

P. CHARDON

INRA-CEA, Laboratoire de Radiobiologie  
et d'Etude du Génome,  
78352 Jouy-en-Josas cedex

e-mail : Patrick.Chardon@biotec.jouy.inra.fr

## 2 - Polymorphismes génétiques

# Le polymorphisme du complexe majeur d'histocompatibilité

**Résumé.** Le complexe majeur d'histocompatibilité est une région chromosomique riche en locus qui témoigne de l'évolution des génomes. Des duplications et réarrangements de cinq exons ancestraux ont donné naissance, chez les vertébrés, à plusieurs dizaines de gènes, dont les gènes d'histocompatibilité, et ont favorisé l'apparition de nouvelles fonctions biologiques comme la réponse immunitaire adaptative. Après l'émergence des mammifères, les gènes d'histocompatibilité ont continué à évoluer et certains ont acquis une fonction spécifique. Les gènes d'histocompatibilité sont caractérisés par un polymorphisme exceptionnel résultant de mécanismes actifs de génération de la diversité. Les multiples copies des gènes constituent un réservoir de séquences favorisant la création de nouveaux allèles. La variabilité est aussi entretenue par une transmission des haplotypes qui favorise les hétérozygotes.

## 1 / Structure et fonction des molécules d'histocompatibilité

Les molécules d'histocompatibilité ont été découvertes dans les années cinquante lors des premières transplantations d'organes. Ce sont les structures très polymorphes qui définissent la compatibilité tissulaire entre donneur et receveur. Des molécules identiques induisent une tolérance du greffon, tandis que des molécules issues d'allèles différents provoquent un rejet immédiat. A ce titre, elles constituent de véritables marqueurs du soi. Dans des conditions physiologiques normales, les molécules d'histocompatibilité participent à la défense de l'individu contre l'envahissement d'organismes étrangers (virus, bactéries, champignons et autres parasites). Elles présentent les peptides antigéniques aux cellules spécialisées du système immunitaire et déclenchent les réactions qui aboutissent à l'élimination de l'intrus. Elles signalent la présence de molécules inhabituelles, contribuant ainsi à la neutralisation des cellules anormales ou transformées.

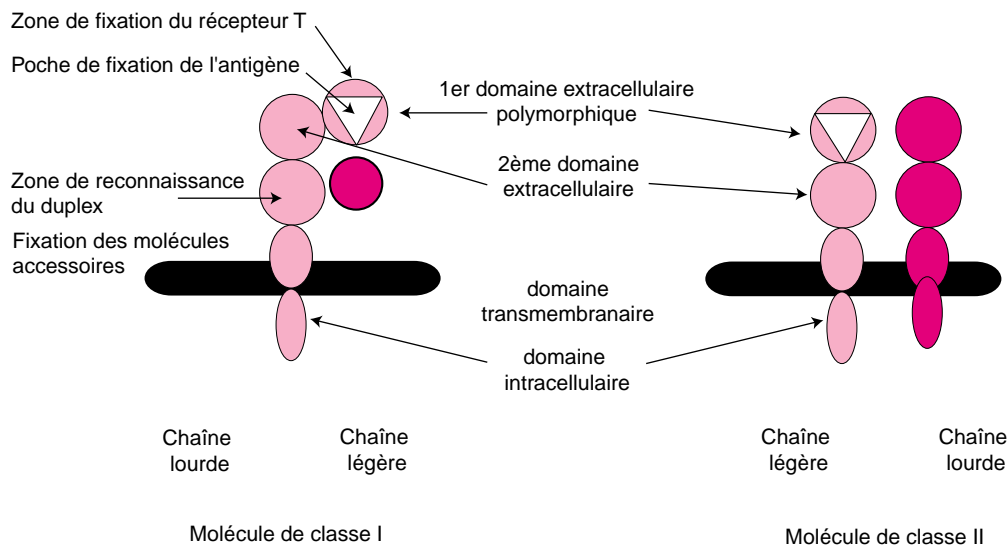
Les molécules d'histocompatibilité possèdent une origine évolutive commune et sont rattachées au groupe phylogénétique des immunoglobulines. Schématiquement, ce sont des glycoprotéines transmembranaires formées de deux chaînes polypeptidiques, l'une constante, l'autre variable (figure 1). Les molécules sont organisées en domaines. Le premier domaine, extra-cellulaire, porte l'essentiel du polymorphisme. C'est le lieu de fixation de l'antigène et du récepteur spécifique du lymphocyte. Les autres domaines sont très conservés. Ils favorisent les contacts cellulaires en fixant des molécules accessoires de la membrane des lymphocytes.

Suivant des critères biochimiques, on regroupe les molécules d'histocompatibilité en deux familles distinctes, les molécules de classe I et de classe II. La fonction de chacune des deux classes est sensiblement différente. Les molécules de classe I présentent principalement des antigènes d'origine endogène, c'est-à-dire des particules qui pénètrent seules dans les cellules, comme les virus ; les molécules de classe II servent à éliminer des antigènes captés par des cellules spécialisées de la surveillance immunitaire.

Si l'on va plus loin dans les comparaisons de séquences, on peut subdiviser chaque classe en sous-groupes. Cependant, le nombre et la fonction de ces sous-groupes varient chez les mammifères. Il existe ainsi, chez l'Homme, trois groupes fonctionnels de classe II : DR, DQ et DP, alors que le porc et la souris ne possèdent que les séries DR et DQ. Les molécules de classe I se scindent en trois sous-groupes : classique (I-a), non classique (I-b) et divergente (I-c). Cette distinction tient compte à la fois des séquences, de la distribution cellulaire et du polymorphisme. En général, les molécules I-a sont très polymorphes, les I-c ont un assez grand nombre d'allèles et les I-b sont monomorphes. Le nombre et la fonction de ces molécules varient sensiblement selon les espèces. Les molécules I-b de l'Homme n'ont pas d'équivalent fonctionnel chez le porc et la souris, et inversement. La souris ne possède pas de molécules I-c.

## 2 / Génétique et évolution de la région du CMH

La quasi-totalité des gènes qui spécifient les molécules d'histocompatibilité sont réunis sur un segment

**Figure 1.** Les molécules de classe I et de classe II du CMH.

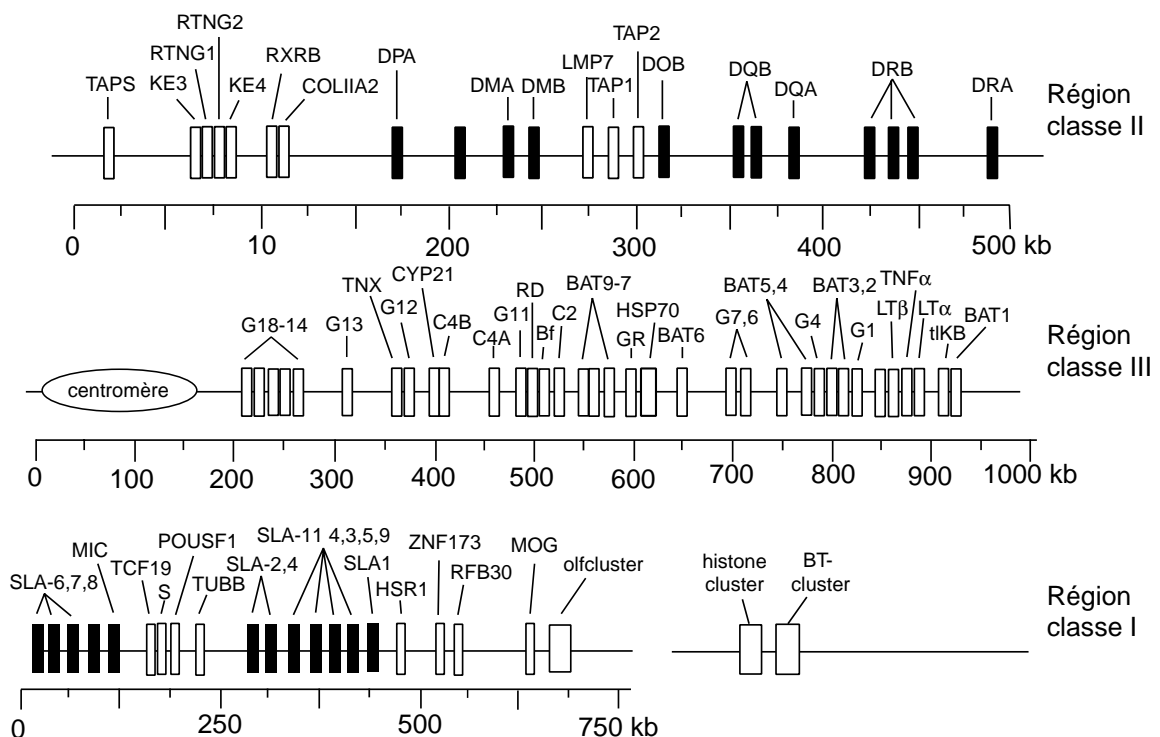
chromosomique unique qui forme le Complexe Majeur d'Histocompatibilité (CMH), appelé HLA (Human Leucocyte Antigen) chez l'Homme, H2 (Histocompatibility 2) chez la souris ou SLA (Swine Leucocyte Antigen) chez le porc.

Le CMH s'étend sur plusieurs millions de nucléotides et comporte plusieurs dizaines de gènes d'histocompatibilité intercalés entre de nombreux autres locus (figure 2). Les séquences d'histocompatibilité sont clairement le produit de duplications plus ou moins récentes comprenant les locus fonctionnels mais aussi de nombreux gènes tronqués et des pseudogènes (Chardon *et al* 1999).

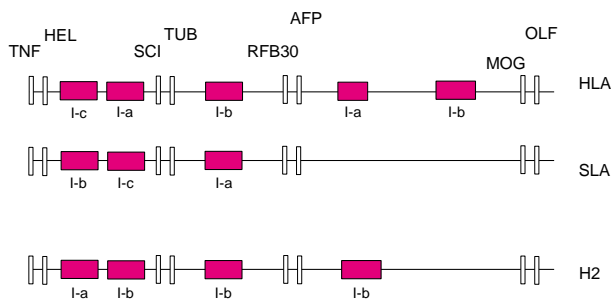
L'organisation générale du complexe est conservée chez les mammifères et, dans une moindre

mesure, dans l'ensemble des vertébrés. Il comporte trois régions : la région I qui contient les gènes de classe I, la région II qui abrite les gènes de classe II et la région III localisée entre les deux premières. L'ordre des gènes est le même, à l'exception notable des gènes de classe I. Il n'y a pas, en effet, de relation d'orthologie entre ces séquences : des loci situés dans une position homologue peuvent avoir une fonction différente (figure 3).

La phylogénie des gènes non associés à l'histocompatibilité donne des éléments pour reconstituer l'histoire de la région et ouvre des pistes pour comprendre l'origine du polymorphisme. La trace d'un chromosome ancestral, formé de gènes indépendants d'un point de vue évolutif, est retrouvé chez

**Figure 2.** Carte du complexe SLA du porc. Les rectangles noirs représentent les gènes d'histocompatibilité, les blancs

**Figure 3.** Distributions des familles de gènes de classe I dans les complexes HLA, SLA et H2. Les clusters de gènes de classe I classiques (I-a), non classiques (I-b) ou divergents (Ic) sont représentés en couleur, les gènes d'ancrages en blanc.



les vertébrés et même chez des animaux plus primitifs (figure 4). A l'intérieur de cette trame, on trouve chez les vertébrés toute une série de gènes qui sont le produit de séries de duplication et d'assemblage de cinq séquences codant pour des domaines protéiques (motifs à doigt de zinc RING, domaine constant et variable d'immunoglobuline, motif de fixation de peptides, motif B30). La juxtaposition de ces domaines protéiques a créé de nouvelles molécules et de nouvelles fonctions, comme les molécules d'histocompatibilité et la fonction immunitaire adaptative (Abi-Rached *et al* 1999).

Chez les mammifères, on trouve un grand nombre de séquences ou de groupe de séquences dupliquées dans la région du CMH. Ce sont, au premier chef, les gènes d'histocompatibilité mais aussi la butyrophiline (BUT), des protéasomes (LMP) des transporteurs de peptides (TAP), des protéines de choc thermique (HSP), des cytokines TNF, du complément (C4), les récepteurs olfactifs et bien d'autres loci. Le nombre de ces séquences est variable d'une espèce à l'autre.

Enfin, à l'intérieur de chaque espèce, il n'est pas rare de trouver des amplifications de groupes de gènes. Ainsi, chez l'Homme, le nombre de locus DRB peut aller de un à cinq en fonction des haplotypes. Ces séquences additionnelles ne sont pas forcément toutes fonctionnelles. Chez le porc, on peut détecter de un à quatre gènes de classe I-a exprimés. Le bloc C4-CYP21 (facteur du complément - enzyme du métabolisme des stéroïdes) dans la région de classe III est dupliqué en tandem chez la plupart des mammifères étudiés, mais pas chez le porc. A l'inverse, on observe de nombreuses délétions dans ce même bloc. Ces événements peuvent induire des troubles métaboliques importants.

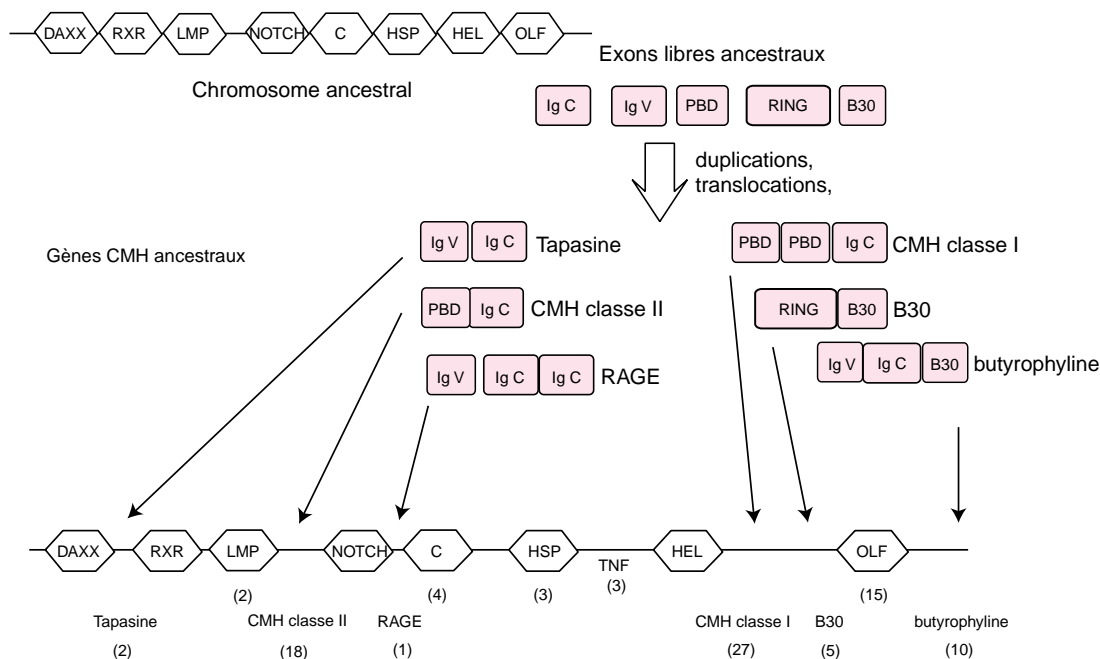
Le segment chromosomique du CMH paraît donc soumis à différentes forces évolutives, l'une semblant maintenir la structure, l'autre entraînant une diversification par fixation des séquences dupliquées.

### 3 / Polymorphisme des gènes d'histocompatibilité : impact des mutations ponctuelles

Une des caractéristiques du CMH est son extraordinaire polymorphisme. On compte, chez l'Homme, des dizaines d'allèles aux loci de classe I-c, de classe II DRB, DQB et plusieurs centaines aux trois principaux loci de classe I-a. Chez les animaux domestiques, des études moins exhaustives révèlent aussi une variabilité très importante. La règle n'est cependant pas généralisable à toutes les espèces. Il existe des cas où le polymorphisme MHC est extrêmement faible, en particulier pour quelques espèces qui ont été menacées d'extinction à un moment de leur histoire.

Les comparaisons de séquences donnent des informations sur l'origine du polymorphisme. Tout

**Figure 4.** Evolution possible de la région CMH. Les gènes (fond blanc) qui ne contiennent pas un exon IgC, IgV, RING, B30 ou PBD constituent le chromosome ancestral. Les gènes créés à partir de ces exons sont en couleur. (n) : nombre de gènes dans HLA.



d'abord, l'analyse des séquences des gènes de classe I-a montre que la diversité est aussi importante à l'intérieur d'une série allélique qu'entre les séries et qu'il n'y a pas de véritable motif nucléotidique caractéristique d'une série. Il est par conséquent difficile de distinguer des allèles vrais et des pseudo-allèles. Par ailleurs, la répartition des mutations n'est pas homogène tout au long de la séquence et certains segments de la molécule sont très polymorphes, tandis que d'autres sont très conservés. Ainsi, les domaines trans-membranaires, intracellulaires et le troisième domaine extra-cellulaire sont très conservés et ne diffèrent que par un nombre restreint de mutations ponctuelles. Ces dernières ont généralement peu d'impact sur le fonctionnement de la molécule, sauf lorsqu'une substitution de nucléotides empêche l'accrochage des molécules accessoires. Il existe un exemple de mutation qui génère un codon de terminaison prématuré au niveau du domaine intra-membranaire. La molécule produite n'est pas ancrée dans la membrane et figure sous forme soluble dans la circulation générale.

L'essentiel de la variabilité est en réalité apporté par le premier et deuxième domaine extra-cellulaire qui fixent l'antigène. Le nombre de résidus différents entre les séquences est souvent très élevé et, en certaines positions clés, le nombre de substitutions possibles est considérable. On peut trouver ainsi plus de 10 acides aminés alternatifs aux positions les plus polymorphes (figure 5). Par ailleurs, dans ces domaines, le rapport mutation synonyme/non synonyme est inversé par rapport au reste de la molécule. Toutes ces observations indiquent que le nombre et la diversité des allèles ne résultent pas de mutations spontanées. Le polymorphisme dans la région du CMH est vraisemblablement la conséquence d'un mécanisme actif de génération du polymorphisme.

## 4 / Génération de la diversité

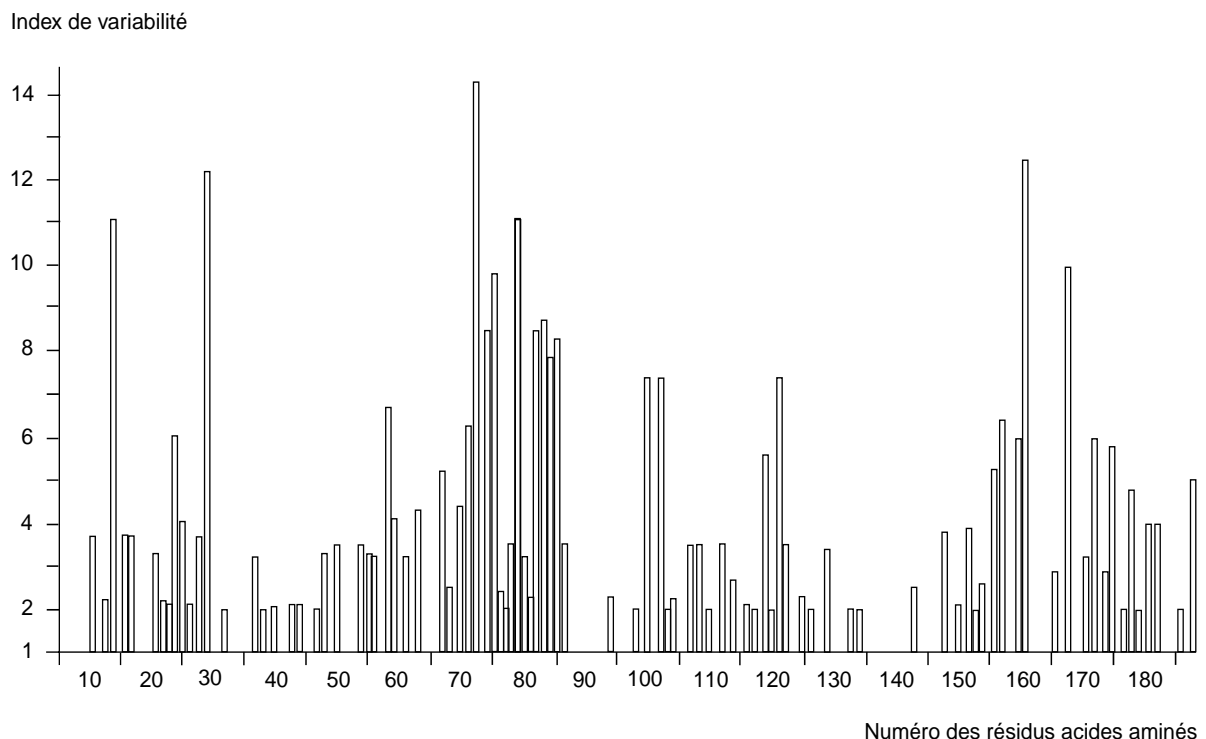
Les comparaisons des séquences des gènes de classe I-a de l'Homme et de leurs équivalents chez des grands singes anthropoïdes ont montré qu'il existe des allèles communs aux différentes espèces. Dans une moindre mesure, on trouve aussi des séquences remarquablement voisines entre l'Homme et les bovins. Ces relations peuvent être la conséquence d'une évolution convergente, mais le passage de groupes d'allèles durant la spéciation des mammifères est une hypothèse retenue (Miko *et al* 1999).

Un autre mécanisme de génération du polymorphisme fait appel à des recombinaisons non équilibrées, ou conversions géniques. Il n'est pas rare de trouver un motif nucléotidique commun à plusieurs allèles d'un locus dans la séquence d'un allèle à un autre locus. On imagine que de tels motifs sont transposés d'un locus à l'autre à la suite de recombinaisons non équilibrées. Il y a ainsi une création de nouveaux allèles très différents des séquences d'origine. Les nombreuses copies de gènes d'histocompatibilité fonctionnels, pseudogènes ou même les fragments de gènes constituent un réservoir de séquences qui contribuent à l'émergence rapide de nouveaux allèles (Hogstrand *et al* 1999).

De façon surprenante, les gènes de classe I de la série I-b, structures très proches des gènes I-a, sont pour l'essentiel monomorphes. En ce qui concerne les molécules de classe II, la chaîne légère a une variabilité élevée tandis, qu'à l'inverse, la chaîne lourde est monomorphe ou très peu polymorphe.

Les multiples loci et allèles forment par combinaison un très grand nombre d'haplotypes chez l'Homme. Dans les populations de mammifères

**Figure 5.** Variabilité de la molécule SLA de classe I-a. La variabilité est indiquée sous forme d'un index qui tient compte du nombre d'acides aminés possibles pour chaque position et de leur fréquence. La figure regroupe l'analyse de 34 allèles.



d'élevage analysées, le nombre d'haplotypes est plus limité, sans doute à cause d'un effet fondateur, mais peut-être aussi de la pression de sélection. La fréquence des haplotypes SLA du porc varie fortement dans les différentes races et évolue dans le temps. Par ailleurs, il existe un déficit en homozygotes et une distorsion de ségrégation dans la descendance pour certains haplotypes SLA.

Enfin, les recombinaisons ne se font pas au hasard dans le CMH. Il y a des points préférentiels de recombinaisons dans les complexes HLA et H2, mais aussi, vraisemblablement, dans SLA. Ces changements de fréquence de recombinaison modifient dans des proportions importantes les relations entre distance génétique et distance physique à l'intérieur du CMH.

## 5 / Variabilité des fonctions

La majorité des gènes de la région CMH interviennent à différents niveaux de la réponse immunitaire. Plusieurs codent pour des facteurs de la réponse innée, ou non spécifique, comme des cytokines, des facteurs du complément sanguin et de l'inflammation. D'autres agissent dans la réponse spécifique. C'est le cas des molécules d'histocompatibilité et d'un certain nombre de molécules qui font partie du processus de préparation des peptides antigéniques. Il est donc tout à fait logique de considérer la région du CMH comme une véritable entité fonctionnelle.

Par ailleurs, il existe dans le CMH de nombreux gènes actifs dans d'autres domaines de la physiologie, comme le développement embryonnaire, la croissance et la reproduction. Le rôle exact de ces gènes est souvent difficile à identifier. Le TNF « tumor necrosis factor », par exemple, a une activité pléiotropique : il intervient dans l'élimination des cellules tumorales, dans la différenciation des lymphocytes et des adipocytes, dans le métabolisme énergétique, la mobilisation des réserves corporelles et l'appétit. La molécule AIF1, surtout connue comme agent inflammatoire lors de rejet de greffes, a un rôle très important dans la régulation du taux d'insuline et le catabolisme des sucres. Il existe enfin dans le CMH de nombreux facteurs de transcription, activant de multiples voies métaboliques.

Depuis longtemps, le polymorphisme de la région du CMH a été utilisé comme marqueur pour repérer des QTL, notamment chez la souris et le porc. Plusieurs QTL de composition corporelle, de croissance, de reproduction ont été associés à des haplotypes du CMH, mais les gènes réellement impliqués dans ces associations restent à identifier. La diversité des loci dans le CMH ne facilite pas la tâche, et les gènes candidats ne font pas défaut. Reste à savoir si l'abondance des QTL dans cette région du génome est fortuite ou si elle est dépendante de la fonction d'histocompatibilité.

## Références

Abi-Rached L., McDermott M. F., Pontarroti P., 1999. The MHC big bang. *Immunological Reviews*, 167, 33-44

Chardon P., Renard C., Vaiman M., 1999. The major histocompatibility complex in swine. *Immunological Reviews*, 167, 179-192.

Miko S., Roed K., Schmutz S., Andersson L., 1999. Monomorphism and polymorphism at mhc DRB loci in domestic and wild ruminants. *Immunological Reviews*, 167, 169-178.

Hogstrand K., Bohme J., 1999. Gene conversion can create new MHC alleles. *Immunological Reviews*, 167, 305-318.

