorphisms for those regions is high, with 4 polymorphisms, including a 2 bp deletion, in the upstream 62 bp and 6 polymorphisms, including an 8 bp deletion, in the downstream 51 bp. This high number of polymorphisms could be somehow related to the mechanisms involved in the evolution of the repeat numbers. In this study, we chose to analyse the relationship between the different alleles by looking primarily at the flanking polymorphisms. This was mainly because point mutations such as SNPs and particular indels are considered to show slower rates of evolution than repeat numbers. Thus, the issue of homoplasy is more relevant for number of repeats than for point mutations, indeed a number of LEI0258 alleles have the same size and combination of repeats but different flanking sequences. In the network analysis, however, it appears that some of those mutations must have happened at least twice in order to explain the observed alleles. It is hard to tell which ones are the recurrent ones as this depends of the phylogenetic method chosen (data not shown). This is why we chose a network analysis representing all the possible events. Thus, this must be taken as a representation of the relationships between alleles, not a phylogeny. The choice made to analyse only the point mutations seems relevant as different types of combinations of the R13 and R12 repeats seemed to be associated with different parts of the network. If one

considers that the 8 bp and the 2 bp deletion have more significance than the other point mutations, then the network can be considered as divided into three parts: three clusters without deletion (at the top of figure 2), five clusters with the 2 bp deletion (bottom right), 4 clusters with the 8 bp deletion (bottom left), and a single allele cluster having those two deletions.

It should be noted that alleles found in wild species of Junglefowls, namely Grey Junglefowl, Red Junglefowl, and Sri-Lanka Junglefowl, can be found in seven of the 13 clusters defined by the point mutations, and in all the 4 groups defined above (Fig. 2). This means that an important part of the diversity observed in domestic chickens with the LEI0258, which has been linked to the *Mhc* diversity, can be found in the wild ancestors of those chickens. Furthermore, it has been shown that several domestication events took place in Asia, mainly from the Red Junglefowl (International Chicken Genome Sequencing Consortium, 2004) but also that the Grey Junglefowl contributed also to the domestication of chickens (Eriksson *et al.*, 2008). Thus, the ancestral *Mhc* diversity is explained and was maintained in domestic chickens, almost certainly due to the peculiar characteristics of the *Mhc* region.

As shown, the genotyping of the LEI0258 marker can be done by gel electrophoresis, in any kind of chicken populations and even in wild species of fowls. These data provide a suitable framework for studying *Mhc* genes but also disease resistance or susceptibility in any population of chickens. We recommend this typing to be done in order to allow comparison between results obtained in different populations, including experimental chickens with *Mhc* defined in serology as suggested in Fulton *et al.* 2006, as comparisons between serology reagents or stocks have not been done recently between the reference laboratories. We also encourage researchers to sequence the different alleles found in their chicken populations or lines at the beginning, since we have demonstrated that a great part of the information provided by the LEI0258 marker is based on the combination of repeats and polymorphisms in the flanking sequences. Those molecular techniques are also useful for the study of heterozygotes, this is relevant for the *Mhc* region where we have shown that the heterozygosity is particularly high. The present study and the study by Fulton *et al.* (2006) provide a valuable reference, as a wide reference set of LEI0258 alleles, which can be used to identify and classify MHC genotypes in yet undescribed populations, and particularly local populations.

## Acknowledgements

We would like to thank Anne-Marie Chaussé (INRA, Nouzilly, France), Yannick Roman (Zoo de Clères, Clères, France), and Cheng Yu-Shin (Livestock Research Institute, Sinhua, Taiwan) for bird samples, Xavier Rognon (INRA, Jouy en Josas, France) for his help getting African samples, and Marie-Hélène Pinard-van der Laan (INRA, Jouy en Josas, France) for helpful discussions and for providing access to the I.R. lines.

## References

- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37-48
- Belkhir K, Borsa P, Goudet J, Chikhi L, Raufasate N, Bonhomme F (2000) GENETIX, logiciel sous Windows TM pour la génétique des populations In *Laboratoire Génome, Populations, Interactions: CNRS UMR5117 Université Montpellier II, Montpellier, France*
- Berthouly C, Bed'Hom B, Tixier-Boichard M, Chen CF, Lee YP, Laloë D, Legros H, Verrier E, Rognon X (2008) Using molecular markers and multivariate methods to study the genetic diversity on local European and Asian chickens breeds. *Anim Genet* 39:121-129
- Berthouly C, Leroy G, Van TN, Thanh HH, Bed'Hom B, Nguyen BT, Vu CC, Monicat F, Tixier-Boichard M, Verrier E, Maillard JC, Rognon X (2009) Genetic analysis of local Vietnamese chickens provides evidence of gene flow from wild to domestic populations. *BMC Genet* 10:1
- Chang CM, Coville JL, Coquerelle G, Gourichon D, Oulmouden A, Tixier-Boichard M (2006) Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genomics* 7:19
- Delany ME, Robinson CM, Goto RM, Miller MM (2009) Architecture and organization of chicken microchromosome 16: order of the NOR, MHC-Y, and MHC-B subregions. *J Hered* 2009 100:507-514
- Downing T, Lynn DJ, Connell S, Lloyd AT, Bhuiyan AK, Silva P, Naqvi AN, Sanfo R, Sow RS, Podisi B, O'Farrelly C, Hanotte O, Bradley DG (2009) Contrasting evolution of diversity at two disease-associated chicken genes. *Immunogenetics* 61:303-314
- Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Strömstedt L, Wright D, Jungerius A, Vereijken A, Randi E, Jensen P, Andersson L (2008) Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet* 4:e1000010
- Fotsa J-C (2008) Caractérisation des populations de poules locales (*Gallus gallus*) au Cameroun. Thesis, AgroParistech, France
- Fulton JE, Juul-Madsen HR, Ashwell CM, McCarron AM, Arthur JA, O'Sullivan NP, Taylor RL Jr (2006) Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* 58:407-421

Goto RM, Wang Y, Taylor RL Jr, Wakenell PS, Hosomichi K, Shiina T, Blackmore CS, Briles WE, Miller MM (2009) BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proc Natl Acad Sci U S A* 106:16740-16745

Granevitze Z, Hillel J, Chen GH, Cuc NTK, Feldam M, Eding H, Weigend S (2007) Genetic diversity within chicken populations from different continents and management histories. *Anim Genet* 38:576-583

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41:95-98

Hillel J, Groenen MA, Tixier-Boichard M, Korol AB, David L, Kirzhner VM, Burke T, Barre-Dirie A, Crooijmans RP, Elo K, Feldman MW, Freidlin PJ, Mäki-Tanila A, Oortwijn M, Thomson P, Vignal A, Wimmers K, Weigend S (2003) Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genet Sel Evol* 35:533-557

Huson DH, Bryant D (2006) Application of phylogenetics networks in evolutionary studies. *Mol Bio Evo* 23:254-267

International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695-716

Kaufman J, Milne S, Göbel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S (1999) The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 28:923-925

Lee LF, Bacon LD, Yoshida S, Yanagida N, Zhang HM, Witter RL (2004) The efficacy of recombinant fowlpox vaccine protection against Marek's disease: its dependence on chicken line and B haplotype. *Avian Dis* 48:129-137

Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, Baig M, Beja-Pereira A, Ding ZL, Palanichamy MG, Zhang YP (2006) Multiple maternal origins of chickens: out of the Asian jungles. *Mol Phylogenet Evol* 38:12-19

Loukou NE, Yapi-Gnaoré CV, Touré G, Coulibaly Y, Rognon X, Kayang B, Youssao I, Tixier-Boichard M, N'guetta ASP (2009) Assessing the diversity of indigenous chicken from two agro-ecological zones of Côte d'Ivoire using microsatellite markers. *J Anim Plant Sci* 5:425 - 436

Loywyck V, Bed'Hom B, Pinard-van der Laan M-H, Pitel F, Verrier E, Bijma P (2008) Evolution of polymorphisms at molecular markers in QTL and non-QTL regions in selected lines of chicken. *Genet Sel Evol* 40:639-661

McConnell SK, Dawson DA, Wardle A, Burke T (1999) The isolation and mapping of 19 tetranucleotide microsatellite markers in the chicken. *Anim Genet* 30:183-189

Miller MM, Bacon LD, Hala K, Hunt HD, Ewald SJ, Kaufman J, Zoorob R, Briles WE (2004) 2004 Nomenclature for the chicken major histocompatibility (B and Y) complex. *Immunogenetics* 56:261-279

Minozzi G, Parmentier HK, Mignon-Grasteau S, Nieuwland MG, Bed'hom B, Gourichon D, Minvielle F and Pinard-van der Laan MH (2008) Correlated effects of selection for immunity in White Leghorn chicken lines on natural antibodies and specific antibody responses to KLH and *M. butyricum*. *BMC Genetics* 9:5

Muir WM, Wong GK, Zhang Y, Wang J, Groenen MA, Crooijmans RP, Megens HJ, Zhang H, Okimoto R, Vereijken A, Jungerius A, Albers GA, Lawley CT, Delany ME, MacEachern S, Cheng HH (2008) Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci U S A* 105:17312-17317

O'Neill AM, Livant EJ, Ewald SJ (2009) The chicken BF1 (classical MHC class I) gene shows evidence of selection for diversity in expression and in promoter and signal peptide regions. *Immunogenetics* 61:289-302

Osei-Amponsah R, Kayang BB, Naazie A, Osei YD, Tixier-Boichard M, Rognon X (2009) Genetic diversity of Forest and Savannah chicken populations of Ghana as estimated by microsatellite markers *Anim Sci J* in press

Pinard-van der Laan MH (2002) Immune modulation: the genetic approach. *Vet Immunol Immunopathol* 87:199-205

Pinard MH, Hepkema BG (1993) Biochemical and serological identification of major histocompatibility complex antigens in outbred chickens. *Vet Immunol Immunopathol* 39:407-419

Pinard-van der Laan M-H, Siegel PB, Lamont SJ (1998) Lessons from selection experiments on immune response in the chicken. *Poult Avian Biol Rev* 9:125-141

Rogers SL, Kaufman J (2008) High allelic polymorphism, moderate sequence diversity and diversifying selection for B-NK but not B-1c, the pair of lectin-like receptor genes in the chicken MHC. *Immunogenetics* 2008 60:461-475

Rosenberg NA, Burke T, Elo K, Feldman MW, Freidlin PJ, Groenen MA, Hillel J, Mäki-Tanila A, Tixier-Boichard M, Vignal A, Wimmers K, Weigend S (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* 159:699-713

Twito T, Weigend S, Blum S, Granevitze Z, Feldman MW, Perl-Treves R, Lavi U, Hillel J (2007) Biodiversity of 20 chicken breeds assessed by SNPs located in gene regions. *Cytogenet Genome Res* 117:319-326

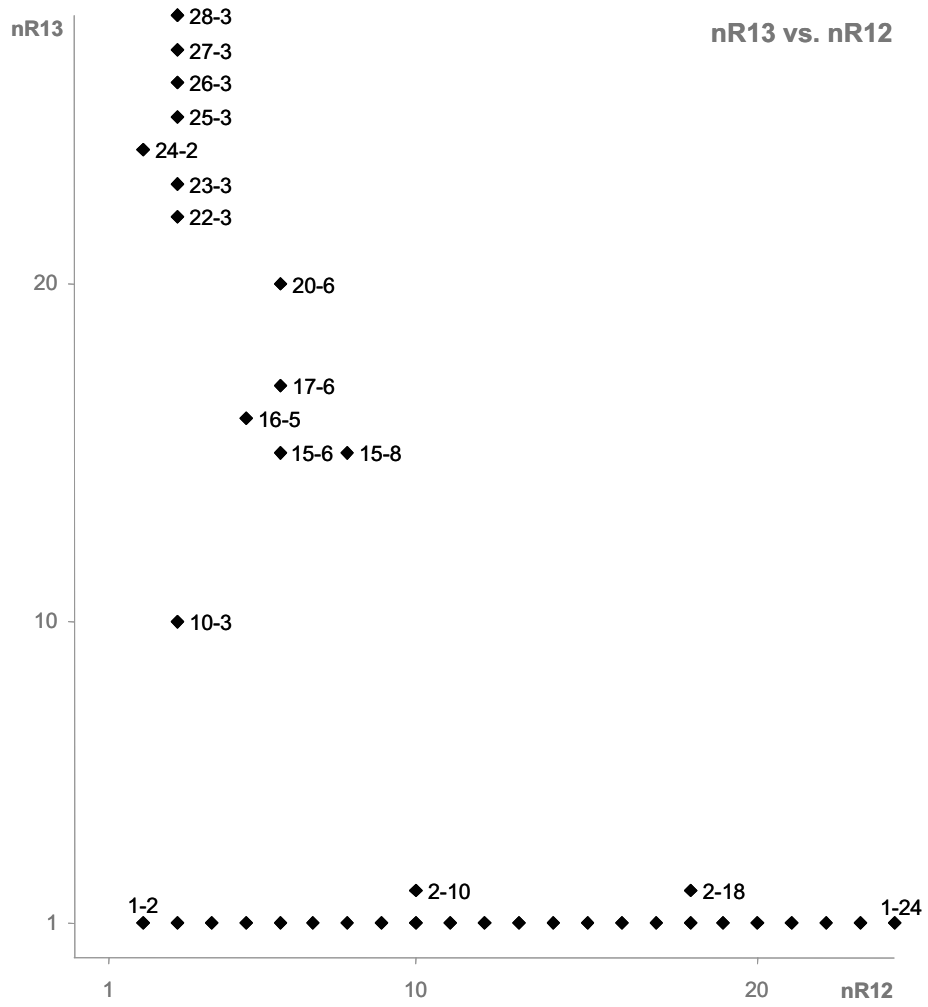
Youssao IAK, Dahouda M, Kpodékon MT, Idrissou N-D, Bonou GA, Tougan UP, Ahounou S, Yapi-Gnaoré V, Kayang B, Rognon X, Tixier-Boichard M (2009) Phenotypic characterisation and molecular polymorphism of indigenous poultry populations of the species *Gallus gallus* of North and South ecotypes of Benin. *African J Biotech* in press

## Figures legends

**Fig. 1** The different combinations of repeats number observed in the 81 alleles reported for the LEI0258 marker. nR13, number of 13 bp repeats; nR12, number of 12bp repeats

**Fig. 2** Median-joining network of the 82 alleles defined for the LEI0258 marker. The transformations implicating the ten binary characters are indicated by the corresponding number of the character as reported Table 2.





**Figure 1**

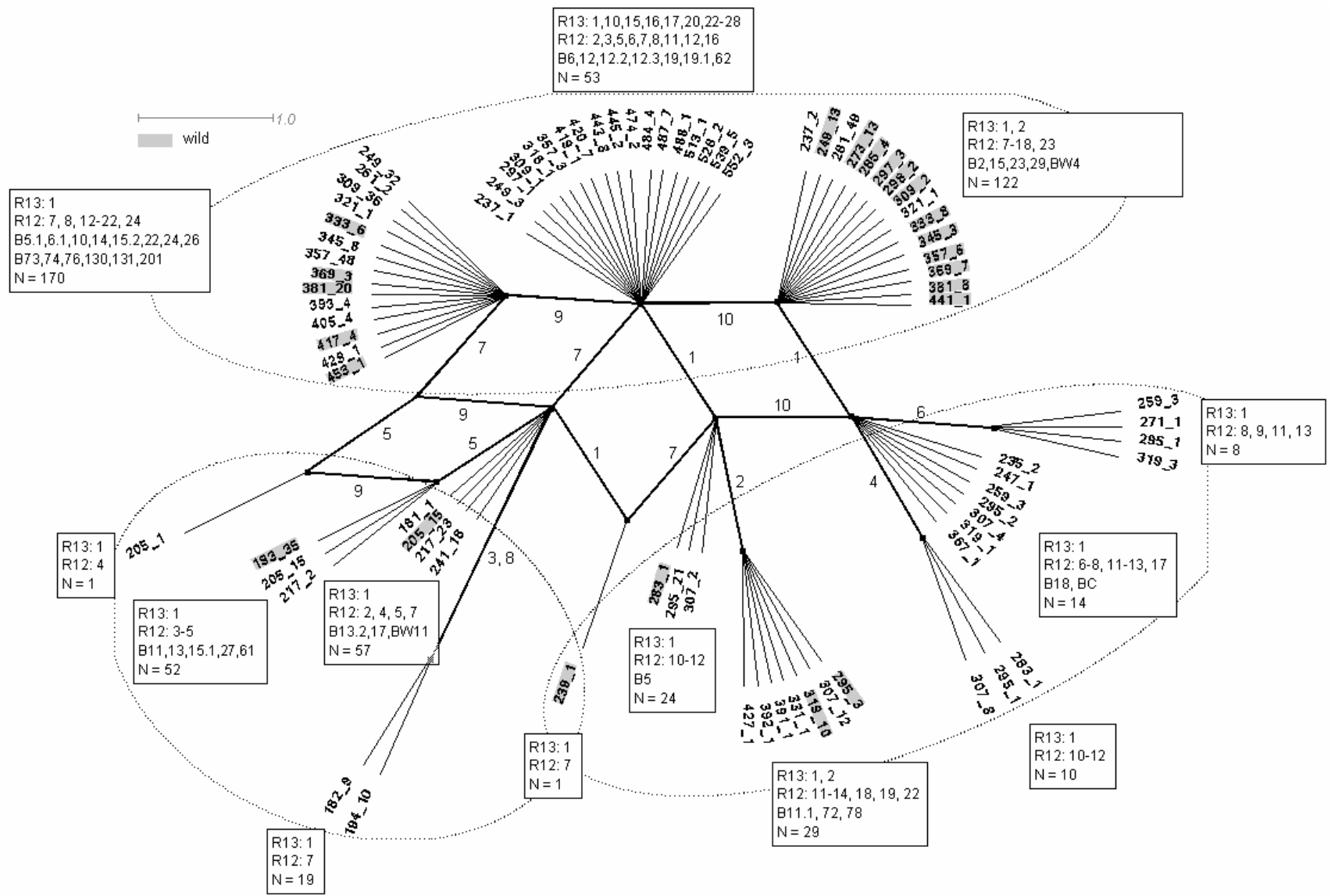


Figure 2

**Table 1** Description and diversity measures of the studied populations or lines of chicken

| Code  | Name                    | N  | N <sub>all</sub> | N <sub>geno</sub> | H <sub>exp</sub> | H <sub>obs</sub> | F <sub>is</sub> | H <sub>obs</sub><br>ref 1 | H <sub>obs</sub><br>ref 2 | H <sub>obs</sub><br>ref 3 | H <sub>obs</sub><br>ref 4 |
|-------|-------------------------|----|------------------|-------------------|------------------|------------------|-----------------|---------------------------|---------------------------|---------------------------|---------------------------|
| BAZ   | Barbezieux              | 18 | 9                | 13                | 0.86             | 0.94             | -0.10           |                           |                           |                           | 0.526                     |
| BBL   | Bresse blanche          | 24 | 5                | 12                | 0.82             | 0.88             | -0.07           |                           |                           |                           | 0.555                     |
| BEN   | Benin                   | 26 | 18               | 23                | 0.94             | 0.96             | -0.02           |                           |                           |                           | 0.568                     |
| BNA   | Bourbonnaise            | 18 | 6                | 7                 | 0.65             | 0.67             | -0.02           |                           |                           |                           | 0.487                     |
| CAM   | Cameroon                | 23 | 15               | 19                | 0.92             | 0.91             | 0.01            |                           |                           |                           | 0.626                     |
| CI    | Cote d'Ivoire           | 24 | 9                | 17                | 0.86             | 0.71             | 0.18            |                           |                           |                           | 0.587                     |
| COU   | Coucou rennes           | 24 | 10               | 16                | 0.87             | 0.63             | 0.28            |                           |                           |                           | 0.583                     |
| CRC   | Creve-coeur             | 24 | 6                | 9                 | 0.61             | 0.63             | -0.02           |                           |                           |                           | 0.528                     |
| GAS   | Gasconne                | 24 | 9                | 13                | 0.83             | 0.88             | -0.05           |                           |                           |                           | 0.538                     |
| GHA   | Ghana                   | 25 | 14               | 20                | 0.90             | 0.96             | -0.07           |                           |                           |                           | 0.568                     |
| GLD   | Gauloise doree          | 11 | 4                | 4                 | 0.40             | 0.18             | 0.55            |                           |                           |                           | 0.409                     |
| GLG   | Gauloise grise          | 18 | 6                | 9                 | 0.71             | 0.83             | -0.17           |                           |                           |                           | 0.479                     |
| GLN   | Gauloise noire          | 17 | 5                | 7                 | 0.73             | 0.76             | -0.04           |                           |                           |                           | 0.473                     |
| GLT   | Geline de Touraine      | 24 | 6                | 10                | 0.71             | 0.67             | 0.06            |                           |                           |                           | 0.581                     |
| GOU   | Gournay                 | 21 | 6                | 10                | 0.71             | 0.76             | -0.08           |                           |                           |                           | 0.542                     |
| MAR   | Maroc                   | 21 | 10               | 16                | 0.86             | 0.90             | -0.06           |                           |                           |                           |                           |
| NCH   | Noire de Challans       | 18 | 7                | 10                | 0.82             | 0.67             | 0.19            |                           |                           |                           | 0.445                     |
| P04   | Fayoumi                 | 12 | 5                | 6                 | 0.71             | 0.75             | -0.06           | 0.35                      | 0.44                      | 0.33                      |                           |
| P05   | Bedouin                 | 23 | 13               | 20                | 0.89             | 0.96             | -0.07           | 0.58                      | 0.59                      | 0.54                      |                           |
| P06   | Westfaeliche Totleger   | 10 | 6                | 7                 | 0.80             | 0.60             | 0.25            | 0.44                      | 0.32                      |                           |                           |
| P07   | Sundheimer              | 14 | 3                | 3                 | 0.40             | 0.43             | -0.06           | 0.49                      | 0.48                      |                           |                           |
| P08   | Light Brown Leghorn     | 12 | 4                | 6                 | 0.71             | 0.83             | -0.17           | 0.41                      |                           |                           |                           |
| P09   | Owl-bearded             | 24 | 8                | 13                | 0.84             | 0.54             | 0.35            | 0.54                      |                           |                           |                           |
| P10   | Friesian fowl           | 12 | 6                | 7                 | 0.79             | 0.58             | 0.26            | 0.39                      |                           |                           |                           |
| P11   | Bresse noire            | 18 | 8                | 12                | 0.80             | 0.72             | 0.10            | 0.50                      |                           |                           |                           |
| P12   | Houdan                  | 11 | 3                | 5                 | 0.61             | 0.45             | 0.25            | 0.22                      |                           |                           | 0.426                     |
| P13   | Marans                  | 11 | 5                | 7                 | 0.75             | 0.82             | -0.09           | 0.42                      | 0.50                      | 0.51                      | 0.492                     |
| P14   | Dorking                 | 17 | 5                | 6                 | 0.49             | 0.53             | -0.07           | 0.50                      |                           |                           |                           |
| P15   | Cochin                  | 20 | 5                | 9                 | 0.65             | 0.60             | 0.08            | 0.54                      |                           |                           |                           |
| P16   | Icelandic landrace      | 18 | 9                | 13                | 0.83             | 0.67             | 0.20            | 0.48                      | 0.51                      | 0.50                      |                           |
| P17   | Finnish Landrace        | 24 | 6                | 14                | 0.82             | 0.79             | 0.04            | 0.59                      |                           |                           |                           |
| P18   | Old Scand. Ref. Pop.    | 18 | 9                | 12                | 0.85             | 0.89             | -0.05           | 0.45                      | 0.46                      | 0.46                      |                           |
| P19   | Jaerhoens               | 12 | 2                | 3                 | 0.52             | 0.42             | 0.20            | 0.34                      | 0.45                      | 0.35                      |                           |
| P20   | Sicilienne Buttercup    | 18 | 4                | 6                 | 0.65             | 0.67             | -0.02           | 0.50                      |                           |                           |                           |
| P21   | Padovana                | 12 | 3                | 4                 | 0.54             | 0.58             | -0.08           | 0.17                      | 0.36                      | 0.34                      |                           |
| P22   | Black Castellana        | 12 | 5                | 6                 | 0.68             | 0.67             | 0.03            | 0.44                      |                           |                           |                           |
| P23   | Red Villafrantina       | 18 | 4                | 5                 | 0.57             | 0.72             | -0.26           | 0.48                      |                           |                           |                           |
| P24   | Czech Golden Pencilled  | 19 | 7                | 8                 | 0.67             | 0.63             | 0.05            | 0.49                      |                           |                           |                           |
| P25   | Oravka hen              | 24 | 6                | 12                | 0.79             | 0.88             | -0.11           | 0.52                      |                           |                           |                           |
| P26   | Transsylvania N. Neck   | 24 | 5                | 7                 | 0.60             | 0.67             | -0.12           | 0.53                      | 0.44                      | 0.42                      |                           |
| P27   | Green-legged Partridge  | 24 | 5                | 9                 | 0.71             | 0.67             | 0.06            | 0.53                      | 0.44                      | 0.46                      |                           |
| P28   | Orlov                   | 24 | 3                | 4                 | 0.51             | 0.63             | -0.23           | 0.60                      | 0.59                      | 0.56                      |                           |
| P29-1 | Yurlov crower (Russia)  | 13 | 4                | 5                 | 0.58             | 0.54             | 0.07            | 0.62                      |                           |                           |                           |
| P29-2 | Yurlov crower (Ukraine) | 12 | 6                | 7                 | 0.75             | 0.83             | -0.12           | 0.58                      |                           |                           |                           |
| P30   | Ukrainian bearded       | 23 | 9                | 14                | 0.83             | 1.00             | -0.20           | 0.57                      |                           |                           |                           |
| P31   | Poltava clay            | 18 | 4                | 9                 | 0.75             | 0.83             | -0.11           | 0.51                      |                           |                           |                           |
| P32   | C line                  | 12 | 2                | 2                 | 0.52             | 0.00             | 1.00            | 0.05                      | 0.05                      | 0.06                      |                           |
| P33   | Godollo Nhx             | 23 | 8                | 14                | 0.84             | 0.91             | -0.08           | 0.60                      | 0.61                      | 0.59                      |                           |
| P34   | Line Sarcoma R or S     | 12 | 1                | 1                 | 0.00             | 0.00             | NA              | 0.37                      | 0.34                      | 0.37                      |                           |
| P37   | White egg layer A       | 12 | 1                | 1                 | 0.00             | 0.00             | NA              | 0.33                      | 0.35                      | 0.35                      | 0.27                      |
| P38   | Brown egg layer A       | 18 | 3                | 4                 | 0.30             | 0.22             | 0.26            | 0.46                      | 0.44                      |                           |                           |

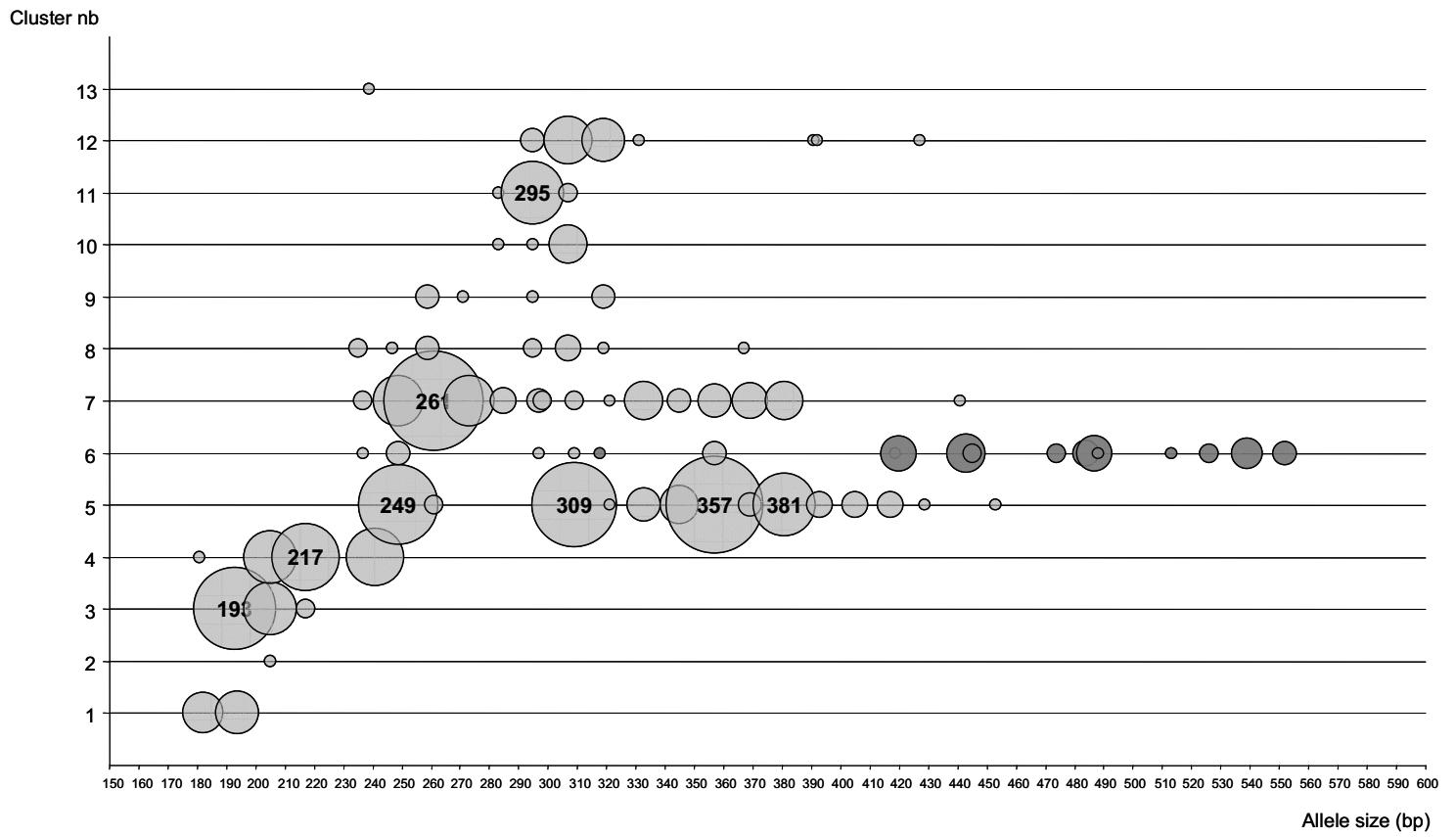
|       |                            |    |    |    |        |        |        |      |      |      |       |
|-------|----------------------------|----|----|----|--------|--------|--------|------|------|------|-------|
| P39   | Broiler dam line A         | 18 | 3  | 4  | 0.45   | 0.50   | -0.11  | 0.57 | 0.58 |      |       |
| P40   | Broiler sire line A        | 23 | 6  | 11 | 0.80   | 0.78   | 0.03   | 0.55 | 0.53 |      |       |
| P41   | Broiler dam line B         | 18 | 5  | 6  | 0.50   | 0.50   | 0.01   | 0.49 |      |      | 0.47  |
| P42   | Broiler sire line B        | 18 | 6  | 14 | 0.82   | 0.83   | -0.02  | 0.47 | 0.51 | 0.47 |       |
| P43   | Brown egg layer B          | 18 | 4  | 5  | 0.53   | 0.56   | -0.05  | 0.45 |      |      | 0.36  |
| P44   | Brown egg layer C          | 12 | 2  | 2  | 0.08   | 0.08   | 0.00   | 0.44 | 0.43 | 0.41 |       |
| P45   | Brown egg layer D          | 18 | 3  | 4  | 0.25   | 0.17   | 0.34   | 0.47 | 0.49 | 0.43 |       |
| P46   | Brown egg layer E          | 12 | 5  | 8  | 0.78   | 0.75   | 0.04   | 0.39 |      |      |       |
| P47   | Broiler sire line C        | 12 | 6  | 6  | 0.76   | 0.83   | -0.10  | 0.44 |      |      | 0.49  |
| P48   | Broiler dam line C         | 11 | 2  | 2  | 0.25   | 0.27   | -0.10  | 0.42 |      |      |       |
| P49   | Broiler sire line D        | 12 | 6  | 8  | 0.79   | 0.83   | -0.06  | 0.44 |      |      | 0.45  |
| P50   | Broiler dam line D         | 22 | 10 | 17 | 0.86   | 0.68   | 0.21   | 0.62 | 0.61 | 0.63 |       |
| P51   | Ab line, high              | 20 | 3  | 6  | 0.63   | 0.50   | 0.21   | 0.51 | 0.49 | 0.45 |       |
| P52   | Ab line, low               | 20 | 3  | 4  | 0.54   | 0.65   | -0.21  | 0.46 |      |      |       |
| P53   | Ab line, control           | 17 | 4  | 5  | 0.37   | 0.29   | 0.20   | 0.51 |      |      |       |
| P101  | <i>G. gallus gallus</i>    | 25 | 8  | 11 | 0.86   | 0.92   | -0.07  | 0.60 | 0.63 |      | 0.66  |
| P102  | <i>G. gallus spadiceus</i> | 24 | 14 | 15 | 0.84   | 0.79   | 0.06   | 0.64 |      | 0.61 | 0.58  |
| RJF   | <i>G. gallus</i>           | 6  | 4  | 3  | 0.76   | 1.00   | -0.32  |      |      |      |       |
| GJF   | <i>G. sommeratii</i>       | 5  | 5  | 4  | 0.82   | 0.80   | 0.03   |      |      |      |       |
| SLJF  | <i>G. lafayetteii</i>      | 4  | 2  | 2  | 0.54   | 0.75   | -0.40  |      |      |      |       |
| SLK   | Silkie                     | 25 | 4  | 8  | 0.71   | 0.44   | 0.38   |      |      |      |       |
| T     | Hsin-Yi                    | 47 | 4  | 7  | 0.61   | 0.51   | 0.16   |      |      |      | 0.534 |
| J     | Ju-Chi                     | 48 | 4  | 5  | 0.38   | 0.40   | -0.05  |      |      |      | 0.458 |
| F     | Hua-Tung                   | 47 | 7  | 18 | 0.84   | 0.85   | -0.01  |      |      |      | 0.584 |
| KM    | Quemoy                     | 47 | 4  | 7  | 0.61   | 0.64   | -0.05  |      |      |      | 0.491 |
| KT    | Shek-Ki                    | 47 | 4  | 9  | 0.73   | 0.70   | 0.04   |      |      |      | 0.423 |
| Ng    | Nagoya                     | 48 | 3  | 5  | 0.53   | 0.58   | -0.09  |      |      |      | 0.438 |
| L2    | Layer Native 2             | 50 | 6  | 9  | 0.45   | 0.42   | 0.07   |      |      |      |       |
| B     | Broiler Native             | 49 | 7  | 14 | 0.74   | 0.69   | 0.07   |      |      |      |       |
| I.R.1 | INRA I.R. line 1           | 49 | 4  | 9  | 0.7042 | 0.7113 | 0.74   |      |      |      |       |
| I.R.2 | INRA I.R. line 2           | 50 | 5  | 9  | 0.6484 | 0.6549 | 0.62   |      |      |      |       |
| I.R.3 | INRA I.R. line 3           | 49 | 4  | 7  | 0.597  | 0.6032 | 0.551  |      |      |      |       |
| I.R.4 | INRA I.R. line 4           | 49 | 4  | 9  | 0.6708 | 0.6777 | 0.6939 |      |      |      |       |

N, number of animals studied;  $N_{\text{all}}$ , number of alleles;  $N_{\text{geno}}$ , number of genotypes;  $H_{\text{exp}}$ , expected heterozygosity ( $H_{\text{n.b.}}$ , unbiased from GENETIX);  $H_{\text{obs}}$ , observed heterozygosity; ref1, Hillel *et al.* 2003; ref2, Granevitz *et al.* 2007; ref3, Rosenberg *et al.* 2001; ref4, Berthouly *et al.* 2008, Berthouly *et al.* 2009, Fotsa 2008, Loukou *et al.* 2009, Loywyck *et al.* 2008, Osei-Amposah *et al.* 2009, Youssao *et al.* 2009

**Table 2** LEI0258 alleles ranked by cluster defined by the network analysis

| name                                   | size (bp) | upstream |     |     |   | R13 | R12 | downstream |    |            |     | N   | cluster n° |    |   |
|--|-----------|----------|-----|-----|---|-----|-----|------------|----|------------|-----|-----|------------|----|---|
| position                               |           | -30-29   | -28 | -10 | 1 |     |     | 5          | 15 | 23-30      | 33  | 39  | 46         |    |   |
| character nb                           |           | 1        | 2   | 3   | 4 |     |     | 5          | 6  | 7          | 8   | 9   | 10         |    |   |
|  |           | TT/Δ     | G/A | G/A | C | n   | n   | C/T        | T  | ATTTTGAG/Δ | Δ/A | A/T | T/A        |    |   |
| B4                                     | 182       |          |     | A   |   | 1   | 2   |            |    | Δ          | A   |     |            | 9  | 1 |
| BW3                                    | 194       |          |     | A   |   | 1   | 3   |            |    | Δ          | A   |     |            | 10 | 1 |
| 1KM                                    | 205       |          |     |     |   | 1   | 4   | T          |    | Δ          |     | T   |            | 1  | 2 |
| B15.1, B11, B61, B27                   | 193       |          |     |     |   | 1   | 3   | T          |    | Δ          |     |     |            | 35 | 3 |
| B13                                    | 205       |          |     |     |   | 1   | 4   | T          |    | Δ          |     |     |            | 15 | 3 |
| 8P18 b                                 | 217       |          |     |     |   | 1   | 5   | T          |    | Δ          |     |     |            | 2  | 3 |
| 1Ng                                    | 181       |          |     |     |   | 1   | 2   |            |    | Δ          |     |     |            | 1  | 4 |
| B13.2, B17, BW11                       | 205       |          |     |     |   | 1   | 4   |            |    | Δ          |     |     |            | 15 | 4 |
| 2B                                     | 217       |          |     |     |   | 1   | 5   |            |    | Δ          |     |     |            | 23 | 4 |
| 2KT                                    | 241       |          |     |     |   | 1   | 7   |            |    | Δ          |     |     |            | 18 | 4 |
| B15.2, B22, B73                        | 249       |          |     |     |   | 1   | 7   |            |    |            |     | T   |            | 32 | 5 |
| 681 ghana b                            | 261       |          |     |     |   | 1   | 8   |            |    |            |     | T   |            | 2  | 5 |
| B10, B24, B26, B76                     | 309       |          |     |     |   | 1   | 12  |            |    |            |     | T   |            | 36 | 5 |
| B74                                    | 321       |          |     |     |   | 1   | 13  |            |    |            |     | T   |            | 1  | 5 |
| 66P101                                 | 333       |          |     |     |   | 1   | 14  |            |    |            |     | T   |            | 6  | 5 |
| B14                                    | 345       |          |     |     |   | 1   | 15  |            |    |            |     | T   |            | 8  | 5 |
| B130, B131, B201, B5.1, B6.1, B21, B75 | 357       |          |     |     |   | 1   | 16  |            |    |            |     | T   |            | 48 | 5 |
| B21.1, BQ, BW1                         | 369       |          |     |     |   | 1   | 17  |            |    |            |     | T   |            | 3  | 5 |
| B13.1                                  | 381       |          |     |     |   | 1   | 18  |            |    |            |     | T   |            | 20 | 5 |
| B1                                     | 393       |          |     |     |   | 1   | 19  |            |    |            |     | T   |            | 4  | 5 |
| B8                                     | 405       |          |     |     |   | 1   | 20  |            |    |            |     | T   |            | 4  | 5 |
| 20P102 a                               | 417       |          |     |     |   | 1   | 21  |            |    |            |     | T   |            | 4  | 5 |
| 513 ghana a                            | 429       |          |     |     |   | 1   | 22  |            |    |            |     | T   |            | 1  | 5 |
| 37P102 a                               | 453       |          |     |     |   | 1   | 24  |            |    |            |     | T   |            | 1  | 5 |
| 2F                                     | 237       |          |     |     |   | 1   | 6   |            |    |            |     |     |            | 1  | 6 |
| 19P5 a                                 | 249       |          |     |     |   | 1   | 7   |            |    |            |     |     |            | 3  | 6 |
| 22P8 b                                 | 297       |          |     |     |   | 1   | 11  |            |    |            |     |     |            | 1  | 6 |
| 534 c.i                                | 309       |          |     |     |   | 1   | 12  |            |    |            |     |     |            | 1  | 6 |
| 104P10                                 | 318       |          |     |     |   | 10  | 3   |            |    |            |     |     |            | 1  | 6 |
| 2006 ben.n b                           | 357       |          |     |     |   | 1   | 16  |            |    |            |     |     |            | 3  | 6 |
| 8F                                     | 419       |          |     |     |   | 15  | 6   |            |    |            |     |     |            | 1  | 6 |
| B62                                    | 420       |          |     |     |   | 16  | 5   |            |    |            |     |     |            | 7  | 6 |
| B6                                     | 443       |          |     |     |   | 15  | 8   |            |    |            |     |     |            | 8  | 6 |
| 6P20                                   | 445       |          |     |     |   | 17  | 6   |            |    |            |     |     |            | 2  | 6 |
| B12.2                                  | 474       |          |     |     |   | 22  | 3   |            |    |            |     |     |            | 2  | 6 |
| 2P11 a                                 | 484       |          |     |     |   | 20  | 6   |            |    |            |     |     |            | 4  | 6 |
| B12                                    | 487       |          |     |     |   | 23  | 3   |            |    |            |     |     |            | 7  | 6 |
| 7B                                     | 488       |          |     |     |   | 24  | 2   |            |    |            |     |     |            | 1  | 6 |
| B12.3                                  | 513       |          |     |     |   | 25  | 3   |            |    |            |     |     |            | 1  | 6 |
| 31P19 a                                | 526       |          |     |     |   | 26  | 3   |            |    |            |     |     |            | 2  | 6 |
| B19                                    | 539       |          |     |     |   | 27  | 3   |            |    |            |     |     |            | 5  | 6 |
| B19.1                                  | 552       |          |     |     |   | 28  | 3   |            |    |            |     |     |            | 3  | 6 |
| 540 ghana b                            | 237       |          |     |     |   | 1   | 6   |            |    |            |     |     | A          | 2  | 7 |
| 132P102                                | 249       |          |     |     |   | 1   | 7   |            |    |            |     |     | A          | 13 | 7 |
| B15, B2, B29                           | 261       |          |     |     |   | 1   | 8   |            |    |            |     |     | A          | 49 | 7 |
| 3T                                     | 273       |          |     |     |   | 1   | 9   |            |    |            |     |     | A          | 13 | 7 |
| 22P102 a                               | 285       |          |     |     |   | 1   | 10  |            |    |            |     |     | A          | 4  | 7 |

|               |     |   |   |   |    |    |   |   |    |    |
|---------------|-----|---|---|---|----|----|---|---|----|----|
| 122P102       | 297 |   |   | 1 | 11 |    |   | A | 3  | 7  |
| 03-44 b       | 298 |   |   | 2 | 10 |    |   | A | 2  | 7  |
| 1P101 b       | 309 |   |   | 1 | 12 |    |   | A | 2  | 7  |
| 16 cam a      | 321 |   |   | 1 | 13 |    |   | A | 1  | 7  |
| BW4           | 333 |   |   | 1 | 14 |    |   | A | 8  | 7  |
| 11P101        | 345 |   |   | 1 | 15 |    |   | A | 3  | 7  |
| B23           | 357 |   |   | 1 | 16 |    |   | A | 6  | 7  |
| 121P102       | 369 |   |   | 1 | 17 |    |   | A | 7  | 7  |
| 46219 G.CL    | 381 |   |   | 1 | 18 |    |   | A | 8  | 7  |
| 04-57 a       | 441 |   |   | 1 | 23 |    |   | A | 1  | 7  |
| 2 cam b       | 235 | Δ |   | 1 | 6  |    |   | A | 2  | 8  |
| B18           | 247 | Δ |   | 1 | 7  |    |   | A | 1  | 8  |
| 12P5 b        | 259 | Δ |   | 1 | 8  |    |   | A | 3  | 8  |
| 34P13         | 295 | Δ |   | 1 | 11 |    |   | A | 2  | 8  |
| 13P23 a       | 307 | Δ |   | 1 | 12 |    |   | A | 4  | 8  |
| 27 n.cha a    | 319 | Δ |   | 1 | 13 |    |   | A | 1  | 8  |
| BC            | 367 | Δ |   | 1 | 17 |    |   | A | 1  | 8  |
| 509 ghana b   | 259 | Δ |   | 1 | 8  | C  |   | A | 3  | 9  |
| 522 c.i a     | 271 | Δ |   | 1 | 9  | C  |   | A | 1  | 9  |
| 515 ghana b   | 295 | Δ |   | 1 | 11 | C  |   | A | 1  | 9  |
| 8P10 a        | 319 | Δ |   | 1 | 13 | C  |   | A | 3  | 9  |
| 10P5 b        | 283 | Δ |   | G | 1  | 10 |   | A | 1  | 10 |
| 235 ge.tou b  | 295 | Δ |   | G | 1  | 11 |   | A | 1  | 10 |
| 14p15         | 307 | Δ |   | G | 1  | 12 |   | A | 8  | 10 |
| 62P101 b      | 283 | Δ |   | 1 | 10 |    |   |   | 1  | 11 |
| B5            | 295 | Δ |   | 1 | 11 |    |   |   | 21 | 11 |
| 7P21 a        | 307 | Δ |   | 1 | 12 |    |   |   | 2  | 11 |
| B11.1         | 295 | Δ | A | 1 | 11 |    |   |   | 3  | 12 |
| B72, B78      | 307 | Δ | A | 1 | 12 |    |   |   | 12 | 12 |
| 6F            | 319 | Δ | A | 1 | 13 |    |   |   | 10 | 12 |
| 5P6 b         | 331 | Δ | A | 1 | 14 |    |   |   | 1  | 12 |
| 3210 barb a   | 391 | Δ | A | 1 | 19 |    |   |   | 1  | 12 |
| 3P06          | 392 | Δ | A | 2 | 18 |    |   |   | 1  | 12 |
| 1739 ge.tou a | 427 | Δ | A | 1 | 22 |    |   |   | 1  | 12 |
| 98-9031 b     | 239 | Δ |   | 1 | 7  |    | Δ |   | 1  | 13 |



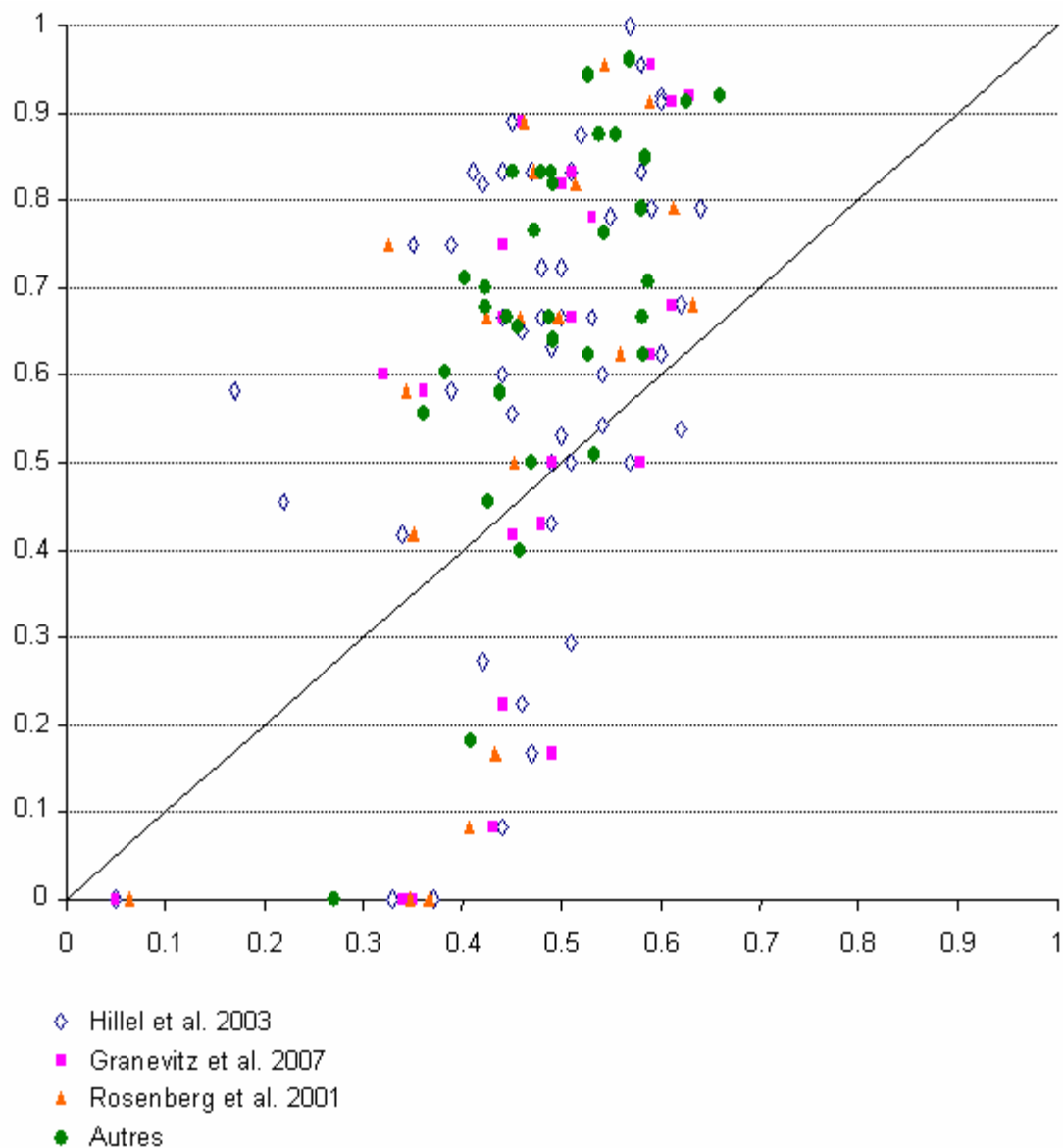
**Supplementary Figure 1** Graphical representation of the LEI0258 alleles (circle size is proportional to the number of observations)

### **C. Résultats complémentaires et Discussion**

L'objectif principal de l'étude était de fournir une image globale de la diversité de la région du CMH du poulet B et d'utiliser les différents allèles identifiés pour évaluer la relation entre les types de CMH. Ceci est particulièrement intéressant quand on sait que le marqueur LEI0258 est très fortement corrélé avec les sérotypes B, associés à des phénotypes de résistance, de susceptibilité et de réponse différentielle aux vaccins.

Au niveau des populations, pour 76 des 84 origines de poulet, nous avons pu trouver une évaluation de l'hétérozygotie observée à partir de marqueurs microsatellites, considérés comme neutres. Les travaux de recherche utilisés pour la comparaison sont ceux de Hillel *et al.*, 2003, Granevitz *et al.*, 2007, Rosenberg *et al.*, 2001, dans ce dernier cas les hétérozygoties ont été recalculées à partir des données utilisées dans l'article, ainsi que les données publiées à partir de génotypes aux marqueurs AvianDiv typés dans le laboratoire de référence INRA, Labogena, de Berthouly *et al.*, 2008, Berthouly *et al.*, 2009, Fotsa 2008, Loukou *et al.*, 2009, Loywyck *et al.*, 2008, Osei-Amposah *et al.*, 2009, Youssao *et al.*, 2009. Dans la majorité des cas, en particulier en ce qui concerne les races locales ou les espèces sauvages, l'hétérozygotie observée au niveau du marqueur LEI0258 était plus élevée que celle observée à l'aide des marqueurs microsatellites répartis ailleurs dans le génome (Figure 9).





**Figure 9.** Représentation graphique des valeurs d'hétérozygotie observée, en ordonnée, au niveau du marqueur LEI0258, en abscisse, au niveau de marqueurs microsatellites répartis ailleurs dans le génome

Afin d'analyser ces résultats nous avons tout d'abord cherché à établir, à l'aide du coefficient de corrélation de Spearman, l'existence de relation entre les deux variables quantitatives (Tableau 4). Ces résultats permettent de dire qu'il existe une corrélation entre les deux variables c'est-à-dire que l'hétérozygotie observée au niveau du CMH est corrélée à celle observée au niveau du génome. Dans un second temps, nous avons voulu tester l'hypothèse d'une différence entre les deux valeurs, à l'aide du test des rangs signés de Wilcoxon, en considérant donc que les échantillons sont appariés. Ces résultats nous permettent d'affirmer qu'il existe une différence significative entre les deux hétérozygoties observées, à savoir que

l'hétérozygotie observée au niveau du CMH est significativement plus élevée que celle observée au niveau du génome.

**Tableau 4.** Résultats de l'analyse de corrélation de Spearman (Rho) et du test des rangs signés de Wilcoxon (W) pour les données en totalité ou les différentes sources, incluant ou excluant les populations expérimentales et commerciales

|   | Données de Comparaison         | Corrélation |      |         | Différence |    |          |
|---|--------------------------------|-------------|------|---------|------------|----|----------|
|   |                                | N           | Rho  | P       | W          | N  | P        |
| <b>Toutes populations</b>                             | Total                          | 77          | 0.42 | < 0.001 | 2047       | 76 | < 0.0001 |
|   | Hillel <i>et al.</i> , 2003    | 51          | 0.38 | < 0.01  | 803        | 51 | < 0.001  |
|   | Granevitz <i>et al.</i> , 2007 | 25          | 0.67 | < 0.001 | 100        | 25 | N.S.     |
|   | Rosenberg <i>et al.</i> , 2001 | 20          | 0.66 | < 0.01  | 90         | 20 | < 0.1    |
|   | Autres                         | 35          | 0.60 | < 0.001 | 478.5      | 34 | < 0.0001 |
| <b>Lignées expérimentales et commerciales exclues</b> | Total                          | 54          | 0.41 | < 0.01  | 1321       | 53 | < 0.0001 |
|   | Hillel <i>et al.</i> 2003      | 32          | 0.36 | < 0.1   | 504        | 32 | < 0.0001 |
|   | Granevitz <i>et al.</i> 2007   | 14          | 0.58 | < 0.1   | 97         | 14 | < 0.01   |
|   | Rosenberg <i>et al.</i> 2001   | 12          | 0.50 | < 0.1   | 78         | 12 | < 0.01   |
|   | Autres                         | 26          | 0.55 | < 0.01  | 289        | 25 | < 0.001  |

Ces résultats sont nouveaux car la plupart des comparaisons d'hétérozygotie faites au niveau du CMH ont été réalisées entre l'hétérozygotie observée et celle théorique attendue en l'absence de sélection par exemple sous les hypothèses de l'équilibre d'Hardy-Weinberg. Alors que dans notre situation, ce qui est comparé est bien l'hétérozygotie observée au niveau du CMH, région qui peut être soumise à des pressions de sélection spécifiques et l'hétérozygotie observée pour le reste du génome.

En ce qui concerne la corrélation entre allèles au marqueur LEI0258 et sérotypes, dans cette étude par rapport à celle de Fulton *et al.*, (2006), l'information de sérologie n'était disponible que pour un petit nombre d'animaux. Globalement, les mêmes résultats, c'est à dire une bonne corrélation entre les deux informations, ont été obtenus.

Enfin, pour les inférences pouvant être faites sur les modes d'évolution du marqueur LEI0258, celles-ci concernent les différents niveaux de variation observés. L'évolution du nombre de répétitions du marqueur est probablement essentiellement basée sur un modèle "step-wise", un pas étant une duplication ou la perte d'un motif. Cependant la combinaison des nombres de répétitions de chaque motif semble être régie par d'autres contraintes qui pourraient notamment limiter la taille totale des allèles.

Quel que soit le mécanisme impliqué dans l'évolution du nombre de répétition, la zone répétée est associée à un certain nombre de SNP et d'indels dans les régions flanquantes. Dans cette étude, nous avons choisi d'analyser les relations entre les différents allèles du marqueur en nous basant uniquement sur ces polymorphismes sous forme d'un réseau représentant tous les événements possibles. Ainsi, cette analyse doit être interprétée comme une représentation de la relation entre les allèles et non une phylogénie. Ce choix semble pertinent puisque les différentes combinaisons de nombre de répétitions ainsi que les tailles d'allèles apparaissent associées à différentes parties du réseau. Si l'on considère que les indels de 2 et 8 pb ont plus de poids que les autres mutations ponctuelles, alors le réseau peut être considéré comme divisé en trois parties: trois clusters sans délétion (en haut), cinq clusters présentant la délétion de 2 pb (en bas à gauche), quatre clusters présentant la délétion de 8 pb (en bas à droite), et un cluster d'un seul allèle présentant les deux délétions.

Il est intéressant d'observer que les allèles provenant des espèces sauvages, sont observés dans sept des 13 clusters définis à partir des mutations ponctuelles, et dans les 4 grands groupes définis ci-dessus. Cela signifie qu'une partie importante de la diversité observée chez les poulets domestiques avec le LEI0258, qui est liée à la diversité du CMH, se trouve chez les ancêtres sauvages de ces poulets. La diversité ancestrale du CMH a donc été maintenue chez les poulets domestiques, presque certainement en raison des caractéristiques particulières de cette région.

La publication de cette étude devrait convaincre la communauté des chercheurs concernés que le génotypage du marqueur LEI0258 peut être fait pour n'importe quel type de poulet et même pour les espèces sauvages. Si l'usage du marqueur devient systématique, cela devrait grandement faciliter les comparaisons entre les différentes études menées et avec les deux référentiels disponibles à ce jour, à savoir celui de Fulton *et al.*, 2006, plutôt axé sur les lignées expérimentales ou commerciales pondeuses oeufs blancs, et le nôtre.

## **II DEUXIEME PARTIE : ETUDE DE LA CORRELATION ENTRE GENOTYPAGE ET SEROLOGIE CHEZ LES LIGNEES EXPERIMENTALES WHITE LEGHORN R.I., A L'AIDE DU MARQUEUR LEI0258 ET D'UN JEU REDUIT DE SNPS**

### **A. Introduction**

Cette étude a bénéficié de l'existence au laboratoire d'un dispositif de sélection expérimentale sur la réponse immunitaire qui a accumulé des données de sérologie du CMH sur dix générations. Ce dispositif a pour but d'étudier les stratégies possibles de sélection de la réponse immunitaire chez le poulet. En effet, les systèmes de production avicole en Europe sont confrontés à une tendance à réduire la consommation d'antibiotiques et d'antiparasitaires, renforçant ainsi l'intérêt d'animaux génétiquement résistants, répondant mieux à la vaccination ou robustes en terme de maintien du niveau de production. Des indicateurs de ce type ont toujours été recherchés comme objectifs de sélection. Toutefois, des critères de sélection fiables et pratiques sont encore en cours de mise au point. En outre, la résistance aux maladies est particulièrement difficile à caractériser dans les nombreuses races locales répertoriées dans le monde pour lesquels des données expérimentales sont rares et généralement seul l'ADN génomique peut-être accessible. C'est pourquoi la disponibilité en outils moléculaires est nécessaire afin de caractériser la diversité fonctionnelle en rapport avec la résistance aux maladies dans ces populations.

Le génotypage moléculaire du CMH du poulet est particulièrement indiqué car il comprend les gènes principalement impliqués dans la présentation de peptides aux cellules T et associés à la résistance aux maladies et à la réponse immunitaire. La division du CMH du poulet en deux régions génétiquement indépendantes (B et Y) est à prendre en compte. Historiquement, c'est principalement la région B qui a été étudiée, par des techniques de sérologie. Par contre des techniques de biologie moléculaires ont pu être appliquées pour les régions B et Y, avec notamment des techniques RFLP et de SSCP. Trois marqueurs atypiques ont été décrits dans la région B : 2 répétitions d'un mononucléotide, MCW0370 et MCW0371, et un VNTR, LEI0258. Leur corrélation avec la sérologie a été étudiée pour des lignées de référence et ils ont été proposés en remplacement des tests sérologiques. Cependant des études supplémentaires sont nécessaires afin de comprendre pleinement les limites de leur corrélation avec les tests sérologiques. En effet, ils ne sont pas suffisamment polymorphes pour décrire toute la variabilité des séquences codantes du CMH, en particulier au niveau des régions impliquées dans la formation des complexes peptide / molécule de classe I ou de classe II. La

rareté des répétitions utilisables comme marqueurs dans la région du CMH chez le poulet rend les marqueurs de type SNP (single nucleotide polymorphism) particulièrement intéressants. Leur utilisation devrait contribuer à établir et valider une corrélation entre la sérologie classique et les polymorphismes au niveau des gènes.

Pour ce faire, l'association entre des marqueurs moléculaires et les typages en sérologie ont été étudiés chez les lignées expérimentales White Leghorn R.I. sélectionnées pour des caractères en relation avec la réponse immunitaire. Différents marqueurs, les trois marqueurs atypiques MCW0370 et MCW0371 et LEI0258 ainsi que 10 SNPs, répartis le long de la séquence du CMH, ont été génotypés. La corrélation entre les allèles des marqueurs atypiques et la sérologie a été étudiée dans plus de 800 animaux et la correspondance entre les SNPs et les allèles de ces marqueurs a été définie. Un tableau complet de correspondance entre les tailles des allèles des marqueurs atypiques, les génotypes aux SNPs et la sérologie a ensuite pu être construit. Cette étude montre que l'intégration de nouveaux outils moléculaires peuvent certainement aider à la caractérisation des haplotypes du CMH.

**Ce travail a été présenté au 1er symposium international AGAH (Animal Genomics for Animal Health) en octobre 2007 et à ce titre est paru dans Developments in Biologicals (Basel), volume 132, pages 271-278.**

## B. Article de recherche

Pinard M-H, Gay C, Pastoret P-P, Dodet B (eds): Animal Genomics for Animal Health. Dev Biol (Basel). Basel, Karger, 2008, vol 132, p 271-278.

.....

# Molecular Genotype Investigation of the *Gallus gallus* Major Histocompatibility Complex

O. Chazara<sup>1</sup>, M.H. Pinard-Van Der Laan<sup>1</sup>, M. Tixier-Boichard<sup>1</sup>, B. Bed'hom<sup>1</sup>

<sup>1</sup> UMR 1236 Génétique et Diversité Animales, INRA/AgroParisTech, 78352 Jouy en Josas, France

**Key words:** *Gallus gallus*, major histocompatibility complex, SNPs, microsatellite

**Abstract:** The association between genetic markers and serological major histocompatibility complex (MHC) definition was studied in experimental White Leghorn lines selected for different immune traits. Different markers, three variable number tandem repeats (VNTRs) and 10 single nucleotide polymorphisms (SNPs), distributed along the MHC sequence were genotyped. Correlation between VNTR marker alleles and serology was studied in more than 800 animals and correspondence between SNP haplotypes and allele size was defined. A complete table of correspondence between VNTR allele sizes, SNP genotypes and serology was then drawn. This research shows that integrating new molecular tools can definitely help in MHC haplotype characterization and that, since the chicken MHC locus has been associated with disease resistance, these tools are becoming more and more needed for disease resistance and immunogenomics studies.

## INTRODUCTION

Chickens are the primary animal protein source of human food at the global level. Today in Europe, poultry production systems are facing a trend towards reduced use of drugs, thus strengthening the need for robust animals. Moreover, disease resistance has always been considered in selection objectives. Genetic variability for disease resistance has been shown to take place both between and within lines or breeds of chickens. However, reliable and convenient selection criteria are still under investigation. Furthermore, disease resistance is still difficult to characterize in the numerous local breeds identified worldwide for which experimental data is scarce and generally only genomic DNA is accessible. Thus, universal molecular

tools are needed to characterize functional diversity regarding disease resistance across populations.

The chicken major histocompatibility complex (MHC) is a genomic region comprising genes coding for proteins mainly involved in peptide presentation to T cells (Class-I and Class-II molecules). The chicken MHC locus has been associated with disease resistance, susceptibility, and differential response to vaccine. A major concern has arisen about the fine dissection of molecular variation in this locus which is divided in two genetically independent regions (B and Y). Molecular MHC genotyping tools in the chicken would permit exploration of the diversity of more breeds at this level. Historically, the B-locus genotype has been studied using serology, and DNA studies of the B- and Y-locus polymorphisms have been based on restriction fragment length polymorphisms (RFLPs) and single-strand conformation polymorphism (SSCP). Several atypical variable number tandem repeat (VNTR) markers, including MCW0370, MCW0371, and LEI0258, are now available to replace the serological tests, but further studies are required in order to fully understand the limits of their correlation with the serology tests. They are also not polymorphic enough to describe the whole variability of coding sequences, particularly at the level of the peptide binding region sequence of the Class-I and Class-II genes. A specific single nucleotide polymorphism (SNP) strategy would help to establish and validate a correlation between classical serology and polymorphism at the gene level.

## MATERIALS AND METHODS

White Leghorn lines were previously selected for 12 generations at INRA for different immune traits (line 1, antibody response to Newcastle vaccine; line 2, delayed hypersensitivity measured by phytohemagglutinin; line 3, carbon clearance as an indicator of phagocytosis; line 4, control [1]). The founders were genotyped with three atypical markers, LEI0258, MCW0370 and MCW0371 as described in Fulton et al. [2] (Table 1, Fig. 1) and studied by serology, to determine the correlation between the markers and serology in those lines. Other animals from the selected lines, including one whole generation (generation 9, with 800 animals born in 2003 from 110 families) were genotyped for LEI0258 (Fig. 2). Twenty-four animals, six per selected line, from another generation of selection (generation 12, animals born in 2006) were studied for SNP using 10 SNPs predicted in the chicken genetic variation map of the International Chicken Polymorphism Map Consortium [3] (Table 1), chosen and typed by Wageningen University in the framework of an Eadgene project focusing on the genetic basis of the immune response within chickens. Eight of those SNPs are located in the B-region and two are in the Y-region (Fig. 1); they were typed using Illumina technology. PHASE 2.1 software [4,5] was used; it implements a Bayesian statistical method for reconstructing haplotypes from population SNP data. A Neighbour-Joining unrooted tree was then constructed, based on the number of differences, using MEGA 3.1 [6].

## RESULTS

LEI0258, MCW0370 and MCW0371 observed allele sizes for the founders were compared to the serology previously determined for those animals (B15, B19, B21, B34, and B124). A strong correlation was found between the MHC haplotypes determined by both techniques, and a correspondence table was made (Table 2). Genotyping selection generation 9 for LEI0258 confirmed those results, but 8.6% incompatibility was observed between the known serology and the LEI0258 alleles (261 bp, 359 bp, 422 bp, 489 bp and 544 bp). One example of a family in which such incompatibilities were found is shown in Table 3; parents are from generation 8 with known serology but no LEI0258 typing. In this family, parents' serology was consistent with that of the progeny, but three offsprings exhibited unexpected LEI0258 genotypes that were incompatible with their own serology and their parents'. This

example only shows incompatibilities regarding B15 and B21 haplotypes, but similar situations of incompatibility occurred for all five observed haplotypes in the generation under study.

*Table 1:* SNP or marker characteristics (name, chromosome, position in bp on reversed sequence published by Shiina et al. [8], region information, known variations, 5' and 3' near 30 bp sequence for SNPs or forward and reverse primers).

| Name       | Chr | Position | Region    | Variations | 5' near 30 bp or forward primer  | 3' near 30 bp or reverse primer |
|------------|-----|----------|-----------|------------|----------------------------------|---------------------------------|
| rs16057122 | 16  | 3386     | NA        | [A/G]      | atggataattagaccatag-gtttcacagc   | tgtgcagaaatccaacc-tatgagtacttc  |
| rs15026791 | 16  | 65849    | BF2       | [A/G]      | gctgatatgcaagaaca-cacggtgaccac   | gggacaaccaggagc-catggggacatgtgg |
| rs15788177 | 16  | 112368   | Blec1     | [A/G]      | gttgactggaaaagcatgaa-gaaagcacc   | catgatgagag-gaacagttcatcccacagc |
| MCW0370    | 16  | 112482   | Blec1     | (A)n       | GAAAGAGGAGAG-TAGTTCACG           | TCCCACCTCAT-CATGCATTCC          |
| MCW0371    | 16  | 114542   | Blec2     | (A)n       | TTTCATGGCATCC-TAAGATGG           | CTGCTCCGAGCT-GTAATCCTG          |
| rs15026773 | 16  | 116738   | Blec2     | [G/T]      | acacggacggaggctgag-gaggcagaaaatt | tgtcagaattccgtcagctc-cgccgttttc |
| LEI0258    | 16  | 124903   | BG1       | 12-13 bp   | TCGGGAAAA-GATCTGAGTCATTG         | TGATTTTCA-GATCGCGTTCCTC         |
| rs15788101 | 16  | 133829   | B-BTN2    | [C/T]      | aggcagcagcctcagctc-caacccttctca  | attcccgtctcac-cgggctttatctcctt  |
| rs15788011 | 16  | 165938   | TRIM2 7.2 | [A/G]      | aacgcagcagggattct-taggggaataagg  | tgtaaaagataaacaca-gatggaagattaa |
| rs15026648 | 16  | 211705   | Bzfp2     | [A/C]      | cagctcccaaaagagcaacacaaccaaac    | gtcctcaaccattcccttgc-aaatatgt   |
| rs15787808 | 16  | 227307   | Bzfp3     | [A/C]      | tgtgaatggatgggt-gagtctcccctctca  | attaaaaatgttggggtcttc-ctgtgagag |
| rs14096793 | 16  | NA       | Y-region  | [C/T]      | tgtttcgttctctcctcctc-cttctctc    | ttccttccctcctttcttctc-cttctctt  |
| rs16742957 | 16  | NA       | Y-region  | [C/T]      | caactgcactatgtgtctgccc-ctccacac  | tctgaaccaccg-taaggcctgaaggaca   |



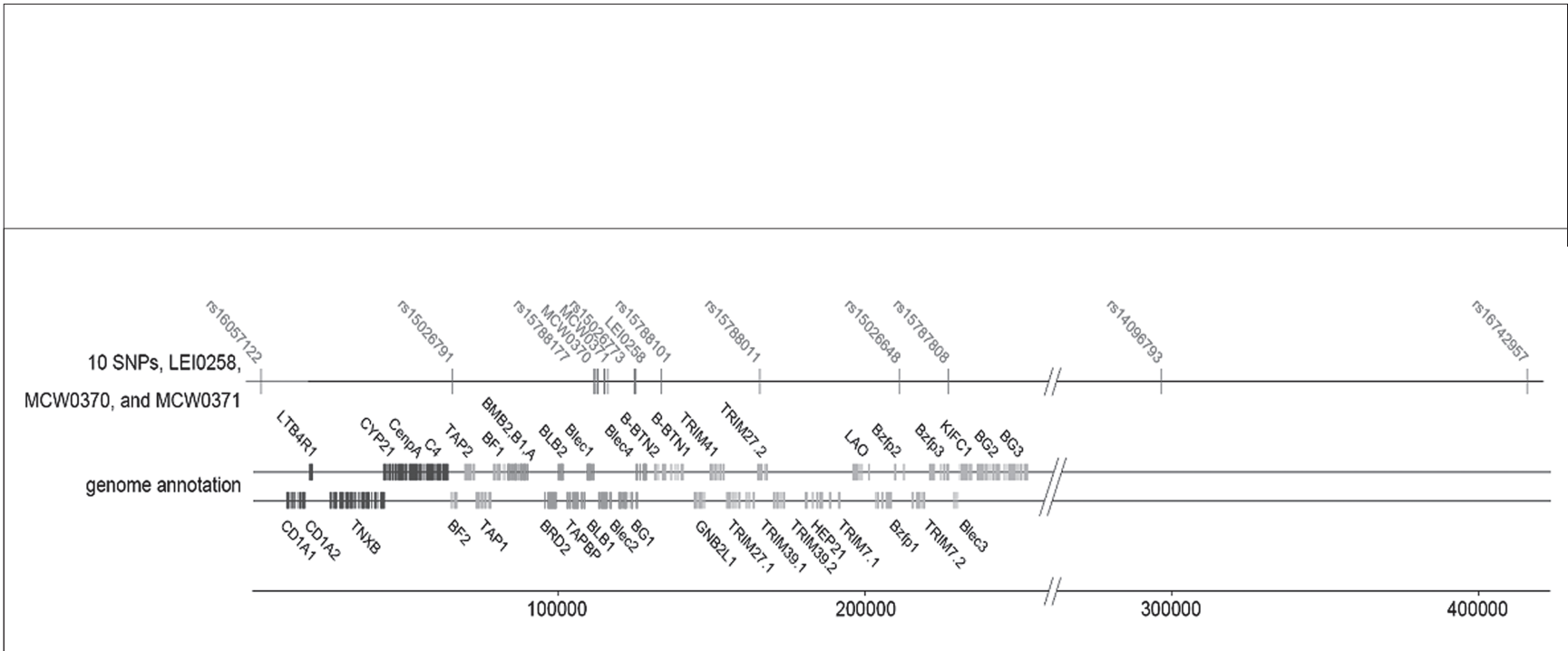


Fig 1: Schema of the 242 kb B-region showing gene content (modified from [8]) and positions of the markers (MCW0370, MCW0371, LEI0258) and the 10 SNPs.

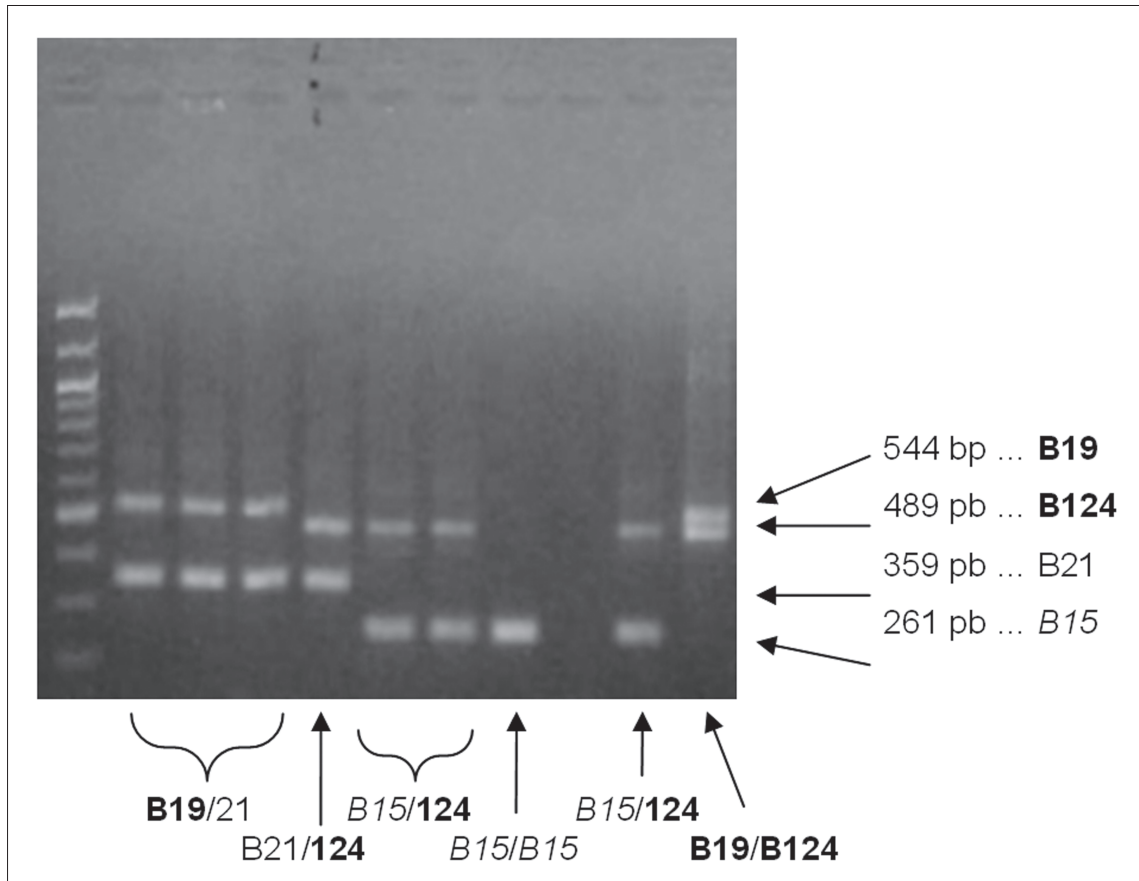


Fig 2: Marker LEI0258 PCR products separated on an agarose gel showing the correspondence between the allele sizes and the MHC B-region defined by serology.

Table 2: Table of correspondence between serology; MCW0370, MCW0371, and LEI0258 allele sizes and haplotypes defined by the eight SNPs located in the MHC B-region.

| Serology    | MCW0370     | MCW0371    | LEI0258    | Haplotype(s)                    |
|-------------|-------------|------------|------------|---------------------------------|
| <i>B15</i>  | <i>null</i> | <i>201</i> | <i>261</i> | <i>ABBBABAA</i> <i>ABBBABAB</i> |
| B21         | 181         | 203        | 359        | BAAAAABA                        |
| <b>B34</b>  | <b>179</b>  | <b>203</b> | <b>422</b> | <b>ABABABBA</b>                 |
| <b>B19</b>  | <b>178</b>  | <b>203</b> | <b>544</b> | <b>BAABBBBB</b> <b>BBABBBBB</b> |
| <b>B124</b> | <b>179</b>  | <b>203</b> | <b>489</b> | <b>BAABBBBB</b> <b>BBABBBBB</b> |

Table 3: Comparison between serology and LEI0258 allele sizes showing incompatibilities in one family from selection generations 8 and 9

|             | Serology   |             | LEI0258          |                   |
|-------------|------------|-------------|------------------|-------------------|
| Male        | <i>B15</i> | B21         |                  |                   |
| Female      | <i>B15</i> | <b>B124</b> |                  |                   |
| Offspring 1 | <i>B15</i> | B21         | 359 = B21        | 359 = B21         |
| Offspring 2 | <i>B15</i> | B21         | 359 = B21        | 359 = B21         |
| Offspring 3 | B21        | <b>B124</b> | 359 = B21        | <b>489 = B124</b> |
| Offspring 4 | B21        | <b>B124</b> | 359 = B21        | <b>489 = B124</b> |
| Offspring 5 | <i>B15</i> | B21         | 261 = <i>B15</i> | 359 = B21         |
| Offspring 6 | B21        | <b>B124</b> | 359 = B21        | <b>489 = B124</b> |
| Offspring 7 | <i>B15</i> | B21         | 359 = B21        | 359 = B21         |

Table 4: Comparison between LEI0258 allele sizes and haplotypes defined by 10 SNPs for 24 individuals from the four selected lines.

| ind | line | LEI0258    | 10 SNP haplotypes |                                       |
|-----|------|------------|-------------------|---------------------------------------|
| 1   | 1    | 359        | <b>489</b>        | BAAAAABA BA <b>BAABBBBB BA</b>        |
| 2   | 1    | 359        | <b>489</b>        | BAAAAABA AA <b>BAABBBBB AA</b>        |
| 3   | 1    | 359        | <b>489</b>        | BAAAAABA AA <b>BAABBBBB BA</b>        |
| 4   | 1    | <b>489</b> | <b>489</b>        | <b>BBABBBBB AA</b> <b>BBABBBBB AA</b> |
| 11  | 2    | <b>489</b> | <b>544</b>        | <b>BBABBBBB BA</b> <b>BBABBBBB AA</b> |
| 23  | 4    | <b>489</b> | <b>489</b>        | <b>BBABBBBB BA</b> <b>BBABBBBB AA</b> |
| 5   | 1    | 422        | <b>489</b>        | <i>ABABABBA AA</i> <b>BBABBBBB AA</b> |
| 6   | 1    | 359        | 359               | BAAAAABA AA BAAAAABA AA               |
| 7   | 2    | <b>489</b> | <b>544</b>        | <b>BBABBBBB BA</b> <b>BBABBBBB BA</b> |
| 8   | 2    | <b>489</b> | <b>544</b>        | <b>BBABBBBB BA</b> <b>BBABBBBB BA</b> |
| 10  | 2    | <b>489</b> | <b>489</b>        | <b>BBABBBBB BA</b> <b>BBABBBBB AA</b> |
| 13  | 3    | 261        | 359               | <i>ABBBABAB BA</i> BAAAAABA AA        |
| 18  | 3    | 261        | 359               | <i>ABBBABAB BA</i> BAAAAABA AA        |
| 22  | 4    | 261        | 359               | <i>ABBBABAB BA</i> BAAAAABA AA        |
| 14  | 3    | 261        | 359               | <i>ABBBABAB BA</i> BAAAAABA BA        |
| 24  | 4    | 261        | 359               | <i>ABBBABAB AA</i> BAAAAABA AA        |
| 15  | 3    | 261        | 261               | <i>ABBBABAA BA</i> <i>ABBBABAB AA</i> |
| 17  | 3    | 261        | 261               | <i>ABBBABAA BA</i> <i>ABBBABAB BA</i> |
| 9   | 2    | 261        | <b>544</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB AA</b> |
| 12  | 2    | 261        | <b>544</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB AA</b> |
| 16  | 3    | 261        | <b>544</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB AA</b> |
| 19  | 4    | 261        | <b>489</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB AA</b> |
| 20  | 4    | 261        | <b>489</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB AA</b> |
| 21  | 4    | 261        | <b>489</b>        | <i>ABBBABAA BA</i> <b>BBABBBBB BA</b> |

Genotyping of the selected lines revealed a new allele size for LEI0258, 557 bp, in one animal from the generation 11 [7]. This new allele is probably mutated from the mother's 544 bp allele, the 557 bp size would be obtained by one additional 13 bp repeat, as LEI0258 is a 12-13 bp repetition.

For the eight typed SNPs located in region B, the PHASE software predicted the most likely haplotype pairs for each individual with a probability ranging from 0.658 to 1.000. Only the information for the eight SNPs located in region B was used because one of the two SNPs located in region Y was monomorphic and the other seemed independent, so the 10 SNP genotypes could not be considered as linked. Predicted haplotypes were compared with independently typed LEI0258 alleles and with the serological typing (Table 4) and used to complete the correspondence table (Table 2). This table shows several discrepancies between sequence polymorphism on one hand and LEI0258 alleles or serology on the other: two SNP haplotypes could be found for a given LEI0258 allele (as with B15) and SNP haplotypes could not be used to discriminate B19 from B124. The haplotypes from the eight B-region SNPs were used to construct a tree (Fig. 3) which, interestingly, clustered on one hand with B15 and B34 and B21, B19 and B124 on the other hand.

## DISCUSSION

In this report, we describe the application of molecular approaches based on differences due to repeat length in three markers, LEI0258, MCW0370, and MCW0371, and 10 SNPs, on the characterization of the MHC region in four lines selected for different immune traits. This approach also helps to determine the extent to which integrating molecular tools can help with MHC haplotype characterization, which is becoming more and more needed for disease resistance and immunogenomics studies.

For example, in this study, we detected incompatibilities between serology and LEI0258 within families that could be explained by serology errors and parentage errors as well as by recombination within the MHC locus.

Typing SNPs yields information on other parts of the MHC B-region, indicating possible recombination spots. The 4<sup>th</sup> and 5<sup>th</sup> SNPs surrounding the LEI0258 locus are an example of this, and may indicate that B15 and B34 (BA) are the result of a recombination between B19 or B124 (BB) and B21 (AA). The 7<sup>th</sup> and 8<sup>th</sup> SNPs are an example of this as we can find the four combination : AA, AB, BA and BB.

This research focused on well-documented experimental lines that allowed for the study of allele segregation. The next step would be to examine various others breeds, including local breeds that have never been explored at this level.

## REFERENCES

- 1 Pinard-van der Laan M-H: Immune modulation: The genetic approach. *Vet Immunol Immunopathol* 2002;87:199-205.
- 2 Fulton JE, Juul-Madsen HR, Ashwell CM, McCarron AM, Arthur JA, O'Sullivan NP, et al.: Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* 2006;58:407-421

- 3 International Chicken Polymorphism Map Consortium: A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 2004;432:717-722.
- 4 Stephens M, Smith NJ, Donnelly PA: New statistical method for haplotype reconstruction from population data. *Am J of Hum Genet* 2001;68:978-989.
- 5 Stephens M, Scheet P: Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am J Hum Genet* 2005;76:449-462.
- 6 Kumar S, Tamura K, Nei M: MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 2004;5:150-163.
- 7 Loywyck V, Bed'Hom B, Pinard-van der Laan M-H, Pitel F, Verrier E, Bijma P: Evolution of polymorphisms at molecular markers in QTL and non-QTL regions in selected lines of chicken. *Genet Select Evol* 2007;submitted.
- 8 Shiina T, Briles WE, Goto RM., Hosomichi K, Yanagiya K, Shimizu S, et al.: Extended Gene Map Reveals Tripartite Motif, C-Type Lectin, and Ig Superfamily Type Genes within a Subregion of the Chicken MHC-B Affecting Infectious Disease. *J Immunol* 2007;178:7162-7172.

Olympe Chazara, UMR INRA/AgroParisTech, Génétique et Diversité Animales, Domaine de Vilvert, 78352 Jouy en Josas, France.  
E-mail: [olympe.chazara@jouy.inra.fr](mailto:olympe.chazara@jouy.inra.fr)

## C. Discussion et Conclusion

Cette étude, basée sur des lignées expérimentales bien décrites, en pedigree et en sérologie, a permis l'étude des corrélations entre marqueurs et sérologie d'une part, et l'analyse de la ségrégation des allèles du marqueur LEI0258. Nous avons appliqué différentes approches de génotypage moléculaire reposant soit sur les différences de taille de trois marqueurs, LEI0258, MCW0370 et MCW0371, soit sur les polymorphismes de séquence de 10 SNPs, afin de caractériser la région du CMH dans quatre lignées sélectionnées pour différents critères de réponse immunitaire. Cette approche montre dans quelle mesure l'intégration des outils moléculaires peut permettre de génotyper plus finement le CMH chez le poulet, ce qui est devenu nécessaire dans les études de résistance aux maladies.

Plus particulièrement, nous avons pu mettre en évidence, des incompatibilités entre la sérologie et les allèles au marqueur LEI0258 au sein de familles, qui pourrait s'expliquer par des erreurs de sérologie et les erreurs de filiation mais également par des recombinaisons intra-CMH. Les résultats obtenus avec le génotypage de SNPs situés dans d'autres zones de la région B du CMH pourraient ainsi indiquer des recombinaisons. Par exemple, l'examen des 4<sup>e</sup> et 5<sup>e</sup> SNPs, qui entourent le locus LEI0258, semble indiquer que B15 et B34 (génotype BA) sont le résultat d'une recombinaison entre B19 ou B124 (génotype BB) et B21 (génotype AA).

Ces résultats indiquent que l'utilisation des marqueurs SNPs est effectivement un apport appréciable à la connaissance de la région, il s'agit maintenant de les caractériser afin de pouvoir les génotyper dans d'autres origines ou races de poulet, notamment les races locales.



### **III TROISIEME PARTIE : IDENTIFICATION ET CARACTERISATION DE SNPs DANS LA REGION B DU CMH DU POULET PAR RESEQUENCAGE DE 11 GENES CIBLES**

#### **A. Introduction**

Le but de ce travail est de pouvoir proposer un typage moléculaire du CMH pour tout type de poulet, y compris les races locales, en utilisant un marqueur appelé LEI0258, combiné à des marqueurs SNPs. Le marqueur LEI0258 a donc été précisément caractérisé et des séquences cibles de la région B du MHC ont été déterminées pour en dériver des SNPs par reséquençage. La corrélation entre ce typage moléculaire et la classification sérologique traditionnelle a également été étudiée, dans des lignées expérimentales White Leghorn. Le panel de marqueurs a été étudié sur un large panel de races expérimentales, commerciales et locales.

La validation de marqueurs permettant le typage moléculaire du CMH chez le poulet autorise l'exploration de la diversité génétique de cette région dans de nouvelles lignées et races. Cela participe au développement de stratégies alternatives et intégrées de lutte contre des pathologies majeures de l'élevage aviaire.

**Ce travail, concernant seulement 10 gènes, a été présenté aux 8èmes JRA (Journées de la Recherche Avicole) en mars 2009. L'article court en français présenté ici est très similaire, sauf que les données concernant le 11e gène (TRIM41) ont été ajoutées.**



## **B. Article**

### **TYPAGE MOLECULAIRE DU COMPLEXE MAJEUR D'HISTOCOMPATIBILITE CHEZ LE POULET : VERS UNE MEILLEURE CARACTERISATION DES GENOTYPES**

**Chazara Olympe, Pinard-van der Laan Marie H el ene, Tixier-Boichard Mich ele,  
Bed'Hom Bertrand**

#### **ABSTRACT**

##### **Molecular characterization of the Major Histocompatibility Complex in Chicken: towards a better description of genotypes**

The major histocompatibility complex (MHC) is a complex genomic region in Vertebrates, still imperfectly known in the chicken and which shows a great genetic variability. The MHC plays a central role in the immune response of an animal to infectious diseases: it contains the genes coding for proteins that present antigens to the lymphocytes.

Characterization of MHC in the chicken is classically performed by serology and tests are available only for already known genotypes. The purpose of this study is to develop a molecular MHC genotyping for any type of chicken, including local breeds, using a marker called LEI0258 combined with SNP markers.

The marker LEI0258 was precisely characterized and SNP markers have been obtained by resequencing specific targets in MHC B region. Correlation between molecular typing and traditional serological classification has been established in experimental lines of White Leghorn. The markers panel has then been studied in a wide range of experimental, commercial and local breeds.

Validation of markers for molecular genotyping of the Chicken MHC will permit the exploration of the genetic diversity of this region in new lines and breeds. This work contributes to the development of alternative strategies and global fight against major diseases in poultry breeding.

**Keywords:** Major Histocompatibility Complex; Chicken; Genotyping; Disease resistance

## **Introduction**

Le complexe majeur d'histocompatibilité (CMH) est une région génomique complexe des Vertébrés, encore imparfaitement connue chez le Poulet (Kaufman *et al.*, 1999) et qui présente à certains loci une très grande variabilité génétique. Le CMH joue un rôle central dans l'organisation de la réponse immune d'un animal aux pathologies infectieuses. Une meilleure connaissance du déterminisme génétique de la réponse immunitaire contre les agents pathogènes est un atout important pour développer une stratégie globale de lutte contre les maladies infectieuses (Pinard-van der Laan, 2002). Les outils de la génomique, notamment le grand nombre de marqueurs génétiques de type SNP (Single Nucleotide Polymorphism) issus de la séquence du génome, renouvellent les possibilités de caractérisation du CMH du poulet. Ce projet contribue donc à l'amélioration des connaissances sur le système immunitaire du poulet.

En premier lieu, la corrélation entre le typage sérologique traditionnel du CMH, par hémagglutination, et le typage moléculaire, établie sur des populations de référence par Fulton *et al.* (2006), a été examinée dans des lignées expérimentales INRA. Le typage moléculaire présenté vise à caractériser la région B du CMH du Poulet, génétiquement indépendante de la région Y. C'est la région B qui a été essentiellement impliquée dans des phénomènes de résistance/susceptibilité aux maladies infectieuses chez le Poulet.

L'architecture génétique de la région B du CMH a ensuite pu être explorée dans différents types de populations de poulets pour lesquelles de l'ADN est disponible, en l'absence de sérologie. Il s'agit en particulier de races locales, qui peuvent présenter des caractères intéressants d'adaptation au milieu, mais pour lesquelles le recueil de données expérimentales est très difficile.

## **Matériel et Méthodes**

### ***Animaux***

Les lignées Réponse Immune (R.I.), White Leghorn, de l'UMR GDA, Jouy-en-Josas, INRA, maintenues à Nouzilly, INRA, ont été sélectionnées depuis 13 générations sur des critères relatifs à la réponse immunitaire : lignée 1, réponse anticorps après vaccination contre le virus de la maladie de Newcastle; lignée 2, hypersensibilité retardée en réponse à la phytohémagglutinine; lignée 3, clairance du carbone comme indicateur de la phagocytose ; lignée 4, lignée témoin (Minozzi *et al.*, 2008).

Un référentiel de 48 populations a été défini pour l'exploration du CMH. Il inclut 6 lignées commerciales (4 lignées poulets de chair et 2 lignées pondeuses œufs bruns), 17 populations White Leghorn (issues de 15 lignées expérimentales dont les lignées R.I. et d'une lignée commerciale pondeuse) et 25 autres races dont 12 races européennes, 7 races locales asiatiques et 6 populations locales africaines. Chaque population est représentée par 2 animaux, un mâle et une femelle quand cette information était disponible.

### ***Marqueurs moléculaires***

Les fondateurs des lignées R.I. ont été typés en sérologie. La sérologie par hémagglutination caractérise les protéines exprimées à la surface des cellules sanguines, essentiellement les gènes *B-G* (sans équivalent chez les Mammifères) ainsi que les gènes *B-F* (gènes de classe I). Ces animaux ont ensuite été génotypés avec le marqueur LEI0258 comme décrit par Fulton *et al.* (2006), afin de déterminer la corrélation entre ce marqueur et la sérologie dans ces lignées. Le marqueur LEI0258 est un marqueur de type Variable Number of Tandem Repeat, complexe, composé principalement de deux motifs répétés consécutifs de 12 et 13 pb. Les 96 animaux du panel défini pour l'exploration du CMH ont été également génotypés avec le marqueur LEI0258 sur séquenceur automatique.

## Régions séquencées

Des régions conservées le long de la séquence du CMH ont été identifiées afin de concevoir des amorces permettant d'amplifier 24 fragments correspondant à 11 gènes cibles (Tableau 1, Figure 1) : *CD1A1*, *TAP2*, *TAP1*, *BMB1*, *BMB2*, *BMA*, *Tapasin*, *Blec1*, *B-BTN2*, *TRIM41*, *TRIM27.1*, chez les 96 animaux issus du panel préalablement défini.

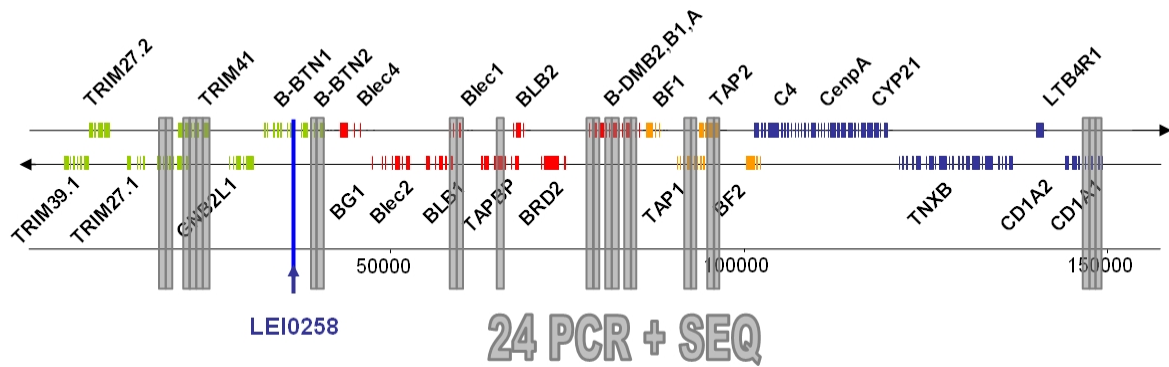
Les gènes étudiés ont été choisis en partie grâce aux données déjà disponibles sur leur polymorphisme (Shiina *et al.*, 2007) mais également en fonction de la faisabilité technique.

Les 24 fragments, de 306 pb à 1249 pb, ont été séquencés au Centre National de Séquençage (CNS).

**Tableau 1.** Amorces utilisées pour amplifier les 24 fragments PCR

| Gene            | Amorce | Séquence                  | Position extrême |
|-----------------|--------|---------------------------|------------------|
| <i>TRIM27.1</i> | AE01F  | GACAGATCTTTAGGGGAGATCTAAA | 93756            |
|                 | AE01R1 | TCAATTGCAAGAGCCAGGTA      | 94845            |
|                 | AG01F  | GCTCTTGCAATTGAAACATCAG    | 94832            |
|                 | AG01R  | CCAAGAAGGAAGGGAAGAGG      | 95804            |
| <i>TRIM41</i>   | AG02F  | GCCTGGACTACTTCGTGGAG      | 99098            |
|                 | AG02R  | TCCTCCTCCATCACATCCTC      | 100000           |
|                 | AG03F  | CTTTGGGGACGATGATTATGA     | 99959            |
|                 | AG03R  | TCCTGCTCATTGCACTTCTG      | 100910           |
|                 | AG04F  | GGAGCCACTGAAGAAGAAGC      | 100854           |
|                 | AG04R  | CCGTCCTGCTTGCTCTTATC      | 101643           |
|                 | AG05F  | CTCCACCAGTTCCTCATTGG      | 101477           |
|                 | AG05R  | GTATTTCTTCCCCACGTCCA      | 102725           |
|                 | AG06F  | TGAGAACATCAAATTCCAGGAG    | 102668           |
|                 | AG06R  | TGATACGAGTGCCCTTAGCC      | 103670           |
| <i>B-BTN2</i>   | AA01F  | CTGCTGCTTCCAGAGCAAAG      | 124464           |
|                 | AA01R  | AGGATGCCGAGGAGAGAAAAT     | 125324           |
|                 | AA02F  | TGTCAGAATTCCGTCAGCTC      | 125177           |
|                 | AA02R  | CCTACAGATGCTCCCTCTTCC     | 126298           |
| <i>Blec1</i>    | AB01F  | TGGAGGAAGTGAGGGAATATC     | 141450           |
|                 | AB01R  | CTGTCAAACACAGCCAAGGA      | 142627           |
|                 | AB02F  | GTGTCCCTTCGACTGGATTG      | 142505           |
|                 | AB02R  | TCAGGTGCTGATGCAGAAGT      | 143522           |
| <i>TAPBP</i>    | AG07F  | GCATCATCCCATAGTTCACC      | 147700           |
|                 | AG07R  | CAGCATCCCAATACCTCCAC      | 147210           |
| <i>B-DMA</i>    | AD01F  | GCCATTGGCTAAACTGGTTG      | 163239           |
|                 | AD01R  | TTGGCTTCAGGCATTGGGTTT     | 164059           |
|                 | AD02F  | ATGGACATCCCAACTGATGC      | 164511           |
|                 | AD02R  | GGATCAACTTATGGGGGACA      | 165379           |
| <i>B-DMB1</i>   | AF02F  | AAGGGGTTTCAGGGTGCT        | 165629           |
|                 | AF02R  | TGTTTGCAATGCTCCACTTG      | 166454           |
|                 | AF03F  | ATCCCCAGCTCTGACCATC       | 166934           |
|                 | AF03R  | AGGAGGTACACCCCACTG        | 167508           |

|               |        |                       |        |
|---------------|--------|-----------------------|--------|
| <i>B-DMB2</i> | AI03F  | AATGCAGGTGCCTTCATGGTG | 168939 |
|               | AI03R  | CCATTTGCAATGCTCACTCCT | 169244 |
|               | AF01F  | TCAGCCCCAGGTCCGCATCGT | 170152 |
| <i>TAP1</i>   | AF01R  | GAGAGATGGGGTCACTCAGC  | 171035 |
|               | AI01F  | ACCCTGTCCATCACATCCAT  | 176578 |
|               | AI01R  | ATACCTACCCCGTCCTGTCC  | 176978 |
| <i>TAP2</i>   | AI02F  | GGACAGGACGGGGTAGGTAT  | 176959 |
|               | AI02R  | GGGGACATTGACAGAACAGG  | 177409 |
|               | AH01F  | GGGTCTGGTGGTTACTGGTG  | 180131 |
| <i>CD1A1</i>  | AH01R  | AGAGGTGGATGGGGTTTTGT  | 181026 |
|               | AC01F  | ATCTCCGCACCAAGACTACG  | 239051 |
|               | AC01R1 | GGGACCATCAGGACCTTCTT  | 239830 |
|               | AC02F  | AGCTTTGGCTGACAAGTGGT  | 239717 |
|               | AC02R1 | AAGGTGACTGGGACGTCATC  | 240839 |
|               | AC03F  | CCTGCTGAAATCCCTCACAT  | 240781 |
|               | AC03R1 | CATGCCATCCTCCTTCTTT   | 241242 |



**Figure 1.** Structure du CMH chez le Poulet et localisation du marqueur LEI0258 et des fragments séquencés (échelle en paire de bases, d'après Shiina *et al.*, 2007)

## Résultats et Discussion

Les tailles d'allèles observées pour le marqueur LEI0258 chez les animaux fondateurs des lignées R.I. ont été comparées à la sérologie déjà établie pour ces animaux (B15, B19, B21, B34 et B124). Une forte corrélation a été mise en évidence entre les types de CMH déterminés par les deux techniques et un tableau de correspondance a pu être établi (Tableau 2).

**Tableau 2.** Correspondance entre le typage du CMH en sérologie et le typage moléculaire du CMH, en taille d'allèle du marqueur LEI0258 (en paires de bases) chez les animaux issus des lignées R. I.

| Sérologie | LEI0258 |
|-----------|---------|
| B15       | 261     |
| B21       | 359     |
| B34       | 422     |
| B124      | 489     |
| B19       | 544     |

**Tableau 3.** Tailles d'allèles identifiées pour le marqueur LEI0258 et correspondance hypothétique avec un ou plusieurs sérotypes B du CMH définis par Fulton *et al.*, pour chacune des catégories de populations étudiées par séquençage.

| Taille allèle LEI0258* (pb)      | Sérotipe B du CMH (Fulton <i>et al.</i> 2006)               | Poules locales Afrique | Races locales Asie | Lignées poulets de chair | Races Europe | Pond. œufs bruns | Lignées White Leghorn |
|----------------------------------|---|------------------------|--------------------|--------------------------|--------------|------------------|-----------------------|
| Nombre de populations étudiées : |   | 6                      | 7                  | 4                        | 12           | 2                | 17                    |
| 183                              | B4  |                        |                    |                          |              |                  | X                     |
| 193-194                          | B15.1, B11, B61, B27, BW3                                   | X                      |                    | X                        | X            |                  |                       |
| 205-206                          | B13.2, B17, BW11  | X                      | X                  | X                        | X            |                  | X                     |
| 217                              |   |                        |                    | X                        | X            |                  |                       |
| 235                              |   | X                      |                    |                          |              |                  |                       |
| 247-249                          | B18, B15.2, B22, B73  | X                      | X                  | X                        | X            |                  | X                     |
| 260-261                          | B15, B2, B29  | X                      |                    | X                        | X            |                  | X                     |
| 273-274                          |   |                        | X                  | X                        |              |                  |                       |
| 295-296                          | B5, B11.1,  | X                      | X                  | X                        | X            |                  |                       |
| 307-308                          | B72, B78  | X                      |                    |                          | X            | X                | X                     |
| 309-310                          | B10, B24, B26, B76  |                        | X                  |                          |              | X                | X                     |
| 311                              |   | X                      |                    |                          | X            |                  |                       |
| 319-320                          |   | X                      | X                  |                          |              |                  |                       |
| 321-322                          | B74   | X                      |                    |                          |              |                  |                       |
| 334                              | BW4   | X                      |                    |                          |              |                  |                       |
| 346                              | B14   |                        | X                  |                          | X            |                  |                       |
| 356-359                          | B5.1, B6.1, B21, B75, B130, B131, B201, B1.1, B75, B23, B77 | X                      | X                  | X                        | X            |                  | X                     |
| 369                              | B21.1, BQ, BW1  | X                      |                    | X                        |              |                  |                       |
| 380-382                          | B13.1   | X                      | X                  | X                        | X            |                  |                       |
| 422                              | B34, B62  |                        |                    |                          |              |                  | X                     |
| 443                              | B6  |                        | X                  |                          |              |                  |                       |

|         |                  |   |   |   |
|---------|------------------|---|---|---|
| 485-488 | B12, B12.1, B124 | X |   | X |
| 526     |                  |   | X |   |
| 538-541 | B19              |   |   | X |

\* les tailles d'allèles ont été regroupées en catégories car la technique de typage utilisée à une limite de résolution d'environ 2 pb

Les tailles d'allèles du marqueur LEI0258, observées pour les animaux du panel, sont présentées au Tableau 3. Les allèles identifiés, d'une taille allant de 183 à 541 paires de bases (pb), peuvent être classés dans 24 catégories. Grâce aux données de la littérature (Fulton *et al.*, 2006) et aux corrélations établies dans les lignées R.I., il est possible, dans la plupart des cas, de proposer une correspondance avec un ou plusieurs sérotype B du CMH (défini en sérologie) à partir de l'information donnée par la taille d'allèle du marqueur LEI0258. On remarque l'importante diversité présente en Afrique et en Asie dans les populations locales ainsi que dans les races européennes, en comparaison avec la diversité observée dans les lignées White Leghorn : 10 catégories d'allèles pour 17 populations étudiées. Ce résultat s'explique notamment par la présence de lignées expérimentales congéniques dans le panel de populations White Leghorn. On remarque cependant que ces lignées montrent la plus grande amplitude de taille d'allèle (183 à 541 pb).

Le séquençage des 24 fragments, représentant 11 gènes, *CDIA1*, *TAP2*, *TAP1*, *BMB1*, *BMB2*, *BMA*, *Tapasin*, *Blecl1*, *B-BTN2*, *TRIM41*, *TRIM27.1*, a permis d'explorer la variabilité génomique d'un total de 20025 paires de bases réparties sur une région de 147500 pb (Tableau 4).

**Tableau 4.** Bilan total des SNPs identifiés pour les 11 gènes étudiés et comparaison avec les résultats obtenus par Shiina *et al.* en 2007 lors de la comparaison de la séquence d'un animal Coq de Bankiva et de celle provenant d'un animal B12

| nb pb analysées | nb SNPs | % SNPs | nb pb analysées Shiina <i>et al.</i> | nb SNPs Shiina <i>et al.</i> | % SNPs Shiina <i>et al.</i> | nb SNPs s. | nb SNPs n.s. | nb SNPs s. Shiina <i>et al.</i> | nb SNPs n.s. Shiina <i>et al.</i> |
|-----------------|---------|--------|--------------------------------------|------------------------------|-----------------------------|------------|--------------|---------------------------------|-----------------------------------|
| 20025           | 593     | 2.96%  | 29219                                | 315                          | 1.08%                       | 152        | 89           | 71                              | 40                                |

nb, nombre ; pb, paires de bases ; s., synonymes ; n.s., non-synonymes

De nombreuses différences d'organisation génomique ont été observées, des insertions, délétions et SNPs. Au total, 593 SNPs ont été caractérisés, ce qui correspond à un taux de polymorphisme de 2,96 % (2,96 SNPs pour 100 pb). Mais la distribution des SNPs n'est pas uniforme, certains gènes présentent un plus fort pourcentage, en particulier le gène *TAP1* (5,52%, Tableau 5). Ce gène est un gène de classe 1 non-classique qui code pour une protéine

intervenant au niveau du chargement des peptides antigéniques par les molécules de classe I, qui a pour conséquence la présentation de ces peptides aux lymphocytes T à la surface de la cellule.

**Tableau 5.** Bilan des SNPs identifiés par séquençage pour les 11 gènes

| Gène            | nb fragments | nb pb analysées | nb seq analysées | nb SNPs    | % SNPs       | nb SNPs ds codant | nb SNPs s. | Nb SNPs non s. | ratio s/n.s |
|-----------------|--------------|-----------------|------------------|------------|--------------|-------------------|------------|----------------|-------------|
| <i>TRIM27.1</i> | 2            | 2063            | 275              | 75         | 3.64%        | 24                | 17         | 7              | 2.43        |
| <i>TRIM41</i>   | 5            | 4897            | 782              | 77         | 1.57%        | 17                | 14         | 3              | 4.67        |
| <i>B-BTN2</i>   | 2            | 1983            | 320              | 69         | 3.48%        | 26                | 10         | 16             | 0.63        |
| <i>Blec1</i>    | 2            | 2197            | 315              | 78         | 3.55%        | 10                | 5          | 5              | 1.00        |
| <i>TAPBP</i>    | 1            | 491             | 159              | 27         | 5.50%        | 23                | 13         | 10             | 1.30        |
| <i>BMA</i>      | 2            | 1690            | 329              | 54         | 3.20%        | 26                | 19         | 7              | 2.71        |
| <i>BMB2</i>     | 2            | 1401            | 321              | 54         | 3.85%        | 34                | 19         | 15             | 1.27        |
| <i>BMB1</i>     | 2            | 1190            | 320              | 39         | 3.28%        | 26                | 20         | 6              | 3.33        |
| <i>TAPI</i>     | 2            | 852             | 369              | 47         | 5.52%        | 12                | 8          | 4              | 2.00        |
| <i>TAP2</i>     | 1            | 896             | 155              | 42         | 4.69%        | 28                | 20         | 8              | 2.50        |
| <i>CD1A1</i>    | 3            | 2365            | 504              | 32         | 1.35%        | 16                | 8          | 8              | 1.00        |
| <b>Total</b>    | <b>24</b>    | <b>20025</b>    | <b>3849</b>      | <b>593</b> | <b>2.96%</b> | <b>241</b>        | <b>152</b> | <b>89</b>      |             |

nb, nombre ; pb, paires de bases ; seq, séquences ; s., synonymes ; n.s., non-synonymes

Une partie des régions codantes de ces 11 gènes a donc été séquencée permettant l'identification de 241 SNPs. Ces SNPs peuvent alors être répartis entre SNPs synonymes (changement d'acide nucléique sans conséquence sur la séquence d'acides aminés de la protéine) ou non-synonymes.

Certains gènes présentent un grand nombre relatif de SNPs non-synonymes, par exemple *Blec1* (Tableau 5). La fonction du gène *Blec1* n'a pas encore été caractérisée mais il code pour une protéine appartenant à la famille des lectines de type C. D'autres gènes, comme *BMB1*, présentent relativement peu de mutations non-synonymes (Tableau 5). *BMB1* est un gène de classe II non-classique. Chez les Mammifères, les molécules codées par ces gènes interviennent au niveau du chargement des peptides antigéniques par les molécules de classe II.

Ces résultats ont pu être comparés aux résultats publiés par Shiina *et al* (2007). Dans cette étude, les auteurs ont séquencé la région du CMH d'un animal d'une lignée expérimentale Coq de Bankiva et l'ont comparé à une séquence déjà publiée du CMH provenant d'un animal de sérologie B12 (Kaufman *et al.* 1999). Grâce au grand nombre de séquences analysées et à la diversité des populations choisies pour notre panel, nous avons pu mettre en évidence 1,9



fois plus de SNPs (Tableau 4). Nous avons également pu mettre en évidence 2 fois plus de SNPs dans des régions codantes (241 vs 111). Mais, proportionnellement, le nombre de SNPs synonymes vs non-synonymes est globalement cohérent entre les deux études (152 et 89 vs 71 et 40, respectivement).

## **CONCLUSION**

Cette étude présente les résultats préliminaires obtenus lors du séquençage d'un nombre important de portions de gènes de la région B du CMH du Poulet dans un large panel de populations d'origines très variées. De nombreux nouveaux SNPs ont été identifiés et en parallèle les 96 animaux du panel de populations choisies pour l'étude ont été typés au niveau du CMH grâce au marqueur moléculaire LEI0258. Les informations de SNP et du marqueur LEI0258 étant maintenant disponibles pour chacun des 96 animaux du panel, leur corrélation le long de la séquence et donc la reconstruction d'haplotypes va permettre de mieux comprendre la structure de la région et déterminer comment le typage avec le marqueur LEI0258 doit être complété afin de mieux caractériser la région B du CMH chez les animaux. Cette étude permettra également de rappeler l'importante diversité génétique présente dans les races locales et la ressource potentielle que représentent ces races locales, étudiées ici pour un complexe de gènes impliqués dans des mécanismes de résistance à des maladies infectieuses chez le Poulet, dont les conséquences économiques ne sont plus à démontrer.

## REFERENCES BIBLIOGRAPHIQUES

Fulton JE, Juul-Madsen HR, Ashwell CM, McCarron AM, Arthur JA, O'Sullivan NP, Taylor RL Jr (2006) Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* **58**:407-421.

Kaufman J, Milne S, Göbel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S (1999) The chicken B locus is a minimal essential major histocompatibility complex. *Nature* **28**:923-925.

Minozzi G, Parmentier HK, Mignon-Grasteau S, Nieuwland MG, Bed'hom B, Gourichon D, Minvielle F and Pinard-van der Laan MH (2008) Correlated effects of selection for immunity in White Leghorn chicken lines on natural antibodies and specific antibody responses to KLH and *M. butyricum*. *BMC Genetics* **9**:5.

Pinard-van der Laan M-H (2002) Immune modulation: the genetic approach. *Vet Immunol Immunopathol* **87**:199–205.

Shiina T, Briles WE, Goto RM, Hosomichi K, Yanagiya K, Shimizu S, Inoko H, Miller MM (2007) Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J Immunol* **178**:7162-7172.

### C. Résultats complémentaires et Discussion

L'objectif principal de ce reséquençage était d'identifier de nombreux polymorphismes dans la région B du CMH du poulet, et la stratégie qui a été adoptée a été de privilégier le nombre de populations étudiées par rapport au nombre d'individus étudiés intra-population. Pour les mêmes raisons, les gènes choisis étaient ceux présentant *a priori* des caractéristiques intéressantes en terme de polymorphisme et rapport dS/dN, la référence disponible au moment du choix étant Shiina *et al.*, 2007. Cependant les gènes de classe I et classe II ont été mis de côté pour des raisons techniques malgré leur intérêt évident. En effet chez le poulet deux gènes de classe I et deux gènes de classe II se trouve dans la zone centrale de la région B, avec des localisations proches bien qu'en sens de transcription inverse. Les différentes amorces de PCR conçues pour ces quatre gènes n'ont jamais passé les tests de spécificité appliqués *in silico*, en terme de nombre de correspondances possibles dans la région ou le génome en BLAST et en terme de nombre de fragments amplifiés en ePCR (PCR virtuelle). De plus, des contraintes techniques étaient imposées par le fait de réaliser et séquencer ces PCR et ces séquences au Centre National de Séquençage (CNS), pour des raisons de coûts. En effet toutes les PCR n'étaient réalisées qu'une seule fois pour tous les animaux, à la même température d'hybridation pour toutes les amorces. De même le séquençage n'était réalisé qu'une seule fois dans chaque sens. C'est également pour cette raison qu'aucun animal d'origine sauvage n'a été choisi pour le panel de population en raison d'une probabilité plus grande de variations au niveau des séquences correspondant aux amorces de PCR.

Les caractéristiques observées pour les onze gènes choisis sont cohérentes avec ce qui avait été publié par Shiina *et al.*, en 2007 ainsi que ce qui avait été trouvé par Hosomichi *et al.*, en 2008. Ainsi, il est maintenant possible de comparer les caractéristiques de ces gènes, calculées globalement à partir des polymorphismes observés lors du reséquençage de 96 animaux, les caractéristiques observées en comparant la séquence de la région (242 kb) pour un animal Coq de Bankiva et pour un animal défini en sérologie comme B12 (Shiina *et al.*) ainsi que les valeurs publiées après reséquençage pour 14 animaux correspondant à des sérologies de référence (B2, B5, B6, B8, B9, B11, B12, B13, B15, B17, B19, B21, B23, B24) mais pour une région plus restreinte (57 kb, Hosomichi *et al.*) (Tableau 5).

**Tableau 5.** Comparaison des caractéristiques en terme de polymorphismes, des 24 gènes composant la zone centrale de la région B du CMH du poulet

| Gène            | nbS | nbN | nbS/nbN | dS/dN | nbS | nbN | nbS/nbN | dS/dN | nbS | nbN | nbS/nbN | dS/dN |
|-----------------|-----|-----|---------|-------|-----|-----|---------|-------|-----|-----|---------|-------|
| <i>TRIM39.1</i> | -   | -   | -       | -     | 3   | 1   | 3.00    | 7.44  | -   | -   | -       | -     |
| <i>TRIM27.1</i> | 17  | 7   | 2.43    | 8.82  | 8   | 3   | 2.67    | 6.64  | -   | -   | -       | -     |
| <i>TRIM41</i>   | 15  | 3   | 5.00    | 18.19 | 2   | 0   | -       | -     | -   | -   | -       | -     |
| <i>GNB2L1</i>   | -   | -   | -       | -     | 2   | 0   | -       | -     | -   | -   | -       | -     |
| <i>B-BTN1</i>   | -   | -   | -       | -     | 1   | 2   | 0.50    | 1.64  | -   | -   | -       | -     |
| <i>B-BTN2</i>   | 9   | 17  | 0.53    | 1.88  | 6   | 4   | 1.50    | 0.49  | -   | -   | -       | -     |
| <i>B-G1</i>     | -   | -   | -       | -     | 9   | 11  | 0.82    | 2.28  | 16  | 40  | 0.40    | 1.14  |
| <i>Blec2</i>    | -   | -   | -       | -     | 0   | 15  | 0.00    | -     | 1   | 35  | 0.03    | 0.14  |
| <i>Blec1</i>    | 5   | 5   | 1.00    | 3.79  | 4   | 1   | 4.00    | 10.24 | 5   | 9   | 0.56    | 3.21  |
| <i>B-LB1</i>    | -   | -   | -       | -     | 6   | 19  | 0.32    | 0.36  | 15  | 68  | 0.22    | 0.71  |
| <i>TAPBP</i>    | 13  | 10  | 1.42    | 4.20  | 7   | 7   | 1.00    | 2.18  | 25  | 19  | 1.32    | 3.63  |
| <i>B-LB2</i>    | -   | -   | -       | -     | 4   | 17  | 0.24    | 0.30  | 8   | 59  | 0.14    | 0.47  |
| <i>BRD2</i>     | -   | -   | -       | -     | 14  | 2   | 7.00    | 16.85 | 43  | 4   | 10.75   | 31.63 |
| <i>B-DMA</i>    | 18  | 7   | 2.57    | 7.75  | 3   | 0   | -       | -     | 17  | 1   | 17.00   | 94.50 |
| <i>B-DMB1</i>   | 18  | 15  | 1.20    | 4.29  | 2   | 3   | 0.67    | 0.80  | 12  | 8   | 1.50    | 3.95  |
| <i>B-DMB2</i>   | 20  | 7   | 2.86    | 9.49  | 5   | 2   | 2.50    | 2.55  | 17  | 5   | 3.40    | 14.13 |
| <i>B-F1</i>     | -   | -   | -       | -     | 11  | 25  | 0.44    | 1.10  | 37  | 86  | 0.43    | 1.31  |
| <i>TAP1</i>     | 8   | 4   | 2.00    | 6.40  | 21  | 7   | 3.00    | 6.53  | 52  | 23  | 2.26    | 6.84  |
| <i>TAP2</i>     | 20  | 8   | 2.50    | 6.63  | 14  | 11  | 1.27    | 2.70  | 57  | 31  | 1.84    | 3.38  |
| <i>B-F2</i>     | -   | -   | -       | -     | 20  | 34  | 0.59    | 1.43  | 29  | 99  | 0.29    | 0.79  |
| <i>C4</i>       | -   | -   | -       | -     | 17  | 14  | 1.21    | 3.03  | -   | -   | -       | -     |
| <i>CenpA</i>    | -   | -   | -       | -     | 2   | 0   | -       | -     | -   | -   | -       | -     |
| <i>CDIA2</i>    | -   | -   | -       | -     | 0   | 0   | -       | -     | -   | -   | -       | -     |
| <i>CDIA1</i>    | 6   | 9   | 0.67    | 2.05  | 1   | 2   | 0.50    | 1.15  | -   | -   | -       | -     |

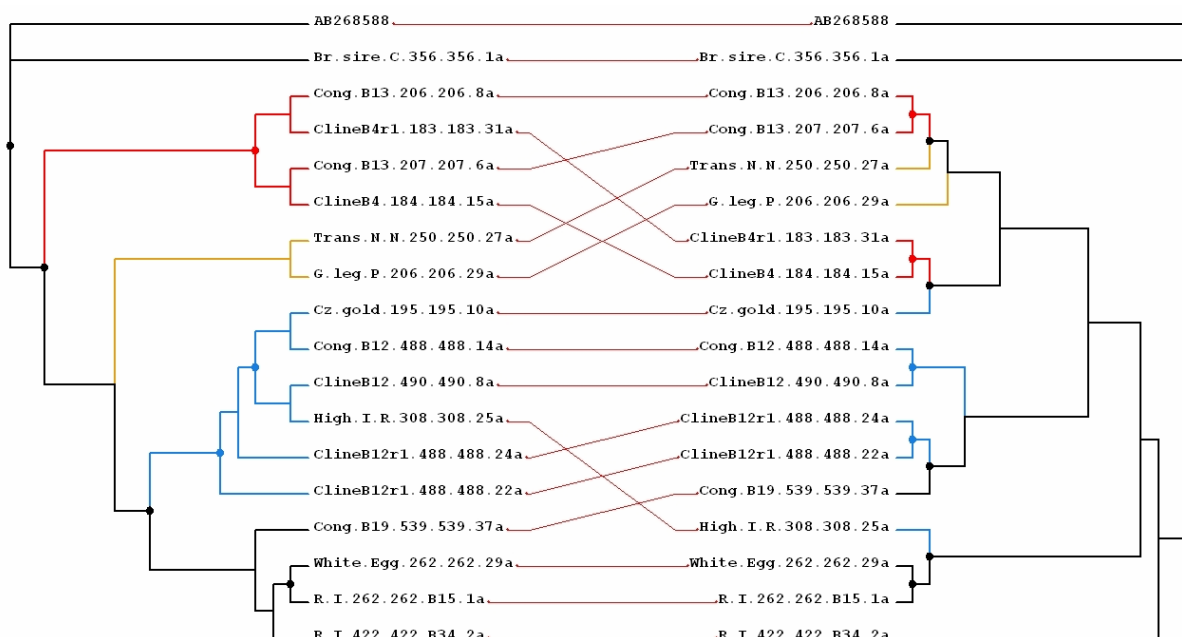
-----Shiina *et al.*, 2007-----

-----Hosomichi *et al.*, 2008-----

Les différentes valeurs observées sont globalement cohérentes et sachant que la comparaison du nombre de polymorphismes de séquences synonymes et non-synonymes est un indicateur des pressions de sélection subies par les gènes, ce type de résultats permet de mieux caractériser cette région centrale du CMH. Il apparaît ainsi que ce n'est pas la région dans sa globalité qui présente un polymorphisme mais plutôt certains gènes, et notamment les gènes de classe I et classe II classiques, respectivement *B-F1* et *B-F2*, *B-LB1* et *B-LB2*.

D'autre part, cette étude basée sur un reséquençage relativement important fournit également des données sur les animaux et les populations choisis. Globalement chaque PCR a été séquencée avec succès, avec en moyenne 90% des animaux séquencés. Dans l'ensemble ce sont les animaux hétérozygotes des races locales asiatiques et africaines qui ont présenté les plus hauts taux d'échec. L'analyse des génotypes individuels n'a pas été réalisée pour tous les gènes, certains gènes ont été traités en priorité. Il s'agit des trois gènes *B-DM*, à la localisation centrale, dans la région de classe II, et qui font l'objet d'un volet particulier de cette thèse. Il s'agit également des gènes *TRIM27.1* et *CDIA1* qui ont été analysés en raison de leur

localisation à chaque extrémité de la zone centrale de la région B du CMH du poulet. Alors que pour l'étude des gènes B-DM, de nombreuses séquences complémentaires ont été obtenues par la suite, l'analyse des gènes *TRIM27.1* et *CDIA1* s'est limitée aux séquences obtenues lors du reséquençage global au CNS. Ainsi les animaux incomplets pour l'un des fragments PCR correspondant aux gènes, en l'occurrence, deux fragments pour *TRIM27.1* et trois fragments pour *CDIA1*, n'ont pas pu être utilisés. Pour chaque gène, tous les animaux « complets » ont été caractérisés au niveau d'un maximum de SNPs, au nombre de 71 pour *TRIM27.1* et de 29 pour *CDIA1*. Les haplotypes ont ensuite pu être reconstruits pour chaque gène séparément, en utilisant le logiciel PHASE sous l'hypothèse d'un taux de recombinaison égal à zéro, cohérent avec la longueur des zones étudiées, 1.8 kb pour *TRIM27.1* et 2 kb pour *CDIA1* (les méthodes utilisées pour l'analyse de ces gènes sont les mêmes que celles utilisées pour l'analyse des gènes *B-DM* qui font l'objet d'une description plus détaillée dans la partie suivante de cette thèse). Au final, 58 animaux ont pu être caractérisés au niveau des deux gènes. L'un des objectifs de cette étude était de rechercher des traces de recombinaison dans la région B, le type d'analyse choisie a donc été la comparaison de dendrogrammes. Cependant, celle-ci s'est révélée trop complexe pour les animaux hétérozygotes, présentant deux haplotypes possibles à chaque gène. Au final, elle n'a été appliquée pour 18 animaux homozygotes, qui étaient donc majoritairement des animaux issus de lignées expérimentales à la sérologie connue.



**Figure 10.** Comparaison des dendrogrammes obtenus en Neighbour Joining pour le gène *CD1A1* (à gauche) et le gène *TRIM27.1* (à droite) situés aux extrémités de la zone centrale de la région B du CMH du poulet

Les résultats obtenus (Figure 10) indiquent que le très bas taux de recombinaison généralement admis pour cette région semble fondé, bien que certains traces de recombinaisons sont mises en évidence en particulier pour deux animaux provenant de races locales, Transylvanian Naked Neck (de génotype 250/250 en terme de pb au marqueur LEI0258) et Green-legged Partridge (206/206), un individu provenant d'une lignée expérimentale appelée High Immune response (308/308) et un individu provenant d'une lignée expérimentale congénique définie comme B19 en sérologie qui, d'une façon intéressante, est similaire aux animaux définis comme B12 en sérologie pour le gène *TRIM27.1* mais pas pour le gène *CD1A1*. Cette différence est particulièrement intéressante au regard de la grande similarité observée pour toute la région B entre les séquences des sérotypes B12 et B19.

En conclusion, la stratégie choisie, c'est à dire le reséquençage de gènes cibles en moyen débit pour un grand nombre de populations, races ou lignées a été un succès en terme d'identification de polymorphismes SNPs exploitables. Même si certains gènes n'ont pas pu être étudiés, notamment pour des critères techniques, il est probable que, au vu du fort déséquilibre de liaison caractérisant la région, les indications données par les autres gènes

soient, dans une certaine mesure, suffisantes. De plus les différents SNPs caractérisés sont en cours de valorisation sous la forme d'une puce dédiée à la région B du CMH du poulet, décrite dans la cinquième partie de ce manuscrit.

Cependant la stratégie d'échantillonnage en elle-même, le grand nombre d'animaux hétérozygotes et le petit nombre d'animaux par origine n'ont pas favorisé l'exploitation des données au niveau individuel pour tous les gènes. En effet du séquençage complémentaire ainsi que des analyses plus poussées ont été nécessaires et n'ont pu être réalisés que pour certains gènes.

## **IV QUATRIEME PARTIE : CARACTERISATION DE LA STRUCTURE GENOMIQUES DES TROIS GENES PARTICULIERS DU CMH DU POULET CORRESPONDANT AUX GENES *DM* DES MAMMIFERES ET ETUDE DE LEUR POLYMORPHISMES DANS TROIS ESPECES DE *GALLUS* SAUVAGES ET UN LARGE PANEL DE RACES EXPERIMENTALES, COMMERCIALES ET LOCALES**

### **A. Introduction**

Chez les Mammifères, les molécules DM sont codées par des gènes du complexe majeur d'histocompatibilité (CMH) et sont impliquées dans des fonctions clé dans la voie de présentation des antigènes exogènes, la voie des molécules de classe II classiques.

Actuellement, chez le poulet, pas moins de quatre classiques gènes de classe II ont été décrits, deux situés dans la région B codant pour des chaînes  $\beta$  (*B-LB1* et *B-LB2*), au moins un dans la région Y, mais également un ADNc, codant pour une chaîne  $\alpha$ . Cet ADNc a été localisé sur le chromosome 16, à proximité de la région B. Alors que chez le poulet, l'existence de gènes de type DM était admise depuis les premières annotations de la séquence de la région, ils n'avaient pas été caractérisés et leur nombre n'était pas toujours précisément rapporté. Par exemple seulement deux gènes *DM* sont prédits par Ensembl.

Ce travail a donc consisté en la description précise des trois gènes de la région B du CMH du poulet, similaires aux gènes *DM* des Mammifères, *B-DMA*, *B-DMB1* et *B-DMB2*. Ils codent pour une chaîne de classe II alpha DM et deux chaînes bêta qui présentent à la fois des caractéristiques des molécules de classe II classiques de poulet mais également des particularités typiques des protéines DM des Mammifères. Afin d'étudier ces similitudes, les différentes séquences nucléotidiques, protéiques ainsi que l'expression des trois gènes ont été prises en compte.

En parallèle, le polymorphisme des trois gènes, évalué par reséquençage pour 96 animaux, a fait l'objet de travaux complémentaires et notamment du séquençage d'animaux sauvages des trois espèces suivantes : Coq de Bankiva (*Gallus gallus*), Coq de Lafayette (*Gallus lafayetii*) et Coq de Sonnerat (*Gallus sonneratii*). Ainsi, nous avons pu reconstituer 124 haplotypes correspondant à une région de 8.8 kb chez 146 individus représentant un large éventail de lignées commerciales et expérimentales, de races locales d'Europe, d'Asie et d'Afrique et trois



espèces sauvages. Nous avons également identifié une répétition dans le second intron de *B-DMA*, rendant possible la conception d'un nouveau marqueur pour le CMH de poule, situé dans la région de classe II.

**Cet article doit être soumis à *Molecular Immunology*.**

## B. Article de recherche

### Class II *B-DM* in the *B* region of the chicken *Mhc*

Olympe Chazara<sup>a,\*</sup>, Véronique Morin<sup>b</sup>, Michèle Tixier-Boichard<sup>a</sup>, Rima Zoorob<sup>b</sup>, Bertrand Bed'Hom<sup>a</sup>

<sup>a</sup>INRA, UMR 1313, Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

<sup>b</sup>CNRS, FRE 2937, Laboratoire Génétique Moléculaire et Intégration des Fonctions Cellulaires, Villejuif, France

\* Corresponding author at: GABI, INRA, Domaine de Vilvert, 78352 Jouy en Josas, France.

Tel.: +33 1 34 65 26 50; fax: 33 1 34 65 29 33.

E-mail address: olympe.chazara@jouy.inra.fr (O. Chazara).

#### ABSTRACT

In Mammals, the DM molecules are encoded by the major histocompatibility complex (*Mhc*) and execute key functions in the class II antigen presentation pathway. Here, we characterized three *DM* genes in the *Mhc B* region of the chicken (*Gallus gallus*): *B-DMA*, *B-DMB1* and *B-DMB2*. They encode one class II DM alpha-chain and two beta chains, exhibiting motifs of chicken's class II antigens as well as specificities of Mammals' DM proteins. We also studied the expression pattern of those three chicken *B-DM* genes; they are expressed in immune related tissues. This is the first description of the genomic sequence of a class II alpha gene in chicken and a valuable description of *DM* genes in a non-mammalian Vertebrate, reinforcing the hypothesis of the existence of the *DM* genes in the primordial *Mhc*, as suggested by previous studies in Mammals. We were able to reconstruct 124 haplotypes corresponding to the 8.8 kb *B-DM* region, in accordance with the 212 SNPs identified in 146 individuals representing a wide range of experimental, commercial, local breeds from Europe, Asia and Africa, and three wild species of fowl. We also discovered a repeat inside *B-DMA* second intron, making possible the design and the typing of a new marker for the chicken *Mhc*, linked to the class II region. Therefore this study not only describes three *DM* genes in the chicken, it also provides an overview of the *Mhc* diversity in chicken.

*Keywords:* Major Histocompatibility Complex; *Gallus gallus*; *DM*; class II; diversity

## 1. Introduction

The *DM* loci, first described in 1991 by Cho *et al.*, in mouse, and Kelly *et al.*, in human, are non-classical *Mhc* class II genes. In chicken, the *Mhc* is located on microchromosome 16, and is organized in two genetically independent gene complexes, *B* and *Y*. A number of *B* serotype, distinguished by classical serological methods (Briles & Briles, 1987) or recent molecular genotyping techniques (Fulton *et al.*, 2006) determines strong differences in resistance or susceptibility to infectious diseases such as Marek's disease. The classical class II molecules of the chicken (BL) are, as for Mammals, composed of two chains,  $\alpha$  and  $\beta$ , generally expressed by B cells and macrophages. They present peptides, principally derived from exogenous molecules, to CD4 + T lymphocytes. In chicken, no less than four classical class II genes have been so far described, two encoding  $\beta$  chains in the *B* region and at least one in the *Y*, plus a cDNA, encoding an  $\alpha$  chain. This cDNA has been described but not mapped in the genomic sequence of chromosome 16, which is only partial (Salomonsen *et al.*, 2003).

In Mammals, the DM molecules are involved in the formation of *Mhc*-II/peptide complexes, before their expression at the surface of the cell. It has been shown that without DM, these complexes are unstable, disrupting antigen presentation (Morris *et al.*, 1994). In fact, class II molecules are assembled in the endoplasmic reticulum and stabilized by a third protein, the invariant chain. In endosomes, the invariant chain is cleaved by proteases, and a peptide from the invariant chain, CLIP, remains attached to the binding site of *Mhc*-II molecules. In the presence of DM, CLIP is cleaved and replaced by antigenic peptides, which are derived from proteolysis of exogenous antigens (Sloan *et al.*, 1995). DM also acts as a chaperone protein for *Mhc*-II molecules by protecting them from denaturation in acidic compartments. In the end, the molecule DM controls the loading of antigenic peptides by inducing the dissociation of *Mhc*-II/peptide complex of low stability. Thus, only *Mhc*-II/peptide complexes of high stability can be presented on the surface of presenting cells (Kropshofer *et al.*, 1997a, 1997b and 1999).

The assembly of the chicken genome, and the extended recent map of the *Mhc* of the chicken (Shiina *et al.*, 2007) have, in the *B* class II region, a segment for which *in silico* gene prediction and comparisons with Mammals indicates the probable presence of two or three genes orthologous to genes characterized as *DM* in Mammals. Ever since, it has been assumed that the *Mhc B* region of the chicken encoded non-classical class II genes, without further

studies, whereas the intracellular localization and oligomerization of the invariant chain with class II molecules has been recently demonstrated in chicken (Ye *et al.*, 2009).

This study was initiated as a step towards improving the knowledge of the gene content of the *B* region, in order to determine the individual contribution of genes to diseases incidence. In this context, we characterized in details the molecular organization and structure of these predicted non-classical class II genes, as to determine the genetic organization, the protein sequences and the pattern of expression of those putative *DM* genes in chicken. As shown below, we demonstrate that three *DM* genes do exist in the *B* region of the *Mhc* of the chicken. But, in order to assess genetic diversity of these key immune system genes, we also characterized for these genes, a large number of SNPs, and a new marker, from a wide range of experimental, commercial, local breeds and wild related species.

## **2. Material et Methods**

### **2.1. The cDNA library and screening**

A set of *B-DM* cDNA were isolated from a serotype B13 spleen cDNA library in the pCDM8 vector (Bernot & Auffray 1991). A clone (C.4.5.6), with high similarities to mouse and human *DMB* gene, had been previously identified by systematic sequencing of Chicken *Mhc* genomic libraries (constructed by subcloning in Puc 19 plasmid vector the *Mhc* cosmid clone C.4.5). The C.4.5.6 fragment was amplified using the GeneAmp PCR (Perkin Elmer) kit, in a total volume of 20  $\mu$ l with 200 ng of DNA from cosmid subclone C.4.5.6, 2  $\mu$ l of 10X PCR buffer (10 mM Tris HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 2  $\mu$ l each of 10 mM dNTP, 0.5  $\mu$ M of each primer (M13 universal) and 1 U of Taq polymerase. The cycling conditions were an initial denaturation step at 94°C for 2 min, followed by 30 cycles at 94°C for 1 min, 60°C for 30 s and 72°C for 1 min followed by a final extension step at 72°C for 7 min. The fragment obtained, encompassing C.4.5.6, was then <sup>32</sup>P-labelled using the Multiprime DNA labelling system (Amersham Biosciences), and used as a probe to screen approximately 200,000 clones from the cDNA library. Six clones gave positive hybridisation signals and their nucleotide sequence was determined for both strands by the enzymatic DyeDeoxy chain termination sequencing approach.

### **2.2. Total RNA isolation**

Total RNA was prepared from tissues of White Leghorn chicks (serotype B12). The tissues were: brain, bursa, heart, lung, and spleen. 30 mg of each tissue was disrupted for 4 minutes at 10-14000 rpm with an IKA yellowline D1 18 dispenser. Total RNA was isolated using the RNeasy mini kit from Qiagen. A purification step was added after disruption by pipetting the clear supernatant into a Qias shredder column (Qiagen) and centrifuging at 14,000 rpm for 5 min. Then, the filtrate was used following the RNeasy protocol, including the optional on-column digestion with DNase (Qiagen RNase-Free DNase). The RNA concentration was measured by UV spectrophotometry.

### **2.3. Genomic DNA origin and extraction**

In the study for genetic polymorphism investigation, 146 animals were used from 47 different populations or lines of domestic chickens (*Gallus gallus*), including 29 populations of various origins and selection history from the European AVIANDIV project (Supplementary Table 1). Detailed information on these populations was reported elsewhere (Hillel *et al.*, 2003). 34 wild *Gallus sp.* were also studied: 25 Red Junglefowl (*Gallus gallus*), 5 Grey Junglefowl

(*Gallus sonneratii*) and 4 Sri Lanka Junglefowl (*Gallus lafayetii*). 19 out of the 25 Red Junglefowl were also from the AVIANDIV project.

For the B12, B13, and B19 inbred lines from INRA (Nouzilly, France, Miller *et al.*, 2004), the Silkie chickens from NCHU (Taichung, Taiwan), and the other wild *Gallus sp.* from a zoological park in Clères, France, high molecular weight genomic DNA had been obtained from blood samples, as described by Chang *et al.*, 2006.

For four additional lines from INRA: the Immune Response lines (I.R. lines kept in Nouzilly, France, Minozzi *et al.*, 2007), and six Taiwanese chickens populations (from NCHU, Taichung, Taiwan), crude DNA extractions were prepared from blood by mixing 2 µl of whole blood with 250 µL of lysis buffer (1M pH 8 Tris 1%, 1M MgCl<sub>2</sub> 0.1%, 3M NaCl 5%, 1% Sigma IGEPAL) before precipitation by centrifugation and incubation with 250 µL of NaOH (0.2 M) at 65°C for at least 2 h followed by neutralization with 250 µL of Tris (0.1M) - HCl (0.2 M).

For South Benin, Cameroon, Cote d'Ivoire and Ghana samples, DNA had been obtained from blood using the DNeasy Blood & Tissue Kit from Qiagen at the Department of Animal Science, University of Ghana (Legon, Ghana).

#### **2.4. Molecular Mhc characterisation of the samples**

Every animal has been genotyped at the *Mhc* level by PCR amplification with the marker LEI0258 (McConnell *et al.*, 1999, Fulton *et al.*, 2006), and with a new marker located in the chicken *Mhc*, called GAB0001. Primers for the marker LEI0258 were LEI0258-F: CACGCAGCAGAACTTGGTAAGG forward and LEI0258-R: AGCTGTGCTCAGTCCTCAGTGC reverse. Primers for the marker GABI0001 were designed using the sequence information obtained during this study and the Primer3 software (Rozen & Skaletsky, 2000),: GABI0001-F: CATTCCCCATTAACGCACTC forward and GABI0001-R: AGAGTGTTGGGGTAGCCGAG reverse. Duplexed amplifications were done in 10 µL with 100 ng of DNA, 1X of QIAGEN Multiplex PCR Kit and primer concentrations of 0.1 µM and the following PCR cycle setup: 94°C for 15 min, then 35 cycles at 94°C for 45 s, 60°C for 1 min 30 s, 72°C for 1 min, with a final extension at 72°C for 15 min. After 1:10 dilution, the PCR products were loaded on an Applied 3730XL automated sequencer for electrophoresis and results were analysed using Genemapper v3.7.

In order to get the real size of the marker LEI0258 amplified fragment, more than half (64%) of the alleles were sequenced, directly for homozygotes or after separation of the two alleles on a 4% agarose gel, bands cutting and purification with QIAquick PCR Purification Kit from

Qiagen. Sequencing was carried out by Eurofins MWG Operon, using their standard protocol for purified PCR products.

### **2.5. RT-PCR**

Seven oligonucleotides were used in order to study the expression of the genes of the region (Supplementary Table 2, Fig. 1). They were designed using the chicken *Mhc* genomic sequence reported by Shiina *et al.*, 2007 (GenBank accession number **AB268588**) and verified by the program Primer3 (Rozen & Skaletsky, 2000).

The RT-PCRs were carried out in two steps, following the Applera protocol for the GeneAmp RT-PCR kit. Approximately 1 µg of total RNA was used with 25 mM MgCl<sub>2</sub>, 2 µL 10X PCR buffer, 0.2 mM of each dNTP, 2.5 µM of oligodT (16), 20 U of RNase Inhibitor and 2.5 U / µL of MulV Reverse Transcriptase in a final volume of 20 µL. The amplification step was carried out by adding to 10 µL of the cDNA products of step 1, 12.5 mM MgCl<sub>2</sub>, 4 µL of 10X PCR buffer, 32.75 µL of H<sub>2</sub>O, 100 ng of each oligonucleotide and 1.25 U of AmpliTaq DNA polymerase in order to obtain a final volume of 50 µL. The following programs were run on a thermal Mastercycler Gradient (Eppendorf) or a GeneAmp PCR System 2700 (Applied Biosystems): for the reverse transcription, 15 min at 42°C, 5 min at 99°C, and 5 min at 5°C, and for the PCR amplification, 1 min 45 s at 95°C, then 35 cycles with 15 s at 95°C, 30 s at 60°C, and a final step of 7 min at 72°C. The RT-PCR products were always checked on agarose gel containing ethidium bromide (BET). The gel was exposed to UV light to reveal the presence or absence of amplified fragments and in order to confirm their expected size.

### **2.6. PCR and Sequencing**

In order to resequence seven parts of the region, including the putative coding sequences, 14 primers were designed (Supplementary Table 2, Fig. 1) using Primer 3 (Rozen & Skaletsky, 2000).

First, for 95 domestic chickens, amplifications of PCR fragments numbers 2 to 7, and their sequencing, were carried out at the “Centre National de Séquençage” (CNS, Evry, France) following their standard protocol, at an annealing temperature of 60°C, fragments were sequenced in the forward and reverse directions. Then, for the same 6 PCR fragments, failed samples were reamplified and resequenced in our laboratory. This was done for approximately 16% of the domestic chickens. In the same time, we amplified and sequenced the same 6 fragments for 34 additional wild *Gallus sp.*. An additional PCR was also performed (numbered 1) on all the 129 samples. All those PCRs were conducted in 40 µL

total volumes with 200 ng of DNA, 1X of Qiagen Master Mix and primer concentrations of 0.1  $\mu$ M. Cycling parameters were 94°C for 15 min, then 35 cycles of 94°C for 45 s, 60°C for 1 min 30 s, 72°C for 1 min, with a final extension of 72°C for 15 min. When amplification wasn't working, which happened especially for Asian local breeds, annealing temperature was decreased at 58°C. Sequencing was carried out by Eurofins MWG Operon, using their standard protocol for purified PCR products. PCR products were always checked on agarose gel containing ethidium bromide (BET). The gel was exposed to UV light in order to check the specificity and quality of the fragment obtained.

### **2.7. Identification of EST**

The EST (Expressed Sequence Tag) corresponding to the studied region and already reported have been searched by BLAST (Altschul *et al.*, 1990) on the dbEST section of GenBank database from NCBI, using the genomic sequence of the Chicken *Mhc* reported by Shiina *et al.*, 2007 as query sequence (**AB268588**). Sequences were then aligned against the genomic sequence with the software EST2GENOME from EMBOSS package (Fig. 1).

### **2.8. Sequence and data analysis**

Sequence quality and variation analyses were carried out for the 129 individuals with NovoSNP (Weckx *et al.*, 2005), and Pregap4 v1.5 and Gap4 v4.10 components of the Staden package (Staden *et al.*, 2000). With NovoSNP, reports were obtained of SNPs characteristics and individuals genotypes using the default parameters. Only the SNPS with high probability of being accurate were kept for further analyses (NovoSNP's score above or equal to 12). Then SNPs were manually confirmed and sample genotypes were manually edited by looking back at the chromatogram, using NovoSNP and the Staden contig editor, especially to check consistency of heterozygote genotypes.

Individual genotypes were obtained for the 129 individuals and completed with the 16 sequences available in GenBank (accession numbers **AL023516** from Kaufman *et al.*, 1999, **AB268588** from Shiina *et al.*, 2007 and **AB426141** to **AB426154** from Hosomichi *et al.*, 2008), and the genome sequence. Genotypes files are available from the authors upon request. Those genotypes were then formatted as an input file for PHASE (Stephens *et al.*, 2001) and it was run under the MR4 model with R the recombination rate set at 0, in order to infer haplotypes. Those parameters seemed the more relevant for this small *Mhc* region (8.8 kb). Then a NeighborNet network was constructed, in order to represent the haplotypes, using



Splitstree version 4.10 (Huson & Bryant 2006) with default parameters. This method combines the principle of the neighbor-joining and split decomposition techniques.

### ***2.9. Phylogenetic analysis***

The online tool SeqPHASE (Flot 2009) was used to convert PHASE output files back into fasta. The nucleotide sequences were aligned and translated using MEGA 4 (Tamura *et al.*, 2007). Maximum parsimony phylogeny was constructed from the protein alignments, also with MEGA 4, with default parameters and 1000 bootstrap replications. Consensus trees were obtained with a 65% cut-off value.

### 3. Results

#### 3.1. Genomic organization of *B-DMA*, *B-DMB1* and *B-DMB2*

##### *B-DMA*

The alignment of the putative genomic sequence for *B-DMA*, the sequence of our two cDNAs, and the different EST identified in the GenBank database, defined four exons with an alternative first exon, 65 bp long (Fig. 1). Considering the most frequently reported first exon, exons were respectively 132, 267, 282 and 238 bp long. Between human and chicken, the size of the second and third exons is conserved. On average, introns were shorter, which is usual for genes within the *Mhc* of the chicken but surprisingly the second intron was larger in the chicken, this was due to the presence of 5 to 7 repeats of a 12 (YYCCACTGATGT) or a 14 (YCCCACYRAYGYRT) nucleotides motif. We used this repeat as a genetic marker and following the nomenclature, we named it GAB0001.

The nucleotide sequence, thus defined for *B-DMA*, encoded a peptide of 263 amino acids, encompassing a *Mhc*-II  $\alpha 1$  domain and an IgC1  $\alpha 2$  domain. Five cysteines, whose positions are conserved in all reported mammalian proteins DMA, made possible two disulfide bridges, one in the  $\alpha 1$  domain and one in the  $\alpha 2$  domain. The only potential site of N-glycosylation identified in chicken, however, didn't coincide with those of Mammals (Fig. 2).

##### *B-DMB1* and *B-DMB2*

The alignment of the putative genomic sequence for *B-DMB2*, the sequence of our four cDNAs, and the different EST identified in the GenBank database defined six exons for *B-DMB2*, respectively, 48, 282, 285, 120, 36 and 99 bp long. There seemed to be an alternative splicing of exons 3 and 4 for these genes as observed on two of our cDNA and two ESTs (Fig. 1). This alternative splicing, however, induced a 2 bp frame shift in the translated protein, leading to a premature stop codon.

Only part of two of our cDNA (cDNA 18A and R4bis) and four EST in Genbank corresponded to the putative genomic sequence for *B-DMB1*, none corresponding to the whole coding sequence. They defined five exons being respectively 103, 282, 276, 116 and 340 bp long (Fig. 1).

The coding sequences, thus defined for *B-DMB1* and *B-DMB2*, encoded for peptides of respectively 310 and 258 amino acids, comprising a *Mhc*-II  $\beta 1$  domain and an IgC1  $\beta 2$  domain. Six cysteines, whose positions are conserved in the majority of mammalian DMb

proteins, made possible three disulphide bridges, two in domain  $\beta 1$  and one in domain  $\beta 2$ . Two and three potential N-glycosylation sites were found in B-DMb1 and B-DMb2, respectively. Within the protein B-DMb2 cytoplasmic tail, the Mammals' DMb proteins endosomal targeting signal: YXXL was identified. This signal was not found in B-DMb1 (Fig. 3).

### **3.2. Genomic sequence features**

Upstream of the first exons identified for *B-DMA* and *B-DMB2*, putative promoters were found: X and Y elements, typical of class II promoter region (Van den Elsen *et al.*, 2004). These sites were similar to the promoters described for the *HLA-DMA* and *HLA-DMB* genes (Kelly *et al.*, 1991, Westerheide *et al.*, 1997), human class IIa and b genes, and chicken class IIb gene. The patterns and the distances between them were conserved for the chicken *B-DMA* and *B-DMB2* genes. At the 3' end of the three genes, we identified a stop codon TGA and a polyadenylation signal sequence ATTA AAA (Supplementary Figure 1A, 1B, and 1C). No putative promoter could be found upstream of *B-DMA* alternative first exon.

The GAB0001 marker was typed for the 95 domestic chickens of the study. Genotypes were obtained for 94 individuals (Supplementary Table 1). Five alleles were found: 144 bp, 159 bp, 161 bp, 174 bp, and 187 bp. This marker was easily duplexed with the LEI0258 marker under standard conditions and led to the discrimination of animals sharing a common LEI0258 allele size, such as 261 bp (serology B2/B15) or 357 bp (serology B21/B23) (Supplementary Table 1).

### **3.3. Expression of *B-DMA*, *B-DMB1* and *B-DMB2***

RT-PCR results indicated that mRNAs of *B-DMA*, *B-DMB1*, and *B-DMB2*, are all present in spleen tissues, whereas, no amplification was obtained, for the five tissues, for the primer corresponding to the alternative first exon of *B-DMA*. Moreover, mRNAs of *B-DMA* could also be amplified from brain and bursa tissues, and mRNAs *B-DMB2*, from brain, bursa and lung tissues. No amplification was obtained from heart tissues. The corresponding ESTs identified from the GenBank database provided additional information about the genes expression: ESTs corresponding to *B-DMA* and *B-DMB2* have been identified in various tissues or cells: ovary, lymphocytes, liver, kidney, small intestine, macrophages, brain, lymphoid organs, testis, eye, and T cells, unlike the four EST corresponding to *B-DMB1*, identified from only two tissues: small intestine and brain (Fig. 1).

### **3.4. Phylogeny analysis of the B-DM molecules**

We studied the relationship between the protein domains  $\alpha 1$ ,  $\alpha 2$  or  $\beta 1$ ,  $\beta 2$  of the three B-DM of the chicken, the three putative DM of the quail and the turkey, the two HLA-DM from human, the three H2-DM from the mouse, the two BOLA-DM of the bovine, the two Mod-DM of the opossum, the two putative DM of the frog, and classical class II proteins of the chicken and the same animals. We found similar results whatever the domain analysed (data not shown). We always observed two clusters: a cluster grouping DM proteins and a cluster grouping class II proteins, as shown the consensus parsimonious tree obtained for  $\beta 2$  domains (Fig. 4).

### **3.5. Genetic polymorphism of the B-DM molecules**

Analysis of the 7 PCR fragments obtained for *B-DMA*, *B-DMB1*, and *B-DMB2*, allowed the identification of 212 SNPs, from which 166 SNPs were polymorph in domestic chickens (Supplementary Table 3). We identified 119 SNPs for *B-DMA*, 49 SNPs for *B-DMB1*, and 44 SNPs for *B-DMB2*. Concerning the coding regions, the distribution was: 32 SNPs in *B-DMA*, 31 SNPs in *B-DMB1*, and 30 SNPs in *B-DMB2*. In the alternative first exon of *B-DMA*, we found 3 SNPs in the 65 bp coding region, all non-synonymous. Due to a high rate of PCR failure for the 7<sup>th</sup> PCR fragment, comprising the alternative first exon of *B-DMA*, we didn't use the genotypes of this fragment for the haplotype reconstruction.

The haplotype reconstruction, based on genotypes of the 146 individuals, at 158 SNPs, has led to the definition of 124 different haplotypes for the whole 8.8kb *B-DM* region. The haplotypes were also inferred gene by gene, leading to the definition of 88 haplotypes for *B-DMA*, 69 haplotypes for *B-DMB1*, and 71 haplotypes for *B-DMB2*. The complete haplotypes were combinations of the partial haplotypes, which is consistent with the hypothesis of a null recombination rate for the region. The removal of the wild *Gallus Sp.* animals and their separate analysis for haplotype reconstruction didn't change the results neither for domestic nor for the wild chickens.

The NeighborNet network illustrates the relationships between the 124 haplotypes, representing 146 individuals (Fig. 5). We were able to grouped those 124 haplotypes into 22 clusters, with two clusters, clusters 8 and 9 only represented by two haplotypes, which weren't linked to another cluster, even with different representation methods (data not shown). If some of the clusters, such as cluster 2, are very compact, some clusters were subdivided into two to four groups, such as cluster 3 with 3a, 3b or cluster 14 with 14a, 14b,

14c, 14d (Supplementary Table 1). Only few clusters could be linked to some characteristic of their corresponding samples, those are cluster 1, grouping some haplotypes of African samples, cluster 2, grouping some Red Junglefowl haplotypes and cluster 21 grouping some Grey Junglefowl haplotypes. But the pattern drawn by the network indicated a structuration mostly based on *Mhc*, not a structuration related to the breeds. This geographic or species-specific structuration was only observed with specific *Mhc* alleles, such as with Grey Junglefowl. Interestingly, Red Junglefowl haplotypes were scattered in the network. There were 17 different serotypes represented in the study, by 43 homozygous animals. Their position in the network is indicated, they are linked to 12 clusters. Some serotypes were grouped together: B12, B124 and B19, grouped in cluster 6, B4 and B13, grouped in cluster 15, and B5, B8, and B11, grouped in cluster 16. We were also able to compare heterozygosity at the *B-DM* level with heterozygosity at the marker LEI0258 level (Supplementary Table 1). Six animals were homozygotes for the marker LEI0258 but heterozygotes concerning *B-DM* haplotypes but in each case, the two *B-DM* haplotypes were grouped in the same cluster and sub-group. There were also two animals homozygotes for *B-DM* haplotypes but heterozygotes for the marker LEI0258, so in total, the heterozygosity matched between *B-DM* region and LEI0258 for 139 of the 147 cases (95 %).

Using the same 124 reconstructed haplotypes, we were also able to predict the consequences on the B-DM proteins of the 28 non-synonymous SNPs. The non-synonymous SNPs in *B-DMA* led to nine amino-acid variations (genome residue-position-alternative residue): G-16-W, S-27-L, A-77-T, A-90-T, P-116-L, V-157-I, V-161-I, T-167-I, and R-246-Q (Fig. 2). The non-synonymous SNPs in *B-DMB1* led to 13 amino-acid variations: G-49-S, V-54-M, D-75-N, I-82-V, I-89-V, A-154-T, D-170-E, A-189-V, T-196-A, A-221-V, V-224-M, A-232-V, and R-250-W (Fig.3). The non-synonymous SNPs in *B-DMB2* led to six amino-acid variations: N-23-S, P-26-S, D-53-N, V-68-M, H-147-R, and A-246-T (Fig. 3). Predicting the consequences of amino-acids variations for those chicken proteins was difficult, in relation with a lack of suited predictive tools. Only few substitutions seemed to affect a conserved residue and there were conservative.

## 4. Discussion

### 4.1. Genomic organization of *B-DMA*, *B-DMB1* and *B-DMB2*

Structurally, the *B-DM* genes have the features of *Mhc* class II loci, and potential additional disulfide bridges in their  $\alpha 1$  or  $\beta 1$  domains. It is worthily to note that the amino-acid sequence of the *B-DMB1* gene lacks a putative endosomal targeting signal, so its function could be different from the Mammals' DM classic function. Interestingly, putative *DMB1* genes found in Quail and Turkey also lacks this signal. We also found a higher rate of non-synonymous SNPs in *B-DMB1* and a more restricted expression. Further studies on functions of this protein would be of high interest because those features don't seem to fully match those of any *Mhc* gene reported in Mammals.

But, the three *B-DM* genes are expressed in immune related tissues and those observations are consistent with the ESTs reported in the databases. Phylogenetic analysis established the clustering of the *B-DM* and the *DM* genes, confirming the orthology of *B-DM* genes with *DM* genes.

Moreover, the phylogenetic analysis suggests that the *DM* genes did separate from classical class II genes at an early stage, at least prior to the divergence of evolutionary lineages leading to Mammals and Birds, some 300 million years ago. This had been suggested before, in Mammals (O'hUigin *et al.*, 1998) and was to be confirmed by the analysis of a non-mammalian Vertebrate, such as the chicken.

It is also relevant to mention that whereas the two *DMB* genes of the Mouse are very similar, the chicken and the two bird species, the quail and the turkey, for which we could find putative *DM* sequences in databases, seem to present two *DMB* genes not so similar, indicating a more ancient duplication.

### 4.2. Genetic polymorphism of the *B-DM* molecules

We resequenced the *B-DM* genes in 95 domestic chickens of different origins and in 25 Red Junglefowl, five Grey Junglefowl, and four Sri Lanka Junglefowl, and we added to our analysis 17 animals from GenBank database. If the polymorphisms are frequent in *B-DM* genes, they aren't as numerous as found in classical class II genes, as reported in Mammals (Carrington *et al.*, 1993, Hermel *et al.*, 1995). The high sequence diversity in the *Mhc* is linked to its important role in the host immune response and evidence of positive selection has been demonstrated recently in chicken at those loci (Hosomichi *et al.*, 2008). The 28 amino acids residues variations reported in this study could result in functional differences for the encoded *B-DM* molecules, predicted to be key proteins of the exogenous antigens

presentation pathway. Since, in chickens, the *Mhc B* region is strongly associated with resistance to some infectious diseases, any protein difference between serotypes is of high importance.

Furthermore, the high number of haplotypes identified, 124, especially under a non-recombinant model, gives for the first time an overview of the worldwide *Mhc* diversity for domestic chickens. Commercial breeds and especially local breeds showed a very high diversity. This leads to some loss of correspondence between the LEI0258 and GAB0001 allele sizes and the *B-DM* haplotypes, in contrast with reference White Leghorns lines.

Finally, the multiple positions of Red Junglefowl haplotypes in the network, and to a lesser extent, of Grey Junglefowl and Sri-Lanka Junglefowl, can be explained by the multiple events of domestication in Asia, mainly from the Red Junglefowl, as shown before (International Chicken Genome Sequencing Consortium, 2004) but also from other wild Junglefowl, as demonstrated recently (Eriksson *et al.*, 2008). The polymorphisms described between the different breeds and *Mhc B* haplotypes studied, are very valuable for taking a step towards understanding the evolutive dynamics of the *B* region, in its class II part but also between the LEI0258 marker, located approximately 50 kb away, close to the class I region. We were able to defined 22 clusters from the identified genotypes, which can be considered as indicating the existence of at least 20 different *Mhc* haplotypes in chicken, whereas 17 common serotypes, defined in serology, only matched 12 clusters.

## **5. Conclusion**

This study demonstrates the existence of three *DM* genes in chicken: *B-DMA*, *B-DMB1*, and *B-DMB2*. They are non-classical class II genes, homologous to the Mammals' *DM* genes and are expressed in various immune related tissues. Their existence in the Chicken is consistent with the hypothesis of their presence in the primordial *Mhc*. The polymorphisms described between the different breeds and serotypes studied indicates a great diversity of the *Mhc* in chicken breeds, in particular in local breeds, only partially represented by experimental lines.

## **Acknowledgements**

We would like to thank Issaka A.K. Youssao (Université d'Abomey-Calavi, Cotonou, Bénin), Valentine Yapi-Gnaore (Centre Nationale de Recherche Agronomique, Abidjan, Côte d'Ivoire) and Jean-Claude Fotsa (Institut de Recherche Agricole pour le Développement, Yaoundé, Cameroun) for bird samples, and Boniface B. Kayang (University of Ghana, Legon, Ghana), Chi-Sheng Chang and Chih-Feng Chen (National Chung-Hsing University, Taichung, Taiwan) for bird samples and DNA extractions.

## **Appendix**

**Supplementary Figure 1**

**Supplementary Table 1**

**Supplementary Table 2**

**Supplementary Table 3**



## References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Bandelt, H-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37-48.
- Bernot, A., Auffray, C., 1991. Primary structure and ontogeny of an avian CD3 transcript. *Proc. Natl. Acad. Sci. U S A.* 88, 2550-2554.
- Briles, W.E., Briles, R.W., 1987. Genetics and classification of major histocompatibility complex antigens of the chicken. *Poult. Sci.* 66, 776-781.
- Carrington, M., Yeager, M., Mann, D., 1993. Characterization of HLA-DMB polymorphism. *Immunogenetics* 38, 446-449.
- Chang, C.M., Coville, J.L., Coquerelle, G., Gourichon, D., Oulmouden, A., Tixier-Boichard, M., 2006. Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genomics* 7, 19.
- Cho, S., Attaya, M., Brown, M.G., Monaco, J.J., 1991. New class II-like genes in the murine MHC. *Nature* 353, 573-576.
- Eriksson, J., Larson, G., Gunnarsson, U., Bed'hom, B., Tixier-Boichard, M., Strömstedt, L., Wright, D., Jungerius, A., Vereijken, A., Randi, E., Jensen, P., Andersson, L., 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.* 4, e1000010.
- Flot, J.F., 2009. Seqphase, a web tool for interconverting phase input/output files and fasta sequence alignments. *Mol. Ecol. Resour.* in press.
- Fulton, J.E., Juul-Madsen, H.R., Ashwell, C.M., McCarron, A.M., Arthur, J.A., O'Sullivan, N.P., Taylor, R.L Jr., 2006. Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* 58, 407-421.
- Hermel, E., Yuan, J., Monaco, J.J., 1995. Characterization of polymorphism within the H2-M MHC class II loci. *Immunogenetics* 42, 136-142.
- Hillel, J., Groenen, M.A., Tixier-Boichard, M., Korol, A.B., David, L., Kirzhner, V.M., Burke, T., Barre-Dirie, A., Crooijmans, R.P., Elo, K., Feldman, M.W., Freidlin, P.J., Mäki-Tanila, A., Oortwijn, M., Thomson, P., Vignal, A., Wimmers, K., Weigend, S., 2003. Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genet. Sel. Evol.* 35, 533-557.
- Hosomichi, K., Miller, M.M., Goto, R.M., Wang, Y., Suzuki, S., Kulski, J.K., Nishibori, M., Inoko, H., Hanzawa, K., Shiina, T., 2008. Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. *J. Immunol.* 181, 3393-3399.

International Chicken Genome Sequencing Consortium, 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695-716.

Kaufman, J., Milne, S., Göbel, T.W., Walker, B.A., Jacob, J.P., Auffray, C., Zoorob, R., Beck, S., 1999. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 28, 923-925.

Kelly, A.P, Monaco, J.J., Cho, S., Trowsdale, J., 1991. A new human HLA class II-related locus, DM. *Nature* 353, 571-573.

Kropshofer, H., Arndt, S.O., Moldenhauer, G., Hammerling, G.J., Vogt, A.B., 1997a. HLA-DM acts as a molecular chaperone and rescues empty HLA-DR molecules at lysosomal pH. *Immunity* 6, 293-302.

Kropshofer, H., Hammerling, G.J., Vogt, A.B., 1997b. How HLA-DM edits the MHC class II peptide repertoire: survival of the fittest? *Immunol. Today*. 18, 77-82.

Kropshofer, H., Hämmerling, G.J., Vogt, A.B., 1999. The impact of the non-classical MHC proteins HLA-DM and HLA-DO on loading of MHC class II molecules. *Immunol. Rev.* 172, 267-278.

McConnell, S.K., Dawson, D.A., Wardle, A., Burke, T., 1999. The isolation and mapping of 19 tetranucleotide microsatellite markers in the chicken. *Anim. Genet.* 30, 183-189.

Miller, M.M., Bacon, L.D., Hala, K., Hunt, H.D., Ewald, S.J., Kaufman, J., Zoorob, R., Briles, W.E., 2004. 2004 Nomenclature for the chicken major histocompatibility (B and Y) complex. *Immunogenetics* 56, 261-279.

Minozzi, G., Parmentier, H.K., Nieuwland, M.G., Bed'hom, B., Minvielle, F., Gourichon, D., Pinard-van der Laan, M.H., 2007. Antibody responses to keyhole limpet hemocyanin, lipopolysaccharide, and Newcastle Disease virus vaccine in F2 and backcrosses of white Leghorn lines selected for two different immune response traits. *Poult. Sci.* 86, 1316-1322.

Morris, P., Shaman, J., Attaya, M., Amaya, M., Goodman, S., Bergman, C., Monaco, J.J., Mellins, E., 1994. An essential role for HLA-DM in antigen presentation by class II major histocompatibility molecules. *Nature* 368, 551-554.

O'hUigin, C., Sülthmann, H., Tichy, H., Murray, B.W., 1998. Isolation of mhc class II DMA and DMB cDNA sequences in a marsupial: the gray short-tailed opossum (*Monodelphis domestica*). *J. Mol. Evol.* 47, 578-585.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365-386.

Shiina, T., Briles, W.E., Goto, R.M., Hosomichi, K., Yanagiya, K., Shimizu, S., Inoko, H., Miller, M.M., 2007. Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J. Immunol.* 178, 7162-7172.

- Sloan, V.S., Cameron, P., Porter, G., Gammon, M., Amaya, M., Mellins, E., Zaller, D.M., 1995. Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* 375, 802-806.
- Staden, R., Beal, K.F., Bonfield, J.K., 2000. The Staden package, 1998. *Methods Mol. Biol.* 132, 115-130.
- Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978-989.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596-1599.
- Van den Elsen, P.J., Holling, T.M., Kuipers, H.F., Van der Stoep, N., 2004. Transcriptional regulation of antigen presentation. *Curr. Opin. Immunol.* 16, 67-75.
- Weckx, S., Del-Favero, J., Rademakers, R., Claes, L., Cruts, M., De Jonghe, P., Van Broeckhoven, C., De Rijk, P., 2005. NovoSNP, a novel computational tool for sequence variation discovery. *Genome Research* 15, 436-442.
- Westerheide, S.D., Louis-Pence, P., Ping, D., He, X.F., Boss, J.M., 1997. HLA-DMA and HLA-DMB gene expression functions through the conserved S-X-Y region. *J. Immunol.* 158, 4812-4821.
- Ye, H., Xu, F.Z., Yu, W.Y., 2009. The intracellular localization and oligomerization of chicken invariant chain with major histocompatibility complex class II subunits. *Poult. Sci.* 88, 1594-1600.

## Figure Captions

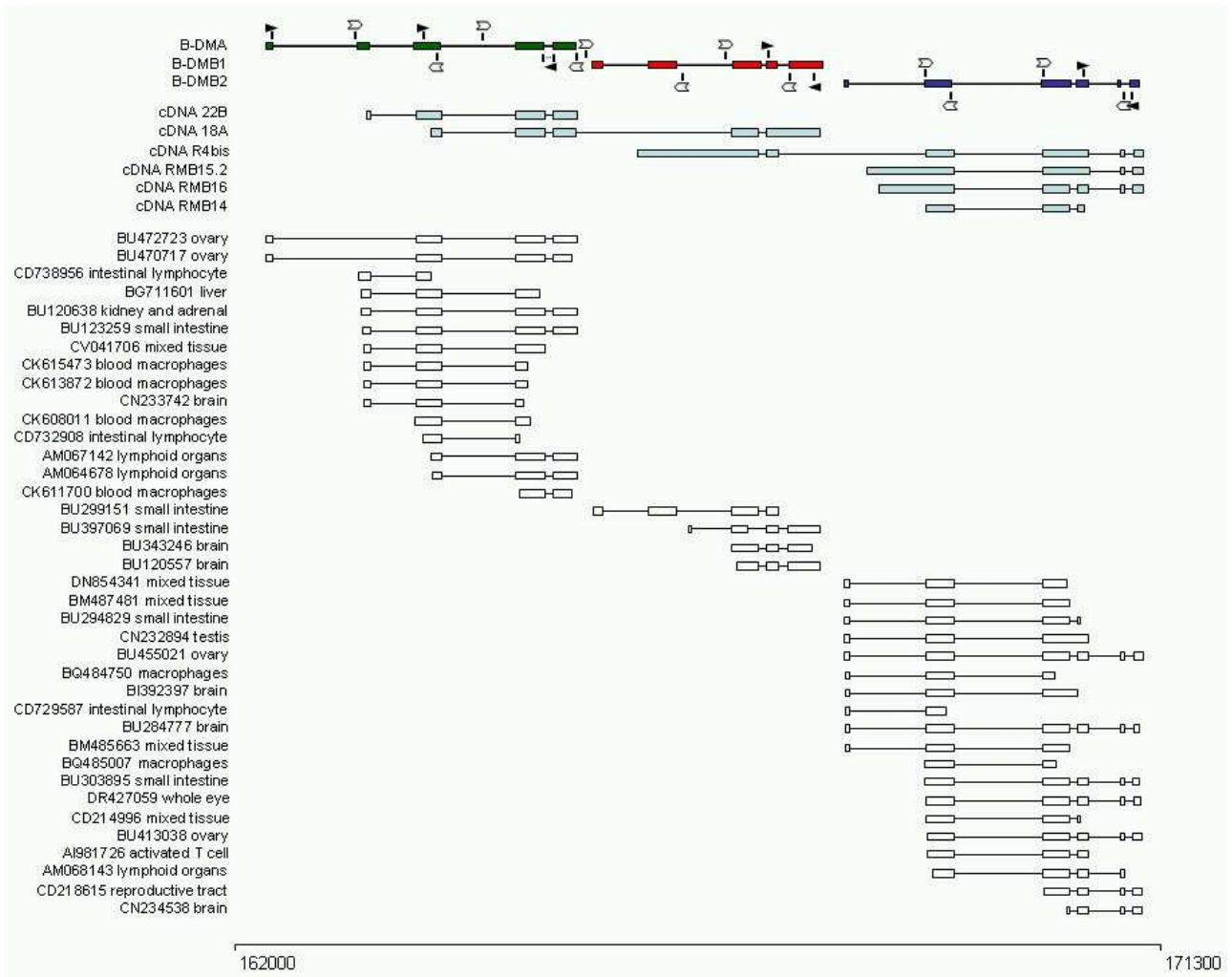
**Fig. 1.** *B-DMA*, *B-DMB1*, and *B-DMB2*, aligned with our seven cDNA and 38 ESTs from GenBank database. Oligonucleotides used for PCR and RT-PCR are positioned respectively as white or black arrows.

**Fig. 2.** Alignment of chicken Gga-B-DMA protein sequence with putative quail and turkey DMA sequence (Cja-DMA, **BAC82512**, Mga-DMA, **ACA64775**), sequences of mammalian proteins DMA (HLA-DMA, H2-DMA, BOLA-DMA, Modo-DMA, respectively: **NP 006111**, **NP 034516**, **NP 001012692**, and **XP 001377359**), putative frog Dma sequence (Xla-DMA, **NP 001083579**), and sequences of Mhc class II  $\alpha$  (chicken Gga-B-La, Mammals' HLA-DRa, HLA-DQa1, H2-Aa, H2-Ea, BOLA-DQa, Modo-DQa, respectively: **AAAY40299**, **NP 061984**, **NP 002113**, **NP 034508**, **NP 034511**, **NP 001013619**, **XP 001376932**, and frog Xla-DRa, **NP 001090247**). Gaps are indicated by dots. Similar amino acids at the same position are shaded in grey, with a 70% threshold. Identical amino acids for the seven Vertebrate DMA proteins are in bold. Potential disulfide bridges are indicated by lines connecting two cysteines. Potential sites of N-glycosylation are underlined. Asterisks indicate variable positions in B-DMA. Sequences are numbered according to the B-DMA protein.

**Fig. 3.** Alignment of chicken protein sequences Gga-B-DMb1 and Gga-B-DMb2 with putative quail and turkey DMb sequence (Cja-DMb1, Cja-DMb2, Mga-DMb1, Mga-DMb2, respectively: **BAC82513**, **BAC82514**, **ACA64776**, **ACA64777**), with sequences of Mammals' proteins DMb (HLA-DMb, H2-DMb1, H2-DMb2, BOLA-DMb, Modo-DMb, respectively: **NP 002109**, **NP 034517**, **NP 034518**, **NP 001035571**, and **XP 001377337**), putative frog Dmb sequence (Xla-DMb, **ABB85336**), and sequences of Mhc class II  $\beta$  (Birds' Gga-B-Lb1, Gga-B-Lb2, Cja-Lb, Mga-Lb, respectively: **NP 001038159**, **NP 001038144**, **BAC82504**, **ACA64772**, and Mammals' HLA-DRb1, HLA-DQb1, H2-Ab1, H2-Eb1, BOLA-DQb, Modo-DAb, respectively: **NP 002115**, **NP 002114**, **NP 996988**, **NP 034512**, **NP 001029840**, **NP 001028163**, and frog Xla-DAb, **BAA08759**). Gaps are indicated by dashes. Similar amino acids in the same position are shaded in grey, with a 70% threshold. Identical amino acids for the eleven DMb proteins are in bold, excluding the sequence of the frog. Potential disulfide bridges are indicated by lines connecting two cysteines, potential sites of N-glycosylation are underlined. Potential endosomal addressing signals are framed. Gga-B-DMb1 is truncated in 3'. Asterisks indicate variable positions in Gga-B-DMb1 and Gga-B-DMb2. Sequences are numbered according to the Gga-B-DMb1 protein.

**Fig. 4.** Bootstrap consensus tree for classical class II and DM  $\beta 2$  domains, derived from the alignment Fig. 3. The percentages indicated are based on the 1000 bootstrap replicates.

**Fig. 5.** NeighborNet network for the 124 B-DM haplotypes defined by 158 SNPs. The scale for branch lengths is indicated in the legend. The clusters are numbered and the presence of an haplotype corresponding to a serotype, previously defined by serology, is indicated by a white arrow.



**Figure 1**

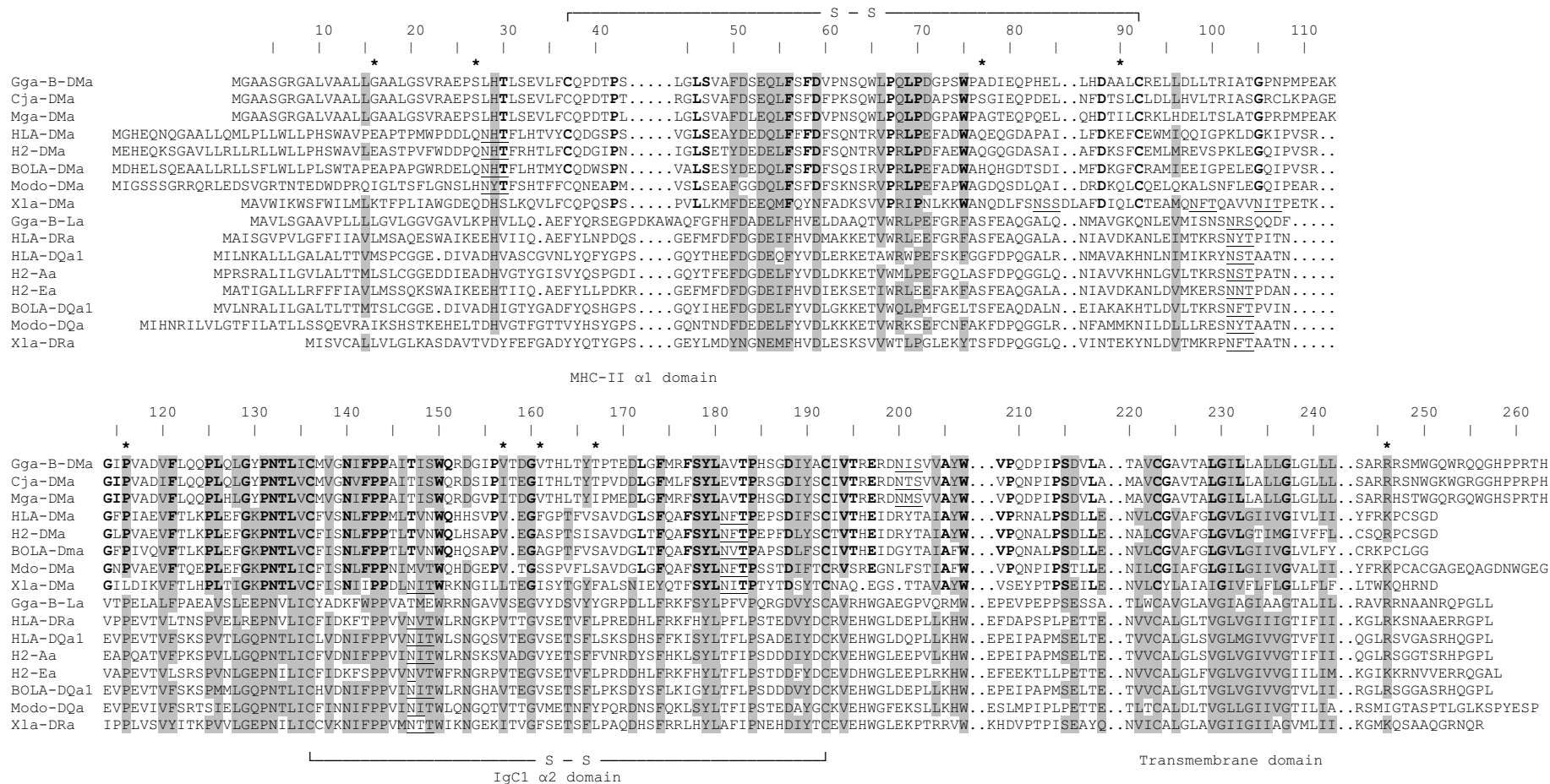


Figure 2

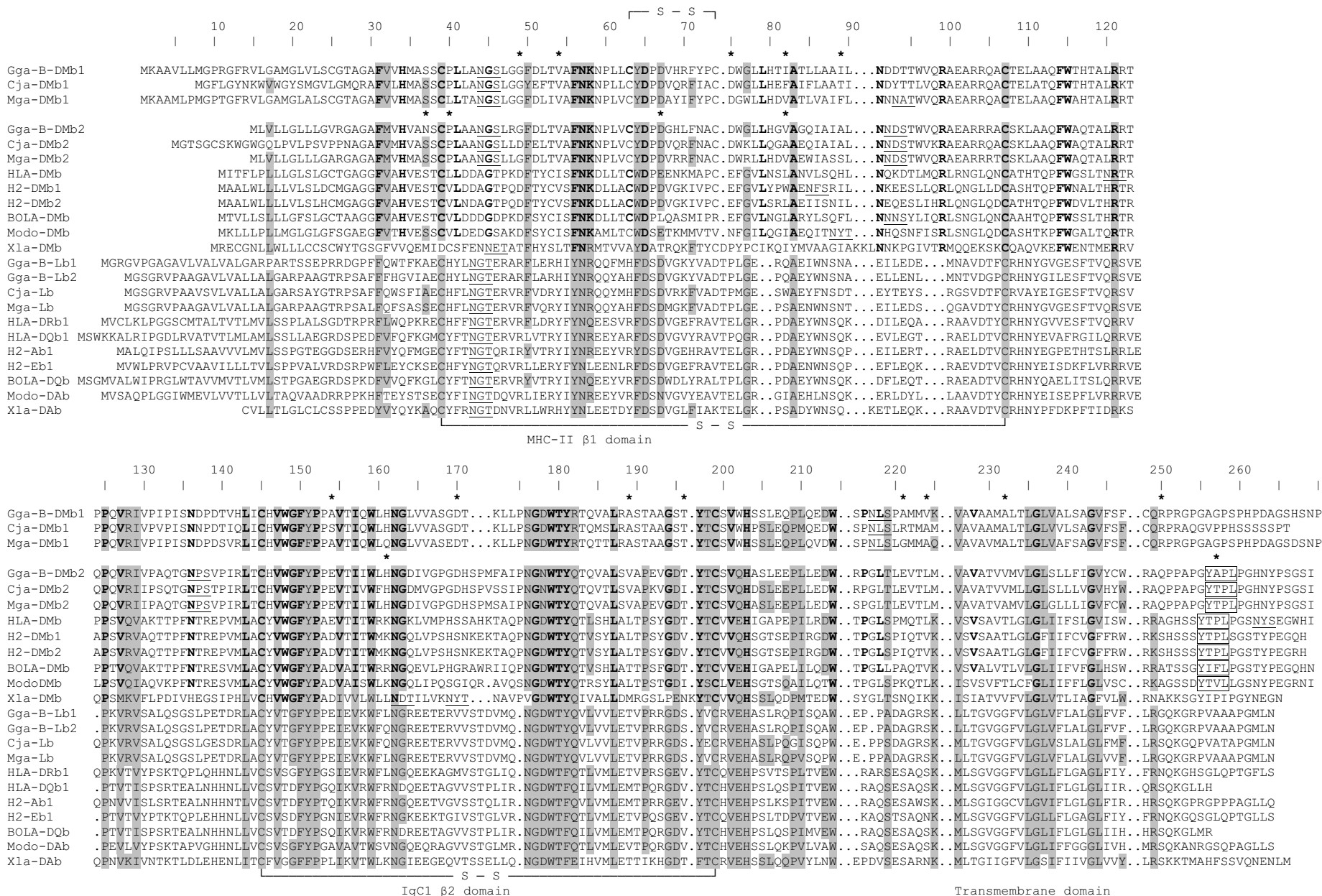
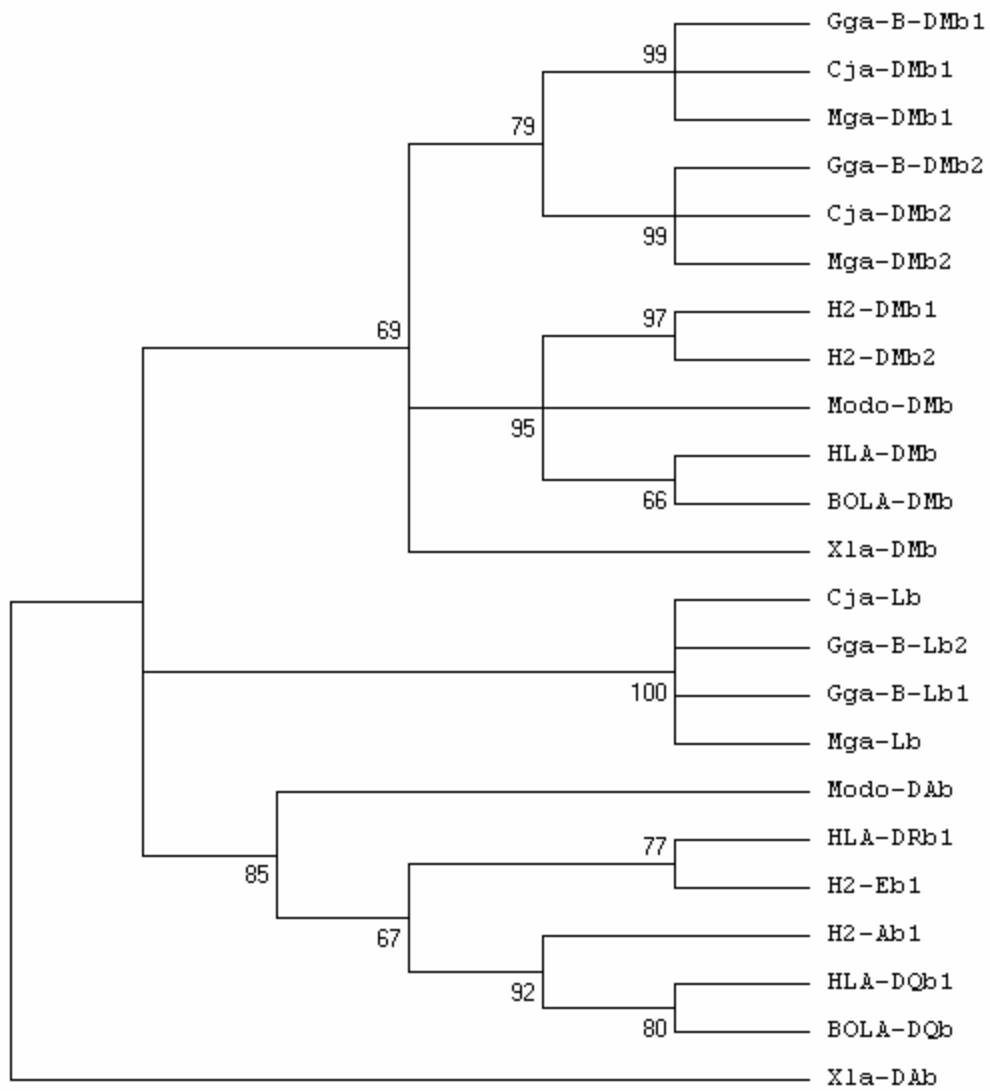


Figure 3





**Figure 4**

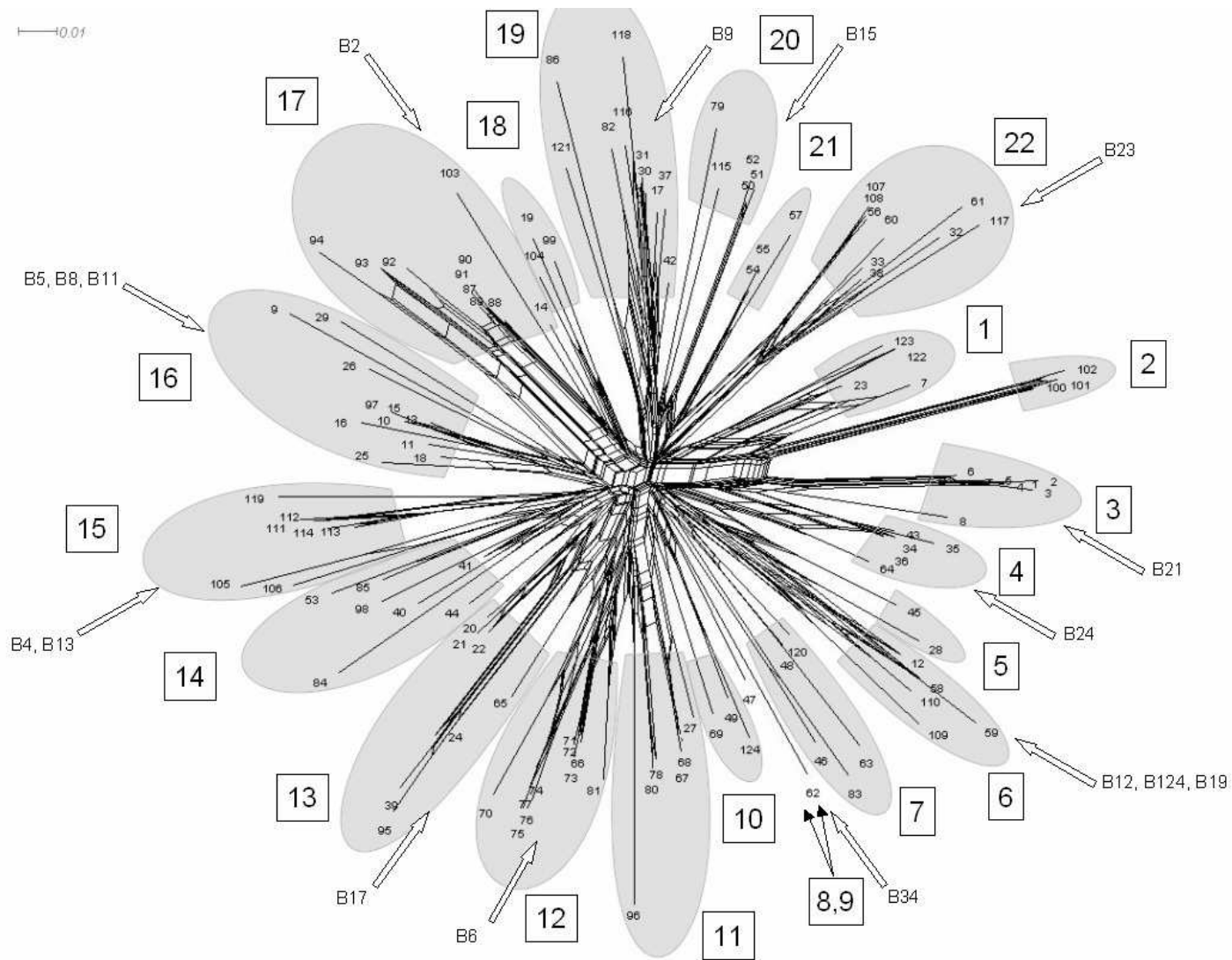


Figure 5

|                                 |  |  |
|---------------------------------|--|--|
|                                 | attaattgcttgcgatggctcagcgcctattacagcgtcggg <b>ggctccc</b> cgctagtttc           |  |
|                                 | ttcta <b>tctagtaacaagtgcgcga</b> aggtaaactgcggagcgcggcc <b>attgg</b> ttgagcccg |  |
|                                 | acgatcatctcctgtcacagcgtggtgttccccgcagatctgttctgctagcaaccg                      |  |
|                                 | ↗ exon 1   |  |
| 5' UT<br>&<br>LP                | atgacgcgtaaagccgcgagggcagccattggctaaactggttgccggtagcagaggg                     |  |
|                                 | <b>ATG</b> GGG GCT GCG AGC GGG CGC GGG GCG CTC GTG GCA GCG CTG CTG 45          |  |
|                                 | M G A A S G R G A L V A A L L 15   |  |
|                                 | ↗ exon 2   |  |
|                                 | GGG GCG GCC CTG GGG AGC GTC AGA GCC GAG CCA TCA CTG CAC ACT 90                 |  |
|                                 | G A A L G S V R A E P S L H T 30   |  |
|                                 | CTG TCC GAG GTG CTC TTC TGC CAG CCG GAC ACG CCG TCG CTG GGG 135                |  |
|                                 | L S E V L F <b>C</b> Q P D T P S L G 45  |  |
|                                 | CTG TCA GTG GCC TTC GAC TCA GAG CAG CTC TTC TCA TTC GAT GTC 180                |  |
|                                 | L S V A F D S E Q L F S F D V 60   |  |
| α1                              | CCC AAC TCG CAG TGG CTG CCG CAG CTC CCC GAT GGC CCC TCG TGG 225                |  |
|                                 | P N S Q W L P Q L P D G P S W 75   |  |
|                                 | CCC GCA GAC ATC GAG CAG CCC CAC GAG CTG CTG CAC GAC GCC GCG 270                |  |
|                                 | P A D I E Q P H E L L H D A A 90   |  |
|                                 | CTG TGC CGT GAG CTG CTC GAT TTG CTC ACC AGA ATC GCC ACC GGG 315                |  |
|                                 | L <b>C</b> R E L L D L L T R I A T G 105                                       |  |
|                                 | ↗ exon 3   |  |
|                                 | CCA AAC CCA ATG CCT GAA GCC AAG GGC ATC CCG GTG GCC GAC GTC 360                |  |
|                                 | P N P M P E A K G I P V A D V 120  |  |
|                                 | TTC CTG CAG CAG CCT CTG CAG CTC GGC TAC CCC AAC ACT CTG ATC 405                |  |
|                                 | F L Q Q P L Q L G Y P N T L I 135  |  |
|                                 | TGT ATG GTG GGC AAC ATC TTC CCC CCA GCC ATC ACT ATC AGC TGG 450                |  |
|                                 | <b>C</b> M V G N I F P P A I T I S W 150                                       |  |
|                                 | CAG CGG GAT GGC ATC CCC GTC ACC GAT GGC GTC ACC CAC CTC ACC 495                |  |
|                                 | Q R D G I P V T D G V T H L T 165  |  |
|                                 | TAC ACC CCC ACC GAG GAC CTG GGC TTC ATG CGC TTC TCC TAC CTG 540                |  |
|                                 | Y T P T E D L G F M R F S Y L 180  |  |
|                                 | GCG GTG ACA CCG CAC TCT GGT GAC ATC TAT GCC TGC ATT GTC ACC 585                |  |
|                                 | A V T P H S G D I Y A <b>C</b> I V T 195                                       |  |
|                                 | ↗ exon 4   |  |
|                                 | CGC GAG AGG GAC AAC ATC TCT GTG GTG GCT TAC TGG GTG CCA CAG 630                |  |
|                                 | R E R D N.....I.....S V V A Y W V P Q 210                                      |  |
|                                 | GAC CCC ATC CCT TCG GAC GTG TTG GCC ACG GCG GTG TGC GGC GCA 675                |  |
|                                 | D P I P S D V L A T A V C G A 225  |  |
|                                 | GTG ACG GCG CTG GGC ATC CTG CTG GCA CTG CTG GGT TTG GGG CTG 720                |  |
|                                 | V T A L G I L L A L L G L G L 240  |  |
|                                 | CTG CTG TCC GCC CGC CGG CGC AGT ATG TGG GGA CAA TGG AGA CAG 765                |  |
|                                 | L L S A R R R S M W G Q W R Q 255  |  |
|                                 | CAG GGA CAC CCG CCC CGT ACT CAC <b>TGA</b> tgtcccccataagttgatccctc 792         |  |
|                                 | Q G H P P R T H * 263  |  |
| CP,<br>TM,<br>CY,<br>&<br>3' UT | ggtgtgggaacggtgatggtgatgtaattaaagcccttcgtttgc (A) <sub>n</sub>                 |  |

**Supplemental Figure 1A:** Sequence features of B-DMA, with coding sequence in upper case and deduced amino acid sequence. Putative promoter boxes are framed. Cysteine residues susceptible to form disulfide bridges are shaded grey. Dotted underlines indicate potential N-glycosylation site. Asterisk indicates stop codon. Polyadenylation signal is underlined. Arrows above the sequence indicate beginning of exons. Region and domain are described on the left with: 5'UT, untranslated 5' region ; LP, leader peptide ; CP, connecting peptide ; TM, transmembrane region ; CY, cytoplasmic region ; 3'UT, untranslated 3' region.

ccataagttgatccctcgggtgtggaacggtgatggtgatgtaattaaagcccttcgt  
 ttgcagcgcggtgtccttgtttgtccccactctggaagggtggcaattaatgggggtg  
 ggcatgtccccatggccccaggtggcaaatctgatccattgcaccgcccattgggggt  
 ↗ exon 1  
 tgacgtgaggaggaggaggagtgaaggagtaggaggaggtggaggtggcacggaggggg  
**5' UT & LP**  
**ATG** AAG GCA GCA GTG CTC CTG ATG GGG CCA AGG GGT TTC AGG GTG 45  
 M K A A V L L M G P R G F R V 15  
 ↗ exon 2  
 CTG GGG GCG ATG GGG CTG GTG CTG AGC TGC GGG ACA GCA GGC GCC 90  
 L G A M G L V L S C G T A G A 30  
**β1**  
 TTC GTG GTG CAC ATG GCC AGC TCC TGC CCA CTG CTG GCC AAT GGC 135  
 F V V H M A S S C P L L A N G 45  
 TCC CTG GGC GGC TTC GAC CTC ACC GTG GCC TTC AAC AAG AAC CCT 180  
 S L G G F D L T V A F N K N P 60  
 CTG CTG TGC TAC GAC CCC GAC GTC CAC CGC TTC TAC CCC TGC GAT 225  
 L L C Y D P D V H R F Y P C D 75  
 TGG GGG CTG CTG CAC ACC ATT GCC ACT TTG CTC GCC GCC ATC CTA 270  
 W G L L H T I A T L L A A I L 90  
 AAT GAT GAT ACC ACA TGG GTG CAG CGT GCA GAG GCA CGC AGG CAG 315  
 N D D T T W V Q R A E A R R Q 105  
 GCG TGC ACT GAG CTG GCT GCA CAG TTC TGG ACA CAC ACA GCA CTG 360  
 A C T E L A A Q F W T H T A L 120  
 ↗ exon 3  
 CGC AGG ACA CCA CCC CAG GTC CGC ATC GTC CCC ATC CCC ACC TCC 405  
 R R T P P Q V R I V P I P I S 135  
 AAC GAC CCC GAC ACC GTC CAC CTC ATC TGC CAT GTT TGG GGC TTC 450  
 N D P D T V H L I C H V W G F 150  
**β2**  
 TAC CCA CCC GCA GTG ACC ATC CAG TGG CTG CAC AAC GGC CTC GTG 495  
 Y P P A V T I Q W L H N G L V 165  
 GTG GCC TCA GGT GAC ACC AAA CTG CTG CCC AAC GGG GAC TGG ACC 540  
 V A S G D T K L L P N G D W T 180  
 TAC AGG ACA CAG GTG GCC CTG AGG GCC AGC ACT GCA GCA GGG AGC 585  
 Y R T Q V A L R A S T A A G S 195  
 ACC TAC ACA TGC TCA GTG TGG CAC TCC AGC CTG GAG CAG CCG CTG 630  
 T Y T C S V W H S S L E Q P L 210  
 ↗ exon 4  
 CAG GAG GAC TGG AGT CCC AAT TTG TCC CCG GCG ATG ATG GTG AAG 675  
 Q E D W S P N L S P A M M V K 225  
 GTG GCA GTG GCG GCC ATG GCG CTG ACG TTG GGG TTG GTG GCA CTC 720  
 V A V A A M A L T L G L V A L 240  
 ↗ exon 5  
 AGC GCC GGG GTT TTC AGC TTC TGT CAG CGG CCA CGG GGT CCT GGC 765  
 S A G V F S F C Q R P R G P G 255  
**CP, TM, CY, & 3' UT**  
 GCT GGT CCC AGT CCC CAT CCT GAT GCG GGT TCT CAC TCC AAT CCT 810  
 A G P S P H P D A G S H S N P 270  
 GGT CCC CAA AAT GAT CCC GGT CCG AGT TCT GGT CCC CAT CCC AGT 855  
 G P Q N D P G P S S G P H P S 285  
 CCT GGA CCC CAT CCC AGT CCT GGT CCC CAT TCT GGT CTT GGT CCT 900  
 P G P H P S P G P H S G L G P 300  
 GGT CCT GGT TCT GCT CCT GGT CCC TAT CCC TGA ctctagtcccgtcc 933  
 G P G S A P G P Y P \* 310  
 ccatacccgatgccagtcccagtcctgggtccccatcctggtcctgctcctcggtttgggg  
 acctcaatgactggaactcccatgtcccaacatggggaccacagtttggggtaggggg  
 ctctcaccucccaataaaaccatctgcagccc (A)<sub>n</sub>

**Supplemental Figure 1B:** Sequence features of B-DMB1 with coding sequence in upper case and deduced amino acid sequence. Putative promoter boxes are framed. Cysteine residues susceptible to form disulfide bridges are shaded grey. Dotted underlines indicate potential N-glycosylation site. Asterisk indicates stop codon. Polyadenylation signal is underlined. Arrows above the sequence indicate beginning of exons. Region and domain are described on the left with: 5'UT, untranslated 5' region ; LP, leader peptide ; CP, connecting peptide ; TM, transmembrane region ; CY, cytoplasmic region ; 3'UT, untranslated 3' region.

ggggtgaggggtctcacccecaataaaaccatctgcagccccaacctcgctccaattc  
 ttcgttccacggtgggtgggtcgggtcccagtgctcccagtgctcccagccggtctat  
 gtcccgtaagcgccggctccaactgcattctgcfccggaacagatgacgctaccacggc  
 gccgcctctgattggctgctccgtgcctctctccgtcccacgtccgtgaaggggggga  
 5' UT & LP  
 tgtgggggtgagggagctgagggggccgccccctcccccccccgctccccctccg  
 ↗ exon 1  
 ATG TTG GTG CTA TTG GGG CTG CTG CTG GGA GTG CGG GGG GCA GGT 45  
 M L V L L G L L L G V R G A G 15  
 ↗ exon 2  
 GCC TTC ATG GTG CAT GTG GCC AAC TCC TGC CCA CTG GCA GCC AAT 90  
 A F M V H V A N S C P L A A N... 30  
 GGC TCT CTG CGC GGC TTC GAC CTC ACC GTG GCC TTC AAC AAG AAC 135  
 G S L R G F D L T V A F N K N 45  
 β1  
 CCT CTG GTG TGC TAC GAC CCC GAT GGC CAC CTC TTC AAC GCC TGC 180  
 P L V C Y D P D G H L F N A C 60  
 GAC TGG GGG CTG CTG CAC GGC GTG GCT GGA CAG ATT GCC ATT GCC 225  
 D W G L L H G V A G Q I A I A 75  
 CTC AAC AAT GAC AGC ACC TGG GTG CAG CGT GCA GAG GCA CGG AGA 270  
 L N N D S T W V Q R A E A R R 90  
 CGG GCG TGC AGC AAA CTG GCT GCA CAG TTC TGG GCA CAG ACG GCG 315  
 R A C S K L A A Q F W A Q T A 105  
 ↗ exon 3  
 CTG CGC AGG ACT CAG CCC CAG GTC CGC ATC GTC CCC GCA CAG ACA 360  
 L R R T Q P Q V R I V P A Q T 120  
 GGG AAC CCC AGC GTG CCC ATC CGC CTC ACC TGC CAC GTG TGG GGC 405  
 G N S V P I R L T C H V W G 135  
 TTC TAC CCC CCC GAG GTG ACC ATC ATC TGG CTA CAC AAT GGG GAC 450  
 F Y P P E V T I I W L H N G D 150  
 β2  
 ATC GTG GGG CCT GGA GAC CAC TCG CCC ATG TTC GCC ATC CCC AAT 495  
 I V G P G D H S P M F A I P N 165  
 GGG AAC TGG ACC TAC CAG ACA CAG GTG GCC CTC TCG GTG GCC CCA 540  
 G N W T Y Q T Q V A L S V A P 180  
 GAG GTG GGG GAC ACC TAC ACG TGC TCG GTG CAG CAT GCT AGC TTG 585  
 E V G D T Y T C S V Q H A S L 195  
 ↗ exon 4  
 GAG GAG CCC CTC CTG GAG GAC TGG CGT CCT GGG CTG ACG CTG GAG 630  
 E E P L L E D W R P G L T L E 210  
 GTG ACG CTG ATG GTG GCT GTG GCC ACT GTA GTG ATG GTG TTG GGG 675  
 V T L M V A V A T V V M V L G 225  
 CP, TM, CY, & 3' UT  
 CTC AGC TTG CTC TTC ATT GGT GTC TAC TGC TGG CGG GCC CAA CCC 720  
 L S L L F I G V Y C W R A Q P 240  
 ↗ exon 5  
 CCT GCC CCA GGT TAC GCC CCG CTT CCC GGT CAC AAC TAC CCT TCA 765  
 P A P G Y A P L P G H N Y P S 255  
 ↗ exon 6  
 GGC AGC ATC TGA tggacacctctgtcaccaactgtccctgtgtgtccccatccc 777  
 G S I \* 258  
 tgactctgcaccgtggtgctgacattaaagacactctgcagcctc (A)<sub>n</sub>

**Supplemental Figure 1C:** Sequence features of B-DMB2, with coding sequence in upper case and deduced amino acid sequence. Putative promoter boxes are framed. Cysteine residues susceptible to form disulfide bridges are shaded grey. Dotted underlines indicate potential N-glycosylation site. Asterisk indicates stop codon. Polyadenylation signal is underlined. Arrows above the sequence indicate beginning of exons. Region and domain are described on the left with: 5'UT, untranslated 5' region ; LP, leader peptide ; CP, connecting peptide ; TM, transmembrane region ; CY, cytoplasmic region ; 3'UT, untranslated 3' region.

## Supplementary Table 1

Detailed description of the samples with indication of the cluster defined with network analysis, breed, line or species, the world continent of origin, the serotype when known, the B-DM region haplotype inferred with PHASE, the number of alleles of the animal corresponding to the indicated haplotype, and the *Mhc* genotype with the allele sizes for marker LEI0258 and marker GAB0001.

| cluster | Breed, Line or Species     | Origin | serotype | B-DM haplotype   | n | LEI0258 all1     | LEI0258 all2     | GAB0001 all1     | GAB0001 all2     |
|---------|----------------------------|--------|----------|------------------|---|------------------|------------------|------------------|------------------|
| 1       | Bedouin                    | Africa | -        | 123 <sup>c</sup> | 1 | 307 <sup>a</sup> | 333 <sup>a</sup> | 159              | 174              |
| 1       | Benin                      | Africa | -        | 122 <sup>c</sup> | 1 | 205 <sup>a</sup> | 333 <sup>a</sup> | 161              | 174              |
| 1       | Cote d'Ivoire              | Africa | -        | 23               | 1 | 322              | 333              | 174              | 174              |
| 1       | Bedouin                    | Africa | -        | 7                | 1 | 333 <sup>a</sup> | 381 <sup>a</sup> | 174              | 174              |
| 2       | R-JF                       | Asia   | -        | 102 <sup>c</sup> | 1 | 367              | 381              | NA               | NA               |
| 2       | R-JF                       | Asia   | -        | 101              | 1 | 367              | 381              | NA               | NA               |
| 2       | R-JF                       | Asia   | -        | 100              | 1 | 295              | 367              | NA               | NA               |
| 3a      | Marans                     | Europe | -        | 6 <sup>c</sup>   | 1 | 295 <sup>a</sup> | 357 <sup>a</sup> | 161              | 174              |
| 3a      | Orlov                      | Asia   | -        | 2                | 1 | 261 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 3a      | WL I.R. line1              | Europe | B21      | 5                | 1 | 357 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | Transsylvanian naked neck  | Europe | -        | 1                | 1 | 249 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | WL AB426152 B21            | Europe | B21      | 1                | 2 | 357 <sup>b</sup> | 357 <sup>b</sup> | 173 <sup>b</sup> | 173 <sup>b</sup> |
| 3a      | WL I.R. line1              | Europe | B21      | 1                | 1 | 357 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | WL I.R. line3              | Europe | B21      | 1                | 2 | 357 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | Cameroon                   | Africa | -        | 1 <sup>c</sup>   | 1 | 235 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | Fayoumi                    | Africa | -        | 1                | 1 | 193 <sup>a</sup> | 381 <sup>a</sup> | 161              | 174              |
| 3a      | Ghana                      | Africa | -        | 1                | 2 | 381 <sup>a</sup> | 381 <sup>a</sup> | 174              | 174              |
| 3a      | Broiler sire line C        | Asia   | -        | 1                | 2 | 357 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3a      | Broiler sire line C        | Asia   | -        | 1                | 1 | 295 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 3a      | Orlov                      | Asia   | -        | 1                | 1 | 261 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 3a      | R-JF P102                  | Asia   | -        | 1 <sup>c</sup>   | 1 | 205 <sup>a</sup> | 381 <sup>a</sup> | NA               | NA               |
| 3a      | R-JF Genome                | Asia   | -        | 1                | 2 | 369 <sup>b</sup> | 369 <sup>b</sup> | 173 <sup>b</sup> | 173 <sup>b</sup> |
| 3a      | R-JF AB268588              | Asia   | -        | 1                | 2 | 369 <sup>b</sup> | 369 <sup>b</sup> | 173 <sup>b</sup> | 173 <sup>b</sup> |
| 3a      | Bedouin                    | Africa | -        | 4                | 1 | 333 <sup>a</sup> | 381 <sup>a</sup> | 174              | 174              |
| 3a      | Fayoumi                    | Africa | -        | 4                | 1 | 319 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 3a      | Green-legged partridge     | Europe | -        | 3                | 1 | 205 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 3b      | Green-legged partridge     | Europe | -        | 8                | 1 | 205 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 4a      | Icelandic landrace         | Europe | -        | 43               | 1 | 217 <sup>a</sup> | 309              | 144              | 159              |
| 4a      | Taiwan KM                  | Asia   | -        | 35               | 1 | 309              | 381              | 159              | 159              |
| 4a      | Brown-egg layer line C     | Asia   | -        | 34               | 2 | 309 <sup>a</sup> | 309 <sup>a</sup> | 159              | 159              |
| 4a      | Low Immune Response line   | Asia   | -        | 34               | 1 | 249 <sup>a</sup> | 309 <sup>a</sup> | 159              | 161              |
| 4a      | New Hampshire AB426154 B24 | Asia   | B24      | 34               | 2 | 309 <sup>b</sup> | 309 <sup>b</sup> | 159 <sup>b</sup> | 159 <sup>b</sup> |
| 4a      | Taiwan J                   | Asia   | -        | 34               | 2 | 309              | 309              | 159              | 159              |
| 4a      | Taiwan KM                  | Asia   | -        | 34               | 2 | 309              | 309              | 159              | 159              |
| 4a      | Taiwan Ng                  | Asia   | -        | 34               | 1 | 249              | 309              | 159              | 161              |
| 4a      | Taiwan T                   | Asia   | -        | 34               | 1 | 309              | 443              | 159              | 159              |
| 4a      | R-JF P101                  | Asia   | -        | 34               | 2 | 273 <sup>a</sup> | 273 <sup>a</sup> | NA               | NA               |
| 4a      | R-JF P101                  | Asia   | -        | 34               | 1 | 273 <sup>a</sup> | 417 <sup>a</sup> | NA               | NA               |
| 4a      | R-JF P102                  | Asia   | -        | 34               | 1 | 273 <sup>a</sup> | 417 <sup>a</sup> | NA               | NA               |
| 4a      | Taiwan J                   | Asia   | -        | 34               | 1 | 205              | 309              | 159              | 161              |
| 4a      | Yurlov crower              | Europe | -        | 36               | 1 | 249              | 309              | 159              | 174              |

|     |                        |        |      |                  |   |                  |                  |                  |                  |
|-----|------------------------|--------|------|------------------|---|------------------|------------------|------------------|------------------|
| 4a  | Brown-egg layer line C | Asia   | -    | 36               | 1 | 307 <sup>a</sup> | 309 <sup>a</sup> | 159              | 187              |
| 4a  | Brown-egg layer line D | Asia   | -    | 36               | 1 | 307 <sup>a</sup> | 309 <sup>a</sup> | 159              | 187              |
| 4a  | Brown-egg layer line D | Asia   | -    | 36               | 1 | 307 <sup>a</sup> | 309 <sup>a</sup> | 159              | 187              |
| 4b  | Silkie                 | Asia   | -    | 64               | 2 | 319 <sup>a</sup> | 319 <sup>a</sup> | NA               | NA               |
| 5   | R-JF P102              | Asia   | -    | 45 <sup>c</sup>  | 1 | 205 <sup>a</sup> | 381 <sup>a</sup> | NA               | NA               |
| 5   | R-JF P102              | Asia   | -    | 28 <sup>c</sup>  | 1 | 193 <sup>a</sup> | 285 <sup>a</sup> | NA               | NA               |
| 6   | Broiler sire line B    | Asia   | -    | 12               | 1 | 195              | 262              | 174              | 174              |
| 6   | Broiler sire line C    | Asia   | -    | 12               | 1 | 295 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 6   | WL AB426147 B12        | Europe | B12  | 58               | 2 | 487 <sup>b</sup> | 487 <sup>b</sup> | 159 <sup>b</sup> | 159 <sup>b</sup> |
| 6   | WL AB426151 B19        | Europe | B19  | 58               | 2 | 539 <sup>b</sup> | 539 <sup>b</sup> | 159 <sup>b</sup> | 159 <sup>b</sup> |
| 6   | WL C line B12          | Europe | B12  | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL C line B12          | Europe | B12  | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL C line B12r1        | Europe | B12  | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL C line B12r1        | Europe | B12  | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL Congenic INRA B12   | Europe | B12  | 58               | 2 | 487              | 487              | 159              | 159              |
| 6   | WL Congenic INRA B12   | Europe | B12  | 58               | 2 | 487              | 487              | 159              | 159              |
| 6   | WL Congenic INRA B19   | Europe | B19  | 58               | 2 | 539              | 539              | 159              | 159              |
| 6   | WL Congenic INRA B19   | Europe | B19  | 58               | 2 | 539              | 539              | 159              | 159              |
| 6   | WL I.R. line1          | Europe | B124 | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL I.R. line2          | Europe | B124 | 58               | 2 | 487 <sup>a</sup> | 487 <sup>a</sup> | 159              | 159              |
| 6   | WL I.R. line2          | Europe | B19  | 58               | 2 | 539 <sup>a</sup> | 539 <sup>a</sup> | 159              | 159              |
| 6   | WL I.R. line2          | Europe | B19  | 58               | 2 | 539 <sup>a</sup> | 539 <sup>a</sup> | 159              | 159              |
| 6   | WL_AL023516_B12        | Europe | B12  | 58               | 2 | 487 <sup>b</sup> | 487 <sup>b</sup> | 159 <sup>b</sup> | 159 <sup>b</sup> |
| 6   | Jaerhoens              | Europe | -    | 59 <sup>c</sup>  | 1 | 261 <sup>a</sup> | 526 <sup>a</sup> | 159              | 174              |
| 6   | R-JF P102              | Asia   | -    | 110 <sup>c</sup> | 1 | 319              | 453 <sup>a</sup> | NA               | NA               |
| 6   | Silkie                 | Asia   | -    | 109              | 2 | 484 <sup>a</sup> | 484 <sup>a</sup> | 159              | 159              |
| 7a  | G-JF                   | Asia   | -    | 120              | 1 | 193 <sup>a</sup> | 441 <sup>a</sup> | NA               | NA               |
| 7a  | G-JF                   | Asia   | -    | 63               | 1 | 239              | 381              | NA               | NA               |
| 7a  | G-JF                   | Asia   | -    | 63               | 1 | 239 <sup>a</sup> | 381              | NA               | NA               |
| 7a  | G-JF                   | Asia   | -    | 48 <sup>c</sup>  | 2 | 381 <sup>a</sup> | 381 <sup>a</sup> | NA               | NA               |
| 7a  | SL-JF                  | Asia   | -    | 83               | 2 | 367 <sup>a</sup> | 367 <sup>a</sup> | NA               | NA               |
| 7a  | SL-JF                  | Asia   | -    | 46               | 1 | 298              | 367              | NA               | NA               |
| 7a  | SL-JF                  | Asia   | -    | 46               | 1 | 298              | 367              | NA               | NA               |
| 7a  | SL-JF                  | Asia   | -    | 46               | 1 | 298 <sup>a</sup> | 367              | NA               | NA               |
| 8   | WL I.R. line1          | Europe | B34  | 62               | 2 | 420 <sup>a</sup> | 420 <sup>a</sup> | 174              | 174              |
| 8   | WL I.R. line1          | Europe | B34  | 62               | 2 | 420 <sup>a</sup> | 420 <sup>a</sup> | 174              | 174              |
| 9   | G-JF                   | Asia   | -    | 47               | 1 | 239              | 381              | NA               | NA               |
| 9   | G-JF                   | Asia   | -    | 47               | 1 | 239 <sup>a</sup> | 381              | NA               | NA               |
| 10  | Bedouin                | Africa | -    | 124 <sup>c</sup> | 1 | 307 <sup>a</sup> | 333 <sup>a</sup> | 159              | 174              |
| 10  | Broiler sire line D    | Asia   | -    | 49 <sup>c</sup>  | 1 | 273              | 295              | 159              | 159              |
| 10  | Red Villafrantina      | Europe | -    | 69 <sup>c</sup>  | 1 | 217 <sup>a</sup> | 249 <sup>a</sup> | 159              | 174              |
| 11a | Taiwan KT              | Asia   | -    | 27               | 1 | 273              | 345              | 159              | 174              |
| 11a | Taiwan KT              | Asia   | -    | 68               | 1 | 273              | 273              | 174              | 174              |
| 11a | Taiwan KT              | Asia   | -    | 67               | 1 | 273              | 273              | 174              | 174              |
| 11b | R-JF P102              | Asia   | -    | 78               | 2 | 297 <sup>a</sup> | 297 <sup>a</sup> | NA               | NA               |
| 11b | R-JF P102              | Asia   | -    | 78               | 1 | 319 <sup>a</sup> | 319 <sup>a</sup> | NA               | NA               |
| 11b | R-JF P102              | Asia   | -    | 78               | 1 | 319 <sup>a</sup> | 319 <sup>a</sup> | NA               | NA               |
| 11b | R-JF P102              | Asia   | -    | 80               | 1 | 297 <sup>a</sup> | 369 <sup>a</sup> | NA               | NA               |
| 11b | R-JF P101              | Asia   | -    | 96               | 1 | 205 <sup>a</sup> | 295 <sup>a</sup> | NA               | NA               |
| 12a | Czech golden pencilled | Europe | -    | 81               | 2 | 193              | 193              | 159              | 159              |
| 12a | Czech golden pencilled | Europe | -    | 81               | 1 | 193 <sup>a</sup> | 249 <sup>a</sup> | 159              | 161              |



|     |                           |        |     |                  |   |                  |                  |                  |                  |
|-----|---------------------------|--------|-----|------------------|---|------------------|------------------|------------------|------------------|
| 12b | Fayoumi                   | Africa | -   | 66               | 1 | 319 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 12b | WL AB426143 B6            | Europe | B6  | 73               | 2 | 443 <sup>b</sup> | 443 <sup>b</sup> | 160 <sup>b</sup> | 160 <sup>b</sup> |
| 12b | Taiwan T                  | Asia   | -   | 72               | 1 | 443              | 443              | 159              | 159              |
| 12b | Taiwan T                  | Asia   | -   | 72               | 1 | 309              | 443              | 159              | 159              |
| 12b | Taiwan T                  | Asia   | -   | 71               | 1 | 443              | 443              | 159              | 159              |
| 12c | R-JF P102                 | Asia   | -   | 74 <sup>c</sup>  | 1 | 205              | 357 <sup>a</sup> | NA               | NA               |
| 12c | Light brown leghorn       | Europe | -   | 76               | 1 | 193 <sup>a</sup> | 309 <sup>a</sup> | 187              | 187              |
| 12c | Brown-egg layer line C    | Asia   | -   | 75               | 1 | 307              | 309              | 159              | 159              |
| 12c | Brown-egg layer line D    | Asia   | -   | 75               | 1 | 307 <sup>a</sup> | 309 <sup>a</sup> | 159              | 187              |
| 12c | Brown-egg layer line D    | Asia   | -   | 75               | 1 | 307 <sup>a</sup> | 309 <sup>a</sup> | 159              | 187              |
| 12c | High Immune Response line | Asia   | -   | 77               | 2 | 307 <sup>a</sup> | 307 <sup>a</sup> | 187              | 187              |
| 12c | High Immune Response line | Asia   | -   | 77               | 1 | 307 <sup>a</sup> | 357 <sup>a</sup> | 159              | 187              |
| 12c | Padova                    | Europe | -   | 70               | 1 | 261 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 12c | Padova                    | Europe | -   | 70 <sup>c</sup>  | 1 | 261 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 13a | WL AB426150 B17           | Europe | B17 | 65               | 2 | 205 <sup>b</sup> | 205 <sup>b</sup> | 174 <sup>b</sup> | 174 <sup>b</sup> |
| 13b | Yurlov crower             | Europe | -   | 24               | 1 | 193              | 249              | 174              | 174              |
| 13b | R-JF P102                 | Asia   | -   | 95 <sup>c</sup>  | 1 | 205              | 357 <sup>a</sup> | NA               | NA               |
| 13b | Yurlov crower             | Europe | -   | 39               | 1 | 249              | 309              | 159              | 174              |
| 13c | Yurlov crower             | Europe | -   | 22               | 1 | 193              | 249              | 174              | 174              |
| 13c | R-JF P101                 | Asia   | -   | 21               | 1 | 309 <sup>a</sup> | 410              | NA               | NA               |
| 13c | Cote d'Ivoire             | Africa | -   | 20               | 1 | 322              | 333              | 174              | 174              |
| 14a | Benin                     | Africa | -   | 44               | 2 | 193 <sup>a</sup> | 307              | 187              | 187              |
| 14b | SL-JF                     | Asia   | -   | 84               | 1 | 298              | 367              | NA               | NA               |
| 14b | SL-JF                     | Asia   | -   | 84               | 1 | 298              | 367              | NA               | NA               |
| 14b | SL-JF                     | Asia   | -   | 84               | 1 | 298 <sup>a</sup> | 367              | NA               | NA               |
| 14b | R-JF P101                 | Asia   | -   | 40               | 1 | 205 <sup>a</sup> | 295 <sup>a</sup> | NA               | NA               |
| 14b | Taiwan KM                 | Asia   | -   | 41               | 1 | 309              | 381              | 159              | 159              |
| 14c | Icelandic landrace        | Europe | -   | 98 <sup>c</sup>  | 1 | 250              | 346              | 159              | 161              |
| 14d | R-JF P101                 | Asia   | -   | 85 <sup>c</sup>  | 1 | 283 <sup>a</sup> | 333 <sup>a</sup> | NA               | NA               |
| 14d | Jaerhoens                 | Europe | -   | 53               | 1 | 250              | 261              | 174              | 174              |
| 15a | Cote d'Ivoire             | Africa | -   | 106              | 1 | 369              | 381              | 144              | 159              |
| 15a | Cameroon                  | Africa | -   | 105 <sup>c</sup> | 1 | 235 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 15b | Light brown leghorn       | Europe | -   | 113              | 1 | 193 <sup>a</sup> | 309 <sup>a</sup> | 187              | 187              |
| 15b | Light brown leghorn       | Europe | -   | 114              | 1 | 193 <sup>a</sup> | 295 <sup>a</sup> | 161              | 187              |
| 15b | WL AB426148 B13           | Europe | B13 | 111              | 2 | 205 <sup>b</sup> | 205 <sup>b</sup> | 187 <sup>b</sup> | 187 <sup>b</sup> |
| 15b | WL C line B4              | Europe | B4  | 111              | 2 | 182 <sup>a</sup> | 182 <sup>a</sup> | 187              | 187              |
| 15b | WL C line B4              | Europe | B4  | 111              | 2 | 182 <sup>a</sup> | 182 <sup>a</sup> | 187              | 187              |
| 15b | WL C line B4r1            | Europe | B4  | 111              | 2 | 183 <sup>a</sup> | 183 <sup>a</sup> | 187              | 187              |
| 15b | WL C line B4r1            | Europe | B4  | 111              | 2 | 182 <sup>a</sup> | 182 <sup>a</sup> | 187              | 187              |
| 15b | WL Congenic INRA B13      | Europe | B13 | 111              | 2 | 205              | 205              | 187              | 187              |
| 15b | WL Congenic INRA B13      | Europe | B13 | 111              | 2 | 205              | 205              | 187              | 187              |
| 15b | Broiler dam line D        | Asia   | -   | 112              | 1 | 205              | 218              | 144              | 144              |
| 15c | R-JF P101                 | Asia   | -   | 119 <sup>c</sup> | 1 | 295              | 345 <sup>a</sup> | NA               | NA               |
| 16a | R-JF P101                 | Asia   | -   | 25 <sup>c</sup>  | 1 | 295              | 345 <sup>a</sup> | NA               | NA               |
| 16a | R-JF P102                 | Asia   | -   | 18               | 1 | 297 <sup>a</sup> | 369 <sup>a</sup> | NA               | NA               |
| 16a | R-JF P102                 | Asia   | -   | 18               | 2 | 369 <sup>a</sup> | 369 <sup>a</sup> | NA               | NA               |
| 16b | Marans                    | Europe | -   | 11 <sup>c</sup>  | 1 | 295 <sup>a</sup> | 357 <sup>a</sup> | 161              | 174              |
| 16b | Light brown leghorn       | Europe | -   | 16               | 1 | 193 <sup>a</sup> | 295 <sup>a</sup> | 161              | 187              |
| 16b | Fayoumi                   | Africa | -   | 10               | 1 | 193 <sup>a</sup> | 381 <sup>a</sup> | 161              | 174              |
| 16b | Taiwan KT                 | Asia   | -   | 13               | 1 | 273              | 345              | 159              | 174              |
| 16b | Taiwan F                  | Asia   | -   | 97               | 1 | 295              | 357              | 159              | 161              |

|     |                           |        |     |                  |   |                  |                  |                  |                  |
|-----|---------------------------|--------|-----|------------------|---|------------------|------------------|------------------|------------------|
| 16b | Ancona AB426144 B8        | Europe | B8  | 15               | 2 | 405 <sup>b</sup> | 405 <sup>b</sup> | 161 <sup>b</sup> | 161 <sup>b</sup> |
| 16b | Ancona AB426146 B11       | Europe | B11 | 15               | 2 | 193 <sup>b</sup> | 193 <sup>b</sup> | 161 <sup>b</sup> | 161 <sup>b</sup> |
| 16b | WL AB426142 B5            | Europe | B5  | 15               | 2 | 295 <sup>b</sup> | 295 <sup>b</sup> | 161 <sup>b</sup> | 161 <sup>b</sup> |
| 16b | Cameroon                  | Africa | -   | 15               | 1 | 249 <sup>a</sup> | 295 <sup>a</sup> | 161              | 161              |
| 16b | Taiwan J                  | Asia   | -   | 15               | 1 | 205              | 309              | 159              | 161              |
| 16c | Red Villafrantina         | Europe | -   | 26               | 2 | 217 <sup>a</sup> | 217 <sup>a</sup> | 174              | 174              |
| 16c | Red Villafrantina         | Europe | -   | 26 <sup>c</sup>  | 1 | 217 <sup>a</sup> | 249 <sup>a</sup> | 159              | 174              |
| 16d | Ghana                     | Africa | -   | 9                | 2 | 261 <sup>a</sup> | 319 <sup>a</sup> | 159              | 159              |
| 16d | Broiler sire line D       | Asia   | -   | 29 <sup>c</sup>  | 1 | 273              | 295              | 159              | 159              |
| 17a | Broiler sire line B       | Asia   | -   | 94               | 1 | 195              | 262              | 174              | 174              |
| 17a | Friesian fowl             | Europe | -   | 93               | 1 | 193              | 261              | 159              | 159              |
| 17a | Friesian fowl             | Europe | -   | 93               | 1 | 193              | 261 <sup>a</sup> | 159              | 159              |
| 17a | Friesian fowl             | Europe | -   | 92               | 1 | 193              | 261              | 159              | 159              |
| 17b | Friesian fowl             | Europe | -   | 89               | 1 | 193              | 261 <sup>a</sup> | 159              | 159              |
| 17b | Padova                    | Europe | -   | 89               | 1 | 261 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 17b | Broiler sire line B       | Asia   | -   | 89               | 1 | 195              | 262              | 174              | 174              |
| 17b | Orlov                     | Asia   | -   | 87               | 1 | 261 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 17b | Orlov                     | Asia   | -   | 91               | 1 | 261 <sup>a</sup> | 357 <sup>a</sup> | 159              | 174              |
| 17b | Broiler sire line B       | Asia   | -   | 90               | 1 | 195              | 262              | 174              | 174              |
| 17b | Padova                    | Europe | -   | 88 <sup>c</sup>  | 1 | 261 <sup>a</sup> | 381 <sup>a</sup> | 159              | 174              |
| 17c | R-JF                      | Asia   | -   | 103              | 1 | 295              | 367              | NA               | NA               |
| 17d | WL AB426141 B2            | Europe | B2  | 14               | 2 | 261 <sup>b</sup> | 261 <sup>b</sup> | 161 <sup>b</sup> | 161 <sup>b</sup> |
| 18  | R-JF                      | Asia   | -   | 104 <sup>c</sup> | 1 | 367              | 381              | NA               | NA               |
| 18  | R-JF                      | Asia   | -   | 104              | 1 | 367              | 381              | NA               | NA               |
| 18  | R-JF                      | Asia   | -   | 19               | 1 | 381 <sup>a</sup> | 393 <sup>a</sup> | NA               | NA               |
| 18  | R-JF                      | Asia   | -   | 19               | 1 | 381              | 393              | NA               | NA               |
| 18  | R-JF                      | Asia   | -   | 19               | 1 | 381              | 393              | NA               | NA               |
| 18  | Transsylvanian naked neck | Europe | -   | 99               | 2 | 249              | 249              | NA               | NA               |
| 18  | Transsylvanian naked neck | Europe | -   | 99               | 1 | 249 <sup>a</sup> | 357 <sup>a</sup> | 174              | 174              |
| 19a | Benin                     | Africa | -   | 121 <sup>c</sup> | 1 | 205 <sup>a</sup> | 333 <sup>a</sup> | 161              | 174              |
| 19a | R-JF                      | Asia   | -   | 86               | 1 | 381 <sup>a</sup> | 393 <sup>a</sup> | NA               | NA               |
| 19a | R-JF                      | Asia   | -   | 86               | 1 | 381              | 393              | NA               | NA               |
| 19a | R-JF                      | Asia   | -   | 86               | 1 | 381              | 393              | NA               | NA               |
| 19a | Czech golden pencilled    | Europe | -   | 82               | 1 | 193 <sup>a</sup> | 249 <sup>a</sup> | 159              | 161              |
| 19b | Icelandic landrace        | Europe | -   | 116 <sup>c</sup> | 1 | 250              | 346              | 159              | 161              |
| 19b | Broiler dam line D        | Asia   | -   | 118              | 1 | 249 <sup>a</sup> | 369 <sup>a</sup> | 159              | 161              |
| 19b | WL AB426145 B9            | Europe | B9  | 31               | 2 | NA               | NA               | 161 <sup>b</sup> | 161 <sup>b</sup> |
| 19b | Low Immune Response line  | Asia   | -   | 31               | 1 | 249 <sup>a</sup> | 249 <sup>a</sup> | 161              | 161              |
| 19b | Taiwan Ng                 | Asia   | -   | 31 <sup>c</sup>  | 2 | 249              | 249              | 161              | 161              |
| 19b | Low Immune Response line  | Asia   | -   | 30               | 1 | 249 <sup>a</sup> | 249 <sup>a</sup> | 161              | 161              |
| 19b | Cameroon                  | Africa | -   | 17               | 1 | 249 <sup>a</sup> | 295 <sup>a</sup> | 161              | 161              |
| 19b | Low Immune Response line  | Asia   | -   | 37               | 1 | 249 <sup>a</sup> | 309 <sup>a</sup> | 159              | 161              |
| 19b | Taiwan Ng                 | Asia   | -   | 42               | 1 | 249              | 309              | 159              | 161              |
| 20a | R-JF P101                 | Asia   | -   | 79               | 1 | 273 <sup>a</sup> | 417 <sup>a</sup> | NA               | NA               |
| 20a | R-JF P101                 | Asia   | -   | 79               | 1 | 309 <sup>a</sup> | 410              | NA               | NA               |
| 20a | R-JF P102                 | Asia   | -   | 79               | 1 | 273 <sup>a</sup> | 417 <sup>a</sup> | NA               | NA               |
| 20a | R-JF P102                 | Asia   | -   | 115 <sup>c</sup> | 1 | 193 <sup>a</sup> | 285 <sup>a</sup> | NA               | NA               |
| 20b | Jaerhoens                 | Europe | -   | 50 <sup>c</sup>  | 1 | 261 <sup>a</sup> | 526 <sup>a</sup> | 159              | 174              |
| 20b | White-egg layer line      | Europe | -   | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | White-egg layer line      | Europe | -   | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL AB426149 B15           | Europe | B15 | 52               | 2 | 261 <sup>b</sup> | 261 <sup>b</sup> | 174 <sup>b</sup> | 174 <sup>b</sup> |

|     |                             |        |     |                  |   |                  |                  |                  |                  |
|-----|-----------------------------|--------|-----|------------------|---|------------------|------------------|------------------|------------------|
| 20b | WL I.R. line2               | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL I.R. line4               | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL Line Sarcoma-resistant   | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL Line Sarcoma-resistant   | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL Line Sarcoma-susceptible | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | WL Line Sarcoma-susceptible | Europe | B15 | 52               | 2 | 261 <sup>a</sup> | 261 <sup>a</sup> | 174              | 174              |
| 20b | Jaerhoens                   | Europe | -   | 51               | 1 | 250              | 261              | 174              | 174              |
| 21  | G-JF                        | Asia   | -   | 54               | 1 | 193 <sup>a</sup> | 441 <sup>a</sup> | NA               | NA               |
| 21  | G-JF                        | Asia   | -   | 55               | 1 | 333 <sup>a</sup> | 441              | NA               | NA               |
| 21  | G-JF                        | Asia   | -   | 57               | 1 | 333 <sup>a</sup> | 441              | NA               | NA               |
| 22a | Cote d'Ivoire               | Africa | -   | 107              | 1 | 369              | 381              | 144              | 159              |
| 22a | Broiler dam line D          | Asia   | -   | 108              | 1 | 205              | 218              | 144              | 144              |
| 22a | Icelandic landrace          | Europe | -   | 108 <sup>c</sup> | 1 | 217 <sup>a</sup> | 309              | 144              | 159              |
| 22a | Marans                      | Europe | -   | 56 <sup>c</sup>  | 1 | 217 <sup>a</sup> | 307              | 144              | 159              |
| 22b | R-JF P102                   | Asia   | -   | 60               | 2 | 249 <sup>a</sup> | 249 <sup>a</sup> | NA               | NA               |
| 22b | R-JF P102                   | Asia   | -   | 60               | 1 | 249 <sup>a</sup> | 333              | NA               | NA               |
| 22b | R-JF P102                   | Asia   | -   | 33 <sup>c</sup>  | 1 | 319              | 453 <sup>a</sup> | NA               | NA               |
| 22b | Marans                      | Europe | -   | 38 <sup>c</sup>  | 1 | 217 <sup>a</sup> | 307              | 144              | 159              |
| 22b | R-JF P101                   | Asia   | -   | 61 <sup>c</sup>  | 1 | 283 <sup>a</sup> | 333 <sup>a</sup> | NA               | NA               |
| 22b | R-JF P101                   | Asia   | -   | 61               | 2 | 333              | 333              | NA               | NA               |
| 22b | R-JF P102                   | Asia   | -   | 61               | 1 | 249 <sup>a</sup> | 333              | NA               | NA               |
| 22b | R-JF P102                   | Asia   | -   | 32               | 1 | 319 <sup>a</sup> | 319 <sup>a</sup> | NA               | NA               |
| 22b | R-JF P102                   | Asia   | -   | 32               | 1 | 319 <sup>a</sup> | 319 <sup>a</sup> | NA               | NA               |
| 22b | Green-legged partridge      | Europe | -   | 117              | 2 | 205 <sup>a</sup> | 205 <sup>a</sup> | NA               | NA               |
| 22b | Broiler dam line D          | Asia   | -   | 117              | 1 | 249 <sup>a</sup> | 369 <sup>a</sup> | 159              | 161              |
| 22b | New Hampshire AB426153 B23  | Asia   | B23 | 117              | 2 | 357 <sup>b</sup> | 357 <sup>b</sup> | 159 <sup>b</sup> | 159 <sup>b</sup> |
| 22b | High Immune Response line   | Asia   | -   | 117              | 1 | 307 <sup>a</sup> | 357 <sup>a</sup> | 159              | 187              |
| 22b | Taiwan F                    | Asia   | -   | 117              | 2 | 357              | 357              | 159              | 159              |
| 22b | Taiwan F                    | Asia   | -   | 117              | 1 | 295              | 357              | 159              | 161              |

<sup>a</sup> indicates sequenced allele, <sup>b</sup> indicates deduced allele size, for LEI0258: from Fulton *et al.*, 2006 or from GenBank sequence, for GAB001: from Hosomichi *et al.*, 2008, <sup>c</sup> indicates probabilities of haplotype inference with PHASE < 0.70.

## Supplementary Table 2

Characteristics and sequences of oligonucleotides used for PCR and RT-PCR. Location are given on the recent chicken *Mhc* from Shiina *et al.*, 2007, accession number **AB268588**.

|        | Gene          | Name        | Fragment length (bp) | Start  | End    | Sequence (5'-3')      |
|--------|---------------|-------------|----------------------|--------|--------|-----------------------|
| PCR    | <i>B-DMA</i>  | MAalt-F     | 1226                 | 162129 | 162143 | GATTGGCCACCCGCC       |
|        |               | MAalt-R     |                      | 163337 | 163354 | TCAGAGCCGGTAGGGGAC    |
|        |               | MA-1F       | 821                  | 163239 | 163258 | GCCATTGGCTAAACTGGTTG  |
|        |               | MA-2R       |                      | 164059 | 164039 | TTGGCTTCAGGCATTGGGTTT |
|        |               | MA-3Fbis    | 869                  | 164511 | 164530 | ATGGACATCCCAACTGATGC  |
|        |               | MA-4R       |                      | 165379 | 165360 | GGATCAACTTATGGGGGACA  |
|        | <i>B-DMB1</i> | MB2-1F      | 826                  | 165629 | 165646 | AAGGGGTTTCAGGGTGCT    |
|        |               | MB2-2R      |                      | 166454 | 166435 | TGTTTGCAATGCTCCACTTG  |
|        |               | MB2-3F      | 575                  | 166934 | 166952 | ATCCCCAGCTCTGACCATC   |
|        |               | MB2-5R      |                      | 167508 | 167489 | AGGAGGTACACCCACACTG   |
|        | <i>B-DMB2</i> | MB1-2F      | 306                  | 168939 | 168959 | AATGCAGGTGCCTTCATGGTG |
|        |               | MB1-2R      |                      | 169244 | 169224 | CCATTTGCAATGCTCACTCCT |
|        |               | MB1-3F      | 884                  | 170152 | 170172 | TCAGCCCCAGGTCCGCATCGT |
|        |               | MB1-4R      |                      | 171035 | 171016 | GAGAGATGGGGTCACTCAGC  |
| RT-PCR | <i>B-DMA</i>  | RT-BMAalt-F | 579                  | 162352 | 162370 | CATGAAGATTCACGGCATT   |
|        |               | RT-BMA-F    | 466                  | 163889 | 163908 | CTCATTGATGTCCCAACT    |
|        |               | RT-BMA-R    |                      | 165091 | 165200 | GTCCTGTGGCACCCAGTAAG  |
|        | <i>B-DMB1</i> | RT-R4-2F    | 369                  | 167339 | 167359 | GATGATGGTGAAGGTGGCAGT |
|        |               | RT-18B-5R   |                      | 167786 | 167806 | CATGGGAGTTCAGTCATTGA  |
|        | <i>B-DMB2</i> | RT-BMB1-F   | 124                  | 170573 | 170592 | GGCTCAGCTTGCTCTTCATT  |
|        |               | RT-BMB1-R   |                      | 171064 | 171083 | TTGGTGACAGAAGGTGTCCA  |

### Supplementary Table 3

The 212 SNPs observed in the *B-DM* region of the chicken *Mhc*. SNPs position are indicated for the genome and for the recent sequence of the *Mhc* of the chicken from Shiina *et al.*, 2007, accession number **AB268588**.

| Genome position | AB268588 position | allA*>allB | DOM "A"<br>n = 220 | DOM "B"<br>n = 220 | RJF "A"<br>n = 54 | RJF "B"<br>n = 54 | GJF "A"<br>n = 10 | GJF "B"<br>n = 10 | SLJF "A"<br>n = 8 | SLJF "B"<br>n = 8 | Coding | aa1 | aa2  |
|-----------------|-------------------|------------|--------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------|-----|------|
| 62328           | 162195            | C>A        | 146                | 2                  | 52                | 0                 | 0                 | 0                 | 4                 | 0                 | N      |     |      |
| 62292           | 162231            | T>G        | 141                | 9                  | 45                | 7                 | 0                 | 0                 | 4                 | 0                 | N      |     |      |
| 62284           | 162239            | C>A        | 140                | 10                 | 43                | 11                | 0                 | 0                 | 2                 | 2                 | N      |     |      |
| 62250           | 162273            | C>A        | <b>12</b>          | <b>138</b>         | <b>8</b>          | <b>46</b>         | <b>0</b>          | <b>2</b>          | <b>0</b>          | <b>8</b>          | N      |     |      |
| 62215           | 162308            | C>T        | 146                | 0                  | 48                | 4                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 62188           | 162335            | T>C        | <b>47</b>          | <b>101</b>         | <b>19</b>         | <b>31</b>         | <b>0</b>          | <b>2</b>          | <b>0</b>          | <b>2</b>          | Y      | V   | A    |
| 62183           | 162340            | C>T        | 138                | 8                  | 50                | 0                 | 2                 | 0                 | 2                 | 0                 | Y      | R   | stop |
| 62177           | 162346            | T>C        | 146                | 0                  | 51                | 1                 | 2                 | 0                 | 2                 | 0                 | Y      | C   | R    |
| 62144           | 162379            | G>T        | 138                | 0                  | 45                | 5                 | 1                 | 1                 | 2                 | 0                 | N      |     |      |
| 62132           | 162391            | G>A        | 126                | 12                 | 46                | 4                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 62124           | 162399            | T>C        | <b>48</b>          | <b>90</b>          | <b>20</b>         | <b>30</b>         | <b>0</b>          | <b>2</b>          | <b>2</b>          | <b>0</b>          | N      |     |      |
| 62105           | 162418            | T>G        | 137                | 1                  | 45                | 5                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 62102           | 162421            | G>A        | 136                | 2                  | 41                | 9                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 62072           | 162451            | G>A        | 140                | 0                  | 45                | 5                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 62069           | 162454            | C>T        | 140                | 0                  | 52                | 0                 | 1                 | 1                 | 2                 | 0                 | N      |     |      |
| 62002           | 162521            | T>C        | 126                | 14                 | 51                | 1                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61997           | 162526            | A>G        | 140                | 0                  | 51                | 3                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61982           | 162541            | C>T        | 144                | 0                  | 53                | 1                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61969           | 162554            | G>C        | 141                | 1                  | 54                | 0                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61940           | 162583            | G>C        | 155                | 1                  | 53                | 1                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61889           | 162634            | G>A        | 124                | 34                 | 43                | 11                | 1                 | 1                 | 2                 | 0                 | N      |     |      |
| 61876           | 162647            | A>T        | 158                | 0                  | 48                | 4                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61875           | 162648            | G>A        | 150                | 8                  | 49                | 3                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61868           | 162655            | G>A        | 156                | 0                  | 51                | 1                 | 4                 | 0                 | 2                 | 0                 | N      |     |      |
| 61849           | 162674            | A>G        | 147                | 11                 | 47                | 5                 | 2                 | 0                 | 2                 | 0                 | N      |     |      |
| 61810           | 162713            | T>C        | 154                | 2                  | 51                | 1                 | 4                 | 0                 | 4                 | 0                 | N      |     |      |
| 61793           | 162730            | C>T        | 154                | 2                  | 52                | 0                 | 5                 | 1                 | 4                 | 0                 | N      |     |      |
| 61791           | 162732            | C>T        | 106                | 50                 | 25                | 27                | <b>1</b>          | <b>5</b>          | 4                 | 0                 | N      |     |      |
| 61782           | 162741            | T>G        | 148                | 8                  | 49                | 3                 | 4                 | 0                 | 4                 | 0                 | N      |     |      |
| 61773           | 162750            | G>A/C      | 146                | 8/2                | 51                | 0/1               | 6                 | 0/0               | 6                 | 0/0               | N      |     |      |
| 61765           | 162758            | C>T        | 158                | 2                  | 49                | 3                 | 6                 | 0                 | 6                 | 0                 | N      |     |      |
| 61705           | 162818            | G>A        | 135                | 27                 | 46                | 6                 | 6                 | 0                 | 6                 | 0                 | N      |     |      |
| 61688           | 162835            | G>C        | 166                | 2                  | 50                | 2                 | 6                 | 0                 | 6                 | 0                 | N      |     |      |
| 61686           | 162837            | T>C        | 166                | 2                  | 51                | 1                 | 6                 | 0                 | 6                 | 0                 | N      |     |      |
| 61641           | 162882            | C>A        | 123                | 51                 | 24                | 28                | <b>0</b>          | <b>8</b>          | <b>0</b>          | <b>8</b>          | N      |     |      |
| 61610           | 162913            | T>C        | 169                | 3                  | 53                | 1                 | 6                 | 0                 | 8                 | 0                 | N      |     |      |
| 61605           | 162918            | C>T/A      | 140                | 32/0               | 43                | 6/3               | 6                 | 0/0               | 8                 | 0/0               | N      |     |      |
| 61583           | 162940            | C>T        | 144                | 32                 | 48                | 6                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61568           | 162955            | A>G        | 174                | 0                  | 54                | 0                 | 8                 | 0                 | 3                 | 5                 | N      |     |      |
| 61532           | 162991            | A>G        | 172                | 2                  | 53                | 1                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61521           | 163002            | A>C        | <b>30</b>          | <b>144</b>         | <b>12</b>         | <b>42</b>         | <b>0</b>          | <b>8</b>          | <b>0</b>          | <b>8</b>          | N      |     |      |
| 61520           | 163003            | T>C        | 125                | 49                 | 30                | 24                | 3                 | 5                 | 3                 | 5                 | N      |     |      |
| 61484           | 163039            | T>C        | 156                | 18                 | 50                | 4                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61471           | 163052            | A>G        | <b>15</b>          | <b>159</b>         | <b>9</b>          | <b>45</b>         | <b>0</b>          | <b>8</b>          | <b>0</b>          | <b>8</b>          | N      |     |      |
| 61462           | 163061            | C>T        | 165                | 9                  | 49                | 5                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61461           | 163062            | G>A        | 162                | 10                 | 54                | 0                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61389           | 163134            | C>T        | 159                | 11                 | 38                | 16                | 4                 | 4                 | 8                 | 0                 | N      |     |      |
| 61371           | 163152            | C>T        | 168                | 0                  | 51                | 3                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61370           | 163153            | G>A        | 164                | 4                  | 50                | 4                 | 4                 | 4                 | 8                 | 0                 | N      |     |      |
| 61301           | 163222            | T>C        | 163                | 1                  | 54                | 0                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61290           | 163233            | G>A        | 164                | 0                  | 51                | 3                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61274           | 163249            | A>G        | 126                | 38                 | 28                | 26                | 3                 | 5                 | 3                 | 5                 | N      |     |      |
| 61260           | 163263            | T>C        | 161                | 1                  | 49                | 5                 | 8                 | 0                 | 5                 | 3                 | N      |     |      |
| 61259           | 163264            | A>T        | 157                | 1                  | 54                | 0                 | 8                 | 0                 | 8                 | 0                 | N      |     |      |
| 61205           | 163318            | G>T        | 210                | 4                  | 54                | 0                 | 9                 | 1                 | 8                 | 0                 | Y      | G   | W    |
| 61191           | 163332            | G>A        | 182                | 34                 | 51                | 3                 | 10                | 0                 | 8                 | 0                 | Y      | G   | G    |
| 61163           | 163360            | C>A        | 216                | 0                  | 48                | 6                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61162           | 163361            | G>A        | 213                | 3                  | 53                | 1                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61160           | 163363            | G>T        | 203                | 13                 | 54                | 0                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61141           | 163382            | C>T        | 200                | 16                 | 52                | 2                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61116           | 163407            | A>G        | 216                | 0                  | 52                | 2                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61113           | 163410            | T>C        | 205                | 11                 | 54                | 0                 | 10                | 0                 | 8                 | 0                 | N      |     |      |
| 61100           | 163423            | C>T        | <b>86</b>          | <b>130</b>         | <b>20</b>         | <b>34</b>         | <b>0</b>          | <b>10</b>         | <b>0</b>          | <b>8</b>          | N      |     |      |
| 61099           | 163424            | G>A        | 216                | 0                  | 51                | 3                 | 10                | 0                 | 8                 | 0                 | N      |     |      |





### C. Discussion et Conclusion

Les gènes *DM* du poulet comme les gènes *DM* des mammifères se distinguent des gènes de classe II classiques par les potentiels ponts disulfures supplémentaires de leurs domaines  $\alpha 1$  et  $\beta 1$ . Au cours de cette étude des gènes ressemblant aux gènes *DM* du poulet ont également pu être identifiés, chez la caille et la dinde, deux autres Galliformes. Les analyses phylogénétiques réalisées ont permis de regrouper les gènes *DM* des trois Galliformes étudiés avec les gènes *DM* des Mammifères, par rapport aux gènes de classe II classiques ce qui confirme le caractère ancestral de ces gènes, qui avait été suggéré auparavant.

Sachant que d'autres gènes de classe II non-classiques existent chez les Mammifères (*DO*, *DN*), on peut s'interroger sur leur existence chez le poulet.

L'étude des polymorphismes de ces gènes dans un grand nombre d'animaux d'origines variées a permis d'identifier un grand nombre de sites variables, dont 28 non-synonymes, bien que ces gènes soient moins polymorphes que les gènes de classe II classiques, probablement en relation avec le fait qu'ils ne présentent pas directement de peptides à la surface cellulaire. Étant donné que chez le poulet, la région B du CMH est fortement associée à la résistance à certaines maladies infectieuses, tout polymorphisme non-synonyme distinguant deux sérotypes de référence est potentiellement intéressant. En outre, le nombre élevé d'haplotypes identifiés, 124, donne un aperçu de la diversité de la région de classe II du CMH chez le poulet au niveau mondial. Il est à noter que cette diversité trouve essentiellement sa source dans les lignées commerciales et les populations locales, conduisant à une certaine perte de correspondance entre les tailles d'allèles aux marqueurs LEI0258 et GAB0001 et les haplotypes *DM* identifiés, contrairement à ce qui est observé pour les lignées de référence White Leghorn. Ainsi le fait que nous ayons pu identifier 22 clusters à partir de ces 124 haplotypes signifie, dans une certaine mesure, qu'au moins une vingtaine de grands groupes d'haplotypes de la région B du CMH existent chez le poulet, alors que les 17 sérotypes B de référence inclus dans l'analyse ne correspondent qu'à 12 clusters.





## **V. CINQUIEME PARTIE : ÉTUDE EXPLORATOIRE DES GENES DE CLASSE I CLASSIQUES PAR CLONAGE**

### **A. Introduction**

Les molécules de classe I classiques sont des protéines membranaires exprimées sur pratiquement toutes les cellules, impliquées dans la présentation de peptides d'origine endogène. On a vu qu'en raison de l'existence de deux gènes très similaires de classe Ia dans la région B du CMH du poulet, l'étude de ces gènes devait faire l'objet d'approches spécifiques et ils n'ont pas pu être reséquencés directement. C'est pour ces raisons qu'une collaboration avec un laboratoire danois, du "Department of Animal Health, Welfare and Nutrition" du centre de recherche de Foulum, a été mise en place. Une équipe expérimentée dans l'étude des gènes de classe I du poulet, dirigée par Helle Juul-Madsen a accueilli un séjour de courte durée de 6 semaines, financé par le réseau d'excellence européen Eadgene. L'objectif de ce séjour était de cloner et séquencer les gènes de classe Ia de la région B du CMH du poulet pour divers animaux. L'utilisation des mêmes techniques pour les gènes de classe II devait également être testée et l'analyse des ADN des 96 animaux formant le panel de populations étudiées par reséquençage au niveau d'autres gènes devait être tentée.

**En raison de la collaboration établie, le rapport concernant cette étude a été rédigé en anglais, mais ne pourra pas faire l'objet d'une publication sans travaux complémentaires.**

## **B. Rapport**

### **GENETIC DIVERSITY OF CHICKEN MHC CLASSICAL CLASS Ia GENES: PRELIMINARY RESULTS**

**Chazara Olympe<sup>1</sup>, Svenstrup Hanne<sup>2</sup>, Bed'Hom Bertrand<sup>1</sup>, Juul-Madsen Helle<sup>2</sup>**

<sup>1</sup>*INRA / AgroParisTech, UMR1313 Génétique Animale et Biologie Intégrative 78350 Jouy en Josas, France*

<sup>2</sup>*Department of Animal Health, Welfare and Nutrition, Danish Institute of Agricultural Sciences, Research Centre Foulum, Tjele, Denmark*

#### **Background**

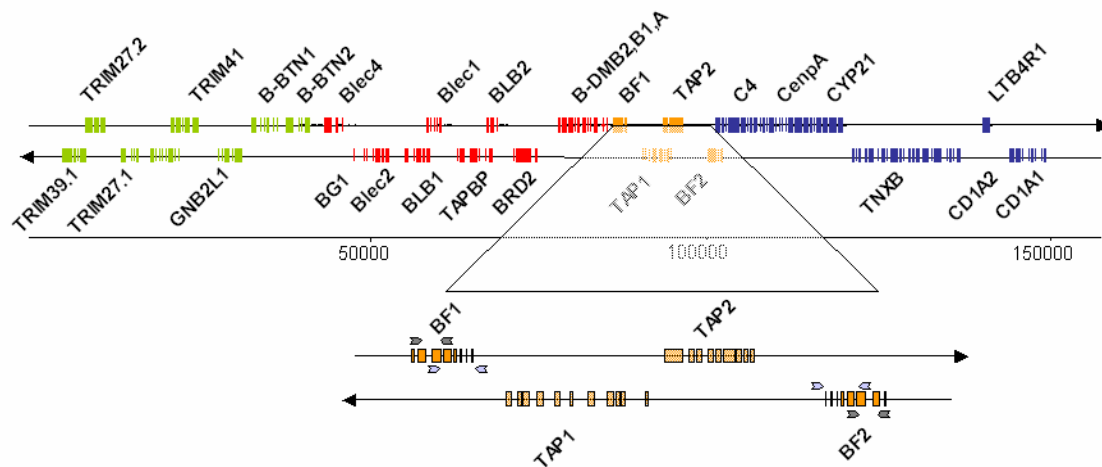
Classical MHC class I molecules are membrane proteins expressed on virtually all cells. They are heterodimers comprised of a highly polymorphic trans-membrane glycoprotein ( $\alpha$ /heavy chain) and a  $\beta$ 2-microglobulin ( $\beta$ /light-chain). They bind small peptides derived from the intracellular compartment: the peptides are bound in a groove of the molecule formed by the two polymorphic domains  $\alpha$ 1 and  $\alpha$ 2, respectively encoded by exons 2 and 3 of the class Ia gene. The binding is sequence specific and mediated by interactions between anchor residues of the peptide and complementary pockets of the groove.

The loading of the peptide takes place in the endoplasmic reticulum, and then the heterodimer reaches the cell surface where the complex is recognized by the T cell receptor (TCR) of cytotoxic T lymphocytes. Classical MHC class I molecules also interact with receptors on natural killer cells. Thus those molecules play important roles in the process of self versus non-self recognition.

The chicken MHC B encodes two class Ia chains: *B-F1* (*B-FI* or *B-F* minor) and *B-F2* (*B-FIV* or *B-F* major). The *B-F2* gene is described as the major expressed locus. Both genes are located in the MHC B class I region, with the *TAP1* and *TAP2* genes, interestingly, *B-F1* and *B-F2*, *TAP1* and *TAP2* are in opposite transcriptional orientation (Figure 1).

The aim of this study was to evaluate the feasibility of full cloning and sequencing (from exon 1 to 3' UTR) the two classical class Ia genes of the chicken MHC B region expressed in

individuals from three different breeds, White-Leghorn, Rhode Island Red and Fayoumi, in order to explore allelic variation and correlation with LEI0258 genotypes.



**Figure 1:** Map of the core area of the B region of the chicken MHC showing details of the class I region and the PCR primers. PCR1 primers are dark gray arrows (exon 1 to exon 3 for both genes), PCR2 primers are light gray arrows (exon 3 to 3'UTR ofr both genes)

## Material and Methods

### *Animals*

For this study, ten animals have been used from a challenge experiment with coccidiosis (*Eimeria tenella*). Four chickens were Rhode Island Red (RIR) and three were White Leghorn (WL), from different experimental lines. These experimental lines were the N line from Compton, UK, fixed at the MHC-level for the serotype B21, the I.R. lines from INRA, France, selected for different immune traits, and the R+/R- lines from INRA, France, selected for feed efficiency. The three remaining chickens were Fayoumi, an Egyptian local breed. The animals had been genotyped at the MHC level with the LEI0258 marker, this VNTR marker is located within the MHC B region of the chicken, and its allele sizes are in most cases correlated with MHC serotypes. In terms of allele size, in base pairs, four animals were homozygous 357/357, two 381/381, one 309/309, one 307/307 and two heterozygous, one 307/309 and one 261/357. Four animals were infected and six were controls (Table 1).

**Table 1:** The ten chickens studied, informations and experimental results

| <b>Id</b>   | <b>Breed</b>       | <b>LEI0258<br/>all1</b> | <b>LEI0258<br/>all2</b> | <b>serotype</b> | <b>Infected</b> | <b>PCR</b> | <b>Tot<br/>clones</b> | <b>Clones<br/>+</b> | <b>Seq</b> |
|-------------|--------------------|-------------------------|-------------------------|-----------------|-----------------|------------|-----------------------|---------------------|------------|
| <b>7833</b> | WL<br>(N line)     | 357                     | 357                     | B21             | N               | 1_564 bp   | 20                    | 11                  | 6          |
|             |                    |                         |                         |                 |                 | 2_850 bp   | 18                    | 6                   | 4          |
| <b>4605</b> | Fayoumi            | 381                     | 381                     | B13.1 ?         | N               | 1          | 10                    | 8                   | 8          |
|             |                    |                         |                         |                 |                 | 2          | 22                    | 6                   | 6          |
| <b>4821</b> | WL<br>(I.R. lines) | 357                     | 357                     | B21             | N               | 1          | 10                    | 7                   | 7          |
|             |                    |                         |                         |                 |                 | 2          | 15                    | 3                   | 3          |
| <b>4203</b> | RIR<br>(R- line)   | 309                     | 309                     | B24 ?           | N               | 1          | 20                    | 6                   | 6          |
|             |                    |                         |                         |                 |                 | 2          | 25                    | 6                   | 5          |
| <b>4394</b> | RIR<br>(R+ line)   | 357                     | 357                     | B21 ?           | N               | 1          |                       |                     |            |
|             |                    |                         |                         |                 |                 | 2          |                       |                     |            |
| <b>4649</b> | Fayoumi            | 381                     | 381                     | B13.1 ?         | Y               | 1          |                       |                     |            |
|             |                    |                         |                         |                 |                 | 2          |                       |                     |            |
| <b>4375</b> | RIR<br>(R+ line)   | 307                     | 309                     | B78/B24 ?       | Y               | 1          |                       |                     |            |
|             |                    |                         |                         |                 |                 | 2          |                       |                     |            |
| <b>4123</b> | RIR<br>(R- line)   | 307                     | 307                     | B78 ?           | Y               | 1          | 18                    | 10                  | 10         |
|             |                    |                         |                         |                 |                 | 2          | 26                    | 3                   | 2          |
| <b>4472</b> | Fayoumi            | 357                     | 357                     | B21 ?           | Y               | 1          | 10                    | 6                   | 6          |
|             |                    |                         |                         |                 |                 | 2          | 10                    | 2                   | 1          |
| <b>4867</b> | WL<br>(I.R. lines) | 261                     | 357                     | B15/B21         | N               | 1          | 18                    | 15                  | 14         |
|             |                    |                         |                         |                 |                 | 2          | 26                    | 3                   | 0          |

WL, White Leghorn; RIR, Rhode Island Red

#### *RNA extraction and cDNA synthesis*

Caecal tonsils had been conserved in RNA later (Qiagen) at  $-20^{\circ}\text{C}$  prior to extraction. Total RNA was isolated using the Rneasy Mini Kit (Qiagen) according to the manufacturer's instructions. Approximately half a caecal tonsil was disrupted and homogenized in 600  $\mu\text{L}$  of buffer RLT by an ultrasound stick on ice and the first centrifugation was done at 13,200 rpm, 3 min. The reverse transcription and first strand cDNA synthesis was carried out in 33  $\mu\text{L}$  using the Amersham Pharmacia Biotech kit from 2.5  $\mu\text{g}$  RNA with DTT solution and pd(N)<sub>6</sub> at 0.2  $\mu\text{g}/\mu\text{L}$  as primers.

#### *PCR amplification*

PCRs were carried out from 2  $\mu\text{L}$  of cDNA in a final volume of 25  $\mu\text{L}$  containing 2.5  $\mu\text{L}$  10X buffer, 2  $\mu\text{L}$  dNTP (2 mM each), 1.25  $\mu\text{L}$  MgCl<sub>2</sub> (25mM), 1  $\mu\text{L}$  of each primer (at 10  $\mu\text{M}$ ), and 0.5 U DNA polymerase. The PCR cycle setup was 40 cycles of 94°C for 1 min, 95°C for 1 min, 55°C for 1 min, 72°C for 2 min, followed by 72°C for 10 min. Two PCR primers pairs were used, they are designed to amplify the two *B-F* genes in the same time. The first one

amplified 564 bp: BF $\alpha$ -1-8 F GAGCTCCATACCCTGCGGTACATC and BF $\alpha$ -180-189 R CCCCCACACTCGCACCTCGGGCCGCTC, and the second one amplified 850 bp: BF-exon-3A F TTCCAGAGGCAGTTCCCACCAA and BF-EXON-8b R AGCCAAACTGGGACACGGTC (Figure 1). The total volume of PCR product was then loaded on 1.5% agarose 200 mL gel containing 10  $\mu$ L RedGel (Biotum). The gel was exposed to UV light to reveal the amplified fragments and their size, by comparison with the 100 Base-Pair Ladder (GE Healthcare). In order to insure specificity, bands at the right size were cut using BandPick (Elchrom Scientific AG) and purified with QIAquick PCR Purification Kit (Qiagen).

### *Cloning and sequencing*

Two  $\mu$ L of the purified PCR products were cloned into the pCR 2.1-TOPO vector using the TOPO TA Cloning Kit and One Shot Mach1<sup>TM</sup>-T1R chemically competent *E. coli* (Invitrogen) according to the manufacturer's instructions. Clones were tested by PCR amplification using custom ordered primers: TOPO M13 forward GTAAAACGACGGCCAG and TOPO M13 reverse CAGGAAACAGCTATGAC, in 25  $\mu$ L final volume, from 1  $\mu$ L liquid culture, with 2.5  $\mu$ L 10X buffer, 2.5  $\mu$ L dNTP (2 mM each), 2.5  $\mu$ L of each primer (at 3  $\mu$ M), and 0.5 U DNA polymerase. The PCR cycle setup was 95°C for 3 min followed by 40 cycles of 95°C for 45 s, 55°C for 1 min, 72°C for 1 min, followed by 72°C for 10 min. When needed, the resulting PCR were purified with the QIAquick PCR Purification Kit (Qiagen), then the sequencing was carried out in both directions by Eurofins MWG Operon, using their standard protocol for purified PCR products, with the custom ordered TOPO M13 primers.

### *Sequence Analysis*

Analysis of sequence quality and variation analyses were carried out with NovoSNP (Weckx *et al.*, 2005), and Pregap4 v1.5 and Gap4 v4.10 components of the Staden package (Staden *et al.*, 2000).

## Results and Discussion

### *mRNA expression of MHC classical class Ia genes*

Purified PCR fragments have been obtained for the ten animals but further analysis were only conducted for seven animals (Table 1). 63 positive clones from 106 tested clones were obtained for the PCR 1 (564 bp) and 29 from 142 clones for the PCR 2 (850 bp), resulting in success percentages of respectively 59% and 20% probably reflecting the decreased efficiency of the cloning but also of the final PCR used to test the clones, with the increase of the size of the insert. Finally 57 sequences corresponding to a classical class Ia gene of the B region of the MHC of the chicken were obtained for PCR1 and 21 for PCR2. It is worth mentioning that for PCR2, three sequences of the class Ia gene of the Y region of the MHC of the chicken were obtained, one for animal 4472 and two for animal 4867, corresponding to the predicted gene XM\_425314 of the chromosome 16 reference assembly (NC\_006103).

### *Nucleotide sequence variability*

We identified 141 polymorphic positions corresponding to 123 biallelic SNPs, 17 triallelic SNPs, and one 1 bp indel. Those polymorphisms defined a minimum number of eight sequences. We also observed two important indels, concerning respectively 33 bp and 63 bp, corresponding to animal 4203. Because those deletions affected the same gene of the same animal, they are probably the consequences of splicing variations. If the 33 bp deletion affecting the entire exon 7 had been observed before (Dalgaard *et al.*, 2005), the 63 bp deletion starting 3' of exon 3 until 5' of exon 4 was never reported and included the site matching the reverse primer of PCR1, BF $\alpha$ -180-189 R.

### *Sequences comparison and analyses*

In order to compare the obtained sequences with sequences of standard chicken serotypes, we aligned our nucleotide sequences with the genomic sequences published by Hosomichi *et al.* in 2008, corresponding to ten White-Leghorn serotypes: B2, B5, B6, B9, B12, B13, B14, B15, B17, B19, and B21, two Ancona breed serotypes: B8 and 11B, and two New Hampshire breed serotypes B23 and B24. The genes were then called *B-F1* or *B-F2* according to their annotation in those sequences (Table 2).

The most frequent sequence, sequence 1, encountered in 34 clones and in four animals, corresponded to the *B21-BF2* gene. Those four animals were homozygous 357/357 with the LEI0258 marker (WL from the N line and the I.R. lines, Fayoumi) or 381/381 (Fayoumi). As described by Fulton *et al.*, in 2006, the 357bp LEI0258 allele corresponds, among others, to the B21 serotype whereas the 381 bp allele corresponds to the B13.1 serotype (defined only in White-Leghorn).

The second sequence, observed in 8 clones, matched the *B23-BF1* gene and was found in one RIR animal, whose genotype was 307/307. Another sequence, sequence 5, was found in three clones from this chicken, no close match was found, even in the GenBank database. The 307 bp LEI0258 allele corresponds to serology of RIR or White Plymouth Rock commercial chickens (defined as B72 and B78).

The third sequence, observed in 10 clones, matched the *B24-BF2* gene and was found in another RIR animal, with a 309/309 genotype. The 309 bp LEI0258 allele corresponds, among others, to the RIR B24 serology.

The sequence 4 and 4bis, observed in 13 clones in a heterozygous 261/357 WL chicken from IR lines, corresponded, for sequence 4, to the *B15-BF2* gene, and, for sequence 4bis, to one *B15-BF2* gene in the GenBank database (accession number Z54316). The 261 bp LEI0258 allele corresponds to the B15 or B2 serology for WL. For this animal, another sequence, sequence 7 was found, in one clone, and no close match was found.

Finally, sequence 6 was found in one clone, for the 309/309 RIR chicken, and matched *B13*, *B15*, *B21*, and *B24-BF1*, and sequence 8 was found in three clones, for the I.R. WL 357/357 chicken and the 381/381 Fayoumi chicken, and corresponded to *B15* or *B21-BF1*.

Besides, for five clones, we observed recombinant sequences composed of the two sequences found for the concerned animal, four cases were in B21, between the genes *B21-BF1* and *BF2*, and one case was between the gene *B23-BF1* and the unknown sequence 5. Such recombinant sequences have been reported previously in MHC genes families, where an amplified gene fragment can act as a PCR primer for another gene, leading to artefactual sequences (Ennis *et al.*, 1990, Bradley and Hillis 1997).



**Table 2:** The 8 *B-F* sequences identified from the seven chickens studied

| Animal Id | LEI0258 | PCR | <i>B-F</i> sequence | Best match                        | No |
|-----------|---------|-----|---------------------|-----------------------------------|----|
| 7833      | 357/357 | 1   | sequence 1          | <i>B21-BF2</i>                    | 6  |
|           |         | 2   |                     |                                   | 3  |
| 4605      | 381/381 | 1   | sequence 1          | <i>B21-BF2</i>                    | 8  |
|           |         | 2   |                     |                                   | 2  |
| 4821      | 357/357 | 1   | sequence 1          | <i>B21-BF2</i>                    | 6  |
|           |         | 2   |                     |                                   | 2  |
| 4472      | 357/357 | 1   | sequence 1          | <i>B21-BF2</i>                    | 6  |
|           |         | 2   |                     |                                   | 1  |
| 4123      | 307/307 | 1   | sequence 2          | <i>B23-BF1</i>                    | 7  |
|           |         | 2   |                     |                                   | 1  |
| 4203      | 309/309 | 1   | sequence 3          | <i>B24-BF2</i>                    | 5  |
|           |         | 2   |                     |                                   | 3  |
|           |         | 2   | sequence 3          | <i>B24-BF2</i> + del 33 bp        | 1  |
|           |         | 2   | sequence 3          | <i>B24-BF2</i> + del 63bp & 33 bp | 1  |
| 4867      | 261/357 | 1   | sequence 4          | <i>B15-BF2</i>                    | 9  |
|           |         | 1   | sequence 4b         | <i>B15-BF2</i> (Z54316)           | 4  |
| 4123      | 307/307 | 1   | sequence 5          | no match                          | 2  |
|           |         | 2   |                     |                                   | 1  |
| 4203      | 309/309 | 1   | sequence 6          | <i>B13, B15, B21, B24-BF1</i>     | 1  |
| 4867      | 261/357 | 1   | sequence 7          | no match                          | 1  |
| 4605      | 381/381 | 2   | sequence 8          | <i>B21-B15-BF1</i>                | 2  |
| 4821      | 357/357 | 2   | sequence 8          | <i>B21, B15-BF1</i> (partial)     | 1  |
| 7833      | 357/357 | 2   | recomb              | <i>B21-BF1</i> & <i>BF2</i>       | 1  |
| 4605      | 381/381 | 2   | recomb              | <i>B21-BF2</i> & <i>BF1</i>       | 2  |
| 4123      | 307/307 | 1   | recomb              | <i>B23-BF1</i> & sequence 5       | 1  |

Globally, those results showed that, in most cases, the *B-F* gene found corresponded to the MHC typing obtained with the LEI0258 marker, even in non-WL animals (Table 2). We also found, excepted in one case, that the majority of the clones obtained have inserted the sequence described as the major gene (91% for B24, 92% for B21). This wasn't true for the 307/307 RIR chicken (4123) for which we have mostly found the sequence described as the *B-F1* gene for B23. As the B23 serotype correspond to an 357 bp allele of the LEI0258 marker, this chicken has probably not a B23 serotype, and *B-F1* and *B-F2* must be well annotated in the Hosomichi *et al.*, sequences as they have been obtained by Long-Range PCR. One hypothesis could be that the *B-F2* gene corresponding to this 307/307 chicken is very similar to the *B23-BF1* gene.

## Conclusion

Determination of the sequences of the class Ia genes of the B region of the MHC of the chicken has been explored in three different breeds of chicken: the White Leghorn breed, the Rhode Island Red breed, and a local breed of high interest in disease resistance studies, the Fayoumi breed. This contributed to the identification of two novel *B-F* sequences (sequence 5 and 7), corresponding probably to the *B-F1* genes. The results showed the feasibility of this approach to estimate the correlation between LEI0258 marker alleles and *B-F* genes.

This small scale study illustrates that the sequencing of classical class Ia genes is feasible but involves relatively costly and time consuming methods. Yet, it is a necessary approach to take into account the complexity of the gene organization and the similarity of the different alleles or genes sequences.

Furthermore, the fact that the minor expressed or major expressed genes can be sorted by simple sequence analysis remains to be discussed, as the alignment of sequences obtained with the Hosomichi *et al.*, sequences and the observation of only *B15-BF* genes in the B15/B21 heterozygote indicated that the level of expression might be determined by other factors than the localisation of the genes.

It also shows that the LEI0258 marker is linked to the *B-F* genes, which is not surprising considering the high level of linkage disequilibrium of the region. This has been reported for local Brazilian Caipira chickens (Lima-Rosa *et al.*, 2005). This means that, to some extent, LEI0258 provides indication about the *B-F* genes and can be used as a preliminary screening in any population or breed of chickens, but also that in a well characterized population, the LEI0258 genotype is probably sufficient.

## References

- Bradley RD, Hillis DM (1997) Recombinant DNA sequences generated by PCR amplification. *Mol Biol Evol* **14**:592-593.
- Dalgaard TS, Vitved L, Skjødt K, Thomsen B, Labouriau R, Jensen KH, Juul-Madsen HR (2005) Molecular characterization of major histocompatibility complex class I (*B-F*) mRNA variants from chickens differing in resistance to Marek's disease. *Scand J Immunol* **62**:259-270.
- Ennis PD, Zemmour J, Salter RD, Parham P (1990) Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. *Proc Natl Acad Sci U S A* **87**:2833-2837.
- Fulton JE, Juul-Madsen HR, Ashwell CM, McCarron AM, Arthur JA, O'Sullivan NP, Taylor RL Jr (2006) Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* **58**:407-421.
- Hosomichi K, Miller MM, Goto RM, Wang Y, Suzuki S, Kulski JK, Nishibori M, Inoko H, Hanzawa K, Shiina T (2008) Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. *J Immunol* **181**:3393-3399.
- Lima-Rosa CA, Canal CW, Fallavena PRV, Freitas LB, Salzano FM (2005) LEI0258 microsatellite variability and its relationship to B-F haplotypes in Brazilian (blue-egg Caipira) chickens. *Genet Mol Biol* **28**:386-389.
- Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* **132**:115-130.
- Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, Van Broeckhoven C, De Rijk P (2005) NovoSNP, a novel computational tool for sequence variation discovery. *Genome Research* **15**:436-442.

## Addendum

### *Classical class II genes*

In order to make possible similar strategies for the two classical class IIb genes of the B region of the chicken MHC: *B-LB1* (minor) and *B-LB2* (major), two pairs of primers were designed and tested. The criteria were: amplification of the two genes, no polymorphism within sequences corresponding to the primers, and an overlapping sequence with specific polymorphisms in order to enable sequence reconstruction from the two fragments for any serotype. Two PCR primers pairs were designed: for the first PCR (514 bp) BLB-1F GGCCGTGCTGGTGGCACT and BLB-1R GTACGTCCAGTCCCCGTTCT, and for the second PCR (604 bp) BLB-2F GACAGCGACGTGGGGAAA and BLB-2R GGGGCAGCAGCTAATTCA. The predicted length were based on the reference mRNA sequence of *B-LB2* (GenBank accession number NM\_001044679).

We tested the primers for two animals: one Fayoumi chicken (4605, 381/381 with the LEI0258 marker) and one RIR chicken (4203, 309/309 with the LEI0258 marker).

PCR were carried out from 0.5 µL of cDNA in a final volume of 25 µL containing 2.5 µL 10X buffer, 2 µL dNTP (2 mM each), 0.5 µL of each primer (at 10 µM), and 0.1 U DNA polymerase. The PCR cycle setup was 40 cycles of 94°C for 1 min, 95°C for 1 min, 55°C for 1 min, 72°C for 2 min, followed by 72°C for 10 min. In order to insure specificity, bands at the right size were cut using BandPick (Elchrom Scientific AG) and purified with QIAquick PCR Purification Kit (Qiagen). Direct sequencing was carried out in the forward direction by Eurofins MWG Operon, using their standard protocol for purified PCR products, with the same primers: BLB-1F or BLB-2F.

Four good quality sequences were obtained and aligned with the genomic sequences published by Hosomichi *et al.*, in 2008, the sequences obtained for the animal 4605 matched only *B21-BLB2* and the sequences obtained for the animal 4203 matched *B24-BLB2* but also *B8-BLB2* as they are identical at 99.8% for this fragment. As no polymorphism was found in the primer corresponding sequences for those serotypes, to explain why only the *B-LB2* gene was obtained, this is probably due to the different level of expression of the two genes, *B-LB2* been described as the major expressed locus.

Those preliminary results show that the primers amplify the correct region and that the overlap obtained is sufficient in order to reconstruct the sequence. They also interestingly show that direct sequencing of *B-LB* genes, in those cases, is sufficient in order to obtain the sequence of the major expressed locus.

*From DNA instead of RNA*

The feasibility of cloning and sequencing the same class Ia genes has been also evaluated, for two animals, from genomic DNA (obtained by the rapid extraction procedure). The PCR carried out under similar or optimized conditions, with the same two primers pairs, were non-specific and even after cutting and purifying the bands, no sequence was obtained.

### **C. Discussion et Conclusion**

Le clonage et séquençage des gènes de classe Ia classiques de la région B du poulet a pu être réalisé pour sept animaux provenant de trois races de poulet : la race Fayoumi, la race Rhode Island Red et la race White Leghorn. Cela a contribué à l'identification de deux nouveaux allèles correspondant probablement au gène *B-F1*, minoritairement exprimé. Cette étude exploratoire n'a pas pu être réalisée sur d'autres animaux car les techniques utilisées nécessitent une disponibilité en ARN. Cependant elle a permis de démontrer que les génotypes au niveau du marqueur LEI0258 sont de bons indicateurs des allèles *B-F* présents et que donc le génotypage d'un grand nombre de marqueurs SNPs à proximité de ces gènes peut permettre de les caractériser en première approche.

De premiers résultats encourageants ont été obtenus pour les gènes de classe II classiques.



## **VI SIXIEME PARTIE : MISE AU POINT D'UNE PUCE DE 96 SNPs DEDIEE POUR LA REGION B DU CMH DU POULET**

### **A. Introduction**

En 2004, 2.8 millions de SNPs ont été caractérisés par reséquençage partiel d'un poulet de chair, un animal White-Leghorn et un animal Silkie, et leur comparaison avec le génome du poulet nouvellement assemblé (SNPs du Beijing Genomics Institute, BGI). Cependant, comme illustré dans la deuxième partie de cette thèse, un très petit nombre de SNPs ont été identifiés pour le chromosome 16 et encore moins de SNPs étaient correctement décrits et localisés. C'est ainsi que de nombreux dispositifs QTL, y compris en relation avec la réponse immunitaire, n'ont pas pu être génotypés au niveau du chromosome 16.

La mise au point de SNPs pour le chromosome 16 du poulet nécessite certaines précautions. En raison du grand nombre d'erreur d'assemblage et répétitions de la séquence du génome à ce niveau, nous avons choisi de limiter ces SNPs à une région de 242 kb, bénéficiant d'une annotation détaillée, correspondant à la région B du CMH.

Les SNPs identifiés lors du reséquençage de 11 gènes réalisé pendant cette thèse ont servis de première source de données, ainsi que les SNPs du BGI mais également des SNPs dérivés de la comparaison des 14 séquences de sérotypes B de référence publiées par Hosomichi *et al.*, en 2008.

**Cette publication est en cours de préparation est devra être complétée par les résultats des premiers typages obtenus pour 500 animaux.**



## **B. Article en préparation**

### Development of a 96 SNPs chip dedicated to the chicken MHC B region

Olympe Chazara, Bertrand Bed'Hom

O. Chazara, B. Bed'Hom  
INRA, UMR 1313, Génétique Animale et Biologie Intégrative,  
Domaine de Vilvert, 78352 Jouy en Josas, France  
e-mail: [olympe.chazara@jouy.inra.fr](mailto:olympe.chazara@jouy.inra.fr)  
Tel.: +33-1-34652720  
Fax.: +33-1-34652933

#### **Abstract**

The chicken major histocompatibility complex (MHC) is actually defined by serology or by a single atypical marker, named LEI0258. We report how we succeeded in the development of a dedicated chip of 96 SNPs (Single nucleotide polymorphisms) for the chicken MHC, despite the features of this region, such as duplications and a very high rate of polymorphism. This was done, amongst other things, by applying a strict criteria of unique match for the flanking sequences of each SNP. Only 179 SNPs were therefore selected and submitted to Illumina, leading to the obtention of 96 SNPs mapped in 214,312 pb of the *B* region.

**Keywords:** *Gallus gallus*, SNP, MHC, Illumina GoldenGate

## Introduction

Single nucleotide polymorphisms (SNPs) are polymorphisms concerning only 1 base pair (bp) of the genome sequence but are the most frequently occurring variations described. Thus they facilitate higher resolution and coverage, compare to microsatellite markers. This is especially true for the MHC *B* region of the chicken MHC where only one atypical repeat has been reported for a region at least 240 kb long (Fulton *et al.*, 2006). This is to be linked to the MHC of the chicken being “minimal and essential” (Kaufman *et al.*, 1999), in other words, very compact, with small intergenic and intronic sequences. But the MHC also contains, for the chicken as for other Vertebrates, some of the most polymorphic genes of the genome, such as the class I and class II genes, coding for proteins involved in peptide presentation to T cells. The large number of SNPs available in the MHC could therefore be a solution for typing this region in the chicken. This is of high interest as the MHC of the chicken, especially the *B* region, has been associated with resistance, susceptibility, and differential response to vaccine (Pinard-van der Laan *et al.*, 1998, Lee *et al.*, 2004). This has been demonstrated with serology based-typing, an heavy technique hardly standardizable and which allowed no interbreed comparison.

SNP have been described in the chicken genome since its release in 2004, with the Beijing Genomics Institute (BGI) SNPs discovery project. The highest density of SNPs then reported was for the microchromosome 16, the chromosome where the MHC is located. But this particularity of the region has counterparts, which are genes families and duplications, and small distances between SNPs, especially for sequences showing very high rates of polymorphisms. This is a problem as low cost SNP typing technologies, such as the Illumina GoldenGate, require no polymorphism in the 60 bp flanking sequences of the targeted SNP.

Here we described the methods used to select 96 SNPs, mapped to the *B* region of the chicken MHC, including some we discovered by resequencing a large selection of local, commercial and experimental chicken lines or breeds. The SNPs were genotyped using the Illumina platform BeadXpress, and we provide information on the quality, characteristics and location of the SNPs, and on the genotyping results obtained, such as heterozygosity values and linkage of the markers.

## Materials and Methods

### SNP discovery

Conserved regions along the sequence of the MHC have been identified in order to design primers amplifying 24 fragments corresponding to 11 target genes: *TRIM27.1*, *TRIM41*, *B-BTN2*, *Blec1*, *Tapasin*, *B-DMA*, *B-DMB2*, *B-DMB1*, *TAP1*, *TAP2*, and *CD1A1*. The genes studied were chosen partly from the data already available on their polymorphism (Shiina *et al.*, 2007) but also in terms of technical feasibility.

The 24 fragments of 306 bp to 1179 bp were sequenced at Centre National du Sequencage (CNS), for 96 individuals from 48 populations. This selection included 6 commercial lines (4 broilers lines and 2 brown egg layer lines), 17 White Leghorn populations (15 experimental lines including INRA I.R. lines and a commercial layer line) and 25 other breeds including 12 European breeds, 7 Asian local breeds and 6 African local populations. Each origin was represented by 2 animals, one male and one female when this information was available.

For each one of the 24 PCR, sequence quality and variation analyses were carried out for the 96 individuals with NovoSNP (Weckx *et al.*, 2005). Only the SNPs with a high probability of being accurate were kept for further analyses (NovoSNP's score above or equal to 12). Then SNPs were manually confirmed by looking back at the chromatogram alignment generated by NovoSNP.

### SNPs from other sources

On one hand, all the SNPs located within the *B* region of the chicken MHC, defined by the 2007 sequence of Shiina *et al.*, (AB268588, 241 kb), were extracted from the GenBank database and their polymorphism characteristics compiled from the UCSC server, when available. This was the case for the majority of those SNPs, the BGI SNPs, identified by comparison of a Layer, a Broiler and a Silkie chickens.

On the other hand, the 14 sequences described by Hosomichi *et al.*, in 2008, corresponding to 14 chicken homozygotes for specific *B* region, were aligned with the AB268588 sequence and the positions of the polymorphisms and their genotypes distribution reported using BioEdit (Hall, 1999).

## SNP pre-selection

First, only biallelic SNPs were kept, then each SNP's flanking sequences (100 bp upstream and 100 bp downstream) were blasted against the genome sequence and the AB268588 sequence. Only SNPs with unique hits for both flanking sequences were kept at this stage. Then, only SNP with no other polymorphism (including indels identified with resequencing and Hosomichi *et al.*, (2008) sequences alignment) in the 40 bp around the position were selected.

As some sub-regions were poorly covered, especially regions exhibiting high rates of polymorphisms or repeats, those regions were used to search by BLAST (Altschul *et al.*, 1990) the dbEST section and the nucleotide section of the GenBank database from NCBI. We were able to identify additional SNPs from the obtained alignments.

## SNP selection by Illumina

The SNPs that showed no other SNPs in their 60 bp flanking region, were submitted to Illumina, in order to obtain their designability scores for the GoldenGate technology. As Illumina recommends using SNPs with final scores  $> 0.4$  (Illumina technical note for designing custom GoldenGate genotyping assays, 2009), we took into account the final scores and not the design scores for further selection.

## SNP final selection

After removal of the SNPs showing a final score smaller than 0.4, the remaining SNPs had to be selected in order to keep only 96 SNPs. This was mainly done based on their position on the AB268588 sequence and their spacing. Characteristics related to their known polymorphism, obtained with the resequencing method or the sequences alignments, were used as a last resort.

## Results

### SNPs

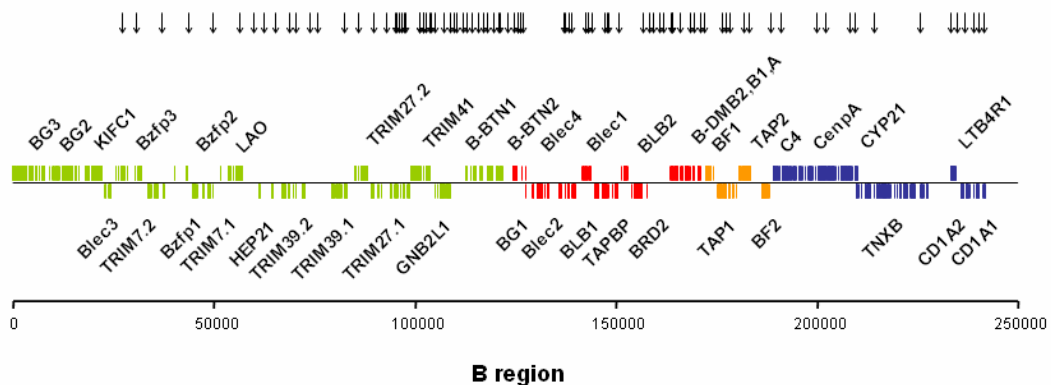
A total of 594 SNPs were identified from the resequencing project, from 20,025 bp analysed, covering not uniformly more than 147 kb. Those SNPs were identified in 48 populations, by comparison with the 4201 SNPs identified from the alignment of the Hosomichi *et al.*, (2008) and the AB268588 sequences, corresponding to 14 serology serotypes, covering 57 kb. But only 238 SNPs matched the selection criteria, covering almost 223 kb.

We were then able to submit 179 SNPs to Illumina. They provided SNP scores (ranging from 0 to 1) and designability rank (which categorized SNP scores as 0 for SNP score <0.4 ; 0.5 for SNP score between 0.4 and 0.6 ; 1 for SNP score >0.6). 5 SNPs were assigned a rank of 0 (very low designability), corresponding to scores between 0.245 and 0.372, 28 SNPs a rank of 0.5 (moderate designability), corresponding to scores between 0.400 and 0.599, and 146 SNPs a rank of 1 (highly designable), corresponding to scores between 0.600 and 0.993.

We kept 96 SNPs, with scores ranging from 0.450 to 0.993, covering 214,312 pb, this represents, on average, 1 SNP every 2232 bp. In fact, the largest spacing was 11,763 bp, due to regions where no SNPs respecting the criteria used could be identified, especially the no duplication criteria (unique flanking sequences) (Figure 1).

Those regions corresponded to gene families such as the *B-G1*, *B-G2* and *B-G3* genes, the *B-LB1* and *B-LB2* genes and the *B-F1* and *B-F2* genes, but also to genes with repeated structures such as the *C4* gene and the *TNXB* gene.

**Figure 1.** Schema of the 242 kb *B* region of the chicken MHC showing gene content and positions of the 96 SNPs



## **C. Discussion**

Actuellement les SNPs ont été choisis et validés par Illumina pour un génotypage par technologie GoldenGate sur BeadExpress. Tous ces SNPs sont précisément localisés dans la région et pour la plupart, on dispose d'informations complémentaires, telles que des informations de polymorphisme et de contexte génomique. Un premier lot de 500 animaux va être génotypé rapidement et vont permettre d'évaluer la qualité des SNPs choisis. Cette évaluation permettra de compléter la courte publication prévue pour accompagner le développement de cet outil. Ces animaux sont des animaux de sérotypes de références, conservés au laboratoire, par J. Fulton ou H. Juul-Madsen, ainsi que des animaux de dispositifs INRA de challenge infectieux, anciens ou en cours.



## DISCUSSION GENERALE & PERSPECTIVES

Les objectifs de ce travail de thèse étaient d'explorer la variabilité génétique du CMH chez le poulet, d'améliorer les connaissances sur cette région si particulière du génome, de mettre au point et de tester les limites du génotypage moléculaire de la région B, en tenant compte de la grande variabilité des races, populations ou lignées existantes, et d'évaluer les possibilités d'implémentation de ces techniques dans des dispositifs expérimentaux. Différentes techniques moléculaires ont été utilisées, afin de considérer les différents aspects de cette problématique et ses différentes échelles, de la région B dans son ensemble, avec 242 kb décrits actuellement, à des sous-régions plus spécifiques, voire des locus simples. Il en découle une assez importante quantité de données, permettant de nombreux types d'analyses, le CMH se prêtant particulièrement bien à l'identification de polymorphismes en raison de sa relativement grande diversité.

### **I. APPORT DES DIFFERENTES STRATEGIES CHOISIES POUR L'ETUDE DU CMH DU POULET**

#### **Construction d'un référentiel basé sur le marqueur LEI0258**

En premier lieu, il a été démontré que le marqueur LEI0258 était informatif chez tout type de poulets, y compris les différentes espèces sauvages proches. L'analyse approfondie de ce marqueur a mis en lumière des caractéristiques atypiques et particulièrement intéressantes d'un point de vue évolutif. Son utilisation est donc pertinente pour un génotypage du CMH chez le poulet, notamment grâce à l'existence du référentiel établi par Fulton *et al.*, 2006, que nous avons complété lors de cette thèse. Ces deux référentiels, en particulier celui de Fulton *et al.*, récapitulent les correspondances entre les allèles du LEI0258 et les sérotypes caractérisés au niveau mondial, ce qui devrait grandement améliorer les possibilités de comparaison entre études et entre populations. Ainsi, l'utilisation systématique de ce marqueur a été testée dans des lignées expérimentales dédiées à l'étude de la réponse immunitaire. Le suivi des génotypes à ce marqueur permet un premier niveau de suivi de la région B du CMH chez les animaux et accompagne l'abandon des techniques de sérologie. L'information apportée par la sérologie et le typage moléculaire n'est évidemment pas identique mais ce dernier permet d'y suppléer tout en gardant un bon niveau de contrôle de la ségrégation des allèles. Il présente



également des avantages techniques évidents tels que la rapidité, la pérennité de la méthode (sans le problème des réactifs de sérologie) et donc un contrôle *a posteriori* possible, ainsi qu'une plus grande robustesse (moindre sensibilité au facteur expérimentateur). De plus, les différents allèles de ce marqueur ont pu être classifiés sous forme d'un réseau basé sur les polymorphismes des régions flanquantes, et de ce fait proposer des regroupements des différents haplotypes CMH qui leur sont associés. Cependant ces regroupements ne permettent pas de faire d'inférence fonctionnelle concernant les haplotypes associés.

Enfin, l'étude des lignées expérimentales a permis d'évaluer l'apport d'une puce SNPs 20K qui comprenait 10 SNPs de la région du CMH. Cette étude pilote a montré que le typage SNP pouvait également être utilisable pour caractériser la région B et que ces marqueurs permettaient d'y évaluer l'existence de recombinaisons et de conduire des analyses de type phylogénétique.

Le développement d'un plus grand nombre de SNPs dans cette région avait donc un intérêt certain et a été entrepris par une approche de séquençage systématique de gènes choisis dans la région B du MCH.

### **Polymorphisme de gènes**

Un large panel d'animaux a été défini, composé de deux animaux par population échantillonnée. En effet, le but était d'explorer la variabilité de la région et d'identifier un maximum de variations et non pas de caractériser les populations. D'une façon qui peut paraître contradictoire *a priori*, les gènes de classe I et de classe II classiques ont été écartés de la liste des gènes candidats au reséquençage. En effet, leur exploration, dans des populations non caractérisées pour le CMH est rendue difficile par l'existence de deux gènes très similaires pour chacun de ces deux types qui rend l'interprétation des données de séquençage classique inextricable.

De plus, la plupart des variations observées au niveau des gènes de classe I et de classe II classiques sont d'ordre inter-individuel, avec de nombreuses homoplasies possibles, ce qui n'en fait pas de bons candidats pour la définition de grands groupes d'haplotypes du CMH.

Le choix des gènes cibles du reséquençage s'est donc fait sur des critères de positionnement et de variabilité en utilisant notamment les résultats publiés par Shiina *et al.*, 2007 issus de la comparaison de la séquence disponible du génome, c'est à dire d'origine Coq de Bankiva et

de séquences disponibles issues d'animaux d'haplotypes B12. Les gènes présentant un fort taux de SNP et/ou un faible rapport dS/dN ont été privilégiés. Les résultats de séquençage obtenus ont été particulièrement bons, puisque sur les 34 cibles proposées dans un premier temps, 33 ont pu être amplifiées et séquencées dans des conditions standards de haut-débit, ainsi qu'une cible remplaçante. Au final, les 11 gènes sélectionnés ont été séquencés, avec 2 PCR par gène au moins pour 9 de ces gènes, permettant une bonne couverture de leur séquence codante. Une grande quantité de données a été obtenue lors de cette étude, et si l'objectif premier d'identification de polymorphismes a pu être atteint, une analyse fine des différents génotypes n'a pu être réalisée que pour trois gènes de classe II non classiques, ainsi que présentée dans cette thèse, mais également pour les deux gènes situés aux extrémités de la région analysée, les gènes *CD1A1* et *TRIM27*.

L'étude ciblée sur les trois gènes *B-DMA*, *B-DMB1* et *B-DMB2* a permis de caractériser ces trois gènes, de tester leur expression et leur polymorphisme afin de les définir comme des gènes de classe II non-classiques de type DM, par homologie avec les gènes *DM* des Mammifères et d'apporter des indications sur leur fonctionnalité potentielle. C'est également une comparaison avec les Mammifères qui a permis, dans un second temps, de démontrer le caractère ancestral de ces gènes et donc de renforcer l'hypothèse selon laquelle ils feraient partie du CMH « primordial ». Enfin, l'étude du polymorphisme de ces gènes a été réalisée dans un large panel de races expérimentales, commerciales et locales de poulet et a permis d'établir une classification en 22 clusters des différents haplotypes identifiés et donc pour la première fois de proposer une classification des haplotypes du CMH existants chez le poulet, basée sur une partie de leur région de classe II.

En complément du projet de reséquençage de 11 gènes, une étude exploratoire des deux gènes de classe I classiques de la région B a été menée. Cette étude avait pour but de tester les possibilités mais également les limites de ces techniques pour des populations dont le CMH n'avait jamais été caractérisé. Il est apparu que les techniques actuellement utilisées dans un laboratoire familial de l'étude du CMH sont réalisables, mais sous certaines conditions. En effet, des résultats encourageants ont pu être obtenus à la condition d'utiliser des animaux homozygotes au marqueur LEI0258 et de travailler à partir d'ARN. Ces critères limitent donc la possibilité d'étendre cette approche à une échelle plus large et pour des échantillons plus diversifiés. Cependant l'accès aux informations de séquence des gènes de classe I mais également de classe II classiques va s'avérer indispensable, du moins pour certains génotypes

**Tableau 6.** Récapitulatif des informations recueillies pour vingt principaux sérotypes ou haplotypes

| Sérologie de référence | Races <sup>a</sup> | Autres sérologies possibles | Autres races possibles | Présence en lignées com. | LEI0258  | Cluster LEI0258 | GAB0001   | Cluster B-DM | Fréquence exp. | Fréquence autres races | MHC (séquence) <sup>b</sup> | Réfs <sup>c</sup> |
|------------------------|--------------------|-----------------------------|------------------------|--------------------------|----------|-----------------|-----------|--------------|----------------|------------------------|-----------------------------|-------------------|
| B2                     | WL                 |                             |                        | Oui                      | 261      | 7               | 161*      | 17d          | ++             | ?                      |                             | H. <i>et al.</i>  |
| B4                     | WL                 |                             |                        |                          | 182, 183 | 1               | 187       | 15b          | +              | +                      | B4 ~ B13                    |                   |
| B5                     | WL                 |                             |                        | Oui                      | 295      | 11              | 161*      | 16b          | +              | ++                     | B5 ~ B8/B11                 | H. <i>et al.</i>  |
| B6                     | WL                 |                             |                        |                          | 443      | 6               | 160*      | 12b          | +              | +                      |                             | H. <i>et al.</i>  |
| B8                     | ANC                |                             |                        |                          | 405      | 5               | 161*      | 16b          | +              | +                      | B8 = B11                    | H. <i>et al.</i>  |
| B9                     |                    | B22?, B73?                  | NH, WPR                | Oui ?                    | 249      | 5               | 161*      | 19b          | +              | ++                     |                             | H. <i>et al.</i>  |
| B11                    | ANC                |                             |                        |                          | 193      | 3               | 161*      | 16b          | +              | ++                     | B8 = B11                    | H. <i>et al.</i>  |
| B12                    | WL                 |                             |                        |                          | 487      | 6               | 159       | 6            | ++             | .                      | B12 = B19                   | H. <i>et al.</i>  |
| B13                    | WL                 |                             |                        | Oui                      | 205      | 3               | 187       | 15b          | +              | ++                     | B4 ~ B13                    | H. <i>et al.</i>  |
| B15                    | WL                 |                             |                        | Oui                      | 261      | 7               | 174       | 20b          | ++             | ?                      |                             | H. <i>et al.</i>  |
| B17                    | WL                 |                             |                        |                          | 205      | 4               | 174*      | 13a          | +              | ++                     |                             | H. <i>et al.</i>  |
| B19                    | WL                 |                             |                        | Oui                      | 539, 552 | 6               | 159       | 6            | ++             | +                      | B12 = B19                   | H. <i>et al.</i>  |
| B21                    | WL                 | B130, B131                  | WC                     | Oui                      | 357      | 5               | 173*, 174 | 3a           | +++            | +++                    | B21                         | H. <i>et al.</i>  |
| B23                    | NH                 |                             |                        | Oui                      | 357      | 7               | 159*      | 22b          | +              | +                      |                             | H. <i>et al.</i>  |
| B24                    | NH, RIR, WPR       | B26?, B76?                  |                        | Oui                      | 309      | 5               | 159*      | 4a           | +              | +++                    |                             | H. <i>et al.</i>  |
| n.a.                   |                    | B34                         | WL                     |                          | 420      | 6               | 174       | 8            | .              | .                      |                             |                   |
| n.a.                   |                    | B72?, B78?                  |                        | Oui ?                    | 307      | 12              | 187       | 12c          | .              | ++                     |                             |                   |
| n.a.                   | RJF, CN            | BQ, BW1                     |                        |                          | 369      | 5               | 173       | 3a           | .              | .                      | B21                         | Génome            |
| n.a.                   |                    | n.a.                        |                        |                          | 381      | 5               | 174       | 3a           | .              | ++                     | B21                         |                   |
| n.a.                   |                    | n.a.                        |                        |                          | 217      | 4               | n.a.      | n.a.         | .              | ++                     |                             |                   |

ANC, Ancona; CN, Black Castellana; NH, New Hampshire; RIR, Rhode Island Red; RJF, Red Jungle Fowl; WC, White Cornish; WL, White Leghorn; WPR, White Plymouth Rock

\* Déduit des séquences publiées par Hosomichi *et al.*, 2008

a, Dans Fulton *et al.*, 2006

b, Synthèse des informations de séquences connues (11 gènes reséquencés, Hosomichi *et al.*, 2008, clonage des gènes de classe Ia)

c, H. *et al.*, Hosomichi *et al.*, 2008

? pour B2 et B15 car ces deux sérologies correspondent au même allèle LEI0258, de 261bp qui est très fréquent

d'intérêt. Les autres techniques utilisables semblent être actuellement les techniques de PCR Long-Range, avec notamment des couples d'amorces publiés par Hosomichi *et al.*, en 2008. Dans le cadre de cet article, les PCR ont été séquencées par technique de Sanger classique en utilisant un certain nombre d'amorces internes pour le séquençage. Les techniques de seconde génération de séquençage vont probablement permettre de simplifier cet aspect mais n'ont pas encore fait leur preuve dans ce contexte. Une autre technique récente devrait pouvoir être appliquée avec succès pour des régions telles que le CMH du poulet, il s'agit de la technique de capture de séquence couplée à un séquençage haut-débit actuellement en cours de mise au point à l'INRA (Projet ANR CapSeqAn, B. Bed'hom, com. pers.).

Enfin, les différents polymorphismes de type SNPs identifiés au cours de la thèse, combinés aux différentes sources de données disponibles pour la région B du CMH, ont permis l'établissement d'un panel dédié de 96 SNPs. Les différentes données et observations faites au cours de la thèse ont grandement aidé à l'établissement de ce panel et à la phase de choix des marqueurs. Ce panel va être testé dans un premier temps sur 500 animaux, les 96 animaux utilisés au cours de la thèse pour le reséquençage, des animaux références pour le CMH au niveau international, mais également différents animaux impliqués dans des dispositifs d'épreuve par un agent pathogène, notamment la grippe aviaire (races locales taiwanaises), la maladie de Marek (lignées commerciales) et la coccidiose (races Fayoumi et White Leghorn).

Afin de résumer les différentes informations recueillies sur vingt principaux sérotypes ou haplotypes, un tableau récapitulatif a été établi (Tableau 6).

Ces différents résultats peuvent être discutés à la lumière de deux grandes thématiques, la première étant l'évolution des différents locus composant cette région si particulière du génome qu'est le MHC, la deuxième étant les différentes applications pouvant découler de l'identification de facteurs génétiques influençant la réponse immunitaire chez le poulet.

## **II. L'ETUDE DU CMH COMME MOYEN D'ANALYSE DE LA DIVERSITE GENETIQUE CHEZ LE POULET**

### **Difficultés liées aux particularités du CMH du poulet**

Le CMH du poulet se situe sur un microchromosome, le chromosome 16. Comme les autres microchromosomes, il est riche en gènes et difficile à cloner sous forme de BAC (Bacterial Artificial Chromosome), ce qui est à mettre en relation avec la proximité des télomères et du centromère, due à la petite taille (estimée à 12-15Mb pour le 16), mais également avec un fort taux de GC caractéristique de certains de ces chromosomes. En 1999, la première séquence d'un BAC contenant environ 90kb de la région B du CMH d'un poulet B12 a été publiée par Kaufman *et al.*. Il semblerait que les autres BAC produits, séquencés et assemblés depuis, dans d'autres laboratoires, à partir d'autres animaux, correspondent également à cette région. En 2007, une carte étendue a été proposée par Shiina *et al.*, pour un poulet provenant de la même lignée consanguine Coq de Bankiva que l'individu utilisé pour la séquence du génome. Cette carte correspondait à trois BAC et deux PCR Long-Range, couvrant 242 kb la région B du CMH du poulet. Par comparaison avec les efforts déployés, les particularités du chromosome 16 freinent son séquençage, ce qui est également observé pour les autres microchromosomes, ainsi les séquences courtes du génome du poulet obtenues avec un séquençage *de novo* grâce à un séquenceur Roche 454 GS FLX, ne semblent pas contenir de nouvelles zones (B. Bed'hom, com. pers.). Si des BAC correspondant à la région Y du CMH du poulet existent et ont pu être séquencés, aucune séquence assemblée n'a été proposée, ce qui est à mettre en relation avec une probable duplication voir triplification de cette région, ainsi qu'à un très grand nombre de séquences répétées. De ce fait, l'étude du CMH du poulet concerne principalement la région B, ce qui est probablement suffisant dans une certaine mesure, puisqu'elle contient des régions de classe I, II, III et de classe I étendue comparables avec celles des Mammifères. Cependant l'existence de la région Y ne doit pas être ignorée, elle contient également des gènes considérés comme centraux dans l'organisation du CMH et qui sont exprimés.

En ce qui concerne la région B, étudiée au cours de cette thèse, elle présente également des caractéristiques rendant complexe l'utilisation des techniques actuelles de biologie moléculaire, notamment un fort taux de GC et des familles de gènes présentant de fortes

similitudes. Ces particularités ont rendu certaines stratégies impossibles ou compliquées. Ainsi, chaque conception d'amorces d'amplification est délicate et les essais de PCR Long-Range n'ont pas été poursuivis, cette technique favorisant les amplifications non-spécifiques. De même les amorces conçues pour l'amplification des gènes de classe I n'ont pas pu être utilisées pour les animaux utilisés pour le reséquençage dont seul l'ADN était disponible.. De plus, les animaux étudiés, choisis pour représenter la diversité de l'espèce, étaient pour la plupart hétérozygotes au CMH. Si cette caractéristique est théoriquement idéale pour l'observation de polymorphismes, elle ne facilite pas l'exploitation des données du séquençage direct. En effet peu d'outils informatiques sont capables de détecter les pics de séquençage hétérozygotes et surtout de les différencier des pics artefactuels obtenus en présence d'une insertion ou d'une délétion à l'état hétérozygote, qui décale une séquence par rapport à l'autre, ou des artefacts obtenus à proximité des homopolymères. Pour ces raisons, la validation visuelle de chaque génotype, permettant de définir les génotypes individuels, n'a été réalisée que pour 5 gènes (les trois gènes *B-DM*, le gène *TRIM 27.1* et le gène *CD1A1*).

La stratégie choisie, en accord avec les membres du comité de thèse, de privilégier un grand nombre de source d'échantillons, avec peu d'individus par source, a favorisé l'identification de SNPs et le recueil d'informations sur la région et non pas la caractérisation individuelle des animaux. En raison de ce choix, les génotypes individuels obtenus n'ont pas permis d'analyse de type populationnelle. Celles-ci n'ont pu être réalisées que pour le marqueur LEI0258. La reconstruction des haplotypes s'est également avérée problématique car elle s'accommode mal des positions hétérozygotes et des allèles rares, or les individus hétérozygotes ont tendance à l'être à toutes les positions polymorphes et la grande diversité existante implique de nombreux allèles rares. Cependant, une autre caractéristique observée au cours de la thèse favorise la reconstruction d'haplotype, c'est le fait que les différents haplotypes, même rares, puissent être présents dans deux populations, même présentant des histoires distinctes, par exemple des origines géographiques très différentes.

Ainsi la région du CMH comporte des particularités structurales à prendre en compte dans les différents choix technologiques, afin de pouvoir répondre aux objectifs prédéfinis. La collaboration avec d'autres équipes travaillant également sur le CMH a ainsi été déterminante afin de ne pas commettre d'erreurs stratégiques tout en développant de nouvelles procédures.

## **Les différents intérêts présentés par l'étude du CMH**

Une fois les différents problèmes techniques contournés ou pris en compte, la région du CMH est une région particulièrement intéressante à étudier. En premier lieu, il y est facile de trouver des polymorphismes de séquences, notamment au sein de la diversité des races, populations ou lignées de poulet existantes, alors qu'aucun marqueur de type microsatellite classique n'a pu être mis au point pour le chromosome 16. L'identification de SNPs est donc d'autant plus nécessaire. Il est également possible de disposer d'informations sur la région à différents niveaux, d'autres équipes travaillant sur la caractérisation de gènes d'intérêt, sur des dispositifs de challenge infectieux et sur la description de races ou de lignées particulières. Comme discuté précédemment, la disponibilité de la séquence des deux régions composant le CMH du Poulet et leur cartographie n'est pas encore résolue mais des moyens sont actuellement développés pour les résoudre, et le microchromosome 16, malgré la présence de la région NOR, est l'un des microchromosomes les plus étudiés du poulet. L'étude du CMH du poulet nécessite donc d'interagir avec des équipes de recherche aux motivations et compétences différentes, ce qui est toujours enrichissant : chercheurs en immunologie, en génétique des populations, en génétique moléculaire ou en génétique quantitative.

La présence de famille de gènes fortement similaires, même si elle ne facilite pas l'étude du CMH, permet d'étudier et de comparer ces gènes entre eux, chez le poulet, ou interspécifiquement, avec ceux d'autres oiseaux ou ceux des Mammifères. Cela permet de comprendre l'histoire d'une famille de gènes comme on l'a vu avec les gènes de classe II non classiques *B-DM* et les gènes de classe I classiques du poulet, ce type d'étude permet de proposer des hypothèses sur l'état « primordial » du CMH chez les Gnathostomes.

Un autre intérêt du CMH est qu'il est considéré comme étant, quel que soit l'organisme, soumis à sélection. Le poulet étant un animal d'intérêt agronomique, il est possible d'étudier différents types de situations, de l'espèce sauvage à la lignée expérimentale sélectionnée pour un type de CMH, via la race locale non sélectionnée, la race en conservation, la lignée commerciale sélectionnée ou non pour le CMH et la lignée expérimentale sélectionnée pour des critères relatifs à la réponse immunitaire. Ainsi, un certain nombre de races, populations ou lignées abordées au cours de cette thèse, notamment grâce à l'existence de la banque d'ADN issue du projet européen AvianDiv, avaient également été étudiées à l'aide de marqueurs neutres, ce qui a permis de comparer les hétérozygoties observées.

### III. L'EVOLUTION DU CMH

La majorité des études relatives à la diversité du CMH concerne les Mammifères et en particulier l'Homme et la Souris. Cependant, quel que soit le modèle utilisé, le même type d'observation est récurrent, il s'agit de la mise en évidence de forces de sélection favorisant de haut taux de polymorphismes au niveau des gènes du CMH et notamment des signatures de sélection équilibrante, favorisant les hétérozygotes. Deux grands types de mécanismes peuvent expliquer le maintien voire l'accroissement de ces taux de polymorphismes: des mécanismes en relation avec la résistance aux maladies et des mécanismes impliqués dans la reproduction (Bernatchez & Landry, 2003).

Dans la première catégorie peuvent se classer les phénomènes de sélection équilibrante (*balancing selection*) mais également les phénomènes de sélection des allèles rares (*rare-allele advantage*). Les hypothèses sous-jacentes à la sélection équilibrante sont que les hétérozygotes au CMH peuvent présenter plus d'antigènes (car ils possèdent un plus grand nombre d'allèles au niveau des gènes impliqués, c'est à dire les gènes de classe I et de classe II) et donc pourraient reconnaître et traiter une plus grande variété de type de pathogènes que les individus homozygotes (Penn *et al.*, 2002). L'environnement pathogénique étant fluctuant, spatio-temporellement, aucun génotype ne serait toujours favorisé. Le terme «sélection équilibrante» est en général associé à ce type de phénomènes et c'est celui qui est le plus fréquemment mentionné en ce qui concerne le CMH. Ces hypothèses sont particulièrement intéressantes au regard du petit nombre de gène de classe I et II présents chez le poulet, et correspondent au concept de CMH minimal et essentiel (Kaufman *et al.*, 1999).

Le deuxième principe pouvant également être considéré comme sélection équilibrante est le principe de l'avantage de l'allèle rare : les pathogènes cherchent constamment à s'adapter aux allèles les plus fréquents des populations et donc les allèles rares et nouveaux échapperaient à cette «course à l'armement»(Takahata et Nei 1990). Il en découle que les allèles rares pourraient devenir fréquents et donc induire des dynamiques plutôt cycliques des fréquences alors que la sélection équilibrante plus classique conduit à polymorphisme élevé mais stable (Slade et McCallum 1992).

Dans la seconde grande catégorie de mécanismes, ceux impliquant la reproduction sexuée, on retrouve des principes favorisant la reproduction entre individus présentant des MHC différents, notamment en relation avec les odeurs individuelles et leur perception par les



récepteurs olfactifs, situés à proximité des gènes du CMH chez les Mammifères, cependant, aucun récepteur olfactif n'a été mis en évidence sur le chromosome 16 chez le poulet. On doit également envisager les mécanismes défavorisant les produits homozygotes au CMH, notamment l'accumulation d'allèles récessifs (délétères à l'état homozygote) par effet d'autostop dans les régions proches du CMH (van Oosterhot 2009).

Les observations réalisées lors de ce travail de thèse sont à mettre en perspective par rapport à ces différentes forces évolutives, sachant qu'elles agissent probablement toutes à différents degrés. Les résultats obtenus avec la comparaison des hétérozygoties observées au marqueur LEI0258, lié au CMH, et les hétérozygoties précédemment rapportées avec des marqueurs microsatellites pouvant être considérés comme neutres, montre une bonne corrélation entre ces valeurs pour les populations considérées comme non soumises à sélection artificielle pour le CMH, s'accompagnant cependant de valeurs d'hétérozygotie plus élevées pour la région du CMH. Ceci indique donc que le CMH est bien plus polymorphe que les autres régions du génome. Les autres observations et notamment le partage des allèles, même rares, entre les différentes populations, ne font pas du CMH un bon locus pour l'étude des différenciations entre populations. Ceci correspond à ce qui avait pu être prédit sous sélection équilibrante, la différenciation entre populations est réduite par rapport aux marqueurs neutres, en raison de deux phénomènes : d'une part, le maintien des différents allèles à des fréquences proches et donc une diversité intra-population importante par rapport à la diversité totale, et, d'autre part, la rapidité potentielle d'introgression d'un allèle positivement sélectionné dans une population en l'absence de gestion (Bernatchez et Landry 2003).

Par ailleurs, les résultats obtenus chez les espèces sauvages du genre *Gallus* s'apparentent à ce qui a pu être observé chez l'Homme et le Chimpanzé par Parham *et al.*, 1989, c'est à dire le partage d'un grand nombre d'allèles et donc une origine ancestrale de la diversité contemporaine. L'observation d'allèles « ancestraux » valide plutôt une sélection équilibrante classique avec le maintien d'un grand nombre d'allèles. Cependant, l'observation d'une grande diversité d'allèles chez les poulets domestiques avec potentiellement de nouveaux allèles indique également une importante dynamique de la région.

En ce qui concerne les mécanismes impliquant la reproduction, le poulet présente, parmi les animaux d'élevage, des temps de génération relativement courts et potentiellement des familles de taille importante, il pourrait donc être un bon modèle d'étude. Au cours de cette

thèse, un suivi de plusieurs générations de lignées expérimentales présentant une bonne variabilité au CMH a été réalisé, et pourrait être exploité pour étudier la problématique de la proportion d'hétérozygotes à la naissance dans une portée.

Enfin, globalement, les résultats de cette thèse sont cohérents avec les études présentant le CMH du Poulet comme minimal et compact puisque malgré la diversité des animaux étudiés, aucune variation dans le nombre de gènes n'a été mise en évidence. Cependant, ces variations potentielles n'ont pas été spécifiquement recherchées, donc elles ne sont pas exclues.

#### IV. RESISTANCE AUX MALADIES ET CMH CHEZ LE POULET

L'obtention d'animaux en bonne santé dans les élevages ne peut être obtenue que par un faisceau de moyens, comprenant les bonnes pratiques d'élevages, la vaccination et des traitements vétérinaires mais également la prise en compte du potentiel génétique des animaux. La composante génétique n'est donc qu'un des facteurs à considérer, cependant elle représente une alternative ou un complément à l'utilisation de molécules anti-infectieuses (antibiotiques et antiparasitaires), une solution potentielle à la résistance des pathogènes aux traitements et une solution intéressante pour les pays du Sud où la mise en oeuvre des mesures sanitaires est souvent difficile. Toutefois les différents aspects que peuvent recouvrir le terme résistance, et notamment le phénomène des porteurs sains, doivent être pris en compte ainsi que la nécessité de préserver la diversité génétique dans toute stratégie à long terme.

Deux grands types de travaux de recherche étudient cet aspect génétique de la résistance, chez les animaux d'élevage : ceux se concentrant sur la réponse à un pathogène particulier et ceux évaluant la réponse immunitaire d'une façon globale. En ce qui concerne les facteurs génétiques impliqués, ils sont de plusieurs types, des gènes particuliers, tel le gène Mx ou des groupes de gènes comme le CMH ou simplement des marqueurs liés à des caractères d'intérêt. Ce que ces diverses études ou expériences de sélection nous ont appris, c'est que les caractères liés à la résistance aux maladies sont complexes, avec différents aspects tels que les taux d'anticorps, la réponse cellulaire, les taux de cytokines, l'efficacité de phagocytose... Ceci est à mettre en parallèle avec les différents pathogènes concernés, bactéries, virus ou parasites. Chez le poulet, différentes lignées ont pu être sélectionnées avec succès, dont celles étudiées dans le cadre de cette thèse, comme les lignées R.I., sélectionnées en parallèle, chacune sur un critère unique (Pinard-van der Laan *et al.*, 1998) et dont le croisement peut présenter un intérêt pour certains de ces critères (Minozzi *et al.*, 2008), mais également les lignées résistances ou susceptibles au Sarcome de Rous (dérivées de la lignée Cornell K, maintenues en Allemagne chez S. Weigend) ou les lignées divergentes pour le taux d'anticorps (maintenues à Wageningen, Pinard *et al.*, 1992).

Dans de nombreux cas, le rôle du CMH a été pris en compte, notamment quand ses effets forts ont été mis en évidence pour différents pathogènes ou caractères chez le poulet. Cela ne signifie pas qu'il joue un rôle central dans tous les aspects de résistance aux maladies chez le

poulet, mais cela implique qu'il soit systématiquement pris en compte dans tous ces types d'études chez le poulet. C'est probablement le manque d'outils moléculaires disponibles et standardisés qui a freiné cet aspect dans de nombreux dispositifs expérimentaux. On est en droit d'espérer que la mise au point récente de marqueurs, à laquelle ce travail de thèse a contribué, va permettre d'améliorer grandement la prise en compte de la variabilité génétique du CMH dans ces études, permettre une meilleure connaissance des lignées commerciales et la prise en compte des particularités des populations locales, et faciliter la comparaison des résultats obtenus.

Le cas des populations commerciales montre l'effet d'une sélection directionnelle sur le CMH, visant à l'homogénéisation du génotype intra-lignée. Cet objectif est indirectement lié au choix d'un haplotype et donc d'un pathogène, ou d'un groupe de pathogènes, prioritaire. Ainsi, les sélectionneurs de poules pondeuses ont toujours visé la résistance à l'herpesvirus de la maladie de Marek. Ce choix est susceptible de diminuer la capacité des lignées à réagir vis-à-vis d'autres agents pathogènes et ne paraît pas favorable à l'obtention d'une robustesse générale des lignées. Les pondeuses commerciales sont issues d'un croisement à 2 étages entre des lignées grand-parentales peu variables pour le CMH, il serait donc souhaitable que ces lignées portent des haplotypes différents, afin que le croisement terminal présente une certaine variabilité au CMH. Les sélectionneurs disposent maintenant des outils permettant de suivre plus facilement le polymorphisme du CMH, encore faut-il que leurs populations de base aient encore une certaine variabilité au CMH.

La disponibilité en outils de typage moléculaire pour le CMH devrait également permettre de mieux évaluer les ressources génétiques disponibles à ce niveau chez les poulets domestiques. C'est une étape nécessaire aux prises de décision concernant la conservation mais également l'utilisation de ces ressources. Il reste à espérer qu'en faveur de stratégies de production à long terme, ces évaluations auront pour effet de préserver la diversité génétique, voire de l'accroître, que ce soit pour les populations locales ou pour les stocks commerciaux.

## **V. PERSPECTIVES**

### **Evaluation de la diversité génétique du CMH chez le poulet**

Il ressort des résultats de la thèse que le marqueur LEI0258 est un bon indicateur du type de CMH présent, en ce qui concerne la région B et un bon indicateur de l'hétérozygotie des gènes de cette région, en particulier lorsque ses allèles sont séquencés. L'étude assez globale réalisée permet donc d'avoir une bonne estimation de la diversité génétique présente à ce niveau chez le poulet.

Dans le cas de certaines populations ou lignées présentant un intérêt particulier pour d'autres travaux du laboratoire, cette étude sera probablement complétée par le génotypage d'un plus grand nombre d'animaux par exemple ou par un séquençage plus étendue des séquences flanquant la zone répétée. Certains haplotypes différents de CMH étant maintenant connus pour partager les mêmes allèles ou des allèles proches au niveau du LEI0258 (par exemple B21 et B23, B2 et B15, ou les allèles 307 et 309), certains cas indéterminés bénéficieront du typage du marqueur GAB001, situé dans la région de classe II. L'accès récent à une plateforme de génotypage pour des typages en moyen débit a permis la mise en place d'une procédure et d'un protocole permettant le typage en duplex des deux marqueurs.

La publication du marqueur GAB001 via l'article BM pourra permettre sa diffusion, le partage des informations de typage est en cours avec le laboratoire d'H. Juul-Madsen.

Les premiers résultats du typage du premier lot de 500 animaux à l'aide de la puce dédiée de 96 SNPs vont permettre d'évaluer rapidement la qualité de cette puce et de publier ses caractéristiques. Parmi les 500 animaux, un grand nombre, provenant des échantillons conservés par J. Fulton ou dont dispose H. Juul-Madsen, seront des échantillons définis en sérologie, afin d'établir les corrélations avec cette méthode de typage. Parmi ces échantillons, certains n'étaient pas disponibles pour le projet de thèse.

### **Connaissance de la structure de la région B du CMH du poulet**

Actuellement il n'existe pas vraiment de consortium pour l'étude du CMH chez le poulet. Les différentes équipes sont en compétition bien que des accords tacites existent. Il est donc difficile de savoir qui est en train de caractériser quel gène. Ainsi, M. Miller et ses collaborateurs ont un fort intérêt pour les gènes *B-G* et leur récent article publié démontrant

une association entre une insertion rétrovirale dans la partie 3' UTR du gène *B-G1* et la résistance au virus de Marek va probablement accroître l'intérêt porté à ces gènes. C'est également la première fois qu'une fonction est proposée pour un gène *B-G* alors que cette famille de gènes est une particularité remarquable du poulet et joue un rôle dans la détermination de la sérologie. Cependant le nombre exact de ces gènes chez le poulet n'est pas défini et ils présentent une structure complexe avec de nombreux exons très ressemblants, un projet d'étude devrait donc se faire en collaboration avec des laboratoires déjà expérimentés. D'autres gènes présents dans la région B du CMH du poulet pourrait faire l'objet de projets intéressants, comme les gènes *TAP1* et *TAP2*, en relation avec les gènes de classe I et les hypothèses de co-évolution proposées (Kaufman 2000), le gène *HEP21* exprimé dans l'oviducte ou les gènes *B-BTN1* et *B-BTN2* aux caractéristiques intéressantes mais à la structure complexe.

### **Association entre haplotypes du CMH et résistance aux maladies infectieuses**

Les nouvelles connaissances acquises sur la région B du CMH du Poulet ont permis de définir de nouveaux outils de génotypage et de caractériser des gènes pouvant jouer un rôle dans la détermination de la composante génétique de la réponse immunitaire. Cependant, le premier obstacle à la détermination fine des zones du MHC en rapport avec un phénotype est l'exceptionnel déséquilibre de liaison régissant la région. Il est ainsi actuellement impossible d'isoler les effets d'une mutation, on ne peut connaître que les effets d'un haplotype complet. Dans de nombreux cas, la détermination des haplotypes au CMH des animaux utilisés dans les challenges expérimentaux sera déjà un progrès, notamment quand l'on sait que certains dispositifs ont été phénotypés avec un grand soin, à plusieurs niveaux, alors que le génotype du CMH n'a pu être déterminé, comme dans le cas du challenge infectieux réalisé pour la maladie de Marek sur des lignées commerciales au laboratoire. Par la suite il devrait être possible, grâce à la puce 96 SNPs d'identifier des recombinants. Il a déjà été montré que des lignées recombinantes existent pour la région B du CMH même si les points de cassure sont mal connus, sauf pour un exemple récent (Goto *et al.*, 2009). Certaines rares observations faites pendant la thèse pourraient également indiquer des recombinaisons entre haplotypes. Cependant les caractéristiques de ces recombinaisons et notamment leurs fréquences sont inconnues, il est donc difficile de prédire si ce type d'animal pourra être identifié par chance dans un croisement expérimental ou si ces animaux devront être expressément recherchés puis testés.



## REFERENCES

- Afanassieff M, Goto RM, Ha J, Sherman MA, Zhong L, Auffray C, Coudert F, Zoorob R, Miller MM (2001) At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. *J Immunol* **166**:3324-3333.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**:403-410.
- Arakawa H, Hauschild J, Buerstedde JM (2002) Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* **295**:1301-1306.
- Bacon LD, Hunt HD, Cheng HH (2001) Genetic resistance to Marek's disease. *Curr Top Microbiol Immunol* **255**:121-141.
- Bacon LD, Sundick RS, Rose NR (1977) Genetic and cellular control of spontaneous autoimmune thyroiditis in OS chickens. *Adv Exp Med Biol* **88**:309-318.
- Bacon LD, Witter RL, Crittenden LB, Fadly A, Motta J (1981) B-haplotype influence on Marek's disease, Rous sarcoma, and lymphoid leukosis virus-induced tumors in chickens. *Poult Sci* **60**:1132-1139.
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**:37-48.
- Belkhir K, Borsa P, Goudet J, Chikhi L, Raufasate N, Bonhomme F (2000) GENETIX, logiciel sous Windows TM pour la génétique des populations In Laboratoire Génome, Populations, Interactions: CNRS UMR5117 Université Montpellier II, Montpellier, France.
- Bernot A, Auffray C (1991) Primary structure and ontogeny of an avian CD3 transcript. *Proc Natl Acad Sci U S A* **88**:2550-2554.
- Berthouly C, Bed'Hom B, Tixier-Boichard M, Chen CF, Lee YP, Laloë D, Legros H, Verrier E, Rognon X (2008) Using molecular markers and multivariate methods to study the genetic diversity on local European and Asian chickens breeds. *Anim Genet* **39**:121-129.
- Berthouly C, Leroy G, Van TN, Thanh HH, Bed'Hom B, Nguyen BT, Vu CC, Monicat F, Tixier-Boichard M, Verrier E, Maillard JC, Rognon X (2009) Genetic analysis of local Vietnamese chickens provides evidence of gene flow from wild to domestic populations. *BMC Genet* **10**:1.
- Blackburn HD (2006) The National Animal Germplasm Program: challenges and opportunities for poultry genetic resources. *Poult Sci* **85**:210-5.
- Bonneaud C, Sorci G, Morin V, Westerdahl H, Zoorob R, Wittzell H (2004) Diversity of MHC class I and IIB genes in House Sparrows (*Passer domesticus*). *Immunogenetics* **55**:855-865.



- Bradley RD, Hillis DM (1997) Recombinant DNA sequences generated by PCR amplification. *Mol Biol Evol* **14**:592-593.
- Briles WE, Briles RW (1987) Genetics and classification of major histocompatibility complex antigens of the chicken. *Poult Sci* **66**:776-781.
- Briles WE, Bumstead N, Ewert DL, Gilmour DG, Gogusev J, Hála K, Koch C, Longenecker BM, Nordskog AW, Pink JR, Schierman LW, Simonsen M, Toivanen A, Toivanen P, Vainio O, Wick G (1982) Nomenclature for chicken major histocompatibility (B) complex. *Immunogenetics* **15**:441-447
- Briles WE, McGibbon WH, Irwin MR (1948) Studies of the time of development of cellular antigens in the chicken. *Genetics* **33**:97.
- Briles WE, Stone HA, Cole RK (1977) Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science* **195**:193-195.
- Calenge F, Lecerf F, Demars J, Feve K, Vignoles F, Pitel F, Vignal A, Velge P, Sellier N, Beaumont C (2009) QTL for resistance to Salmonella carrier state confirmed in both experimental and commercial chicken lines. *Anim Genet* **40**:590-597.
- Carrington M, Yeager M, Mann D (1993) Characterization of HLA-DMB polymorphism. *Immunogenetics* **38**:446-449.
- Chang CM, Coville JL, Coquerelle G, Gourichon D, Oulmouden A, Tixier-Boichard M (2006) Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genomics* **7**:19.
- Chaves LD, Krueth SB, Reed KM (2007) Characterization of the turkey MHC chromosome through genetic and physical mapping. *Cytogenet Genome Res* **117**:213-220.
- Chaves LD, Krueth SB, Reed KM (2009) Defining the turkey MHC: sequence and genes of the B locus. *J Immunol* **183**:6530-6537.
- Chen Y, Carpenter SL, Lamont SJ (2000) Detection of protein binding to a glucocorticoid response element-like sequence in a chicken major histocompatibility complex class II promoter. *Vet Immunol Immunopathol* **30**:1-7.
- Cho S, Attaya M, Brown MG, Monaco JJ (1991) New class II-like genes in the murine MHC. *Nature* **353**:573-576.
- Cock van Oosterhout (2009) A new theory of MHC evolution: beyond selection on the immune genes. *Proc R Soc B* **276**:657-665.
- Coquerelle G (2000) Les poules Diversité Génétique Visible. INRA Editions, France.
- Cotter PF, Ayoub J, Parmentier HK (2005) Directional selection for specific sheep cell antibody responses affects natural rabbit agglutinins of chickens. *Poult Sci* **84**:220-225.

- Crooijmans RP, Dijkhof RJ, van der Poel JJ, Groenen MA (1997) New microsatellite markers in chicken optimized for automated fluorescent genotyping. *Anim Genet* **28**:427-437.
- Dalgaard TS, Højsgaard S, Skjødt K, Juul-Madsen HR (2003) Differences in chicken major histocompatibility complex (MHC) class I alpha gene expression between Marek's disease-resistant and -susceptible MHC haplotypes. *Scand J Immunol* **57**:135-143.
- Dalgaard TS, Vitved L, Skjødt K, Thomsen B, Labouriau R, Jensen KH, Juul-Madsen HR (2005) Molecular characterization of major histocompatibility complex class I (*B-F*) mRNA variants from chickens differing in resistance to Marek's disease. *Scand J Immunol* **62**:259-270.
- Davison TF (2003) The immunologists' debt to the chicken. *Br Poult Sci* **44**:6-21.
- Delany ME, Robinson CM, Goto RM, Miller MM (2009) Architecture and organization of chicken microchromosome 16: order of the NOR, MHC-Y, and MHC-B subregions. *J Hered* **100**:507-514.
- Downing T, Lynn DJ, Connell S, Lloyd AT, Bhuiyan AK, Silva P, Naqvi AN, Sanfo R, Sow RS, Podisi B, O'Farrelly C, Hanotte O, Bradley DG (2009) Contrasting evolution of diversity at two disease-associated chicken genes. *Immunogenetics* **61**:303-314.
- Edwards SV, Gasper J, Garrigan D, Martindale D, Koop BF (2000) A 39-kb sequence around a Blackbird MHC class II gene: Ghost of selection past and songbird genome architecture. *Mol Biol Evol* **17**:1384-1395.
- Ennis PD, Zemmour J, Salter RD, Parham P (1990) Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. *Proc Natl Acad Sci U S A* **87**:2833-2837.
- Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Strömstedt L, Wright D, Jungerius A, Vereijken A, Randi E, Jensen P, Andersson L (2008) Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet* **4**:e1000010.
- Flot JF (2009) Seqphase, a web tool for interconverting phase input/output files and fasta sequence alignments. *Mol Ecol Resour* in press.
- Fotsa J-C (2008) Caractérisation des populations de poules locales (*Gallus gallus*) au Cameroun. Thesis, AgroParistech, France.
- Frangoulis B, Park I, Guillemot F, Séverac V, Auffray C, Zoorob R (1999) Identification of the Tapasin gene in the chicken major histocompatibility complex. *Immunogenetics* **49**:328-337.
- Fulton JE, Juul-Madsen HR, Ashwell CM, McCarron AM, Arthur JA, O'Sullivan NP, Taylor RL Jr (2006) Molecular genotype identification of the *Gallus gallus* major histocompatibility complex. *Immunogenetics* **58**:407-421.

- Fumihito A, Miyake T, Takada M, Shingu R, Endo T, Gojobori T, Kondo N, Ohno S (1996) Monophyletic origin and unique dispersal patterns of domestic fowls. *Proc Natl Acad Sci U S A* **93**:6792-6795.
- Gandon S, Mackinnon MJ, Nee S, Read AF (2001) Imperfect vaccines and the evolution of pathogen virulence. *Nature* **414**:751-756.
- Goto RM, Afanassieff M, Ha J, Iglesias GM, Ewald SJ, Briles WE, Miller MM (2002) Single-strand conformation polymorphism (SSCP) assays for major histocompatibility complex B genotyping in chickens. *Poult Sci* **81**:1832-1841.
- Goto RM, Wang Y, Taylor RL Jr, Wakenell PS, Hosomichi K, Shiina T, Blackmore CS, Briles WE, Miller MM (2009) BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proc Natl Acad Sci U S A* **106**:16740-16745.
- Granevitze Z, Hillel J, Chen GH, Cuc NTK, Feldam M, Eding H, Weigend S (2007) Genetic diversity within chicken populations from different continents and management histories. *Anim Genet* **38**:576-583.
- Grimholt U, Larsen S, Nordmo R, Midtlyng P, Kjoeglum S, Storset A, Saebø S, Stet RJ (2003) MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics* **55**:210-219.
- Groenen MA, Cheng HH, Bumstead N, Benkel BF, Briles WE, Burke T, Burt DW, Crittenden LB, Dodgson J, Hillel J, Lamont S, de Leon AP, Soller M, Takahashi H, Vignal A. A consensus linkage map of the chicken genome. *Genome Res* **10**:137-147.
- Guillemot F, Billault A, Pourquié O, Béhar G, Chaussé AM, Zoorob R, Kreibich G, Auffray C (1988) A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer. *EMBO J* **7**:2775-2785.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41**:95-98.
- Hermel E, Yuan J, Monaco JJ (1995) Characterization of polymorphism within the H2-M MHC class II loci. *Immunogenetics* **42**:136-142.
- Hill AV (2001) The genomics and genetics of human infectious disease susceptibility. *Annu Rev Genomics Hum Genet* **2**:373-400.
- Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, Brewster D, McMichael AJ, Greenwood BM (1991) Common west African HLA antigens are associated with protection from severe malaria. *Nature* **352**:595-600.

- Hillel J, Groenen MA, Tixier-Boichard M, Korol AB, David L, Kirzhner VM, Burke T, Barre-Dirie A, Crooijmans RP, Elo K, Feldman MW, Freidlin PJ, Mäki-Tanila A, Oortwijn M, Thomson P, Vignal A, Wimmers K, Weigend S (2003) Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genet Sel Evol* **35**:533-557.
- Hofmann A, Plachy J, Hunt L, Kaufman J, Hala K (2003) v-src oncogene-specific carboxy-terminal peptide is immunoprotective against Rous sarcoma growth in chickens with MHC class I allele B-F12. *Vaccine* **21**:4694-4699.
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, Wain HM, Trowsdale J, Ziegler A, Beck S (2004) Gene map of the extended human MHC. *Nat Rev Genet* **5**:889-899.
- Hosomichi K, Miller MM, Goto RM, Wang Y, Suzuki S, Kulski JK, Nishibori M, Inoko H, Hanzawa K, Shiina T (2008) Contribution of mutation recombination, and gene conversion to chicken MHC-B haplotype diversity. *J Immunol* **181**:3393-3399.
- Hudson JC, Hoerr EJ, Parker SH, Ewald SJ (2002) Quantitative measures of disease in broiler breeder chicks of different major histocompatibility complex genotypes after challenge with infectious bursal disease virus. *Avian Dis* **46**:581-592.
- Hughes AL, Nei M (1990) Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol Biol Evol* **7**:491-514.
- Huson DH, Bryant D (2006) Application of phylogenetics networks in evolutionary studies. *Mol Bio Evo* **23**:254-267.
- International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**:695-716.
- International Chicken Polymorphism Map Consortium (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* **432**:717-722.
- Janeway CA, Travers M, Walport M, Shlomchik M (2008) *Immunobiology*. Garland Publishing, New York.
- Jarvi SI, Goto RM, Briles WE, Miller MM (1996) Characterization of MHC genes in a multigenerational family of ring-necked pheasants. *Immunogenetics* **43**:125-135.
- Jeffery KJ, Usuku K, Hall SE, Matsumoto W, Taylor GP, Procter J, Bunce M, Ogg GS, Welsh KI, Weber JN, Lloyd AL, Nowak MA, Nagai M, Kodama D, Izumo S, Osame M, Bangham CR (1999) HLA alleles determine human T-lymphotropic virus-I (HTLV-I) proviral load and the risk of HTLV-I-associated myelopathy. *Proc Natl Acad Sci U S A* **96**:3848-3853.
- Jones Dukes MD, Locklear CC, Buchholz R, Hecht SJ, Findley AM (2003) Production of an MHC class II B molecular probe in the turkey, *Meleagris gallopavo*. *J Appl Genet* **44**:369-373.

- Kaiser P, Staeheli P (2008) Avian cytokines and chemokines p203 in *Avian immunology*. Elsevier, Amsterdam.
- Kaslow RA, Duquesnoy R, VanRaden M, Kingsley L, Marrari M, Friedman H, Su S, Saah AJ, Detels R, Phair J, *et al.*, (1990) A1, Cw7, B8, DR3 HLA antigen combination associated with rapid decline of T-helper lymphocytes in HIV-1 infection. A report from the Multicenter AIDS Cohort Study. *Lancet* **335**:927-930.
- Kaufman J (2000) The simple chicken major histocompatibility complex: life and death in the face of pathogens and vaccines. *Philos Trans R Soc Lond B Biol Sci* **29**:1077-1084
- Kaufman J (2008) The Avian MHC p159 in *Avian immunology*. Elsevier, Amsterdam.
- Kaufman J, Milne S, Gobel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S (1999) The chicken *B* locus is a minimal essential major histocompatibility complex. *Nature* **28**:923-925.
- Kaufman J, Salomonsen J (1997) The "minimal essential MHC» revisited: both peptide-binding and cell surface expression level of MHC molecules are polymorphisms selected by pathogens in chickens. *Hereditas* **127**:67-73.
- Kaufman J, Völk H, Wallny HJ (1995) A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunol Rev* **143**:63-88.
- Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. *Immunogenetics* **56**:683-695.
- Kelly AP, Monaco JJ, Cho S, Trowsdale J (1991) A new human HLA class II-related locus, DM. *Nature* **353**:571-573.
- Kroemer G, Zoorob R, Auffray C (1990) Structure and expression of a chicken MHC class I gene. *Immunogenetics* **31**:405-409.
- Kroner BL, Goedert JJ, Blattner WA, Wilson SE, Carrington MN, Mann DL (1995) Concordance of human leukocyte antigen haplotype-sharing, CD4 decline and AIDS in hemophilic siblings. Multicenter Hemophilia Cohort and Hemophilia Growth and Development Studies. *AIDS* **9**:275-280.
- Kropshofer H, Arndt SO, Moldenhauer G, Hammerling GJ, Vogt AB (1997a) HLA-DM acts as a molecular chaperone and rescues empty HLA-DR molecules at lysosomal pH. *Immunity* **6**:293-302.
- Kropshofer H, Hammerling GJ, Vogt AB (1997b) How HLA-DM edits the MHC class II peptide repertoire: survival of the fittest? *Immunol Today* **18**:77-82.
- Kropshofer H, Hämmerling GJ, Vogt AB (1999) The impact of the non-classical MHC proteins HLA-DM and HLA-DO on loading of MHC class II molecules. *Immunol Rev* **172**:267-278.

- Kulski JK, Shiina T, Anzai T, Kohara S, Inoko H (2002) Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity, and complexity from shark to man. *Immunol Rev* **190**:95–122.
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**:150-163.
- Lahti JM, Chen CL, Tjoelker LW, Pickel JM, Schat KA, Calnek BW, Thompson CB, Cooper MD (1991) Two distinct alpha beta T-cell lineages can be distinguished by the differential usage of T-cell receptor V beta gene segments. *Proc Natl Acad Sci U S A* **88**:10956-10960.
- Lakshmanan N, Lamont SJ (1998) Rfp-Y region polymorphism and Marek's disease resistance in multitrait immunocompetence-selected chicken lines. *Poult Sci* **77**:538-541.
- Lambourne MD, Si W, Niemiec PK, Read LR, Kariyawasam S, Sharif S (2005) Identification of novel polymorphisms in the B-LB locus of *Gallus lafayettei*. *Anim Genet* **36**:445-448.
- Lamont SJ (1998) Impact of genetics on disease resistance. *Poult Sci* **77**:1111-1118.
- Lamont SJ (2003) Unique population designs used to address molecular genetics questions in poultry. *Poult Sci* **82**:882-884.
- Lamont SJ, Gerndt BM, Warner CM, Bacon LD (1990) Analysis of restriction fragment length polymorphisms of the major histocompatibility complex of 15I5-B-congenic chicken lines. *Poult Sci* **69**:1195-1203.
- Lee LF, Bacon LD, Yoshida S, Yanagida N, Zhang HM, Witter RL (2004) The efficacy of recombinant fowlpox vaccine protection against Marek's disease: its dependence on chicken line and B haplotype. *Avian Dis* **48**:129-137.
- Lekstrom-Himes JA, Hohman P, Warren T, Wald A, Nam JM, Simonis T, Corey L, Straus SE (1999) Association of major histocompatibility complex determinants with the development of symptomatic and asymptomatic genital herpes simplex virus type 2 infections. *J Infect Dis* **179**:1077-1085.
- Lima-Rosa CA, Canal CW, Fallavena PRV, Freitas LB, Salzano FM (2005) LEI0258 microsatellite variability and its relationship to B-F haplotypes in Brazilian (blue-egg Caipira) chickens. *Genet Mol Biol* **28**:386-389.
- Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, Baig M, Beja-Pereira A, Ding ZL, Palanichamy MG, Zhang YP (2006) Multiple maternal origins of chickens: out of the Asian jungles. *Mol Phylogenet Evol* **38**:12-19.
- Livant EJ, Brigati JR, Ewald SJ (2004) Diversity and locus specificity of chicken MHC B class I sequences. *Anim Genet* **35**:18-27.

- Longenecker BM, Mosmann TR (1981) Structure and properties of the major histocompatibility complex of the chicken. Speculations on the advantages and evolution of polymorphism. *Immunogenetics* **13**:1-23.
- Loukou NE, Yapi-Gnaoré CV, Touré G, Coulibaly Y, Rognon X, Kayang B, Youssao I, Tixier-Boichard M, N'guetta ASP (2009) Assessing the diversity of indigenous chicken from two agro-ecological zones of Côte d'Ivoire using microsatellite markers. *J Anim Plant Sci* **5**:425 – 436.
- Loywyck V, Bed'Hom B, Pinard-van der Laan M-H, Pitel F, Verrier E, Bijma P (2008) Evolution of polymorphisms at molecular markers in QTL and non-QTL regions in selected lines of chicken. *Genet Sel Evol* **40**:639–661.
- Loywyck V, Pinard-van der Laan MH, Goldringer I, Verrier E (2006) On the need for combining complementary analyses to assess the effect of a candidate gene and the evolution of its polymorphism: the example of the Major Histocompatibility Complex in chicken. *Genet Res* **87**:125-131.
- Martin C, Ohki-Hamazaki H, Corbel C, Coltey M, Le Douarin N.M. (1991) Successful xenogeneic transplantation in embryos: induction of tolerance by extrathymic chick tissue grafted into quail. *Dev Immunol* **1**:265-77.
- McConnell SK, Dawson DA, Wardle A, Burke T (1999) The isolation and mapping of 19 tetranucleotide microsatellite markers in the chicken. *Anim Genet* **30**:183-189.
- Mesa CM, Thulien KJ, Moon DA, Veniamin SM, Magor KE (2004) The dominant MHC class I gene is adjacent to the polymorphic TAP2 gene in the duck, *Anas platyrhynchos*. *Immunogenetics* **56**:192-203.
- Miller MM (1995) US Patent n°5451670.
- Miller MM, Bacon LD, Hala K, Hunt HD, Ewald SJ, Kaufman J, Zoorob R, Briles WE (2004) 2004 Nomenclature for the chicken major histocompatibility (*B* and *Y*) complex. *Immunogenetics* **56**:261-79.
- Miller MM, Goto R, Young S, Chirivella J, Hawke D, Miyada CG (1991) Immunoglobulin variable-region-like domains of diverse sequence within the major histocompatibility complex of the chicken. *Proc Natl Acad Sci U S A* **88**:4377-4381.
- Miller MM, Wang C, Parisini E, Coletta RD, Goto RM, Lee SY, Barral DC, Townes M, Roura-Mir C, Ford HL, Brenner MB, Dascher CC (2005) Characterization of two avian MHC-like genes reveals an ancient origin of the CD1 family. *Proc Natl Acad Sci U S A* **102**:8674-8679.
- Minozzi G, Parmentier HK, Mignon-Grasteau S, Nieuwland MG, Bed'hom B, Gourichon D, Minvielle F, Pinard-van der Laan MH (2008) Correlated effects of selection for immunity in White Leghorn chicken lines on natural antibodies and specific antibody responses to KLH and M. butyricum. *BMC Genet* **14**:9-5.

- Minozzi G, Parmentier HK, Nieuwland MG, Bed'hom B, Minvielle F, Gourichon D, Pinard-van der Laan MH (2007) Antibody responses to keyhole limpet hemocyanin lipopolysaccharide, and Newcastle Disease virus vaccine in F2 and backcrosses of white Leghorn lines selected for two different immune response traits. *Poult Sci* **86**:1316-1322.
- Monteyne P, Bureau JF, Brahic M (1997) The infection of mouse by Theiler's virus: from genetics to immunology. *Immunol Rev* **159**:163-176.
- Moon D.A, Veniamin S.M, Parks-Dely J.A, Magor K.E. (2005) The MHC of the Duck (*Anas platyrhynchos*) contains five differentially expressed class I genes. *J Immunol* **175**:6702-12.
- Morris P, Shaman J, Attaya M, Amaya M, Goodman S, Bergman C, Monaco JJ, Mellins E (1994) An essential role for HLA-DM in antigen presentation by class II major histocompatibility molecules. *Nature* **368**:551-554.
- Muir WM, Wong GK, Zhang Y, Wang J, Groenen MA, Crooijmans RP, Megens HJ, Zhang H, Okimoto R, Vereijken A, Jungerius A, Albers GA, Lawley CT, Delany ME, MacEachern S, Cheng HH (2008) Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci U S A* **105**:17312-17317.
- Nanda I, Sick C, Münster U, Kaspers B, Scharl M, Staeheli P, Schmid M (1998) Sex chromosome linkage of chicken and duck type I interferon genes: further evidence of evolutionary conservation of the Z chromosome in birds. *Chromosoma* **107**:204-210.
- Nau F, Guérin-Dubiard C, Désert C, Gautron J, Bouton S, Gribonval J, Lagarrigue S (2003) Cloning and characterization of HEP21, a new member of the uPAR/Ly6 protein superfamily predominantly expressed in hen egg white. *Poult Sci* 2003 **82**:242-50.
- O'hUigin C, Sültsmann H, Tichy H, Murray BW (1998) Isolation of mhc class II DMA and DMB cDNA sequences in a marsupial: the gray short-tailed opossum (*Monodelphis domestica*). *J Mol Evol* **47**:578-585.
- O'Neill AM, Livant EJ, Ewald SJ (2009) The chicken BF1 (classical MHC class I) gene shows evidence of selection for diversity in expression and in promoter and signal peptide regions. *Immunogenetics* **61**:289-302.
- Osei-Amponsah R, Kayang BB, Naazie A, Osei YD, Tixier-Boichard M, Rognon X (2009) Genetic diversity of Forest and Savannah chicken populations of Ghana as estimated by microsatellite markers *Anim Sci J* in press.
- Parham P, Lawlor DA, Lomen CE, Ennis PD (1989) Diversity and diversification of HLA-A,B,C alleles. *J Immunol* **142**:3937-3950.
- Payne LN, Jaffe P (1962) Genetic basis for graft-against-host immunological reactions between two inbred lines of chickens. *Nature* **190**:373-374.



- Penn DJ, Damjanovich K, Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci U S A* **99**:11260–11264.
- Pinard MH & van der Zijpp AJ (1993) Effects of major histocompatibility complex on antibody response in F1 and F2 cross of chicken lines. *Genet Sel Evol* **25**:283-296.
- Pinard MH, Hepkema BG (1993) Biochemical and serological identification of major histocompatibility complex antigens in outbred chickens. *Vet Immunol Immunopathol* **39**:407-419.
- Pinard-van der Laan MH (2002) Immune modulation: the genetic approach. *Vet Immunol Immunopathol* **87**:199-205.
- Pinard-van der Laan MH, Siegel PB, Lamont SJ (1998) Lessons from selection experiments on immune response in the chicken. *Poult Avian Biol Rev* **9**:125–141.
- Pinard-van der Laan MH, Soubieux D, Mérat L, Bouret D, Luneau G, Dambrine G, Thoraval P (2004) Genetic analysis of a divergent selection for resistance to Rous sarcomas in chickens. *Genet Sel Evol* **36**:65-81.
- Plachy J, Kaiser P, Hála K (2003) Genetics of the Immune System. p293 in *Poultry Genetics, Breeding and Biotechnology*. CABI Publishing, Cambridge.
- Ratcliffe MJ (2005) Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development. *Dev Comp Immunol* **30**:101-118.
- Rogers SL, Göbel TW, Viertlboeck BC, Milne S, Beck S, Kaufman J (2005) Characterization of the chicken C-type lectin-like receptors B-NK and B-lec suggests that the NK complex and the MHC share a common ancestral region. *J Immunol* **174**:3475-3483.
- Rogers SL, Kaufman J (2008) High allelic polymorphism, moderate sequence diversity and diversifying selection for B-NK but not B-lec, the pair of lectin-like receptor genes in the chicken MHC. *Immunogenetics* **60**:461-475.
- Rosenberg NA, Burke T, Elo K, Feldman MW, Freidlin PJ, Groenen MA, Hillel J, Mäkitanila A, Tixier-Boichard M, Vignal A, Wimmers K, Weigend S (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* **159**:699-713.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**:365-386.
- Ruby T, Bed'Hom B, Wittzell H, Morin V, Oudin A, Zoorob R (2005) Characterisation of a cluster of TRIM-B30.2 genes in the chicken MHC B locus. *Immunogenetics* **57**:116-128.
- Salomonsen J, Marston D, Avila D, Bumstead N, Johansson B, Juul-Madsen H, Olesen GD, Riegert P, Skjødt K, Vainio O, Wiles MV, Kaufman J (2003) The properties of the single chicken MHC classical class II alpha chain (B-LA) gene indicate an ancient origin for the DR/E-like isotype of class II molecules. *Immunogenetics* **55**:605-614.

- Salomonsen J, Sørensen MR, Marston DA, Rogers SL, Collen T, van Hateren A, Smith AL, Beal RK, Skjødt K, Kaufman J (2005) Two CD1 genes map to the chicken MHC, indicating that CD1 genes are ancient and likely to have been present in the primordial MHC. *Proc Natl Acad Sci U S A* **102**:8668-8673.
- Schultz U, Magor KE (2008) Comparative Immunology of Agricultural Birds p395 in *Avian immunology*. Elsevier, Amsterdam.
- Shaw I, Powell TJ, Marston DA, Baker K, van Hateren A, Riegert P, Wiles MV, Milne S, Beck S, Kaufman J (2007) Different evolutionary histories of the two classical class I genes BF1 and BF2 illustrate drift and selection within the stable MHC haplotypes of chickens. *J Immunol* **178**:5744-5752.
- Shiina T, Briles WE, Goto RM, Hosomichi K, Yanagiya K, Shimizu S, Inoko H, Miller MM (2007) Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J Immunol* **178**:7162-7172.
- Shiina T, Hosomichi K, Hanzawa K (2006) Comparative genomics of the poultry major histocompatibility complex. *Anim Sci J* **77**:151-162.
- Shiina T, Shimizu S, Hosomichi K, Kohara S, Watanabe S, Hanzawa K, Beck S, Kulski JK, Inoko H (2004) Comparative genomic analysis of two avian (Quail and Chicken) MHC regions. *J Immunol* **1**:6751-6763.
- Simonsen M, Crone M, Koch C, Hála K (1982) The MHC haplotypes of the chicken. *Immunogenetics* **16**:513-532.
- Sironi L, Lazzari B, Ramelli P, Stella A, Mariani P (2008) Avian TAP genes: detection of nucleotide polymorphisms and comparative analysis across species. *Genet Mol Res* **7**:1267-1281.
- Slade RW, and McCallum HI (1992). Overdominant vs. frequency-dependent selection at MHC loci. *Genetics* **132**:861-864.
- Sloan VS, Cameron P, Porter G, Gammon M, Amaya M, Mellins E, Zaller DM (1995) Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* **375**:802-806.
- Spike CA, Lamont SJ (1995) Genetic analysis of three loci homologous to human G9a: evidence for linkage of a class III gene with the chicken MHC. *Anim Genet* **26**:185-187.
- Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* **132**:115-130.
- Stear MJ, Bishop SC, Mallard BA, Raadsma H (2001) The sustainability, feasibility and desirability of breeding livestock for disease resistance. *Res Vet Sci* **71**:1-7.
- Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am J Hum Genet* **76**:449-462.

- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**:978-989.
- Takahata N, and Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**:967-978.
- Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* **130**:925-938.
- Takami K, Zaleska-Rutczynska Z, Figueroa F, Klein J (1997) Linkage of LMP, TAP, and RING3 with Mhc class I rather than class II genes in the zebrafish. *J Immunol* **159**:6052-6060.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**:1596-1599.
- Tixier-Boichard M, Durand L, Morisson M, Ricard F, Coquerelle G (1994) Comparative analysis of avian leukosis related endogenous viral genes in experimental strains of the domestic chicken. *Genet Sel Evol* **26**:53-66.
- Twito T, Weigend S, Blum S, Granevitze Z, Feldman MW, Perl-Treves R, Lavi U, Hillel J (2007) Biodiversity of 20 chicken breeds assessed by SNPs located in gene regions. *Cytogenet Genome Res* **117**:319-326.
- Van den Elsen PJ, Holling TM, Kuipers HF, Van der Stoep N (2004) Transcriptional regulation of antigen presentation. *Curr Opin Immunol* **16**:67-75.
- Wain HM, Toye AA, Hughes S, Bumstead N (1998) Targeting of marker loci to chicken chromosome 16 by representational difference analysis. *Anim Genet* **29**:446-452.
- Wallny HJ, Avila D, Hunt LG, Powell TJ, Riegert P, Salomonsen J, Skjødt K, Vainio O, Vilbois F, Wiles MV, Kaufman J (2006) Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. *Proc Natl Acad Sci U S A* **103**:1434-1439.
- Wang X, Erf GF (2004) Apoptosis in feathers of Smyth line chickens with autoimmune vitiligo. *J Autoimmun* **22**:21-30.
- Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, Van Broeckhoven C, De Rijk P (2005) NovoSNP, a novel computational tool for sequence variation discovery. *Genome Research* **15**:436-442.
- Westerdahl H, Wittzell H, von Schantz T (2000) MHC diversity in two passerine birds: no evidence for a minimal essential MHC. *Immunogenetics* **52**:92-100.
- Westerheide SD, Louis-Plence P, Ping D, He XF, Boss JM (1997) HLA-DMA and HLA-DMB gene expression functions through the conserved S-X-Y region. *J Immunol* **158**:4812-4821.

- Wittzell H, Bernot A, Auffray C, Zoorob R (1999) Concerted evolution of two MHC class II B loci in pheasants and domestic chickens. *Mol Biol Evol* **16**:479-490.
- Ye H, Xu FZ, Yu WY (2009) The intracellular localization and oligomerization of chicken invariant chain with major histocompatibility complex class II subunits. *Poult Sci* **88**:1594-1600.
- Youssao IAK, Dahouda M, Kpodékon MT, Idrissou N-D, Bonou GA, Tougan UP, Ahounou S, Yapi-Gnaoré V, Kayang B, Rognon X, Tixier-Boichard M (2009) Phenotypic characterisation and molecular polymorphism of indigenous poultry populations of the species *Gallus gallus* of North and South ecotypes of Benin. *African J Biotech* in press.
- Zhou H, Lamont SJ (2003) Chicken MHC class I and II gene effects on antibody response kinetics in adult chickens. *Immunogenetics* **55**:133-140.
- Zinkernagel RM, Pfau CJ, Hengartner H, Althage A. Susceptibility to murine lymphocytic choriomeningitis maps to class I MHC genes--a model for MHC/disease associations. *Nature* **316**:814-817.
- Zöller B, Ozato K, Kroemer G, Auffray C, Jungwirth C (1992) Interferon induction of chicken MHC class I gene expression: phylogenetic conservation of the interferon-responsive element. *Virology* **191**:141-149.
- Zoorob R, Auffray C, Chaussé AM (1998) Réactifs et méthodes pour la détection des genes lies au CMH d'oiseaux d'élevage tels que le poulet. Patent PCT/FR98/02501, France.



---

## GENETIC RESISTANCE – WITH FOCUS ON MAJOR HISTOCOMPATIBILITY COMPLEX

---

**Helle R. Juul-Madsen<sup>1</sup>**  
**Tina S. Dalgaard<sup>1</sup>**  
**Jan Salomonsen<sup>2</sup>**  
**Dan E. Heller<sup>3</sup>**

<sup>1</sup>*Danish Institute of Agricultural Sciences, DK-8830 Tjele, Denmark*

<sup>2</sup>*Royal Veterinary and Agricultural University, DK-1870 Frederiksberg, Denmark*

<sup>3</sup>*The Hebrew University of Jerusalem, Zip 76100, City Rehovot, Israel*

### TABLE OF CONTENTS

|   |   |
|---|---|
| INTRODUCTION .....  | 2 |
| Serological typing of MHC                                     |   |
| Introduction.....   | 3 |
| Equipment .....   | 3 |
| Reagents .....  | 3 |
| Preparation of Normal Chicken Plasma.....                     | 3 |
| Preparation of B-F and B-G allo-antisera.....                 | 3 |
| Preparation of “ready-to-use” allo-antiserum.....             | 4 |
| Haemagglutination test.....                                   | 5 |
| Serological MHC typing of offspring.....                      | 5 |
| Restriction Fragment Length Polymorphism (RFLP) typing of MHC |   |
| Introduction.....   | 5 |
| Equipment .....   | 6 |
| Reagents .....  | 6 |
| Purification of genomic DNA with Sodium Chloride.....         | 6 |
| Genomic DNA restriction digestion.....                        | 7 |
| Alkali blotting.....  | 7 |
| Micro-satellite fragment typing of MHC                        |   |
| Introduction.....   | 8 |
| Equipment .....   | 8 |
| Reagents .....  | 8 |
| PCR amplification of micro-satellite .....                    | 8 |
| Fragment analyses .....                                       | 9 |
| Suppliers .....   | 9 |
| Concluding remarks.....                                       | 9 |
| Acknowledgements.....   | 9 |
| References.....   | 9 |

## Introduction

Avian infectious diseases are a major problem in intensive poultry production systems. Avian infectious diseases are also a matter of concern to the human population as chickens may be carriers of human bacterial pathogens such as *Salmonella* or *Escherichia coli*. Furthermore, viral pathogens may evolve in chickens, cross species barriers, and infect humans. The outbreaks of avian influenza in Hong Kong in 1997 and 2003 are good examples (Kuhnlein et al., 2003).

The major tools to combat avian diseases have so far been antibiotics and vaccines. However, there are major concerns that both antibiotics and vaccines may contribute to the evolution of more resistant pathogens, which require the development of new and more efficient vaccines. This has been so for Infectious Bursal Disease Virus (IBDV) for several years. Increased response to vaccines in combination with better hygiene may therefore be a favourable production principle of choice.

The immune system is the natural means by which animals resist infection, and immunological parameters may reflect the functional capacity or immuno-competence of the immune system and, in turn, the ability of the animal to resist infection. So far only a few genes have been implicated in affecting disease resistance. Genes from the Major Histocompatibility Complex (MHC) involved in antigen presentation were among the earliest genes to be identified. These genes encode some of the most polymorphic molecules known in the animal kingdom. Diversity at the population level is thought to evolve via the interaction between host MHC molecules and pathogens leading to a balanced evolution of both species.

In chickens, the MHC genes are located in two chromosomal regions (the B system and the Rfp-Y system) that assort in a genetically independent way, although located on the same micro-chromosome (Briles et al., 1993, Miller et al., 1994; Fillon et al., 1996). The B system is considered the strongest histocompatibility antigen delivering locus and contains at least two B-F genes, which are homologous to mammalian MHC class I genes, and at least two B-LB genes which are homologous to mammalian MHC class IIB genes. In addition, the B system contains genes encoding a third group of molecules, namely the B-G genes. The B-G molecules have so far only been found in birds, and are just as polymorphic as the B-F and the B-L molecules. However, their function is still unknown (Guillemot et al., 1989; Kaufman et al., 1989; Miller et al., 1990). B-F antigens are expressed on almost all chicken cells including nucleated chicken erythrocytes whereas the B-L antigens primarily are expressed on B-cells, activated T-cells, and antigen presenting cells. The B-G molecules are primarily expressed on erythrocytes and thrombocytes although white blood cells also have B-G epitopes (Salomonsen et al., 1991). The Rfp-Y system is considered to be an MHC-like region rather than a second MHC since there is no report on rapid allograft rejection due to polymorphism in Rfp-Y (Kaufman and Wallny, 1996).

Genetic disease resistance is known to influence several diseases in poultry. In many cases, this resistance has been mapped to the MHC region. One of the most significant associations between MHC and disease resistance is Marek's disease (Lamont, 1998), but resistance to other virus-induced diseases such as Rous sarcoma virus (Schierman and Collins, 1987) and avian leukosis (Yoo and Sheldon, 1992) have been linked to the MHC region as well. In addition to viral diseases, other diseases have been shown to be influenced by the MHC. These include fowl cholera (Lamont et al., 1987), coccidiosis (Lillehoj et al., 1989) and salmonella infections (Cotter et al., 1998).

In relation to IBDV, disease resistance between different chicken lines has been reported (Bumstead et al., 1993), however, no disease resistance has so far been associated with the B system. But recently Juul-Madsen et al. (2002) showed a difference in response to live attenuated IBDV vaccine that in part was related to the MHC.

Analysis of genetic variation in the chicken MHC can be approached by different methods. Originally MHC typing was performed with a haemagglutination technique based on allospecific polyclonal antisera. The serological agents, however, often cross-react between different MHC haplotypes and typing of outbred chicken populations can be difficult (Kroemer et al., 1990). DNA may easily be isolated from the nucleated chicken erythrocytes, and MHC genotyping by restriction fragment length analysis (RFLP) or by a PCR based test of a polymorphic micro-satellite located in the B complex (Zoorob et al., 1998) are useful supplements.

## Serological typing of MHC

### Introduction

Serological typing of MHC is performed by a haemagglutination test at room temperature in which the allo-antibody react (agglutinate) with the allo-antigen on the surface of erythrocytes. The haemagglutination is performed on open glass slides supplemented with 20% normal chicken plasma as a source of haemagglutination-enhancing factor (Simonsen 1975). Before a MHC haemagglutination test can be performed, specific B antisera must be produced by immunisation of chickens (recipients) with donor cells expressing different MHC antigens than the recipient. B-G allo-antiserum is produced by immunisation of chickens with Red Blood Cells (RBC) whereas B-F allo-antiserum is produced by immunisation of chickens with White Blood Cells (WBC) to avoid the simultaneous formation of anti-B-G specific antibodies.

### Equipment

- Glass slides with 4x10 holes (Figure 1)
- A humid chamber
- A light table
- Table centrifuge e.g. Hereus Omnifuge 2.0
- Rotator
- Syringes (1, 2, and 20 mL)
- Needles (20G, 23G, 25G, and 21G x 31/4)
- CPD (Citrate Phosphate Dextrose adenine) tubes (code VT-050SCPD17) (TE)
- Glass Pasteur pipettes
- 13 mL centrifuge tubes
- 4 mL polystyren test tubes
- 50 mL polypropylen tubes

### Reagents

- 3.8 % Sodium citrate
- PBS buffer pH 7.4
- 10 % Na-azide
- 4 mL of Normal Chicken Plasma (NCP)
- Lymphoprep (Me)
- "Ready-to-use" allo-antisera (the antibody) reacting with different MHC haplotypes.

### Preparation of Normal Chicken Plasma (NCP)

1. Centrifuge blood from a chicken stabilised in 1/5 vol. 3.8 % Sodium Citrate at 500xg for 8 min at 20°C.
2. Remove the upper phase and store the plasma in 4 mL aliquots at -20°C.

### Preparation of B-F and B-G allo-antisera (antibodies)

1. Bleed a donor chicken for 8 mL of blood and stabilise the blood in 1/5 vol. 3.8% Natrium Citrate.
2. Purify the white blood cells (WBC) on a Lymphoprep gradient according to the manufacturer's description.
3. Collect the WBC from the interphase and transfer the cells to a new 13 mL tube.
4. Pipette 100 µL of packed peripheral Red Blood Cells (RBC) from the bottom of the gradient and transfer the cells to a new 13 mL tube containing 5 mL PBS buffer.
5. Gently resuspend both kinds of cells in PBS buffer up to 10 mL and centrifuge the cells at 200xg for 10 min at 4°C.
6. Wash the cells gently 2 times in 5 mL cold PBS buffer and repeat the centrifugation.



7. For the RBC aspirate the supernatant down to 0.5 mL and add 4.5 mL cold PBS – mix gently.
8. For the WBC aspirate the supernatant down to 0.5 mL and add 0.5 mL cold PBS – mix gently.
9. Place the tubes on a rotator at room temperature until immunisation.
10. Immunise each recipient chicken in the wing vein with either 1 mL WBC or 1 mL RBC according to your immunisation plan.
11. For each kind of donor cells two recipients having different MHC haplotypes are chosen (two for WBC and two for RBC).
12. The immunisation is carried out once a week during 8 weeks, and bleeding for collection of allo-antibodies is initiated after the 3rd immunisation and then once a week.
13. Test-typing of the bled sera, diluted 1:2 in PBS buffer, must be analysed for specificity using the haemagglutination test (see below).
14. Add 0.2% Na-azide to the produced allo-antisera and store at - 20°C in aliquots with batch numbers.
15. One animal may be bled for 20 mL for at least 4 times.

## Preparation of “ready-to-use” allo-antiserum

1. Defrost an allo-antiserum diluted 1:2 in PBS buffer containing 0.2% Na-azide.
2. Prepare a panel of 2% blood suspension from animals with known MHC haplotypes (see below).
3. Make a test dilution series of the allo-antiserum in PBS buffer (e.g. 1:2 to 1:128)
4. Test all the dilutions against the panel of known MHC haplotypes using the haemagglutination test described below. The test is performed to find the correct dilution for optimal reaction (agglutination in the highest dilution) and to test the allo-antiserum for cross-reaction to other MHC antigens than the one the allo-antiserum was raised against.
5. If a cross-reaction is found, the allo-antiserum has to be **absorbed** to remove antibodies reacting with unwanted MHC antigens. Below an example of absorption is described.

Example: the allo-antiserum anti-B-F19 cross-reacts with blood cells from animals carrying the B21 haplotype. These antibodies therefore have to be removed by absorption.

1. Bleed 10-20 mL blood from an animal carrying the B21 haplotype and stabilise the blood with 1/5 vol. 3.8% sodium-citrate.
2. Collect the blood cells by centrifugation at 500xg in 8 min at 20° C.
3. Wash the blood cells 2-3 times in 5-10 mL PBS buffer by vortexing followed by centrifugation at 500xg in 8 min at 20° C until the supernatant is clear.
4. After the last wash, the white layer of cells (the lymphocytes on top of the red blood cells) is removed with a Pasteur pipette and discarded. 20 mL blood gives about 5–6 mL packed RBC.
5. Make a test-absorption by mixing 1 mL anti-B-F19 allo-antiserum and 1 mL packed RBC from the B21 animal.
6. Rotate the tube for 30 min on a rotator.
7. Centrifuge the solution at 500xg for 8 min at 20° C.
8. Transfer the supernatant to a new tube. Discard the RBC.
9. Test the anti-B-F allo-antiserum for specificity by haemagglutination (see below) of various dilutions of the absorbed sera (1:2 to 1:128) against 2% blood suspensions from an animal carrying the B19 haplotype (positive control) and an animal carrying the B21 haplotype (negative control).
10. If there still seems to be a cross-reaction to B21 cells, it may be enough to dilute the absorbed allo-antiserum further to solve the problem. Otherwise, the test-absorption must be repeated.
11. If the absorption is acceptable, larger portions of absorbed allo-antiserum may be prepared. It is important that the antibody/antigen ratio is the same as in the test-absorption.
12. If an antibody is to be absorbed with various haplotypes, it may be an advantage to absorb it with one or two haplotypes first, and then test the allo-antiserum against the panel of cross-reactive haplotypes. Sometimes the cross-reactive unspecific antibodies can be removed at the first absorption.
13. The absorbed allo-antiserum is now “ready-to-use” in the haemagglutination test. The batches of “**ready-to-use**” allo-antisera are kept at -20°C.

## Haemagglutination test

1. Bleed 1-2 mL blood from animals to be serotyped and stabilise the blood in 1/5 vol. 3.8 % Sodium Citrate.
2. Bleed 1-2 mL blood from animals with known MHC haplotypes (controls) and stabilise the blood in 1/5 vol. 3.8 % Sodium Citrate. The control blood may be kept for 2 weeks in the refrigerator if the blood is transferred to a CPD (Citrate Phosphate Dextrose adenine) tube.
3. Produce a 2 % suspension of the blood samples in PBS buffer in 4 mL test tubes (2-3 drops of blood + 2 ml PBS buffer - i.e. the blood colour turns into light orange). All the blood suspensions must have exactly the same colour. The suspensions may be kept for a maximum of 2 days in the refrigerator.
4. Drop onto the glass slide one drop of "ready-to-use" allo-antiserum (the antibody) and one drop of 2 % blood suspension (antigen).
5. Mix the two solutions gently by rocking the glass slide back and forth.
6. Incubate the glass slide in a humid chamber for 30 min.
7. Defrost a 4 mL NCP tube and dilute the plasma with PBS buffer up to 10 mL (20% dilution).
8. Filtrate the solution through a piece of gaze placed at the bottom of a 20 mL syringe into a 50 ml tube.
9. Adjust the vol. to 20 mL again with PBS buffer and repeat the filtration step using a new piece of gaze and the same 50 mL tube.
10. After incubation for 30 min take the glass slide out of the humid chamber and add one drop of 20% NCP. Mix the solution by rocking the glass slide back and forth.
11. Leave the glass slide for another 30 min in the humid chamber.
12. The haemagglutination is read on the light table and categorised as very positive (++), positive (+), or negative (-). Please, be aware that a positive reaction may stick to the bottom so rock the glass slide back and forth when reading.
13. The slide may, if necessary, be placed in the humid chamber for another 10-15 min in order to have a more clear reaction, however; this may result in a fainter reaction.

## Serological MHC typing of offspring

As the antibody titre in the allo-antiserum may change when frozen, "ready-to-use" allo-antisera are always test-titrated before use. Testing is carried out using a panel of 6 – 10 different haplotypes with known positive and negative reactions. The typing of animals with unknown haplotypes can then be performed on the glass slides. The schema below shows a simple result of a serological MHC typing.

| Antibody/Antigen | B15 animal | B19 animal | B21 animal | Unknown | Unknown | Unknown |
|------------------|------------|------------|------------|---------|---------|---------|
| Anti-B-F15       | +++        | -          | -          | +++     | -       | -       |
| Anti-B-F19       | -          | +++        | -          | -       | +++     | -       |
| Anti-B-F21       | -          | -          | +++        | -       | -       | +++     |
| Anti-B-G15       | +++        | -          | -          | +++     | -       | -       |
| Anti-B-G19       | -          | +++        | -          | -       | +++     | -       |
| Anti-B-G21       | -          | -          | +++        | -       | -       | +++     |
| Typing result    | OK         | OK         | OK         | B15     | B19     | B21     |

## Restriction Fragment Length Polymorphism (RFLP) typing of MHC

### Introduction

The use of RFLP analysis to molecular characterisation of genetic variation in the MHC is a valuable supplement to serological analysis. The RFLP analysis can examine sub-regions of genotypes for which the production of serological reagents has been elusive, analyse MHC chromosomal recombinants, and categorise newly identified genotypes. Identification of new MHC haplotypes by RFLP analysis makes it easier to make an immunisation plan for generation of new allo-antisera to unknown MHC antigens. Actually,

this was the method of choice when we distinguished between 5 to 6 new MHC haplotypes in a commercial breeding flock from the Lohmann Company (unpublished results).

The method is very time consuming and expensive compared to the serological typing but we use the method routinely in the lab as a supplement when we have problems with the serological typing. In our lab all our chicken lines containing different MHC haplotypes have been analysed by this method and a system has been worked out to distinguish between the different MHC haplotypes using one single restriction enzyme (Bgl II) and cDNA probes for B-F, B-L and B-G (Juul-Madsen et al., 1993 and 2000).

## Equipment

- Micro-tube centrifuge
- High Speed centrifuge
- 37°C incubator
- Tubes (1.5 mL, 2.2 mL, 50 mL)
- Tea filters
- 10 ml pipettes
- 13 mL Sarsted centrifuge tubes
- Pasteur pipette
- Hybond N+ nylon membranes (RPN203B, **AM**)

## Reagents

- TKM-1 buffer (10mM Tris-HCl\*\* pH 7.6; 10mM KCl; 10mM MgCl<sub>2</sub>; 2mM EDTA)
- TKM-2 buffer (10mM Tris-HCl\*\* pH 7.6; 10mM KCl; 10mM MgCl<sub>2</sub>; 2mM EDTA; 0.4M NaCl)
- NP-40
- SDS
- 6M NaCl (It is OK with some undissolved crystals in the bottle)
- 2-propanol
- TE buffer (Sambrook et al., 1989)
- 10 mg/mL BSA
- 0.2 M Spermidine
- 0.5 M EDTA (Sambrook et al., 1989)
- 7.5M NH<sub>4</sub>-O-A
- 99% Ethanol
- DNA loading buffer (Sambrook et al., 1989)
- 0.25 M HCl
- 4M NaOH
- 10 mg/ml carrier DNA (Sambrook et al., 1989)
- <sup>32</sup>P labelled probe (Sambrook et al., 1989)

\*\* Trizma Base 99.9% from Sigma # T-8524

## Purification of genomic DNA with Sodium Chloride

This method is included because it is a very easy, safe and cheap method to get large amounts of genomic DNA but the method is not very known. The method is adapted from the paper of Lahiri and Nurnberg (1991).

1. Transfer a blood sample to a micro tube with a Pasteur pipette and centrifuge in a micro centrifuge at 300xg for 5 min.
2. Mix 200µl packed cells with 25mL TKM-1 buffer containing 2.5% NP-40 in 50mL tubes.
3. Centrifuge at 1000xg for 10 min at room temperature (RT).
4. Discard the supernatant and resuspend the pellet (the nuclei's) in 20 mL TKM-1 buffer and vortex thoroughly.
5. Centrifuge at 1000xg for 10 min at RT.

6. Discard the supernatant and immediately resuspend the pellet thoroughly in 300 $\mu$ L TKM-1 buffer.
7. Add 10 mL TKM-2 buffer containing 0.625% SDS and shake the tube vigorously – do not vortex at this step as it will break the DNA into smaller pieces.
8. Incubate for a minimum of 1 hour (possibly o/n) at 60°C in a water bath and shake a couple of times with your hands. The solution has to be viscously but clear for best results.
9. Transfer the DNA solution to a 13mL Sarstedt centrifuge tube containing 3.75mL 6M NaCl - mix well by shaking with your hands.
10. Centrifuge in a high-speed centrifuge at 10.844xg for 10 min at RT.
11. Filter the supernatant through a tea filter into a new 50 ml tube.
12. Add 1 vol. 2-propanol and mix carefully until the DNA precipitates.
13. Spool the DNA onto a Pasteur pipette, wash in 70% EtOH and dry the DNA by rotating the pipette on the inside of a clean tube.
14. Resuspend the DNA in 2 mL TE buffer in a 2.2 mL micro tube.
15. Incubate for 30 min at 65°C, then o/n at 4°C and shake until the DNA dissolutes.
16. If the DNA has not dissolved, filter the DNA solution through a tea filter.
17. Measure OD 260 and 280 nm.
18. Store the DNA in 2.2mL micro tubes at -20°C.

## Genomic DNA restriction digestion

1. Use 1.5 mL micro tubes.
2. 7.5 - 10 $\mu$ g DNA is digested in an end concentration of 0.075 - 0.1 $\mu$ g/ $\mu$ L.
3. Add in the following order:
  - dH<sub>2</sub>O
  - 0.1 $\mu$ g/ $\mu$ L BSA
  - 2 mM Spermidine
  - 1/10 end volume of 10 x restriction buffer
  - the DNA
4. Digest the DNA with a given restriction enzyme - 2U/ $\mu$ g DNA for 6-8 hours at an enzyme determined temperature.
5. Add an extra 2U/ $\mu$ g DNA and continue the digestion o/n.
6. Terminate the digestion by adding EDTA up to 10 mM, ½ volume 7.5M NH<sub>4</sub>-O-Ac, and 3 volume 99% Ethanol. Mix and incubate the samples for a minimum of 1 hour at -20°C.
7. Centrifuge the samples at 26384xg for a minimum of 15 min at 4°C.
8. Discard the supernatant and dry the pellet. The pellet must not dry completely as it then is difficult for genomic DNA to dissolve again.
9. Resuspend the pellet thoroughly in 20 $\mu$ L TE buffer containing 1:5 amount DNA loading buffer.
10. Incubate the samples for 30 min at 37°C with the lid open to evaporate the alcohol.
11. The samples are ready to be loaded on an 0.6 – 1% agarose gel containing 0.5  $\mu$ g/mL Etidium Bromide – the percentage of agarose depends on how often the restriction enzyme cut the DNA.

## Alkali Blotting - DNA hybridisation

1. Photograph the gel to check for successful digestion of the samples.
2. Place the gel in ½ L 0.25 M HCl on a shaking table for 20 min (gentle shaking). This treatment depurinizes the DNA and breaks the DNA into smaller pieces which makes the transfer to the Hybond N+ membrane easier.
3. Pour off the acid and rinse the gel several times in MilliQ H<sub>2</sub>O to remove all the acid.
4. Make a Southern Blot according to the manufacturer's description for alkali blotting using Hybond N+ and 0.4M NaOH.
5. Prehybridise in 10 - 20 mL prehybridisation buffer (6 X SSC; 1% SDS; 10 x Denhardt; 200 -400  $\mu$ L Carrier DNA\*\* (10mg/mL) at 67°C in shaker water bath for 2 – 4 hrs.
6. Hybridise in 10 – 20 mL hybridisation buffer (prehybridisation buffer + <sup>32</sup>P-labelled probe\*\*) at 67°C in shaker water bath over night.
7. Wash the filter in a 55°C hot water bath after the following program:

- 2 times in 1 x SSC, 0.1% SDS for 25 min
- 2 times in 0.1 x SSC, 0.1% SDS for 25 min.
- 2 times in 0.1 x SSC for 25 min.

8. The filter is then ready for autoradiography.

\*\* Boil the probe and the carrier DNA for 5 min and place on ice for 5 min. Add <sup>32</sup>P-labelled probe until there is 6-8 \* 10<sup>6</sup> cpm/ml.

## Micro-satellite fragment typing of MHC

### Introduction

Micro-satellites are DNA sequences consisting of stretches of repeated short nucleotide motifs that may show polymorphism in the number of repeats. Micro-satellites are usually embedded in DNA with a unique sequence and can therefore easily be amplified by PCR. The LEI0258 micro-satellite (Zoorob et al., 1998; EMBL Accession no. Z83781) is a polymorphic 12-mer located within the B region of the chicken MHC on chromosome 16. The length of the LEI0258 was found to differ in 13 different MHC haplotypes from inbred lines with the exception of B2 that was identical with B15 (unpublished results). The use of fluorescent-labelled primers and automatic gel scanning makes this technique an attractive and rapid alternative to the other two MHC typing techniques.

### Equipment

- PCR instrument
- 0.2 µL PCR tubes
- Automatic Gel Scanner (e.g. ALFexpress® DNA Sequencer AM)
- Analysis software (e.g. ALFwin Fragment Analyser 1.0 AM)

### Reagents

- QIAamp®DNA Mini Kit (**QI**)
- PCR primers: [f]: 5'-(Cy5)-CACGCAGCAGAACTTGGTAAGG-3'  
[r]: 5'-AGCTGTGCTCAGTCCTCAGTGC-3'
- Taq DNA Polymerase (27-0799-05, **AM**)
- PCR nucleotide mix (US77170, **AM**)
- PCR buffer (50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 10 mM Tris-HCl)
- Polyacrylamide (e.g. Reprogel High resolution, **AM**)
- Loading Buffer (5 mg/ml of dextran blue in 100 % deionised formamide)
- Fluorescent 50 bp fragment ladder (e.g. ALFexpress™Sizer, 27-4539-01, **AM**)

### PCR amplification of micro-satellite

1. Purify genomic DNA from tissue or blood samples with the QIAamp® DNA Mini Kit according to the manufacturer's instructions (15 - 40µg DNA).  
Alternatively genomic DNA can be purified from red blood cells by the sodium chloride method mentioned above.
2. Set up a PCR reaction in a total volume of 25 µL PCR buffer including:
  - 1 µL of genomic DNA
  - 0.05 µM of each primer
  - 0.2 mM of each dNTP (PCR nucleotide mix)
  - 1 U of Taq DNA polymerase
3. Apart from samples, a panel of DNA from known MHC haplotypes can be included in the PCR step.

4. Perform PCR with a reaction profile of 94°C/5 min followed by 94°C/1 min, 56°C/1 min, 72°C/2 min for 25 cycles and finally 72°C/10 min.

## Fragment analysis

5. Use 1.5 µL of the PCR reaction and mix it with 3-4 µL of Loading Buffer
6. Denature samples at 95°C for 2 min.
7. Keep samples on ice for 2 min.
8. Separate PCR products on polyacrylamide gel along with 50 bp fragment ladder and controls.
9. Evaluate results on automatic gel scanner.
10. Deduce MHC type of unknown samples by comparing fragment sizes with known MHC controls.

## Suppliers

**AM:** Amersham Biosciences, Little Chalfont, UK  
**ME:** Medinor ASA, Oslo, Norway  
**QI:** Qiagen, GmbH, Hilden, Germany  
**TE:** Terumo Europe NV, Paris, France

## Concluding remarks

- The three typing methods do not replace but complement each other depending on the purpose of the typing.
- The fragment analysis may also be run on a 3% metaphor agarose gel or a 20 cm 6 – 9% polyacrylamid gel stained with ethidium bromide. A 40 cm 4% denaturing polyacrylamide gel visualised by silver staining may also be used.
- A table of inbreed and partly inbreed chicken lines with known MHC haplotype in the whole world will be published by Marcia Miller, USA (manuscript in preparation).

## Acknowledgements

The author's wish to thanks the technicians Lene Rosborg Dal and Hanne Svenstrup - the real experts in the lab for critical reviewing of the methods.

## References

Briles W.E., Goto R.M., Auffray C. & Miller M.M. (1993) A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. *Immunogenetics* **37**, 408-14.

Bumstead N, Reece RL, and Cook JKA (1993). Genetic difference in susceptibility of chicken lines to infectious bursal disease virus. *Poultry Science* **72**: 189-193.

Cotter P F, Taylor RL, and Abplanalp H (1998). B-complex associated immunity to Salmonella enteritidis challenge in congenic chickens. *Poultry Sci.* **77**:1846-1851.

Fillon V., Zoorob R., Yerle M., Auffray C. & Vignal A. (1996) Mapping of the genetically independent chicken major histocompatibility complexes B<sup>Q</sup> and RFP-Y<sup>Q</sup> to the same microchromosome by two-color fluorescent in situ hybridisation. *Cytogenetics Cell Genetics* **75**, 7-9.

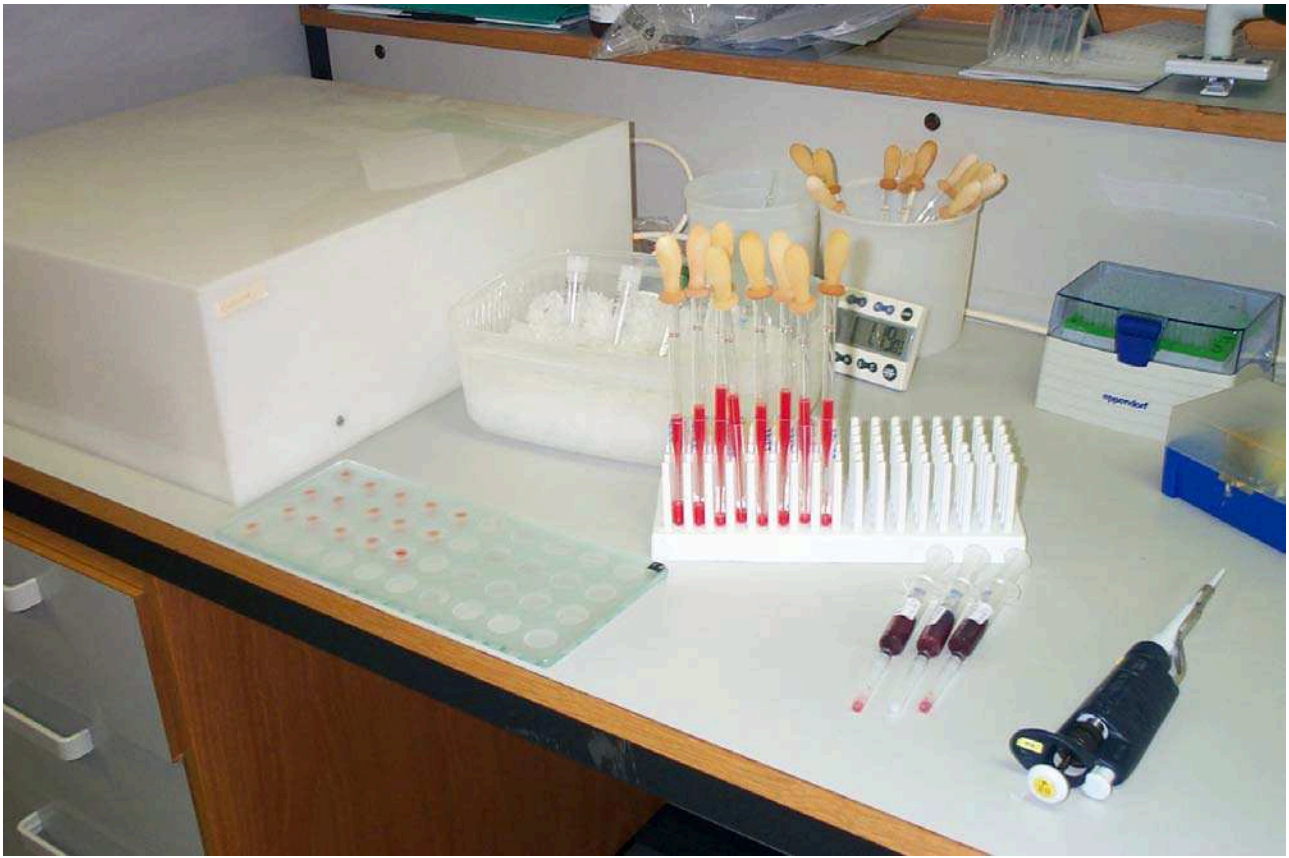
Guillemot F., Kaufman J.F., Skjødt K. & Auffray C. (1989) The major histocompatibility complex in the chicken. *Trends in Genetics* **5**, 300-04.

Juul-Madsen HR, Hedemand JE, Salomonsen J, and Simonsen M (1993). RFLP analysis of the chicken B-F and B-L genes and their association with serologically defined B haplotypes. *Animal Genetics* **24**: 243-247.

Juul-Madsen HR, Afanassieff M, and Dalgaard TS (2000). Molecular characterisation of major and minor MHC class I and II genes in B21-like haplotypes in chickens. *Animal Genetics*, **31**: 252-261.

- Juul-Madsen HR, Nielsen OL, Krogh-Meibom T, Rontved CM, Dalgaard TS, Bumstead N, and Jørgensen PH (2002). Major histocompatibility complex-linked response of young chickens vaccinated with an attenuated live infectious bursal disease virus vaccine followed by an infection. *Poultry Science* 81: 649-656.
- Kaufman J., Salomonsen J. & Skjødt K. (1989) B-G cDNA clones have multiple small repeats and hybridize to both chicken *MHC* regions. *Immunogenetics* 30, 440-51.
- Kaufman J and Wallny HJ (1996). Chicken MHC molecules, disease resistance and the evolutionary origin of birds. *Current Topics in Microbiological Immunology* 212: 129-141.
- Kroemer G, Guillemot F and Auffray C (1990). Genetic organisation of the Chicken MHC. *Immunol Res* 9: 8-19.
- Lahiri DK and Nurenberg Jr. JI (1991). A rapid non-enzymatic method for the preparation of HMW DANN from blood for RFLP studies. *Nucl Acid Res* 19: 5444.
- Lamont SJ (1998). The chicken major histocompatibility complex and disease. *Rev. Sci. Tech.* 17:128-142.
- Lamont S. J., C. Bolin, and N. Cheville (1987). Genetic resistance to fowl cholera is linked to the major histocompatibility complex. *Immunogenetics* 25:284-289.
- Lillehoj H S, Ruff MD, Bacon LD, Lamont SJ, and Jeffers TK (1989). Genetic control of immunity to *Eimeria tenella*. Interaction of MHC genes and non-MHC linked genes influences levels of disease susceptibility in chickens. *Vet. Immunol. Immunopathol.* 20:135-148.
- Miller M.M., Goto R., Bernot A., Zoorob R., Auffray C., Bumstead N. & Briles W.E. (1994). Two *MHC* class I and two *MHC* class II genes map to the chicken Rfp-Y system outside the B complex. *Proceedings of the National Academy of Science USA* 91, 4397-401.
- Salomonsen J, Dunon D, Skjødt K, Thorpe D., Vainio O, and Kaufman J (1991). Chicken MHC encoded B-G antigens are found on many cell types that are important for the immune system. *Proc Natl Acad Sci* 88: 1359-1363.
- Sambrook J, Fritsch EF, and Maniatis T (1989). *Molecular Cloning – A Laboratory Manual*, second edition Cold Spring Harbor Laboratory Press.
- Schierman LW, and Collins LW (1987). Influence of the major histocompatibility complex on tumor regression and immunity in chickens. *Poultry Sci.* 66:812-818.
- Simonsen M (1975). Serological identification of the major genetic locus. *Acta Pathol Microbiol Scand* 83: 1-14.
- Yoo B H, and Sheldon BL (1992). Association of the major histocompatibility complex with avian leukosis virus infection in chickens. *Br. Poult. Sci.* 33:613-620.
- Zoorob R, Auffray C, Chaussé AM (1998). Réactifs et méthodes pour la détection des gènes liés au CMH d'oiseaux d'élevage tels que le poulet. Patent PCT/FR98/02501, France.

Figure 1: Glass plates







## ANNEXE 2. Liste des amorces utilisées en PCR et en RT-PCR

| Cible        | Amorce          | Séquence              | Position extrême          |        |
|--------------|-----------------|-----------------------|---------------------------|--------|
| PCR          | <i>TRIM27.1</i> | AE01F                 | GACAGATCTTTAGGGGAGATCTAAA | 93756  |
|              |                 | AE01R1                | TCAATTGCAAGAGCCAGGTA      | 94845  |
|              |                 | AG01F                 | GCTCTTGCAATTGAAACATCAG    | 94832  |
|              |                 | AG01R                 | CCAAGAAGGAAGGGGAAGAGG     | 95804  |
|              | <i>TRIM41</i>   | AG02F                 | GCCTGGACTACTTCGTGGAG      | 99098  |
|              |                 | AG02R                 | TCCTCCTCCATCACATCCTC      | 100000 |
|              |                 | AG03F                 | CTTTGGGGACGATGATTATGA     | 99959  |
|              |                 | AG03R                 | TCCTGCTCATTGACTTCTG       | 100910 |
|              |                 | AG04F                 | GGAGCCACTGAAGAAGAAGC      | 100854 |
|              |                 | AG04R                 | CCGTCTGCTTGCTCTTATC       | 101643 |
|              |                 | AG05F                 | CTCCACCAGTTCCTCATTGG      | 101477 |
|              |                 | AG05R                 | GTATTTCTTCCCCACGTCCA      | 102725 |
|              |                 | AG06F                 | TGAGAACATCAAATTCCAGGAG    | 102668 |
|              |                 | AG06R                 | TGATACGAGTGCCCTTAGCC      | 103670 |
|              | <i>B-BTN2</i>   | AA01F                 | CTGCTGCTTCCAGAGCAAAG      | 124464 |
|              |                 | AA01R                 | AGGATGCCGAGGAGAGAAAT      | 125324 |
|              |                 | AA02F                 | TGTCAGAATTCCGTCAGCTC      | 125177 |
|              |                 | AA02R                 | CCTACAGATGCTCCCTCTTCC     | 126298 |
|              | <i>Blec1</i>    | AB01F                 | TGGAGGAAGTGAGGGAATATC     | 141450 |
|              |                 | AB01R                 | CTGTCAAACACAGCCAAGGA      | 142627 |
|              |                 | AB02F                 | GTGTCCCTTCGACTGGATTG      | 142505 |
|              |                 | AB02R                 | TCAGGTGCTGATGCAGAAGT      | 143522 |
|              | <i>TAPBP</i>    | AG07F                 | GCATCATCCCATAGTTCACC      | 147700 |
|              |                 | AG07R                 | CAGCATCCCAATACCTCCAC      | 147210 |
|              | <i>B-DMA</i>    | AD01F                 | GCCATTGGCTAAACTGGTTG      | 163239 |
|              |                 | AD01R                 | TTGGCTTCAGGCATTGGGTTT     | 164059 |
|              |                 | AD02F                 | ATGGACATCCCAACTGATGC      | 164511 |
|              |                 | AD02R                 | GGATCAACTTATGGGGGACA      | 165379 |
|              | <i>B-DMB1</i>   | AF02F                 | AAGGGGTTTCAGGGTGCT        | 165629 |
|              |                 | AF02R                 | TGTTTGCAATGCTCCACTTG      | 166454 |
|              |                 | AF03F                 | ATCCCCAGCTCTGACCATC       | 166934 |
|              |                 | AF03R                 | AGGAGGTACACCCCACACTG      | 167508 |
|              | <i>B-DMB2</i>   | AI03F                 | AATGCAGGTGCCTTCATGGTG     | 168939 |
| AI03R        |                 | CCATTTGCAATGCTCACTCCT | 169244                    |        |
| AF01F        |                 | TCAGCCCCAGGTCCGCATCGT | 170152                    |        |
| AF01R        |                 | GAGAGATGGGGTCACTCAGC  | 171035                    |        |
| <i>TAP1</i>  | AI01F           | ACCCTGTCCATCACATCCAT  | 176578                    |        |
|              | AI01R           | ATACCTACCCCGTCTGTCC   | 176978                    |        |
|              | AI02F           | GGACAGGACGGGGTAGGTAT  | 176959                    |        |
|              | AI02R           | GGGACATTGACAGAACAGG   | 177409                    |        |
| <i>TAP2</i>  | AH01F           | GGGTCTGGTGGTTACTGGTG  | 180131                    |        |
|              | AH01R           | AGAGGTGGATGGGGTTTTGT  | 181026                    |        |
| <i>CD1A1</i> | AC01F           | ATCTCCGCACCAAGACTACG  | 239051                    |        |
|              | AC01R1          | GGGACCATCAGGACCTTCTT  | 239830                    |        |
|              | AC02F           | AGCTTTGGCTGACAACTGGT  | 239717                    |        |
|              | AC02R1          | AAGGTGACTGGGACGTCATC  | 240839                    |        |
|              | AC03F           | CCTGCTGAAATCCCTCACAT  | 240781                    |        |
|              | AC03R1          | CATGCCATCCTCCTTCTCTT  | 241242                    |        |

|               | <b>Cible</b>  | <b>Amorce</b>          | <b>Séquence</b>             | <b>Position extrême</b> |
|---------------|---------------|------------------------|-----------------------------|-------------------------|
|               | LEI0258       | LEI0258-F              | CACGCAGCAGAACTTGGTAAGG      | 116745                  |
|               |               | LEI0258-R              | AGCTGTGCTCAGTCCTCAGTGC      | 117113                  |
|               | MCW0370       | MCW0370-F              | GAAAGAGGAGAGTAGTTCACG       | 129262                  |
|               |               | MCW0370-R              | TCCCACCTCATCATGCATTCC       | 129442                  |
|               | MCW0371       | MCW0371-F              | TTTCATGGCATCCTAAGATGG       | 127190                  |
|               |               | MCW0371-R              | CTGCTCCGAGCTGTAATCCTG       | 127394                  |
|               | GABI0001      | GABI0001-F             | CATTCCCCATTAACGCACTC        | 164706                  |
|               |               | GABI0001-R             | AGAGTGTTGGGGTAGCCGAG        | 164878                  |
| <b>PCR</b>    | <i>B-F</i>    | BF $\alpha$ -1-8 F     | GAGCTCCATACCCTGCGGTACATC    | 187620                  |
|               |               | BF $\alpha$ -180-189 R | CCCCCACACTCGCACCTCGGGCCGCTC | 172966                  |
|               |               | BF-exon-3A F           | TTCCAGAGGCAGTTCCCACCAA      | 186991                  |
|               |               | BF-EXON-8b R           | AGCCAAACTGGGACACGGTC        | 185540                  |
|               | <i>B-LB</i>   | BLB-1F                 | GGCCGTGCTGGTGGCACT          | 151209                  |
|               |               | BLB-1R                 | GTACGTCCAGTCCCCGTTCT        | 152015                  |
|               |               | BLB-2F                 | GACAGCGACGTGGGGAAA          | 151585                  |
|               |               | BLB-2R                 | GGGGCAGCAGCTAATTCA          | 152541                  |
|               | <i>B-DMA</i>  | MAalt-F                | GATTGGCCACCCGCC             | 162129                  |
|               |               | MAalt-R                | TCAGAGCCGGTAGGGGAC          | 163337                  |
|               |               | MA-1F                  | GCCATTGGCTAAACTGGTTG        | 163239                  |
|               |               | MA-2R                  | TTGGCTTCAGGCATTGGGTTT       | 164059                  |
|               |               | MA-3Fbis               | ATGGACATCCCAACTGATGC        | 164511                  |
|               |               | MA-4R                  | GGATCAACTTATGGGGGACA        | 165379                  |
|               | <i>B-DMB1</i> | MB2-1F                 | AAGGGGTTTCAGGGTGCT          | 165629                  |
|               |               | MB2-2R                 | TGTTTGCAATGCTCCACTTG        | 166454                  |
|               |               | MB2-3F                 | ATCCCCAGCTCTGACCATC         | 166934                  |
|               | MB2-5R        | AGGAGGTACACCCCACACTG   | 167508                      |                         |
| <i>B-DMB2</i> | MB1-2F        | AATGCAGGTGCCTTCATGGTG  | 168939                      |                         |
|               | MB1-2R        | CCATTTGCAATGCTCACTCCT  | 169244                      |                         |
|               | MB1-3F        | TCAGCCCCAGGTCCGCATCGT  | 170152                      |                         |
|               | MB1-4R        | GAGAGATGGGGTCACTCAGC   | 171035                      |                         |
| <b>RT-PCR</b> | <i>B-DMA</i>  | RT-BMAalt-F            | CATGAAGATTCACGGCATT         | 162352                  |
|               |               | RT-BMA-F               | CTCATTGATGTCCCCAACT         | 163889                  |
|               |               | RT-BMA-R               | GTCCTGTGGCACCCAGTAAG        | 165200                  |
|               | <i>B-DMB1</i> | RT-R4-2F               | GATGATGGTGAAGGTGGCAGT       | 167339                  |
|               |               | RT-18B-5R              | CATGGGAGTTCAGTCATTGA        | 167806                  |
|               | <i>B-DMB2</i> | RT-BMB1-F              | GGCTCAGCTTGCTCTTCATT        | 170573                  |
|               |               | RT-BMB1-R              | TTGGTGACAGAAGGTGTCCA        | 171083                  |

### ANNEXE 3. Activités connexes

#### BOURSES ET FINANCEMENTS

- 2009** Candidature acceptée pour la GMOD Summer School (Oxford, Royaume-Uni)
- 2008** Financement EADGENE, réseau d'excellence européen pour un séjour de courte durée (Foulum, Danemark)
- 2008** Programme jeune du XXIIIe World's Poultry Congress: voyage, inscription, logement (Brisbane, Australie)
- 2007** Financement EADGENE pour le Summer Institute in Statistical Genetics européen (Liège, Belgique)
- 2006-2009** Financement du Ministère de la Recherche pour la thèse

#### ARTICLES

Solinhac R, Leroux S, Galkina S, **Chazara O**, Feve K, Vignoles F, Morisson M, Derjusheva S, Gaginskaya E, Bed'hom B, Vignal A, Fillon V, Pitel F (**2009**) Integrative analysis of chicken microchromosome 16: reconsidering the MHC organization. en cours de soumission.

**Chazara O**, Minvielle F, Roux D, Bed'hom B, Feve K, Coville JL, Kayang BB, Lumineau S, Vignal A, Boutin JM, Rognon X (**2009**) Evidence for introgressive hybridization of wild common quail (*Coturnix coturnix*) by domesticated Japanese quail (*Coturnix japonica*). *Conservation Genetics* in press.

**Chazara O**, Pinard-Van Der Laan MH, Tixier-Boichard M, Bed'hom B (**2008**) Molecular genotype investigation of the *Gallus gallus* Major Histocompatibility Complex. *Developments in Biologicals (Basel)* **132**:271-8.

**Chazara O**, Lumineau S, Minvielle F, Roux D, Feve K, Kayang BB, Boutin JM, Vignal A, Coville JL, Rognon X (**2006**) Etude des risques d'introgession génétique de la caille des blés (*Coturnix coturnix coturnix*) par la caille japonaise (*C. c. japonica*) : comparaison et intégration des données comportementales et moléculaires obtenues dans le Sud-Est de la France. *Les Actes du BRG* **6**:317-334.

#### PRESENTATIONS ORALES

Typage moléculaire du complexe majeur d'histocompatibilité chez le poulet : vers une meilleure caractérisation des génotypes. Huitièmes Journées de la Recherche Avicole, 25 et 26 mars **2009**, St Malo, France.

Genetic diversity of the *Gallus gallus* Major Histocompatibility Complex. 23rd World's Poultry Congress, 30 Juin-4 Juillet **2008**, Brisbane, Australie.

Diversité génétique structurale et fonctionnelle du CMH chez le poulet : implication pour la résistance aux maladies. 10<sup>e</sup> Séminaire des thésards du Département de Génétique Animale de l'INRA, 20-21 Mars **2008**, Toulouse, France.

Exploration de la diversité génétique du CMH chez le Poulet : vers une meilleure caractérisation des haplotypes. Séminaire Doc'J, 10-11 Mai **2007**, Jouy-en-Josas, France.

Etude des risques d'introggression génétique de la caille des blés (*Coturnix coturnix coturnix*) par la caille japonaise (*C. c. japonica*) : comparaison et intégration des données comportementales et moléculaires obtenues dans le Sud-Est de la France. 6e colloque national du BRG, 3-4 Octobre **2006**, La Rochelle, France.

#### **POSTERS**

**Chazara O**, Pinard-Van Der Laan M.H, Tixier-Boichard M, Bed'hom B. Molecular genotype investigation of the *Gallus gallus* Major Histocompatibility Complex. International Symposium on Animal Genomics for Animal Health, 23-25 Octobre **2007**, Paris, France.

**Chazara O**, Pinard-Van Der Laan M.H, Tixier-Boichard M, Bed'hom B. Comment mieux connaître le CMH chez le Poulet ? En analysant la variabilité génétique et fonctionnelle de la région. 9<sup>e</sup> Séminaire des thésards du Département de Génétique Animale de l'INRA, 25-26 Avril **2007**, Jouy-en-Josas, France.

#### **ENSEIGNEMENT**

**Octobre 2009** Cours pour les étudiants du Master II Erasmus Mundus d'AgroParisTech, Markers and Mapping

**Janvier 2009** Cours pour les étudiants du Master II d'AgroParisTech, Nouvelles techniques de séquençage : application à l'étude du CMH

**Automne 2008** TP, Modèles moléculaires: acides nucléiques, Université de Versailles St-Quentin (UVSQ)

**Automne 2008** TP, Modèles moléculaires: acides aminés, Université de Versailles St-Quentin (UVSQ)

#### **AUTRES EXPERIENCES**

**2008-2009** Présidente d'ADABIES, l'association des étudiants en thèse de l'école doctorale ABIES

**Nov-Dec 2008** Séjour (6 semaines) dans le « Laboratory of Disease mechanisms, markers and prevention » d'Helle Risdahl Juul-Madsen à Foulum, Danemark.