

UNIVERSITÉ DE NANTES

FACULTÉ DE MÉDECINE

**ANTIGÈNES TISSULAIRES DE GROUPE SANGUIN ABH ACTEURS DE LA  
PROTECTION INNÉE ANTIVIRALE : EXEMPLES DES CALICIVIRUS (NV  
ET RHDV) ET DU CORONAVIRUS SRAS-CoV**

THÈSE DE DOCTORAT

Ecole doctorale Chimie-Biologie  
Discipline Sciences de la Vie et de la Santé  
Spécialités Glycobiologie et Virologie

Présentée et soutenue  
publiquement par  
**Patrice GUILLON**

le 23 Octobre 2008, devant le jury ci-dessous

**Président :**

**M. Antoine BLANCHER**, Professeur des Universités, Hôpital de Rangueil, Toulouse

**Rapporteurs :**

**Mme Evelyne KOHLI**, Professeur des Universités, Université de Bourgogne, Dijon

**M. Antoine BLANCHER**, Professeur des Universités, Hôpital de Rangueil, Toulouse

**Co-directeur de thèse :** **Mme Nathalie RUVOËN-CLOUET**, Maître de conférence, ENVN, Nantes

**Directeur de thèse :** **M. Jacques LE PENDU**, Directeur de recherche, INSERM U892, Nantes

« L'animal qui a le plus profité de la  
compagnie de l'homme est le microbe »

Malcolm de Chazal, *Penser par étapes*

« Les espèces qui survivent ne sont pas les  
espèces les plus fortes, ni les plus  
intelligentes, mais celles qui s'adaptent le  
mieux aux changements »

Charles Darwin

<b>CHAPITRE</b>	<b>1 :</b>	
<b>BIOCHIMIE DES GLYCANNES.....</b>		<b>1</b>
<b>1. Les glycoconjugués.....</b>		<b>1</b>
1.1. Les glycolipides.....		2
1.2. Les glycoprotéines.....		3
a - Les glycannes N-liés.....		3
b - Les glycannes O-liés.....		5
<b>2. Les glycosyltransférases.....</b>		<b>6</b>
2.1. Caractéristiques communes.....		6
2.2. Classification.....		8
a - Classification de Campbell.....		8
b - Classification CAZy (Carbohydrate Active enZymes).....		8
<b>CHAPITRE</b>	<b>2 :</b>	
<b>LES ANTIGÈNES DE GROUPES SANGUINS ABH.....</b>		<b>11</b>
<b>1. Historique.....</b>		<b>11</b>
<b>2. Structures et biosynthèse.....</b>		<b>13</b>
2.1. Les chaînes précurseurs.....		13
2.2. Biosynthèse des antigènes A, B et H.....		14
<b>3. Les glycosyltransférases du système ABH.....</b>		<b>15</b>
3.1. Les $\alpha$ 1,2 fucosyltransférases.....		15
3.2. Les $\alpha$ 1,3 glycosyltransférases A et B.....		17
<b>4. Gènes du système ABH et évolution.....</b>		<b>18</b>
4.1. Polymorphisme des gènes FUT1 et FUT2.....		18
4.2. Polymorphisme du gène ABO.....		20
4.3. HBGA ABH chez les autres espèces animales.....		24
a - Primates.....		26
b - Rat.....		27
c - Souris.....		27
d - Porc.....		28
e - Lapin.....		28
f - Vache.....		29

4.4. Histoires évolutives des gènes liés aux HBGA ABH.....	29
--	----

<b>CHAPITRE</b>	<b>3 :</b>
<b>DES GLYCANNES ET DES VIRUS : INTER-RELATIONS.....</b>	<b>35</b>

<b>1. Des sucres comme récepteurs aux virus.....</b>	<b>35</b>
1.1. Glycannes chargés.....	37
a - Acide sialique et virus Influenza.....	38
b - Héparane sulfate et Herpes Simplex Virus.....	39
1.2. Glycannes non chargés.....	40
<b>2. Glycosylation des virus.....</b>	<b>41</b>
2.1. Echappement par masquage des épitopes viraux.....	42
2.2. Facilitation par les lectines de l'hôte.....	43
2.3. Protection par les CBA.....	44
2.4. Reconnaissance par les anticorps naturels de l'hôte.....	46

<b>CHAPITRE</b>	<b>4 :</b>
<b>VIRUS ET HBGA : EXEMPLES DES CALICIVIRUS ET DU SRAS-CoV.....</b>	<b>47</b>

<b>1. Calicivirus et HBGA.....</b>	<b>47</b>
1.1. Taxonomie.....	47
a - Vesivirus.....	48
b - Lagovirus.....	48
c - Norovirus.....	50
d - Sapovirus.....	51
1.2. Génome viral et protéines associées.....	52
1.3. Morphologie et structure.....	54
a - La protéine majeure de capsid.....	55
b - Structure de la capsid.....	56
1.4. Interaction avec les HBGA.....	57
a - RHDV et HBGA.....	57
b - Norovirus et HBGA.....	59
<b>2. SRAS-CoV et HBGA.....</b>	<b>61</b>
2.1. Taxonomie.....	61
2.2. Génome viral et protéines associées.....	64
2.3. Morphologie et structure.....	67

2.4. SRAS-CoV : interactions virus/récepteurs.....	68
2.5. Sensibilité et résistance génétique au SRAS-CoV.....	71

<b>CHAPITRE</b>	<b>5 :</b>	
<b>TRAVAUX ET PUBLICATIONS.....</b>		<b>73</b>
<b>1. Objectifs.....</b>		<b>73</b>
<b>2. Articles.....</b>		<b>75</b>
2.1. Article 1 : La lipase stimulée par les sels biliaires ainsi que les mucines du lait de mères « sécréteur » inhibent la fixation du NV sur ses récepteurs glycaniques .....		75
<i>a - Introduction.....</i>		75
<i>b - Conclusions et perspectives.....</i>		84
2.2. Article 2 : Association entre expression de l'antigène tissulaire de groupe sanguin H, polymorphisme des $\alpha$ 1,2fucosyltransférases chez le lapin de garenne et sensibilité au RHDV.....		86
<i>a - Introduction.....</i>		86
<i>b - Conclusions et perspectives.....</i>		109
2.3. Article 3 : Inhibition de l'interaction entre la protéine Spike du SRAS-CoV et son récepteur cellulaire par des anticorps anti-antigènes tissulaires de groupe sanguin.....		111
<i>a - Introduction.....</i>		111
<i>b - Conclusions et perspectives.....</i>		139
<b>DISCUSSION GÉNÉRALE.....</b>		<b>141</b>
<b>BIBLIOGRAPHIE.....</b>		<b>154</b>

## Figures

FIGURE 1 : FORMULE TOPOLOGIQUE D'UN GLYCOSPHINGOLIPIDE.....	2
FIGURE 2 : VOIES DE SYNTHÈSE DES DIFFÉRENTES FAMILLES DE GLYCOSPHINGOLIPIDES.....	3
FIGURE 3 : EXEMPLE DE GLYCANNES N-LIÉS DE TYPE OLIGOMANNOSIDIQUES (A), COMPLEXES (B) OU HYBRIDES (C).....	4
FIGURE 4 : STRUCTURE DES HUIT CŒURS DES O-GLYCANNES COMPLEXES DE TYPE MUCINE.....	6
FIGURE 5 : STRUCTURE TRIDIMENSIONNELLE EN « RUBAN » D'UNE GLYCOSYLTRANSFÉRASE (TRITERPÈNE UDP-GLUCOSYLTRANSFÉRASE UGT71G1 [MEDICAGO TRUNCATULA]).....	7
FIGURE 6 : SCHÉMA DE SYNTHÈSE DES ANTIGÈNES ABH À PARTIR DU PRÉCURSEUR DE TYPE 2.....	15
FIGURE 7 : LOCALISATION DES GÈNES DES $\alpha$ 1,2 FUCOSYLTRANSFÉRASES SUR LE GÉNOME HUMAIN, STRUCTURE DES GÈNES FUT1 ET FUT2 ET REPRÉSENTATION SCHÉMATIQUE DES ENZYMES FUT1 ET FUT2.....	16
FIGURE 8 : REPRÉSENTATION SCHÉMATIQUE DU LOCUS ABO, DE L'ADN COMPLÉMENTAIRE DES ALLÈLES A, B ET O, AINSI QUE DE LEUR PRODUIT DE TRADUCTION RESPECTIF.....	18
FIGURE 9 : FRÉQUENCES DES DIFFÉRENTS ALLÈLES FUT2 (Se ET se) DANS PLUSIEURS POPULATIONS.....	20
FIGURE 10 : RÉPARTITION DES ALLÈLES A, B ET O SUR UNE ÉCHELLE MONDIALE.....	21
FIGURE 11 : DISTRIBUTION TISSULAIRE DES ANTIGÈNES DE GROUPES SANGUINS CHEZ DIFFÉRENTES ESPÈCES DE VERTÉBRÉS PLUS OU MOINS ÉLOIGNÉES D'UN POINT DE VUE PHYLOGÉNÉTIQUE.....	25
FIGURE 12 : ÉVOLUTION HYPOTHÉTIQUE DES GÈNES CODANT LES $\alpha$ 1,2 FUCOSYLTRANSFÉRASES.....	31
FIGURE 13 : LES DIFFÉRENTS MODES D'ÉVOLUTION DES FAMILLES MULTIGÉNIQUES.....	34
FIGURE 14 : STRUCTURE DE L'ACIDE SIALIQUE ET DE L'HÉPARANE SULFATE.....	37
FIGURE 15 : SCHÉMA DESCRIPTIF DES DIFFÉRENTS MODES DE TRANSMISSION DES IAV ENTRE LES ESPÈCES.....	39
FIGURE 16 : REPRÉSENTATION SCHÉMATIQUE DES HUIT PROFILS DE FIXATION OBSERVÉS CHEZ LES NoV ET MODÉLISATION DES GLYCANNES PRÉFÉRENTIELLEMENT RECONNUS PAR CES SOUCHES.....	41
FIGURE 17 : ACTION ANTIVIRALE DIRECTE DE LA PROTÉINE D DU SURFACTANT. ....	45
FIGURE 18 : ARBRE PHYLOGÉNÉTIQUE DU RHDV REPRÉSENTANT LES DIFFÉRENTS GÉNOGROUPE CIRCULANT OU AYANT CIRCULÉ EN FRANCE.....	49
FIGURE 19 : ARBRE PHYLOGÉNÉTIQUE DES NoV ET REPRÉSENTATION DES DIFFÉRENTS GÉNOGROUPE.....	51

<b>FIGURE 20 : ARBRE PHYLOGÉNÉTIQUE DES SaV ET REPRÉSENTATION DES DIFFÉRENTS GÉNOGROUPEs.</b>	<b>52</b>
<b>FIGURE 21 : REPRÉSENTATION DU GÉNOME DES CALICIVIRUS ET DE SON ORGANISATION : EXEMPLES DU NV ET DU RHDV.</b>	<b>54</b>
<b>FIGURE 22 : LA PROTÉINE DE CAPSIDE VP1 DU NV ET STRUCTURES TRI- ET BIDIMENSIONNELLES DE LA CAPSIDE DU NV.</b>	<b>56</b>
<b>FIGURE 23 : REPRÉSENTATION ET ORGANISATION DES DIMÈRES A/B ET C/C FORMANT LA CAPSIDE DES NoV.</b>	<b>57</b>
<b>FIGURE 24 : LOCALISATION MOLÉCULAIRE DES DOMAINES DE LIAISON AUX GLYCANNES DU VIRUS VA387 ET DU NV.</b>	<b>61</b>
<b>FIGURE 25 : ARBRE PHYLOGÉNÉTIQUE ÉTABLI SUITE À LA COMPARAISON DE 28 SÉQUENCES COMPLÈTES CODANT L'HÉLICASE DES CoV.</b>	<b>64</b>
<b>FIGURE 26 : ORGANISATION DES 14 ORF DU GÉNOME DU SRAS-CoV (RNA1) ET DE SES 8 ARN SUBGÉNOMIQUES.</b>	<b>66</b>
<b>FIGURE 27 : STRUCTURE SCHÉMATIQUE D'UN VIRION DE SRAS-CoV.</b>	<b>68</b>
<b>FIGURE 28 : REPRÉSENTATION DES DIFFÉRENTS DOMAINES DE LA PROTÉINE S ET LOCALISATION DE SES DIFFÉRENTS SITES POTENTIELS DE N-GLYCOSYLATION.</b>	<b>70</b>
<b>FIGURE 29 : COMPARAISON DE LA FIXATION DE LA PROTÉINE S DE DIFFÉRENTES SOUCHES DE SRAS-CoV SUR L'ACE2 D'ORIGINE HUMAINE OU DE RAT.</b>	<b>71</b>
<b>FIGURE 30 : REPRÉSENTATION SCHÉMATIQUE DU MÉCANISME DE COÉVOLUTION HÔTE/PATHOGÈNE.</b>	<b>143</b>
<b>FIGURE 31 : EFFET REINE ROUGE SUR LA DIVERSIFICATION DES GLYCANNES DE L'HÔTE EN RÉPONSE AUX ADAPTATIONS DES PATHOGÈNES.</b>	<b>145</b>
<b>FIGURE 32 : MÉCANISMES D'ÉCHAPPEMENT DE LA RECONNAISSANCE EXOGÈNE DES GLYCANNES SANS PERTE DE LA RECONNAISSANCE ENDOGÈNE.</b>	<b>146</b>
<b>FIGURE 33 : MODÉLISATIONS MATHÉMATIQUES DU RÔLE DES PATHOGÈNES DANS LE MAINTIEN DU POLYMORPHISME ABH ET DE LEURS FRÉQUENCES DANS LES POPULATIONS.</b>	<b>147</b>

## Tableaux

<b>TABLEAU 1 : CLANS DES GLYCOSYLTRANSFÉRASES DE LA BASE DE DONNÉES CAZY.....</b>	<b>9</b>
<b>TABLEAU 2 : LES SIX CHÂÎNES PRÉCURSEURS DES HBGA ABH.....</b>	<b>13</b>
<b>TABLEAU 3 : EXPRESSION DES HBGA ABH CHEZ LES PRIMATES NON-HUMAINS.....</b>	<b>27</b>
<b>TABLEAU 4 : LISTE DES DIFFÉRENTS GLYCANNES SERVANT DE RÉCEPTEURS OU CO-RÉCEPTEURS AUX VIRUS.....</b>	<b>36</b>
<b>TABLEAU 5 : DONNÉES ÉPIDÉMIOLOGIQUES MONDIALES DE L'ÉPIDÉMIE DE SRAS-CoV DE 2002/2003.....</b>	<b>63</b>
<b>TABLEAU 6 : DONNÉES STATISTIQUES COMPARANT LA PROBABILITÉ D'INFECTION PAR LE SRAS-CoV ET LE PHÉNOTYPE ABH DES INDIVIDUS AYANT ÉTÉ AU CONTACT DIRECT DU VIRUS.....</b>	<b>72</b>

## Abréviations

**AAV** : Virus associés aux Adénovirus  
**ACE2** : Enzyme 2 de conversion de l'angiotensine  
**Ala** : Alanine  
**Arg** : Arginine  
**Asn** : Asparagine  
**Asp** : Acide aspartique  
**BSSL** : Lipase stimulée par les sels biliaires  
**CBA** : Agents de fixation aux glycanes  
**Cer** : Céramide  
**CoV** : Coronavirus  
**Cys** : Cystéine  
**DC-SIGN** : protéine non-intégrine des cellules dendritiques liant ICAM3  
**EBHSV** : Virus du syndrome hémorragique du lièvre brun européen  
**EGFP** : Protéine à fluorescence verte augmentée  
**FCV** : Calicivirus félin  
**Fuc** : Fucose  
**Gal** : Galactose  
**GalNAc** : N-acétylgalactosamine  
**Glc** : Glucose  
**GlcNAc** : N-acétylglucosamine  
**Gly** : Glycine  
**GnT** : Glucosaminyltransférase  
**HBGA** : Antigène(s) tissulaire(s) de groupe sanguin  
**His** : Histidine  
**HIV** : Virus d'immunodéficience humaine  
**HLA** : Antigènes des leucocytes humains  
**HSV** : Herpes simplex virus  
**IAV** : Virus Influenza A  
**ICAM3** : Molécule 3 d'adhésion intercellulaire  
**Ig** : Immunoglobuline  
**L-SIGN** : Intégrine spécifique du foie et du nœud lymphatique liant iCAM3  
**Leu** : Leucine  
**LSECTin** : Lectine de type C des cellules endothéliales sinusoidales du foie et du nœud lymphatique  
**LTR** : Séquence terminale longue répétée  
**Man** : Mannose  
**Met** : Méthionine  
**NAb** : Anticorps naturels  
**NeuAc** : Acide N-acétylneuraminique  
**NoV** : Norovirus  
**NV** : Virus Norwalk  
**ORF** : Cadre ouvert de lecture  
**Pro** : Proline  
**RBD** : Domaine de liaison au récepteur  
**RBM** : Motif de liaison au récepteur  
**RE** : Réticulum endoplasmique  
**RHDV** : Virus de la maladie hémorragique du lapin  
**SaV** : Sapovirus  
**Se/se** : sécréteur/non sécréteur  
**Ser** : Sérine  
**SIDA** : Syndrome d'immunodéficience acquise  
**SINE** : Eléments nucléaires courts alternés = séquence répétitive de l'ADN  
**SIR** : Sensible/Infecté/Résistant  
**SMSV** : Virus San Miguel du lion de mer  
**SP** : Protéine du surfactant  
**SRAS** : Syndrome respiratoire aigu sévère  
**Thr** : Thréonine  
**Trp** : Tryptophane  
**UEA-I** : Agglutinine I d'*Ulex europaeus*  
**UTR** : Région non-traduite de l'ARN  
**VESV** : Virus porcin de l'exanthème vésiculeux  
**VLP** : Pseudo particule virale

## Avant propos

Les antigènes tissulaires de groupe sanguin ABH sont des glycanes portés par les glycoprotéines et glycolipides présents à la surface des cellules épithéliales ou dans les sécrétions de nombreuses espèces animales. Ces structures sont le résultat de glycosylations en chaîne à l'aide d'enzymes codées par des gènes extrêmement polymorphes. Depuis leur découverte par Landsteiner au début du 20<sup>e</sup> siècle, aucun rôle physiologique n'a pu être mis à jour pour expliquer leur maintien au cours de l'évolution. Ce mystère a fait naître nombre de spéculations des plus farfelues. Chacun des groupes sanguins a notamment été associé à des comportements et tempéraments psychologiques particuliers (Bourdel 1960) ou encore à des pratiques alimentaires spécifiques, adaptées d'après nos modes de vie préhistoriques de chasseur, agriculteur ou barbare (D'Adamo et coll. 1996) et utilisant hors contexte les données phénotypiques mondiales A, B, AB et O à l'origine de la caractérisation des flux migratoires humains ancestraux (Mourant et coll. 1983).

Les premières hypothèses solides pouvant expliquer le rôle des antigènes tissulaires de groupe sanguin sont liées à l'implication de ces glycanes dans les interactions hôte/pathogène (Mourant et coll. 1978; Blackwell 1989; Le Pendu 1989). Au cours de ma thèse, je me suis particulièrement intéressé aux inter-relations virus/glycanes et notamment à l'implication des antigènes tissulaires de groupe sanguin ABH dans la protection innée contre les infections virales. J'ai notamment étudié : Le rôle inhibiteur des glycoprotéines du lait maternel sur la fixation tissulaire du virus Norwalk ; L'impact des épidémies de la maladie hémorragique du lapin sur les profils phénotypiques et génétiques des populations sauvages de lapins de garenne ; L'action des anticorps naturels anti-A et/ou -B sur la fixation cellulaire du Coronavirus du syndrome respiratoire aigu sévère et modélisé leur impact sur la dynamique épidémique du virus.

Ce manuscrit a pour ambition de dresser un tableau général concernant la synthèse et l'évolution des antigènes tissulaires de groupe sanguin ABH. Il présente également comment les glycanes et les virus interagissent et les conséquences de ces interactions sur la transmission de ces derniers. Les rôles possibles des antigènes tissulaires de groupe sanguin ABH y sont détaillés et illustrés en partie par les résultats obtenus au cours de ma thèse. Finalement, leur rôle dans la protection innée antivirale, conséquence directe de leur coévolution avec les pathogènes et en particulier avec les virus, est discuté.

---

# **CHAPITRE 1 :**

# **BIOCHIMIE DES GLYCANNES**

## **1. LES GLYCOCONJUGUÉS**

Les glycoconjugués sont des molécules sur lesquelles ont été greffées des chaînes de sucres de manière covalente, grâce à l'action d'une ou d'un ensemble de glycosyltransférases. Ces modifications peuvent avoir lieu sur deux types de précurseur. D'une part, des aglycones protéiques qui une fois glycosylés donneront des glycoprotéines ou des protéoglycannes. D'autre part, des protéines membranaires glypiées ou des aglycones lipidiques qui donneront des glycolipides après glycosylation.

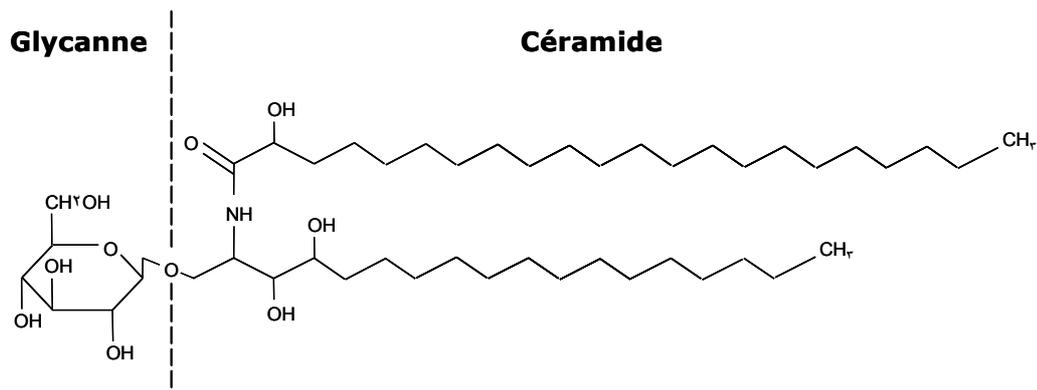
Les glycoconjugués ont une distribution ubiquitaire. Ils sont présents à l'extérieur de la cellule (glycocalix), au niveau de tous les compartiments intra-cellulaires, mais également dans le cytosol et le noyau (Hart et coll. 1989). Ces modifications physico-chimiques confèrent à ces molécules de nouvelles fonctions, extrêmement riches et variées. Elles vont jouer un rôle clef dans la modulation fine de certains processus biologiques comme la prolifération et la différenciation cellulaire, la croissance des tumeurs et la métastase, la coagulation, la réponse et la reconnaissance immune, la communication cellule/cellule, les interactions hôte/pathogène.

Au même titre que les gènes sont regroupés au sein du génome et les protéines au sein du protéome, les glycoconjugués sont répertoriés dans une grande banque de données, appelée le glycome (Hirabayashi et coll. 2001). Cette bibliothèque contient l'ensemble des chaînes glucidiques greffées sur les lipides et protéines synthétisées par un organisme. Chez certaines espèces, le glycome est 10 à 10<sup>4</sup> fois plus vaste que leur protéome (Raman et coll. 2005). En effet, lorsque la combinaison de 3 acides aminés

différents offre la possibilité d'obtenir 6 tripeptides distincts, la combinaison de 3 sucres différents peut quant à elle aboutir à l'obtention de milliers de glycannes distincts.

## 1.1. Les glycolipides

Les glycolipides synthétisés par les cellules de mammifères sont exclusivement des glycosphingolipides. Ils se composent d'une extrémité glycannique et d'une extrémité céramide (Figure 1).



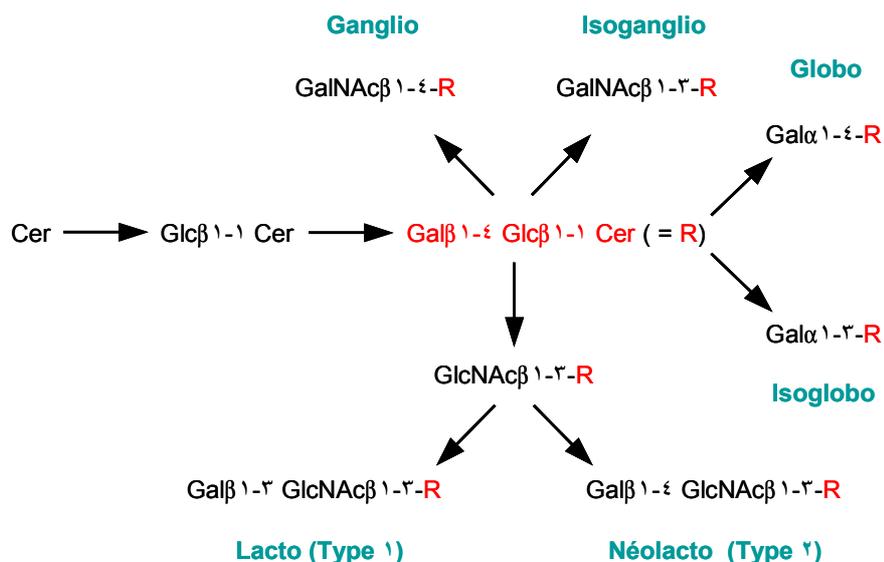
**Figure 1 :** Formule topologique d'un glycosphingolipide.

*Ce glycosphingolipide (glucosylcéramide) est composé d'un résidu glucidique ( $\beta$ -D-glucopyranose) et d'un groupement de type céramide. (D'après Holgersson et coll. 1992)*

Les glycosphingolipides sont synthétisés à partir d'un aminodiol, la sphingosine. Pour obtenir un céramide, un acide gras est ajouté au niveau de la fonction amine de cet alcool gras (amidification), via une liaison amide. Ensuite, un mono ou oligosaccharide est ajouté au niveau d'une des fonctions alcool du céramide (hémiacétalisation) via une liaison glycosidique.

*In vivo*, la synthèse des céramides a lieu au niveau de la face cytosolique du reticulum endoplasmique. L'ajout de la partie glycannique à quant à elle lieu au niveau de l'appareil de Golgi, ceci grâce à l'action séquentielle de différentes glycosyltransférases (Trinchera et coll. 1991; Sandhoff et coll. 1992).

Six familles de glycosphingolipides peuvent être distinguées : les ganglio, les globo, les isoglobo, les lacto et les isolacto (Figure 2). De nombreuses modifications périphériques peuvent intervenir au sein de chaque famille, augmentant ainsi la diversité structurale des glycolipides.



**Figure 2 : Voies de synthèse des différentes familles de glycosphingolipides.**

(D'après Holgersson et coll. 1992).

## 1.2. Les glycoprotéines

Les glycoprotéines peuvent être subdivisées en deux groupes selon le type de liaison qui existe entre le corps protéique et le sucre. On distingue de ce fait les glycoprotéines avec glycanes N-liés et celles avec glycanes O-liés.

### a - Les glycanes N-liés

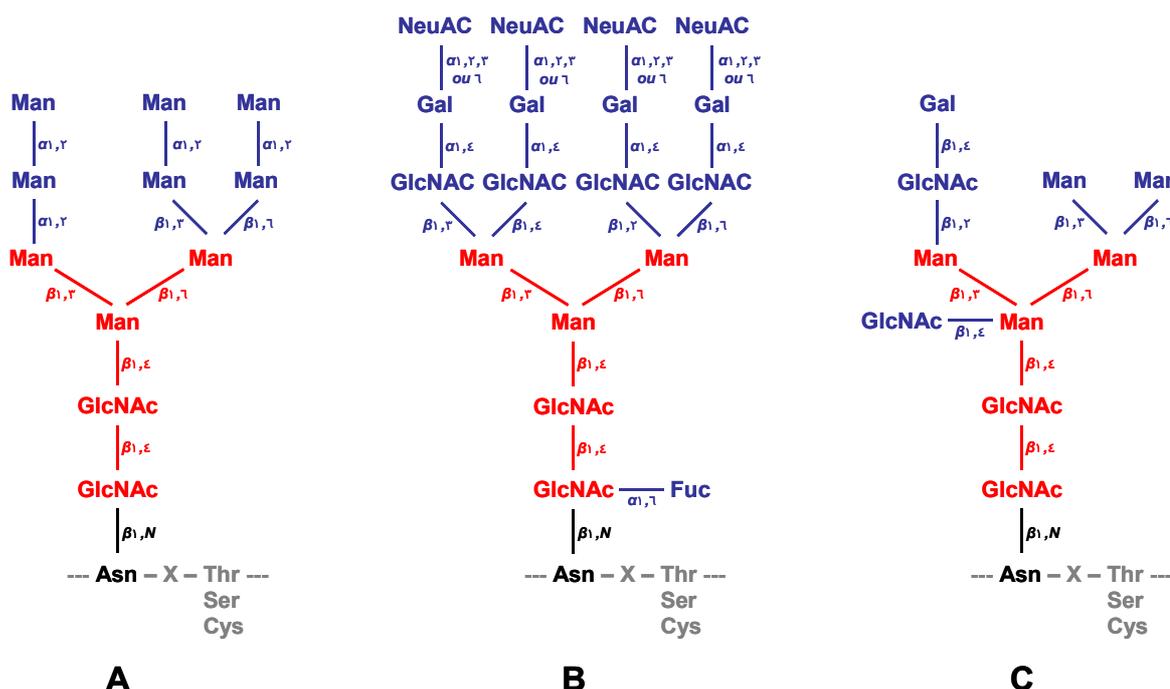
Les protéines N-glycosylées sont les glycoprotéines les plus représentées dans le monde du vivant. Elles sont très largement représentées chez les eucaryotes ainsi que chez les Archaea, mais ne sont que très rarement présentes chez les Eubactéries. Leur caractéristique commune est de posséder au moins une asparagine dont la fonction amine radicalaire est couplée à un polysaccharide par une liaison amine. Pour que ce branchement puisse avoir lieu, l'asparagine doit appartenir par ordre de préférence à l'une de ces séquences peptidiques consensus : Asn-X-Thr, Asn-X-Ser, et plus rarement Asn-X-Cys. X représente ici n'importe lequel des acides aminés, exceptés la proline et l'acide aspartique dans une moindre mesure (Kasturi et coll. 1997).

La biosynthèse des glycanes N-liés commence dans le lumen du RE par l'ajout sur l'asparagine d'un tétradécasaccharide précurseur ((GlcNAc)<sub>2</sub>-(Man)<sub>9</sub>-(Glc)<sub>3</sub>). Cette réaction, catalysée par une oligosaccharyltransférase, va conduire à la formation d'une liaison β1-N entre le GlcNAc terminal du précurseur et la fonction amine radicalaire de l'asparagine. Il s'en suit jusqu'au Golgi médian, une série d'élagages des sucres effectués par une batterie de glycosidases. Les modifications finales vont avoir lieu du Golgi

médian jusqu'au trans Golgi, par ajouts successifs de monosaccharides sur un cœur pentasaccharidique trimannosylé (Man<sub>3</sub>-GlcNAc<sub>2</sub>) commun à tous les glycanes N-liés.

Trois groupes de N-glycanes peuvent être distingués en fonction de la nature des sucres additionnels du coeur (Figure 3) :

- Le type oligomannosidique où seuls des résidus mannose sont ajoutés sur le précurseur commun.
- Le type complexe ou N-acétyllactosaminique qui résulte chez les mammifères de l'action de quatre glycosyltransférases (GnT I, II, III, IV) pour former des structures bi-, tri- ou tétra-antennées. Ces ramifications peuvent ensuite être complétées par l'addition de résidus Gal, GalNAc, Fuc, NeuAc, sulfate ou encore par l'accrochage d'unités répétitives (Galβ1-4GlcNAcβ1-3).
- Le type hybride dont la structure est la combinaison des deux précédents.



**Figure 3 :** Exemple de glycanes N-liés de type oligomannosidiques (A), complexes (B) ou hybrides (C).

Le cœur (en rouge) correspond au pentasaccharide commun à tous les N-glycanes. Les ramifications (en bleu) correspondent aux différents types d'antennes pouvant être greffées sur ce cœur.

Les glycanes oligomannosidiques et hybrides peuvent également subir l'action de la GnT III sur le mannose libre du cœur commun pour lui rajouter une N-acétylglucosamine en β (Figure 3C).

Les N-glycannes ont un potentiel de diversité structurale très grand. Cette variété est la combinaison du nombre variable de leurs antennes, de la nature des différents sucres les composant, ainsi que les différents types de liaisons qui peuvent exister entre eux.

### ***b - Les glycannes O-liés***

Les glycannes O-liés sont présents sur les glycoprotéines des membranes cellulaires et très largement représentés sur les mucines, protéines de haut poids moléculaire qui tapissent le tractus gastro-intestinal, ainsi que d'autres épithélium. Ils sont également retrouvés dans le sérum sur les immunoglobulines (Ig) A et D, ainsi que sur l'inhibiteur C1. Les autres protéines qui circulent dans le sang ne portent habituellement pas ce type de structure.

Les O-glycosylations sont des modifications exclusivement post-traductionnelles qui ont lieu au niveau de l'appareil de Golgi. Lorsqu'une protéine est O-glycosylée, le glycanne est lié sur la fonction hydroxyle radicalaire d'une sérine ou d'une thréonine. Il n'existe pas de motif peptidique consensus caractérisant les sites de O-glycosylation. Toutefois, les structures secondaires et tertiaires de la protéine semblent importantes dans les régions riches en sérine, thréonine et proline (Julenius et coll. 2005). De plus, le radical OH des acides aminés cibles ne doit pas porter de modifications, comme par exemple une phosphorylation. Le premier résidu O-lié sur le squelette peptidique peut être un fucose, une N-acétylglucosamine, un mannose ou une N-acétylgalactosamine (Peter-Katalinić 2005). Ce dernier est le plus fréquemment rencontré et caractérise les glycoprotéines de type mucine. Les glycannes des mucines peuvent être retrouvés sous différentes formes :

- Monosaccharidiques comme l'antigène Tn
- Disaccharidiques, comme les antigènes Sialyl-Tn et T
- Complexes, caractérisés par huit cœurs distincts (Yamashita et coll. 1995; Hounsell et coll. 1996) (Figure 4).

**Figure 4 : Structure des huit cœurs des O-glycannes complexes de type mucine.**

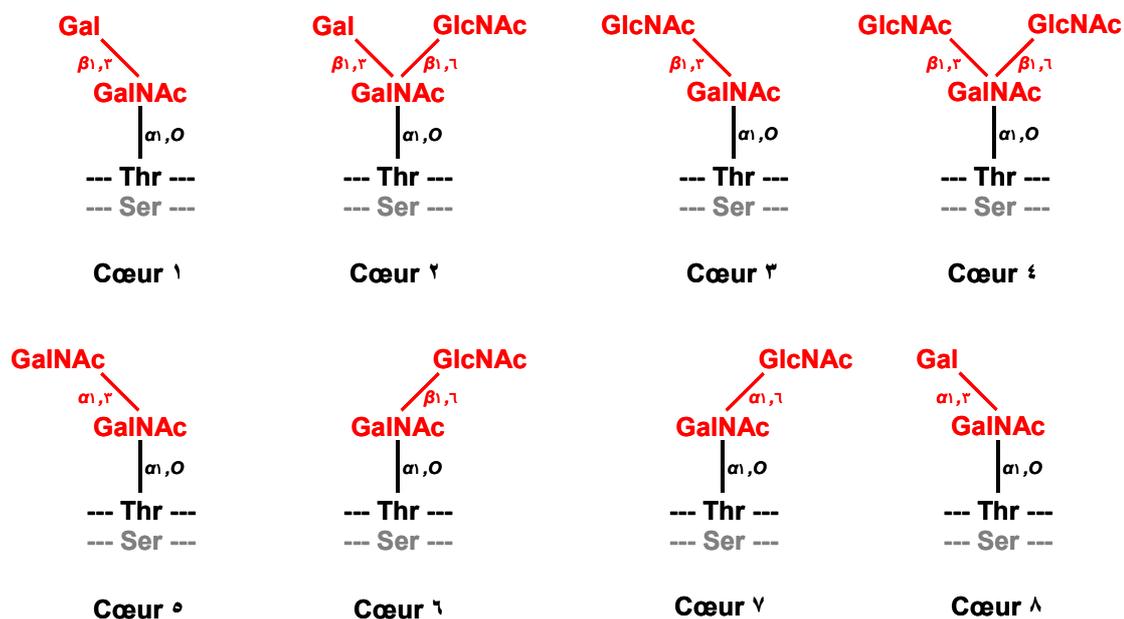
(D'après Peter-Katalinić 2005).

## 2. LES GLYCOSYLTRANSFÉRASES

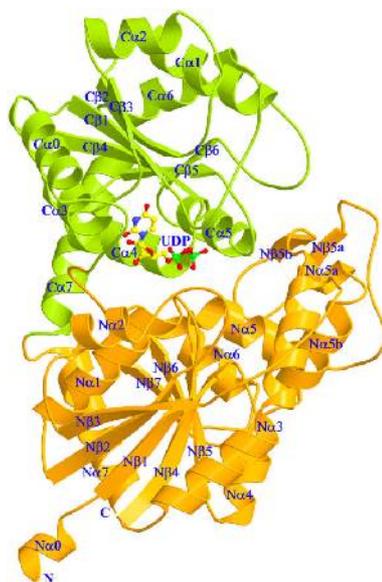
### 2.1. Caractéristiques communes

Les glycosyltransférases permettent de catalyser la synthèse des chaînes glucidiques sur les glycolipides et les glycoprotéines. Il en existe des centaines et chez l'homme, 250 gènes de glycosyltransférases ont été identifiés. Ces enzymes ont des spécificités de substrat strictes. Elles assurent le transfert d'un monosaccharide provenant d'un nucléotide de haute énergie donneur activé (glycosylnucléotide) vers un substrat accepteur en cours de synthèse. Elles sont regroupées en familles sur la base du sucre qu'elles vont ajouter au cœur osidique, mais également en fonction de la nature de la liaison utilisée (par exemple : la famille des  $\beta$ 1,4-galactosyltransférases ou des  $\alpha$ 1,3-galactosyltransférases).

Les glycosyltransférases vont avoir des propriétés différentes en fonction de leur localisation dans la cellule. Celles qui sont présentes dans le RE possèdent plusieurs domaines transmembranaires (Oriol et coll. 2002). Les glycosyltransférases golgiennes sont quant à elles des protéines de Type II, caractérisées par un domaine N-terminal intracellulaire généralement court, un domaine d'ancrage membranaire d'une vingtaine d'acides aminés, une tige flexible de longueur variable et un domaine C-terminal portant les deux sous-domaines catalytiques (Figure 5). Le domaine d'ancrage permet d'orienter le domaine catalytique vers la lumière du RE et du Golgi. Ces enzymes possèdent



également un site de clivage protéolytique au niveau de leur tige. L'action de protéases au niveau de ce site permet le relargage de ces enzymes dans la salive, le plasma ou encore le lait maternel.



**Figure 5 : Structure tridimensionnelle en « ruban » d'une glycosyltransférase (triterpène UDP-glycosyltransférase UGT71G1 [*Medicago truncatula*]).**

*La première hélice  $\alpha$  ( $N\alpha 0$ ) correspond au domaine transmembranaire de la protéine, elle est aussitôt suivie par la structure tige. Le reste de la molécule est organisé afin de former le site catalytique de l'enzyme. Il est divisé en deux sous-domaines, l'un reconnaissant le sucre accepteur (jaune), l'autre reconnaissant le glycosynucléotide donneur (vert). (D'après Shao et coll. 2005)*

Les gènes codant ces enzymes peuvent présenter deux types d'organisation. Certains vont être monoexonique alors que les autres seront polyexoniques (Joziassé 1992). Entre deux familles de glycosyltransférases d'une même espèce, on n'observe que très peu à pas d'homologie entre les séquences primaires. Or, à l'intérieur d'une même famille de glycosyltransférases, l'alignement des séquences peptidiques montre une très grande homologie et ceci bien qu'elles proviennent d'organismes très éloignés sur l'échelle de l'évolution. Par exemple, les sialyltransférases, assurant le transfert d'un acide sialique terminal sur les glycanes, sont très conservées entre la souris et l'homme (Harduin-Lepers et coll. 2005). Cette conservation est non seulement effective sur leurs séquences primaires, mais également sur la spécificité enzymatique qu'elles ont pour leurs substrats accepteurs.

Malgré cette conservation, on retrouve des variations importantes du profil d'expression des glycanes intra et inter-espèces. En effet, ceux-ci vont changer dynamiquement en réponse à de faibles variations de l'environnement extracellulaire,

mais aussi à cause de régulations post-transcriptionnelles liées à la stabilité des ARN messagers ainsi qu'à l'efficacité de leur traduction. Ces différences sont également liées à la nature même des voies de synthèse des glycanes qui est dépendante de réactions enzymatiques multiséquentielles et compétitives.

## **2.2. Classification**

Au vu du nombre croissant de glycosyltransférases découvertes, de la richesse de leurs propriétés enzymatiques, structurales et biochimiques, il est devenu très rapidement urgent d'instituer une classification précise de ces enzymes.

La première en date est celle de 1961, basée sur la nomenclature de l'IUBMB (International Union of Biochemistry and Molecular Biology). Celle-ci a permis de définir six familles auxquelles ont été attribuées un numéro EC (Enzyme Commission) allant de EC1 à EC6. Ces six familles sont : les oxydoréductases (EC1), les transférases (EC2), les hydrolases (EC3), les lyases (EC4), les isomérases (EC5), et les ligases (EC6). Récemment, d'autres modèles de classification ont été proposés.

### ***a - Classification de Campbell***

Campbell et collaborateurs ont proposé en 1997 un nouveau mode de classification, basé sur la similarité des séquences peptidiques primaires entre les différentes glycosyltransférases et s'inspirant du classement des hydrolases glycosidiques (J. A. Campbell et coll. 1997). Cette méthode a permis d'isoler vingt-sept familles distinctes à partir de l'étude de plus de 600 séquences primaires. Cependant cette méthode de classification a très vite trouvé ses limites, du fait de l'essor des outils bioinformatiques. En effet, ceux-ci ont permis d'identifier très rapidement de nombreuses nouvelles séquences par comparaison à celles utilisées par Campbell, sans pouvoir toutefois les rattacher à l'une des six classes.

### ***b - Classification CAZy (Carbohydrate Active enZymes)***

En 2003, Coutinho et collaborateurs proposent une nouvelle classification des glycosyltransférases du vivant. Accessible en ligne (<http://www.cazy.org/>) et régulièrement mise à jour, elle regroupe toutes les glycosyltransférases permettant une O-glycosylation et libérant une molécule phosphorylée à partir d'un donneur activé, ceci indépendamment de la spécificité de donneur ou d'accepteur (Coutinho et coll. 2003).

L'appartenance d'une glycosyltransférase à une classe particulière prend en compte différents critères :

- Sa spécificité de donneur ou d'accepteur. Une famille peut comprendre des enzymes aux spécificités différentes. On parle dans ce cas de famille polyspécifique.
- Le mécanisme moléculaire, « retaining » ou « inverting », par lequel le sucre est greffé sur l'accepteur.
- Son homologie avec d'autres séquences primaires.
- Son organisation structurale, en feuillets A ou B
- La nature de son domaine catalytique
- Sa modularité

A l'origine, cette base comportait 65 familles de glycosyltransférases. Aujourd'hui, elle en compte 91, identifiées GT1 à GT91 (environ 20 000 séquences) mais 597 séquences ne sont toujours pas rattachées à l'une d'entre d'elles. Les familles GT1 à GT27 correspondent à celles définies précédemment par Campbell.

Le classement des enzymes obéit à une hiérarchisation. Elles sont divisées en quatre clans selon qu'elles inversent ou conservent la conformation anomérique du sucre donneur et selon leur organisation structurale en feuillet A ou B (Tableau 1). Le feuillet A a été défini suite à l'identification de la protéine SpsA chez *Bacillus subtilis* (famille GT2). Le feuillet B a quant à lui été défini d'après l'étude de l'organisation structurale de la glycosyltransférase b du phage T4 et du domaine catalytique de la phosphorylase du glycogène.

	Mécanisme moléculaire	Organisation structurale
<b>Clan I</b>	inverting	feuillet A
<b>Clan II</b>	inverting	feuillet B
<b>Clan III</b>	retaining	feuillet A
<b>Clan IV</b>	retaining	feuillet B

**Tableau 1 : Clans des glycosyltransférases de la base de données CAZy.**

*La classification CAZy regroupe quatre grands clans de glycosyltransférases, caractérisés d'après l'organisation structurale de l'enzyme et la conformation anomérique du sucre donneur.*

La base de données CAZy permet de classer les glycosyltransférases en s'affranchissant de la définition de leur activité enzymatique. Elle est donc plus adaptée au classement des séquences primaires générées par analyse bioinformatique. De plus, les caractéristiques définissant l'appartenance à telle ou telle famille répondent à des critères plus stringents que ceux définis par Campbell. Cependant, elle ne permet pas de

prédire les caractéristiques enzymatiques d'une glycosyltransférase à la simple observation de sa séquence ou de son appartenance à une famille car celles-ci peuvent être polyspécifiques.

Dans la suite de ce manuscrit, nous nous intéresserons plus particulièrement aux glycannes correspondant aux antigènes tissulaires de groupes sanguins (HBGA) ABH, antigènes dont la synthèse implique des  $\alpha$ 1,2 fucosyltransférases de la famille GT11 et des  $\alpha$ 1,3 glycosyltransférases de la famille GT6.

# **CHAPITRE 2 :**

## **LES ANTIGÈNES DE GROUPES**

### **SANGUINS ABH**

#### **1. HISTORIQUE**

L'identification des différents HBGA ABH date du début du XX<sup>ème</sup> siècle. A cette époque, Karl Landsteiner, biologiste et médecin, s'interroge sur le fait que certaines transfusions entre humains sont de réels succès alors que d'autres conduisent à de graves complications pouvant aller jusqu'à la mort des patients.

C'est en 1900 que Landsteiner observe qu'en mélangeant le sérum de chacun de ses six collaborateurs avec les érythrocytes provenant de chacun des cinq autres, on obtient parfois une agglutination des globules rouges (Landsteiner 1901). Il émet alors l'hypothèse que les érythrocytes portent à leur surface une molécule qu'il appelle « agglutinogène », molécule qui réagit avec une substance présente dans le plasma de certains individus, les « agglutinines ». Il pose alors les bases de la future classification moderne des antigènes de groupes sanguins ABH, en concluant qu'il existe dans l'espèce humaine trois groupes d'individus différents lorsqu'on s'attache à regarder les propriétés d'agglutination de leur sérum vis a vis des globules rouges d'autres individus.

En effet, il montre que certaines personnes ont des érythrocytes qui ne sont jamais agglutinés alors que leurs sérums agglutinent les globules de tous les individus à l'exception de ceux possédant des sérums aux propriétés d'agglutination comparables aux leurs. Il nomme ce groupe « O » en référence au mot allemand « ohne » qui veut

dire « sans » et en l'occurrence sans agglutinogène. Il définit ensuite deux autres sous-groupes chez les personnes non-O, les groupes A et B de la classification moderne. Il montre qu'ils possèdent tous les deux des érythrocytes agglutinables par le sérum des individus O, mais également par les sérums des individus du second sous-groupe et inversement, ce qui sous-entend deux agglutinogènes différents.

C'est en 1902 que Decastello et Sturli, sous l'égide de Landsteiner, mettent en évidence un quatrième nouveau groupe, le groupe AB (Decastello et Sturli 1902). Ils décrivent que le sérum de ces individus n'est pas capable d'agglutiner les globules des trois autres groupes connus et que leurs érythrocytes sont agglutinés avec tous les sérums. Ils en concluent que chez ces personnes, le sérum ne contient aucune agglutinine alors que les globules rouges portent les deux agglutinogènes.

Dans les mêmes années, deux autres biologistes établissent une classification des individus en fonction des mêmes caractéristiques d'agglutinations des sérums décrites par Landsteiner et ses collaborateurs. Le premier, Jan Janský, sérologiste, neurologue et psychiatre, aurait été le premier à classer les antigènes de groupes sanguins en quatre catégories, les groupes I, II, III et IV (Janský 1907), mais cette découverte passa inaperçue à l'époque. Le second, William Lorenzo Moss, établissait la même classification que Janský, à l'exception que les groupes I et IV de Moss étaient inversés par rapport à ceux de Janský. Cette différence conduisit à des confusions lors des transfusions sanguines, ce qui conduisit à adopter la classification A, B, AB et O de Landsteiner comme standard universel (Crow 1993).

Landsteiner reçut en 1930 le prix Nobel de physiologie ou médecine pour l'ensemble de ses travaux, notamment ceux portant sur l'identification des antigènes de groupes sanguins ABH et Rhésus. En hommage à Jan Janský, les fréquents donateurs volontaires de sang tchèques se voient remettre une médaille en bronze à son effigie pour les récompenser de leur générosité.

Dans les années qui suivirent ces découvertes majeures, von Dungern et Hirschfeld mettent en évidence le caractère héréditaire de la transmission des antigènes ABH (von Dungern et Hirschfeld 1910, 1911) et c'est Felix Bernstein qui démontre formellement que cette transmission à la descendance est corrélée à la présence de multiples allèles (A,B et O) sur un même locus (F Bernstein 1924; F. Bernstein 1925). Quelques années après la découverte des antigènes de groupes sanguins, trois auteurs publient que ceux-ci peuvent être portés par d'autres cellules que les érythrocytes (Kritschewski et Schwartzmann 1927; Witebsky 1927) et c'est Hartmann qui définit le plus clairement comment ces antigènes sont exprimés au niveau de différents organes (Hartmann 1941). Watkins et Morgan montrent ensuite que les épitopes A et B sont des groupements sucrés, respectivement la N-acétylgalactosamine et le galactose (Watkins

et Morgan 1959). En 1990, Yamamoto identifie la séquence du locus *ABO* en clonant et séquençant l'allèle codant la forme soluble de la glycosyltransférase A (F. Yamamoto, Marken, et coll. 1990). La découverte de l'allèle B sera ultérieure et mettra en évidence que celui-ci n'entraîne qu'une différence de quatre acides aminés comparé au peptide codé par l'allèle A.

## 2. STRUCTURES ET BIOSYNTHÈSE

La synthèse des antigènes de groupes sanguins ABH passe par l'ajout successif de monosaccharides (Fuc, Gal et GalNAc) sur des chaînes précurseurs sucrées présentes sur les glycoprotéines et des glycolipides cellulaires (C. Green 1989). Il s'agit de modifications périphériques des glycolipides et des glycanes O- ou N-liés des glycoprotéines. Bien que découverts sur les érythrocytes, ces épitopes sont également exprimés par les cellules de différents tissus et excrétés sous forme libre ou complexée dans les fluides biologiques. Etant donné cette expression tissulaire, il est plus approprié de parler alors d'antigènes tissulaires de groupe sanguin ou « histo-blood group antigens » en anglais (HBGA).

### 2.1. Les chaînes précurseurs

Il existe six chaînes précurseurs connues pour pouvoir porter les HBGA ABH (Tableau 2).

Classe du précurseur	Structure et Composition
Type 1	Gal $\beta$ 1,3GlcNAc $\beta$ 1,R
Type 2	Gal $\beta$ 1,4GlcNAc $\beta$ 1,R
Type 3	Gal $\beta$ 1,3GalNAc $\alpha$ 1,R
Type 4	Gal $\beta$ 1,3GalNAc $\beta$ 1,R
Type 5	Gal $\beta$ 1,3Gal $\beta$ 1,R
Type 6	Gal $\beta$ 1,4Glc $\beta$ 1,R

**Tableau 2 : Les six chaînes précurseurs des HBGA ABH.**

*Les chaînes précurseurs des HBGA ABH sont toutes des disaccharides portant un Gal en position terminale. Le R correspond ici à un radical glycoprotéique ou glycolipidique.*

A l'origine, les disaccharides précurseurs de type 1 et 2 ont été découverts sur des glycoprotéines purifiées de kystes ovariens. La chaîne de type 3 correspond au cœur 1

des glycoprotéines O-liées. Les chaînes de type 4 ne sont présentes que sur les glycosphingolipides et particulièrement ceux exprimés au niveau du rein (King 1994). Les précurseurs de type 1 à 4 ont également été décrits pour être présents dans d'autres tissus (Le Pendu et coll. 1986; Oriol 1987). Le précurseur de type 5 est un pur produit de synthèse de la chimie organique. Il n'a jamais été observé sur du tissu humain, mais a notamment permis la caractérisation d'anticorps monoclonaux dirigés contre les HBGA ABH. Le précurseur de type 6, dont le sucre correspond à une molécule de lactose, a été mis en évidence sur les cellules intestinales et le tissu rénal. On le retrouve également dans les sécrétions exocrines comme le lait et l'urine (sous formes d'oligosaccharides libres) (Oriol 1990).

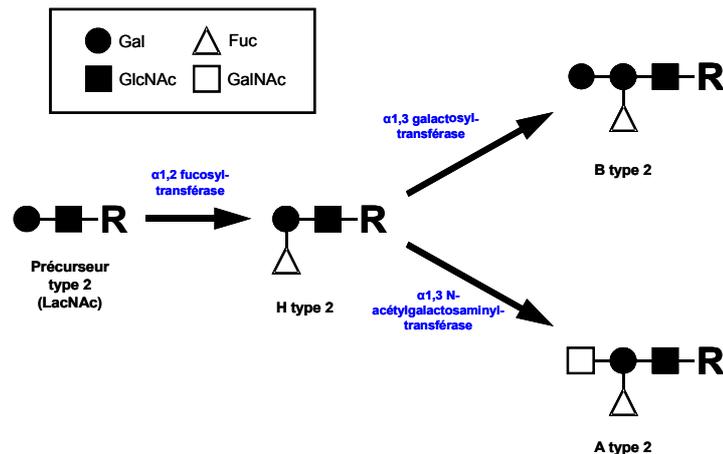
## **2.2. Biosynthèse des antigènes A, B et H**

L'antigène H peut être synthétisé à partir des six précurseurs décrits précédemment. Il est le produit d'une réaction enzymatique qui conduit à l'ajout d'un fucose sur le galactose terminal du précurseur, ceci grâce à l'action d'une  $\alpha$ 1,2 fucosyltransférase (Figure 6). Ces enzymes sont codées par deux gènes chez l'homme, *FUT1* et *FUT2*. Le produit de leur traduction correspond respectivement aux  $\alpha$ 1,2 fucosyltransférases FUT1 et FUT2. Chez d'autres espèces, un troisième gène, *Sec1*, a été caractérisé comme pouvant coder une enzyme (Sec1) possédant les mêmes propriétés catalytiques que FUT1 et FUT2 (Barraud et coll. 2000; Domino et coll. 2001; Hitoshi et coll. 1995). Ce gène est considéré comme un pseudogène chez les Catarhiniens et le porc. On en retrouve la trace dans leur génome mais il n'est pas traduit en protéine active (Apoil et coll. 2000).

L'expression de l'antigène H à la surface des érythrocytes et des cellules de l'endothélium vasculaire se traduit par le phénotype appelé « O », alors que son expression sur d'autres cellules ou dans les sécrétions se traduit par le phénotype « Sécréteur » (Se)

La production des antigènes A et B nécessite la synthèse préalable de l'antigène H. Ils sont obtenus respectivement suite à l'action d'une  $\alpha$ 1,3 N-galactosaminyltransférase et d'une  $\alpha$ 1,3 galactosyltransférase. La première ajoute une N-acétylgalactosamine sur le galactose de l'antigène H pour donner l'antigène A. La seconde ajoute un galactose sur le galactose de l'antigène H pour donner l'antigène B. Les séquences nucléiques codant ces deux enzymes sont situées sur le locus *ABO* et correspondent à des variants alléliques de ce gène.

L'expression des antigènes A ou B se traduit respectivement par les phénotypes A ou B, ceci indépendamment de la nature du glycoconjugués exprimant ces sucres. La co-expression des antigènes A et B est responsable du phénotype AB.

**Figure 6 : Schéma de synthèse des antigènes ABH à partir du précurseur**

**de type 2.**

L'antigène H est obtenu grâce à la fucosylation du Gal terminal d'un précurseur, catalysée par une  $\alpha 1,2$  fucosyltransférase. La lettre R correspond au radical glycoprotéique ou glycolipidique.

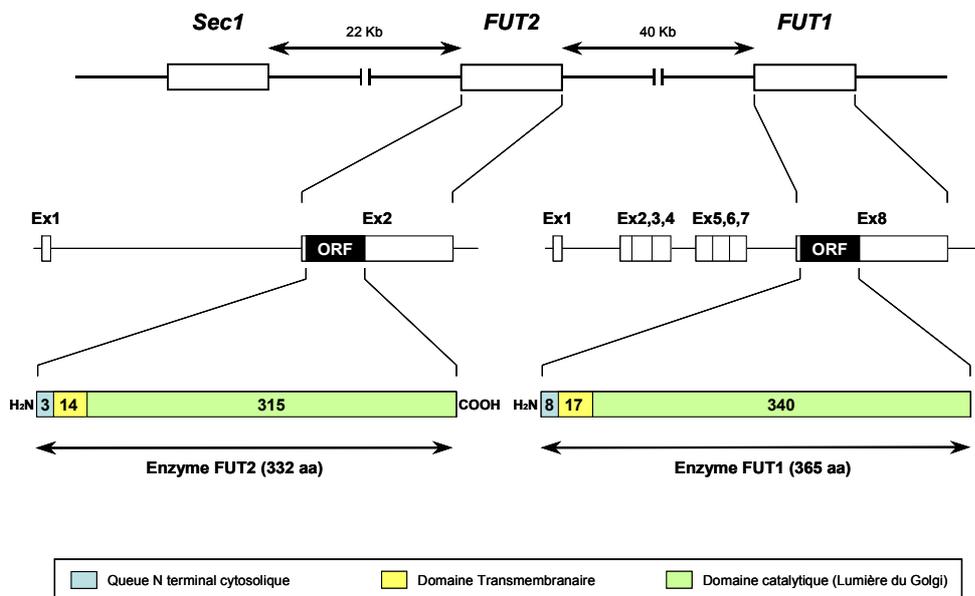
### 3. LES GLYCOSYLTRANSFÉRASES DU SYSTÈME ABH

#### 3.1. Les $\alpha 1,2$ fucosyltransférases

Plusieurs modèles génétiques ont été proposés pour expliquer les variations d'expression observées entre les phénotypes O et Se dans la population. C'est Oriol, qui en 1981 a proposé le modèle génétique actuel, considérant que les phénotypes O et Se sont la résultante de l'expression de deux gènes de structure indépendants (Oriol et coll. 1981), définis par la suite comme les gènes *H* (*FUT1*) et *Se* (*FUT2*). L'hypothèse de deux enzymes distinctes fut renforcée par l'observation de deux activités  $\alpha 1,2$ fucosyltransférases différentes (Kumazaki et coll. 1984; Le Pendu et coll. 1985). Suite à la purification de ces deux enzymes à partir de plasma humain, Sarnesto réussit à confirmer que l'enzyme H (*FUT1*) montre une plus grande affinité pour le précurseur de type 2, alors que l'enzyme Se (*FUT2*) préfère ceux de types 1 et 3 (Sarnesto et coll. 1990, 1992).

Le gène *FUT1* a été cloné et séquencé par Larsen en 1990 (Larsen, Ernst, et coll. 1990). L'utilisation de cette séquence a permis ensuite d'identifier le locus de ce gène dans la région q13.3 du chromosome 19 (Rouquier et coll. 1994). Le clonage et le séquençage de cette région ont ensuite permis de mettre en évidence la présence de deux autres gènes proches physiquement et génétiquement de *FUT1* : *Sec1* et *Sec2* (Rouquier et coll. 1995). *Sec1* s'est avéré être un pseudogène codant une protéine inactive de 246 acides aminés, tandis que le gène *Sec2*, ou *FUT2*, a été identifié comme

codant l'enzyme Se (R. J. Kelly et coll. 1995). *FUT1* se décompose en 8 exons, présente 3 sites possibles d'initiation de la transcription et plusieurs variants d'épissage alors que *FUT2* n'est constitué que de 2 exons, ne montre qu'un seul site d'initiation ATG et n'est pas épissé (Figure 7). Toutefois, pour ces deux gènes la région codant l' $\alpha$ 1,2 fucosyltransférase est comprise dans un exon unique : l'exon 8 pour *FUT1* et l'exon 2 pour *FUT2* (Koda, Soejima, et Kimura 1997; Koda, Soejima, Johnson, et coll. 1997).



**Figure 7 : Localisation des gènes des  $\alpha$ 1,2 fucosyltransférases sur le génome humain, structure des gènes *FUT1* et *FUT2* et représentation schématique des enzymes *FUT1* et *FUT2*.**

Le pseudogène *Sec1* et les gènes *FUT1* et *FUT2* sont localisés dans la région q13.3 du chromosome 9 humain. Très proche physiquement, ils sont regroupés sur une zone d'environ 65 kb. La traduction des gènes *FUT1* et *FUT2* produit deux  $\alpha$ 1,2 fucosyltransférases distinctes ayant chacune une masse d'environ 40 kDa. (D'après Koda et coll. 2001).

Chacun de ces deux gènes possède des allèles dominants, H et Se, codant pour l' $\alpha$ 1,2 fucosyltransférase et des allèles récessifs, h et se, non fonctionnels. Ainsi, un individu possédant un gène Se fonctionnel est défini comme « sécréteur » et présente les HBGA dans ses sécrétions (ex : salive) et au niveau de certains de ses tissus (ex : cellules épithéliales duodénales). Une personne héritant du double allèle récessif se/se est dite "non sécréteur" et ne présente que les sucres précurseurs à l'antigène H dans ses sécrétions et à la surface de certaines cellules.

Les individus homozygotes pour l'allèle nul h (h/h) sont rares et ne portent pas d'antigène H à la surface de leurs érythrocytes. Pour ces individus, deux cas de figure peuvent se présenter : s'ils possèdent en plus le statut de « non sécréteur », ils sont dits de phénotype « Bombay » et l'épitope H est totalement absent de leur organisme ; si par

contre ils possèdent le statut de « sécréteur », ils sont dits de phénotype « para-Bombay » et expriment l'antigène H uniquement au niveau de leurs épithéliums et dans leurs sécrétions. Toutefois, il existe une deuxième définition pour le phénotype « para-Bombay », caractérisée par une très faible expression d'antigène H à la surface des globules rouges associé à une enzyme FUT1 partiellement active et une absence d'expression dans les sécrétions (se/se). Le phénotype « Bombay » a été décrit pour la première fois par Bhende en 1952 dans la ville de Bombay (Mumbai, Inde) (Bhende et coll. 1952).

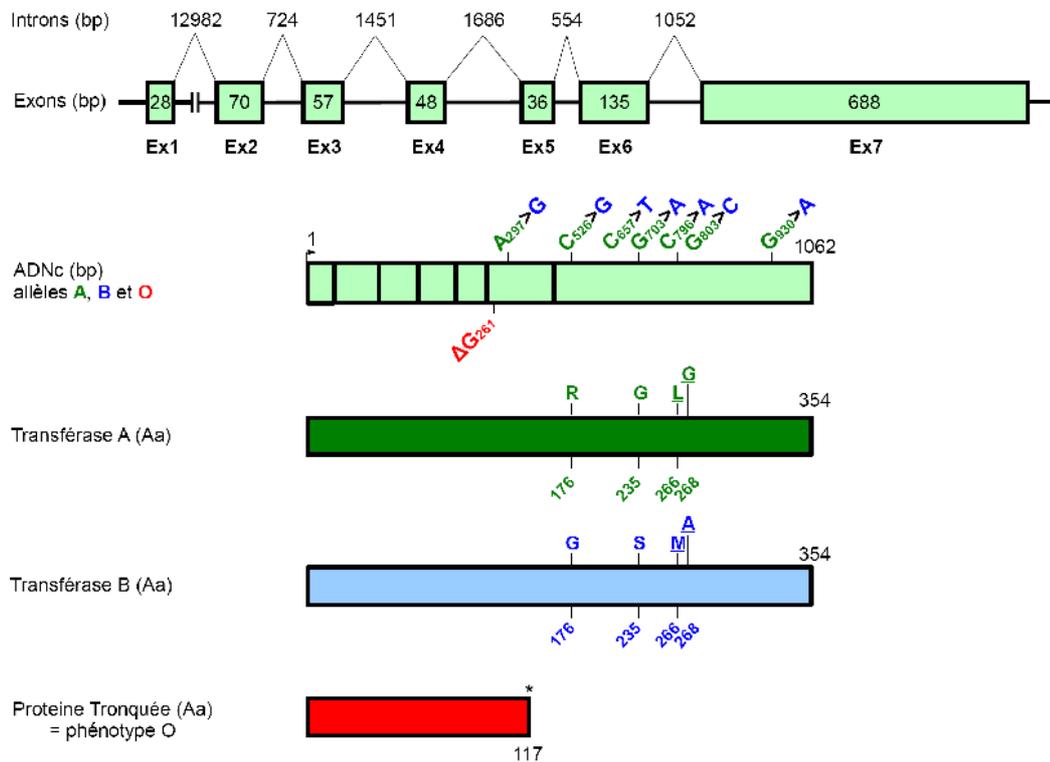
### **3.2. Les $\alpha$ 1,3 glycosyltransférases A et B**

Suite à l'identification des sucres composant les HBGAs A et B par Watkins et Morgan, la caractérisation des enzymes catalysant leur synthèse est passée par la purification des protéines présentes dans les tissus exprimant ces antigènes (Carne et coll. 1977; Nagai et coll. 1978). En 1990, c'est une séquence peptidique partielle de la glycosyltransférase A qui a été isolée à partir d'homogénats de poumon ou d'intestin (Clausen et coll. 1990; Navaratnam et coll. 1990). La même année, Yamamoto réussit le premier à cloner l'ADN complémentaire de l'enzyme A en utilisant des amorces dégénérées définies d'après cette séquence peptidique partielle (F. Yamamoto, Marken, et coll. 1990). Il isole et clona ensuite les allèles O et B, posant ainsi les bases génétiques du système ABH (F. Yamamoto, Clausen, et coll. 1990). Dans son analyse, Yamamoto observe que les séquences nucléotidiques des allèles A, B et O ne diffèrent que très peu. En effet, il n'identifie qu'une seule mutation entre l'allèle O et l'allèle A, alors que l'allèle B présente sept positions nucléotidiques différentes comparé à l'allèle A. La mutation identifiée pour l'allèle O est une délétion du nucléotide positionné en 261 de la séquence de l'allèle A, entraînant un décalage du cadre ouvert de lecture et provoquant ainsi l'apparition d'un codon stop précoce, responsable de la synthèse d'une protéine inactive de 117 acides aminés (Figure 8). Les modifications pour l'allèle B entraînent le changement de quatre acides aminés (positions 176, 235, 266, 268 de la séquence peptidique de l'allèle A) suffisant à changer la nature du substrat donneur de la glycosyltransférase. L'impact de ces mutations entre allèle A et B a ensuite été évalué par mutagenèse dirigée (F. Yamamoto et Hakomori 1990). Cette étude a permis de mettre en évidence que la nature des acides aminés aux positions 266 et 268 de la séquence peptidique de l'allèle A est critique quant à la spécificité de l' $\alpha$ 1,3 glycosyltransférase. En effet, la modification Leu/Met en position 266 et plus encore la modification Gly/Ala en position 268 permettent de faire basculer la spécificité de l'enzyme vers l'un ou l'autre des substrats, mais autorise également leur utilisation combinée dans certains cas. En effet, une autre étude a montré que lorsque la position 268 est occupée par une Ala, Ser, Gly ou Cys, elle confère une activité A et B

transférase. Quand celle-ci est occupée par une His ou une Phe elle engendre une activité A transférase faible, tandis que la présence d'une Asn ou d'une Thr donne une activité B faible (F. Yamamoto et coll. 1996).

**Figure 8 : Représentation schématique du locus ABO, de l'ADN complémentaire des allèles A, B et O, ainsi que de leur produit de traduction respectif.**

Le locus ABO à une taille d'environ 19,5 kb et est localisé sur le chromosome 9 en position q34.1-q34.2. La majeure partie de la séquence codante est représentée par les exons 6 et 7, exons portant les mutations



différenciant les allèles A, B et O. Parmi celles-ci, les positions 796 et 803 jouent un rôle fondamental pour la spécificité du substrat. La délétion ΔG en 261 de la séquence codante est responsable d'un décalage du cadre de lecture entraînant l'apparition d'un codon stop précoce et donc de la synthèse d'une protéine tronquée inactive.

## 4. GÈNES DU SYSTÈME ABH ET ÉVOLUTION

### 4.1. Polymorphisme des gènes FUT1 et FUT2

FUT1, FUT2 et Sec1 présentent entre 67% et 76% d'homologie lorsque l'on compare leurs séquences nucléotidiques (Rouquier et coll. 1995). Ces similarités et leur proximité sur le chromosome 19 suggèrent qu'ils sont le résultat d'au moins deux

duplications suivies d'une dérive génétique vers la pseudogénéisation ou l'évolution vers l'acquisition de nouvelles propriétés catalytiques. Malgré ces ressemblances, il existe des différences importantes concernant le polymorphisme des gènes *FUT1* et *FUT2*.

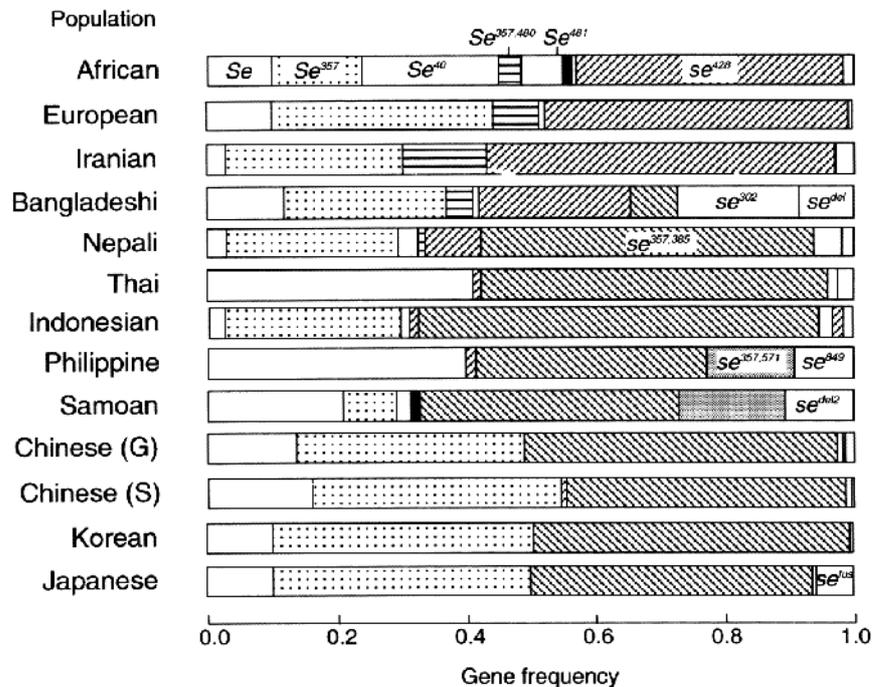
Tout d'abord, l'expression du gène *FUT1* au cours de l'hématopoïèse est restreinte à l'homme et à quelques espèces de grands singes. Au contraire, *FUT2* est exprimé au niveau de l'épithélium intestinal chez un large spectre d'espèces allant des amphibiens aux mammifères supérieurs. La mise en évidence d'une séquence *Alu* et d'une séquence LTR unique (LTR13) en 5' de l'exon 2 de *FUT1* pourrait jouer un rôle important dans la régulation de son expression, mais aussi être à l'origine de son caractère espèce-spécifique. En effet, le site d'initiation de la transcription de l'exon 2 est situé dans ce LTR. De plus, il a été montré que le promoteur rétroviral du LTR est suffisant pour agir seul en tant que promoteur du gène (Koda, Soejima, et Kimura 1997). Cette insertion serait relativement récente car peu d'espèces expriment l'antigène H au niveau érythrocytaire, de plus celles-ci sont très proches d'un point de vue phylogénétique.

Chez l'homme, les allèles mutés de *FUT1* sont très rares et chaque mutation décrite a été isolée chez un seul individu ou chez un petit groupe d'individus apparentés (dbRBC - Blood Group Antigen Gene Mutation Database). Dans la population, la fréquence des individus de phénotype Bombay ou para-Bombay est de l'ordre de 1 pour 13 000 en Inde et de 1 pour 312 000 dans la population allemande et européenne en général (F. F. Wagner et coll. 1997). Plus récemment, un grand nombre d'individu H déficient ont été identifiés sur l'île de la Réunion, avec une fréquence relativement importante de 1 pour 1 000. Dans cette population, deux phénotypes distincts ont été identifiés : le « Bombay » classique (substitution T725G) issu de certaines familles indiennes Tamoul immigrantes et un phénotype partiellement déficient appelé phénotype « Réunion » (substitution C349T) (Fernandez-Mateos et coll. 1998). Ces deux mutations sont celles retrouvées le plus fréquemment en France à l'origine d'une déficience érythrocytaire d'antigène H, comparées à celles des autres variants inactivés de *FUT1*.

Dans le cas de *FUT2*, la fréquence des mutations induisant le phénotype non sécréteur est assez élevée. Celle-ci est corrélée à des spécificités ethniques très hétérogènes (Figure 9). Ainsi la substitution G428A induisant le remplacement du tryptophane par un codon Stop précoce en position 143 de la protéine *FUT2* représente la mutation inactivante la plus fréquemment rencontrée dans les populations caucasiennes (européenne et iranienne) et africaines. Pour ces peuplades, le variant *FUT2* se<sup>428</sup> est retrouvé à l'état homozygote chez la quasi totalité des individus de phénotype non sécréteur, soit environ 20% de leur population.

En Asie, c'est l'association des substitutions C357T et A385T qui est à l'origine de l'allèle nul prédominant chez les individus de phénotype non sécréteur. Ces substitutions

induisent respectivement une mutation synonyme et le remplacement critique d'une isoleucine par une phénylalanine en position 129 de la séquence peptidique de l'enzyme FUT2. Le variant *FUT2*  $se^{357,385}$  est retrouvé chez la majorité des asiatiques de phénotype non sécréteur, le plus souvent à l'état homozygote, mais parfois associé à d'autres variants non rares de *FUT2* comme  $se^{357,571}$ ,  $se^{849}$  et  $se^{del2}$  dans certaines îles du Pacifique.



**Figure 9 : Fréquences des différents allèles *FUT2* (*Se* et *se*) dans plusieurs populations.**

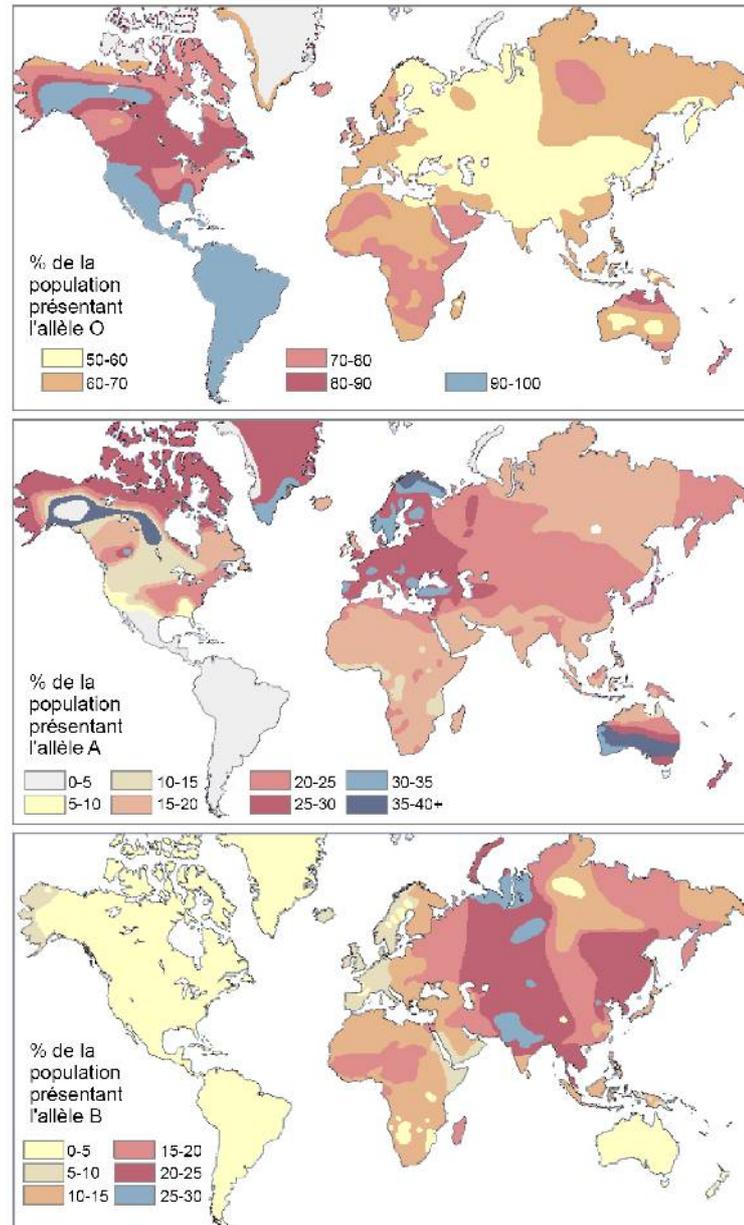
(D'après Y. Koda et coll. 2001).

## 4.2. Polymorphisme du gène ABO

Tout comme les gènes codant les  $\alpha$ 1,2 fucosyltransférases, le gène *ABO* présente lui aussi un polymorphisme important. Depuis le séquençage du locus *ABO*, de nombreuses mutations additionnelles ont été décrites et ont permis d'expliquer les variations d'expression des HBGA A, B et O (dbRBC - Blood Group Antigen Gene Mutation Database). La plupart de ces mutations sont présentes et conservées dans tous les groupes ethniques, ce qui sous-entend qu'elles ont toutes eu lieu avant les grands flux migratoires de la population humaine. Toutefois, il existe des profils génétiques et phénotypiques ABH propres à chaque groupe ethnique (Figure 10).

**Figure 10 : Répartition des allèles A, B et O sur une échelle mondiale.**

(D'après Mourant et coll. 1976). Les allèles A, B et O sont présents dans l'ensemble des groupes ethniques avec toutefois des variations de fréquence propres à chacun d'eux. Prenons l'exemple des populations d'Amérique Centrale et d'Amérique du Sud où le phénotype O est



hyperdominant alors que le phénotype B y est quasiment absent, tout comme sur le continent australien. En Europe, les phénotypes A, B et O sont respectivement retrouvés à hauteur de 47%, 13% et 44% (M. E. Reid et coll. 2003).

Parmi les individus de phénotype A, deux sous-groupes majoritaires se distinguent : le groupe dit A<sub>1</sub> (≈80% de la population caucasienne de phénotype A) et le groupe dit A<sub>2</sub> (≈20% de la population caucasienne de phénotype A). La distinction entre ces individus se fait d'une part au niveau de la densité antigénique qu'ils présentent à la

surface de leur globules rouges, ceci en raison d'une différence d'efficacité entre les enzymes  $A_1$  ( $\approx 1\ 000\ 000$  sites) et  $A_2$  ( $\approx 200\ 000$  à  $400\ 000$  sites). Ils se démarquent d'autre part par le type de structures qu'ils sont capable de synthétiser, ainsi les individus  $A_1$  expriment les antigènes A types 1, 2, 3 et 4, alors que les individus  $A_2$  ne présentent que les antigènes A types 1 et 2 (Clausen et coll. 1984, 1985).

De nombreux variants ont été identifiés comme étant à l'origine du phénotype  $A_1$ . Le premier décrit est l'allèle *A101* de Yamamoto (F. Yamamoto, Marken, et coll. 1990). Le second fut l'allèle *A102* (Ogasawara et coll. 1996) qui présente la mutation non-synonyme C467T (P156L) lorsqu'il est comparé à l'allèle *A101*. Toutefois, cette mutation ne semble pas affecter l'activité enzymatique de la transférase A. Les variants *A103* et *A104* présentent quant à eux une mutation synonyme lorsqu'ils sont comparés respectivement aux allèles *A102* et *A101* (Ogasawara et coll. 1996). La mutation observée pour l'allèle *A104* est une position généralement rencontrée sur l'allèle B (*B101*), signant ainsi un mécanisme de diversification des gènes *ABO* par recombinaison homologue. Ce mécanisme semble relativement fréquent car il a été également mis en cause dans la formation des allèles *A105* et *A106*, allèles A observés beaucoup plus rarement dans la population (Ogasawara et coll. 2001).

L'allèle *A201* responsable du phénotype  $A_2$  a été mis en évidence pour la première fois par Yamamoto (F. Yamamoto et coll. 1992). Il diffère de l'allèle *A101* par une mutation non-synonyme C467T (L156P) ainsi que par la délétion d'une cytosine ( $\Delta C1061$ ) présente dans un motif répété de trois cytosines. Cette délétion entraîne un décalage du cadre ouvert de lecture aboutissant à l'élongation de 21 acides aminés du côté carboxy-terminal de l'enzyme. Cette élongation est à l'origine du phénotype  $A_2$  car un autre variant (*A206*) impliqué dans ce phénotype diffère du variant *A101* uniquement au niveau de cette délétion de la cytosine 1061 (S P Yip 2000). Ogasawara a mis en évidence les 5 autres variants responsables du phénotype  $A_2$  : les allèles *A202*, *A203*, *A204* (Ogasawara et coll. 1996) et *A205* (Ogasawara et coll. 1998). Les allèles *A202* et *A203* diffèrent de l'allèle *A101* respectivement sur les positions C1054T (R352T) et C1054G (R352G). L'allèle *205* varie lui sur 6 positions induisant 3 mutations non-synonymes : G829A (V277M), A297G, C526G (R176G), C657T, G703A (G235S), C771T.

Concernant le phénotype B, il n'existe qu'un seul phénotype de haute fréquence. Malgré cela, de nombreux variants alléliques lui sont associés. Le premier décrit et le plus fréquemment rencontré est l'allèle *B101* de Yamamoto (F. Yamamoto, Clausen, et coll. 1990). D'autres variants beaucoup plus rares ont également été isolés. Les variants *B102* et *B103* qui diffèrent de l'allèle *B101* respectivement sur les positions A930G et T657C, mutations qui ne modifient pas la séquence peptidique de la transférase B (Ogasawara et coll. 1996). L'allèle *B107* présente lui une mutation non synonyme (G526C) qui induit le remplacement d'une glycine par une arginine en position 176 de

l'enzyme (Ogasawara et coll. 1998). Enfin, l'allèle *B108* présente la mutation synonyme G297A. Cette mutation est partagée avec le variant *A101* et l'étude de l'intron 6 a montré clairement un évènement de recombinaison homologue entre l'allèle *A101* et *B101* (Ogasawara et coll. 2001).

D'autres allèles responsables de phénotypes A faibles ou B faibles ont été identifiés. Ceux-ci sont rares et présentent une activité enzymatique inférieure à celle des variants induisant les phénotypes  $A_2$  et B. Il s'agit des phénotypes  $A_3$ ,  $A_x$ ,  $A_{end}$ ,  $A_m$ ,  $A_{el}$ ,  $B_3$ ,  $B_x$ ,  $B_m$ ,  $B_{el}$ ,  $B_{mos}$  (Yamamoto 2000).

Plusieurs mutations responsables du phénotype O ont été identifiées. Les allèles *O01* (56% des individus caucasiens de phénotype O) et *O02* (40% des individus caucasiens de phénotype O) sont ceux les plus fréquemment rencontrés dans les populations. Le premier est celui décrit par Yamamoto en 1990 et différant de l'allèle *A101* par la simple délétion d'une guanine ( $\Delta G261$ ) induisant un décalage du cadre ouvert de lecture et l'apparition d'un codon stop précoce à l'origine d'une protéine tronquée inactive (117aa). Le second, également mis en évidence par Yamamoto (F. Yamamoto, Marken, et coll. 1990), possède lui aussi cette délétion mais diffère du variant *O01* par 9 mutations présentes sur les exons 6 et 7 : G106T (V36F), G188A (R63H), C189T, C220T (P74S), A297G, T646A (R176G), G681A, C771T et G829A (V277M). Les autres allèles O sont beaucoup plus rares et résultent d'autres mutations non synonymes de l'allèle *A101* ou de recombinaisons homologues avec d'autres allèles du gène *ABO*.

Parmi les phénotypes AB, des allèles particuliers *cis-AB* ont été définis comme générant une  $\alpha 1,3$  glycosyltransférase capable de catalyser aussi bien le transfert d'un galactose ou d'une N-actylgalactosamine sur le résidu galactose de l'antigène tissulaire de groupe sanguin H (F. Yamamoto et coll. 1993). Depuis de nombreux allèles associés à ce phénotype ont été isolés (Mifsud et coll. 2000; Roubinet et coll. 2002; Ding-Ping Chen et coll. 2004)

### 4.3. HBGA ABH chez les autres espèces animales

L'expression des HBGA ABH à la surface de certaines cellules et/ou dans les sécrétions n'est pas une caractéristique retrouvée uniquement chez l'homme (Figure 11). En effet, ces glycanes ont été mis en évidence sur certains tissus de nombreux primates, félins, ovins, caprins, marsupiaux, reptiles, rongeurs et oiseaux (Oriol et coll.

		Epithélium digestif	Épiderme	Récepteurs olfactifs	Endothélium vasculaire	Erythrocytes
Anoures		+	-	-	-	-
Reptiles		+	-	-	-	-
Rat		+	+	+	-	-
Lapin		+	-	-	-	-
Marmousets		+	+	+	-	-
Babouins		+	+	+	+	-
Grands singes anthropoïdes		+	+	+	+	+

1999). Néanmoins, l'homme et certains grands singes anthropoïdes (chimpanzés, gorilles, orangs-outans et gibbons) ont l'exclusivité de présenter ces carbohydrates à la surface de leurs érythrocytes (Kermarrec et coll. 1999).

**Figure 11 : Distribution tissulaire des antigènes de groupes sanguins chez différentes espèces de vertébrés plus ou moins éloignées d'un point de vue phylogénétique.**

(D'après Oriol et coll. 1986, 1992).

La présence d'orthologues du gène *ABO* chez d'autres espèces que l'homme est mise en évidence pour la première fois en 1992 par Kominato. Il a utilisé pour ceci une technique d'hybridation entre l'ADN génomique de différents organismes et la séquence d'ADN complémentaire codant la transférase A humaine (Kominato et coll. 1992). Il observa alors une très bonne fixation de cette sonde sur le génome des singes de

l’Ancien Monde (= Catarhiniens), du marmouset, du chien, du chat, du lapin, du bœuf, du mouton, du rat, de la souris et du hamster. Il observa également une plus faible hybridation sur le génome du poulet et une absence d’hybridation sur le génome des espèces situées plus bas sur l’arbre phylogénétique de l’Evolution (bactérie, levure, nématode, palourde, homard, mouche, oursin, grenouille).

### **a - Primates**

Chez les primates non-humains, toutes les espèces présentent exclusivement le phénotype sécréteur (Koda et coll. 2000). Les différents allèles *Fut2* identifiés chez ces espèces codent tous une  $\alpha$ 1,2 fucosyltransférase active, très proche de son homologue humaine. Chez certaines espèces, le gène *Fut2* peut être associé à un variant du gène *Sec1* codant également une  $\alpha$ 1,2 fucosyltransférase active (singes vert et singes du Nouveau Monde (= Platyrrhiniens)). Tous les primates expriment les antigènes ABH dans leur salive, à l’exception d’un Orang-outan (Blancher et coll. 1997). Tout comme l’homme, les Catarhiniens expriment les antigènes ABH à la surface de leur endothélium vasculaire, site d’expression non partagé avec les Platyrrhiniens (Mollicone et coll. 1986; Oriol et coll. 1999). Contrairement au caractère sécréteur, les phénotypes A et B ne présentent pas une répartition homogène entre les différentes espèces de singes. En effet hormis le macaque crabier, aucune lignée simienne non-humaine ne possède la combinaison simultanée des quatre phénotypes A, B, AB et O (Tableau 3). Ceci peut être illustré par l’exemple du gorille qui n’exprimerait que l’antigène B, alors que son proche cousin le chimpanzé n’exprime lui que les antigènes A et O.

<b>Espèces</b>	<b>Phénotypes ABO</b>
Chimpanzé <i>Pan troglodytes</i>	A (113), O (17)
Gorille <i>Gorilla gorilla</i>	B (2)
Orang Outan <i>Pongo pygmaeus</i>	A (22), B (1), AB (3)
Gibbon à mains blanches <i>Hylobates lar</i>	A (2), B (5), AB (4)
Babouin olive <i>Papio anubis, P. cyanocephalus</i>	A (36), B (17), AB (34)
Macaque rhésus <i>Macaca mulatta</i>	B (10)
Macaque à queue de cochon <i>Macaca nemestrina</i>	B (5)
Macaque crabier <i>Macaca irus</i>	A (8), B (1), AB (3), O (1)
Macaque à crête <i>Macaca nigra</i>	A (8), B (2)
Singe écureuil <i>Saimiri sciureus</i>	A (3), O (1)
Capucin à front blanc <i>Cebus albifrons</i>	B (3), O (1)

**Tableau 3 : Expression des HBGA ABH chez les primates non-humains.**

Les chiffres représentés entre parenthèses correspondent aux nombres de cas observés par espèce. (D'après Saitou et coll. 1997).

**b - Rat**

L'expression des gènes *Fut1*, *Fut2*, *Sec1* et *ABO* a également été bien caractérisée chez le rat. C'est Piau qui est le premier à isoler deux gènes codant une  $\alpha$ 1,2 fucosyltransférase : *FTA* (=FUT1) et *FTB* (=FUT2) (Piau et coll. 1994). Par la suite, Bureau et collaborateurs mettent en évidence un troisième gène codant potentiellement cette enzyme : le gène *FTC* (=Sec1). Ils montrent que chez cette espèce, ces gènes présentent la même organisation chromosomique que chez l'homme mais que ceux-ci sont localisés en position q22-q31 du chromosome 1 (Bureau et coll. 2001). L'expression *in vitro* de leur ADN complémentaire respectif a permis de mettre en évidence que seuls les gènes *FTA* et *FTB* étaient capables d'une activité fucosyltransférase. En effet, *FTC* présente une délétion sur son premier codon ATG, décalant de ce fait l'initiation de la traduction et générant une enzyme inactive ne contenant ni queue intracytoplasmique, ni domaine transmembranaire.

Chez le rat, le gène *ABO* est localisé sur le chromosome 3 en position q11-q12. L'expression du gène *ABO* a été documentée par Cailleau-Thomas chez le rat BDIX. Elle montre que les antigènes A et B peuvent être exprimés chez un même animal dans de nombreux organes et parfois co-exprimés dans un même organe mais avec en général une distribution cellulaire différente (Cailleau-Thomas et coll. 2002). Plus tard, il a été mis en évidence que le rat possède en réalité plusieurs copies du gène *ABO*, en tandem sur le chromosome 3, dont une code une transférase B (Iwamoto et coll. 2002; Turcot et coll. 2003)

**c - Souris**

Chez la souris, c'est Domino qui est le premier à cloner un gène codant une  $\alpha$ 1,2 fucosyltransférase (Domino et coll. 1997). Il détermine que ce gène présente 55% d'homologie avec la séquence humaine du gène *FUT2* et 75% d'homologie avec la séquence humaine du gène *FUT1*, gène auquel il l'associe. Il localise le gène *Fut1* murin sur le chromosome 7 en position 4D et montre également que son ARN messager est transcrit dans de nombreux tissus (thymus, poumon, estomac, pancréas, intestin grêle, colon, utérus et épидидyme). En 2001, Lin complète la famille des gènes codant une  $\alpha$ 1,2 fucosyltransférase murine en isolant deux nouveaux gènes : *MFut-II* (=Fut2) et *MFut-III* (= Sec1) (B Lin et coll. 2001). Ceux-ci présentent respectivement 79% et 74,9% d'homologie avec leurs orthologues humains. De plus, il met en évidence que le gène *MFut-III* ne code pas une enzyme active et que l'expression du gène *MFut-II* au niveau

de l'intestin grêle est dépendante de la colonisation de l'organe par la flore bactérienne commensale du tube digestif.

Cette même année, Yamamoto identifia l'équivalent murin du gène *ABO* (M. Yamamoto et coll. 2001). Il met alors en évidence une enzyme capable des deux activités transférases A et B, et nomme cet allèle Cis (AB). Le gène *ABO* de la souris est localisé sur le chromosome 2 en position A3.

### **d - Porc**

L'expression des HBGA porcins a été étudiée avec attention en raison du fort potentiel de cette espèce dans l'optique des xénotransplantations. Le porc peut présenter le phénotype A ou O.

Cohney est le premier à cloner un gène porcine d' $\alpha$ 1,2 fucosyltransférase, orthologue du gène *FUT1* humain et présentant 81% d'homologie entre leurs séquences peptidiques (Cohney et coll. 1996). L'année suivante, Meijerink localise ce gène sur le chromosome 6 du porc en position q11. Il identifie également dans la même région 2 autres gènes : le gène *Sec1* et le gène *Fut2*, présentant 85% d'homologie avec la séquence nucléique *FUT2* (Meijerink et coll. 1997).

La description du gène *ABO* chez le porc a été faite par Yamamoto et Meijerink en 2001 (F. Yamamoto et coll. 2001; Meijerink et coll. 2001). Ils montrent que ce gène présente 77% d'homologie avec la séquence nucléique du gène *ABO* humain et qu'il est localisé sur le chromosome 1 entre les positions q2.10 et q2.11. Yamamoto met de plus en évidence que le phénotype O est toujours associé à une délétion du gène *ABO* chez cette espèce.

### **e - Lapin**

Tous les lapins testés jusqu'à présent sont de phénotype sécréteur. Les premiers gènes codant une  $\alpha$ 1,2 fucosyltransférase mis en évidence sont les gènes *RFT-I* (= *Fut1*) et *RFT-II* (= *Sec1*) (Hitoshi et coll. 1995). Hitoshi met en évidence que *RFT-I* présente 80% d'homologie avec la séquence humaine de *FUT1*. Pour les 2 produits de ces gènes, il met en évidence une activité fucosyltransférase. De plus, il détecte de l'ARN messager de *RFT-I* uniquement au niveau du cortex cérébral, alors que ceux de *RFT-II* sont retrouvés dans les glandes salivaires et mammaires. L'activité enzymatique et l'expression tissulaire de *RFT-II* le conduit à proposer que ce gène soit un homologue du gène humain *FUT2*. Toutefois en 1996, la même équipe met en évidence un troisième gène, *RFT-III*, codant une  $\alpha$ 1,2 fucosyltransférase active (Hitoshi et coll. 1996). Sa séquence peptidique présente plus d'homologie avec son homologue *FUT2* humain (83%) que *RFT-III* (67%). Ce dernier correspond en réalité à un homologue actif du pseudogène humain *Sec1*.

L'expression tissulaire de *RFT-III* semble être essentiellement localisée dans le colon et dans une moindre mesure dans l'estomac et l'intestin grêle. Le lapin possède une expression tissulaire restreinte et spécifique pour chacun de ces trois gènes.

Concernant le(s) gène(s) *ABO*, celui-ci n'a pas encore été cloné et séquencé pour le lapin. Cependant, certains animaux sont capables d'exprimer l'antigène A à la surface de leurs cellules épithéliales intestinales alors que d'autres en sont dépourvus, mais tous expriment l'antigène B (Oriol et coll. 1977 et données personnelles).

### ***f - Vache***

Les gènes codant les orthologues de *FUT1*, *FUT2* et *Sec1* ont été mis en évidence par Barreaud et collaborateurs (Barreaud et coll. 2000), avec respectivement 81%, 84% et 83% d'homologie dans leurs séquences nucléiques. Ils déterminent que ces gènes sont localisés sur le chromosome 18 du bœuf en position q24. Ils montrent également que chacun de ces trois gènes code une enzyme possédant une activité  $\alpha$ 1,2 fucosyltransférase. Toutefois, ils ne détectent que des ARN messagers issus du gène *Fut2* par la technique de northern blot sur un large panel de tissus (rein, intestin, poumon, cerveau, cœur et rate). Néanmoins grâce à la RT-PCR, une technique beaucoup plus sensible, ils mettent en évidence dans certains tissus de l'ARN messager issu de *Fut1* (rein, intestin, cerveau) et *Sec1* (rein). Ils suggèrent également l'existence d'un quatrième gène codant une  $\alpha$ 1,2 fucosyltransférase, au vu de leurs résultats. L'année suivante, Saunier caractérise l'organisation génomique de ces gènes (Saunier et coll. 2001), organisation similaire à celle retrouvée chez l'homme.

Le gène *ABO* n'a pas encore été cloné chez les bovins. Toutefois, une séquence présentant une forte homologie avec le gène *ABO* humain est ressortie lors du séquençage du génome du bœuf (NM\_001077926).

## **4.4. Histoires évolutives des gènes liés aux HBGA ABH**

Comme nous l'avons vu précédemment, la synthèse des HBGA ABH est sous la dépendance de deux familles de gènes, aux histoires évolutives différentes.

L'homme compte neuf gènes de fucosyltransférases (*FUT1* à *FUT9*) dont seulement trois sont associés à une activité  $\alpha$ 1,2 fucosyltransférase effective (*FUT1*, *FUT2*) ou ancienne (*Sec1*). Chez cette famille de gènes, la tendance évolutive a été d'aller vers une séquence codante monoexonique, mais aussi de catalyser l'ajout d'un fucose toujours plus en périphérie des glycanes O et N-liés, ceci jusqu'à aboutir à la présentation du fucose en position terminale (Javaud et coll. 2003).

Les fucosyltransférases semblent avoir évolué par duplications de gènes ou rétrotranspositions. Ces mécanismes offrent la perspective d'acquisition de nouvelles

fonctions et de profils d'expression différents. Dans cette optique, trois stratégies évolutives ont été décrites (Lynch et coll. 2000) :

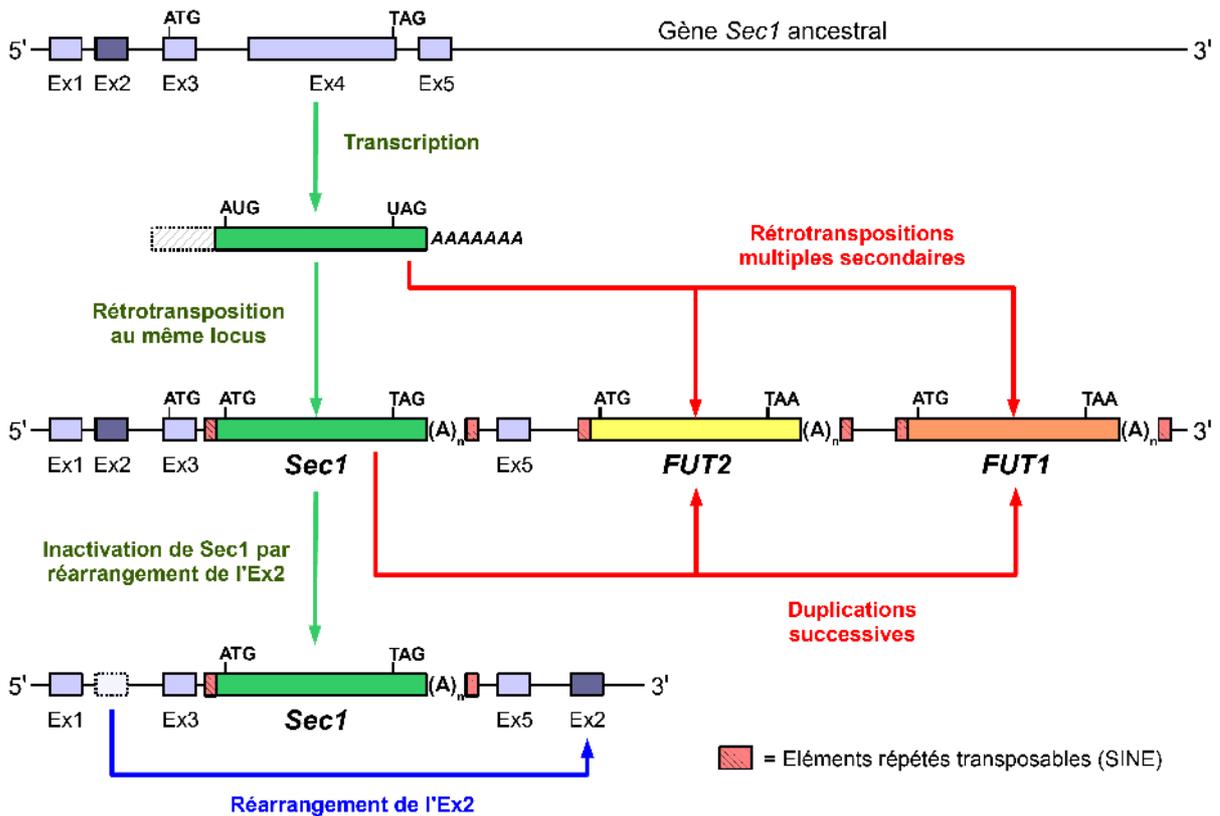
- La néo-fonctionnalisation, où la nouvelle copie du gène peut acquérir de nouvelles fonctions grâce à l'apport de nouvelles mutations. Ce gain s'il est bénéfique à l'organisme induira le maintien de ce nouveau gène selon les lois de la sélection naturelle, alors que l'autre copie maintiendra sa fonction originelle.
- La non-fonctionnalisation (= pseudogénéisation) où une des deux copies peut être inactivée suite à la multiplication de mutations dégénératives (délétion ou codon stop précoce) dans sa séquence codante ou dans une séquence régulant son expression.
- La sub-fonctionnalisation où les deux copies sont maintenues mais deviennent moins actives au cours de leur évolution, ceci du fait de l'accumulation de nouvelles mutations. Au final, les deux copies sont individuellement moins actives que la copie originelle du gène.

A ces trois processus peut venir se greffer le mécanisme de conversion génique, comme cela a été suggéré par certains auteurs (Ogasawara et coll. 2001; Soejima et coll. 2008). Celui-ci est caractérisé par une recombinaison non-homologue entre un gène qui est dit donneur et un gène qui est dit accepteur. Ce mécanisme a bien été caractérisé pour d'autres gènes, comme ceux codant les antigènes du CMH et dont il contribue à augmenter la variabilité en même temps qu'il les homogénéise en les faisant évoluer de manière convergente (Martinson et coll. 1999). Ce mécanisme pourrait également avoir un autre rôle en temps que processus de réparation de l'ADN, le gène donneur représentant une copie de sauvegarde en cas de mutations délétères du gène accepteur.

Concernant l'histoire évolutive des gènes d' $\alpha$ 1,2 fucosyltransférases, celle-ci aurait débuté par une première rétrotransposition d'une  $\alpha$ 1,2 fucosyltransférase ancestrale sur son même locus, formant alors *Sec1* (Figure 12). La genèse des gènes *FUT1* et *FUT2* pourrait quant à elle avoir deux origines. D'une part, elle pourrait être le fruit de nouvelles rétrotranspositions du gène ancestral de *Sec1*. Mais d'autre part, ces gènes pourraient être les produits de deux duplications successives du gène *Sec1*. Ces deux hypothèses sont valides et étayées par la présence d'éléments répétés transposables (SINE) flanquant chacun des gènes *Sec1*, *FUT1* et *FUT2*. La comparaison des séquences nucléiques de ces trois gènes semble montrer que l'apparition de *FUT1* est antérieure à celle de *FUT2*, ceci en raison de la plus forte homologie de séquence entre *FUT2* et *Sec1* (Saunier et coll. 2001). Toutefois, cette hypothèse peut être nuancée, voir remise en cause, par la mise en évidence récente de conversion génique entre *FUT2* et *Sec1*

(homogénéisation des gènes) alors que *FUT1* semble en être dispensé (Soejima et coll. 2008).

La pseudogénéisation de *Sec1* chez l'homme et certains grands singes serait intervenue plus tard au cours de l'évolution, bien après l'apparition de *FUT1* et *FUT2*. La comparaison des séquences humaine et bovine de *Sec1* (*Sec1* actif chez les bovins) a permis de mettre en évidence chez l'homme la transposition d'un exon non-codant (exon 2) en 3' de sa séquence codante (Saunier et coll. 2001). Cette modification semble être à l'origine de l'initiation de la pseudogénéisation de *Sec1* et souligne l'importance de la région non-codante du gène (5'UTR) pour la régulation de son expression. *Sec1* a ensuite accumulé de nombreuses autres mutations au cours du temps (délétions et apparition codons stop prématurés) conduisant irrémédiablement à l'inactivation complète du gène chez l'homme et d'autres espèces de grands singes (Apoil et coll. 2000).



**Figure 12 : Evolution hypothétique des gènes codant les  $\alpha$ 1,2 fucosyltransférases.**

(D'après Saunier et coll. 2001). Ne sont représentés ici que les événements majeurs de la genèse des gènes *FUT1*, *FUT2* et *Sec1*, ne sont pas représentés les mutations ponctuelles ayant eu lieu sur chacun d'eux et influant sur leur fonctionnalité.

A propos du gène *ABO* codant les transférases A et B humaines, comprendre son évolution c'est tout d'abord s'intéresser à celui de nos plus proches voisins sur l'arbre phylogénétique de l'évolution : les Primates.

Kominato est le premier à montrer que des séquences partielles du gène *ABO* de primates présentent les mêmes acides aminés que celles de l'homme aux positions 266 et 268 (Kominato et coll. 1992). Ces acides aminés ont été décrit précédemment comme jouant un rôle majeur dans l'orientation de la spécificité de l' $\alpha$ 1,3 glycosyltransférase pour le Gal ou le GalNAc. Cette observation lui permet de suggérer que ces mutations sont apparues bien avant que la lignée humaine ne se sépare de celles des autres singes. Cette hypothèse fut corroborée l'année suivante par Martinko et collaborateurs (Martinko et coll. 1993). Malgré ses allures solides, celle-ci fut pourtant invalidée par Saitou et Yamamoto lorsqu'ils construisirent un arbre phylogénétique avec les différents gènes *ABO* identifiés chez les primates. En effet, ils mirent alors en évidence qu'au cours de l'évolution, l'allèle B est apparu trois fois de manière indépendante dans des lignées simiennes distinctes (Saitou et coll. 1997). Ils conclurent alors que le gène *ABO* évoluait d'une manière convergente plutôt que sur le principe d'un héritage d'allèles ancestraux entre les espèces. Ironiquement, cette hypothèse fut validée la même année par l'équipe qui corrobora celle de Kominato en 1993 (O'hUigin et coll. 1997).

Concernant l'apparition des variants O du gène *ABO*, Kermarrec montra très clairement que chaque espèce de singe possédait sa propre version de l'allèle O. Elle expliqua ce phénomène comme le résultat de mutations indépendantes intervenues après les différentes spéciations, afin de converger vers une forme éteinte du gène *ABO* (Kermarrec et coll. 1999). Cette hypothèse est d'autant plus forte qu'elle va dans le même sens que celle émise par Saitou et Yamamoto.

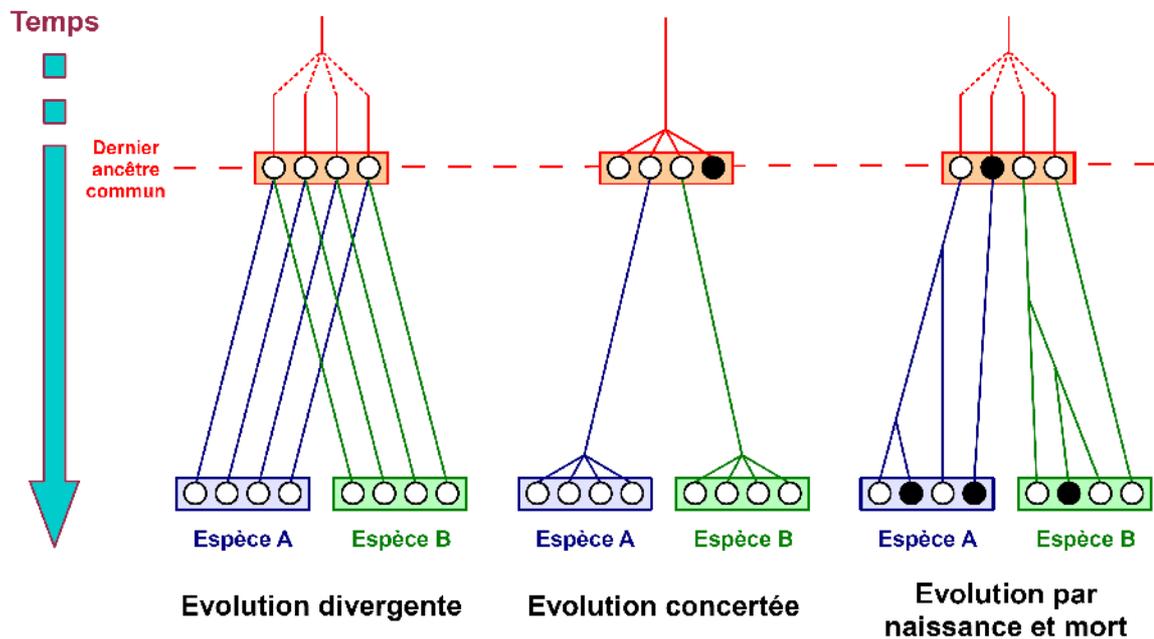
La comparaison de la séquence humaine à celles d'espèces phylogénétiquement plus éloignées a montré de fortes similitudes avec d'autres gènes n'étant pas rattachés directement au gène *ABO*. En effet, l'homme ne possède qu'un seul gène capable de générer une  $\alpha$ 1,3 Gal/GalNAc transférase active. Or chez d'autres espèces, d'autres gènes apparentés mais distincts du gène *ABO* ont été mis en évidence, signant ainsi leur appartenance à une même famille de gène, connue aujourd'hui comme la famille GT6 du système de classification CAZy. Ce regroupement a permis de mieux envisager les origines possibles du gène *ABO*. Parmi ces gènes, nous pouvons citer celui codant l' $\alpha$ 1,3 galactosyltransférase (*Hgta1*), gène actif chez tous les mammifères hormis l'homme et les catarhiniens (Oriol et coll. 1999). Cette enzyme permet la synthèse d'un glycanne connu sous le nom de xénoantigène (Gal $\alpha$ 1,3Gal $\beta$ 1,4GlcNAc). Très proche de l'antigène B, sa synthèse diffère de celle de ce dernier du fait de la spécificité de l' $\alpha$ 1,3 galactosyltransférase pour un sucre accepteur présentant un galactose terminal libre et donc ne portant pas de fucose en  $\alpha$ 1,2 (antigène H). Cet antigène pose un problème

majeur dans le domaine des xénotransplantations du fait de la présence de forts titres en anticorps anti-Gal $\alpha$ 1,3Gal dans le plasma humain. On retrouve toutefois trace chez l'homme d'un pseudogène homologue de celui de l' $\alpha$ 1,3 galactosyltransférase et d'une copie correspondant à un rétrotransposon, signe d'une activité ancestrale perdue au cours de l'évolution (Joziase et coll. 1989; Larsen, Rivera-Marrero, et coll. 1990; F. Yamamoto et coll. 1991). D'autres gènes associés à la famille GT6 ont été recherchés dans le génome humain, notamment ceux codant l'iGb3 synthase et la Forssmann synthase. Le gène de l'iGb3 synthase semble être un pseudogène dans la mesure où il ne semble pas être correctement transcrit (Speak et coll. 2007; Le Pendu et coll. non publié). Cependant, certains auteurs considèrent qu'il serait fonctionnel au moins dans le thymus (Y Li et coll. 2008). Le gène de la Forssmann synthase est inactivé chez l'homme (H. Xu et coll. 1999). En position q34 du chromosome 9 humain, on retrouve également un autre gène de la famille GT6. L'activité enzymatique de l'éventuelle protéine correspondante n'est pas connue, mais un codon non-sens semble être présent chez tous les humains testés. Il s'agit probablement d'un autre pseudogène, pour le moment dénommé GT6m7 pour membre 7 de la famille GT6 (Turcot-Dubois et coll. 2007).

Une étude récente a permis de mieux caractériser le schéma évolutif des gènes *ABO* et plus largement des glycosyltransférases de la famille GT6 (Turcot-Dubois et coll. 2007). Elle étudia les gènes de la famille chez les différents vertébrés dont le génome entier avait été séquencé. Elle montra notamment que l'homme, la souris, le porc et le bœuf ne possèdent qu'une seule copie du gène *ABO*, et qu'à l'inverse le génome du rat, du chien et du xénope en disposent de plusieurs. Parmi les espèces qu'elle étudia, les deux extrêmes concernant le nombre de copies du gène *ABO* sont le xénope avec huit gènes *ABO-like* et le poulet, tout comme trois espèces de poisson, qui ne possèdent pas de gène *ABO* dans leur génome. Ces observations concernant l'expansion ou la régression du nombre de gènes de la famille GT6 chez les différentes espèces sont caractéristique de ce qui est attendu dans le cas d'un modèle d'évolution dit de « naissance et de mort » (birth-and-death) (Nei et coll. 1997, 2005) (Figure 13). L'évolution des  $\alpha$ 1,2 fucosyltransférases pourrait également être rattachée à ce type de schéma évolutif, quoique la conversion génique existant entre les gènes *FUT2* et *Sec1* plaide plutôt pour ces gènes en faveur d'un modèle d'évolution concertée.

Le modèle « birth and death » s'appuie sur la création de nouveaux gènes par duplication. Ces gènes seront soit maintenus pendant de longues périodes dans le génome, soit délétés ou inactivés par des mutations (pseudogénéisation). Ce schéma évolutif est particulièrement adapté aux familles multigéniques impliquées dans des interactions entre la cellule et son environnement. Ce type d'évolution permet aux organismes d'être plus adaptables et de mieux répondre à de nouvelles pressions de

sélection, comme la modification de leur milieu (cas des récepteurs olfactifs et gustatifs) ou l'apparition de nouveaux pathogènes (cas du système HLA).



**Figure 13 : Les différents modes d'évolution des familles multigéniques.**

(D'après Nei et coll. 2005).

Dans le contexte des systèmes ABH et sécréteur, aucun des différents phénotypes ne semble donner un avantage physiologique par rapport aux autres. En effet, un individu Bombay vit tout aussi normalement qu'un individu A, B ou sécréteur. Mais dans ce cas, pourquoi les HBGA sont-ils maintenus et pourquoi se sont-ils tant diversifiés à l'échelle de l'espèce, mais également entre les espèces ? Une réponse semble pouvoir être apportée au regard des inter-relations qui existent entre glycanes et pathogènes, et en particuliers entre les HBGA ABH et certains virus.

# **CHAPITRE 3 :**

## **DES GLYCANNES ET DES VIRUS :**

### **INTER-RELATIONS**

#### **1. DES SUCRES COMME RÉCEPTEURS AUX VIRUS**

Les virus sont des microorganismes situés à la frontière du monde du vivant et du monde de l'inerte. Responsables des plus grands fléaux jamais rencontrés dans la population humaine (grippe espagnole, SIDA, variole, etc ...), ils sont les prototypes du parasitisme obligatoire et ne peuvent de ce fait exister qu'au travers de leur hôte.

Afin de se multiplier, un virus doit faire entrer son matériel génétique à l'intérieur d'une cellule hôte qui est alors dite permissive. Celle-ci commence à répliquer le génome viral et met à la disposition du virus l'ensemble de sa machinerie cellulaire, nécessaire à la production des composants des futures particules filles. Avant de pouvoir se répliquer, le virus doit tout d'abord reconnaître de manière spécifique la cellule permissive et s'y fixer selon un mécanisme de type « clé-serrure ». Les glycanes portés par les différents glycoconjugués de la matrice extracellulaire peuvent jouer ce rôle de « serrure » et constituer des récepteurs ou co-récepteurs aux virus (Tableau 4). La grande majorité des glycanes impliqués sont fortement chargés négativement comme par exemple l'héparane sulfate (glycosaminoglycane sulfaté) ou les glycanes composés d'au moins un acide sialique en position terminale. Toutefois, certains glycanes neutres peuvent également être utilisés comme récepteurs viraux et notamment les HBGA (antigène A, B, H, Lewis, ... etc.).

Récepteur	Famille	Sous famille ou genre	Virus	Références	
Glycosaminoglycane	Parvoviridae	Dependovirus	Virus associés aux Adenovirus, AAV type 2	(Opie et coll. 2003)	
Héparane sulfate ( <sup>a</sup> ) Chondroïtine sulfate	Adenoviridae		Adénovirus 2, 5	(Dececchi et coll. 2000)	
	Flaviviridae	Hepacivirus	Virus hépatite C	(Barth et coll. 2003)	
		Flavivirus	Virus de la dengue	(Barth et coll. 2003)	
	Herpesviridae			Herpes simplex type 1 et 2 <sup>(b)</sup>	
		$\alpha$ Herpesvirus		Virus de la varicelle	(Z. Zhu et coll. 1995)
		$\beta$ Herpesvirus		Cytomegalovirus	(Kari et coll. 1993)
		$\gamma$ Herpesvirus		Herpes virus humain type 6 et 7	(Secchiero et coll. 1997)
	Papillomaviridae	Papillomavirus	Papillomavirus humain type 11,16 et 33	(Giroglou et coll. 2001; Shafti-Keramat et coll. 2003)	
	Paramyxoviridae	Pneumovirus	Virus respiratoire syncytial <sup>(a)</sup>	(Krusat et coll. 1997; Feldman et coll. 1999; Shields et coll. 2003)	
		Metapneumovirus	Metapneumovirus humain	(Wyde et coll. 2004)	
	Poxviridae	Orthopoxvirus	Virus de la vaccine <sup>(a)</sup>	(C. S. Chung et coll. 1998; J. C. Hsiao et coll. 1999)	
Héparane sulfate, Chondroïtine sulfate, Sulfatide, Galactosylcéramide	Retroviridae	Lentivirus	HIV 1	(Bomsel et coll. 2002; Bobardt et coll. 2004)	
Acide sialique <sup>(b)</sup> ( $\alpha$ 2-3) <sup>(c)</sup> ( $\alpha$ 2-6) <sup>(d)</sup> 9-O-acetyl acide sialique	Poxviridae Reoviridae	Orthovirus	Réovirus type 3	(Pacitti et coll. 1987; Gentsch et coll. 1987; Paul et coll. 1989)	
	Polyomaviridae		JC <sup>(c)</sup> BK virus	(Sinibaldi et coll. 1987; C K Liu et coll. 1998)	
			Enterovirus	Enterovirus 70	(Alexander et coll. 2002)
	Picornavirus		Rhinovirus	Rhinovirus 87	(Uncapher et coll. 1991; Blomqvist et coll. 2002)
		Parvoviridae	Dependovirus	Virus associés au Adenovirus AAV type 4 <sup>(b)</sup> AAV Type 5 <sup>(b) ou (c)</sup>	(Walters et coll. 2001; Kaludov et coll. 2001)
	Paramyxoviridae	Respirovirus	Paramyxovirus 1-3	(Ah-Tye et coll. 1999; T. Suzuki et coll. 2001)	
	Orthomyxoviridae		Influenza A	Influenza A aviaire <sup>(b)</sup> Influenza A humain <sup>(c)</sup>	(Rogers et coll. 1983; Webster et coll. 1992; M. N. Matrosovich et coll. 2004)
			Influenza B	Influenza B <sup>(b) ou (c)</sup>	(Y. Suzuki et coll. 1987)
			Influenza C	Influenza C <sup>(d)</sup>	(Rosenthal et coll. 1998)
	Coronaviridae		Coronavirus OC43 <sup>(d)</sup>	(Künkel et coll. 1993)	
	Adenoviridae		Adeno 37 <sup>(b)</sup>	(Arnberg et coll. 2002)	
Dystroglycane	Arenaviridae		Virus de Lassa	(W Cao et coll. 1998)	
Antigène tissulaire de groupe sanguin (ABH, Le)	Caliciviridae		Norovirus	(Marionneau et coll. 2002; A M Hutson et coll. 2002; P Huang et coll. 2003)	
Antigène de groupe sanguin (globoside, antigène P)	Parvoviridae	Erythrovirus	B19	(K. E. Brown et coll. 1993)	

**Tableau 4 : Liste des différents glycanes servant de récepteurs ou co-récepteurs aux virus.**

(D'après Olofsson et coll. 2005)

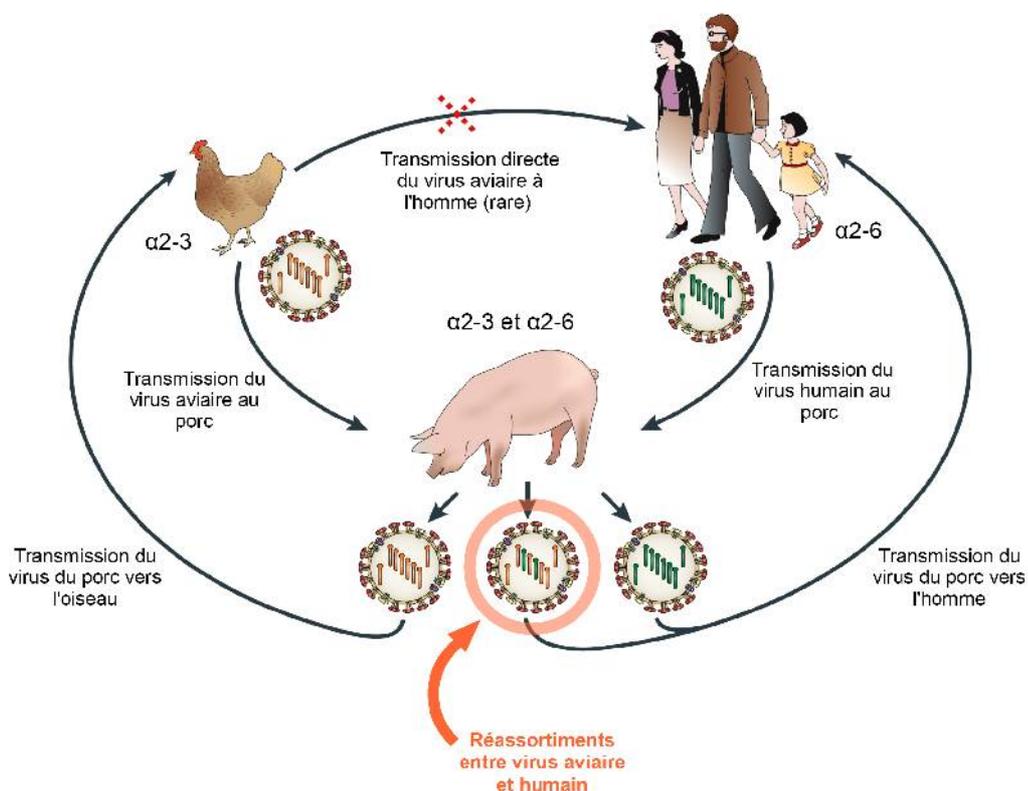


## **a - Acide sialique et virus Influenza**

Le IAV est un virus enveloppé à ARN simple brin de polarité négative. Son génome est découpé en huit segments indépendants qui codent la plupart du temps chacun pour une seule protéine virale. Il appartient à la famille des *Orthomyxoviridae* et au genre *Influenza virus*. On distingue plusieurs sous-types de IAV en fonction de la nature de deux glycoprotéines exprimées à sa surface : l'hémagglutinine (H1 à H16) et la neuraminidase (N1 à N9) (Fouchier et coll. 2005). Celles-ci sont enchâssées dans la matrice protéique du virus et traversent la bicouche lipidique de l'enveloppe afin de présenter la plus grande partie de leur architecture vers l'extérieur. Ces deux protéines jouent des rôles importants dans la physiologie du virus. L'hémagglutinine constitue la protéine virale se liant au récepteur cellulaire. Elle possède en effet à son sommet la poche d'attache à l'acide sialique. La neuraminidase est quant à elle une sialidase qui facilite la libération des virus néoformés en hydrolysant la liaison glycosidique reliant l'acide sialique au reste du glycanne.

Le réservoir naturel des IAV comprend différentes espèces d'oiseaux aquatiques (Webster et coll. 1992; Zambon 1999). L'homme est toutefois infecté par certains virus grippaux de type A, tout comme nombre d'autres vertébrés. Néanmoins, il a été montré que la transmission inter-espèce n'est pas toujours possible et que les virus restreints à certains organismes possédaient chacun une hémagglutinine avec une affinité forte et parfois exclusive pour un type d'acide sialique particulier (Y. Suzuki et coll. 2000). Il a notamment été mis en évidence que la nature de la liaison entre l'acide sialique et le galactose sous-jacent est critique pour la reconnaissance par telle ou telle hémagglutinine (Figure 14). Cette liaison glycosidique peut être de deux types :  $\alpha 2,3$  ou  $\alpha 2,6$ . Dans le cas des IAV aviaires, ceux-ci ne sont capables de reconnaître l'acide sialique que lorsqu'il est lié en  $\alpha 2,3$  sur le Gal. De leur côté, les IAV humains reconnaissent avec une grande affinité l'acide sialique lorsqu'il est lié en  $\alpha 2,6$  sur un Gal. Ils reconnaissent également les acides sialiques liés en  $\alpha 2,3$  mais avec une affinité faible. Les oiseaux ne sont capables d'exprimer que de l'acide sialique lié en  $\alpha 2,3$  au niveau de leur épithélium intestinal, site majeur de réplication du virus chez ces espèces. Chez l'homme, le virus se réplique dans les cellules non ciliées du tractus respiratoire (M. N. Matrosovich et coll. 2004). Or, ces cellules ne sont capables de présenter que des acides sialiques liés en  $\alpha 2,6$ . La transmission directe d'un virus d'origine aviaire à l'homme ne semble donc pas possible bien qu'elle ait eue lieu dans de rares occasions par le passé et notamment lors de l'épidémie de grippe espagnole de 1918. Toutefois, le passage par un hôte intermédiaire comme le porc peut faciliter le passage inter-espèce. En effet, le porc exprime au niveau de son tractus respiratoire des acides sialiques fixés par ces deux types de liaison, ce qui signifie qu'il peut être infecté par un spectre large de virus influenza. De par sa structure génomique segmentée, il est théoriquement possible de

réassortir chacun des gènes du IAV lors de la co-infection d'une cellule avec deux sous-types différents. Ce transfert peut être dramatique lorsque sont réassortis ceux codant l'hémagglutinine et la neuraminidase, réassortiments qui ont déjà été associés à une très forte augmentation de la virulence par le passé.



**Figure 15 : Schéma descriptif des différents modes de transmission des IAV entre les espèces.**

Dans ce modèle, le porc présente la caractéristique de pouvoir être infecté par les IAV aviaires et humains. Cette capacité à la co-infection permet aux IAV humains de pouvoir gagner en virulence au gré des réassortiments avec les virus aviaires. (D'après Stevens et coll. 2006)

### **b - Héparane sulfate et Herpes Simplex Virus**

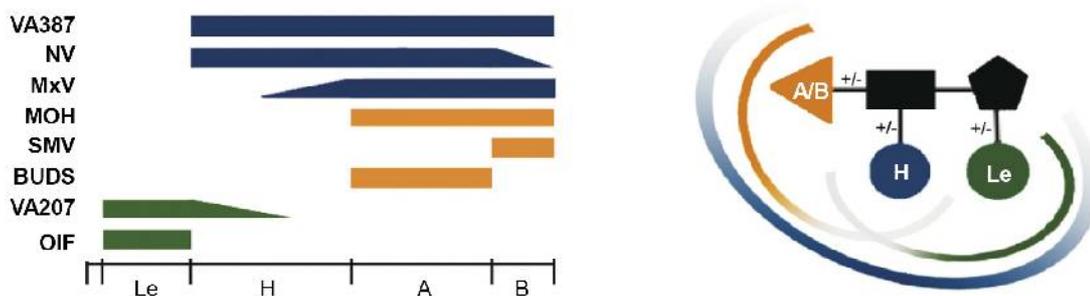
Les virus de l'herpès, appelés Herpes simplex virus (HSV), sont des virus enveloppés à ADN double brins non segmenté. Ils font partie de la famille des *Herpesviridea* qui comprend également le virus de la varicelle et du zona, le cytomégalovirus et le virus d'Epstein-Barr. Les HSV sont uniquement capables d'infecter l'homme, chez qui ils sont responsables de pathologies tels que l'herpès buccal (HSV de type 1) et l'herpès génital (HSV de type 2). Ce sont des virus dermo-neurotropes qui après la primo infection persistent de manière latente tout du long de la vie dans le ganglion de Gasser (HSV1) ou les ganglions sacrés (HSV2).

Parmi les onze glycoprotéines identifiées à la surface du virus, cinq participent à l'entrée du virus dans les cellules permissives : gB, gC, gD, gH et gL (P. G. Spear 1993). Le premier contact entre HSV et cellule hôte est réalisé par la glycoprotéine gC qui reconnaît certains sulfates présents sur les longues chaînes disaccharidiques constituant l'héparane sulfate (WuDunn et coll. 1989). La glycoprotéine gB peut également intervenir dans l'accrochage du virus mais avec une moindre importance pour HSV1 (Herold et coll. 1991) que pour HSV2, chez qui elle semble tenir une place importante dans le mécanisme de fixation (Cheshenko et coll. 2002). L'affinité de gC pour son récepteur est dépendante de l'interaction avec des sulfates spécifiques et correctement organisés sur la chaîne d'héparane sulfate. Les sulfates liés en 2-O sur le GlcA et 6-O sur le GlcNAc semblent plus importants que les sulfates liés en N sur le GlcNAc (Herold et coll. 1995). Pour le HSV1, la molécule minimum requise pour la fixation de gC sur l'héparane sulfate est un dodécamère N sulfaté initié par un GlcNAc portant un sulfate supplémentaire en 6-O, suivi par un GlcA portant un sulfate en 2-O (Figure 14) (Feyzi et coll. 1997). Bien que leurs séquences protéiques soit très proches, la gC de HSV2 montre une plus grande avidité pour son récepteur glycanique que la gC de HSV1. Ceci pourrait être en partie expliqué par la délétion d'une séquence de type mucine à proximité du domaine de liaison de gC chez HSV2 (Trybala et coll. 2000; Adamiak et coll. 2007). De plus, des expériences d'inhibition de la fixation réalisées à l'aide d'héparine et de variants sélectivement désulfatés ont montré des différences en fonction de l'origine de gC. En effet, celle d'HSV2 semble plus sensible à l'inhibition que celle de l'HSV1, ce qui suggère qu'elles reconnaissent des parties différentes de la structure de l'héparane sulfate (Herold et coll. 1996; Trybala et coll. 2000) ou que des régions différentes des gC soient impliquées dans cette reconnaissance, comme par exemple le domaine mucine de HSV1 (Adamiak et coll. 2007).

## **1.2. Glycanes non chargés**

Très peu de sucres non chargés (dits « neutres ») sont utilisés comme récepteurs par les virus. A ce jour, seules les familles des *Caliciviridae* et des *Parvoviridae* comptent des membres capables de reconnaître respectivement les antigènes ABH et l'antigène P (voir Tableau 4). Tous ces glycanes appartiennent à la grande famille des HBGA. Comme dans le cas des glycanes chargés décrit précédemment, la liaison de la protéine virale d'attachement sur les sucres neutres s'effectue de façon hautement spécifique en établissant des liaisons faibles entre certains atomes radicalaires d'acides aminés de la poche d'attache et des atomes particuliers du glycanne. Toutefois ici, les forces électrostatiques ne peuvent intervenir pour faciliter ou stabiliser la liaison virus/glycanne. Parmi les virus utilisant les HBGA comme récepteurs, le cas des Norovirus (NoV) est particulièrement bien documenté.

Les Norovirus (NoV) sont des virus non enveloppés à ARN simple brin de polarité positive et polyadénylé. Ils appartiennent à la famille des *Caliciviridae* et sont responsables chez l'homme de la majorité des cas de gastro-entérites hivernales d'origine non bactérienne. Son génome est divisé en trois cadres ouverts de lecture dont le deuxième (ORF2) code VP60, l'unique protéine de capsid. Cette protéine est impliquée dans la reconnaissance de la cellule hôte par le virus. Celle-ci peut être divisée en trois sous domaines structuraux : S correspondant à la partie coque, P1 et P2 formant la partie saillante de la capsid. Les antigènes H type 1 et 3/4 sont les premiers à avoir été identifiés comme étant reconnus par le virus de Norwalk (NV), souche prototype du genre NoV (Marionneau et coll. 2002). Depuis cette découverte, huit profils différents de fixation ont été mis en évidence pour l'ensemble des souches de NoV, avec pour chacun d'eux des affinités différentes pour les HBGA A, B, H ou Lewis (P Huang et coll. 2005; Le Pendu et coll. 2006) (Figure 16). L'interaction HBGA/NoV sera abordée plus en détails dans le b(p.57).



**Figure 16 : Représentation schématique des huit profils de fixation observés chez les NoV et modélisation des glycanes préférentiellement reconnus par ces souches.**

(D'après Le Pendu et coll. 2006)

## 2. GLYCOSYLATION DES VIRUS

Lors de leur réplication, les virus sont capables d'utiliser la machinerie enzymatique de la cellule infectée pour modifier leurs protéines. Ces modifications peuvent jouer un rôle critique dans la fonction de certaines glycoprotéines virales, impactant notamment sur leur stabilité, leur antigénicité et la reconnaissance de la cellule hôte. Chez les virus enveloppés, la glycosylation représente une part importante de ces modifications.

Au cours de son évolution, un virus peut augmenter ou diminuer le nombre de ses sites de N-glycosylation en modifiant ponctuellement sa séquence peptidique. Toutefois,

quelques uns semblent essentiels à la survie de certains virus, en participant notamment au trafic des protéines virales à l'intérieur de la cellule hôte et au bon repliement de certaines de leurs glycoprotéines (Meunier et coll. 1999; Land et coll. 2001; Slater-Handshy et coll. 2004). La glycosylation des virus va également jouer un rôle important dans différents mécanismes liés au maintien ou à l'augmentation de la virulence, mais également dans la réponse de l'hôte contre ces agents pathogènes.

## **2.1. Echappement par masquage des épitopes viraux**

Le masquage de certains épitopes viraux par des glycanes correspond à un mécanisme mis en place par les virus afin d'échapper à leur reconnaissance par les anticorps neutralisants de l'hôte. Il consiste en la formation d'un « bouclier » de glycanes qui bloque tout accès au domaine protéique adjacent. Pour certains virus, ce mécanisme semble être aussi efficace que la modification de la séquence primaire des domaines protéiques exposés, autre moteur de l'échappement des virus à la réponse immune de l'hôte. Elle a notamment été bien décrite pour le IAV et le virus de l'immunodéficience humaine (HIV).

Il a été montré pour le IAV que l'augmentation du nombre de sites de N-glycosylation sur l'hémagglutinine est associée à une diminution de l'accessibilité et de la reconnaissance de certains épitopes impliqués dans la mise en place de la réponse immunitaire adaptative de l'hôte (Abe et coll. 2004). Toutefois, ces modifications ont un impact sur la qualité de la fixation de la glycoprotéine sur l'acide sialique en entraînant une diminution de sa capacité de reconnaissance du sucre. Cette dégradation de l'affinité du virus pour son récepteur est néanmoins insuffisante pour compromettre sa survie. L'état de glycosylation de l'hémagglutinine a également un impact important sur le relargage des virus néoformés. En effet, moins elle sera glycosylée plus la liaison à l'acide sialique sera forte et par conséquent plus l'activité sialidase de la neuraminidase devra être élevée pour libérer les particules virales filles (Klenk et coll. 2002). Dans le cas précis du IAV, la cinétique de réplication est dépendante d'une balance entre échappement à la réponse immune et virulence et il a été montré récemment que l'état de glycosylation de l'hémagglutinine est inversement proportionnel à la virulence (Vigerust, Ulett, et coll. 2007).

Dans le contexte de l'échappement du virus à la réponse immune, le HIV semble avoir adopté la voie d'une hyperglycosylation de certaines de ces protéines et notamment de la glycoprotéine d'enveloppe gp120 (Kwong et coll. 2002; X. Wei et coll. 2003). En effet, gp120 est à ce jour l'une des protéines les plus glycosylées jamais identifiée, les N-glycanes représentant environ la moitié de sa masse moléculaire (J. S. Allan et coll. 1985). Il a été montré qu'elle peut posséder entre 18 et 33 sites de N-glycosylation avec une moyenne de l'ordre de 25 positions possibles (Korber et coll. 2001). L'impact de

l'augmentation du nombre de N-glycosylations de gp120 sur l'échappement du HIV à la reconnaissance par les anticorps sériques de l'hôte a été évalué et confirmé par de nombreux auteurs (Sagar et coll. 2006; Wolk et coll. 2006). Il a de plus été montré que l'augmentation du nombre de N-glycanes dans les régions voisines de celle permettant la production d'anticorps neutralisants était suffisante pour induire cet échappement tout en préservant l'intégrité du domaine protéique impliqué dans la liaison au récepteur cellulaire (X. Wei et coll. 2003).

## **2.2. Facilitation par les lectines de l'hôte**

Les lectines sont des protéines qui reconnaissent et se fixent spécifiquement sur un ou des motifs oligosaccharidiques présent à la surface de certaines cellules eucaryotes, procaryotes ou bien sur les virus. Celles-ci existent aussi bien dans le règne animal que dans le règne végétal. Elles participent notamment chez l'homme à l'immunité innée en se liant aux oligosaccharides de certains pathogènes, permettant ainsi leur élimination via l'activation de la cascade du complément lectine-dépendante (Ikeda et coll. 1987). Parmi les lectines identifiées chez l'homme, certaines appartenant au type C (liaison calcium-dépendante) jouent un rôle essentiel dans de nombreuses étapes de la réponse immune de l'hôte (Cambi et coll. 2003). Toutefois, certains virus ont su tirer profit de leur caractère « nécessaire » pour infecter ou faciliter l'infection de certains types cellulaires.

Parmi ces molécules, la protéine membranaire DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin = CD209) est utilisée comme porte d'entrée par de nombreux virus comme le virus West Nile (Davis et coll. 2006), le virus Ebola (Alvarez et coll. 2002), le virus de l'hépatite C (Lozach et coll. 2003) et le HIV (Geijtenbeek et coll. 2000) pour ne citer qu'eux. Exprimée à la surface des cellules dendritiques et de certains macrophages, elle participe normalement au mécanisme d'endocytose des pathogènes richement mannosylés et à l'activation des lymphocytes T par les cellules dendritiques via l'interaction avec la molécule ICAM3 lymphocytaire. Une des fonctions principales des cellules dendritiques est à la mise en place de la réponse immunitaire adaptative. Pour cela, elle capture les antigènes par endocytose, les dégrade par protéolyse, puis apprête les peptides exogènes sur les molécules du complexe majeur d'histocompatibilité afin de les présenter aux lymphocytes T présents au niveau des organes lymphoïdes secondaires. Dans le cas du HIV, DC-SIGN va reconnaître les multimères de mannose présents sur la glycoprotéine de capsid gp120 (Hong et coll. 2007). Suite à leur endocytose, les virions ne se répliquent pas mais utilisent la cellule dendritique comme « cheval de Troie » afin d'atteindre les organes lymphoïdes secondaires. Ici, de très nombreux lymphocytes T CD4 sont concentrés. Ceux-ci vont alors être infectés lors de la formation de la synapse immunitaire avec les cellules

dendritiques séquestrant le HIV (Arrighi et coll. 2004). Ce mécanisme appelé *trans* infection par Geijtenbeek permet au HIV d'augmenter fortement l'efficacité de sa transmission en étant véhiculé directement vers les cellules T CD4 permissives, tout en restant imperceptible par l'immunité de l'hôte à l'intérieur des cellules dendritiques (Geijtenbeek et coll. 2000).

Il a également été suggéré que la molécule DC-SIGN puisse participer à la *cis* infection, c'est à dire l'infection directe des cellules T CD4 capables de répliquer efficacement le virus (B Lee et coll. 2001). Cela a été montré *in vitro* pour le virus de la rougeole où est observée une corrélation entre augmentation de la *cis* infection des cellules dendritiques et expression de la lectine DC-SIGN à leur surface. Pour ce virus, il a également été mis en évidence que DC-SIGN est impliquée dans la *trans* infection des lymphocytes au niveau des organes lymphoïdes secondaires (de Witte et coll. 2008).

D'autres lectines interviennent dans la *cis* ou la *trans* infection comme la lectine fixant le mannose (mannose binding lectin) dans le cas du virus influenza (Thielens et coll. 2002), la lectine hépatique L-SIGN qui pourrait être impliquée dans le tropisme du virus de l'hépatite C (Lozach et coll. 2003) et la lectine LSEctin qui permet d'augmenter l'infectivité du coronavirus du SRAS et des Filovirus (Gramberg et coll. 2005).

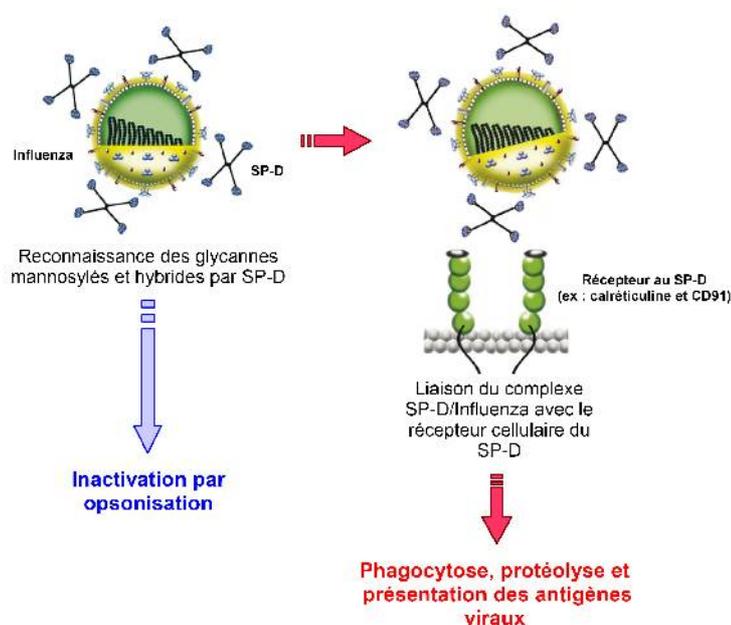
### **2.3. Protection par les CBA**

Bien que les glycanes arborés par les virus puissent constituer des atouts contre la mise en place de l'immunité de l'hôte et favoriser l'infection des cellules permissives, ceux-ci peuvent également les desservir en constituant des cibles de choix pour des leurres solubles endogènes ou exogènes ainsi que des épitopes spécifiques pour certains anticorps neutralisants. Ces molécules fixant les glycanes viraux (Carbohydrate-binding agents = CBA) vont avoir deux types d'action antivirales (Balzarini 2007) :

- Elles vont avoir une activité antivirale directe en opsonisant les virions, empêchant ainsi leur entrée dans la cellule par masquage du domaine de liaison au récepteur. En fonction de la nature du CBA, l'opsonisation pourra être suivie de la prise en charge des virions par les cellules présentatrices d'antigènes en vue de leur dégradation, suivie de l'initiation de la réponse immunitaire adaptative.
- Elles vont avoir une activité antivirale indirecte en exerçant une pression de sélection négative sur les sites de N-glycosylation. Elles favorisent ainsi l'apparition de mutations et la création progressive de délétions dans la carapace glycanique de certaines glycoprotéines virales. Il en résulte l'exposition ou réexposition d'épitopes permettant l'action neutralisante de certains anticorps.

Les CBA comprennent certaines lectines endogènes solubles comme les protéines A et D du surfactant pulmonaire (SP-A et SP-D) mais également des anticorps dirigés contre des glycanes viraux comme l'anticorps monoclonal anti-HIV 2G12.

La lectine SP-D fait partie de la famille des collectines. Elle participe activement à la réponse immunitaire innée dirigée contre de nombreux pathogènes dont le IAV (Kishore et coll. 2006). En effet, celle-ci induit une agrégation massive du IAV qui est alors neutralisé du fait de l'inhibition de l'interaction acide sialique/hémagglutinine (Hartshorn et coll. 1996; Reading et coll. 1997). La capacité d'inactivation de SP-D dépend fortement de l'état de glycosylation de l'asparagine 165 de l'hémagglutinine, position située à proximité du domaine de reconnaissance de l'acide sialique (Hartshorn et coll. 1994; Reading et coll. 1997; Hawgood et coll. 2004). De plus, les macrophages possèdent des récepteurs de surface, comme la calréticuline et CD91, capables de reconnaître les complexes SP-D/virus et d'augmenter leur phagocytose et la production de molécules pro-inflammatoires (LeVine et coll. 2001; Gardai et coll. 2003) (Figure 17).



**Figure 17 : Action antivirale directe de la protéine D du surfactant.**

(D'après Vigerust et Shepherd 2007).

Dans la grande majorité des cas, la glycosylation des glycoprotéines virales est dépendante de la machinerie enzymatique de la cellule hôte, si bien que les glycanes présentés par les virus sont perçus le plus souvent comme des molécules du « Soi » par les acteurs de la réponse immunitaire. Toutefois, il a été montré pour le virus du SIDA que certains oligosaccharides terminaux mannosylés sont la cible d'anticorps neutralisant, comme l'anticorps monoclonal 2G12 qui est l'un des premiers à avoir été bien caractérisé (Trkola et coll. 1996). Il a été montré que cet anticorps reconnaît spécifiquement les sucres présents sur les asparagines 295, 332 et 392 de la

glycoprotéine gp120 (Calarese et coll. 2003) et que les N-glycosylations des asparagines 295, 392 et/ou 406 étaient nécessaires pour qu'il puisse neutraliser le virus (Huskens et coll. 2007). Il a également été montré que les structures minimales qu'il est capable de reconnaître sont le Man $\alpha$ 1,2Man $\alpha$ 1,3Man et le (Man $\alpha$ 1,3)(Man $\alpha$ 1,6)Man (Astronomo et coll. 2008). Cet anticorps anti-glycanes présente un avantage certain comparé aux anticorps neutralisants dirigés contre les épitopes protéiques. En effet, les asparagines mannosylées décrites précédemment sont très conservées entre les différentes clades de HIV. Néanmoins, il a été montré que l'anticorps 2G12 est capable de neutraliser 41% des souches de HIV mais aucune des clades C et E (Binley et coll. 2004).

## **2.4. Reconnaissance par les anticorps naturels de l'hôte**

Les anticorps naturels (NAb) sont des anticorps sériques dont l'origine n'est pas liée à une immunisation induite volontairement (vaccination) ou résultant d'une primo infection. Ce sont majoritairement des immunoglobulines de type M qui présentent le plus souvent un large spectre de reconnaissance antigénique associé à une faible affinité pour les différents épitopes. Ils constituent ainsi l'une des premières lignes de défense contre les pathogènes. Les plus célèbres de ces NAb sont ceux dirigés contre les HBGA, et notamment contre les antigènes A, B, et  $\alpha$ Gal. Leur origine diffère pourtant des NAb conventionnels. En effet, elle a été décrite comme pouvant être le résultat de contacts répétés avec des antigènes bactériens ou végétaux de type HBGA via les muqueuses du tube digestif (Springer et coll. 1969; Hammer et coll. 1992).

Le rôle possible de ces NAb anti-HBGA dans la protection antivirale et dans l'initiation de la réponse immunitaire T cytotoxique a été particulièrement bien documenté par les publications de l'équipe de Kenth Gustafsson. Le premier argument allant dans le sens de cette hypothèse est que certains virus en se répliquant dans des cellules capables de synthétiser des HBGA arborent à leur tour sur leurs glycoprotéines ces même HBGA (Y. Takeuchi et coll. 1996; Preece et coll. 2002). De plus, lorsque les virus ainsi décorés sont mis en contact avec un sérum contenant des NAb dirigés contre ces motifs, ils voient leur infectivité très fortement diminuée. Il a également été mis en évidence qu'associés aux facteurs du complément, les complexes immuns NAb/virus ainsi formés étaient capables d'induire la prolifération des cellules T CD8, acteurs de la réponse immunitaire antivirale T cytotoxique (Dürrbach et coll. 2007).

# **CHAPITRE 4 :**

## **VIRUS ET HBGA : EXEMPLES**

### **DES CALICIVIRUS ET DU SRAS-**

### **CoV**

## **1. CALICIVIRUS ET HBGA**

### **1.1. Taxonomie**

Le nom *Caliciviridae* (Calicivirus) dérive du mot latin *Calix* qui signifie « coupe » ou « calice ». Il a été choisi pour désigner cette famille de virus en raison de l'aspect caractéristique des motifs observés à la surface de la capsidie en microscopie électronique. Pour appartenir à la famille des Calicivirus, un virus doit posséder les caractéristiques suivantes (Francki et coll. 1991) :

- Génome de 7 à 8 kb, constitué d'ARN simple brin de polarité positive, polyadénylé à son extrémité 3' et ne présentant pas de coiffe méthylée à son extrémité 5'.
- Virion non enveloppé de 35 à 39 nm de diamètre, organisé selon une symétrie icosaédrique et présentant 32 dépressions de surface en forme de cupule.
- Protéine de capsidie unique de 60 à 70 kDa.

Cette famille comprend les genres Lagovirus, Vesivirus, Norovirus (anciennement Norwalk-like virus) et Sapovirus (anciennement Sapporo-like virus) (K. Y. Green et coll. 2000). D'autres virus ayant des caractéristiques morphologiques et physico-chimiques proches de celles des *Caliciviridae* ont été identifiés mais n'ont pas encore pu être rattachés à l'un de ces quatre genres. Il a été notamment mis en évidence des Calicivirus chez le chimpanzé (A. W. Smith, Skilling, Ensley, et coll. 1983), la vache (A. W. Smith et coll. 1983), le vison (Seal et coll. 1995), le chien (Schaffer et coll. 1985), le poulet (Cubitt et coll. 1985), le lion (Kadoi et coll. 1997) et la sterne blanche (Poet et coll. 1996).

Récemment, un nouveau genre a été proposé chez les Calicivirus : le genre Becovirus (Oliver et coll. 2006). Genre regroupant exclusivement des virus infectant les espèces bovines, sa souche prototype est le virus Newbury-1.

### **a - Vesivirus**

Les Vesivirus sont essentiellement des virus infectant les animaux. Ce genre contient les premiers membres caractérisés de la famille des Calicivirus : le Vesicular Exanthema of Swine Virus (VESV), le Feline Calicivirus (FCV) et le San Miguel Sea lion Virus (SMSV).

Le VESV est le virus prototype des *Caliciviridae*. Apparue en 1932, il a été responsable d'exanthèmes vésiculeux épizootiques chez le porc, avant d'être éradiqué en 1959 (Traum 1934; A. W. Smith et coll. 1976). Le FCV quant à lui a été isolé en 1957 en Nouvelle Zélande (Fastier 1957). Il provoque chez le chat une maladie respiratoire appelée coryza. Virus à tendance ubiquiste, le FCV va également pouvoir entraîner des affections buccales et des problèmes locomoteurs. Le SMSV fut lui isolé en 1972 suite à un prélèvement réalisé sur une otarie californienne (A. W. Smith et coll. 1973). L'origine porcine (VESV) du SMSV est probable du fait de sa transmission possible au porc et de la zone géographique où il est apparu (élevage intensif de porc) (Prato et coll. 1974). Il a été depuis retrouvé chez d'autres mammifères marins et terrestres de la côte ouest américaine (A. W. Smith, Skilling, et Ridgway 1983; Neill et coll. 1995; A. W. Smith et coll. 1998). Les Vesivirus présentent la caractéristique d'être les seuls Calicivirus à pouvoir être cultivés par infection de cultures cellulaires.

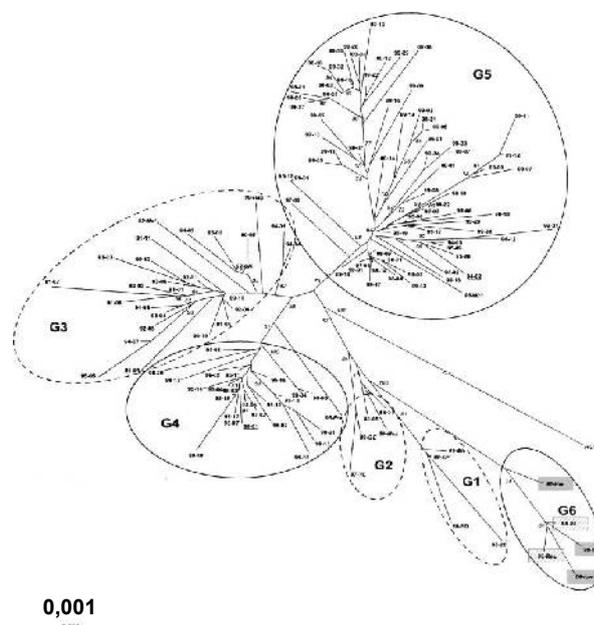
### **b - Lagovirus**

Les Lagovirus sont des virus propres à l'ordre des lagomorphes et en particuliers à certaines espèces de lapin et de lièvre. Les prototypes de ce genre sont le Rabbit Haemorrhagic Disease Virus (RHDV) et l'European Brown Hare Syndrome Virus (EBHSV) qui vont respectivement infecter le lapin et le lièvre.

Apparu en 1984 en Chine (S. J. Liu et coll. 1984), le RHDV est responsable chez le lapin de garenne et le lapin domestique (*Oryctolagus cuniculus*) d'un syndrome caractérisé par une hépatite fulminante, accompagnée d'hémorragies disséminées dans tout l'organisme (Marcato et coll. 1991). Ces lésions entraînent la mort de l'animal dans les 24 à 48 heures post-infection et le taux de mortalité généralement rencontré pour ce virus est de l'ordre de 70 à 100% (S. J. Liu et coll. 1984; Z. J. Xu et coll. 1989).

L'EBHSV a été décrit pour la première fois en Suède en 1980 (Gavier-Widén et coll. 1991), mais certains chasseurs scandinaves semblaient connaître la maladie depuis le début des années 1970 (Lenghaus et coll. 2001). L'EBHSV provoque chez le lièvre brun européen (*Lepus europeus*) et le lièvre variable (*Lepus timidus*) une pathologie comparable à celle du RHDV chez le lapin. Toutefois, son taux de mortalité reste beaucoup moins élevé que celui observé pour le RHDV.

Concernant le RHDV, de nombreuses études phylogénétiques ont été menées depuis l'apparition du virus. Elles ont permis de suivre son évolution au cours des différentes épidémies. En France, elles font état d'une très faible variation des séquences nucléotidiques pour les souches circulant durant la période de 1988 à 2000 (Le Gall-Reculé et coll. 2003). Néanmoins, ces modifications sont suffisantes pour permettre de les séparer en six génogroupes (G1 à G6), avec un maximum de divergence de 8,5% entre les récents variants du GV et la souche de référence française 89-SD (Figure 18).



**Figure 18 : Arbre phylogénétique du RHDV représentant les différents génogroupes circulant ou ayant circulé en France.**

Les génogroupes symbolisés par un trait plein coexistent actuellement en France. Ceux symbolisés par des tirets n'ont à ce jour plus été mis en évidence sur le territoire français. (D'après Le Gall-Reculé et coll. 2003)

Une étude hongroise sur un panel de souches circulant entre 1988 et 2003 confirme ces observations avec un maximum de divergence de 9,5% pour les séquences nucléotidiques et de 2,6% pour les séquences peptidiques (Matiz et coll. 2006).

L'existence de recombinaisons entre souches différentes de RHDV a été récemment mise en évidence. Elles pourraient ainsi participer fortement à l'augmentation du polymorphisme du génome viral (Abrantes et coll. 2008; Forrester et coll. 2008).

### ***c - Norovirus***

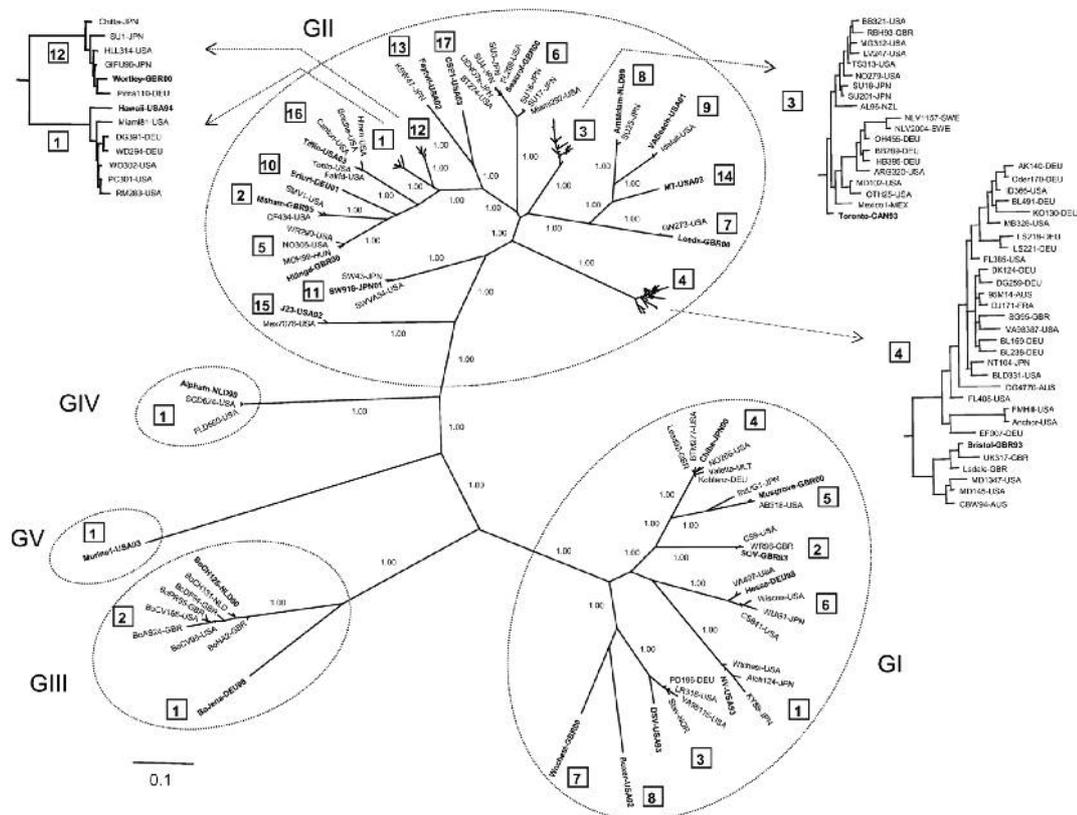
Les Norovirus (NoV) infectent principalement l'homme. Ils sont responsables chez lui de gastro-entérites aiguës pouvant parfois être fatales pour les individus les plus faibles (nourrissons, personnes âgées, personnes malnutries) en raison de la très forte déshydratation engendrée.

Les NoV ont été découverts et décrits pour la première fois suite à une épidémie de gastro-entérite ayant eue lieu en 1968 dans une cantine scolaire de la ville de Norwalk, Ohio (Kapikian et coll. 1972). Ces virus sont responsables de 60 à 85% des cas de gastro-entérites hivernales. Aux Etats Unis, ils ont été mis en évidence dans 96% des épidémies de gastro-entérites d'origine virale (Fankhauser et coll. 1998). Très contagieux, ils infectent les personnes indépendamment de leur âge et de leur niveau social. Leur mode de transmission utilise la voie féco-orale. Celle-ci est favorisée par des contacts rapprochés avec des personnes malades ou en rémission (aérosols et selles) et la consommation de nourriture (huître, salade, framboise, pâtisserie, etc...) ou d'eau contaminées (Rockx et coll. 2002).

Les NoV peuvent être subdivisés en 5 génogroupes (GI à GV) d'après l'analyse phylogénétique de la séquence peptidique de leur capsid (D-P Zheng et coll. 2006) (Figure 19): les NoV du GI et du GIV vont infecter exclusivement l'homme, ceux du GII peuvent infecter l'homme et le porc, alors que les NoV du GIII infectent les bovins et que le NoV du GV infecte spécifiquement la souris.

Les NoV présentent entre eux une très grande variabilité dans leur séquence génomique. Chacun des génogroupes a donc été subdivisé en clusters pour rendre compte de cette diversité. La variabilité peut atteindre dans la séquence codant la capsid, de 0 à 14,1% de variabilité entre les souches d'un même cluster, de 14,3 à 43,8% entre les clusters d'un même génogroupe et de 44,9 à 61,4% entre les différents génogroupes (D-P Zheng et coll. 2006).

Plus récemment, certains auteurs ont proposé de différencier deux nouveaux génogroupes : le GVI correspondant à la souche J23/99/US et GVII correspondant à la souche Chiba (Phan et coll. 2007).



**Figure 19 : Arbre phylogénétique des NoV et représentation des différents génogroupes.**

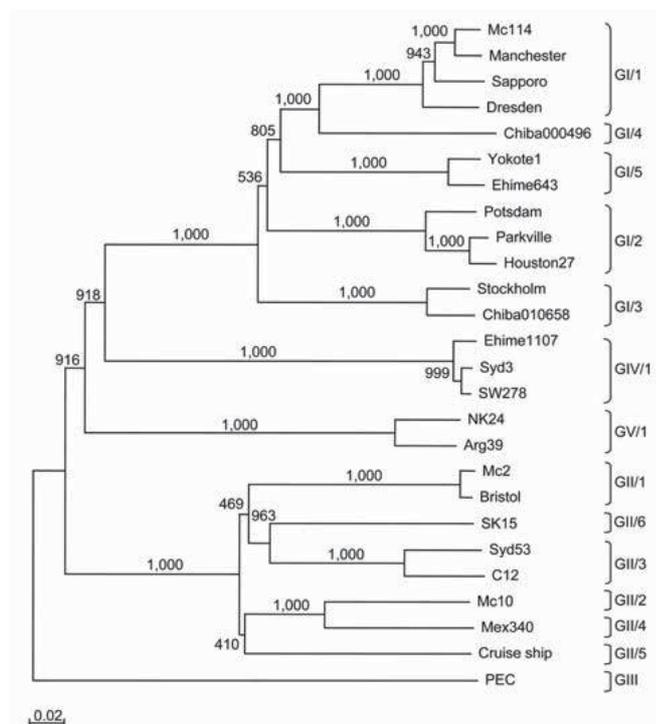
Réalisé d'après la comparaison des séquences entières de la protéine de la capside. Sont représentés les cinq génogroupes ainsi que les différents clusters (nombre encadré). (D'après Zheng et coll. 2006)

### d - Sapovirus

Les Sapovirus (SaV) infectent principalement l'homme. Ils sont également responsables de gastro-entérites, mais moins sévères, affectant surtout le jeune enfant.

Ils ont été découverts et décrits suite à une épidémie ayant eu lieu en 1977 dans un orphelinat de la ville japonaise de Sapporo (Chiba et coll. 1979). Les données épidémiologiques des SaV sont beaucoup moins pertinentes que celles des NoV, du fait de leur plus faible pathogénicité, de la tranche d'âge des individus infectés et surtout du moins grand nombre de laboratoires les étudiant. Toutefois, une étude portant sur une cohorte d'enfants mexicains a montré que les SaV peuvent représenter jusqu'à 40% des Calicivirus détectés dans les selles d'enfants malades (Farkas et coll. 2000). De plus, les études de séro-prévalences montrent que virtuellement tous les enfants sont infectés par les SaV avant d'atteindre l'âge de 5 ans. Leur transmission s'effectue par voie féco-orale et est favorisée par les mêmes facteurs que pour les NoV (Hansman et coll. 2007).

Les SaV sont divisés en 5 génogroupes (GI à GV) (Figure 20) : les virus du GI, GII, GIV et GV vont infecter exclusivement l'homme alors que ceux du GIII vont infecter spécifiquement le porc.



**Figure 20 : Arbre phylogénétique des SaV et représentation des différents génogroupes.**

Réalisé d'après la comparaison des séquences entières de la protéine de la capsid. (d'après Hansman et coll. 2005)

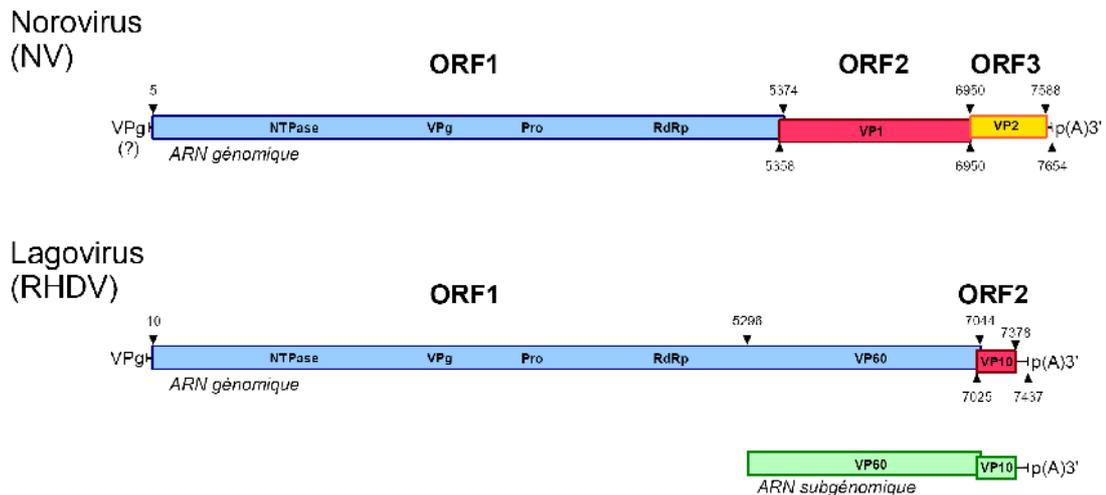
## **1.2. Génome viral et protéines associées**

Tous les Calicivirus possèdent un génome constitué d'une molécule d'ARN simple brin de polarité positive, polyadénylée à l'extrémité 3'. Sa taille est comprise entre 7,3 et 8,3 kb, queue polyadénylée exclue.

Le génome des Calicivirus peut présenter un nombre variable de cadres ouverts de lecture (Figure 21) :

- Les Vesivirus et les Lagovirus en possèdent deux : Le premier (ORF1) code la protéine structurale majeure de capsid et la polyprotéine à l'origine des différentes protéines virales non-structurales. Le second (ORF2) code la protéine structurale mineure de la capsid.

- Les Norovirus et les Sapovirus en possèdent trois : Le premier code la polyprotéine à l'origine des différentes protéines virales non-structurales. Le second (ORF2) code la protéine structurale majeure de la capside. Le troisième (ORF3) code la protéine structurale mineure de la capside.



**Figure 21 : Représentation du génome des Calicivirus et de son organisation : exemples du NV et du RHDV.**

*NTPase = nucléoside triphosphatase 2C-like; Pro = Protéinase 3C-like; RdRp = ARN polymérase ARN dépendante; VP1 = VP60 = capside = protéine structurale majeure ; VP2 = VP10 = protéine structurale mineure.*

Les Vesivirus et les Lagovirus présentent également la caractéristique de posséder un ARN subgénomique colinéaire de la fin de l'ORF1 (capside) et de l'ensemble de l'ORF2 de leurs ARN génomiques respectifs. Ces deux types d'ARN sont reliés covalamment en 5' à l'aide d'une petite protéine virale d'environ 13 kDa, la VPg.

La majorité des protéines virales matures proviennent de l'autocatalyse de la polyprotéine précurseur (ORF1) par la protéase 3C-like (Figure 21). Ces clivages vont au final donner les protéines non-structurales : l'hélicase 2C-like, VPg, la protéase 3C-like et l'ARN polymérase ARN dépendante 3D-like. Dans le cas des Lagovirus et des Vesivirus, le clivage de la polyprotéine est également à l'origine de la protéine de capside. Ce clivage génère également trois autres protéines dont la fonction n'a pas encore été déterminée.

### **1.3. Morphologie et structure**

A l'exception des Vésivirus, les Calicivirus ne peuvent pas être amplifiés par infection de cultures cellulaires. Toutefois, une technique permettant d'obtenir des pseudo particules virales (VLP) a été développée pour nombre d'entre eux (X Jiang et coll. 1992, 1999; Nagesha et coll. 1995; M. Guo et coll. 2001; Di Martino et coll. 2007). Copies fidèles de la capside du virus sauvage, les VLP possèdent les mêmes

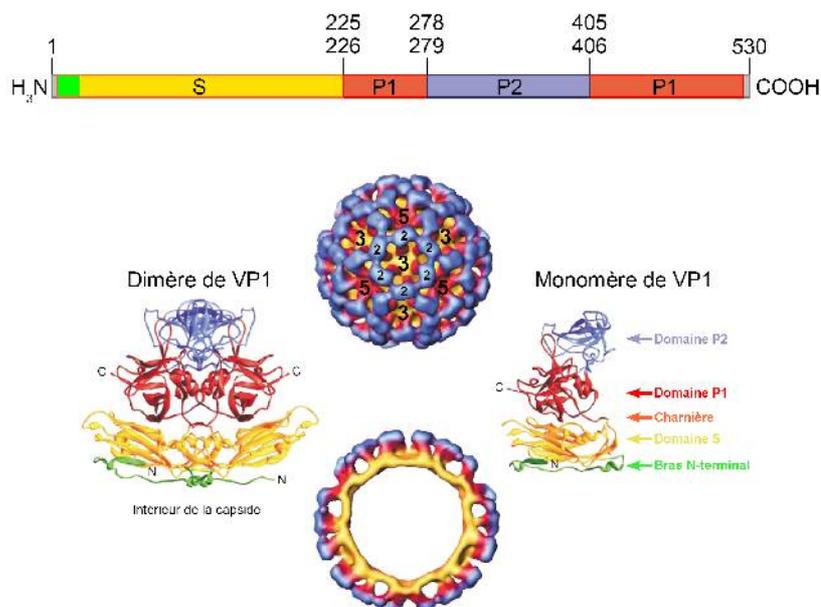
caractéristiques antigéniques et morphologiques que celle-ci. Elles présentent également l'avantage de ne pas être infectieuses étant privées de tout génome viral. Celles-ci sont produites en cellules d'insecte par infection avec un vecteur baculovirus recombinant contenant le gène de la capsid. L'étude des VLP en cristallographie rayon X a notamment permis à l'équipe de Venkataram Prasad de bien caractériser la structure et l'assemblage de la capsid du NV (Prasad et coll. 1999). Dans la suite de ce manuscrit, j'ai choisi de décrire en détail les caractéristiques structurales et morphologiques de la capsid des NoV présentées dans cet article, ces données pouvant être facilement transposables aux capsides des autres membres de la famille des Calicivirus.

### ***a - La protéine majeure de capsid***

L'une des caractéristiques des Calicivirus est de ne posséder qu'une seule protéine structurale, la protéine VP1 (VP60 pour le RHDV). Une fois associée en dimères, elles sont utilisées comme bases élémentaires pour l'assemblage de leur capsid. Cette singularité implique que les premiers contacts physiques avec les cellules permissives à l'infection se font via des domaines particuliers de VP1.

VP1 peut être décomposée en 2 domaines : un domaine S (pour Shell) et un domaine P (pour Protruding) (Figure 22). Ceux-ci sont reliés par une charnière peptidique conférant une certaine flexibilité à la protéine. Le domaine S est localisé à l'extrémité N-terminale et s'étend sur 225 acides aminés chez le NV. Il participe essentiellement à la formation de la partie coque de la capsid. Le domaine P est quant à lui constitué du reste de la protéine et forme la partie saillante de cette dernière. Il s'étend de la position 225 jusqu'à l'extrémité C-terminale de la protéine et participe à la dimérisation de la protéine VP1.

Le domaine P peut à son tour être subdivisé en deux sous-domaines : le sous-domaine P1 et le sous-domaine P2. Le sous-domaine P1 s'étend entre les acides aminés en position 226 à 278 et 406 à 520. Le sous-domaine P2 présente en effet la particularité d'être intégré entre les deux séquences composant le sous-domaine P1, c'est à dire entre les positions 279 et 405 de la séquence peptidique de la capsid.



**Figure 22 : La protéine de capsid VP1 du NV et structures tri- et bidimensionnelles de la capsid du NV.**

(D'après Hutson et coll. 2004)

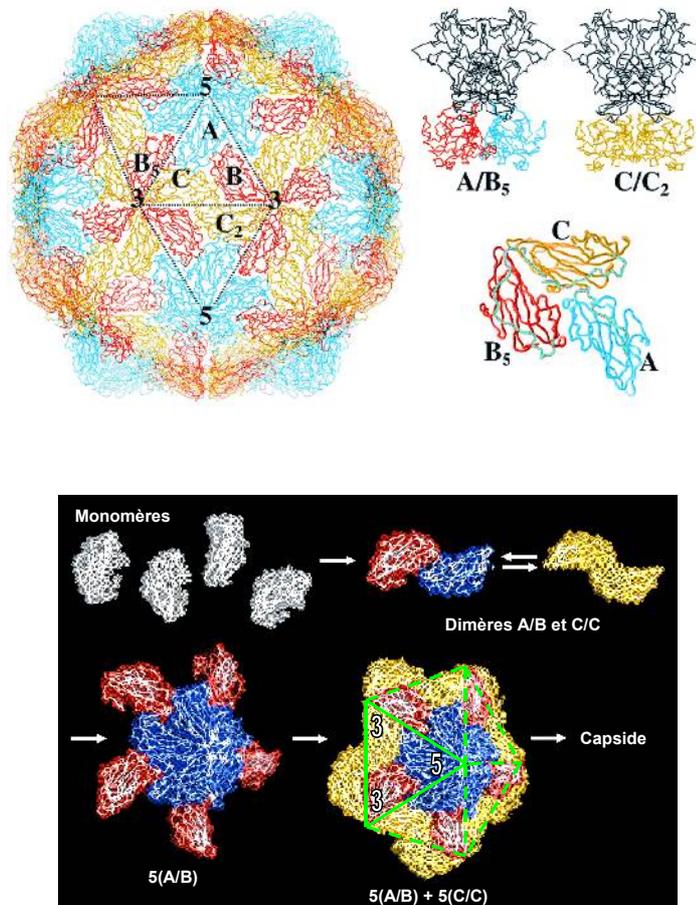
La comparaison des séquences nucléiques et peptidiques de différentes souches de NoV montre que le domaine S est relativement bien conservé, alors que le domaine P1 ne l'est que modérément et que le domaine P2 montre une très grande variabilité (Chakravarty et coll. 2005). Ces différences peuvent être expliquées par le niveau d'exposition à l'hôte des différents domaines et sous-domaines de la capsid. En effet, plus une région protéique est exposée, plus elle est soumise à la pression de sélection positive induite par l'immunité de l'hôte. C'est notamment ce domaine qui porte les caractères propres à chaque souche de Norovirus.

### **b - Structure de la capsid**

La capsid présente une symétrie icosaédrique T=3. Elle est composée de 180 protéines VP1, organisées en 90 dimères. Il en résulte la formation de 32 capsomères, caractérisés par des protubérances sur les axes 2 de l'icosaèdre et de larges dépressions les axes 3 et 5, leur donnant une allure générale caractéristique en forme de « calice » (Figure 22). La partie interne de la capsid est formée par une coque contiguë de domaine S et présente un rayon de 100 à 145 Å. Les protubérances sur les axes 2 de l'icosaèdre vont augmenter ce rayon jusqu'à environ 190 Å.

La formation des dimères puis des capsomères correspond à un auto-assemblage de la protéine VP1 ne nécessitant ni énergie, ni l'intervention de l'ARN génomique ou d'autres protéines. Pour former la structure icosaédrique T=3, la protéine de capsid doit adopter trois positions quasi-équivalentes dénommées A, B et C. Leurs associations peuvent générer deux types de dimères, dits A/B et C/C, qui se différencient

respectivement par une structure courbe ou plane formée par la juxtaposition de leurs domaines S (Figure 23).



**Figure 23 : Représentation et organisation des dimères A/B et C/C formant la capsid des NoV.**

*Le fenêtré au fond noir illustre la méthode adoptée par le virus pour initier la formation de sa capsid. (D'après Prasad et coll. 1999)*

Le mécanisme d'assemblage de la capsid proposé par Prasad débute par la formation d'un pentamère de dimères A/B (Figure 23). Ce pentamère forme alors un des sommets de l'icosaèdre sur l'axe 5 au niveau du point de convergence des sous-unités A des dimères A/B. Ensuite, l'ajout de cinq dimères C/C complète la formation des cinq faces de l'icosaèdre initiées par le pentamère. Les angles opposés au sommet sur l'axe 5 initient alors la formation des cinq sommets sur l'axe 3 de l'icosaèdre.

## **1.4. Interaction avec les HBGA**

### **a - RHDV et HBGA**

L'une des premières caractéristiques biologiques à avoir été mise en évidence pour le RHDV est sa capacité à agglutiner les érythrocytes humains (Pu et coll. 1985). Cette propriété est à la base du premier test diagnostique efficace pour la détection de la présence de virions dans des homogénats d'organes de lapins suspectés d'être infectés (Pu et coll. 1985; Capucci et coll. 1991). Pendant une dizaine d'année, la compréhension du mécanisme d'hémagglutination par le virus n'a pas été approfondie et c'est en 1995 que les premières données concernant un récepteur viral porté par les érythrocytes humains commencent à émerger.

Cette année là sont publiés des résultats préliminaires qui renseignent pour la première fois sur la nature du composé reconnu par le virus (Ruvoën-Clouet et coll. 1995). Il y est notamment décrit que le RHDV est capable d'agglutiner les globules rouges humains et ceci indépendamment de leur phénotype A, B ou O. Les tests d'inhibition de l'hémagglutination réalisés avec les glycoprotéines et glycolipides érythrocytaires humains purifiés ont notamment permis de mettre en évidence que la fixation du RHDV est fortement inhibée par les glycolipides richement glycosylés (polyglycosylcéramides), mais également assez bien inhibée par les sucres neutres portés par les glycoprotéines des globules rouges. Dans un même temps, il est mis en évidence que le virus n'est pas capable d'agglutiner les érythrocytes provenant du sang fœtal ou de cordon ombilical. Cette observation est peut être le résultat le plus important de cette étude. En effet, elle met en évidence que le composé reconnu par le virus n'apparaît que tardivement au cours du développement humain. Cette observation est également en accord avec les observations faites *in naturo*, où les jeunes lapereaux de moins de deux mois ne semblent pas être sensible à la maladie (Z. J. Xu et coll. 1989).

C'est en 2000, que le ligand érythrocytaire humain du RHDV est caractérisé. Il est montré qu'il correspond à l'HBGA H de type 2 (Ruvoën-Clouet et coll. 2000). Cette hypothèse d'un récepteur appartenant aux HBGA du système ABH avait pourtant été jugée peu probable du fait que l'hémagglutination par le virus n'était pas dépendante d'un phénotype ABH particulier. Dans cette nouvelle étude, c'est la non-hémagglutination des globules rouges de phénotype Bombay qui met clairement en évidence la reconnaissance spécifique d'un glycanne de type A, B ou H par le RHDV. En effet, le phénotype Bombay est synonyme de non-présentation des antigènes A,B et H (tous types) à la surface des érythrocytes. Cette caractéristique est partagée avec les cellules provenant du sang de cordon ombilical et explique les résultats obtenus en 1995. En effet, le niveau d'expression des HBGA ABH n'est pas le même au cours des différents

stades du développement. Le fœtus et le nouveau né n'expriment que très peu ces structures (Fukuda et coll. 1983) ; leur expression va augmenter de manière croissante au cours des mois qui suivent la naissance. Il a également été montré dans ces travaux que les HBGA de type 2 sont les seuls à être reconnus par le virus. En effet, la pré-incubation du virus avec des glycanes synthétiques A ou H type 2 immobilisés sur des billes de silicate inhibe complètement la reconnaissance des érythrocytes présentant ces mêmes HBGA. Des résultats similaires ont été obtenus en co-incubant les érythrocytes avec la lectine UEA-I (*Ulex europaeus* agglutinin I) reconnaissant spécifiquement l'HBGA H type 2. Il a de plus été montré que la fixation du virus au niveau de l'épithélium intestinal et trachéal de lapin adulte est corrélée à l'expression de ces HBGA. La même expérience réalisée sur l'épithélium de lapin juvénile montre que celui-ci exprime peu l'HBGA H type 2 et que cette faible expression est corrélée à une faible fixation du virus. Cette observation pourrait expliquer pourquoi les jeunes lapereaux semblent résistants à l'infection par le RHDV. En conclusion de toutes ces observations, les auteurs ont défini que la structure minimum reconnue par le virus est le fucose porté par le galactose du précurseur type 2.

Très récemment, une étude portant sur la reconnaissance de l'HBGA H type 2 par le RHDV a été menée d'un point de vue atomique en utilisant la technique de résonance magnétique nucléaire (Rademacher et coll. 2008). Elle a permis de confirmer que le fucose de l'HBGA H type 2 est la molécule reconnue préférentiellement par le virus.

### ***b - Norovirus et HBGA***

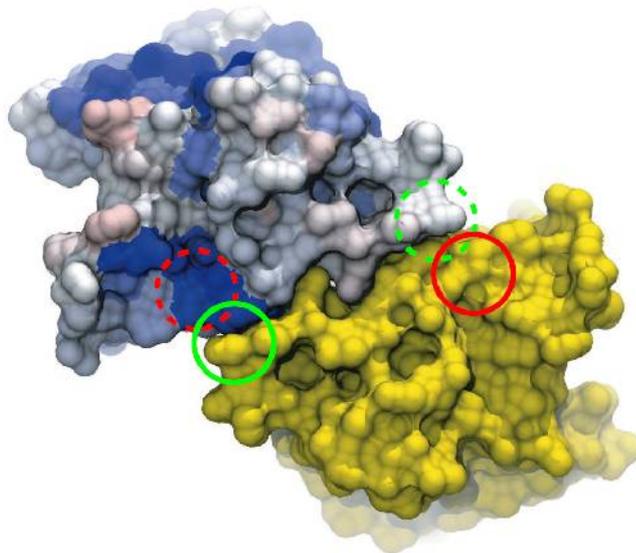
La recherche d'un HBGA potentiellement reconnu par les NoV a été stimulée par la découverte de la nature glycanique du récepteur cellulaire au RHDV et au fait qu'ils appartiennent tous les deux à la même famille de virus. Dans cette optique, la fixation des VLP du NV a été testée sur des coupes histologiques de la jonction gastroduodénale et des mucines salivaires humaines, phénotypées pour les HBGA A, B, H et Lewis (Marionneau et coll. 2002). Il a été ainsi mis en évidence que les VLP se fixent sur les HBGA H et A type 1 et 3/4. Il a été démontré que cette fixation est spécifique en l'inhibant par une pré-incubation des VLP avec des trisaccharides synthétiques H type 1 ou 3, mais également en pré-incubant les coupes histologiques avec des anticorps monoclonaux anti-H type 1 ou 3/4. Cette spécificité a également été mise en évidence en microscopie confocale en évaluant la fixation des VLP de NV sur des cellules transfectées pour exprimer une  $\alpha$ 1,2 fucosyltransférase, enzyme à l'origine de la synthèse de l'antigène H. Il a alors été prouvé que les VLP ne se fixent à la membrane des cellules que lorsque celles-ci sont transfectées avec le gène de fucosyltransférase, mais aussi que cette fixation s'accompagne d'une internalisation des VLP. Dans les mois qui suivirent cette découverte, deux autres équipes ont abouti à la même conclusion que l'HBGA H

type 1 ou 3 représente un candidat sérieux au titre de récepteur cellulaire du NV (Harrington et coll. 2002; A M Hutson et coll. 2003). Ces hypothèses ont été validées *in vivo* en inoculant le NV à des volontaires dont les phénotypes ABH, Sécréteur et Lewis étaient connus (Lindesmith et coll. 2003). Cette expérience a mis en évidence que parmi les individus inoculés avec le NV, 50 à 69% de ceux présentant le phénotype sécréteur déclarent la pathologie alors qu'aucun symptôme de la maladie n'est détecté chez les volontaires de phénotype non-sécréteur. Ces données apportent pour la première fois la preuve que le NV a besoin que l'hôte exprime l'antigène H à la surface de ses cellules épithéliales intestinales pour pouvoir l'infecter, ce qui revient à dire que les individus de phénotype non sécréteur sont génétiquement résistants à l'infection par le NV, du fait de leur homozygotie pour des allèles *FUT2* non-fonctionnels.

Les essais de fixation des NoV sur les HBGA ont pourtant très vite montré certaines subtilités. En effet, tous les NoV ne reconnaissent pas les mêmes HBGA. Si par exemple le NV (génogroupe I) reconnaît préférentiellement les HBGA H et A type 1, le virus Snow Mountain (génogroupe II) montre quant à lui une affinité restreinte à l'HBGA B type 1 (Harrington et coll. 2002). Cette observation a été faite pour d'autres souches comme le virus MOH qui se fixe sur les mucines salivaires des individus « sécréteur » de phénotype A ou B, le virus VA207 qui lui semble ne reconnaître que l'HBGA Le<sup>a</sup>, ou encore le virus VA387 qui se fixe sur les mucines de tous les individus « sécréteur », indépendamment de leur phénotype ABH (P Huang et coll. 2003). Le fait troublant dans ces observations est que même si deux souches de NoV appartiennent à un même génogroupe, voire à un même cluster de virus, elles peuvent tout aussi bien présenter des profils de fixation sur les HBGA différents, comme par exemple les virus Boxer et VA115 (Figure 16 p.40) (P Huang et coll. 2005; Le Pendu et coll. 2006). La très grande variabilité génétique de la capsidie des NoV peut être à l'origine de ces modulations, notamment le polymorphisme observé dans le domaine P2. Ce dernier semble en effet contenir la poche de liaison au récepteur cellulaire car il est possible d'inhiber la reconnaissance du récepteur des VLP de NV en les pré-incubant avec un anticorps monoclonal spécifique de la région comprise entre les acides aminés 300 et 384 de la capsidie (L. J. White et coll. 1996). De plus, des expériences de mutagenèse dirigée réalisées dans cette région (modification de la thréonine 338) ont montré une abolition complète de la reconnaissance des antigènes A, B et H par les VLP du virus VA387 (M Tan et coll. 2003).

Une réponse concrète concernant la localisation du domaine de liaison au récepteur a été apportée récemment grâce à la co-cristallisation des trisaccharides synthétiques A et B sur les dimères de protéine de capsidie du virus VA387 (S Cao et coll. 2007), mais également par la co-cristallisation du pentasaccharide H type 1 et du trisaccharide A sur les dimères de la protéine de capsidie du NV (Bu et coll. 2008; J. Choi

et coll. 2008). Il a notamment été mis en évidence pour la souche VA387 que l'HBGA vient s'insérer au sommet du domaine P2 de la capsid, à l'interface des protéines du dimère (Figure 24). Pour cette souche, la formation du dimère n'a donc pas qu'une fonction structurale mais permet également la création du domaine de liaison aux HBGA. Il a également été montré que les acides aminés Thr344, Arg345, Asp374, et Gly442 sont essentiels au virus VA387 pour la formation des liaisons hydrogène permettant l'interaction glycanne/protéine. Dans le cas du NV, cette poche est différente mais reste proche de celle du VA387 (Figure 24). Pour cette souche, ce sont les acides aminés Asp327, His329, Trp375, Ser377 et Pro378 qui sont essentiels à la formation des liaisons hydrogènes et hydrophobes permettant la fixation de la protéine de capsid sur les antigènes A et H. Les acides aminés de ces différents domaines de liaison au récepteur sont très conservés à l'intérieur de leur génogroupe respectif, alors qu'ils ne le sont pas lorsque l'on compare les séquences peptidiques des NoV du GI et du GII sur ces positions. Il semble donc que chacun des génogroupes subit une pression de sélection négative afin de maintenir les acides aminés essentiels à la fixation du virus sur les HBGA. La localisation différente des poches de liaison signe une évolution convergente des NoV pour fixer les HBGA, mais avec la mise en place de solutions différentes dans le choix du domaine de liaison au récepteur. Ceci pourrait expliquer les différents profils de fixation observés pour ces souches.



**Figure 24 : Localisation moléculaire des domaines de liaison aux glycanes du virus VA387 et du NV.**

*Est modélisé ici un dimère de la protéine de capsid du NV. Sur cette vue de dessus, les domaines de liaison aux HBGA sont représentés en Rouge pour le NV et en vert pour le virus VA387. (D'après J. Choi et coll. 2008)*

## 2. SRAS-CoV ET HBGA

### 2.1. Taxonomie

Les Coronavirus (CoV) font partie de la famille *Coronaviridae*. Ce nom a été attribué à cette famille de virus du fait de leur aspect caractéristique en couronne (*corona* en latin) lorsqu'ils sont observés par microscopie électronique (Tyrrel et coll. 1968). Ils font partie de l'ordre des *Nidoviriales* et en 2005 cette famille a été divisée en deux sous-familles : les Coronavirus et les Torovirus.

Les CoV sont des virus enveloppés présentant un génome composé d'un brin ARN monocaténaire de polarité positive, polyadénylée en 3'. Infectant l'homme ainsi que d'autres espèces mammifères et aviaires, ils sont responsables la plupart du temps d'infections respiratoires ou intestinales, et plus rarement d'atteintes neurologiques et hépatiques. Ils comptent à ce jour 36 membres divisés en trois génogroupes établis d'après la comparaison des séquences nucléiques codant l'hélicase (V. C. C. Cheng et coll. 2007) (Figure 25) :

- Le génogroupe 1 comprend notamment le CoV humain NL63, le CoV canin, le virus de la péritonite infectieuse féline, le virus de la gastro-entérite transmissible du porc et le virus respiratoire porcin.
- Le génogroupe 2 est constitué entre autres du CoV bovin, du virus des hépatites murines, du virus de la silodacryonite du rat et du CoV humain OC43.
- Le génogroupe 3 compte des virus restreints exclusivement aux oiseaux comme le virus de la bronchite infectieuse aviaire et un CoV du dindon.

Jusqu'à l'hiver 2002, les souches de CoV infectant l'homme étaient dans la majorité des cas responsables de maladies respiratoires généralement bénignes (symptômes de rhume). Mais à l'automne 2002, un nouveau CoV humain extrêmement virulent a été responsable de la première grave pandémie du XXI<sup>e</sup> siècle : le CoV du syndrome respiratoire aigu sévère (SRAS-CoV). Le foyer originel de l'épidémie a été localisé dans la province chinoise du Guangdong, où les premiers cas de SRAS-CoV ont été rapportés le 16 novembre 2002. Ce n'est que le 14 février 2003 que l'épidémie chinoise de SRAS est annoncée puis gérée par l'organisation mondiale pour la santé (WHO 2003). De par l'ampleur de la crise, celle-ci déclenche une alerte mondiale le 12 mars 2003. Entre novembre 2002 et juillet 2003, le SRAS-CoV a infecté 8096 personnes dans 29 pays ou régions. Parmi ces malades, 774 sont décédés du fait de la grande détresse respiratoire engendrée par l'infection (Tableau 5).

Pays ou Régions touchés	Nombre de cas cumulés			Age moyen (ages extrêmes)	Nombre de morts <sup>a</sup>	Taux de mortalité (%)	Nombre de cas importés (%)	Nombre de cas dans le personnel soignant (%)	Date du premier cas probable	Date du dernier cas probable
	Femme	Homme	Total							
Australie	4	2	6	15 (1-45)	0	0	6 (100)	0 (0)	28/02/03	01/04/03
Canada	151	100	251	49 (1-98)	43	17	5 (2)	109 (43)	23/02/03	12/06/03
Chine	2674	2807	5327 <sup>b</sup>	Pas de données	349	7	Origine de l'épidémie	1002 (19)	16/11/02	03/06/03
Chine (région administrative d'Hong Kong)	977	778	1755	40 (0-100)	299	17	Origine de l'épidémie	386 (22)	15/02/03	31/05/03
Chine (région administrative de Macao)	0	1	1	26	0	0	1 (100)	0 (0)	05/05/03	05/05/03
China, Taïwan	218	128	346 <sup>c</sup>	42 (0-83)	37	11	21 (6)	88 (20)	25/02/03	15/06/03
France	1	6	7	49 (26 - 61)	1	14	7 (100)	2 (29) <sup>d</sup>	21/03/03	03/05/03
Allemagne	4	5	9	44 (4-73)	0	0	9 (100)	1 (11)	09/03/03	06/05/03
Inde	0	3	3	25 (25-30)	0	0	3 (100)	0 (0)	25/04/03	06/05/03
Indonésie	0	2	2	56 (47-65)	0	0	2 (100)	0 (0)	06/04/03	17/04/03
Italie	1	3	4	30,5 (25-54)	0	0	4 (100)	0 (0)	12/03/03	20/04/03
Koweït	1	0	1	50	0	0	1 (100)	0 (0)	09/04/03	09/04/03
Malaisie	1	4	5	30 (26-84)	2	40	5 (100)	0 (0)	14/03/03	22/04/03
Mongolie	8	1	9	32 (17-63)	0	0	8 (89)	0 (0)	31/03/03	06/05/03
Nouvelle Zélande	1	0	1	87	0	0	1 (100)	0 (0)	20/04/03	20/04/03
Philippines	8	6	14	41 (29-73)	2	14	7 (50)	4 (29)	25/02/03	05/05/03
République d'Irlande	0	1	1	56	0	0	1 (100)	0 (0)	27/02/03	27/02/03
République de Corée	0	3	3	40 (20-80)	0	0	3 (100)	0 (0)	25/04/03	10/05/03
Roumanie	0	1	1	52	0	0	1 (100)	0 (0)	19/03/03	19/03/03
Fédération Russe	0	1	1	25	0	0	Pas de données	0 (0)	05/05/03	05/05/03
Singapour	161	77	238	35 (1-90)	33	14	8 (3)	97 (41)	25/02/03	05/05/03
Afrique du Sud	0	1	1	62	1	100	1 (100)	0 (0)	03/04/03	03/04/03
Espagne	0	1	1	33	0	0	1 (100)	0 (0)	28/03/03	26/03/03
Suède	3	2	5	43 (33-55)	0	0	5 (100)	0 (0)	28/03/03	23/04/03
Suisse	0	1	1	35	0	0	1 (100)	0 (0)	09/03/03	09/03/03
Thaïlande	5	4	9	42 (2-79)	2	22	9 (100)	1 (11) <sup>d</sup>	11/03/03	27/05/03
Royaume Uni	2	2	4	59 (28-74)	0	0	4 (100)	0 (0)	01/03/03	01/04/03
Etats Unis d'Amérique	13	14	27	36 (0-83)	0	0	27 (100)	0 (0)	24/02/03	13/07/03 <sup>e</sup>
Viet Nam	39	24	63	43 (20-76)	5	8	1 (2)	36 (57)	23/02/03	14/04/03
<b>Total</b>			<b>8096</b>		<b>774</b>	<b>9,6</b>	<b>142</b>	<b>1706</b>		

a. Inclus uniquement les cas où la mort est attribuée au SARS.

b. Classification non connue pour 46 cas.

c. Depuis le 11 juillet 2003, 325 cas ont été annulés pour Taïwan. Les informations étaient insuffisantes ou incomplètes pour 135 de ces cas, dont 101 correspondaient à des décès.

d. Inclus le personnel soignant qui a acquis la maladie dans un autre pays.

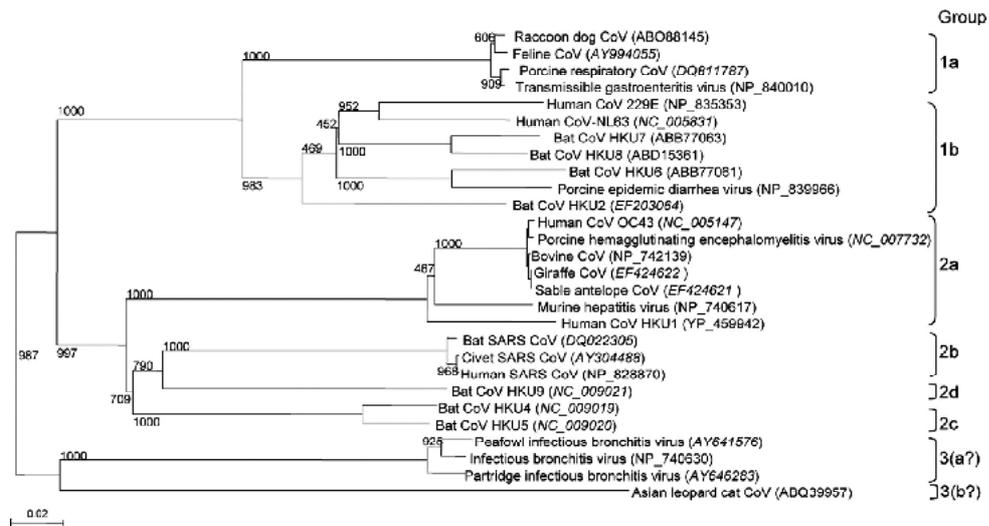
e. A cause de différences dans la définition des cas de SARS, les USA ont rapporté plusieurs cas probables de début de maladie après le 5 juillet 2003.

**Tableau 5 : Données épidémiologiques mondiales de l'épidémie de SRAS-CoV de 2002/2003.**

(Données WHO)

Le SRAS-CoV humain est très proche phylogénétiquement des SRAS-CoV de la chauve-souris et du SRAS-CoV de la civette. Ils constituent ensemble le génogroupe 2b des CoV (Figure 25) (S. K. P. Lau et coll. 2005; Ren et coll. 2006). Le SRAS-CoV étant très différent des autres CoV humains, le passage de la barrière d'espèce vers l'homme par l'une de ces souches animales est très fortement suspectée (Y. Guan et coll. 2003; S. K. P. Lau et coll. 2005).

**Figure 25 : Arbre phylogénétique établi suite à la comparaison de 28 séquences complètes codant l'hélicase des CoV.**



(D'après Cheng et coll. 2007).

## **2.2. Génome viral et protéines associées**

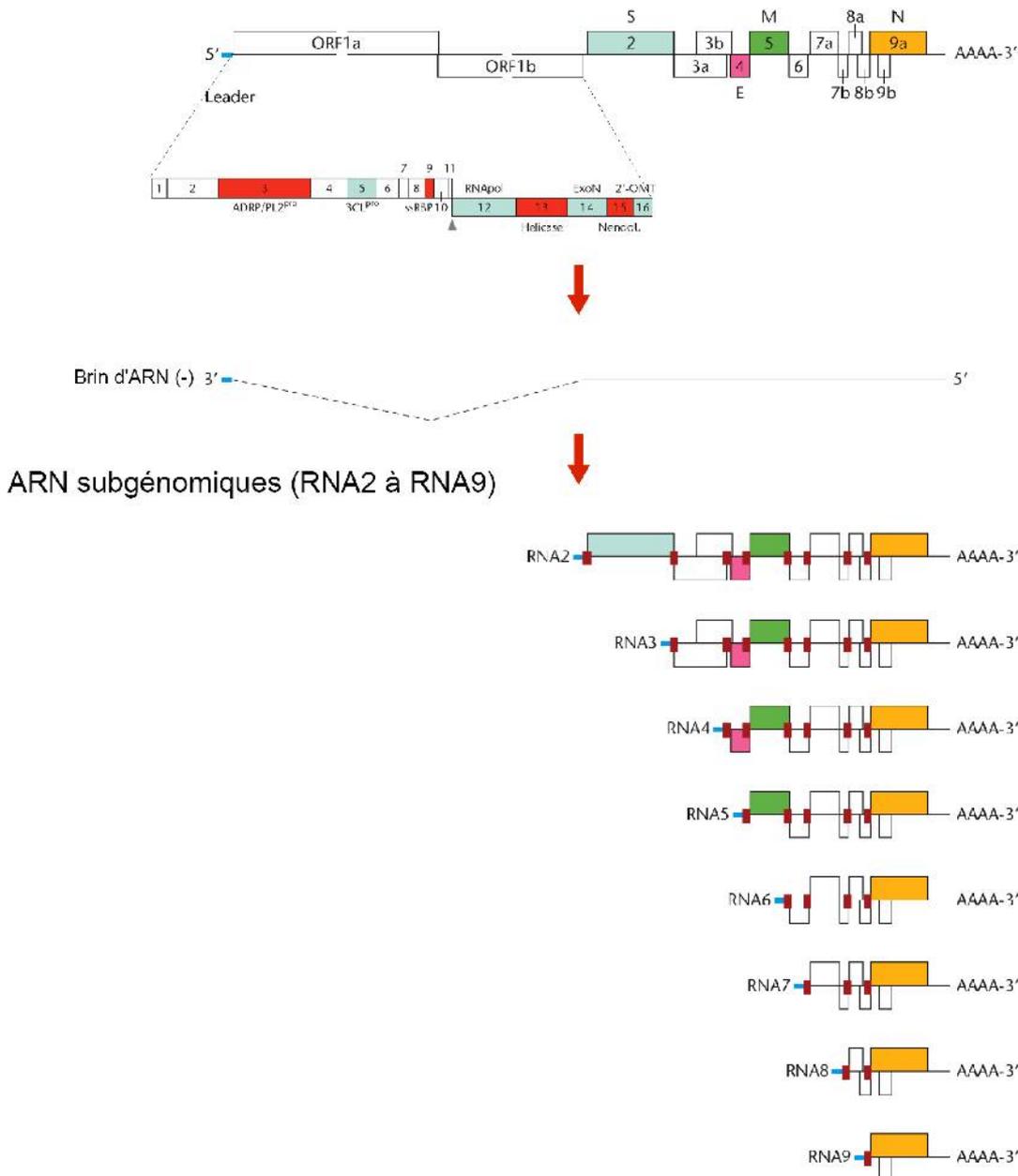
Le génome du SRAS-CoV est un ARN simple brin de polarité positive composé de 29727 nucléotides, queue polyadénylée exclue. Ses 265 premiers nucléotides (3') et 342 derniers nucléotides (5') constituent des régions non-traduites (Marra et coll. 2003; Rota et coll. 2003). Son génome compte un total de 14 ORF (Figure 26).

Les deux premiers ORF (ORF 1a, ORF 1b) représentent environ les deux tiers du génome (21kb). Ils codent pour les polyprotéines pp1a (4382 résidus) et pp1ab (7073 résidus). La polyprotéine pp1ab est une version étendue de pp1a. Elle est le résultat d'un décalage d'un nucléotide (-1) lors de la traduction de l'ORF1a. Ce décalage a lieu un peu en amont du codon stop de l'ORF1a. La protéine pp1ab donne après clivage 16 protéines non-structurales qui forment ensemble la batterie de réplicases du virus. Parmi celles-ci sont retrouvées une ADP-ribose-1-phosphatase/Protéase 2 papaïne-like (nsp3), une protéase 3C-like (nsp5), une protéine se liant à l'ARN simple brin (nsp9), une ARN polymérase ARN dépendante (nsp12), une hélicase (nsp13), une exoribonucléase 5'>3' (nsp14), une endoribonucléase Uridylate-spécifique (nsp15) et une possible 2'-O-ribose méthyltransférase (nsp16) (V. C. C. Cheng et coll. 2007). Les autres protéines issues du clivage de pp1ab n'ont pas encore de rôle bien défini mais certaines semblent favoriser la réplication du virus (R. L. Graham et coll. 2005; Kamitani et coll. 2006). D'autres ORF codent également des protéines non-structurales (ORF 3b, 6, 8a, 8b et 9b). Cependant

leurs fonctions sont méconnues du fait de l'absence de ces protéines chez les autres CoV ou virus en général.

## SARS

### ARN génomique (RNA1)



**Figure 26 : Organisation des 14 ORF du génome du SRAS-CoV (RNA1) et de ses 8 ARN subgénomiques.**

(D'après Dandekar et coll. 2006)

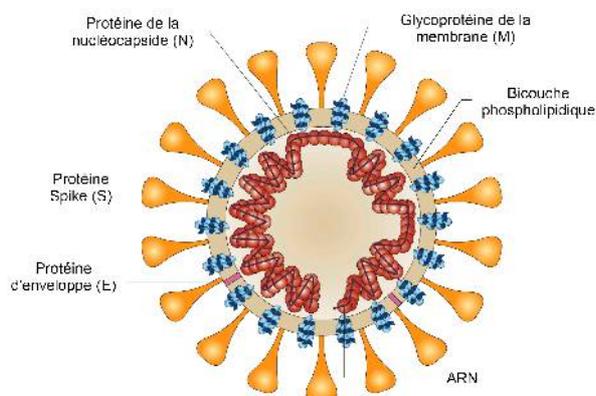
Les protéines structurales du SRAS-CoV sont au nombre de sept et sont codées par les ORF 2, 3a, 4, 5, 7a, 7b et 9a. Les protéines majeures de la capsidie sont la protéine S (Spicule ou « Spike » = ORF 2), la protéine E (Enveloppe = ORF 4), la protéine M (Membranaire = ORF 5) et la protéine N (Nucléocapside = ORF 9a). Les

autres ORF codent pour des protéines mineures de la capsidie dont les rôles ne sont pas encore bien définis (N Ito et coll. 2005; C Huang et coll. 2006; Schaecher et coll. 2007). Néanmoins, certaines semblent jouer un rôle important lors du cycle de réplication du virus (Akerström et coll. 2007).

Le SRAS-CoV présente également la caractéristique de posséder huit ARN subgénomiques différents (Figure 26). Ceux-ci portent une coiffe et chacun d'eux code pour l'ORF situé au plus près de son extrémité 5'. Ils ont également la caractéristique de tous posséder à leur extrémité 5' une séquence « leader » de 72 nucléotides correspondant à la séquence « leader » de l'ARN génomique. Le mécanisme de la synthèse discontinue permettant l'ajout de ces 72 nucléotides n'est pas encore connu. Toutefois, plusieurs études laissent penser qu'elle aurait lieu lors de la synthèse du brin d'ARN négatif par la batterie de réplicases du virus (S. G. Sawicki et coll. 2005).

### **2.3. Morphologie et structure**

Le SRAS-CoV est un virus enveloppé qui apparaît en microscopie électronique avec une forme sphérique plus ou moins régulière. Ses virions ont généralement un diamètre compris entre 80 et 140 nm. L'ARN génomique des CoV a la particularité d'être complexé avec une protéine virale : la protéine de nucléocapside. Cette protéine très basique possède à ses extrémités N et C terminales des domaines lui permettant de reconnaître et de se lier spécifiquement à l'ARN viral (Hsieh et coll. 2005). Son domaine C terminal lui permet également de s'oligomériser pour former une nucléocapside de forme hélicoïdale protégeant l'ARN viral (Surjit et coll. 2008). Ce complexe nucléocapside/ARN se localise à proximité de l'enveloppe à laquelle il est relié par le biais d'interactions entre les protéines de la nucléocapside et la glycoprotéine de membrane intégrée dans l'enveloppe du virus (R He et coll. 2004) (Figure 27).



**Figure 27 : Structure schématique d'un virion de SRAS-CoV.**

(D'après Peiris et coll. 2004)

L'enveloppe des CoV est constituée d'une bicouche phospholipidique ayant pour origine la membrane du réticulum endoplasmique (RE) de la cellule hôte. Le virion s'en entoure lors du bourgeonnement du complexe nucléocapside/ARN vers la lumière du RE. Au cours de la réplication du SRAS-CoV, la membrane de ce compartiment cellulaire se voit intégrer trois protéines virales :

- La protéine M. C'est la protéine structurale la mieux conservée chez les CoV (Lai et coll. 1997). Elle est essentielle au bon assemblage du virion en interagissant avec le complexe nucléocapside/ARN cytosolique qui est ensuite internalisé dans le RE.
- La protéine E. Petite protéine membranaire, elle forme des canaux ioniques au travers de la membrane du RE de la cellule hôte. Ceux-ci pourraient être impliqués dans les mécanismes de bourgeonnement et de relargage du virus (L. Wilson et coll. 2004).
- La protéine S. C'est elle qui donne la forme caractéristique en « couronne » au virus. Elle est impliquée dans la reconnaissance et la liaison au récepteur cellulaire

Il a également été montré lors de la production de VLP de SRAS-CoV que les protéines virales M et E jouent un rôle critique dans la formation et le maintien de la structure des particules virales (Y. Ho et coll. 2004).

Contrairement aux autres CoV du génogroupe 2, le SRAS-CoV a la caractéristique de ne pas posséder d'hémagglutinine estérase membranaire.

## **2.4. SRAS-CoV : interactions virus/récepteurs**

Le SRAS-CoV peut utiliser différentes molécules et stratégies pour infecter les pneumocytes de l'homme. Il a notamment été mis en évidence qu'il est reconnu et se lie aux lectines DC-SIGN, L-SIGN et LSECtin (Z. Yang et coll. 2004; Jeffers et coll. 2004; Gramberg et coll. 2005). Ces différentes lectines permettent la facilitation de l'infection par le mécanisme de *trans* infection. La protéine L-SIGN a de plus été identifiée comme un récepteur potentiel de *cis* infection du SRAS-CoV (Jeffers et coll. 2004). Elle est en effet exprimée à la surface des cellules alvéolaire de type II et des cellules endothéliales du poumon, cellules cibles du virus.

Toutefois, le récepteur cellulaire indispensable à l'infection par le SRAS-CoV semble être une autre protéine : l'enzyme 2 de conversion de l'angiotensine (ACE2). En effet, la fixation de la protéine S sur l'ACE2 est nécessaire à l'infection de la cellule hôte par le SRAS-CoV et au développement de la cytopathie (Kuba et coll. 2005; C-T K Tseng et coll. 2007). Il a en effet été découvert qu'elle permet non seulement la reconnaissance

de la cellule cible, mais également la fusion des membranes virus/cellule à l'origine de l'entrée du virus dans la cellule permissive (W Li et coll. 2003).

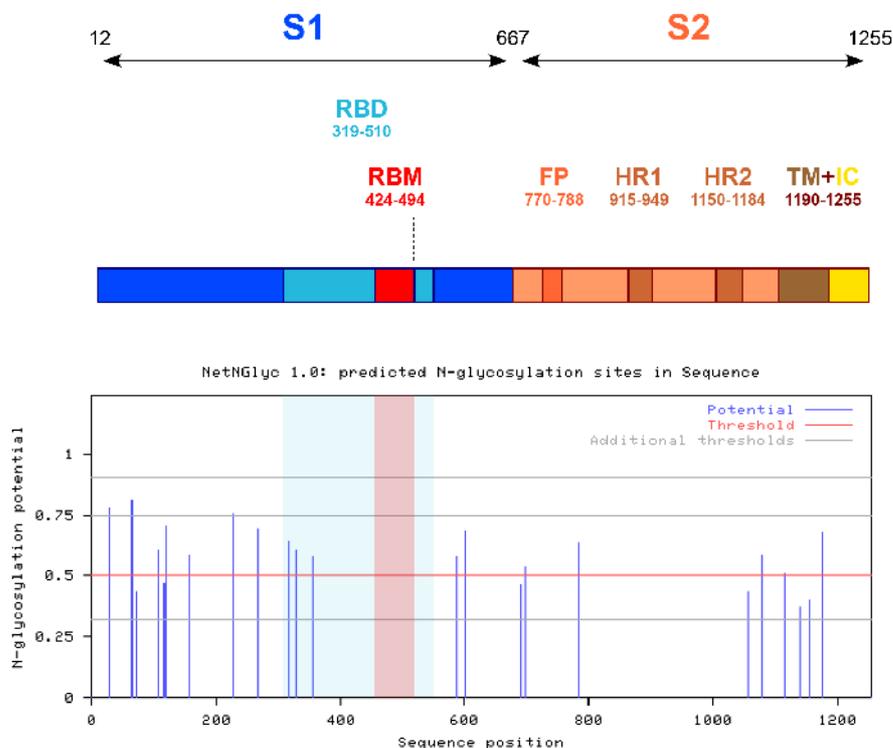
La protéine ACE2 est une protéase qui a également été identifiée comme un récepteur du HCoV-NL63, CoV humain du génogroupe I (Hofmann et coll. 2005). Toutefois, son activité protéolytique n'est pas nécessaire à l'entrée du SRAS-CoV dans la cellule (W Li et coll. 2005). L'ACE2 est très conservée au cours de l'évolution. Elle est exprimée dans de multiples tissus et organes. On la retrouve notamment au niveau du poumon, du cœur, du rein et sur les cellules du tractus gastro-intestinal (Hamming et coll. 2004). Son expression tissulaire est en accord avec le tropisme du virus. En effet, de grandes quantités d'ARN viral sont retrouvées chez les individus infectés au niveau du poumon et de l'intestin, et dans de plus faibles proportions au niveau du rein, de la rate et du foie (Peiris et coll. 2004). Toutefois, les effets cytopathiques sont majoritairement observés au niveau du poumon, ce qui veut dire que le tropisme du virus n'est peut être pas dû qu'à l'expression de la protéine ACE2. L'expression de la lectine L-SIGN sur les pneumocytes pourrait agir de concours avec l'ACE2 dans la localisation de l'infection.

La protéine S du SRAS-CoV est une glycoprotéine de type I composée de 1255 acides aminés. Sa séquence peptidique ne présente que 20 à 27% de similitude avec celles des autres CoV (Rota et coll. 2003). Elle compte 23 sites potentiels de N-glycosylation dont 17 sont considérés comme privilégiés (Krokhin et coll. 2003). Structurellement, elle peut être décomposée en deux domaines : S1 et S2 (Figure 28).

Le domaine S2 est présent du côté C-terminal de la protéine virale. C'est ce domaine qui est impliqué dans le mécanisme de fusion des membranes virus/cellule. Il comprend pour cela deux régions de sept acides aminés répétés (heptade) (HR1 et HR2) ainsi qu'un peptide de fusion (FP) de 18 acides aminés. Lors de la fixation de la protéine S sur l'ACE2 s'opère un changement de conformation dans les heptades qui induit la formation du cœur de fusion et le rapprochement des membranes virus/cellule (Tripet et coll. 2004). Le domaine S2 comprend également sur ses 52 derniers acides aminés la région transmembranaire (TM) de la protéine S, suivie d'une queue cytoplasmique (IC) très riche en cystéines.

Le domaine S1 est localisé du côté N-terminal de la protéine S. Il participe à la liaison sur l'ACE2. Sur S1, une région de 193 acides aminés (résidus 319 à 510) a été définie comme étant la séquence minimale pour que la protéine S puisse se lier sur l'ACE2 (S K Wong et coll. 2004). Cette région est appelée domaine de liaison au récepteur (RBD). Li et collaborateurs ont ensuite identifié un motif de liaison au récepteur (RBM) dans ce RBD en co-cristallisant celui-ci avec des molécules d'ACE2 solubles (résidus 19 à 615) (F Li et coll. 2005). Ils ont ainsi mis en évidence que le RBM correspond au peptide compris entre les positions 424 et 494 de la protéine S. De

nombreuses études ont également mis en évidence que le RBD de la protéine S est l'épitope majoritairement reconnu par les anticorps anti-SRAS-CoV de l'hôte et que ceux-ci exercent un fort effet neutralisant en empêchant la fixation et la réplication du virus (Z Chen et coll. 2005; Y He et coll. 2004, 2005; Traggiai et coll. 2004; van den Brink et coll. 2005; S Wang et coll. 2005; T. Zhou et coll. 2004).



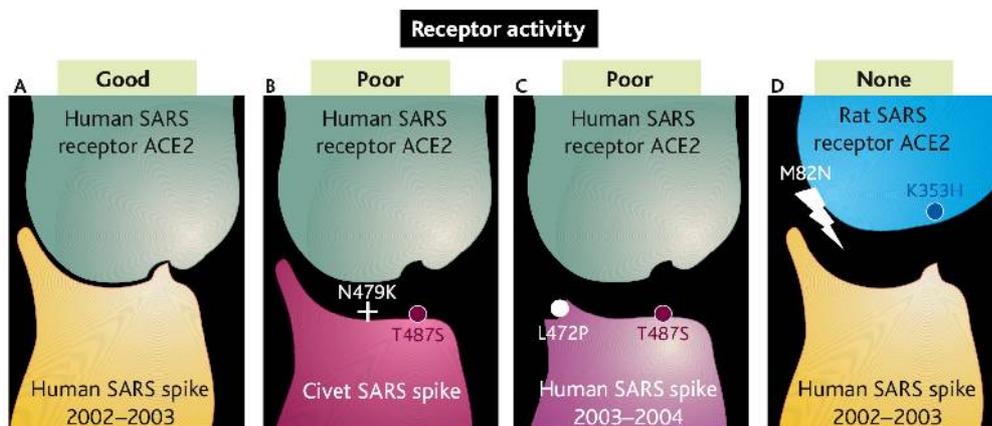
**Figure 28 : Représentation des différents domaines de la protéine S et localisation de ses différents sites potentiels de N-glycosylation.**

(D'après F Li et coll. 2005)

De façon surprenante, la reconnaissance de l'ACE2 humaine par la protéine S du SRAS-CoV humain ne tolère que très peu de modifications dans leurs séquences protéiques respectives. Deux études ont notamment comparé la fixation de la protéine S du SRAS-CoV humain sur l'ACE2 de la souris, du rat, de la civette et de l'homme (W Li et coll. 2004, 2005). Elles ont mis en évidence que la protéine ACE2 de la souris et de la civette est beaucoup moins efficacement reconnue par la protéine virale et que la protéine S du SRAS-CoV humain est incapable de se fixer sur la protéine ACE2 de rat.

Dans une autre étude, la fixation de la protéine S du SRAS-CoV humain a été comparée à celle du CoV de la civette. Leurs RBD respectives ne présentent que quatre positions différentes qui sont malgré tout suffisantes pour induire un différentiel de 1000 entre leurs capacités de fixation sur l'ACE2 (W Li et coll. 2005). Parmi ces variations, deux sont localisées dans le RBM (substitution N479K et T487S) et sont critiques pour l'établissement de la liaison virus/récepteur (Qu et coll. 2005). La position 487 du RBM

de la civette est de plus identique à celle de la souche de SRAS-CoV humain émergent en 2003/2004. Cette nouvelle souche humaine ne reconnaît l'ACE2 qu'avec une affinité faible et sa non-transmissibilité d'homme à homme pourrait être en partie due à cette position partagée avec le virus d'origine animale.



**Figure 29 : Comparaison de la fixation de la protéine S de différentes souches de SRAS-CoV sur l'ACE2 d'origine humaine ou de rat.**

*Les acides aminés identifiés sur les protéines virales et l'ACE2 de rat sont les acides aminés clés permettant ou non l'établissement de la liaison virus/récepteur. (D'après K. V. Holmes 2005)*

## **2.5. Sensibilité et résistance génétique au SRAS-CoV**

Les investigations concernant la recherche d'une résistance ou d'une sensibilité génétique au SRAS-CoV ont débuté dès l'émergence du SRAS-CoV.

Ces recherches ont permis de mettre en évidence que certains haplotypes HLA étaient associés à une augmentation de la sévérité de l'infection par le SRAS-CoV. Parmi eux, le HLA-B\*4601 a été identifié comme un facteur aggravant du SRAS chez un groupe de patients taiwanais (M Lin et coll. 2003). Toutefois, cette observation n'a pas été corroborée lors de la recherche de cet haplotype chez des individus infectés Hongkongais (M H L Ng et coll. 2004). Cependant, une association entre augmentation de la sensibilité au SRAS-CoV et expression des HLA-B\*0703 et HLA-DRB1\*0301 a été mise en évidence pour ces individus. Une étude a également montré une possible corrélation entre un taux faible ou nul de lectines sériques liant le mannose et une plus grande sensibilité à l'infection (Ip et coll. 2005).

Concernant une résistance génétique au SRAS-CoV, certains auteurs ont observé sur une cohorte d'individus Hongkongais qu'elle pouvait être corrélée à une homozygotie du gène *CLEC4M* codant la lectine L-SIGN (V S F Chan et coll. 2006). Toutefois ces résultats ont été controversés car deux autres équipes ont montré qu'ils n'étaient pas statistiquement valides (Tang et coll. 2007; Zhi et coll. 2007) Néanmoins, un autre

caractère génétique a été corrélé statistiquement à une résistance au SRAS-CoV : l'expression des HBGA ABH.

Une étude épidémiologique réalisée sur du personnel soignant exposé directement au virus sans protection appropriée a permis de mettre en évidence que les individus de phénotype O sont beaucoup moins sujets à l'infection par le SRAS-CoV que les individus possédant les phénotypes A et/ou B (Y. Cheng et coll. 2005) (Tableau 6). Les auteurs de cette étude concluent à une potentielle utilisation des HBGA par le SRAS-CoV comme le font les NoV ou la bactérie *Helicobacter pilori*. Toutefois, d'autres auteurs comme Gustafsson émettent l'hypothèse d'une possible protection contre le virus associée aux anticorps naturels anti-ABH (K Gustafsson et coll. 2005; Dürrbach et coll. 2007).

**Tableau 6 : Données statistiques comparant la probabilité d'infection par**

Phénotype ABH	Nombre de participants		Probabilité d'infection (intervalle de confiance 95%)	Valeur statistique
	Infectés par les SRAS-CoV (n=34)	Non-infectés par les SRAS-CoV (n=11)		
O	11	8	<b>0,18 (0,04-0,81)</b>	0,03
A	8	0	Indéfini	0,17
B	12	3	1,46 (0,32-6,53)	0,73
AB	3	0	Indéfini	0,57

**le SRAS-CoV et le phénotype ABH des individus ayant été au contact direct du virus.**

(D'après Y. Cheng et coll. 2005)

# CHAPITRE 5 :

## TRAVAUX ET PUBLICATIONS

### 1. OBJECTIFS

L'ensemble des travaux effectués durant ma thèse a été consacré à l'étude des inter-relations qui peuvent exister entre les HBGA ABH et les virus. Dans ce but, nous nous sommes particulièrement intéressés aux Calicivirus (NoV et RHDV) et au SRAS-CoV ; l'objectif principal étant de mieux comprendre pourquoi les gènes du système ABH (*ABO*, *FUT2*, *Sec1*) sont si polymorphes et quels avantages sélectifs cela pourrait apporter à l'hôte vis-à-vis de l'infection par ces virus.

Ces recherches ont donné lieu à trois articles dont deux sont actuellement en cours de soumission ou de révision :

- **Article 1 : La lipase stimulée par les sels biliaires ainsi que les mucines du lait de mères « sécréteur » inhibent la fixation du NV sur ses récepteurs glycaniques**

Le but de cette étude a été d'identifier les glycoprotéines solubles du lait de mères « sécréteur » capables d'inhiber la fixation des VLP de NV sur son récepteur : l'HBGA H type 1. Pour cela plusieurs échantillons de lait ont été traités (écrémage, dénaturation des protéines, élimination des oligosaccharides libres) afin de répondre aux questions suivantes : **1/** Quelle est la nature des molécules inhibant la fixation du virus sur son récepteur ? **2/** Dans quelle(s) fraction(s) du lait celles-ci sont-elles présentes ? **3/** Quel est leur impact individuel sur l'inhibition de la fixation du NV ?

- **Article 2 : Association entre expression de l'antigène tissulaire de groupe sanguin H, polymorphisme des  $\alpha$ 1,2fucosyltransférases chez le lapin de garenne et sensibilité au RHDV**

L'objectif de cette étude a été de déterminer si la pression de sélection exercée par le RHDV sur les populations de lapins de garenne pouvait engendrer dans celles-ci une augmentation de la fréquence d'un allèle associée à un éventuel phénotype non sécréteur entraînant une résistance à l'infection. Grâce à une collaboration avec l'Office National de la Chasse et de la Faune Sauvage, nous avons pu avoir accès à des échantillons provenant de trois populations sauvages de lapins présentant des passés épidémiques différents vis-à-vis du RHDV. Nous avons pu ainsi tester : **1/** Si la fréquence du phénotype non sécréteur augmente dans les populations soumises à de très fortes pressions de sélection par le RHDV. **2/** Si cette augmentation de fréquence peut être associée à un variant particulier codant une  $\alpha$ 1,2fucosyltransférase inactive. **3/** Si le phénotype non sécréteur chez le lapin de garenne peut être associé à une mutation particulière dans la séquence codante des gènes d' $\alpha$ 1,2fucosyltransférases.

- **Article 3 : Inhibition de l'interaction entre la protéine Spike du SRAS-CoV et son récepteur cellulaire par des anticorps anti-HBGA**

Dans cette étude, nous avons voulu tester si les données épidémiologiques concernant une plus grande résistance des individus O à l'infection par le SRAS-CoV pouvaient être expliquées par la reconnaissance des glycanes de la protéine S par les anticorps naturels (NAb) anti-ABH. Pour cela, nous avons utilisé un modèle cellulaire de fixation de la protéine S sur son récepteur cellulaire, l'ACE2. Nous avons modifié ce système en faisant co-exprimer la glycoprotéine virale et la machinerie enzymatique nécessaire à la synthèse de l'HBGA A dans les mêmes cellules. Nous avons ensuite évalué : **1/** Si la fixation pouvait être inhibée par un anticorps monoclonal anti-A. **2/** Si la fixation pouvait être inhibée par les NAb anti-A du plasma d'individus O. **3/** L'impact épidémiologique du polymorphisme ABH, associé à de forts titres en NAb anti-ABH.

## 2. ARTICLES

### **2.1. Article 1 : La lipase stimulée par les sels biliaires ainsi que les mucines du lait de mères « sécréteur » inhibent la fixation du NV sur ses récepteurs glycaniques**

#### **a - Introduction**

Les NoV sont responsables chez l'homme de la majorité des épidémies de gastro-entérites virales à caractère hivernal. Ils infectent les individus indépendamment de leur âge ou de leur niveau social. Ces virus reconnaissent et se lient à certains HBGA. Il a notamment été montré que la souche NV utilise l'HBGA H type 1 (Fuc $\alpha$ 1,2Gal $\beta$ 1,3GlcNAc) pour se fixer aux entérocytes de l'hôte (Marionneau et coll. 2002). Certains individus n'expriment pas cet HBGA au niveau de leurs épithéliums et dans leurs sécrétions. Cette carence est attribuée à la présence de deux copies inactives du gène *FUT2* dans leur génome. Ces personnes sont alors dites de phénotype non-sécréteur. D'ailleurs, deux études d'inoculation du NV chez des volontaires ont montré que le caractère non-sécréteur est associé à une résistance génétique à l'infection par ce virus (Lindesmith et coll. 2003; A M Hutson et coll. 2005).

Les études épidémiologiques concernant les gastro-entérites infantiles montrent que l'allaitement par les mères de phénotype sécréteur peut avoir un effet bénéfique en préservant les nouveau-nés de l'infection par certains pathogènes, dont les NoV (A L Morrow, G M Ruiz-Palacios, et coll. 2004). Contrairement au lait de vache, le lait maternel de femme sécréteur contient un grand nombre de molécules glycosylées solubles capables d'inhiber la fixation du NV sur l'HBGA H type 1 (Marionneau et coll. 2002; X Jiang et coll. 2004). *In vivo*, ces molécules joueraient le rôle de leurre. Reconnues par le NV, elles pourraient séquestrer les virions, les empêchant ainsi d'atteindre leur récepteur sur les entérocytes (Le Pendu 2004). Cette hypothèse est d'autant plus plausible qu'une grande partie des oligosaccharides du lait maternel sont préservés au cours du transit digestif du nouveau-né (Chaturvedi, Warren, Buescher, et coll. 2001).

Bien que les effets inhibiteurs du lait maternel sur la fixation du NV soient confirmés, les molécules impliquées dans cette inhibition ne sont toujours pas clairement identifiées. Les travaux présentés dans l'article suivant ont eu pour objectif de les caractériser et d'évaluer leur capacité inhibitrice respective.

## Bile-salt-stimulated lipase and mucins from milk of 'secretor' mothers inhibit the binding of Norwalk virus capsids to their carbohydrate ligands

Nathalie RUVOËN-CLOUET\*†, Eric MAS‡, Séverine MARIONNEAU\*, Patrice GUILLON\*, Dominique LOMBARDO‡ and Jacques LE PENDU\*<sup>1</sup>

\*INSERM U601, Institute of Biology, 9 Quai Moncousu, 44093 Nantes Cedex 01, France, †National Veterinary School of Nantes, Route de Gachet, BP 40706, 44307 Nantes Cedex 03, France, and ‡INSERM U559, IPHM Faculté de Médecine, 27 Blv Jean Moulin, 13385 Marseille Cedex 05, France

Breast-feeding-associated protection against calicivirus diarrhoea is associated with the presence of high levels of 2-linked oligosaccharides in mother's milk, and human calicivirus strains including the NV (Norwalk virus) use gut 2-linked fucosylated glycans as receptors, suggesting the presence of decoy receptors in milk. Our aim was to analyse the ability of human milk to inhibit the attachment of rNV VLPs (recombinant NV-like particles) to their carbohydrate ligands and to characterize potential inhibitors found in milk. Milk from women with the secretor phenotype was strongly inhibitory, unlike milk from women that are non-secre-

tors, which is devoid of 2-linked fucosylated structures. At least two fractions in human milk acted as inhibitors for the NV capsid attachment. The first fraction corresponded to BSSL (bile-salt-stimulated lipase) and the second to associated mucins MUC1 and MUC4. These proteins present tandem repeat O-glycosylated sequences that should act as decoy receptors for the NV, depending on the combined mother/child secretor status.

**Key words:** bile-salt-stimulated lipase (BSSL), mucin, norovirus, fucosyltransferase, histo-blood-group, gastroenteritis.

### INTRODUCTION

Gastroenteritis constitutes a major cause of morbidity and mortality in developing countries and remains a major health problem in developed countries. Recent epidemiological studies revealed that the majority of outbreaks of non-bacterial gastroenteritis were caused by members of the norovirus genus in the calicivirus family [1,2]. These viruses, which infect people of all ages have also been reported to be involved in a large number of sporadic cases, particularly among children [3,4]. Although it is considered that norovirus gastroenteritis is of moderate severity, recent reports suggest that in people under stressful conditions or in young children it could be more severe than expected [5].

We and others have recently shown that some caliciviruses attach to host epithelial cells through carbohydrates of the histo-blood-group family. Thus NV (Norwalk virus), the prototypical norovirus, binds to 2-fucosylated structures based on the blood-group-antigen H type 1 trisaccharide  $Fuc\alpha 2Gal\beta 3GlcNAc$  [6–8]. Synthesis of this ligand requires an active  $\alpha 1,2$ -fucosyltransferase, which in most epithelial cells is the FUT2 enzyme. The presence of this enzymatic activity characterizes the so-called secretor phenotype. Inactivating mutations in the *FUT2* gene, occurring in approx. 20% of the population with European ancestry, are responsible for the non-secretor phenotype, characterized by an absence of 2-fucosylated structures, and hence of ABO(H) blood-group antigens in secretions such as saliva and on most epithelial cells [9]. Consistent with the attachment of NV to H type 1-containing glycans, volunteers challenged with the virus become infected only if they are secretors [10,11].

Breast-feeding has been known for a long time to protect infants from diarrhoea and a number of potential protective factors in human milk have been characterized. These include maternal antibodies, lactoferrin, BSSL (bile-salt-stimulated lipase), lactadherin, lysozyme, cytokines, oligosaccharides and mucins

[12,13]. Unlike cows' milk, human milk contains large amounts of fucosylated oligosaccharides, the amount of 2-linked fucose depending upon the *FUT2* status of the mother, with large variations during the course of lactation [14–16]. Human milk additionally contains glycans in the form of glycolipids or glycoproteins that can be fucosylated in a genetically determined manner. We observed previously that the milk from secretor mothers could inhibit the binding of rNV VLPs (recombinant NV virus-like particles) to their carbohydrate ligands [7] and hypothesized that the 2-fucosylated milk component could have a protective effect, acting as a decoy receptor, depending upon the combined *FUT2* status of the mother and child [17]. Thus a non-secretor child would be genetically resistant to NV whereas a secretor child that is potentially sensitive would be protected by his mother's milk providing she is of the secretor-type. As a first step towards testing this hypothesis, the present study was undertaken in order to characterize the components of human milk that inhibit NV-binding to its carbohydrate ligand.

### MATERIALS AND METHODS

#### Collection and treatment of milk samples

Human milk was obtained from the milk bank of the Nantes University Hospital in accordance with its ethical guidelines. Milk was collected from 15 anonymous, healthy mothers at 3 weeks lactation, labelled from 1 to 15 and stored at  $-20^{\circ}\text{C}$  prior to analysis. Samples labelled A, from an additional mother, were collected at 3 and 5 weeks post-partum. Inactivation of the milk antibodies and enzymes was performed by boiling  $100\ \mu\text{l}$  samples in a water bath in the presence of 5% (v/v)  $\beta$ -mercaptoethanol.

To prepare fractions, frozen human milk was thawed and centrifuged at  $20\,000\ \text{g}$  for 1 h. The fat layer and the pellet were removed and the remaining skimmed milk was filtered. In order to

Abbreviations used: BSDL, bile-salt-dependent lipase; BSSL, bile-salt-stimulated lipase; Le, Lewis blood-group-antigen; mAb, monoclonal antibody; NV, Norwalk virus; rNV VLP, recombinant NV virus-like particle; TMB, 5-tetramethylbenzidine.

<sup>1</sup> To whom correspondence should be addressed (email jlependu@nantes.inserm.fr).

remove small molecules such as oligosaccharides, skimmed milk was dialysed against PBS. Human fat globule membranes were prepared by sonicating the fat layer after addition of PBS followed by ultracentrifugation at 100 000 *g* for 1 h and recovery of the pellet as previously described [18].

#### Determination of the secretor character

Milk samples were diluted at 1:10 in 100 mmol/l carbonate buffer (pH 9.6), coated on to NUNC immunoplates (Roskilde, Denmark) and stored overnight at 37 °C. After blocking with 5% dried cows' milk in PBS, mAb (monoclonal antibody) LM137, which is specific for H type 1 and Le<sup>b</sup> (Lewis b histo-blood-group) antigens, was added. After incubation for 2 h at 37 °C, horseradish peroxidase conjugated anti-goat IgG and anti-mouse IgG (Uptima, Montluçon, France) at 1:2000 dilutions were added. Between each step, the plates were washed three times with PBS, 5% (v/v) Tween-20. The enzyme signals were detected using TMB (5-tetramethylbenzidine) as substrate (BD Bioscience, San Jose, CA, U.S.A.) then read at A<sub>450</sub>. Strong reactivity with the anti-II type 1/Le<sup>b</sup> mAb was detected in 11 out of 16 cases. To confirm that this did correspond to the difference between secretor- and non-secretor-mothers, genotyping for the *FUT2* gene was performed. The G428A mutation is responsible for nearly all cases of non-secretor phenotype in the French population, thus it was the only mutation that was screened. Genomic DNA from total milk that contains epithelial cells was extracted from the first pellet using the Qiagen QIAmp DNA mini kit according to the manufacturer's instructions. A fragment of the *FUT2* gene was amplified with the following primers: GAGGAATACCGCCACATCCCGGGGAGTAC, forward; ATGGACCCCTACAAAGGTGCCCGGCCGGCT, reverse, and digested using *Av*II. The G428A mutation eliminates this restriction site.

#### Inhibition of rNV VLP-binding to saliva or synthetic H type 1 antigen

Saliva from secretor individuals [secretor (active *FUT2* allele), blood type O, Le positive] at a dilution of 1:1000, or H type 1 synthetic oligosaccharide bound to polyacrylamide at 50 µg/ml was coated on to Maxisorb NUNC plates, stored overnight at 37 °C and saturated with 5% dried milk in PBS. rNV VLPs were produced using a baculovirus expression system as previously described [19]. rNV VLPs were pre-incubated with serially diluted total milk samples or milk fractions (v/v) at a final concentration of 0.6 µg/ml for 60 min at room temperature and were then incubated for 1 h on the coating plate. Binding was detected using a rabbit anti-rNV VLP antiserum diluted to 1:2000 in 5% dried milk in PBS, followed by incubation with peroxidase-conjugated anti-(rabbit IgG). Both reagents were incubated for 1 h at 37 °C. Reactions were revealed by using the TMB substrate and then read at A<sub>450</sub>.

#### Immunohistochemical analyses

Paraffin embedded tissue sections of the gastroduodenal junction from adult secretor individuals of blood-group O type were rehydrated in graded ethanol and washed in PBS. Endogenous peroxidase was inhibited using 0.3% methanol or H<sub>2</sub>O<sub>2</sub> for 15 min. After saturation with PBS containing 3% BSA for 30 min, sections were covered for 90 min at room temperature with a 1:1 (v/v) mixture of 5 µg/ml rNV VLPs and human milk samples at 1:10 dilution, in PBS that had been pre-incubated at 4 °C overnight. The anti-rNV capsid protein 9C3 antiserum, diluted 1:1000 in PBS and 1% (w/v) BSA was added and the mixture was incubated at room temperature for 1 h. Sections were then rinsed

three times with PBS and incubated with peroxidase-conjugated anti-(mouse IgG) diluted to 1:2000 in PBS, 1% BSA (Vector Labs, Burlingame, CA, U.S.A.) for 45 min, and reactions were visualized with 3-amino-9-ethylcarbazol (Vector Labs). Counterstaining was performed using Mayer's hemalun.

#### Western blotting

Protein concentrations were measured using bicinchoninic acid. Proteins (30 µg) were subjected to SDS/PAGE (5–10% gel). Separated proteins were electrophoretically transferred to Immobilon membranes (Millipore, Milford, MA, U.S.A.) in 25 mM Tris, 192 mM glycine, 20% methanol at 350 mA for 1 h. After transfer, membrane strips were incubated in PBS containing 3% skimmed milk for 1 h at 37 °C. Membranes were then incubated with 3 µg/ml rNV VLPs diluted in PBS containing 1% skimmed milk overnight at 4 °C. Binding was detected using a rabbit anti-rNV VLP antiserum diluted to 1:2000 in 5% dried milk in PBS, followed by peroxidase-conjugated anti-(rabbit IgG) at a dilution of 1:1000. After washing twice in PBS containing 5% (v/v) Tween-20 and once in PBS, the bound antibodies were revealed by chemiluminescence using the ECL<sup>®</sup> (enhanced chemiluminescence) Kit from Amersham (Little Chalfont, U.K.). The following primary antibodies were also used: LM137 anti-H type 1/Le<sup>b</sup>, L64 rabbit anti-BSDL (bile-salt-dependent lipase) [20], or HMFG1 and 8G7 directed against the MUC1 and MUC4 human mucins respectively [21,22]. After washes, membranes were incubated for 1 h with peroxidase-labelled anti-(mouse IgG) diluted to 1:2000 in PBS containing 1% skimmed milk and reactions were visualized as described above. Alternatively, membranes that were treated with polyclonal L64 antibody were incubated for 1 h with alkaline phosphatase-labelled anti-(rabbit IgG) and reactions were carried out as previously described [20].

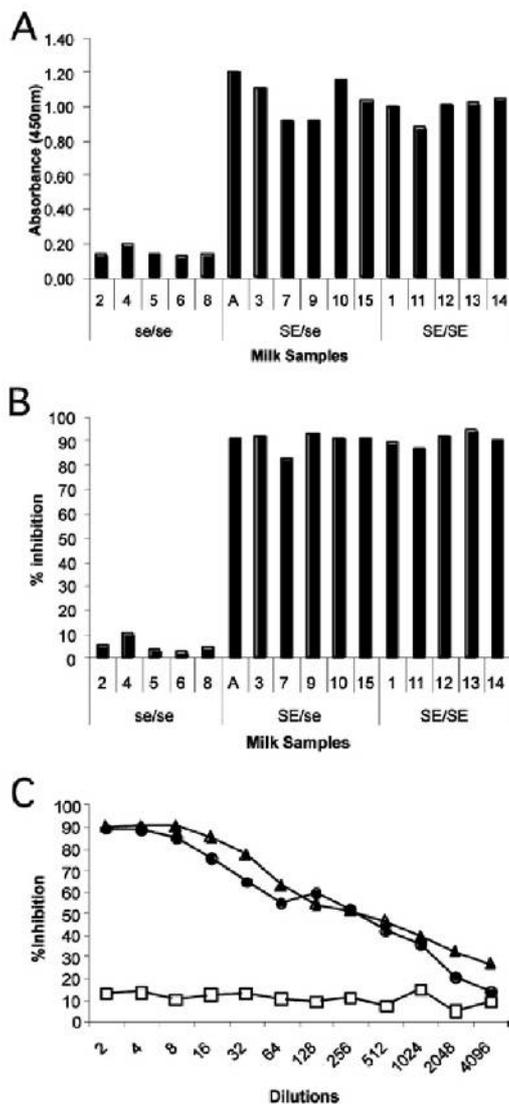
#### Preparation of the mucin fraction and purification of BSSL and BSDL

The mucin fraction from human milk was prepared by preparative electrophoresis. Skimmed milk samples (750 µl) containing 30 mg/ml protein were loaded on to SDS/PAGE (5%) gel and allowed to migrate for 1 h at 150 V until Coomassie Blue stained bands were no longer detectable. The upper third of the gel was cut and minced in to small fragments and incubated in PBS for 48 h at 4 °C before elution. After centrifugation at 800 *g* for 5 min, the soluble material was dialysed against PBS and concentrated to the original volume of milk on 100 kDa cut-off Centricon membranes (Millipore). BSSL and its pancreatic counterpart BSDL were purified from human milk and human pancreatic juice respectively [20]. The esterolytic activity of these enzymes was recorded on 4-nitrophenyl caproate as described by Gjellesvik et al. [23].

## RESULTS

#### Inhibition of rNV VLP ligand-binding by human milk

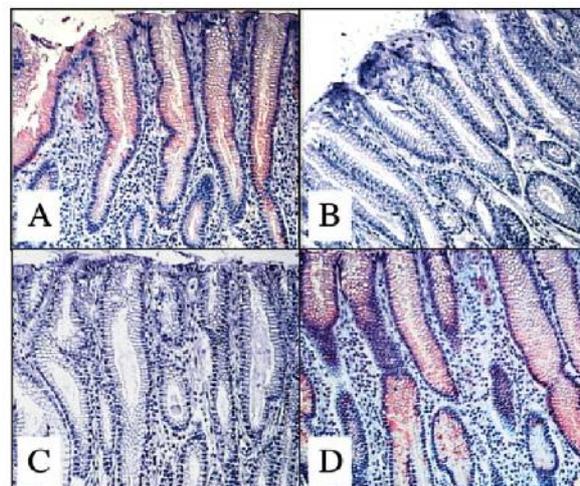
We have shown previously that human milk could inhibit rNV VLP-binding to their histo-blood-group ligand [7]. We thus sought to determine if this inhibition could be influenced by the secretor character of the mother. Since the composition of milk varies during the course of lactation, all samples used were collected at the same time-point post-partum. The mothers' secretor phenotype was determined from their milk samples (Figure 1A). A total of five non-secretors were found, all of whom were homozygous for the G428A mutation. Serially diluted total milk samples were used and as can be seen in Figure 1(B) at a dilution of 1:10, none



**Figure 1** FUT2-dependent molecules from human milk inhibit rNV VLP attachment to H type 1 histo-blood-group antigen

(A) Milk samples from women genotyped for the *FUT2* locus were coated on to ELISA plates at a dilution of 1:1000 and the presence of H type 1-reactive molecules was determined by reactivity of the anti-H type 1/Le<sup>x</sup> mAb, LM137. A total of five women were considered as non-secretors (2, 4, 5, 6 and 8) and the remaining as secretors. SE = active *FUT2* allele; se = inactive *FUT2* allele. (B) Inhibition of rNV VLP attachment to polyacrylamide-conjugated H type 1 by milk samples from secretor and non-secretor women diluted to 1:10. (C) Examples of the inhibitory potency of milk samples, from 2 secretors (●, ▲) and a non-secretor (□), on rNV VLP attachment to saliva from an O type secretor individual. The inhibition assay was performed as described in the Materials and methods section. The percentage of inhibition is shown as a function of the reciprocal milk dilution.

of the five non-secretor samples inhibited rNV VLP-binding to immobilized synthetic H type 1, in contrast with the strong inhibition given by all samples from secretor mothers, irrespective of whether they were homozygotes or heterozygotes. A representative example of the strength of the inhibition by secretor milk is shown in Figure 1(C). Inhibition (50%) was obtained at a dilution of 1:500, whereas no significant inhibition was observed even at a



**Figure 2** Binding of rNV VLPs to human gastroduodenal tissue and inhibition by human milk

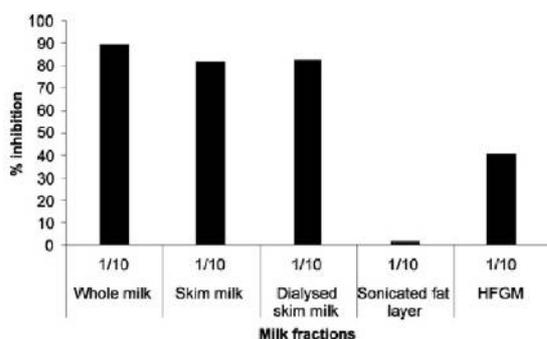
Tissue sections from the gastroduodenal junction of a secretor (A, C and D) and a non-secretor (B) were incubated with rNV VLPs and the binding was detected as described in the Materials and methods section. In (C, D) VLPs were co-incubated with milk samples from secretor and non-secretor women respectively, at a dilution of 1:10. All pictures taken are from the pyloric area. A complete absence of VLP binding was observed in the presence of milk from a secretor, as in the tissue from a non-secretor.

dilution of 1:2 in non-secretor milk. Similar results were obtained when saliva from an O type secretor individual was used, instead of synthetic H type 1, for attachment of rNV VLPs (results not shown).

In order to determine if the inhibition by the synthetic ligand could be physiologically relevant, it was tested on gastroduodenal junction tissue sections. We have previously shown that rNV VLPs attach to the surface of epithelial cells only in secretors [7]. Thus rNV VLPs were co-incubated with milk samples on tissue sections from secretor individuals. As illustrated in Figure 2, for the pyloric area, a strong binding of recombinant capsids occurred in the presence of milk from non-secretor mothers, whereas no binding to tissue sections could be seen in the presence of milk from secretor mothers. The same results were obtained in tissue sections from the duodenum (results not shown). This indicates that human milk from secretors has the potential to block the attachment of NV to its carbohydrate ligand on epithelial cells.

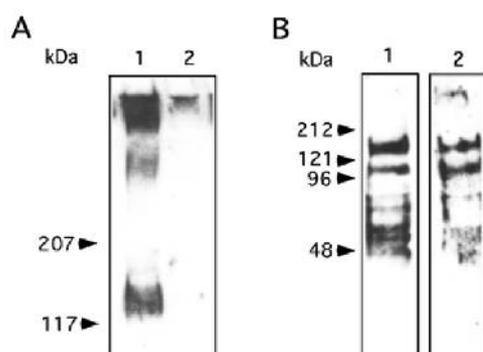
#### Milk fractions containing inhibitory activity

Human milk possesses large amounts of free oligosaccharides that could potentially inhibit the attachment of rNV VLPs to their carbohydrate ligands. In addition, similar carbohydrate structures can be complexed to proteins or lipids. Blocking-antibodies could also be present. In order to determine which of these components is mostly responsible for the inhibition observed above, various milk fractions were prepared. We first compared milk samples that had been boiled in the presence of  $\beta$ -mercaptoethanol with their untreated counterpart in the inhibition assay. No difference was observed between the two types of samples, ruling out the possibility that blocking-antibodies or milk enzyme activities were responsible for the inhibition of rNV VLP-binding to carbohydrate ligands. Fat-globule membranes showed some inhibition (Figure 3). Skimmed milk, as well as dialysed skimmed milk were nearly as potent as inhibitors as unfractionated milk. Since



**Figure 3** Inhibition of rNV VLP attachment to saliva from a secretor individual by milk fractions from a secretor donor

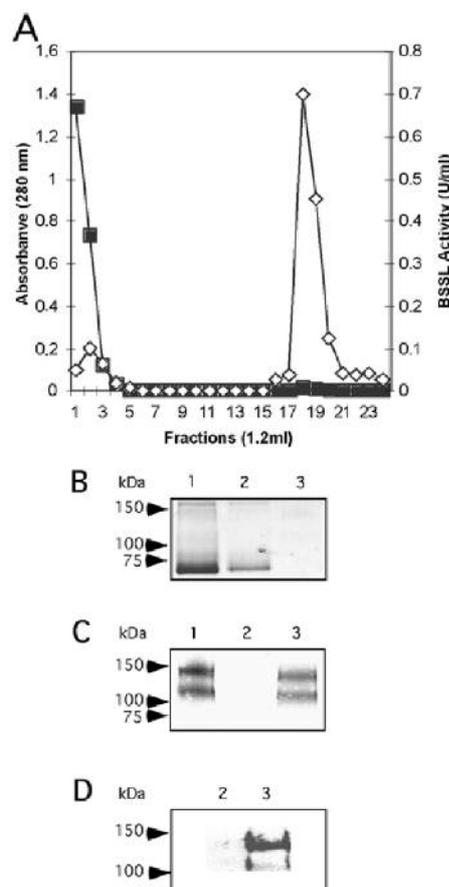
A fresh secretor milk-sample was centrifuged at 1500 *g* for 20 min to separate the cream from the milk. The skimmed milk was dialysed to remove free oligosaccharides and the cream was washed by centrifugation in PBS to remove soluble proteins. Washed cream was resuspended in PBS and sonicated to disrupt fat globules. The human fat globule membranes (HFGM) were pelleted by centrifugation at 100 000 *g* for 1 h and separated from the fat layer. Each fraction corresponding to the original milk dilution at 1:10 was co-incubated with rNV VLPs in the inhibition assay as described above.



**Figure 4** Detection of human milk rNV VLP ligands by Western blotting

(A) Skimmed milk samples were run on SDS/PAGE (5% gel under reducing conditions) and probed with rNV VLPs; lane 1, secretor sample number 3; lane 2, non-secretor sample number 4. (B) Skimmed milk from a secretor (sample A) was run on an SDS/PAGE 5–10% gradient gel under reducing conditions and probed either with the anti-H type 1/Le<sup>b</sup> mAb, LM137 (lane 1) or with rNV VLPs (lane 2). Positions of the molecular mass markers are shown by arrows.

dialysis did not decrease inhibitory potency, it is likely that free oligosaccharides are not involved. This was confirmed by using purified oligosaccharides. Neither the pentasaccharide Fuca2Gal-β3GlcNAcβ3Galβ4Glc (LNF-I) nor the tetrasaccharide Galβ3-GlcNAcβ3Galβ4Glc (LN<sup>T</sup>) was able to inhibit ligand-binding even at concentrations as high as 22 mM. Unlike the tetrasaccharide, the fucosylated pentasaccharide is terminated by the II type 1 epitope and is characteristic of milk from secretor mothers, yet it did not inhibit binding. The inhibitory material was thus searched for in the protein fraction. Western blotting indicated that rNV VLPs could detect two sets of bands in milk samples exclusively from secretor mothers. These corresponded to a smear of very high molecular mass and to a broad band of approx. 130 kDa (Figure 4A). Most milk samples showed only one band at that level, but some samples showed two bands. Yet since larger samples were obtained from woman A whose milk showed two bands of approx. 130 kDa, most experiments were performed



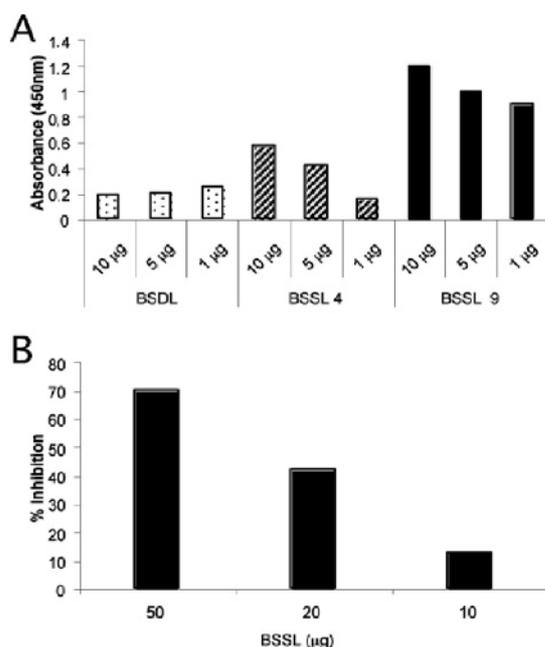
**Figure 5** Purification of BSSL from the milk of secretor woman A

(A) Total proteins were loaded on to an immobilized anti-BSDL pAbL64 (polyclonal antibody L64) column equilibrated with PBS and incubated overnight at 4°C. Unbound fractions were eluted and the column was washed with 0.02 M Tris/HCl and 0.5 M NaCl, pH 7.4 until zero absorbance. Bound fractions were eluted with 0.1 M glycine, pH 2.5. Elution of material was recorded by the absorbance at 280 nm (■). Enzyme activity (◇) was determined with 4-nitrophenyl hexanoate as substrate. (B) Coomassie Blue stained SDS/PAGE gels and (C, D) Western blotting of fractions eluted from the immobilized pAbL64 column, (lane 1, 1 μg of starting material), (lane 2, 1 μg of unbound material), (lane 3, 1 μg of bound material). Immunodetection was performed with mAb pAbL64 (C) or rNV VLPs (D).

using this milk. Additional bands corresponding to material of lower size could be detected using an anti-H type 1/Le<sup>b</sup> mAb but they were less clearly revealed with rNV VLPs (Figure 4B).

#### BSSL and mucins from secretor mothers are inhibitory

BSSL corresponds to a major milk protein. It can be found in a heavily glycosylated form and has a molecular mass of approx. 130 kDa. It therefore appeared to be a good candidate for the inhibition of rNV ligand-binding. In addition, Western blotting of our human milk samples using an antiserum raised against a BSSL isoform purified from pancreatic juice (BSDL) revealed bands that co-migrated with those detected by rNV VLPs (as described below). We thus purified BSSL from the milk of a secretor and a non-secretor mother. The purification was performed by affinity chromatography of immobilized antibodies as shown in Figure 5(A). The milk fraction from woman A, which contained the lipase activity, yielded two bands specifically recognized by

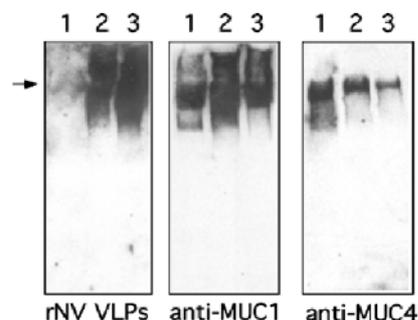


**Figure 6** Binding of rNV VLPs to purified BSDL and BSSL

(A) Purified glycoproteins (1, 5 or 10 µg) from secretor donors (BSDL from the pancreatic juice of an O type secretor patient and BSSL from mother number 9) or from a non-secretor milk donor (mother number 4) were coated on to ELISA plates and the binding of rNV VLPs was tested. (B) Inhibition of rNV VLP attachment to the saliva of a secretor individual by purified BSSL from secretor A.

the anti-BSDL antiserum, as well as by rNV VLPs (Figures 5B, 5C and 5D). The binding of rNV VLPs to purified BSSL from the secretor mother was compared with that of the non-secretor mother and to BSDL purified from the pancreatic juice of an O type secretor individual. A strong binding occurred to BSSL from the secretor mother whereas binding was much lower to BSSL from the non-secretor mother. More surprisingly, no binding was observed to purified BSDL from an O type secretor individual, suggesting a difference in glycosylation between the pancreatic and milk isoforms of the lipase that can be detected by the virus capsids (Figure 6A). The potency of BSSL was then tested in the inhibition assay. BSSL from the secretor mother was indeed able to inhibit the binding of rNV VLPs to either salivary mucins from an O type secretor or to synthetic H type 1 antigen. Final amounts of 30 or 2 µg of BSSL were necessary to produce 50% inhibition of rNV VLP-binding to salivary mucins (Figure 6B) and to synthetic H type 1 antigen respectively (results not shown).

The diffuse high molecular mass band stained by rNV VLPs on Western-blotting suggested recognition of a mucin component. This high molecular mass fraction was isolated by preparative electrophoresis. Western-blotting from 5% polyacrylamide gels showed that rNV VLPs bind equally well to the mucin preparation as to the high molecular mass fraction from the total milk of a secretor mother. They do not bind to similar material from the milk sample of a non-secretor donor. Yet, irrespective of the donor's secretor character, milk and high molecular mass preparations contained both MUC1 and MUC4 mucins (Figure 7). The intensity of staining is weaker when using the anti-MUC4 mAb than the anti-MUC1 mAb but it should be noted that the anti-MUC4 mAb was raised against an unglycosylated synthetic peptide [22] and



**Figure 7** Binding of rNV VLPs to human milk mucins

Skimmed milk samples from non-secretor (lane 1) or secretor (lane 2) donors and the mucin fraction from milk of a secretor donor prepared as described in the Materials and methods section (lane 3) were run on SDS/PAGE (5%) gel with a 4% stacking gel. After transfer, membranes were probed with either rNV VLPs, an anti-MUC1 mAb or an anti-MUC4 mAb. The part of the gels shown is above the 200 kDa marker and the arrow indicates the limit of the stacking gel.

might not optimally recognize the native glycosylated milk mucin. Similar to that which we observed using purified BSSL, the mucin fraction from the milk of secretor mothers was able to strongly inhibit the binding of rNV VLPs whereas the mucin fraction from a non-secretor woman was not (results not shown). In order to determine the relative importance of BSSL versus the mucins, we compared the inhibitory capacity of milk from a secretor woman with that of the same BSSL-depleted milk. BSSL depletion was performed by loading 0.5 ml of skimmed milk on to the anti-BSSL pAb64 column. This corresponded to 12 mg of total protein with 290 units/mg of BSSL specific activity. The unretained fraction was collected and 12 mg of protein was recovered containing 9 units/mg of BSSL specific activity, indicating that most of the BSSL had been removed by the immobilized antibodies. This was further confirmed by Western blotting. The ratio of dilutions for the two samples (BSSL-depleted/total) required to reach 50% inhibition in the rNV VLP binding assay was then determined. The ratio was  $0.40 \pm 0.07$  from four independent experiments performed in triplicate, showing that BSSL contributes 60% of the total inhibitory activity, whereas the remaining 40% corresponds to the contribution from the mucins and possibly from some other minor components.

## DISCUSSION

Recent studies have provided evidence that fucosylated oligosaccharides from human milk confer protection against diarrhoea to breast-fed children [24]. More specifically, the presence of 2-linked fucose residues is associated with protection against both *Campylobacter* and calicivirus diarrhoea [25]. *Campylobacter* as well as various calicivirus strains have been shown to attach to 2-linked fucosylated glycoconjugates of the digestive tract [6,7,26–28] and non-secretor individuals are resistant to NV infection, indicating that the 2-linked fucosylated ligands act as receptors [10,11]. Thus it is likely that the protection conferred by breast-feeding results from the presence of decoy receptors in milk. In previous work, we have shown that binding of rNV VLPs to epithelial cells was inhibited by the milk from a secretor woman, indicating that NV decoy receptors were present in human milk [7]. Recently, Jiang et al. [29] confirmed this observation in NV and other strains of noroviruses. We have now extended these findings by characterizing the decoy receptors.

Fucosylated oligosaccharides in human milk are synthesized by the products of the *FUT2* and *FUT3* genes that are responsible for the secretor and Le<sup>b</sup> blood types. These two fucosyltransferases catalyse the transfer of fucose residues on to positions two, three or four. The presence of fucosylated oligosaccharides in human milk depends upon genetic polymorphism at the *FUT2* and *FUT3* loci and milk from non-secretor women who lack the *FUT2* enzyme is almost entirely devoid of 2-linked fucosylated oligosaccharides [16]. Since blocking of rNV VLP-binding to its ligand is tightly controlled by the *FUT2* locus, it was tempting to assume that the NV decoy receptors in human milk were 2-linked fucosylated oligosaccharides. However, we observed that dialysis did not decrease the blocking activity and that purified LNF-I (Fuc $\alpha$ 2Gal $\beta$ 3GlcNAc $\beta$ 3Gal $\beta$ 4Glc) was not inhibitory. In a previous study, Jiang et al. [29] observed that total oligosaccharides had no inhibitory activity. Thus the milk's oligosaccharide fraction does not contain the decoy receptors. In a previous study, we observed that a synthetic H type 1 trisaccharide could block the binding of rNV VLPs to saliva from a secretor individual [7]. This could be due to a structural difference between the synthetic trisaccharide and the milk oligosaccharides. The synthetic trisaccharide was coupled to a short aliphatic chain that might increase its affinity by contributing a non-specific hydrophobic interaction with the viral capsid protein. The affinity of the monomeric milk oligosaccharides for their binding site is probably too low to bring about detectable inhibition in our assay conditions. Western blotting indicated that rNV VLPs recognize high molecular mass glycoproteins which turned out to correspond to BSSL and mucins. Purified BSSL from secretor mothers was inhibitory, as was a preparation of mucins that contained a mixture of MUC1 and MUC4. The degree of fucosylation of milk glycoproteins is likely to parallel that of free oligosaccharides, since both depend upon the presence and amount of the *FUT2* and *FUT3* enzymes. Thus, in the association recently noted between protection against calicivirus diarrhoea and the level of 2-linked fucosylated oligosaccharides [25], in all likelihood, these oligosaccharides were only a marker of the degree of fucosylation of decoy receptors that correspond to the above mentioned high molecular mass glycoproteins.

BSSL is one of the major glycoproteins from human milk, which is absent from cows' milk. This lipase is not expressed in the newborn immature pancreas and it is thought that its presence in milk compensates for this lack of pancreatic enzyme [30]. It can participate in the degradation of a large spectrum of lipids and thus may have an essential role in nutrition [31] but also in protection in the newborn [32]. Regardless, the blocking activity of BSSL from secretor milk strongly suggests that it could have a dual function. The first, corresponding to the catalytic domain of the lipase is located at the N-terminus of the protein. The second, involving innate protection against certain infectious diseases, would correspond to the C-terminal mucin-like domain that could act as a decoy receptor depending upon its glycosylation status. A protective role against *Giardia lamblia* trophozoites has previously been attributed to BSSL [33,34]. This pathogen provokes severe diarrhoea in the newborn [35] and its toxic effect is counteracted by BSSL upon the release of *cis*-unsaturated fatty acids from milk triacylglycerol [34]. However, the inhibitory effect of BSSL against attachment of rNV VLPs to their carbohydrate ligands cannot be attributed to a similar mechanism since its catalytic lipase activity is destroyed by heating [32] whereas boiling milk samples did not affect the inhibitory activity. However, BSSL, like BSDL, its pancreatic isoform, possesses a mucin-type C terminal domain with 16 tandem repeats rich in potential O-glycosylation sites. By the ELISA method, both BSSL and BSDL from O type secretor

individuals reacted with the anti-H type 1/Le<sup>b</sup> mAb, whereas those from non-secretor individuals reacted with an anti-Le<sup>a</sup> (Lewis a histo-blood-group-antigen) mAb, indicating that the two isoforms from the mammary gland and the pancreas respectively, carried histo-blood-group carbohydrates as previously reported (results not shown). However, only the milk isoform could be recognized by rNV VLPs and recognition was dependent upon the presence of 2-linked fucose residues since BSSL from a non-secretor woman was not recognized. The total carbohydrate content of BSSL ranges from 19 to 26% of the glycoprotein mass, whereas that of BSDL has been estimated at approx. 7% [36,37]. The lipase possesses a single site of N-glycosylation and its N-linked sugar chain corresponds to about 10% of the total oligosaccharides in the milk isoform [38]. Consistent with its greater degree of glycosylation, we noted, as have previous authors, that BSSL from most individuals displayed a higher apparent molecular mass by gel electrophoresis than did BSDL. Thus the presence of complex O-glycans on the milk isoform could explain why rNV VLPs could bind to BSSL and not to BSDL. In one individual, we observed two distinct forms of BSSL with apparent molecular masses of approx. 130 and 110 kDa. Such variants have already been described but their significance is still unknown. Given the difference in size between the two forms of BSSL in that individual's milk, it is possible that they result from a polymorphism of the number of tandem repeats of the C-terminal domain [39]. This type of polymorphism is well described for mucins, particularly for MUC1, and its existence for BSSL warrants further study regarding the inhibitory potency of the variants. Changes in BSSL glycosylation have been observed during lactation [38]. It will thus be interesting to determine whether the inhibitory activity of BSSL on rNV VLP binding, as observed here, varies during the course of lactation.

MUC1 has for a long time been known as a component of human fat globule membranes and as a soluble mucin. Its presence in fat globule membranes could explain the inhibition of rNV VLP attachment by the fat globule membrane fraction. A second mucin, distinguished by its higher mass, was called MUCX [40]. Although we were not able to clearly distinguish the two types of mucins by gel electrophoresis based on their size and using a specific antibody, we could show that our mucin preparations contained MUC1 as expected, regardless of the secretor character of the donor. Since lactating rat mammary gland has recently been shown to express MUC4 [41], which like MUC1 can be either membrane bound or soluble, we used a recently developed anti-human MUC4 mAb to assay the presence of MUC4 in human milk [22]. A clear reactivity was observed in all samples irrespective of the secretor character. MUC4 is a much larger mucin than MUC1 [42]. Thus the mucin previously termed MUCX, which was larger than MUC1, probably corresponds to MUC4. Our preparations of high molecular mass material containing the two mucins strongly blocked the attachment of rNV VLPs to their carbohydrate ligands, indicating that these mucins can act as decoy receptors, in accordance with their glycosylation state controlled by polymorphism in the fucosyltransferase genes. MUC1 from milk binds to rotavirus and to enteropathogenic strains of *Escherichia coli* and for this reason is already considered as a component of innate immunity from milk [43,44]. Our inability to separate MUC1 from MUC4 did not allow us to determine which of them is the best inhibitor but our results raise the possibility that MUC4 could also be a decoy receptor for some pathogens.

In conclusion, the results presented here show that human milk blocks the binding of NV capsids to their receptor and that the blocking results from the interaction with 2-linked fucose residues complexed on high molecular mass glycoproteins that

possess mucin-type tandem repeats. This repetitive structure in the core peptide, allowing a multimeric presentation of the glycan, probably increases the avidity of the interaction. Thus it should be possible to design artificial decoy receptors against noroviruses based on repeatedly complexed 2-linked fucosylated oligosaccharides mimicking the mucin-type arrangement. Such artificial decoy receptors could be included in milk formulations.

This work was supported by INSERM and by a grant (02/2) from the Direction de la Recherche Clinique from Nantes University Hospital. S. M. was supported by a grant from Danone. The authors thank Dr R. F. Fraser (Blood Transfusion Centre, Carlisle, Scotland, U.K.), J. Burchell (Breast Cancer Biology Group, Guy's Hospital, London, U.K.) and S. K. Batra (Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, NE, U.S.A.) for their generous gifts of antibodies LM137, HFGM-1 and 8G7. We also thank Dr G. Strocker (Université de Sciences et Technologies de Lille, France) for his gift of purified milk LND and LNF-1 and Mme C. Delalande (Nantes University Hospital, France) for her help with the collection of milk samples.

## REFERENCES

- Fankhauser, R. L., Nool, J. S., Monroe, S. S., Ando, T. and Glass, R. I. (1998) Molecular epidemiology of Norwalk-like viruses in outbreaks of gastroenteritis in the United States. *J. Infect. Dis.* **178**, 1571–1578.
- Lopman, B. A., Reacher, M. H., van Duynhoven, Y., Hanon, F.-X., Brown, D. W. and Koopmans, M. (2003) Viral gastroenteritis outbreaks in Europe, 1995–2000. *Emerg. Infect. Dis.* **9**, 90–96.
- Bon, F., Fasola, P., Dauvergne, M., Tenenbaum, D., Planson, H., Parion, A. M., Pothier, P. and Köhli, E. (1999) Prevalence of group A rotavirus, human calicivirus, astrovirus and adenovirus type 40 and 41 infections among children with acute gastroenteritis in Dijon, France. *J. Clin. Microbiol.* **37**, 3055–3058.
- Pang, X.-L., Honma, S., Nakata, S. and Vesikari, T. (2000) Human caliciviruses in acute gastroenteritis of young children in the community. *J. Infect. Dis.* **81**, S288–S294.
- Parashar, U. D., Li, J. F., Cama, R., DeZalia, M., Monroe, S. S., Taylor, D. N., Figueroa, D., Gilman, R. H. and Glass, R. I. (2004) Human caliciviruses as a cause of severe gastroenteritis in peruvian children. *J. Infect. Dis.* **190**, 1088–1092.
- Harrington, P. R., Lindensmith, L., Yount, B., Moe, C. L. and Baric, R. S. (2002) Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J. Virol.* **76**, 12325–12343.
- Marionneau, S., Ruvoën-Clouet, N., Le Moullac-Vaidye, B., Clément, M., Cailleau-Thomas, A., Ruiz-Palacios, G., Huang, P. W., Jiang, X. and Le Pendu, J. (2002) Norwalk virus binds to H types 1/3 histo-blood group antigens on gastro-duodenal epithelial cells of 'secretor' individuals. *Gastroenterology* **122**, 1967–1977.
- Hutson, A. M., Atmar, R. L., Graham, D. Y. and Estes, M. K. (2002) Norwalk virus infection and disease is associated with ABO histo-blood group type. *J. Infect. Dis.* **185**, 1335–1337.
- Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B., Ruvoën-Clouet, N., Clément, M. and Le Pendu, J. (2001) ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie* **83**, 565–573.
- Lindensmith, L., Moe, C. L., Marionneau, S., Ruvoën-Clouet, N., Jiang, X., Lindblad, L., Stewart, P. A., Le Pendu, J. and Baric, R. S. (2003) Human susceptibility and resistance to Norwalk virus infection. *Nat. Med.* **9**, 548–553.
- Hutson, A. M., Airaud, F., Le Pendu, J., Estes, M. K. and Atmar, R. L. (2005) Norwalk virus infection associates with secretor status genotyped from sera. *J. Med. Virol.* **77**, 116–120.
- Hamosh, M. (1996) Protective function of proteins and lipids in human milk. *Biol. Neonate* **74**, 163–176.
- Bernt, K. and Walker, W. A. (2001) Human milk and the response of intestinal epithelium to infection. *Adv. Exp. Med. Biol.* **501**, 11–30.
- Brand Miller, J., Bull, S., Miller, J. and McVeagh, P. (1994) The oligosaccharide composition of human milk: Temporal and individual variations in monosaccharide components. *J. Pediatr. Gastroenterol. Nutr.* **19**, 371–376.
- Thurl, S., Henker, J., Siegel, M., Tovar, K. and Sawatzki, G. (1997) Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides. *Glycoconj. J.* **14**, 795–809.
- Chaturvedi, P., Warren, C. D., Altaye, M., Morrow, A. L., Ruiz-Palacios, G. M., Pickering, L. K. and Newburg, D. S. (2001) Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation. *Glycobiology* **11**, 365–372.
- Le Pendu, J. (2004) Histo-blood group antigens and milk oligosaccharides. *Adv. Exp. Med. Biol.* **554**, 135–143.
- Patton, S. and Huston, G. E. (1986) A method for isolation of milk fat globules. *Lipids* **21**, 170–174.
- Jiang, X., Wang, M., Graham, D. Y. and Estes, M. K. (1992) Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *J. Virol.* **66**, 6527–6532.
- Abouakil, N., Rogalska, E., Bonicel, J. and Lombardo, D. (1988) Purification of pancreatic carboxylic-ester hydrolase by immunoaffinity and its application to the human bile-salt-stimulated lipase. *Biochim. Biophys. Acta* **961**, 299–308.
- Burchell, J., Cendler, S., Taylor-Papadimitriou, J., Girling, A., Lewis, A., Millis, R. and Lampert, D. (1987) Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of human milk mucin. *Cancer Res.* **47**, 5476–5482.
- Moniaux, N., Varshney, G. C., Chauhan, S. C., Copin, M. C., Jain, M., Wittel, U. A., Andrianifahanana, M., Aubert, J. P. and Batra, K. (2004) Generation and characterization of anti-MUC4 monoclonal antibodies reactive with normal and cancer cell lines in humans. *J. Histochem. Cytochem.* **52**, 253–261.
- Gjellesvik, D. R., Lombardo, D. and Walther, B. T. (1992) Pancreatic bile salt dependent lipase from cod (*Gadus morhua*): purification and properties. *Biochim. Biophys. Acta* **1124**, 123–134.
- Newburg, D. S., Ruiz-Palacios, G., Altaye, M., Chaturvedi, P., Meinzen-Derr, J., Guerrero, I. and Morrow, A. L. (2004) Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhoea in breastfed infants. *Glycobiology* **14**, 253–263.
- Morrow, A., Ruiz-Palacios, G., Altaye, M., Jiang, X., Guemero, M., Meinzen-Derr, J., Farkas, T., Chaturvedi, P., Pickering, L. K. and Newburg, D. S. (2004) Human milk oligosaccharides are associated with protection against diarrhoea in breast-fed infants. *J. Pediatr.* **145**, 297–303.
- Ruiz-Palacios, G. M., Cervantes, L. E., Farnes, P., Chavez-Munguia, B. and Newburg, D. S. (2003) *Campylobacter jejuni* binds intestinal H(O) antigen and fucosyloligosaccharides of human milk inhibit its binding and infection. *J. Biol. Chem.* **278**, 14112–14120.
- Huang, P. W., Farkas, T., Marionneau, S., Zhong, W. M., Ruvoën-Clouet, N., Morrow, A., Pickering, L. K., Newburg, D. S., Le Pendu, J. and Jiang, X. (2003) Norwalk-like viruses bind to ABO, Lewis and secretor histo-blood group antigens but different strains bind to distinct antigens. *J. Infect. Dis.* **188**, 19–31.
- Hutson, A. M., Atmar, R. L., Marcus, D. M. and Estes, M. K. (2003) Norwalk virus-like particles hemagglutination by binding to H histo-blood group antigens. *J. Virol.* **77**, 405–415.
- Jiang, X., Huang, P., Zhong, W., Tan, M., Farkas, T., Morrow, A. L., Newburg, D. S., Ruiz-Palacios, G. and Pickering, L. K. (2004) Human milk contains elements that block binding of noroviruses to human histo-blood group antigens in saliva. *J. Infect. Dis.* **190**, 1850–1859.
- Sbarra, V., Mas, E., Henderson, T. R. and Walther, B. T. (1996) Digestive lipases of the newborn ferret: compensatory role of milk bile-salt dependent lipase. *Pediatr. Res.* **40**, 263–268.
- Lombardo, D. (2001) Bile salt-dependent lipase: its pathophysiological implications. *Biochim. Biophys. Acta* **1533**, 1–28.
- Gillin, F. D., Reiner, D. S. and Wang, C. S. (1963) Human milk kills parasitic intestinal protozoa. *Science (Washington, D.C.)* **221**, 1290–1292.
- Reiner, D. S., Wang, C. S. and Gillin, F. D. (1966) Human milk kills *Giardia lamblia* by generating toxic lipolytic products. *J. Infect. Dis.* **154**, 825–832.
- Hennell, O., Ward, H., Blackberg, L. and Pereira, M. (1986) Killing of *Giardia lamblia* by human milk lipases: an effect mediated by lipolysis of milk lipids. *J. Infect. Dis.* **153**, 715–720.
- Ali, S. A. and Hill, D. R. (2003) *Giardia intestinalis*. *Curr. Opin. Infect. Dis.* **16**, 453–460.
- Mas, E., Abouakil, N., Roudani, S., Miralles, F., Guy-Crotte, O., Figarella, C., Escribano, M. J. and Lombardo, D. (1993) Human foetal pancreatic protein: an oncofetal glycolipid of the normally secreted pancreatic bile-salt-dependent lipase. *Biochem. J.* **289**, 609–615.
- Mas, E., Sadoulet, M. O., El Ballari, A. and Lombardo, D. (1997) Glycosylation of bile-salt-dependent lipase (cholesterol esterase). *Methods Enzymol.* **284**, 340–353.
- Landberg, E., Huang, Y., Strömquist, M., Mechel, Y., Hansson, L., Lundblad, A., Novotny, M. V. and Pahlsson, P. (2000) Changes in glycosylation of human bile-salt-stimulated lipase during lactation. *Arch. Biochem. Biophys.* **377**, 246–254.
- Lindquist, S., Blackberg, L. and Hennell, O. (2002) Human bile salt-stimulated lipase has a high frequency of size variation due to a hypervariable region in exon 11. *Eur. J. Biochem.* **269**, 759–767.

- 40 Patton, S. (2001) MUC1 and MUC-X, epithelial mucins of breast and milk. *Adv. Exp. Med. Biol.* **501**, 35–45.
- 41 Price-Schiavi, S. A., Andrechek, E., Idris, N., Li, P., Rong, M., Zhang, J., Carraway, C. A., Muller, W. J. and Carraway, K. L. (2004) Expression, location and interactions of ErbB2 and its intramembrane ligand Muc4 (sialomucin complex) in rat mammary gland during pregnancy. *J. Cell. Physiol.* **203**, 44–53.
- 42 Moniaux, N., Escande, F., Porchet, N., Aubert, J. P. and Batra, S. K. (2001) Structural organization and classification of the human mucin genes. *Front. Biosci.* **6**, D1192–D1206.
- 43 Yolken, R. H., Peterson, J. A., Vonderfecht, S. L., Fouts, E. T., Midthun, K. and Newburg, D. S. (1992) Human milk mucin inhibits rotavirus replication and prevents experimental gastroenteritis. *J. Clin. Invest.* **90**, 1984–1991.
- 44 Schrotten, H., Hanisch, F. G., Plogmann, R., Hacker, J., Uhlenbruck, G., Nobis-Bosch, R. and Wahn, V. (1992) Inhibition of adhesion of S-fimbriated *Escherichia coli* to buccal epithelial cells by human fat globule membrane components: a novel aspect of the protective function of mucins in the nonimmunoglobulin fraction. *Infect. Immun.* **60**, 2898–2899.

Received 6 June 2005/19 September 2005; accepted 2 November 2005

Published as BJ Immediate Publication 2 November 2005; doi:10.1042/EJ20050898

## ***b - Conclusions et perspectives***

La présence d'oligosaccharides  $\alpha$ 2-fucosylés dans le lait de mère « sécréteur » a été corrélée à un risque plus faible d'infection à NoV chez les nouveau-nés allaités. Le rôle protecteur de ces molécules pourrait être expliqué par leur capacité à se lier aux virions. Elles joueraient ainsi le rôle de leurre et limiteraient l'infection en favorisant l'élimination des particules virales. Notre étude a eu pour vocation de mieux caractériser les molécules capables de se lier aux VLP de NV et d'estimer leur potentiel protecteur respectif.

Pour cela, le caractère sécréteur ou non sécréteur des différents laits étudiés a été déterminé grâce au génotypage de la position 428 du gène *FUT2*. La mutation G428A est en effet la mutation inactivante du gène *FUT2* la plus fréquemment rencontrée dans la population caucasienne. La présence de l'HBGA H type 1 dans les laits a ensuite été vérifiée par ELISA en utilisant un anticorps monoclonal anti-H type 1/Le<sup>b</sup> et par l'utilisation des laits dans des essais d'inhibition de la fixation des VLP de NV sur des salives et des coupes histologiques d'individus de phénotype sécréteur.

Dans un premier temps, la dénaturation des protéines des laits a permis d'écarter toute inhibition de la fixation induite par des interactions protéine/protéine, et en particulier la reconnaissance potentielle des épitopes viraux par des immunoglobulines A (IgA). Les inhibitions observées par la suite ont donc pu être attribuées uniquement à des interactions protéine/glycane. Les oligosaccharides libres, ainsi que la fraction crémeuse du lait n'ont montré aucune propriété inhibitrice sur la fixation du NV. Les globules gras micellaires possèdent quant à eux un certain potentiel inhibiteur mais qui est relativement faible comparé à celui de la fraction du lait contenant les glycoprotéines. Cette fraction inhibe en effet la fixation des VLP avec une efficacité comparable à celle observée par l'utilisation du lait non-traité. Les western blots réalisés sur cette dernière fraction ont montré que le NV se fixe sur des glycoprotéines de très hauts poids moléculaires ainsi que sur une autre glycoprotéine d'environ 120 kDa. Identifiées respectivement comme les protéines Muc1, Muc4 et la lipase stimulée par les sels biliaires (BSSL), nous avons mis en évidence que l'inhibition de la fixation des VLP est principalement due à la BSSL (60%) et que la participation des mucines Muc1 et Muc4 semble moindre (40%).

La BSSL est une protéine du lait maternel principalement connue pour faciliter la digestion des lipides par le nouveau-né. Nous avons montré dans cette étude que lorsqu'elle est produite chez des mères de phénotype sécréteur, elle peut en plus participer à l'inhibition de la fixation du NV sur son récepteur. Cette propriété lui confère

donc un second rôle, important cette fois dans l'immunité innée antivirale transmise de la mère à l'enfant.

L'infection par le NV étant génétiquement liée à l'expression de l'HBGA H type 1, les mères de phénotype sécréteur sont les seules à pouvoir disposer de titres élevés en IgA anti-NV dans leur lait (X Jiang et coll. 2004). Ces immunoglobulines pourraient agir de concert avec la BSSL, Muc1 et Muc4 pour protéger le nouveau-né vis-à-vis de l'infection par le NV. Le lait de ces mères présente donc un double avantage comparé à celui des mères de phénotype non sécréteur, celui-ci étant intimement lié à l'expression de l'antigène H type 1.

Toutefois, leurs effets protecteurs respectifs ne sont que temporaire car les titres en IgA diminuent très rapidement au cours du premier mois *post partum* (A L Morrow et Rangel 2004) alors que ceux des oligosaccharides  $\alpha$ 2-fucosylés décroissent progressivement durant des 18 premières semaines de l'allaitement (Chaturvedi, Warren, Altaye, et coll. 2001). Après cette période, le lait maternel ne semble théoriquement plus pouvoir protéger l'enfant contre l'infection par le NV. C'est pourquoi il paraît judicieux de développer la production de glycoprotéines synthétiques de type mucines comme MUC1, MUC4 ou la BSSL, afin de compléter les laits infantiles et ainsi protéger le jeune enfant pendant les premiers mois de sa vie.

## **2.2. Article 2 : Association entre expression de l'antigène tissulaire de groupe sanguin H, polymorphisme des $\alpha$ 1,2fucosyltransférases chez le lapin de garenne et sensibilité au RHDV**

### **a - Introduction**

Le RHDV est responsable chez le lapin (*Oryctolagus cuniculus*) d'une pathologie sévère pouvant conduire jusqu'à 100% de mortalité dans les populations. La mort intervient généralement dans les 24 à 48 heures post-infection, du fait d'une hépatite fulgurante associée à des hémorragies disséminées dans l'organisme (Marcato et coll. 1991). La mise au point rapide d'un vaccin a néanmoins permis de maîtriser la maladie dans les élevages cynicoles, mais l'impossibilité de vacciner les populations sauvages a entraîné localement la régression de l'espèce au fur et à mesure des épidémies (Marchandeau et coll. 2000).

L'étude de la capacité du RHDV à agglutiner les érythrocytes humains a permis d'identifier un récepteur cellulaire potentiel du virus : l'HBGA H type 2 (Ruvoën-Clouet et coll. 2000). Le NV qui est également un Calicivirus utilise une structure assez proche comme récepteur, l'HBGA H type 1 (Marionneau et coll. 2002). Il a été montré pour ce dernier que les individus de phénotype non sécréteur sont génétiquement résistants à l'infection, du fait de la présence dans leur génome de deux copies inactives du gène *FUT2* (Lindesmith et coll. 2003). Il a été observé dans la nature que les jeunes lapereaux (<2 mois) sont naturellement résistants au RHDV (Morisse et coll. 1991) et cette résistance apparente pourrait être liée à la faible expression de l'HBGA H type 2. En effet, les lapereaux acquièrent progressivement l'expression de cet HBGA avec l'âge. De plus, sur des coupes histologiques de duodénum, une corrélation entre l'âge des animaux, le niveau d'expression de l'antigène H type 2 et l'intensité de la fixation du virus a pu être observée. Chez le lapin, l'expression de l'HBGA H type 2 au niveau épithélial (phénotype sécréteur) est sous la dépendance des gènes *Fut2* (homologue du gène *FUT2* humain) et potentiellement *Sec1* (pseudogène chez l'homme) (Hitoshi et coll. 1995, 1996). Chez cette espèce, le gène *Fut1* contribue à la synthèse de ce glycanne au niveau neuronal.

Nous avons alors émis l'hypothèse que les lapins qui résistent à une épidémie de RHDV seraient susceptibles de présenter un phénotype non sécréteur ou sécréteur faible. Pour la vérifier, nous avons phénotypé des populations de lapins aux passés épidémiques différents. Nous avons ensuite séquencé les gènes *Fut2*, *Sec1* et *Fut1* dans la population ayant subi la plus forte pression de sélection par le RHDV, pour y estimer le polymorphisme de ces trois gènes et rechercher s'il existait une association entre

polymorphisme et survie. Nous avons finalement testé *in vivo* l'activité enzymatique du produit des différents allèles *Fut2* et *Sec1* identifiés.

**ASSOCIATION BETWEEN EXPRESSION OF THE H HISTO-BLOOD GROUP  
ANTIGEN,  $\alpha$ 1,2FUCOSYLTRANSFERASES POLYMORPHISM OF WILD RABBITS  
AND SENSITIVITY TO RHDV**

**GLYCOBIOLOGY, 2009 Jan, 19 (1) : 21-8**

Patrice GUILLON<sup>2</sup>, Nathalie RUVOËN-CLOUET<sup>2,3</sup>, Béatrice LE MOULLAC-VAIDYE<sup>2</sup>,  
Stéphane MARCHANDEAU<sup>4</sup>, Jacques LE PENDU<sup>1,2</sup>

<sup>2</sup> INSERM, U892, Nantes, France ; Université de Nantes, Nantes, France.

<sup>3</sup> Ecole Nationale Vétérinaire de Nantes, Nantes, France

<sup>4</sup> Office National de la Chasse et de la Faune Sauvage, Nantes, France

Key words : RHDV / histo-blood group antigens / fucosyltransferase / polymorphism /  
resistance to pathogens

Running title : Genetics of wild rabbit resistance to RHDV

<sup>1</sup> To whom correspondance should be addressed:  
INSERM, U892, Institut de Biologie, 9 Quai Moncoussu  
44093, Nantes, Cedex 01, France  
Tel : 33 240 08 40 99, Fax : 33 240 08 40 82  
E-mail : jlependu@nantes.inserm.fr

## Abstract

RHDV is a highly virulent calicivirus that has become a major cause of mortality in wild rabbit populations (*Oryctolagus cuniculus*). It binds to the histo-blood group antigen (HBGA) H type 2 which requires an  $\alpha$ 1,2fucosyltransferase for its synthesis. In rabbit, three  $\alpha$ 1,2fucosyltransferases genes are known, *Fut1*, *Fut2*, and *Sec1*. Non functional alleles at either of these loci could potentially confer resistance to RHDV, similar to human *FUT2* alleles that determine the nonsecretor phenotype and resistance to infection by various NoV strains. In this study, we looked for the presence of H type 2 on buccal epithelial cells of wild rabbits from two geographic areas under RHDV pressure and from one RHDV free area. Some animals with diminished H type 2 expression were found in the three populations (nonsecretor-like phenotype). Their frequency markedly increased according to RHDV impact, suggesting that outbreaks selected survivors with low expression of the virus ligand. Polymorphisms of the *Fut1*, *Fut2*, and *Sec1* coding regions were determined among animals who either died or survived outbreaks. The *Fut2* and *Sec1* genes presented a high polymorphism and the frequency of one *Sec1* allele was significantly elevated, over 6 fold, among survivors. *Sec1* enzyme variants showed either moderate, low or undetectable catalytic activity, whereas all variant *Fut2* enzymes showed strong catalytic activity. This functional analysis of the enzymes encoded by each *Fut2* and *Sec1* allele suggests that the association between one *Sec1* allele and survival might be explained by a deficit of  $\alpha$ 1,2fucosyltransferase expression rather than by impaired catalytic activity.

## Introduction

RHDV (rabbit hemorrhagic disease virus) is a non envelopped single positive strand RNA virus belonging to the lagovirus genus of the *Caliciviridae* family. It is an emergent virus responsible for a devastating disease of European rabbits (*Oryctolagus cuniculus*), called rabbit hemorrhagic disease (RHD) or rabbit calicivirus disease (RCD), characterized by a fulminating hepatitis accompanied by disseminated intravascular coagulation (Marcato et al. 1991). Death usually occurs within 48 h post-infection. RHD was first reported in 1984 in rabbitries in China where it caused the death of millions of domestic animals. It most likely originated from animals imported from Germany. In 1987, it appeared in Italy and broke out simultaneously in several other European countries, including France in 1988. It rapidly expanded in wild rabbit populations and has become endemic (Mitro and Krauss 1993). Its impact on wild animals has been reported in Spain, France and Australia. In this last country, RHDV was introduced to control rabbit populations since they threaten many endogenous plant and animal species and cause huge economic losses. These studies indicated that during the first epizootic, mortality rates can be extremely high, ranging from over 50% to 90% (Marchandeau et al. 1998, Mutze et al. 1998, Villafuerte et al. 1995, Villafuerte et al. 1994). At present, fighting the disease is possible in husbandry only, prophylaxy being based on systematic individual vaccination. Non pathogenic strains of lagoviruses infecting European rabbits have also been described and evidence of serological relatedness with RHDV has been obtained, although the existence of cross-protection with virulent strains remains uncertain (Capucci et al. 1996, Marchandeau et al. 2005, Moss et al. 2002).

In wild rabbit populations, RHD is considered as a major cause of the species regression, particularly in dry areas of the South of France and the Iberian Peninsula (Marchandeau 2000, Marchandeau et al. 2000). In those areas, the drastic decrease of rabbit numbers threatens several other species that depend upon rabbit for their survival, such as a lynx species (*Lynx pardinus*) and two eagle species (*Aquila aldaberti* and *Hieraaetus fasciatus*). In France, the regression of a lizard species (*Lacerta lepida*)

which uses rabbit warrens during hot summer periods, has also been noticed in areas where wild rabbits have become rare.

We observed earlier that RHDV binds to a trisaccharide present on the surface epithelial cells of the upper respiratory and digestive tracts, which are natural doors of entry for the virus (Ruvoën-clouet et al. 2000). The trisaccharide,  $\text{Fuc}\alpha 2\text{Gal}\beta 4\text{GlcNAc}\beta\text{-R}$ , is a member of the histo-blood group antigen (HBGA) family called H type 2 and exists in other mammalian species, including human, although its tissue distribution may vary in a species-specific manner. Its synthesis requires a fucosyltransferase that catalyses addition of a fucose residue in  $\alpha 1,2$  linkage onto the galactose of the precursor unit (Marionneau et al. 2001). Three rabbit genes encode such enzymes : *Fut1*, *Fut2* and *Sec1*. Only *Sec1* and *Fut2* are expressed in epithelial cells, *Fut1* mRNA being detected in the brain only (Hitoshi et al. 1995, Hitoshi et al. 1996). In human, *Sec1* is a pseudogene and the polymorphism of the *FUT2* gene, orthologous to the rabbit *Fut2* gene, controls sensitivity or resistance to several caliciviruses belonging to the norovirus genus (Le Pendu et al. 2006, Tan and Jiang 2005). It was shown that these viruses, which cause acute gastroenteritis, bind to human digestive surface epithelial cells using  $\alpha 1,2$ fucosylated structures closely related to that used by RHDV for attaching to rabbit epithelial cells. Addition of the  $\alpha 1,2$ -linked fucose residue to human gastric and duodenal surface epithelial cells is catalyzed by the FUT2 enzyme. The *FUT2* gene is polymorphic with some alleles encoding a functional enzyme and other alleles being either null alleles or alleles encoding enzymes with low activity. Functional alleles confer the so-called secretor phenotype, characterized by the presence of  $\alpha 1,2$ fucosylated structures on various epithelial cell types or in secretions such as saliva. By contrast, null or weak alleles in the homozygous state confer the nonsecretor phenotype, characterized by the absence of such structures (Oriol et al. 2000). Nonsecretor individuals are resistant to infection by strains of norovirus which bind to  $\alpha 1,2$ fucosylated ligands (Hutson et al. 2005, Kindberg et al. 2007, Lindesmith et al. 2003, Thorven et al. 2005). Since RHDV binds to a similar glycan, by analogy with the observations on human-norovirus interaction, we hypothesized that rabbits devoid of H type 2 on their epithelial cells

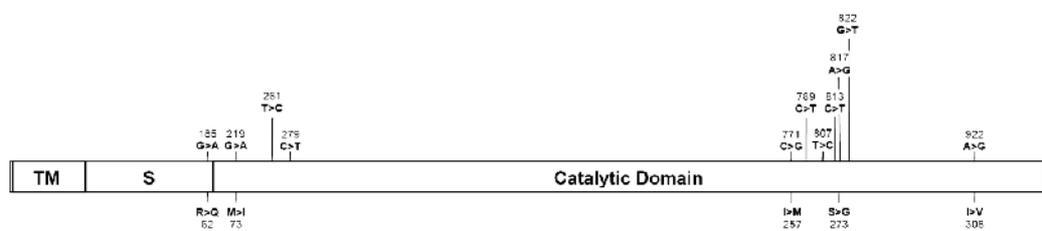
should be resistant to infection by RHDV. Given the very high impact of RHD on some European rabbit populations, if such nonsecretor-like animals exist, one expects that they should be positively selected by the virus and therefore that their frequency would increase in populations that have experienced RHD outbreaks. The  $\alpha$ 1,2 fucosyltransferase polymorphism should thus provide a mechanism of protection against RHDV at the species level. To put this hypothesis to the test, the expression of the H type 2 ligand and the polymorphism of  $\alpha$ 1,2fucosyltransferases genes were determined in wild rabbit populations which have been closely monitored with respect to RHDV impact.

## Results

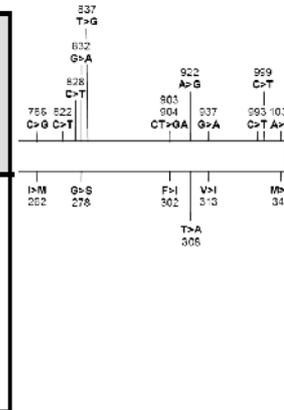
In human, the secretor or nonsecretor phenotypes have classically been defined by the presence of ABH antigens on salivary mucins. Practically, binding of the UEA-I lectin to human saliva characterizes the secretor phenotype, whereas absence of binding characterizes the nonsecretor phenotype. For other mammals, collection of saliva is either unpractical or salivary mucins do not express ABH antigens, even though they may be expressed by various epithelial cell types. Nevertheless, the expression of these antigens can easily be determined from buccal epithelial cells collected using cotton swabs. We thus collected buccal epithelial cells from captured wild European rabbits and looked for the presence of UEA-I binding sites on these cells. Epithelial cells recovered from the swabs are generally mixed up with debris of food stuff and sometimes blood cells due to some bleeding during the sampling process. This precluded a quantitative analysis of UEA-I binding site expression on buccal epithelial cells by flow cytometry. However, under the microscope, such cells have a typical morphology and can be readily distinguished. We thus chose to estimate their UEA-I labeling visually in a semi-quantitative manner. We observed a variable intensity of labeling from sample to sample. Some were clearly strongly positive, others completely negative, while others yet were weakly positive and classified as doubtful. Although the doubtful group was somewhat arbitrarily defined, classification into the 3 groups, positive, doubtful and negative, was consistent among either 2 or 3 independent observers. The positive animals were

considered to have a secretor phenotype, whereas the negative and doubtful animals were pooled into a single nonsecretor-like group. Analysis of the frequency of the two phenotypes in three populations differentially affected by RHD showed significant differences from a population to another. Less than 15% animals from the Dompierre site, where no RHD outbreak was ever recorded, presented a nonsecretor-like phenotype. None were completely UEA-I negative. By contrast, at the Cerisay site, where the epizootic killed about half of the population, almost 50% individuals had the nonsecretor-like phenotype and even more striking, at the Chèvreloup site, where RHD killed about 90% of the population in 1995, over 70% animals presented the nonsecretor phenotype (Table 1). The data are consistent with a significant association between the impact of phenotype.

Polymorphi. **Fut2**



Population	% nonsecretor-like (n samples)	Impact of RHD
Dompierre	14,8 (27)	0
Cerisay	47,1 (70)	~50
Chèvreloup	71,8 (32)	~90



**Table 1** : Individuals with the nonsecretor-like phenotype were defined by either a negative or a weak UEA-I lectin labeling of buccal epithelial cells. The number of samples analyzed for each site is given in parentheses. The impact of RHD is given as estimates of mortality rates due to the disease (Marchandeau et al. 2000, Marchandeau et al. 1998). Analysis of the frequencies of the nonsecretor-like phenotype between populations was performed by a Chi<sup>2</sup> test (p=0,00007); Paired comparisons between populations: \*p<0.01, \*\*p<0.02, \*\*\*p<0.001.

With the aim to determine which genetic polymorphism was responsible for the nonsecretor-like phenotype, we sequenced the coding regions of the three rabbit  $\alpha$ 1,2fucosyltransferases genes, *Fut1*, *Fut2* and *Sec1* from animals of the Chèvreloup population which had been sampled in 1994, most of whom were alive at the start of the 1995 devastating outbreak, as well as from animals sampled soon after the 1995-1996

outbreaks. Sequences from a total of 83 animals were obtained revealing an important polymorphism quite different among the three genes. Thus, over about 1 kb sequenced for each gene, 1, 11 and 26 variable positions were detected for *Fut1*, *Fut2* and *Sec1*, respectively. The unique *Fut1* nucleotide change is silent and all wild rabbit *Fut1* genes differed from the *Fut1* reference sequence (Hitoshi et al. 1995) by one non silent mutation. Five and 14 amino-acid changes were detected in *Fut2* and *Sec1* respectively. Interestingly, the *Fut2* mutations are clustered in two regions of the coding sequence while the *Sec1* mutations were spread more evenly (Fig. 1).

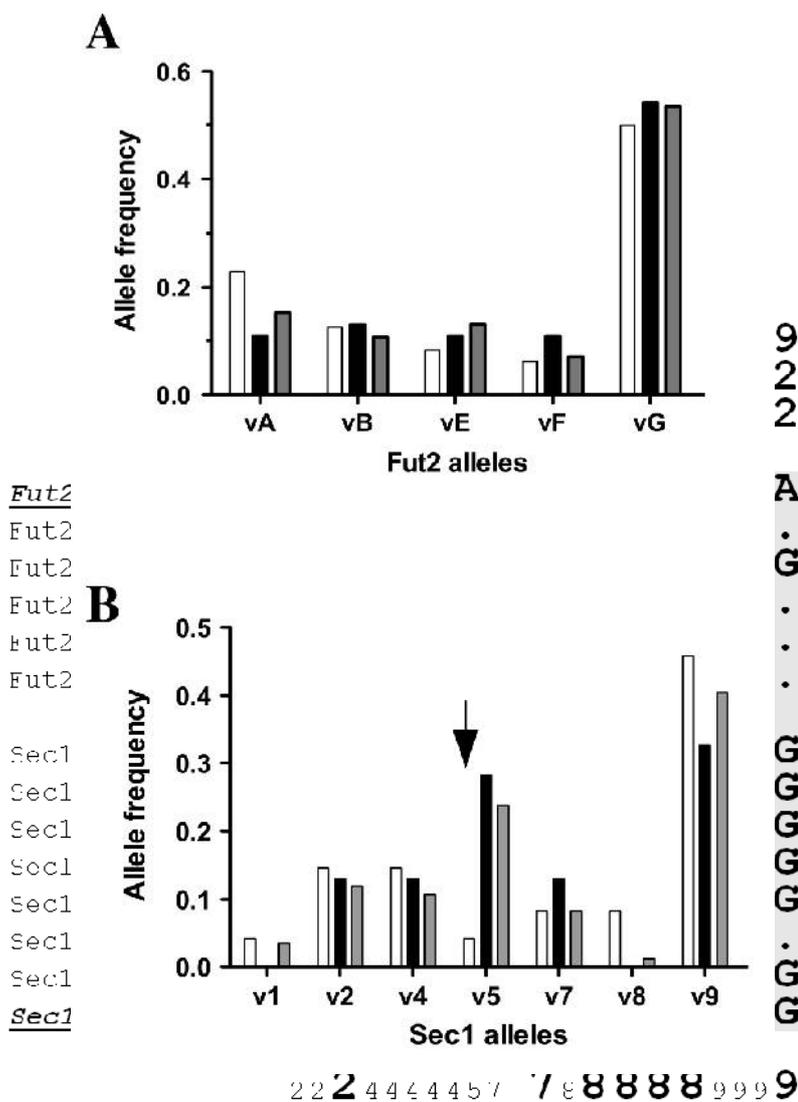
The variable positions define 2 alleles for *Fut1*, 5 alleles for *Fut2* and 7 alleles for *Sec1*. They were denoted alphabetically for *Fut1* (variants  $\alpha$  and  $\beta$ ) and *Fut2* (i.e. vA for variant A), and numerically for *Sec1* (i.e. v1 for variant 1). An alignment of all sequences revealed shared polymorphisms between these 3 genes. Thus, 7 polymorphisms localized in the region coding for the enzymes catalytic domains were shared between *Fut2* and *Sec1* (Fig. 2).

Considering the very high homology between these two genes in that region, the shared polymorphisms are clear signs of gene conversion events. Likewise the single *Fut1* polymorphic site is shared with *Sec1* since sequence alignment indicates that the synonymous *Fut1* 780 C>G mutation corresponds to the *Sec1* 750 C>G mutation.

**Figure 1** : : Distribution of the wild rabbit *Fut2* and *Sec1* polymorphisms along coding region. The 2 genes are aligned and corresponding protein domains are shown in open bars, nucleotide variations are shown above the bars and peptide variations are below the bars. Positions numbering begins at the start codons and first methionine, respectively. TM : transmembrane domain, S : stem region.

The frequency of each *Fut1*, *Fut2* and *Sec1* allele among the animals who either died or survived during the epizootic were then compared (Fig. 3). Allelic frequencies were quite variable with some alleles highly represented, such as *Fut2* vG or *Sec1* v9, and some alleles, such as *Fut2* vF or *Sec1* v1, being of low frequency. Some additional alleles found at other locations on the French territory, such as *Fut2* vC, were completely absent from the Chèvrelope population, indicating that there existed some population differentiation among wild rabbits prior to the emergence of RHD in France (not shown).

When the animals were subdivided into deceased and survivors, some minor differences in allelic frequencies between the two groups occurred for *Fut2*, but none reached statistical significance (Fig. 3A). Likewise, there was no difference in the frequencies of the two *Fut1* alleles between survivors and non survivors (not shown). By contrast, the *Sec1* v5 allele was significantly over-represented among survivors to the 1995-1996 RHD outbreaks (Fig. 3B). The increase was as high as 6.8 fold when considering the survivors alone, or 5.7 fold after pooling the survivors and the probable survivors groups.

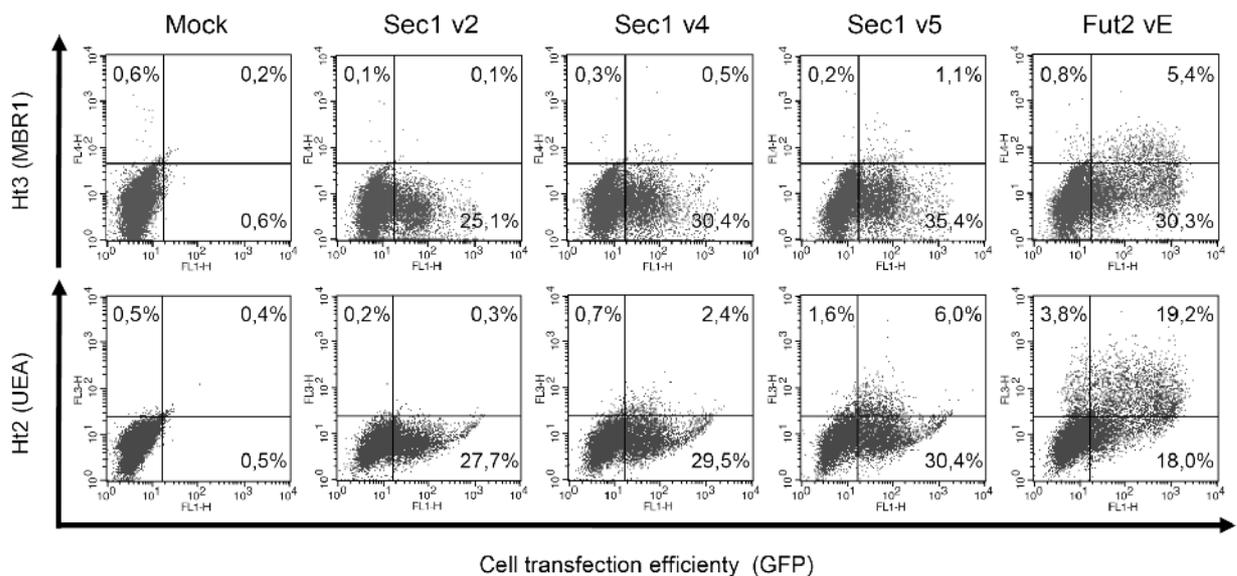


**Figure 3 :** Comparison of the frequencies of *Fut2* alleles (A) and *Sec1* alleles (B) between animals which died during the RHDV outbreak (white bars), survivors (black bars) and a group composed of survivors plus probable survivors (grey bars). Frequency of the *Sec1* v5 allele (arrow) is significantly increased among survivors ( $p=0.0056$ ) or among survivors and probable survivors ( $p=0.0098$ ) compared to non survivors (likelihood ratio  $\chi^2$  with correction for multiple analysis).

**Figure 2 :** Characteristic nucleotide variations of the wild rabbit *Fut2* and *Sec1* alleles. All variants were aligned with published reference sequences and orthologous variable positions are shown. *Fut2* alleles are designated alphabetically (A-E), *Sec1* alleles are designated numerically (1-9). Positions presenting a shared polymorphism are in bold. Position numbering begins at the start codon..

### Functional analysis of the $\alpha$ 1,2fucosyltransferases allelic variants

Since our working hypothesis was that  $\alpha$ 1,2fucosyltransferases allelic variants would encode non functional enzymes unable to synthesize the virus ligand, we analyzed the activity of each Fut2 and Sec1 enzyme variant. *Fut1* is known not to be expressed in epithelial tissues (Hitoshi et al. 1995) and all animals presented the same single variation at the amino-acid level compared to the reference sequence. Thus, the wild rabbit Fut1 catalytic activity was not analysed since it was not informative. All the *Fut2* and *Sec1* variants were cloned into an expression vector so as to generate N-terminal GFP fusion proteins. After transfection into CHO cells, which are devoid of endogenous  $\alpha$ 1,2fucosyltransferase activity, the presence of the  $\alpha$ 1,2 fucosylated structures at the cell surface was measured by flow cytometry using the UEA-I lectin that detects H type 2 and a mAb that detects H type 3. Like most cultured cell lines, CHO cells do not express type 1 precursor and therefore the synthesis of H type 1 was not looked for. The presence of the GFP tail allowed normalization of the transfection and protein expression efficacy (Fig.4).



**Figure 4** : Flow cytometry analysis of the wild rabbit *Sec1* and *Fut2* allelic variant enzymes activities. Enzymes were produced by transfection in CHO cells as N-term GFP fusion proteins. Synthesis of cell surface H type 2 and H type 3 was determined using the UEA-I lectin (lower panel) and the Mbr1 mAb (upper panel), respectively and is given by the percentage of positive cells in the upper right and left quadrants. Fluorescence of the GFP protein allows normalisation of the transfection efficiency (lower and upper right quadrant values). Examples of inactive (*Sec1v2*), very weakly active (*Sec1v4*), moderately active (*Sec1v5*) are shown in comparison to a fully active enzyme *Fut2 vE* that corresponds to the *Fut2* reference sequence.

## Discussion

Numerous studies concerning host-pathogen relationships suggest that genetic diversity of the host population is essential to prevent evolution of the pathogen toward maximal virulence. Direct proofs of a benefit conferred by this diversity have only been obtained in a restricted number of instances such as that of the Major Histocompatibility Complex (MHC) (Woolhouse et al. 2002). Recently, several NoV strains were shown to use  $\alpha$ 1,2fucosylated structures as ligands and polymorphisms of the *FUT2*  $\alpha$ 1,2fucosyltransferase gene and of the *ABO* gene were shown to control sensitivity or resistance to infection (Harrington et al. 2002, Hutson et al. 2005, Hutson et al. 2002, Kindberg et al. 2007, Lindesmith et al. 2003, Marionneau et al. 2002, Thorven et al. 2005). We suggested that this situation resulted from a co-evolution between these caliciviruses and their human host, whereby the host genetic diversity would both have been maintained by the pathogen selective pressure and contributed to force the virus

**Figure 4** : Relative catalytic activities of wild rabbit *Fut2* allelic variants (**A** and **B**) and *Sec1* allelic variant (**C** and **D**) as determined by flow cytometry (see Fig. 4). The ratios between the percentage of GFP fluorescent cells and the percentages of UEA-I (H type 2) fluorescent cells, (A, C) or Mbr1 (H type 3) fluorescent cells (B,D) indicate the relative activity of each variant. In panels C and D the activity of a *Fut2* variant (*Fut2* vE) assayed in the same experiment is added for comparison. The results represent values from one representative experiment out of at least two for each variant.

evolution toward its present low or moderate virulence (Le Pendu et al. 2006). Owing to their historical aspect, testing such hypotheses is particularly difficult, but the emergence of RHDV in wild European rabbit populations could provide the opportunity to validate some assumptions of the hypothesis. Indeed, RHDV binds to fucosylated glycans similar to those used by several NoV strains (Rademacher et al. 2008, Ruvoën-clouet et al. 2000). Because neither human NoV strains, nor RHDV are cultivable, it has been difficult to demonstrate that histo-blood group antigens constitute functional receptors. Thus at this stage, it has not been proven that RHDV binding to H type 2 is necessary for infection. However, the similarity with the NoV-human situation and the observation that young rabbits below two months of age are resistant to RHD (Morisse et al. 1991, Rodak et al. 1991) and express low levels of H type 2 make it highly plausible that it is a good candidate as a primary receptor (Ruvoën-clouet et al. 2000). If this was the case, RHDV virulence is expected to exert a very strong selective pressure on its host so as to

positively select animals with diminished expression of H type 2. Our prediction is that when nonsecretor-like animals will have reached a significant frequency, they will contribute to decrease the virus transmission in rabbit populations, favoring mutant strains with decreased virulence because of a trade-off between virulence and transmission as indicated by mathematical modeling (Fouchet et al. Submitted). Such nonsecretor-like animals were thus searched in wild rabbit populations differentially affected by RHD. Owing to the difficulty to obtain tissues from these animals, hunting in heavily affected populations being now restricted, the phenotyping was performed on buccal epithelial cells. Since it is no longer allowed to recapture animals from these populations, DNA could not be obtained from the same animals as the cellular material collected on swabs was entirely used for phenotyping. Nevertheless, for the first time, a nonsecretor-like phenotype was described in a wild mammalian species. So far a deficit of  $\alpha$ 1,2fucosylated HBGAs in secretions or on epithelial cells had been observed in human and pigs only (Coddens et al. 2007, Oriol et al. 2000). Most interestingly, the frequency of animals with this phenotype was significantly associated with the impact of RHDV outbreaks, providing indirect evidence that the H type 2 ligand is necessary for infection and that its absence or low expression should confer resistance to the disease, similar to what has been observed between NoV and human HBGAs expression. In addition, it indicates that RHDV may have selected animals of the nonsecretor-like phenotype. French wild rabbit populations show only moderate differentiation as determined from microsatellite markers (Queney et al. 2001). Nevertheless, it cannot be completely ruled out that the differential frequency of the nonsecretor-like phenotype among the three populations that we analysed could have preexisted RHDV emergence. It was therefore important to ascertain that polymorphism at the  $\alpha$ 1,2fucosyltransferases loci was associated with survival to RHD. Although not systematically searched in rabbit breeds, we failed to observe the nonsecretor-like phenotype among 30 laboratory animals (data not shown). Fortunately, DNA samples from animals captured either before or soon after the 1995-1996 outbreaks at the Chèvreloup site were available. Monitoring data from these animals were also available, allowing to distinguish survivors and probable

survivors from non survivors (in practice, animals whose tract was lost immediately after an outbreak). The coding region of the three rabbit  $\alpha$ 1,2fucosyltransferases genes were thus sequenced. We chose to sequence those regions only since the non coding regions of the rabbit genes are unknown as yet and since in human all known mutations that generate either null or weak alleles affect the coding regions. Surprisingly, very distinct polymorphisms were found between the three genes. *Fut1* presented almost no variation whereas a large number of mutations were found in *Fut2* and even more in *Sec1*. In addition, the distribution of the mutations was quite different between *Fut2* and *Sec1*. Shared polymorphisms signed the existence of gene conversion which leads to convergent evolution of paralogues and increases genetic diversity (Teshima and Innan 2004). The pattern of diversity that we observed is thus consistent with a possible role for *Fut2* and *Sec1* expressed in epithelial cells in coping with the environment. By contrast, the near absence of variation in *Fut1*, which is expressed in the brain only (Hitoshi et al. 1995), suggests that it undergoes a strong negative selection to maintain functionality. Nevertheless, functional analysis of Fut2 enzyme variants revealed that all were equally active. In accordance with the first description of rabbit  $\alpha$ 1,2fucosyltransferases activity (Hitoshi et al. 1996), we observed that the reference *Sec1* (v5) was functional but that its activity was lower than that of any *Fut2* variant. Other variants of *Sec1* showed even lower activity or no activity at all, suggestin that the *Sec1* gene of rabbits is on its way to pseudogenisation, similar to what happened to the apes, pig and mice orthologues. Yet, the *Sec1*v5 allele was strongly associated with resistance to RHD since its frequency was largely higher among survivors than among non survivors. Along with the association between the nonsecretor-like phenotype and the impact of RHD, this observation suggests that decreased expression of  $\alpha$ 1,2fucosylated ligands determines resistance to RHDV. Nevertheless, the *Sec1*v5 allele is not a null allele and it is always associated with a functional *Fut2*. It is therefore a marker of survival but cannot be the cause of the nonsecretor-like phenotype or of the resistance to RHDV. The three genes *Fut1*, *Fut2* and *Sec1* are closely located in the human, mouse, rat, cow and pig genomes. Their genomic organisation has not been

determined in rabbit as yet, but it is likely similar to that of other species. The ORFs of *Fut2* and *Sec1* are separated by about 22-28 kb with a *Fut2* untranslated 5' exon lying at near equidistance between the two genes. We therefore hypothesize that in the Chèvreloup population, the Sec1v5 allele is in linkage disequilibrium with a polymorphism affecting *Fut2* expression which would be localized in its regulatory region located near *Sec1*. So far polymorphisms of HBGAs expression in human have been linked to mutations in coding regions of glycosyltransferases genes. However, mutations in regulatory regions affecting gene transcription could similarly affect HBGAs expression. The Sec1v5 variant would thus be genetically linked to such a regulatory mutation in the  $\alpha$ 1,2 fucosyltransferases loci rather than directly causative of the nonsecretor-like phenotype. Further work is underway to validate this hypothesis.

Regardless, the present work provides the first evidence of an  $\alpha$ 1,2fucosyltransferase gene polymorphism associated with survival to a viral disease and of a non-secretor like phenotype in a wild mammalian population which frequency is also associated with the impact of the disease. The latter aspect is suggestive of a process of selection of resistant animals taking place *in naturo*. Although limited by the availability of the biological material (wild rabbits cannot be sampled freely and are extremely difficult to breed), these observations represent an important step in the analysis of the evolutionary relationship between HBGAs and caliciviruses.

## Materials and methods

### *Sampling wild rabbits*

Animals were caught by trapping or using ferrets (*Mustela furo*). Buccal epithelial cells were collected using cotton swabs from animals captured at 3 different locations on the French metropolitan territory in 2004. The 3 sites were chosen for their very different situation with regard to RHD impact. When sampled, the Dompierre site had never been affected by the disease. In 2000 at the Cerisay site, an outbreak caused a 50% mortality (Marchandeau et al. 2005). Finally the wild rabbit population of the Chèvreloup site

experienced 2 consecutive outbreaks in 1995 and 1996. In 1995, annual mortality rates at that site were estimated as high as 88% in adults and 99% in juveniles (Marchandeau et al. 1998).

In addition, DNA samples collected in 1994 and 1996, i.e. before and after the RDH outbreaks, during a study on the genetic structure of the Chèvrelope population were used (Queney et al. 2000, Queney et al. 2001). The Chèvrelope population being monitored for a long-term survey, the rabbits were individually marked with reflecting ear tags which allowed individual recognition from a distance and thus a follow up of each animal during and after the outbreaks (Marchandeau et al. 2000). Three methods of recapture (ferreting, trapping and spotlighting) were used. Thus, from that site, 3 groups of animals could be defined. The first group includes rabbits sampled in 1994 which survived the epizootic (n=23). The second group includes animals sampled in 1994 and presumed to have died during either the 1995 or the 1996 outbreak (n=24). The third group includes animals first captured soon after the 1996 outbreak who presented adult morphological features and body weight and therefore considered as probable survivors (n=20). One may notice that in wild populations the fate of each animal is almost impossible to record. Therefore, the classification of the rabbits into three classes imperfectly reflects their sensitivity or resistance to the disease. Some rabbits that survived to the 1995-1996 outbreaks may have not been exposed to the virus and some that died may have died from other causes than RHD.

#### *Determination of the nonsecretor-like phenotype*

Cotton swabs were dipped in 1 ml PBS and then centrifuged at 1500 rpm, room temperature for 5 min. The cell pellet was collected and incubated for 1 h at room temperature with the UEA-I-FITC conjugated lectin (Sigma, St Louis, MO). The lectin specifically binds to the H type 2 motif  $Fuc\alpha 2Gal\beta 4GlcNAc\beta -R$  recognized by RHDV. Cells were then washed thrice with PBS, and nuclei were stained with propidium iodide prior to observation under an epifluorescence microscope (Olympus). Membrane fluorescence of cells with a clear morphology of buccal epithelial cells was recorded by 2 or 3

independent observers without prior knowledge of the animals site of origin and qualitatively graded as strongly positive, weakly positive or negative.

#### *Sequencing of the rFut1, rFut2 and rSec1 open reading frames*

DNA extractions were performed from ear fragments using the QIAamp DNA Mini kit (Qiagen). The coding regions of the *Fut1*, *Fut2* and *Sec1* rabbit genes were amplified using primers deduced from the X80226, X91269 and X80225 sequences, respectively : gen *Fut2* (CCCAAGAGCT AACTAACATC CCTGCTG) and gen *Fut2* inv (GCAAGAAACG GGTCTTGGCC CACGTGAC) for *Fut2*, gen *Sec1* (CTTCAGCCTG CACAGCCCCA GCCGTT) and gen *Sec1* inv (ACTCCGTGTG GGATCCCGGG TCCATGC) for *Sec1*, gen *Fut1* (CTGTCTGGAA AGCCATCATC) and gen *Fut1* inv (GTAGAATCAC TCTGGCTGTG) for *Fut1*. Amplifications were performed in the presence of 2.5% DMSO using GoTaq (Promega). Sequencing was performed on an Applied Biosystems 3730 sequencer (Ouest-Genopole/IFR26 core facility) and sequences were analyzed using the BioEdit v7.0.5.3 program (Hall T.A., North Carolina State University, Department of Microbiology).

#### *Amplification and cloning*

For cloning, amplification were performed using the AmpliTaq Gold kit with GeneAmp 10X PCR Gold Buffer & MgCl<sub>2</sub> Solution (Applied Biosystems) with the following primers : rFut 2.3 (ATGAGCACCG CCCAGGTCCC CTTC) and rFut 2.2 inv (TCAGTGCTTG AGCAATGGGG ACAG ) for *Fut2*, rSec1.3 (ATGAGATTCG CCCCTGACTA TGTCC) and rSec1.2 inv (CTAGAGGCCA CTCCACAAGG C) for *Sec1*. The obtained amplicons are 1044 pb and 1066 bp long for *Fut2* and *Sec1*, respectively. Cloning was performed in the pcDNA 3.1 vector using the NT-GFP Fusion TOPO TA Expression kit (Invitrogen), following the manufacturer's instructions from 2 µl fresh PCR product and a 20 min incubation at room temperature. The bacteria used for transformation were *E. coli* One Shot TOP10F' (Invitrogen). Orientation of the insert was determined using the T7 Pro primer (TAATACGACT CACTATAGGG) and a primer common to the *Fut2* and *Sec1* inserts, rFut3.2 Inv (GTTGAGGTGG TAGTTCTGCC). All inserted sequences were verified before use in transfection experiments.

*Assay of the enzyme activity of the  $\alpha$ 1,2fucosyltransferases variants*

In order to study the functionality of each of the allelic variant of Fut2 and Sec1, the coding regions were cloned into an eukaryotic expression vector so as to obtain a GFP (green fluorescent protein) fusion protein with the GFP domain at the N-terminus. CHO-K1 cells were cultured in RPMI 1640, supplemented with 10% fetal calf serum, 2 mM L-glutamine, free nucleotides (10  $\mu$ g/ml), 100U/ml penicillin, and 100 mg/ml streptomycin (Gibco, Paisley, UK). They were cultured at confluence after dispersal with 0.025% trypsin and 0.02% EDTA. Cells were routinely checked for mycoplasma contamination using Hoëchst 33258 (Sigma, St Louis, MO) labeling. Following amplification and purification with the Miniprep Qiagen kit, the recombinant pcDNA 3.1 constructs were transfected into CHO-K1 cells using Lipofectamin 2000 (Invitrogen) according to the manufacturer's instructions. Forty-eight hours following transfection, cells were collected and the presence of  $\alpha$ 1,2-linked fucose residues was tested by flow cytometry using the UEA-I lectin (Vector Labs) and the MBr1 mAb (Alexis Biochemicals) that detect the H type 2 (Fuc $\alpha$ 2Gal $\beta$ 4GlcNAc $\beta$ -R) and H type 3 (Fuc $\alpha$ 2Gal $\beta$ 3GalNAc $\alpha$ -R) motifs, respectively. To this aim,  $2.5 \times 10^5$  viable transfected cells were incubated in the presence of either biotin-labeled lectin at 5  $\mu$ g/ml (Vector Labs) or the MBr1 mAb at 5  $\mu$ g/ml for 20 min at 4°C. After 3 washings with PBS, cells were incubated in the same conditions in the presence of either PerCP-conjugated streptavidin (BD Biosciences) at 0.5  $\mu$ g/ml or a Cy5-conjugated anti-mouse IgG (BD Biosciences) at a 1/500 dilution. After 3 more washings with PBS, cell fluorescence was measured on a FACScalibur flow cytometer (Becton-Dickinson) and analyzed using the CellQuest program (Becton-Dickinson). The transfected protein expression was quantified by the GFP fluorescence recorded on the FL1 channel. The presence of the H type 2 and H type 3 motifs was detected on the FL3 and FL4 channels, respectively.

*Statistical analysis*

Comparison of frequencies of the nonsecretor-like phenotype between the 3 rabbit populations was performed by a likelihood ratio Chi<sup>2</sup> test with one degree of freedom.

Analysis of the frequencies of the *Fut2* and *Sec1* alleles between the Chèvreloup subgroups was performed by a likelihood ratio  $\chi^2$  with one degree of freedom and correction for multiple comparisons with  $p$  corrected =  $1-(1-p)^n$ , where  $p$  is the uncorrected  $p$  value and  $n$  the number of comparisons (Sham 1998).

## **Funding**

The work was supported by grants from the BRG (Bureau des Ressources Génétiques), the FNC (Fédération Nationale des Chasseurs) and the Région des Pays de la Loire.

## **Acknowledgements**

The authors are thankful to Dr. S. Marionneau for involvement in the initial phase of the work and to Dr. G. Queney for providing DNA samples.

## **Abbreviations**

CHO, chinese hamster ovary ; GFP, green fluorescent protein, HBGA, histo-blood group antigen ; mAb, monoclonal antibody ; PBS, phosphate buffered saline ; RHDV, rabbit hemorrhagic disease virus ; UEA-I, Ulex europaeus agglutinin I.

## References

- Capucci L, Fusi P, Lavazza A, Pacciarini M, L, Rossi C. 1996. Detection and preliminary characterization of a new rabbit calicivirus related to Rabbit Haemorrhagic Disease Virus but non pathogenic. *J Virol*, 70: 8614-8623.
- Coddens A, Verdonck F, Tiels P, Rassschaert K, Goddeeris BM, Cox E. 2007. The age-dependent expression of the F18+ E. coli receptor on porcine gut epithelial cells is positively correlated with the presence of histo-blood group antigens. *Vet Microbiol*, 122: 332-341.
- Harrington PR, Lindensmith L, Yount B, Moe CL, Baric RS. 2002. Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J Virol*, 76: 12325-12343.
- Hitoshi S, Koijima N, Kusunoki S, Inokuchi J, Kanazawa I, Tsuji S. 1996. Expression of the beta-galactoside alpha 1,2-fucosyltransferase gene suppresses axonal outgrowth of neuro2a neuroblastoma cells. *J Neurochem*, 66: 1633-1640.
- Hitoshi S, Kusunoki S, Kanazawa I, Tsuji S. 1995. Molecular cloning and expression of two types of rabbit  $\beta$ -galactoside  $\alpha$ 1,2fucosyltransferase. *J Biol Chem*, 270: 8844-8850.
- Hitoshi S, Kusunoki S, Kanazawa I, Tsuji S. 1996. Molecular cloning and expression of a third type of rabbit GDP-L-fucose:  $\beta$ -D-galactoside 2- $\alpha$ -L-fucosyltransferase. *J Biol Chem*, 271: 16975-16981.
- Hutson AM, Airaud F, Le Pendu J, Estes MK, Atmar RL. 2005. Norwalk virus infection associates with secretor status genotyped from sera. *J Med Virol*, 77: 116-120.
- Hutson AM, Atmar RL, Graham DY, Estes MK. 2002. Norwalk virus infection and disease is associated with ABO histo-blood group type. *J Infect Dis*, 185: 1335-1337.
- Kindberg E, Akerlind B, Johnsen C, Knudsen JD, Heltberg O, Larson G, Böttinger B, Svensson L. 2007. Host genetic resistance to symptomatic norovirus (GGII.4) infections in Denmark. *J Clin Microbiol*, 45: 2720-2722.
- Le Pendu J, Ruvöen-Clouet N, Kindberg E, Svensson L. 2006. Mendelian resistance to human norovirus infections. *Seminars in Immunology*, 18: 375-386.
- Lindensmith L, Moe CL, Marionneau S, Ruvöen-clouet N, Jiang X, Lindblad L, Stewart PA, Le Pendu J, Baric RS. 2003. Human susceptibility and resistance to Norwalk virus infection. *Nat Med*, 9: 548-553.

- Marcato P, S, Benazzi C, Vecchi G, Galeotti M, Salda L, D, Sarli G, Lucidi P. 1991. Clinical and pathological features of viral haemorrhagic disease of rabbits and the European brown hare syndrome. *Rev Sci Tech Off Int Epiz*, 10: 371-392.
- Marchandeau S. 2000. Enquête nationale sur les tableaux de chasse à tir saison 1998-1999. Le lapin de garenne. *Faune Sauvage*, 251: 18-25.
- Marchandeau S, Chaval Y, Le Goff E. 2000. Prolonged decline in the abundance of wild European rabbits and high immunity level over three years following the arrival of Rabbit Haemorrhagic Disease. *Wildlife Biol*, 6: 141-147.
- Marchandeau S, J C, Portejoie Y, Barraud S, Chaval Y. 1998. Impact of viral hemorrhagic disease on a wild population of European rabbits in France. *J Wildl Dis*, 34: 429-435.
- Marchandeau S, Le Gall-Recule G, Bertagnoli S, J A, Botti G, Lavazza A. 2005. Serological evidence for a non-protective RHDV-like virus. *Vet Res* 36: 53-62.
- Marionneau S, Cailleau-Thomas A, Rocher J, Le Moullac-Vaidye B, Ruvoën-clouet N, Clément M, Le Pendu J. 2001. ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie*, 83: 565-573.
- Marionneau S, Ruvöen-Clouet N, Le Moullac-Vaidye B, Clement M, Cailleau-Thomas A, Riuz-Palacios G, Huang PW, Jiang X, Le Pendu J. 2002. Norwalk virus binds to H types 1/3 histo-blood group antigens on gastro-duodenal epithelial cells of "secretor" individuals. *Gastroenterology*, 122: 1967-1977.
- Mitro S, Krauss H. 1993. Rabbit hemorrhagic disease: a review with special reference to its epizootiology. *European Journal of Epidemiology*, 9: 70-78.
- Morisse J-P, Le Gall G, Boilletot E. 1991. Hepatitis of viral origin in Leporidae: introduction and aetiological hypotheses. . *Rev Sci Tech Off Int Epiz*, 10: 283-295.
- Moss SR, Turner SL, Trout RC, White PJ, Hudson PJ, Desai A, Armesto M, Forrester NL, Gould EA. 2002. Molecular epidemiology of rabbit haemorrhagic disease virus. *J Gen Virol*, 83: 2461-2467.
- Mutze G, Cooke B, Alexander P. 1998. The initial impact of rabbit hemorrhagic disease on European rabbit populations in South Australia. *J Wildl Dis*, 34.
- Oriol R, Candelier JJ, Mollicone R. 2000. Molecular genetics of H. *Vox Sang*, 78: 105-118.
- Queney G, Ferrand N, Marchandeau S, Azevodo M, Mougél F, Branco M Monnerot M. 2000. Absence of a genetic bottleneck in a wild rabbit (*Oryctolagus cuniculus*) population exposed to a severe viral epizootic. *Molecular Ecology*, 9: 1253-1264.

- Queney G, Ferrand N, Weiss S, Mougel F, Monnerot M. 2001. Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Mol Biol and Evol*, 18: 2169-2178.
- Rademacher C, Krishna NR, Palcic MM, Parra F, Peters T. 2008. NMR experiments reveal the molecular basis of receptor recognition by a calicivirus. *J Am Chem Soc*, 130: 3669-3675.
- Rodak L, Smid B, Valicek L. 1991. Application of control measures against viral haemorrhagic disease of rabbits in the Czech and Slovak Federal Republic. *Rev Sci Tech Off Int Epiz*, 10: 513-524.
- Ruvoën-clouet N, Ganière JP, André-Fontaine G, Blanchard D, Le Pendu J. 2000. Binding of Rabbit Hemorrhagic Disease Virus to antigens of the ABH histo-blood group family. *J Virol*, 74: 11950-11954.
- Sham P. 1998. *Statistics in human genetics*. Arnold, London.
- Tan M, Jiang X. 2005. Norovirus and its histo-blood group antigen receptors: an answer to a historical puzzle. *Trends Microbiol*, 13: 285-293.
- Teshima KM, Innan H. 2004. The effect of conversion on the divergence between duplicated genes. *Genetics*, 166: 1553-1560.
- Thorven M, Grahn A, Hedlund K-O, Johansson H, Wahlfrid C, Larson G, Svensson L. 2005. A homozygous nonsense mutation (428G>A) in the human FUT2 gene provides resistance to symptomatic norovirus (GGII) infections. *J Virol*, 79: 15351-15355.
- Villafuerte R, Calvete C, Blanco JC, Lucientes J. 1995. Incidence of viral hemorrhagic disease in wild rabbit populations in Spain. *Mammalia*, 59: 651-659.
- Villafuerte R, Calvete C, Gortazar C, Moreno S. 1994. First epizootic of rabbit hemorrhagic disease in free living populations of *Oryctolagus cuniculus* at Doñana national park, Spain. *J Wildl Dis*, 30: 176-179.
- Woolhouse MEJ, Webster JP, Domingo E, Charlesworth B, Levin BR. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet*, 32: 569-577.

## ***b - Conclusions et perspectives***

Depuis son apparition dans les années 1980, le RHDV est devenu le pathogène le plus meurtrier chez les populations sauvages de lapins de garenne. La mise en évidence du rôle potentiel de l'HBGA H type 2 comme récepteur cellulaire du virus a dans un premier temps permis de suggérer que les jeunes lapereaux ne seraient pas sensibles au virus, du fait de la faible expression de cet HBGA à la surface de leurs cellules épithéliales. Le but de notre étude a été de déterminer si dans la nature, les lapins adultes pouvaient avoir à leur tour une résistance associée à un phénotype non sécréteur et si tel était le cas, quelle(s) mutation(s) des gènes codant une  $\alpha$ 1,2fucosyltransférase pouva(en)t en être responsable(s).

Grâce à une collaboration avec les ingénieurs et techniciens de l'Office National de la Chasse et de la Faune Sauvage, nous avons eu accès à des échantillons provenant de trois populations sauvages de lapins de garenne. Parmi ces groupes d'animaux, une distinction a été faite en fonction de leur passé épidémique vis-à-vis du RHDV, mais également en fonction de la sévérité de ces épidémies. Le phénotypage du caractère sécréteur a permis pour la première fois de mettre en évidence l'existence de lapins non sécréteur ou sécréteur faible *in naturo*. Nous avons de plus montré qu'il existe une corrélation entre la sévérité des épidémies de RHDV et l'augmentation relative de la fréquence du caractère non sécréteur et sécréteur faible chez les animaux survivants. Ceci pourrait signifier que la pression de sélection exercée par le RHDV a conduit à l'augmentation de la fréquence d'un allèle d' $\alpha$ 1,2fucosyltransférase non fonctionnel. Chez le lapin, ces gènes correspondent à *Fut2* et *Sec1*. En se focalisant sur la population de lapins la plus sévèrement touchée par le virus, le séquençage de leur région codante respective a tout d'abord montré qu'ils sont très polymorphes chez cette espèce. Ces résultats sont en contradiction avec ceux établis dans les mêmes populations suite à l'étude de marqueurs neutres de l'évolution (séquences microsatellites) (Queney et coll. 2000), ainsi qu'avec ceux que nous avons obtenus pour le gène *Fut1* qui lui ne montre que très peu de polymorphisme. Cette observation pourrait signifier que ces deux gènes sont soumis à une forte pression de sélection positive. Cependant, nous avons observé des signes de conversion génique entre les séquences codant les domaines catalytiques des enzymes *Fut2* et *Sec1*. Dans le cas présent, la conversion génique pourrait être assimilée à un mécanisme contribuant à augmenter la diversité du gène *Fut2*, correspondant ici au gène receveur, ou bien à un mécanisme permettant la correction de mutations trop délétères grâce la copie de sauvegarde *Sec1* qui joue lui ici le rôle de gène donneur.

Afin de vérifier si l'un des allèles *Fut2* (n=5) ou *Sec1* (n=7) pouvait correspondre à un variant inactif de ces gènes, nous avons cloné chacun d'eux dans un vecteur

permettant la production d'une glycosyltransférase fusionnée à l'EGFP. Leur activité enzymatique a ensuite été évaluée *in vivo* en transfectant des cellules CHO avec ces constructions, puis en évaluant l'expression de l'antigène H type 2 et 3 par la technique de cytométrie en flux. L'activité enzymatique relative à chacun des variants a pu être normalisée grâce à l'expression synchrone de l'EGFP. Nous avons alors mis en évidence que tous les variants *Fut2* isolés sont actifs et équivalents entre eux. Les variants *Sec1* codent quant à eux des enzymes aux activités beaucoup plus variables, certaines étant totalement inactives, les autres étant très faiblement à faiblement actives, comparées aux enzymes codées par les variants *Fut2*. Toutefois, la présence chez tous ces lapins d'au moins un variant (*Fut2* ou *Sec1*) capable de produire une enzyme active ne permet pas d'expliquer la fréquence élevée du phénotype non sécréteur dans cette population.

Toutefois, en superposant les données épidémiologiques aux fréquences relatives de chacun des allèles *Fut2* et *Sec1* identifiés, nous avons pu mettre en évidence qu'un variant de *Sec1* est associé à la survie des lapins à l'épidémie de RHDV. Cette résistance ne peut pas être expliquée par son activité enzymatique faible. En revanche, de par leur position en tandem sur le chromosome, une mutation inactivante dans la région promotrice ou non codante du gène *Fut2* pourrait être associée à ce variant de *Sec1* et expliquer pourquoi il est associé à une plus grande résistance au RHDV.

## **2.3. Article 3 : Inhibition de l'interaction entre la protéine Spike du SRAS-CoV et son récepteur cellulaire par des anticorps anti-antigènes tissulaires de groupe sanguin**

### **a - Introduction**

Le SRAS-CoV est responsable de la première pandémie du XXI<sup>e</sup> siècle. Extrêmement pathogène, son origine la plus probable est le passage de la barrière d'espèce par l'une des souches de CoV d'origine animale (civette, chauve-souris) (Y. Guan et coll. 2003; S. K. P. Lau et coll. 2005). Il se réplique préférentiellement dans les pneumocytes et l'épithélium intestinal. Ces cellules lui sont permissives du fait de la reconnaissance de leur protéine ACE2 par la protéine virale Spike (protéine S). Cette liaison est hautement spécifique et la réplication du virus est très fortement diminuée lors de l'apparition de variations dans leur domaine d'interaction respectif (K. V. Holmes 2005).

Une étude épidémiologique concernant l'impact du SRAS-CoV sur le personnel soignant d'un hôpital de Hong Kong a mis en évidence que les individus de groupe sanguin O sont moins sujets à l'infection que les personnes de phénotype A, B ou AB (Y. Cheng et coll. 2005). Cette observation nous a fait nous questionner sur le rôle potentiel des NAb anti-ABH dans cette apparente résistance. En effet, la protéine S est une glycoprotéine qui contient 23 sites potentiels de N-glycosylation dont trois sont localisés dans son RBD. Le SRAS-CoV se réplique dans des cellules possédant la machinerie enzymatique nécessaire à la synthèse des HBGA ABH. Or, il a été mis en évidence pour d'autres virus enveloppés que la réplication dans de telles cellules permet l'ajout des HBGA ABH sur les glycoprotéines virales (Neil et coll. 2005; Preece et coll. 2002; Y. Takeuchi et coll. 1996). Il a de plus été montré pour ces virus que la reconnaissance des motifs A ou B par les NAb induit une diminution de leur infectivité, et que la combinaison NAb/complément permet l'activation de la réponse immunitaire adaptative (Dürnbach et coll. 2007). D'après ces observations et les règles universelles de transfusion, la résistance des individus O au SRAS-CoV pourrait être expliquée par la reconnaissance des HBGA A et/ou B présent sur la protéine S. Afin de tester cette hypothèse, nous avons utilisé un modèle cellulaire d'interaction protéine S/ACE2 permettant la co-expression de la glycoprotéine virale et de l'HBGA A.

Les résultats présentés dans l'article suivant montrent que les NAb anti-A peuvent inhiber la fixation de la protéine S sur l'ACE2, ceci uniquement quand la glycoprotéine virale est synthétisée dans des cellules exprimant également l'HBGA A. Nous avons de

plus évalué à l'aide d'un modèle mathématique de type SIR, l'impact d'une telle protection sur le profil de l'épidémie, ceci en fonction du polymorphisme ABH de différentes populations et de leur titre en NAb anti-HBGA.

**INHIBITION OF INTERACTION BETWEEN SARS-CoV SPIKE  
PROTEIN AND ITS CELLULAR RECEPTOR BY ANTI-HISTO BLOOD  
GROUP ANTIBODIES**

**GLYCOBIOLOGY, 2008 Dec, 18 (12) : 1085-93**

**Patrice Guillon<sup>2,3</sup>, Monique Clément<sup>2,3</sup>, Véronique Sébille<sup>4</sup>, Jean-Gérard Rivain<sup>4</sup>,  
Chih-Fong Chou<sup>5</sup>, Nathalie Ruvoën-Clouet<sup>3,6</sup>, and Jacques Le Pendu<sup>1,3</sup>**

<sup>3</sup> INSERM, U892, Nantes, France ; University of Nantes, Nantes, France.

<sup>4</sup> EA 4275, University of Nantes, Nantes, France.

<sup>5</sup> Institute of Molecular and Cell Biology, Singapore.

<sup>6</sup> Ecole Nationale Vétérinaire de Nantes, Nantes, France.

Key words : SARS/histo-blood group antigens/natural antibodies/ABO/cellular receptor

Running title: SARS protection by natural ABO antibodies

<sup>1</sup>To whom correspondence should be addressed

INSERM U892, Institut de Biologie, 9 Quai Moncoussu

44093, Nantes Cedex 01, France

Tel : 33 240 08 40 99, Fax : 33 240 08 40 82

E-Mail : jlependu@nantes.inserm.fr

<sup>2</sup>These two authors contributed equally to the work

One supplementary material file is provided

## Abstract

SARS-CoV is a highly pathogenic emergent virus which replicates in cells that can express ABH histo-blood group antigens. The heavily glycosylated SARS-CoV spike (S) protein binds to angiotensin-converting enzyme 2 which serves as a cellular receptor. Epidemiological analysis of a hospital outbreak in Hong Kong, revealed that blood group O was associated with a low risk of infection. In this study, we used a cellular model of adhesion to investigate whether natural antibodies of the ABO system could block the S protein-angiotensin-converting enzyme 2 interaction. To this aim, a C-terminally EGFP-tagged S protein was expressed in chinese hamster ovary cells cotransfected with an  $\alpha$ 1,2fucosyltransferase and an A-transferase in order to coexpress the S glycoprotein ectodomain and the A antigen at the cell surface. We observed that the S protein/angiotensin-converting enzyme 2-dependent adhesion of these cells to an angiotensin-converting enzyme 2 expressing cell line was specifically inhibited by either a monoclonal or human natural anti-A antibodies, indicating that these antibodies may block the interaction between the virus and its receptor, thereby providing protection. In order to more fully appreciate the potential effect of the ABO polymorphism on the epidemiology of SARS, we built a mathematical model of the virus transmission dynamics that takes into account the protective effect of ABO natural antibodies. The model indicated that the ABO polymorphism could contribute to substantially reduce the virus transmission, affecting both the number of infected individuals and the kinetics of the epidemic.

## Introduction

Severe acute respiratory syndrome (SARS) is characterized by an acute respiratory disease, often accompanied by gastroenteritis, which is fatal in approximately 10% of infected individuals (Gu and Korteweg 2007). The etiological agent is a novel coronavirus (CoV) designated SARS-CoV which emerged from an animal reservoir during the winter 2002-2003 when it infected over 8000 humans worldwide. It emerged again the next winter and since then, no SARS outbreak has been recorded (Chen and Subbarao 2007). Yet, a serious possibility of reemergence of SARS-CoV or of the introduction of other related viruses from an animal reservoir remains.

Spike (S) proteins of coronaviruses are large transmembrane heavily N-glycosylated proteins that mediate association with a cell surface receptor (Li et al. 2006). The SARS-CoV S protein possesses 23 N-linked glycosylation sites distributed in 3 clusters. The glycosylation of 13 of these sites has been confirmed (Chakraborti et al. 2005, Krokhnin et al. 2003, Ying et al. 2004). Receptor binding domains (RBDs) have been identified in the S1 domain of a number of coronaviruses and a fragment of the SARS-CoV S1 domain, from residues 318-510, binds human angiotensin-converting enzyme 2 (ACE2) with high affinity (Babcock et al. 2004, Wong et al. 2004, Xiao et al. 2003). It has additionally been demonstrated that ACE2 constitutes an obligate cellular receptor, although other receptors may participate in the infection process (Chen and Subbarao 2007, Li et al. 2003). The structure of SARS-CoV RBD complexed with ACE2 revealed that an extended loop of the RBD, comprising residues 424-494, is in direct contact with ACE2 (Li et al. 2005). This receptor binding motif (RBM) is not glycosylated, but it is surrounded by two clusters of glycosylation sites (Han et al. 2007).

Various genetic factors influencing the susceptibility to or the outcome of SARS have been described (Gu and Korteweg 2007). The *ABO* gene stands out among the genes involved since O blood group individuals were shown to have very low odds of infection compared to non O individuals in an hospital outbreak that occurred in March

2003 in Hong Kong (Cheng et al. 2005). Histo-blood group antigens are present not only on erythrocytes but also on many epithelial cells, which are their main site of expression (Marionneau et al. 2001). Since SARS-CoV replicates in epithelial cells of the respiratory and digestive tracts that have the ability to synthesize ABH carbohydrate epitopes, we hypothesized that the S protein of virions produced by either A or B individuals could be decorated with A or B carbohydrate epitopes, respectively. Natural anti-A or B antibodies from blood group O, B and A individuals, respectively, could bind to the S protein and block its interaction with ACE2, thereby preventing infection in accordance with the rules of transfusion.

In order to put this hypothesis to the test, we used a cell binding assay that reconstitutes the interaction between the S protein and ACE2 (Chou et al. 2005). We present data indicating that the S protein/ACE2-mediated adhesion between cells expressing ACE2 and cells coexpressing the S protein and the A histo-blood group antigen can be specifically blocked by anti-A antibodies. To further evaluate the potential effect of the ABO polymorphism on the epidemiology of SARS, we present a model of its transmission dynamics that takes into account the effect of the protection by anti-histo-blood group natural antibodies.

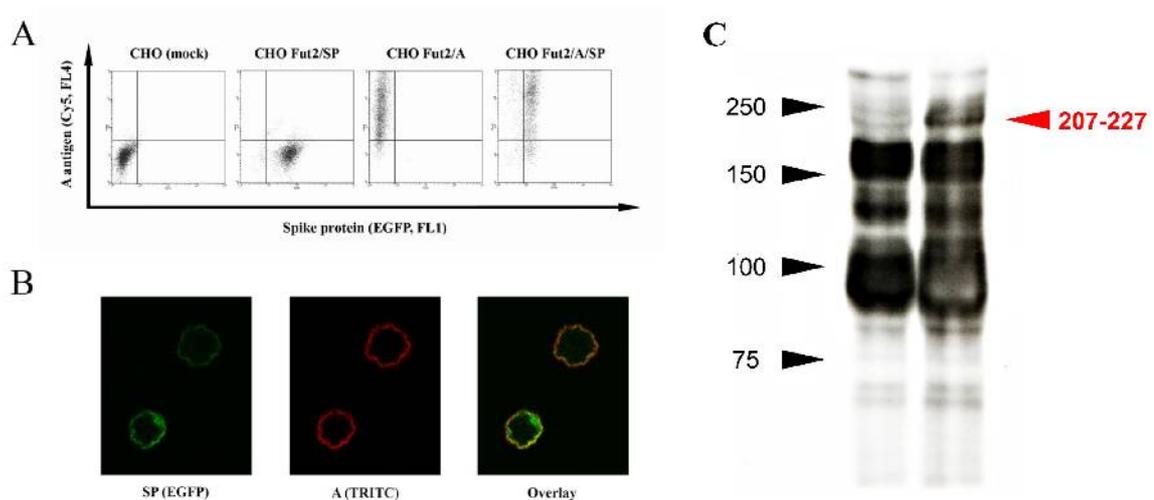
## Results

### *Preparation of cells coexpressing the A antigen and SARS-CoV S protein for the study of ACE2/S protein interaction in a cell adhesion assay*

The interaction between SARS-CoV spike protein and its cellular receptor ACE2 can be studied using a cell based assay, as described previously (Chou et al. 2005). In this assay, the viral S protein expressed by transfection into CHO cells mediates adhesion to Vero E6 cells that possess ACE2. CHO cells do not express ABH antigens because of the lack of an  $\alpha$ 1,2fucosyltransferase activity and of either the A or B histo-blood group enzymes. In order to obtain cells able to express the A antigen, parental CHO cells were stably transfected successively with the rat Fut2 cDNA and a rat A enzyme cDNA. Unlike mock transfected cells, these double transfectants strongly express cell surface A antigen

as detected by flow cytometry. Transfection of the S protein-EGFP fusion construction (S-EGFP) into these cells allowed expression of the S protein together with the histo-blood group A antigen (Fig. 1A). Observation of the triple transfectants by confocal microscopy revealed that, as expected, the A antigen and the S-EGFP fusion protein partially colocalized at the cell surface (Fig. 1B). In addition, western blot analysis revealed that an A antigen positive band at the expected size of the S-EGFP fusion protein, between 210-230 kDa (Chou et al. 2005), was present following transfection with the fusion protein construct. The band was absent from the mock transfectant cell extracts, indicating that the S-protein expressed by A positive CHO cells carried A histo-blood group epitopes (Fig. 1C).

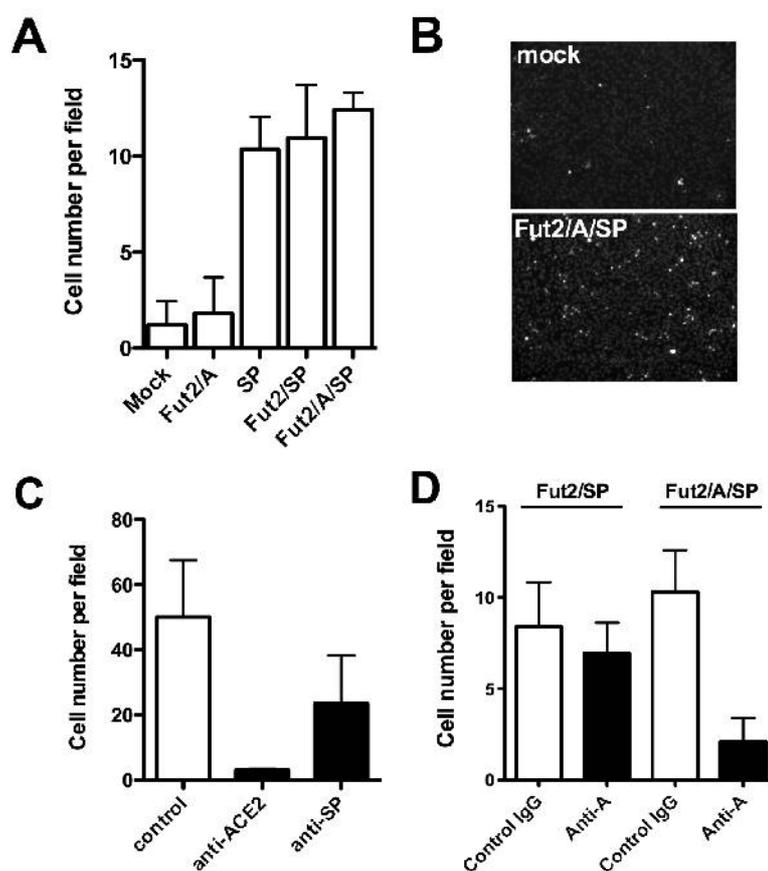
A stable A antigen and S-EGFP expressing clone showed significantly higher



**Figure 1** : CHO cells coexpressing the A histo-blood group antigen and the SARS-CoV S protein. **A/** Flow cytometry analysis of the expression of A antigen and the S protein-EGFP construct on CHO mock transfected cells (mock), double transfectants either with the Fut2 and S protein constructs (Fut2/SP) or the Fut2 and A enzyme cDNAs (Fut2/A), triple transfectants with the Fut2, A enzyme and S protein constructs (Fut2/A/SP). Fluorescence of the S-EFGP molecule was directly recorded on the FL1 channel. Detection of the A antigen was performed using an anti-A mAb followed by Cy5-labeled anti-mouse IgG and recorded on the FL4 channel. **B/** Confocal microscopy analysis of the A antigen (A TRITC) and the S protein-EGFP construct coexpression on CHO cells triple transfectants. Detection of the A antigen was performed using an anti-A mAb followed by TRITC-labeled anti-mouse IgG. **C/** Western blot analysis of transfected CHO cells glycoproteins.

adhesion to Vero cells than either mock transfectants or the A expressing clone devoid of S protein (Fig. 2A, B). Similar results were obtained after transient transfection of the S-EGFP construct (not shown). The presence of the A and H antigens on S protein expressing cells did not affect adhesion since CHO cells only transfected with the S-EGFP

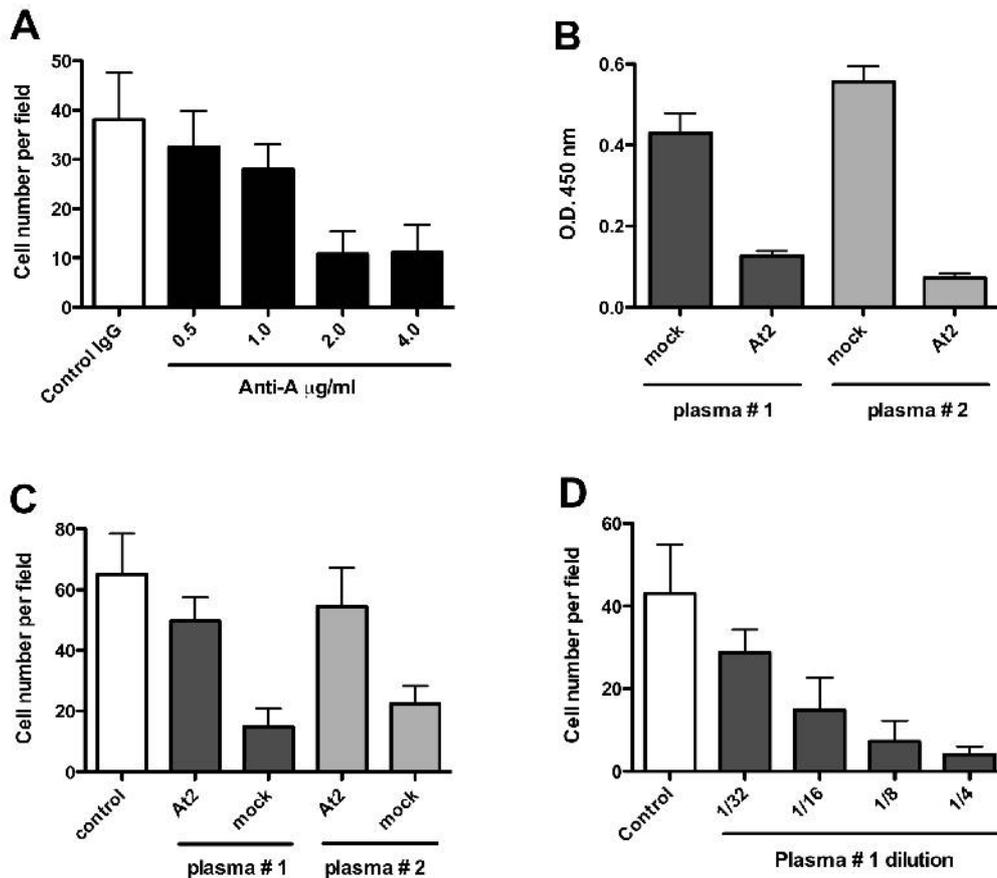
construct as well as CHO cells transfected with both the S-EGFP and Fut2 cDNAs adhered to Vero cells at a similar level as the A antigen S protein triple transfectants (Fig. 2A). In order to control that the cell adhesion was dependent upon the ACE2/S protein interaction, blocking experiments with either an anti-ACE2 or an anti-S protein were performed. Both antibodies significantly inhibited adhesion, although the anti-ACE2 mAb proved more efficient (Fig. 2C).



**Figure 2** : S Protein/ACE2-dependent adhesion of CHO cells to Vero cells. The binding assay between CHO cells and Vero E6 cells was performed as described in the « materials and methods » section. Adherent cells were counted under a fluorescence microscope. Cells from a total of 36 fields from 6 wells were counted. **A/** The results shown correspond to the mean  $\pm$  SD of one representative experiment out of 4 obtained with CHO cells mock transfectants (mock), double transfectants with the Fut2 and A histo-blood group glycosyltransferases (Fut2/A), simple transfectants with the SARS-CoV S protein construct (SP), double transfectants with the Fut2 enzyme and the SARS-CoV S protein construct (Fut2/SP), and triple transfectants with the Fut2, A glycosyltransferases and the SARS-CoV S protein cDNAs (Fut2/A/SP). Adhesion of SP, Fut2/SP and Fut2/A/SP cells is significantly higher than that of either mock or Fut2/A cells ( $p < 0.001$ , Student t test). **B/** Representative fields illustrating the adhesion of either mock transfected CHO cells, or triple transfectants. **C/** Inhibition of triple CHO transfectants to Vero cells by anti-ACE2 or anti-S protein mAbs. The mAbs, were added to the CHO cells suspension at 20 and 25 mg/ml, respectively prior to incubation on the Vero cell layer. Adhesion in the presence of the anti-ACE2 and anti-SP are significantly lower than that of control cells ( $p < 0.001$  and  $p < 0.01$ , respectively). **D/** Inhibition of the adhesion to Vero cells of S protein transfected CHO cells coexpressing either the H or the A antigen by an anti-A mAb or a control isotype matched antibody used at 4 mg/ml. Only the adhesion of the triple transfectant in the presence of the anti-A is differs significantly from the other 3 conditions ( $p < 0.01$ ).

*Inhibition of adhesion by anti-A antibodies*

The effect of anti-A antibodies on the S protein/ACE2 interaction was first tested using a monoclonal anti-A. A clear cut inhibition of the cell adhesion was observed using monoclonal antibody 3-3A at 2 µg/ml. Specificity of the inhibition was confirmed since a control irrelevant antibody failed to inhibit and since the adhesion to Vero cells of S protein transfected CHO cells lacking the A antigen was not inhibited by the anti-A mAb (Fig. 2D). Vero cells do not express the A histo-blood group antigen. Therefore the inhibition of adhesion mediated by the anti-A mAb can only result from a binding to CHO S protein expressing cells and not to the glycans of ACE2. In order to assay the ability of natural human anti-A to inhibit cell adhesion, plasma samples from two O blood group individuals with high anti-A titers were selected. The samples were first adsorbed on silica beads conjugated to the A type 2 tetrasaccharide in order to specifically remove the anti-A natural plasma antibodies. Efficacy of the adsorption was controlled by ELISA which showed that the reactivity to the A type 2 tetrasaccharide was almost completely abolished following adsorption (Fig. 3B). The A type 2 adsorbed and mock adsorbed plasma samples from the two individuals were then added in the cell based assay. Both mock adsorbed samples, containing the anti-A, strongly inhibited the adhesion of A antigen-S protein expressing cells to Vero cells. This inhibition was strongly decreased after A type 2 adsorption, showing that it was specifically mediated by anti-A plasma antibodies (Fig. 3C). Moreover, the inhibition of adhesion by blood group O plasma was dose-dependent and still detected at a plasma dilution as low as 1/32 (Fig. 3D).

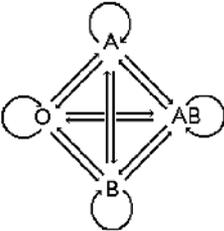
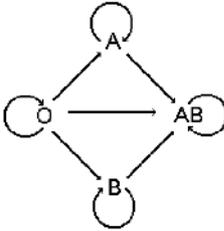
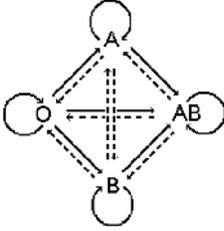


**Figure 3** : Effect of anti-A antibodies on the interaction between SARS-CoV S protein and ACE2. **A/** The anti-A monoclonal antibody 3-3A was added at the indicated concentrations to the triple transfected CHO cells suspension prior to incubation on the Vero cell layer. An irrelevant IgG1 was used as control at 4 mg/ml. The results are presented as mean cell number per field +/- SD of one representative experiment out of 2. From 1.0 mg/ml to 4.0 mg/ml anti-A, values are significantly different from those for the control IgG ( $p < 0.05$  and  $0.001$ , respectively). **B/** Adsorption of the anti-A natural antibody from human O plasmas. Plasma samples from 2 individuals were adsorbed on either control silica beads (mock) or A type 2 tetrasaccharide conjugated to silica beads (At2). The post-adsorption plasma reactivity on A type 2 conjugated to polyacrylamide was tested by ELISA. Results are shown as OD 450 nm values of duplicate wells +/- SD for each plasma sample diluted at 1/4. In absence of A type 2 conjugate, mean OD values were 0.130. **C/** Inhibition of the adhesion of CHO triple transfected cells to Vero cells by mock adsorbed (mock) or A type 2 (At2) adsorbed human blood group O plasma samples from individuals #1 and #2. Plasma samples were diluted at 1/8 in PBS. Control values were obtained in absence of plasma. Values for the mock adsorbed plasma were significantly different from the control value ( $p < 0.001$ ). **D/** Inhibition of adhesion in the cell based assay as in C by serial dilutions of unadsorbed plasma from individual #1. All values obtained in the presence of plasma were significantly different from the control value (from  $p < 0.05$  to  $p < 0.0001$ ).

### Modelling the effect of protection by natural anti-A or B antibodies on the virus transmission in populations

A total population  $N_{TOT} = N_A + N_B + N_O + N_{AB}$  of  $10^6$  individuals including one blood group O infected individual was simulated. Since the model was not affected by the blood group of the index case, only the results obtained with a blood group O index case are shown. Three different patterns of transmission of the virus, with different probabilities of

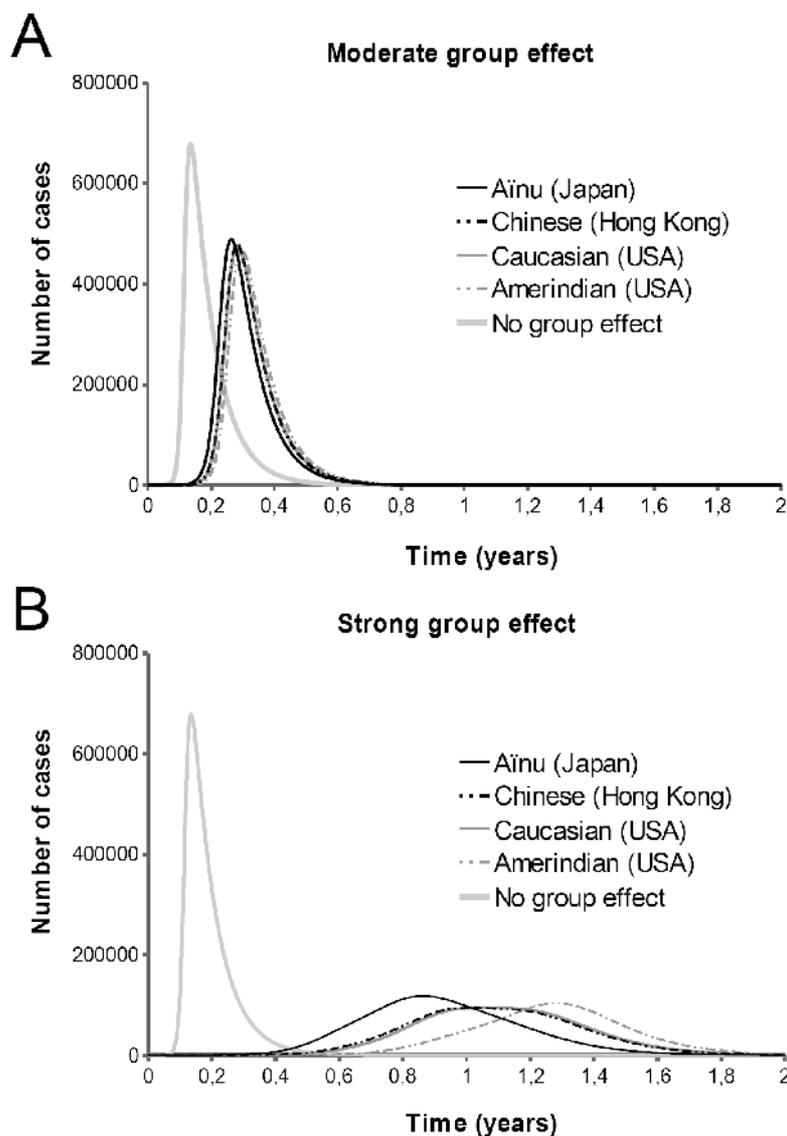
transmission, were assumed according to whether some protection by anti-histo-blood group natural antibodies was present (Fig. 4).

Group effect		
None	Strong	Moderate
		
Transmission coefficients		
$\beta = 0.00944$	$\beta = 0.00639$	$\beta_1 = 0.00639$ (full lines) $\beta_2 = 0.00305$ (dashed lines)

A strong or moderate group effect denotes a strong or moderate protection,

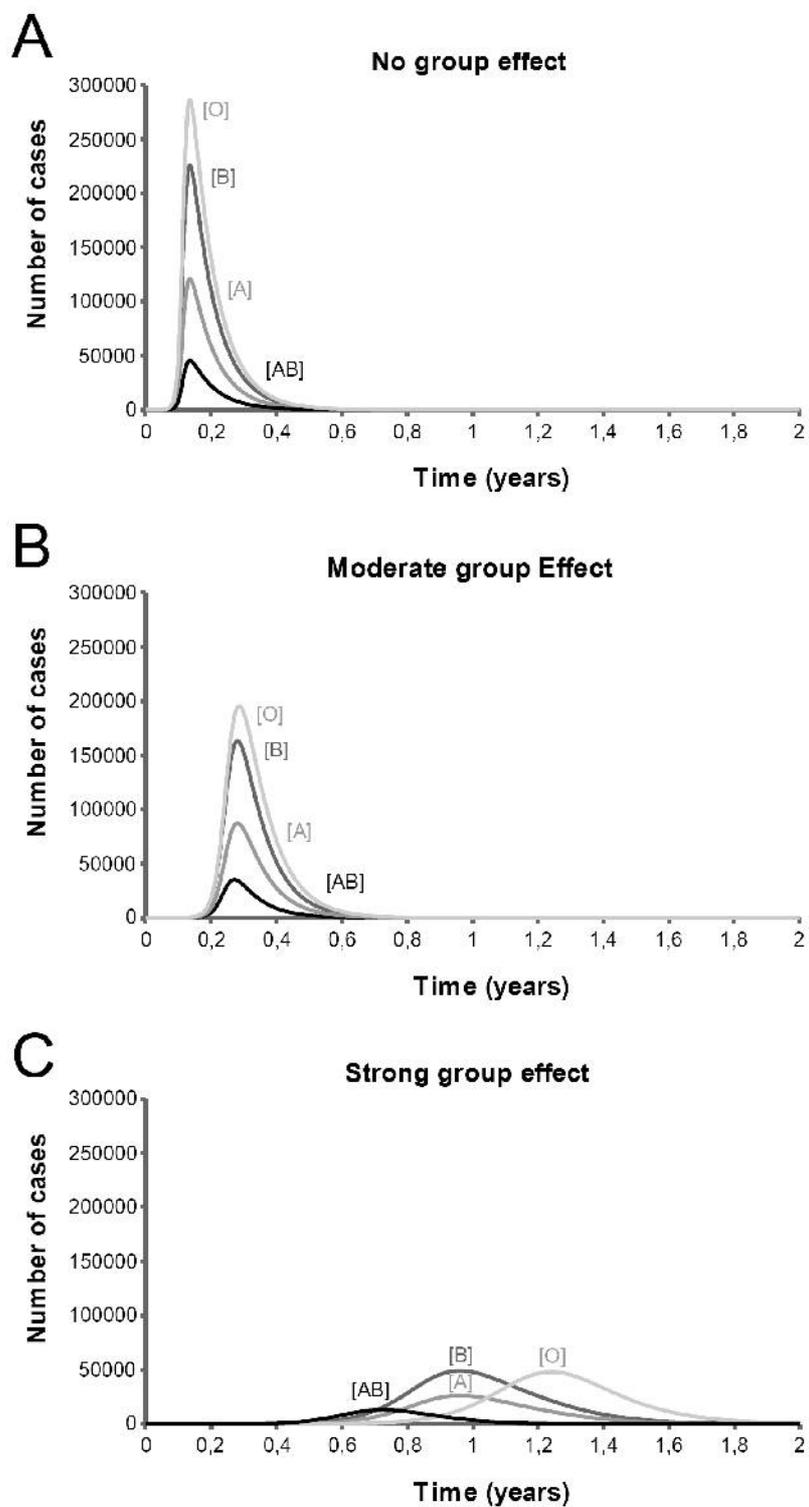
**Figure 4** : Transmission patterns used to model the effect of the ABO polymorphism. In absence of ABO effect, transmission can occur irrespective of the ABO type (full arrows in all directions). In the presence of a strong ABO effect, transmission occurs strictly according to the rules of transfusion, whereas in the case of a moderate ABO effect, some incompatible transmission can occur (dashed arrows). Determination of the values of the transmission coefficients  $\beta$ ,  $\beta_1$ ,  $\beta_2$  has been done based on the Hong Kong hospital outbreak data, as described in the supplemental material. Transmission coefficients correspond to the transmission rate of the disease for each contact.

respectively, whereas no group effect corresponds to an absence of protection. The impact of group effect was assessed on the number of infected individuals over time in four different populations with very different frequencies of ABO phenotypes (Fig. 5). A strong group effect delayed the initiation of the epidemic as well as it decreased importantly the total number of infected individuals whatever the population considered as compared to no group effect and, to a lesser extent, a moderate group effect. Moreover, a strong group effect evidenced different starting times of the epidemic according to the different populations (Fig. 5B).



**Figure 5** : Influence of the ABO polymorphism, with either a moderate group effect (A) or a strong group effect (B) as compared with no group effect, on the number of infected individual over time in four different populations presenting large differences in the frequencies of ABO phenotypes.

The impact of group effect was assessed on the number of infected individuals over time according to their blood groups in the Chinese (Hong Kong) population (Fig. 6). The number of blood group O infected individuals was always the most important one, closely followed by blood group B, A and AB infected individuals whatever the transmission pattern considered. A strong group effect also delayed the initiation of the epidemic as well as it markedly decreased the total number of infected individuals. Furthermore, a strong group effect also postponed the starting time of the epidemic in blood group O individuals as compared with the other blood groups.



**Figure 6** : Influence of different transmission patterns, no group effect (A), moderate group effect (B) or strong group effect (C), on the number of infected individuals over time according to their blood groups in the Chinese (Hong Kong) population.

## Discussion

Probably due to mimicry of A or B antigens by flora or infectious bacteria, individuals acquire anti-A or B blood group antibodies to the antigen that they do not synthesize. These so-called natural anti-histo-blood group antibodies have long been suspected to play a role in anti-viral immunity since viruses may carry ABH structures as terminal carbohydrate motifs of their envelope glycoproteins or possibly as inserted glycolipids (Greenwell 1997). In line with this concept, a monoclonal anti-A was shown to neutralize HIV produced by lymphocytes from blood group A donors only (Arendrup et al. 1991). More recently, anti-A or B from human serum were shown to sensitize HIV to complement-mediated inactivation (Neil et al. 2005). Likewise, measles virus produced by cells expressing either A or B histo-blood group epitopes was neutralized by natural anti-HBGAs in a complement-dependent manner (Preece et al. 2002). Though these *in vitro* data suggest that natural anti-HBGAs may provide protection against some viruses, they have not been substantiated by epidemiological observations so far. If natural anti-A or B serum antibodies provide protection, it is expected that during an outbreak, blood group O individuals should experience a lower risk of infection than non blood group O individuals. This has not been observed for either HIV or measles virus at present. Yet, it is precisely what was observed in the case of a hospital outbreak of SARS in Hong Kong, where O blood group individuals appeared at a much lower risk of being infected by SARS-CoV than subjects of other blood types (Cheng et al. 2005). Interestingly, SARS-CoV infects cells that express ABH antigens according to the individual's ABO phenotype. Indeed, SARS-CoV infection has been documented in pneumocytes, enterocytes of the small intestine as well as in cells of the kidney distal tubular epithelium, all cell types known to be able to synthesize ABH antigens (Chen and Subbarao 2007, Gu and Korteweg 2007). Since the glycosylation of viral glycoproteins necessitates the glycosylation machinery of the infected cell, viral particles synthesized in cells that may express histo-blood group antigens are expected to be tagged with these antigenic motifs

and therefore natural antibodies directed against these carbohydrate tags may have a protective role (Seymour et al. 2004).

In order to determine if anti-HBGAs could block SARS-CoV entry into target cells, we used an experimental model of cell adhesion that has been developed with the aim of screening molecules that block the virus entry without using infectious particles. The model allowed us to show that either a monoclonal anti-A antibody or natural plasma anti-A specifically inhibited the SARS-CoV S protein/ACE2-dependent adhesion. This is in accordance with the hypothesis of the protective role of natural anti-HBGAs and strongly suggests that the low risk of infection of blood group O individuals during the Hong Kong hospital outbreak was due to the presence of these antibodies prior to the outbreak.

Although the RBM of SARS-CoV S protein does not involve any glycan chain, clusters of glycosylation sites are located in its vicinity (Han et al. 2007, Li et al. 2005). Large molecules such as lectins or antibodies binding to these glycans are thus expected to interfere with the S protein/ACE2 interaction. This has already been observed with mannose-specific lectins which have shown anti-viral activity against SARS-CoV by blocking virus attachment to its receptor (Keyaerts et al. 2007). Our data indicate that natural anti-carbohydrate antibodies could have a similar effect. In addition to the blocking of virus attachment to its receptor, natural antibodies could block entry or opsonize viral particles leading to complement-mediated neutralization (Neil et al. 2005). Moreover, it was recently shown that natural antibodies can contribute to help the generation of cytotoxic T cells against the pathogen (Dürnbach et al. 2007, Stäger et al. 2003). These additional mechanisms of protection may also have participated to the protection of blood group O individual during the SARS outbreak.

Since the epidemiology of SARS has been well documented (Anderson et al. 2004, He and Chinese 2004), it was possible to develop a model of the virus transmission that takes into account the effect of the ABO polymorphism. Critical transmission parameters were deduced from those of the global SARS outbreak and from those of the Hong Kong hospital outbreak where the ABO effect was observed. The model considers that no

prophylactic measures were taken, although this is clearly unrealistic, it allows to fully appreciate the impact of the blood group polymorphism itself. Of note, the expression of ABH antigens in epithelial cells where SARS-CoV replicates is also controlled by polymorphisms of the *FUT2* gene. Thus, individuals with two *FUT2* null alleles, the so-called nonsecretors, are unable to synthesize H antigen and hence A or B antigens in these cells (Marionneau et al. 2001). For simplicity, the model did not consider such individuals since with regard to the virus transmission, they would behave as O blood group donors. Including them in the analysis is therefore similar to slightly increase the pool of O individuals. Since anti-A and anti-B titers inter-individual variability is quite high, and since we observed that plasmas from O blood group individuals with low anti-A titers were not inhibitory in the cell adhesion assay (not shown), we considered transmission parameters allowing for a moderate ABO effect only. In this case, virus transmission in incompatible ABO situations remains possible but with a lower probability of occurrence than in compatible situations. That is certainly more realistic than the case of the strong ABO effect where ABO incompatibility completely impairs virus transmission. The latter case was analysed in order to evaluate the maximal potential of the ABO polymorphism. The model indicated that both in the presence of a moderate or of a strong group effect, virus transmission was decreased, supporting the hypothesis that natural anti-A and B antibodies can contribute to protect against selected viral diseases at the population level. Less intuitively, the model shows that the main effect of the natural anti-histo blood group antibodies is to delay and to slow down the epidemic. In the case of a full protection (strong effect), the delay between the occurrence of the first cases and the full development of the outbreak can be very large. It is linked to the frequency of O individuals in the population, but remains very significant even in a population with an unusually low blood group O frequency such as the Ainous of Japan. This delay, already clearly visible when taking into account a moderate ABO effect only, might have had an adaptive value, in past epidemics of other viruses with transmission characteristics similar to those of SARS, since it allows for modifications of behavior limiting the spread of epidemics. That could have contributed to the maintenance of the

ABO polymorphism throughout human evolution and history. Mean natural anti-A or B titers tend to decrease over the years in developed countries, possibly due to improved hygiene (Dr. A Blancher, personal communication). It is thus possible that in the past their protective effect was higher than in most contemporary populations. It could thus be of interest to raise the levels of anti-A or B in all populations so as to slow and limit the spread of some emergent pathogens. This could prove a valuable prevention strategy against SARS but also against other coronaviruses which are responsible for a significant proportion of common colds and can contribute to more severe respiratory tract infections (van der Hoek 2005).

## **Materials and Methods**

### *Cell culture and transfections*

Cell lines CHO-K1 and Vero E6 were purchased from American Type Cell Collection (Manassas, VA, USA). Vero-E6 cells were cultured in Dulbecco's modified Eagle medium supplemented with 10% decompemented fetal-calf serum, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin (Gibco, Paisley, UK).

CHO cells are devoid of  $\alpha$ 1,2fucosyltransferase activity and therefore of ABH antigens. To obtain expression of the A antigen, they were transfected first with the rat Fut2 cDNA and then with an A enzyme cDNA (Bureau et al. 2001, Cailleau-Thomas et al. 2002). The rat Fut2 cDNA was inserted into the pDR2 eukaryotic expression vector (Clontech, St Germain en Laye, France) deleted of the sequences lying between the EcoRV and ClaI sites. This vector possesses an hygromycin selection marker. The rat A enzyme cDNA was inserted into the pcDNA3.1 vector (Invitrogen, Paisley, UK) with a zeocin selection marker. Cells were transfected with the rat Fut2 cDNA using Lipofectamin 2000™ according the manufacturer's instructions (Invitrogen). After selection in hygromycin containing medium (0.4 mg/ml), cells were cloned by limiting dilution and a clone strongly expressing cell surface H antigen, as detected using the FITC-labeled UEA-I lectin, was selected. This clone was further transfected with the rat A

enzyme cDNA. Transfectants were then selected in zeocin containing medium (0.6 mg/ml). After cloning by limiting dilutions, a clone strongly expressing the A histo-blood group antigen was selected. Control transfected cells were prepared by transfection with the empty vectors. These stable transfectants were cultured in RPMI 1640 supplemented with 10% decompemented fetal-calf serum, 2 mM L-glutamine, 10 µg/ml free nucleotides, 100 U/ml penicillin, and 100 µg/ml streptomycin, 0.4 mg/ml hygromycin and 0.6 mg/ml zeocin. To obtain expression of the SARS-CoV S protein, control CHO cells, transfected CHO cells expressing the H antigen, and double transfected CHO cells expressing the A antigen were transfected with the previously described pEGFP-N1 vector containing an S protein-EGFP construction (CMV-SG) (Chou et al. 2005). Stable transfectants were obtained after selection with 1 mg/ml G418 (Gibco). Since expression of the S protein was progressively lost, even when cells were continuously cultured in the presence of G418, part of the experiments were performed 48 h after transient transfection with the CMV-SG vector.

Cells were passaged at confluence after dispersal with 0.025% trypsin in 0.02% EDTA and routinely checked for mycoplasma contamination by Hoechst 33258 (Sigma, St Louis , MO, USA) labeling.

#### *Flow cytometry, immunofluorescence and western blot analyses*

Cells at near confluence were detached by a brief 0.025% trypsin/0.02% EDTA treatment. Viable cells,  $2 \times 10^5$  per well of 96 culture microtiter plates were incubated with primary anti-H type 2 or broad reactive anti-A monoclonal antibodies 19-0LE and 3-3A respectively, in PBS containing 0.1% gelatin for 30 minutes at 4°C (Bara et al. 1988, Mollicone et al. 1996). After three washes with this same buffer, a 30 min incubation with the secondary anti-mouse IgG FITC-labeled antibody (Sigma) was performed at 4°C. After washing, fluorescence analysis was performed on a FACSCalibur (Becton-Dickinson, Heidelberg, Germany). Expression of the EGFP-S protein construct was detected by its autofluorescence on the FL1 channel. CHO cells transfectants, cultivated on glass lamellae, were fixed by addition of 2% formaldehyde in culture medium for 10 minutes.

After washing in PBS, the cell monolayer was incubated with the anti-A mAb 3-3A at 0.5  $\mu\text{g}/\text{ml}$  in PBS for 1 hour, washed thrice in PBS before incubation for 30 minutes with TRITC-labeled anti-mouse IgG (Sigma) diluted at 1/400. After 3 final washings in PBS, slides were mounted in Mowiol and observed on a Leica TCS SP (Heidelberg, Germany) confocal fluorescence microscope. Negative controls were incubated with the secondary antibody alone.

Confluent cells (CHO FUT2 simple transfectants, CHO FUT2/A double transfectant and CHO FUT2/A/PS triple transfectants) were rinsed with ice cold PBS pH 7,2 then recovered by scraping. After washing with ice cold PBS, cells were solubilized in 50 mM potassium phosphate, pH=6.0, containing 2% (v/v) triton X-100 on ice for 30 min. Following a centrifugation at 13,000g for 10 min, the supernatant was collected and protein concentration was determined using the BC assay protein quantitation kit (Uptima, Montluçon, France). Thirty mg total proteins of each extract were separated on 8% SDS-PAGE under reducing conditions and electrotransferred to immobilon P sheets (Millipore, Bedford, MA). Immunoblots were saturated for 1 hour at room temperature with western blocking reagent (Roche Diagnostics GmbH, Mannheim, Germany) and incubated overnight at 4°C with the anti-A 3-3A Mab at 2  $\mu\text{g}/\text{ml}$  in western blocking reagent. Following 3 washes for 15 min with TBS containing 0,05% tween 20, strips were incubated with secondary horseradish peroxidase-labeled anti-mouse IgG (H+L) (Beckman Coulter, CA) diluted at 1/5000, for one hour at room temperature. After three final washes, detection was performed with a chemiluminescence kit (Roche Diagnostics GmbH).

#### *Human plasma preparation*

Plasma samples from two blood group O healthy donors with anti-A titers 1/256 by classical hemagglutination were used. The presence of anti-A natural antibodies was confirmed by ELISA on synthetic A type 2 tetrasaccharide conjugated to polyacrylamide (Lectinity, Moscow, Russia). The polyacrylamide conjugate at 10  $\mu\text{g}/\text{ml}$  in carbonate-bicarbonate buffer, 100 mM, pH 9.6 was coated onto Maxisorp ELISA plates (NUNC,

Roskilde, Denmark) by overnight incubation at 37°C. After 3 washes with PBS containing 0.05% Tween 20 (TPBS), wells were incubated with 5% defatted milk in PBS for 1 h at 37°C. PBS 2-fold serially diluted plasma samples were then incubated for 1 h at 37°C. After washing with TPBS, peroxidase-labeled anti-human IgG (H+L) (Uptima, Montluçon, France) diluted at 1/10 000 were incubated for 1 h at 37°C. After final washes with TPBS, reactivities were detected using TMB (5-tetramethylbenzidine) as substrate (BD Bioscience, San Jose CA, USA) and read at  $A_{450}$  nm. To adsorb the anti-A natural antibodies, 1600  $\mu$ l of plasma diluted 1/4 in PBS were incubated onto 120 mg of silica beads conjugated with either synthetic A type 2 tetrasaccharide or a methyl group (kind gift from the late Pr. R.U. Lemieux, Edmonton, Canada). After a 1 h incubation under gentle agitation and centrifugation at 13,000 g for 10 minutes, the supernatant was collected and kept at 4°C until used.

#### *Cell-based binding assay*

Vero E6 cells were grown to a confluent layer in wells of 48 wells plate (NUNC). The CHO transfected cells were labeled with Hoechst 33258 at 2  $\mu$ M for 5 min, rinsed three times with PBS and then suspended after incubation with 0.025 % trypsin in 0.02 % EDTA.  $1 \times 10^5$  cells suspended in 500  $\mu$ l were gently laid onto the Vero cell layer in at least triplicate wells. After 2 h incubation at 37°C, wells were gently rinsed with culture medium three times and cells were fixed by a 10 minutes incubation with 2% formaldehyde. Adherent CHO cells were counted under an epifluorescence microscope (Zeiss, Jena, Germany) with a 10X objective. Three to six fields/well were counted and results expressed as the mean number of cells per field. In the antibody blocking experiments, monoclonal antibodies or human plasma samples were added to the transfected CHO cells suspension before seeding on the Vero cell layer.

#### *Modelling the effect of protection by natural anti-A or B antibodies on the virus transmission in populations*

A deterministic SIR (susceptible, infectious, recovered) model of the transmission dynamics of SARS that takes into account the effect of the protection by anti-histo-blood

group natural antibodies was developed. In this model, the whole population was divided according to the different blood groups (A, B, O or AB) into four interacting subpopulations  $N_A$ ,  $N_B$ ,  $N_O$  and  $N_{AB}$ . For each subpopulation  $N = S(t) + I(t) + R(t)$  where  $S(t)$ ,  $I(t)$ , and  $R(t)$  represent, for each blood group, the number of individuals susceptible, infected or recovered at time  $t$ , respectively. Four different populations reflecting the variability in blood groups distributions were studied (Table 1): Chinese from Honk Kong, Aïnu from Japan, Amerindians and white Americans from the USA.

	[O]	[A]	[B]	[AB]
Aïnu (Japan)	0.17	0.32	0.337	0.173
Chinese (Hong Kong)	0.422	0.178	0.333	0.067
Caucasians (USA)	0.45	0.40	0.111	0.039
Amerindians (USA)	0.79	0.16	0.046	0.004

**Table 1.** Distribution of ABO blood groups in four different populations used to model the impact of the ABO effect on SARS transmission. The frequencies of each phenotype were obtained from Mourant (Mourant 1983).

Mathematical details are given in the appendix; the program was written in C++, a one-day step was used for all the simulations. In each simulation, changes in the number of infected individuals (cases) over time were determined.

## Funding

PG was supported by a grant from the Office National de la Chasse et de la Faune Sauvage (ONCFS).

## **Acknowledgements**

We thank Dr. J Bara (INSERM, Villejuif, France) for providing the anti-H and anti-A mAbs and Dr. A Blancher for helpful discussions.

## **Abbreviations**

ACE2, angiotensin converting enzyme 2 ; CHO, chinese hamster ovary ; EGFP, enhanced green fluorescent protein ; HBGA, histo-blood group antigen, mAb, monoclonal antibody ; PBS, phosphate buffered saline ; RBD, receptor binding domain ; RBM, receptor binding motif ; S protein, spike protein ; SARS-CoV, severe acute respiratory syndrome coronavirus.

## References

- Anderson RM, Fraser C, Ghani AC, Donnelly CA, Riley S, Fergusson NM, Leung GM, Lam TH, Hedley AJ. 2004. Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. *Phil Trans R Soc Lond B*, 359: 1091-1105.
- Arendrup M, Hansen JE, Clausen H, Nielsen C, Mathiesen LR, Nielsen JO. 1991. Antibody to histo-blood group A antigen neutralizes HIV produced by lymphocytes from blood group A donors but not from blood group B or O donors. *AIDS*, 5: 441-444.
- Babcock GJ, Eshaki DJ, Thomas WD, Ambrosino DM. 2004. Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor. *J Virol*, 78: 4552-4560.
- Bara J, Gautier R, Le Pendu J, Oriol R. 1988. Immunochemical characterization of mucins. Polypeptide (M1) and polysaccharide (A and Le<sup>b</sup>) antigens. *Biochem J*, 254: 185-193.
- Bureau V, Marionneau S, Cailleau-Thomas A, Le Moullac-Vaydie B, Liehr T, Le Pendu J. 2001. Comparison of the three rat GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferases FTA, FTB and FTC. *Eur J Biochem*, 268: 1006-1019.
- Cailleau-Thomas A, Le Moullac-Vaidye B, Rocher J, Bouhours D, Szpirer C, Le Pendu J. 2002. Cloning of a rat gene encoding the histo-blood group A enzyme. Tissue expression of the gene and of the A and B antigens. *Eur J Biochem*, 269: 4040-4047.
- Chakraborti S, Prabakaran P, Xiao X, Dimitrov DS. 2005. The SARS coronavirus S glycoprotein receptor binding domain: fine mapping and functional characterization. *Virol J*, 2: 73.
- Chen J, Subbarao K. 2007. The immunobiology of SARS. *Ann Rev Immunol*, 25: 443-472.
- Cheng Y, Cheng G, Chui CH, Lau FY. 2005. ABO blood group and susceptibility to severe acute respiratory syndrome. *JAMA*, 293: 1450-1451.
- Chou C-F, Shen S, Tan Y-J, Fielding BC, Tan THP, Fu J, Xu Q, Lim SG, Hong W. 2005. A novel cell-based binding assay system reconstituting interaction between SARS-CoV S protein and its cellular receptor. *J Virol Methods*, 123: 41-48.
- Dürrbach A, Baple E, Preece AF, Charpentier B, Gustafsson K. 2007. Virus recognition by specific natural antibodies and complement results in MHC I cross-presentation. *Eur J Immunol*, 37: 1254-1265.

- Greenwell P. 1997. Blood group antigens: molecules seeking a function? *Glycoconj J*, 14: 159-173.
- Gu J Korteweg C. 2007. Pathology and pathogenesis of severe acute respiratory syndrome. *Am J Pathol*, 170: 1136-1147.
- Han DP, Lohani M Cho MW. 2007. Specific asparagine-linked glycosylation sites are critical for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry *J Virol*, 81: 12029-12039.
- He JF Chinese SMEC. 2004. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. *Science*, 303: 1666-1669.
- Keyaerts E, Vijgen L, Pannecouque C, Van Damme E, Peumans W, Egberink H, Balzarini J Van Ranst M. 2007. Plant lectins are potent inhibitors of coronaviruses by interfering with two targets in the viral replication cycle. *Antiviral Res*, 75: 179-187.
- Krokhin O, Li Y, Andonov A, Feldmann H, Flick R, Jones S, Stroehrer U, Bastien N, Dasuri KV, Cheng K, et al. 2003. Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. *Mol Cell Proteomics*, 2: 346-356.
- Li F, Li W, Farzan M Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science*, 309: 1864-1868.
- Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, et al. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, 426: 450-454.
- Li W, Wong S-K, Li F, Kuhn JH, Huang I-C, Choe H Farzan M. 2006. Animal origins of the severe acute respiratory syndrome coronavirus: insights from ACE2-S-protein interactions. *J Virol*, 80: 4211-4219.
- Marionneau S, Cailleau-Thomas A, Rocher J, Le Moullac-Vaidye B, Ruvoën-clouet N, Clément M Le Pendu J. 2001. ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie*, 83: 565-573.
- Mollicone R, Cailleau A, Imberty A, Gane P, Perez P Oriol R. 1996. Recognition of the blood group H type 2 trisaccharide epitope by 28 monoclonal antibodies and three lectins. *Glycoconj J*, 13: 263-271.
- Mourant AE. 1983. Blood relations: blood groups and anthropology. Oxford University Press, Oxford.

- Neil SJ, McKnight A, Gustafsson K, Weiss RA. 2005. HIV-1 incorporates ABO histo-blood group antigens that sensitize virions to complement-mediated inactivation. *Blood*, 105: 4693-4699.
- Preece AF, Strahan KM, Devitt J, Yamamoto FF, Gustafsson K. 2002. Expression of ABO or related antigenic carbohydrates on viral envelopes leads to neutralization in the presence of serum containing specific natural antibodies and complement. *Blood*, 99: 2477-2482.
- Seymour RM, Allan MJ, Pomiankowski A, Gustafsson K. 2004. Evolution of the human ABO polymorphism by two complementary selective pressures. *Proc Biol Sci*, 271: 1065-1072.
- Stäger S, Alexander J, Kirby AC, Botto M, Rooijen NV, Smith DF, Brombacher F, Kaye PM. 2003. Natural antibodies and complement are endogenous adjuvants for vaccine-induced CD8+ T-cell responses. *Nat Med*, 9: 1287-1292.
- van der Hoek L. 2005. Human coronaviruses: what do they cause? *Antivir Ther*, 12: 651-658.
- Wong SK, Li W, Moore MJ, Choe H, Farzan M. 2004. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem*, 279: 3197-3201.
- Xiao X, Chakraborti S, Dimitrov AS, Gramatikoff K, Dimitrov DS. 2003. The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem Biophys Res Commun* 312: 1159-1164.
- Ying W, Hao Y, Zhang Y, Peng W, Qin E, Cai Y, Wei K, Wang J, Chang G, Sun W, et al. 2004. Proteomic analysis on structural proteins of severe acute respiratory syndrome coronavirus. *Proteomics*, 4: 492-504.

## Additional material

The deterministic model is described in detail, and the set of differential equations is given for each subpopulation. The model allows for the computation of the rate of transfer of individuals from each state (S, I or R) to any other state over time. The system of differential equations was numerically solved by using a seventh-order Runge-Kutta method.

Notations

**S<sub>A</sub>(t)** : Number of susceptible blood group A individuals

**I<sub>A</sub>(t)** : Number of infected blood group A individuals

**R<sub>A</sub>(t)** : Number of recovered blood group A individuals

The other blood groups have similar notations.

*Derivation of mortality rate* : From data of the major 2003 outbreak that infected about 8000 people from which 774 died (*He JF. et al., Science, 2004;303:1666–1669*).

$m$ : mortality rate due to the disease =  $\left[ \frac{774}{8000} \right] / 30 = 0,003225$  per day assuming a mean duration of disease of 30 days (*Anderson RM et al., Trans. R. Soc. Lond. B., 2004;359:1091-1105*).

The other parameter values were calculated from the published work on the hospital outbreak that occurred in March 2003 in Hong Kong (*Cheng et al. 2005*) as follows:

1) Total sample

- observed proportion of infected during 8 days =  $34/45 = 0.7555 \Rightarrow$  per day:  $0.7555/8 = 0.0944 = \lambda = c\beta$  representing the mean number of infectious contacts per person per day.

-  $c$ : Mean number of contacts per day assumed to be equal to 10. This value was chosen from two different publications, a ward based study (*Cheng et al. 2005*) and a large scale study (*Mossong J, PloS Medicine, 2008; 5:e74*) where the mean daily number of contacts ranged from 7 to 19.

$\beta_{ij}^i$  : Transmission coefficient (transmission rate of the disease for each contact) from a blood group  $j$  individual to a blood group  $i$  individual =  $1 / c = 0.00944$

## 2) Blood group O sample

- observed proportion of infected per day =  $11/45/8 = 0.0305$
- c remains unchanged
- $\beta_j^i = 0.00305$

## 3) Other blood group sample

- observed proportion of infected per day =  $23/45/8 = 0.0639$
- c remains unchanged
- $= \beta_j^i 0.00639$

$\gamma$  : recovery rate =  $1/\text{average duration of disease} = 1/30 = 0.0333$

## Blood group A individuals

$$\frac{dS_A(t)}{dt} = - \left[ cS_A \left( \beta_A^A \frac{I_A}{N_A} + \beta_B^A \frac{I_B}{N_B} + \beta_{AB}^A \frac{I_{AB}}{N_{AB}} + \beta_0^A \frac{I_0}{N_0} \right) \right]$$

$$\frac{dI_A(t)}{dt} = \left[ cS_A \left( \beta_A^A \frac{I_A}{N_A} + \beta_B^A \frac{I_B}{N_B} + \beta_{AB}^A \frac{I_{AB}}{N_{AB}} + \beta_0^A \frac{I_0}{N_0} \right) \right] - [I_A(\gamma + m)]$$

$$\frac{dR_A(t)}{dT} = [\gamma I_A]$$

## Blood group B individuals

$$\frac{dS_B(t)}{dt} = - \left[ cS_B \left( \beta_A^B \frac{I_A}{N_A} + \beta_B^B \frac{I_B}{N_B} + \beta_{AB}^B \frac{I_{AB}}{N_{AB}} + \beta_0^B \frac{I_0}{N_0} \right) \right]$$

$$\frac{dI_B(t)}{dt} = \left[ cS_B \left( \beta_A^B \frac{I_A}{N_A} + \beta_B^B \frac{I_B}{N_B} + \beta_{AB}^B \frac{I_{AB}}{N_{AB}} + \beta_0^B \frac{I_0}{N_0} \right) \right] - [I_B(\gamma + m)]$$

$$\frac{dR_B(t)}{dT} = [\gamma I_B]$$

## Blood group AB individuals

$$\frac{dS_{AB}(t)}{dt} = - \left[ cS_{AB} \left( \beta_A^{AB} \frac{I_A}{N_A} + \beta_B^{AB} \frac{I_B}{N_B} + \beta_{AB}^{AB} \frac{I_{AB}}{N_{AB}} + \beta_0^{AB} \frac{I_0}{N_0} \right) \right]$$

$$\frac{dI_{AB}(t)}{dt} = \left[ cS_{AB} \left( \beta_A^{AB} \frac{I_A}{N_A} + \beta_B^{AB} \frac{I_B}{N_B} + \beta_{AB}^{AB} \frac{I_{AB}}{N_{AB}} + \beta_0^{AB} \frac{I_0}{N_0} \right) \right] - [I_{AB}(\gamma + m)]$$

$$\frac{dR_{AB}(t)}{dT} = [\gamma I_{AB}]$$

Blood group O individuals

$$\frac{dS_0(t)}{dt} = - \left[ cS_0 \left( \beta_A^0 \frac{I_A}{N_A} + \beta_B^0 \frac{I_B}{N_B} + \beta_{AB}^0 \frac{I_{AB}}{N_{AB}} + \beta_0^0 \frac{I_0}{N_0} \right) \right]$$

$$\frac{dI_0(t)}{dt} = \left[ cS_0 \left( \beta_A^0 \frac{I_A}{N_A} + \beta_B^0 \frac{I_B}{N_B} + \beta_{AB}^0 \frac{I_{AB}}{N_{AB}} + \beta_0^0 \frac{I_0}{N_0} \right) \right] - [I_0(\gamma + m)]$$

$$\frac{dR_0(t)}{dT} = [\gamma I_0]$$

## ***b - Conclusions et perspectives***

Depuis la découverte des NAb anti-HBGA par Landsteiner, aucun rôle précis n'a pu leur être attribué. Jusqu'à récemment, ils étaient systématiquement perçus comme « gênants », notamment à cause des problèmes qu'ils engendrent dans les services d'obstétrique, de transfusion et de transplantation. Toutefois, la mise en évidence de leur rôle potentiel dans la réponse immunitaire antivirale a ouvert la voie à de nouvelles perspectives réhabilitant l'intérêt de leur présence chez l'homme (K Gustafsson et coll. 2005). Dans cette optique, notre étude a eu pour but d'estimer la capacité des NAb anti-ABH à inhiber la fixation de la protéine S du SRAS-CoV sur son récepteur cellulaire et de modéliser l'impact d'une telle reconnaissance sur la dynamique de transmission de ce virus dans des populations aux fréquences du polymorphisme ABO distinctes.

Pour cela, nous avons utilisé un modèle cellulaire de fixation virus/récepteur permettant la co-expression de la protéine S et de l'HBGA A. Nous avons tout d'abord vérifié que la modification du profil de glycosylation des cellules n'avait pas d'impact sur la qualité de la fixation de la protéine virale sur l'ACE2. Nous avons ensuite montré que cette fixation peut être inhibée en co-incubant ces mêmes cellules avec un anticorps monoclonal anti-A, mais aussi en utilisant les plasmas d'individus de groupe sanguin O. Dans ce cas précis, la spécificité de l'inhibition de la fixation par les NAb anti-A a été confirmée en observant de nouveau une fixation lorsque les cellules sont co-incubées avec ces plasmas déplétés de leurs anticorps anti-A à l'aide d'antigène A synthétique immobilisé sur billes de silicate.

Ces résultats pourraient expliquer les données épidémiologiques observées dans l'hôpital chinois pendant l'épidémie de SRAS de 2002/2003. Afin de vérifier ceci, nous avons utilisé un modèle mathématique déterministe considérant l'action des NAb comme inhibitrice de la transmission du virus. Il nous a permis de découvrir que la dynamique de transmission est fortement corrélée au profil ABH de la population considérée, mais surtout aux titres relatifs en NAb anti-ABH des individus la composant. Ces conclusions font pour la première fois le lien entre données épidémiologiques et action antivirale des NAb anti-ABH observée *in vitro*. Ce lien est d'autant plus solide que le SRAS-CoV se réplique dans des cellules capables d'exprimer naturellement les HBGA ABH, ce qui n'était pas le cas des virus étudiés précédemment *in vitro* dans ce contexte, comme le HIV et le virus de la rougeole.

Nous avons également clairement mis en évidence que plus les titres en anticorps anti-ABH sont élevés dans la population, plus l'épidémie met du temps à se propager et qu'elle est alors associée à un nombre moins important d'individus infectés à un temps donné. Ce phénomène ne fait pas diminuer drastiquement le nombre total de personnes

infectées de manière spontanée mais peut faciliter la mise en place de moyens prophylactiques et ainsi contribuer à limiter plus efficacement la flambée épidémique. La vaccination anti-ABH pourrait dans ce cas présenter un grand intérêt pour s'assurer que les titres anti-A et/ou B soient suffisamment élevés, d'autant plus que dans les pays dits développés, une « bonne » qualité de vie rime de nos jours avec quasi-asepsie, limitant probablement ainsi les contacts avec les antigènes exogènes à l'origine des NAb anti-ABH.

Cette protection innée antivirale pourrait s'appliquer à d'autres virus enveloppés possédant des glycoprotéines structurales exposées et se répliquant dans des cellules capables de produire des HBGA ABH ou tout autre glycanne présentant un polymorphisme associé à la production de NAb.

Ce mécanisme de protection pourrait également participer à limiter le passage de la barrière d'espèce par les virus. En effet, chaque espèce possède son propre glycome. Les glycannes exogènes greffés sur les virus permettraient leur élimination précoce sans qu'ils puissent avoir le temps de s'adapter à ce nouvel hôte potentiel. L'extinction chez l'homme du gène codant l' $\alpha$ 1,3 galactosyltransférase et responsable de la synthèse du xénoantigène pourrait trouver son origine au regard d'un tel mécanisme.

# ***DISCUSSION GÉNÉRALE***

Dans la nature, n'importe laquelle des cellules prise au hasard porte à sa surface une strate épaisse de glycanes variés et propres à son identité cellulaire (Bishop et coll. 2007). Outre leurs implications dans de nombreux processus biologiques, ces sucres en contact direct avec l'environnement peuvent représenter des cibles pour les pathogènes et notamment pour les virus (Sharon 1996).

En effet, nombre de pathogènes initient l'infection de leur cellule hôte par la reconnaissance spécifique d'un glycanne afin d'établir le premier contact physique fort (fixation) nécessaire à la mise en place du processus infectieux. Parmi les divers glycanes exprimés dans le monde du vivant, l'acide sialique représente une cible de choix pour les pathogènes. En effet, il est porté en position terminale sur plus d'une cinquantaine de chaînes glycaniques distinctes et est présent dans le glycome d'un très grand nombre d'espèces (deutérostomiens métazoaires, bactéries et quelques mollusques) (Angata et coll. 2002). Ces caractéristiques font de lui l'un des sucres les plus universels du vivant. Son expression représente un caractère vital, notamment chez les mammifères pour qui sa perte d'expression se traduit par une déficience sévère du système immunitaire ou des perturbations graves et létales au cours du développement embryonnaire. Du fait de sa forte expression et de son caractère vital, il est une cible de choix pour de nombreux pathogènes dont une liste non exhaustive a été dressée par Angata et Varki (Angata et coll. 2002). Malgré cette association à un fort désavantage face à ces pathogènes, le maintien des structures sialylées au cours de l'évolution peut être expliqué par le fait que leur perte d'expression ou bien la modification globale du glycanne portant l'acide sialique sont souvent plus néfastes à l'espèce que leur utilisation comme récepteur cellulaire par des microorganismes infectieux.

L'histoire évolutive des HBGA ABH est quant à elle complètement différente. En effet, depuis leur co-découverte par Landsteiner et Janský, aucun rôle physiologique n'a pu leur être clairement attribué. Ces antigènes sont exprimés par les cellules de nombreux organismes très éloignés d'un point de vue phylogénétique (Oriol et coll.

1999). Ils semblent donc être impliqués dans un ou des processus biologiques anciens et utiles aux organismes qui les expriment. Ils ne représentent en rien une « excentricité stérile » de la Nature du fait de leur maintien, de leur diversification ainsi que de l'augmentation du nombre de tissus les exprimant au cours de l'évolution des espèces. La mise en évidence récente de l'utilisation des HBGA ABH comme récepteurs cellulaires par certains Calicivirus, ainsi que nos travaux sur l'immunité innée antivirale induite par l'expression des HBGA ABH soutiennent l'idée que l'évolution des HBGA ABH est intimement liée aux virus et pas seulement aux bactéries comme d'autres l'ont suggéré.

De même que Varki s'est inspiré du titre de l'essai de Dobzhansky pour intituler l'un des sien : « *Rien en Glycobiologie n'a de sens, excepté à la lumière de l'évolution* », je m'inspire de Varki pour dire aujourd'hui que : « Rien en Glycobiologie n'a de sens, excepté à la lumière de la coévolution hôte/pathogène », et en particulier de la coévolution virus/HBGA ABH.

### **Hypothèse de coévolution virus/glycane**

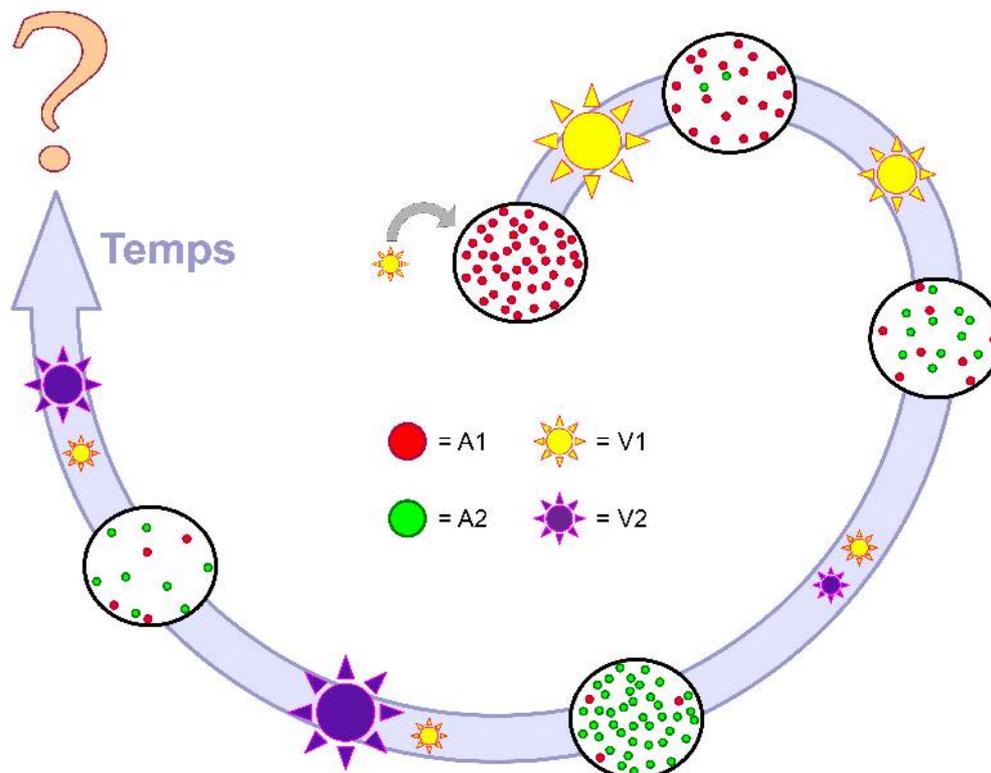
Au vu des nombreux mécanismes impliquant les glycanes dans les interactions hôte/virus, il est naturel de faire l'hypothèse que l'hôte puisse évoluer afin de moduler l'expression de certains de ses glycanes en réponse à la pression de sélection exercée par des virus très pathogènes. Dans ce cas, nous pouvons alors parler d'une coévolution virus/glycanes. Toutefois, celle-ci est difficile à évaluer aujourd'hui chez l'homme, du fait des moyens de prophylaxie que nous avons mis en place pour nous protéger des infections virales. Elle ne représente donc pour l'instant qu'une hypothèse à vérifier.

La coévolution au sens strict du terme correspond à un mécanisme de sélection réciproque entre deux espèces, comme celui qui peut exister entre un parasite et son hôte. Cette sélection réciproque est fréquence dépendante. En effet, un caractère conférant un avantage sélectif ne peut pas être maintenu s'il reste cantonné à l'échelle de l'individu. Il devra donc augmenter sa fréquence dans la population pour pouvoir persister dans le temps et ainsi devenir bénéfique à l'espèce. Afin d'illustrer ceci, prenons l'exemple d'un modèle simple où les caractères sensibilité/résistance de l'hôte et virulence du pathogène ne sont respectivement associés qu'à un seul gène chacun (Figure 30). Ce type d'évolution dit « gène pour gène » peut donner lieu à trois scénaris possibles (Lively et coll. 1995) :

- La coévolution continue indéfiniment avec une accumulation des différents mécanismes mis en place pour augmenter la résistance et la virulence.
- La coévolution est le résultat de cycles évolutifs où chaque adaptation rend obsolète la précédente.

- La coévolution n'est pas équitable et va tendre vers l'extinction du pathogène si l'hôte est plus adaptable que celui-ci ou des deux espèces si le parasite est plus adaptable que l'hôte.

**Figure 30 : Représentation schématique du mécanisme de coévolution**



**hôte/pathogène.**

Si l'hôte possède un allèle A1 à l'origine d'une sensibilité associée au gène de virulence V1 d'un pathogène, alors la population de l'hôte va drastiquement diminuer et la dissémination du pathogène va augmenter. Si un nouvel allèle A2 émerge dans la population de l'hôte et qu'il est associé à une résistance au pathogène codant V1, alors la fréquence relative d'A2 va augmenter dans la population. De son côté, le pathogène se transmettra moins bien du fait de la fragmentation de la population possédant A1 et de l'augmentation relative de la population possédant A2. En réponse, le pathogène va devoir évoluer à son tour vers un variant du gène de virulence V2 qui pourra alors permettre d'infecter la partie de la population de l'hôte précédemment résistante à V1 (= population A2). Dans notre contexte, les allèles A1 et A2 permettent la synthèse d'HBGA qui servent de ligands reconnus spécifiquement par les souches V1 et V2, respectivement.

La taille du symbole représentant le pathogène ( ) est proportionnelle à l'efficacité de sa transmissibilité à un temps donné.

Parmi les hypothèses émises afin d'expliquer l'évolution des glycosyltransférases, et en particulier la coévolution virus/glycane, celle de Gagneux et Varki apparaît être

l'une des mieux adaptée (Gagneux et coll. 1999). Celle-ci repose en effet sur un modèle simple qui peut facilement expliquer le polymorphisme des gènes codant les HBGA ABH comme le résultat de la pression de sélection exercée par les pathogènes les utilisant comme récepteur : l'hypothèse de « la Reine Rouge ».

### **Hypothèse de la « Reine Rouge »**

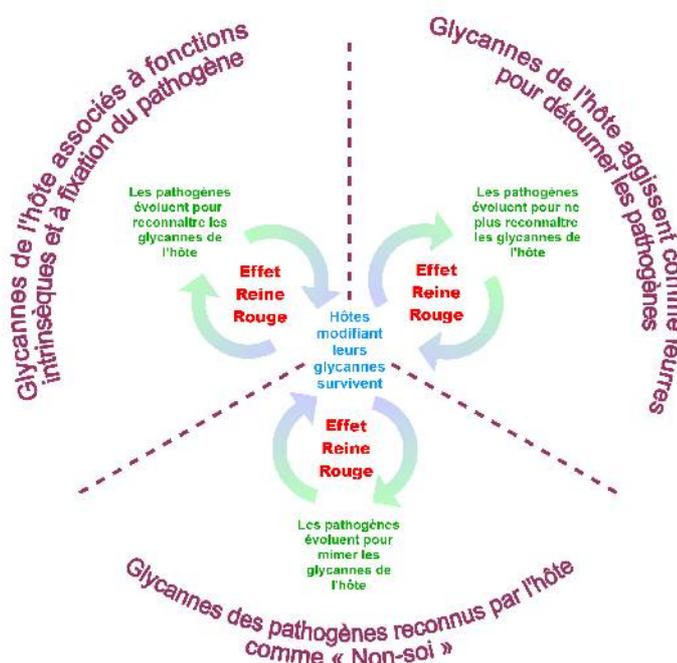
L'hypothèse de la « Reine Rouge » a été proposée par Leigh Van Valen en 1973 (Van Valen 1973). Il l'échafauda à partir de deux observations faites d'après l'étude des fossiles. Tout d'abord, il n'arrivait pas à mettre en évidence de corrélation entre le taux d'extinction des espèces et le temps depuis lequel celles-ci étaient apparues sur Terre. Il en déduisit que bien qu'une espèce soit établie depuis une très longue période, elle peut tout aussi bien disparaître dans un laps de temps très court. Il a nommé ce phénomène la « loi d'extinction constante ». Sa seconde observation fut que le taux d'extinction n'est pas le même entre les différents groupes d'espèces contemporaines, certains disparaissent en effet plus rapidement que d'autres. Il s'est alors demandé pourquoi la probabilité d'extinction semble être aléatoire pour chacune des espèces d'un même groupe alors qu'en même temps celle-ci diffère énormément entre les différents groupes d'espèces.

Pour répondre à cela, il a développé un modèle où deux groupes d'espèces variées sont mis en compétition dans un environnement où chacune des espèces doit évoluer pour s'adapter à ces nouvelles contraintes. Au bout d'un certain temps, l'une d'elles va acquérir un avantage sélectif sur les autres et ainsi pouvoir développer sa population. Les autres espèces vont devoir évoluer à leur tour afin de subsister à l'espèce nouvellement adaptée. Celles ne le pouvant pas sont condamnées à l'extinction à plus ou moins long terme. Pour celles qui vont évoluer, si elles le font suffisamment pour combler l'ensemble des désavantages qu'elles avaient accumulés vis à vis des espèces adaptées, alors elles se retrouveront au même niveau évolutif que ces dernières. Dans ce cas, on peut dire que les espèces ont évolué de manière différente sans pour autant que l'une d'elles ne prennent l'avantage sur les autres. Cette évolution est corrélée à la pression de sélection exercée par l'adaptation des autres espèces et non à une modification de l'environnement. Van Valen a nommé sa théorie l'hypothèse de « la Reine Rouge », ceci en référence à un épisode du livre de Lewis Carroll « De l'autre côté du miroir », au cours duquel Alice et la Reine Rouge se lancent dans une course effrénée. Alice demande alors : « *Mais, Reine Rouge, c'est étrange, nous courons vite et le paysage autour de nous ne change pas ?* » Et la reine de répondre : « *Nous courons pour rester à la même place.* »

Dans les mécanismes participant à l'effet Reine Rouge, la reproduction sexuée représente l'un des moteurs majeurs de la réponse adaptative de l'hôte face aux parades

évolutives développées par les pathogènes. Les virus et les microorganismes parasites sont très compétitifs d'un point de vue adaptatif lorsqu'on les compare aux espèces à longue durée de vie. Ils possèdent en effet un temps de génération très court, parfois des milliards de fois plus court que celui de leur hôte. Ils peuvent de ce fait très rapidement évoluer pour contrer les défenses mises en place par l'hôte. La reproduction sexuée va permettre à l'hôte de pouvoir augmenter la diversité génétique de sa progéniture grâce aux brassages inter-chromosomiques et aux recombinaisons génétiques et par conséquent augmenter sa capacité d'adaptation. Bien que la reproduction asexuée soit plus efficace (pas de partenaires) et n'induit pas d'aberrations génétiques (copie conforme du génome transmise à la descendance), celle-ci a été supplantée par la reproduction sexuée qui prédomine aujourd'hui chez les vertébrés. Toutefois, tout ne peut reposer sur elle dans le processus d'échappement aux pathogènes.

De ce fait, Gagneux et Varki émirent l'hypothèse que les glycanes de surface, de par leur diversité et la modulation plus ou moins grande de leur structure et tropisme d'expression, pouvaient agir de concours avec la reproduction sexuée dans l'adaptation de l'hôte aux pathogènes et ainsi représenter un acteur majeur de l'effet Reine Rouge (Figure 31) (Gagneux et coll. 1999).

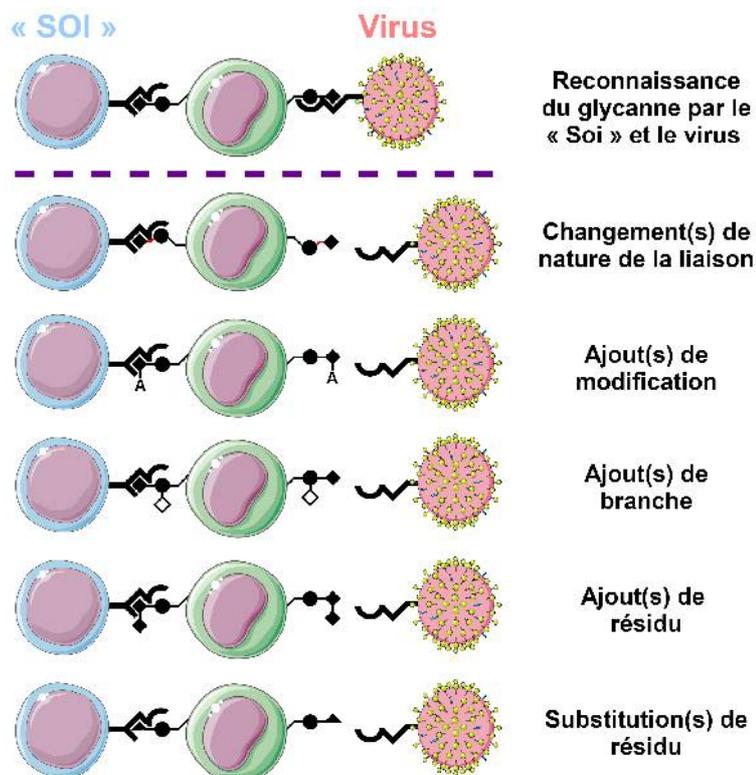


**Figure 31 : Effet Reine Rouge sur la diversification des glycanes de l'hôte en réponse aux adaptations des pathogènes.**

(D'après Varki 2006).

L'hôte possède en effet un champ d'action relativement important permettant l'évolution de ses glycosyltransférases afin d'échapper aux pathogènes sans pour autant

perturber les interactions endogènes impliquant des glycanes (Figure 32). Ces modifications peuvent notamment conduire : au changement de la nature des liaisons osidiques, à l'ajout de modification (ex : N-acetyl), à l'ajout de nouvelles branches, à l'extension par de nouveaux résidus ou à la substitution de résidus



**Figure 32 : Mécanismes d'échappement de la reconnaissance exogène des glycanes sans perte de la reconnaissance endogène.**

*Les différentes modifications de glycosylations (cellules vertes) permettent le maintien de la reconnaissance par les cellules du « soi » (cellules bleues), mais sont suffisantes pour abolir la reconnaissance par le pathogène (virus orange). (D'après Gagneux et coll. 1999)*

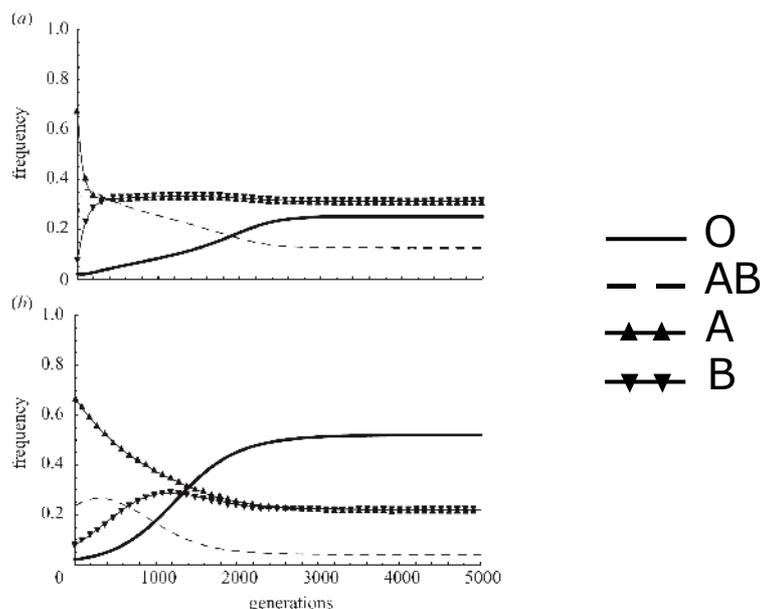
## Virus : moteurs de l'évolution des HBGA ABH ?

Lors de ma thèse, je me suis particulièrement intéressé aux inter-relations virus/glycanes et en particulier à celles concernant les virus et les HBGA ABH. Ces antigènes se prêtent en effet particulièrement à l'étude et à la meilleure compréhension du concept de coévolution virus/glycane développé par Gagneux et Varki.

En raison de l'absence de rôle physiologique avéré, ils peuvent être modifiés à façon en réponse aux pressions de sélection exercées par les pathogènes. Plusieurs modèles mathématiques permettant d'expliquer le maintien de ces structures au cours de l'évolution ont été développés en intégrant comme paramètres la reconnaissance des HBGA ABH par les pathogènes et/ou la reconnaissance des HBGA ABH « décorant » les pathogènes par les anticorps naturels de l'hôte (Figure 33). En effet, le maintien de ces

structures a été proposé depuis longtemps comme la conséquence des inter-relations entre HBGA ABH et pathogènes (Mourant 1982; Mourant et coll. 1978; Le Pendu 1989).

La première étude allant dans ce sens a pris en considération la reconnaissance de l'HBGA A par une bactérie et de l'HBGA B par une autre souche bactérienne (Fischer et coll. 1998) (Figure 33a). Le bilan de ce modèle est que la reconnaissance d'un glycanne confère un avantage sélectif face à un pathogène donné mais qu'en même temps il peut représenter un sérieux désavantage quand il est utilisé comme récepteur par une autre bactérie. Ce modèle permet d'expliquer le maintien du polymorphisme ABH qui s'équilibre dans la population en fonction des avantages/désavantages sélectifs face aux bactéries. Cependant, il ne permet pas d'expliquer les fréquences de chacun des phénotypes ABH dans les populations et notamment leurs fortes proportions en individu O ( $\approx 30\%$ ). Un modèle mathématique permettant d'expliquer les fréquences relatives de chacun des phénotypes ABH a été développé par la suite en intégrant la reconnaissance des HBGA ABH par les bactéries, mais également l'utilisation des NAb anti-ABH comme mécanisme de protection contre les virus glycosylés avec les HBGA ABH lors de leur réplication (Figure 33b) (R M Seymour et coll. 2004).



**Figure 33 : Modélisations mathématiques du rôle des pathogènes dans le maintien du polymorphisme ABH et de leurs fréquences dans les populations.**

(a) modèle n'intégrant que l'action des bactéries. (b) modèle intégrant l'action combinée des bactéries et virus glycosylés. (D'après R M Seymour et coll. 2004).

## Scénario de coévolution homme/NoV

La remarquable situation observée pour les NoV et les HBGA ABH pourrait correspondre à un stade coévolutif avancé dans le modèle de coévolution « gène pour

gène » (Figure 30 p.140). En effet, il est observé que les différentes souches de NoV se fixent sur des HBGA distincts correspondant à des sous-groupes différents de la population. Réciproquement, chacun des individus composant cette population peut être reconnu par au moins une des souches de NoV.

La pression de sélection exercée par les NoV paraît faible à l'heure actuelle dans les pays développés, ces virus y étant considérés comme modérément pathogènes et associés à une très faible mortalité. Toutefois, une méta-analyse épidémiologique toute récente indique que dans les pays en voie de développement, les NoV sont responsables de la mort d'au moins 200 000 enfants par an (Patel et coll. 2008). Ce chiffre indique que leur impact est loin d'être négligeable et que sans les moyens modernes de prise en charge médicale, ils pourraient contribuer au maintien du polymorphisme des HBGA ABH, comme dans le modèle de coévolution « gène pour gène ».

Un autre aspect qui pourrait avoir eu des conséquences importantes sur la coévolution NoV/HBGA ABH apparaît dans ce modèle. En effet, il montre que l'émergence d'une souche très virulente dans une population d'individus sensibles conduit progressivement à sa décimation et par voie de conséquence à l'extinction du pathogène, sauf si celui-ci peut diminuer rapidement sa virulence. Le fractionnement progressif de la population réceptive associé à l'apparition potentielle d'individus résistants contribue en effet à augmenter le coût de la transmission du virus, du fait de l'augmentation du nombre de contacts improductifs avec l'hôte. Par conséquent, le fractionnement de la population contribue à sélectionner positivement les souches de moindre virulence et négativement les souches les plus virulentes qui ne peuvent se maintenir. Ainsi le polymorphisme des HBGA pourrait avoir contraint les NoV à évoluer vers une virulence modérée et continuerait à diminuer la probabilité de réémergence de souches hautement virulentes. Cette hypothèse a été présentée dans une revue récente (Le Pendu et coll. 2006).

Cette hypothèse conduit également à proposer le concept de **protection/immunité innée de groupe**. En effet, un caractère ABO, sécréteur ou Lewis particulier n'est ni bon ni mauvais à l'échelle de l'individu, puisque ces phénotypes confèrent la résistance à une souche donnée et en même temps la sensibilité à une autre souche. C'est la résultante au niveau du groupe qui importerait puisque la diversité de l'hôte protège la population car tous les individus ne sont pas égaux lors d'une épidémie et que le virus est toujours contraint de maintenir une virulence modérée.

## Importance du modèle lapin/RHDV

Il est pourtant difficile de tester et valider ce genre d'hypothèse en raison de leur caractère historique. Il est en effet difficile, voir impossible, d'évaluer rétrospectivement un scénario de coévolution. Toutefois un autre Calicivirus, le RHDV, pourrait représenter un candidat de choix afin de tester le modèle de coévolution virus/HBGA, ainsi que le concept d'immunité innée de groupe.

Virus émergent extrêmement virulent chez le lapin, il partage avec les NoV la propriété de se fixer sur un HBGA, dans son cas l'HBGA H type 2. Il exerce donc depuis un court laps de temps une pression de sélection énorme sur la qualité de l'expression de ce glycanne. Jusqu'à récemment l'expression de cet HBGA était considérée comme identique chez l'ensemble des lapins. Notre travail a permis de montrer qu'il existe effectivement dans la nature des lapins de garenne exprimant fortement cet antigène, néanmoins d'autres ne l'expriment que faiblement à pas du tout (non sécréteur-like). En accord avec le modèle de coévolution qui vient d'être discuté, la fréquence des animaux « non sécréteur-like » est d'autant plus forte dans les populations lorsqu'elles ont été fortement affectées par les épidémies de RHDV. Cette observation suggère que les épidémies successives et rapprochées ont entraîné la sélection d'animaux résistants par défaut d'expression du ligand au virus. Toujours selon le modèle de coévolution « gène pour gène », l'augmentation de la fréquence d'animaux génétiquement résistants devrait conduire, d'une part à l'apparition de souches à la virulence diminuée et capables de reconnaître d'autres structures glycaniques, d'autre part.

Toutefois pour véritablement montrer que les lapins de phénotype « non sécréteur-like » sont bien résistants aux premières souches de RHDV, il est nécessaire de définir le(s) mécanisme(s) génétique(s) responsable(s) de ce phénotype, puis de le(s) valider en testant expérimentalement la résistance au RHDV des animaux ainsi différenciés. Afin de tendre vers cet objectif, nous avons séquencé les parties codantes des gènes d' $\alpha$ 1,2 fucosyltransférases chez le lapin de garenne. Nous avons ainsi mis en évidence qu'un allèle du gène *Sec1* était associé à la survie des animaux lors d'une épidémie de RHDV. Néanmoins, cet allèle ne peut pas être directement mis en cause dans la résistance au virus. Il nous faut donc maintenant rechercher quelle(s) mutation(s) du locus des  $\alpha$ 1,2 fucosyltransférases associée(s) à cet allèle pourrai(en)t altérer l'expression de l'une de ces enzymes et tout particulièrement de Fut2. Parallèlement, il sera important de tester la spécificité de nouvelles souches de RHDV vis-à-vis des HBGA. Ces souches ayant remplacé la souche originale, on peut supposer qu'elles n'auront pas les mêmes spécificités pour les glycannes du fait de la modification de la structure des populations de lapins de garenne et de la forte sélection du phénotype « non sécréteur-like ». Enfin, lorsqu'il aura été possible de sélectionner des lapins

d'élevage en fonction de leurs phénotypes et génotypes pour les HBGA, nous pourrions tester l'effet de leur polymorphisme sur la résistance au virus par des infections en conditions expérimentales. Les résultats exposés dans ma thèse représentent une première étape dans la validation du modèle de coévolution « gène pour gène » suggéré entre les NoV et les HBGA.

### **Impact des glycoprotéines du lait humain sur la fixation des NoV**

Le lait humain semble pouvoir protéger les nourrissons des gastro-entérites à NoV et nous avons mis en évidence qu'il contient des molécules capables d'inhiber la fixation de ces virus sur les HBGA. Les molécules participant à cette inhibition correspondent à des glycoprotéines solubles de type mucine et non pas à des oligosaccharides libres. Parmi ces glycoprotéines, la BSSL s'avère être celle qui possède les meilleures capacités inhibitrices du lait humain. Toutefois, cette capacité d'inhibition et de protection (par extrapolation) dépendra des phénotypes combinés de la mère et de l'enfant, ainsi que de la spécificité de la souche virale pour les HBGA.

En effet, plusieurs cas de figure peuvent se présenter. Par exemple, si la souche de NoV reconnaît l'antigène H, un enfant de phénotype « non sécréteur » ne pourra être infecté du fait de son incapacité à produire le ligand cellulaire du virus. Dans le cas où l'enfant ainsi que sa mère présente le phénotype « sécréteur », le lait de cette dernière pourra le protéger. En revanche si l'enfant présente le phénotype « sécréteur » (Se/se) et sa mère le phénotype « non sécréteur » (se/se), l'enfant ne pourra être protégé. Le phénotype « non sécréteur » étant plus rare que le phénotype « sécréteur » (20% contre 80% dans la population caucasienne), les enfants allaités seront donc majoritairement protégés des infections par les NoV présentant une affinité pour l'HBGA H. A l'inverse, si le virus reconnaît les individus de phénotype « non sécréteur », seules les mères présentant le phénotype « non sécréteur » pourront protéger leurs enfants « non sécréteur », les enfants « sécréteur » étant naturellement résistant au virus.

Un couple mère/enfant où l'une et l'autre présente le phénotype « non sécréteur » est statistiquement plus rare que le cas d'un couple où la mère est de phénotype « sécréteur » (Se/se) et l'enfant « non sécréteur » (se/se). Ainsi, les souches qui reconnaissent le phénotype le plus rare de la population acquièrent un avantage sélectif du fait du peu de chance qu'elles soient reconnues par un leurre maternel. Cette propriété pourrait en partie expliquer pourquoi ces souches continuent de circuler dans la population alors qu'elles ne peuvent infecter qu'une minorité d'individus.

Les glycoprotéines de type mucines du lait humain contribueraient donc d'une part à l'immunité innée individuelle en jouant le rôle de leurre à NoV et d'autre part à l'immunité innée de groupe du fait de l'impact différentiel qu'elles peuvent avoir sur la circulation des différentes souches de NoV. Une application pratique consisterait à ajouter

des molécules mimant les glycoprotéines du lait humain aux préparations lactées destinées aux nouveau-nés. En y incorporant ainsi l'ensemble des glycanes reconnus par les NoV, une protection pourrait être apportée aux très jeunes enfants en cas d'absence d'allaitement ou d'incompatibilité génétique du couple mère/enfant.

### **Protection par les anticorps naturels anti-A et B**

Il a été postulé depuis quelques années que les anticorps naturels anti-A et anti-B pourraient jouer un rôle important dans la protection contre les virus enveloppés. De par le processus répliatif particulier de ces derniers, les glycoprotéines virales peuvent subir les mêmes modifications que les glycoprotéines de la cellule hôte et donc porter les HBGA ABH si la machinerie enzymatique de la cellule le permet. Les glycoprotéines et glycolipides membranaires de la cellule hôte peuvent également être transférés au virus lors du bourgeonnement des virions néoformés.

Un modèle mathématique montrant l'impact que ce phénomène pourrait avoir sur le maintien du polymorphisme ABH a notamment été proposé par le groupe de Kenth Gustafsson. Cependant, il n'y avait jusqu'à présent aucune démonstration montrant que ces anticorps pouvaient réellement avoir un rôle dans l'immunité antivirale *in vivo*. Toutefois, cette même équipe a montré qu'*in vitro* ces anticorps sont capables de diminuer l'infectivité du virus de la rougeole et du HIV, si et seulement si ceux-ci étaient produits dans des cellules capables de synthétiser les HBGA A ou B (Preece et coll. 2002; Neil et coll. 2005). Ce phénomène pourrait également contribuer à l'initiation de la réponse immunitaire à médiation cellulaire lorsque les anticorps naturels sont associés au complément (Dürrbach et coll. 2007). Malheureusement, aucune donnée épidémiologique n'est venue montrer que ces observations faites *in vitro* pouvaient être pertinentes *in vivo*. En effet concernant le HIV, l'équipe d'Antoine Blancher a montré sur un large échantillon de patients infectés qu'il n'y avait aucun biais de fréquence ABO par rapport à une population contrôle (Puissant et coll. 2005) et il n'existe pas de données publiées pour le virus de la rougeole à notre connaissance.

Très récemment, il a été montré une association entre le phénotype O et la résistance à l'infection par le SRAS-CoV durant une épidémie hospitalière à Hong Kong (Y. Cheng et coll. 2005). Cette observation suggère que la présence d'anticorps anti-A et anti-B chez les individus O pourrait les avoir protégés contre des virions produits par les cellules d'individus de phénotype A ou B. Nous basant sur ces données épidémiologiques, nous avons utilisé un modèle expérimental pour valider l'hypothèse de leur origine. En outre, une collaboration avec des mathématiciens nous a permis de montrer l'impact que ce phénomène pourrait avoir sur la dynamique d'une épidémie de SRAS. Pour valider de manière définitive ces observations, nous devons tester si nos résultats peuvent être

corroborés par des essais *in vivo* mais en utilisant un autre modèle que celui impliquant l'homme et le SRAS-CoV.

Il est théoriquement possible que ce même phénomène soit mis en jeu pour d'autres virus et en particulier pour les autres CoV. Le CoV du porc pourrait alors représenter un candidat potentiel pour ces essais car il existe des porcs exprimant l'HBGA A au niveau de leurs épithéliums (A+) et d'autres qui en sont incapables (A-). Il serait donc possible de vérifier qu'un porc A- immunisé pour produire des anticorps anti-A est protégé vis-à-vis des CoV porcins issus d'un porc A+.

## Conclusion générale

Au travers des trois exemples étudiés au cours de ma thèse, nous avons pu relier les HBGA et leur polymorphisme à chacun des trois mécanismes de l'effet « Reine Rouge » (Figure 31 p.142) impliqués dans la diversification des glycannes de l'hôte en réponse à la pression de sélection exercée par les pathogènes et plus particulièrement par des virus.

Nous avons en effet montré que :

1. Les HBGA du lait peuvent servir de leurre aux NoV et qu'ils pourraient de ce fait affecter la coévolution hôte/pathogène.
2. Le pathogène (RHDV) peut reconnaître les glycannes de l'hôte (H type 2) qui sera alors contraint de les modifier pour pouvoir survivre (« non sécréteur-like »). Ici, à la différence du modèle classique « gène pour gène », nous avons ajouté une notion importante, à savoir qu'il s'agit de glycannes qui n'auraient aucune fonction biologique intrinsèque et dont l'individu peut parfaitement se passer. Cela nous a conduit à proposer le concept d'**immunité/protection innée de groupe** par le biais du polymorphisme des HBGA.
3. Les glycannes du pathogène (HBGA) peuvent être reconnus par l'hôte comme du « non soi » (NAb anti HBGA) et induire une protection. Le cas du SRAS-CoV représente à l'heure actuelle le meilleur exemple de cette situation pour la protection attribuée aux NAb anti-A et B.

Toutes ces données expérimentales illustrent clairement l'étroite coévolution virus/HBGA ABH. Sa conséquence principale pour l'hôte est l'apparition d'une immunité innée antivirale complexe et multifactorielle, dépendant exclusivement des HBGA ABH. Leur maintien et diversification au cours de l'évolution témoignent de son caractère vital à l'échelle de l'espèce et de sa mise en place très ancienne. En effet, si les agents pathogènes n'avaient pas eu d'impact sur ces gènes, ceux-ci auraient évolué par dérivation génétique uniquement. Or l'analyse du polymorphisme des gènes *FUT2* (Koda et coll.

2000) et *ABO* (Calafell et coll. 2008) a montré qu'ils ont été l'objet d'une sélection équilibrée (« balancing selection ») contribuant au maintien de leurs divers allèles, y compris ceux qui sont non fonctionnels comme les allèles *O* du locus *ABO* ou les allèles *se* du locus *FUT2*.

# BIBLIOGRAPHIE

## A

- Abe, Y., Takashita, E., Sugawara, K., Matsuzaki, Y., Muraki, Y., et Hongo, S. (2004). Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *J. Virol* 78, 9605-11.
- Abrantes, J., Esteves, P. J., et van der Loo, W. (2008). Evidence for recombination in the major capsid gene VP60 of the rabbit haemorrhagic disease virus (RHDV). *Arch Virol* 153, 329-35.
- Adamiak, B., Ekblad, M., Bergström, T., Ferro, V., et Trybala, E. (2007). Herpes simplex virus type 2 glycoprotein G is targeted by the sulfated oligo- and polysaccharide inhibitors of virus attachment to cells. *J. Virol* 81, 13424-34.
- Ah-Tye, C., Schwartz, S., Huberman, K., Carlin, E., et Moscona, A. (1999). Virus-receptor interactions of human parainfluenza viruses types 1, 2 and 3. *Microb. Pathog* 27, 329-36.
- Akerström, S., Mirazimi, A., et Tan, Y. (2007). Inhibition of SARS-CoV replication cycle by small interference RNAs silencing specific SARS proteins, 7a/7b, 3a/3b and S. *Antiviral Res* 73, 219-27.
- Alexander, D. A., et Dimock, K. (2002). Sialic acid functions in enterovirus 70 binding and infection. *J. Virol* 76, 11265-72.
- Allan, J. S., Coligan, J. E., Barin, F., McLane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H., et Essex, M. (1985). Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science* 228, 1091-4.
- Alvarez, C. P., Lasala, F., Carrillo, J., Muñiz, O., Corbí, A. L., et Delgado, R. (2002). C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *J. Virol* 76, 6841-4.
- Angata, T., et Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem. Rev* 102, 439-69.

- Apoil, P. A., Roubinet, F., Despiau, S., Mollicone, R., Oriol, R., et Blancher, A. (2000). Evolution of alpha 2-fucosyltransferase genes in primates: relation between an intronic Alu-Y element and red cell expression of ABH antigens. *Mol Biol Evol* 17, 337-51.
- Arnberg, N., Kidd, A. H., Edlund, K., Nilsson, J., Pring-Akerblom, P., et Wadell, G. (2002). Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* 302, 33-43.
- Arrighi, J., Pion, M., Garcia, E., Escola, J., van Kooyk, Y., Geijtenbeek, T. B., et Piguet, V. (2004). DC-SIGN-mediated infectious synapse formation enhances X4 HIV-1 transmission from dendritic cells to T cells. *J. Exp. Med* 200, 1279-88.
- Astronomo, R. D. et coll. (2008). A glycoconjugate antigen based on the recognition motif of a broadly neutralizing human immunodeficiency virus antibody, 2G12, is immunogenic but elicits antibodies unable to bind to the self glycans of gp120. *J. Virol* 82, 6359-68.

## B

- Balzarini, J. (2007). Carbohydrate-binding agents: a potential future cornerstone for the chemotherapy of enveloped viruses? *Antivir Chem Chemother* 18, 1-11.
- Barreaud, J. P., Saunier, K., Souchare, J., Delourme, D., Oulmouden, A., Oriol, R., Levéziel, H., Julien, R., et Petit, J. M. (2000). Three bovine alpha2-fucosyltransferase genes encode enzymes that preferentially transfer fucose on Galbeta1-3GalNAc acceptor substrates. *Glycobiology* 10, 611-21.
- Barth, H. et coll. (2003). Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J. Biol. Chem* 278, 41003-12.
- Bernstein, F. (1924). Ergebnisse einer biostatistischen zusammenfassenden Betrachtung über die erblichen Blutstructuren des Menschen. *Klin Wochenschr* 3, 1495-7.
- Bernstein, F. (1925). Zusammenfassende Betrachtungen über die erblichen Blutstrukturen des Menschen. *Molecular Genetics and Genomics* 37, 237-270.
- Bhende, Y. M., Deshpande, C. K., Bhatia, H. M., Sanger, R., Race, R. R., Morgan, W. T. J., et Watkins, W. M. (1952). A "new" blood group character related to the ABO system. *Lancet* 1, 903-4.
- Binley, J. M. et coll. (2004). Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J. Virol* 78, 13232-52.
- Bishop, J., et Gagneux, P. (2007). Evolution of carbohydrate antigens--microbial forces shaping host glycomes? *Glycobiology* 17, 23R-34R.
- Blackwell, C. C. (1989). The role of ABO blood groups and secretor status in host defences. *FEMS Microbiol Immunol* 1, 341-9.

- Blancher, A., et Socha, W. (1997). The ABO, Hh and Lewis blood group in humans and nonhuman primates. Molecular biology of evolution of blood group and MHC in primates. Berlin/Heidelberg/New York: Springer-Verlag.
- Blomqvist, S., Savolainen, C., Råman, L., Roivainen, M., et Hovi, T. (2002). Human rhinovirus 87 and enterovirus 68 represent a unique serotype with rhinovirus and enterovirus features. *J. Clin. Microbiol* 40, 4218-23.
- Bobardt, M. D. et coll. (2004). Contribution of proteoglycans to human immunodeficiency virus type 1 brain invasion. *J. Virol* 78, 6567-84.
- Bomsel, M., et David, V. (2002). Mucosal gatekeepers: selecting HIV viruses for early infection. *Nat. Med* 8, 114-6.
- Bourdel, L. (1960). *Groupes sanguins et tempéraments* (Paris: Maloine).
- van den Brink, E. N. et coll. (2005). Molecular and biological characterization of human monoclonal antibodies binding to the spike and nucleocapsid proteins of severe acute respiratory syndrome coronavirus. *J. Virol* 79, 1635-44.
- Brown, K. E., Anderson, S. M., et Young, N. S. (1993). Erythrocyte P antigen: cellular receptor for B19 parvovirus. *Science* 262, 114-7.
- Bu, W., Mamedova, A., Tan, M., Xia, M., Jiang, X., et Hegde, R. S. (2008). Structural basis for the receptor binding specificity of Norwalk virus. *J Virol* 82, 5340-7.
- Bureau, V., Marionneau, S., Cailleau-Thomas, A., Le Moullac-Vaidye, B., Liehr, T., et Le Pendu, J. (2001). Comparison of the three rat GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferases FTA, FTB and FTC. *Eur J Biochem* 268, 1006-19.

## C

- Cailleau-Thomas, A., Le Moullac-Vaidye, B., Rocher, J., Bouhours, D., Szpirer, C., et Le Pendu, J. (2002). Cloning of a rat gene encoding the histo-blood group A enzyme. Tissue expression of the gene and of the A and B antigens. *Eur J Biochem* 269, 4040-7.
- Calafell, F., Roubinet, F., Ramírez-Soriano, A., Saitou, N., Bertranpetit, J., et Blancher, A. (2008). Evolutionary dynamics of the human ABO gene. *Hum Genet* 124, 123-35.
- Calarese, D. A. et coll. (2003). Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 300, 2065-71.
- Cambi, A., et Figdor, C. G. (2003). Dual function of C-type lectin-like receptors in the immune system. *Curr. Opin. Cell Biol* 15, 539-46.
- Campbell, J. A., Davies, G. J., Bulone, V., et Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J* 326 ( Pt 3), 929-39.
- Cao, S. et coll. (2007). Structural basis for the recognition of blood group trisaccharides by norovirus. *J. Virol* 81, 5949-57.

- Cao, W. et coll. (1998). Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* 282, 2079-81.
- Capucci, L., Scicluna, M. T., et Lavazza, A. (1991). Diagnosis of viral haemorrhagic disease of rabbits and the European brown hare syndrome. *Rev. - Off. Int. Epizoot* 10, 347-70.
- Carne, L. R., et Watkins, W. M. (1977). Human blood group B gene-specified alpha-3-galactosyltransferase: purification of the enzyme in serum by biospecific adsorption onto blood group O erythrocyte membranes. *Biochem Biophys Res Commun* 77, 700-7.
- Chakravarty, S., Hutson, A. M., Estes, M. K., et Prasad, B. V. V. (2005). Evolutionary trace residues in noroviruses: importance in receptor binding, antigenicity, virion assembly, and strain diversity. *J. Virol* 79, 554-68.
- Chan, V. S. F. et coll. (2006). Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. *Nat. Genet* 38, 38-46.
- Chaturvedi, P., Warren, C. D., Altaye, M., Morrow, A. L., Ruiz-Palacios, G., Pickering, L. K., et Newburg, D. S. (2001). Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation. *Glycobiology* 11, 365-72.
- Chaturvedi, P., Warren, C. D., Buescher, C. R., Pickering, L. K., et Newburg, D. S. (2001). Survival of human milk oligosaccharides in the intestine of infants. *Adv. Exp. Med. Biol* 501, 315-23.
- Chen, D., Tseng, C., Wang, W., Wang, M., Tsao, K., Wu, T., et Sun, C. (2004). A novel cis-AB allele derived from the A transferase gene by nucleotide substitution C796A. *Ann Clin Lab Sci* 34, 437-42.
- Chen, Z. et coll. (2005). Recombinant modified vaccinia virus Ankara expressing the spike glycoprotein of severe acute respiratory syndrome coronavirus induces protective neutralizing antibodies primarily targeting the receptor binding region. *J. Virol* 79, 2678-88.
- Cheng, V. C. C., Lau, S. K. P., Woo, P. C. Y., et Yuen, K. Y. (2007). Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clin Microbiol Rev* 20, 660-94.
- Cheng, Y., Cheng, G., Chui, C. H., Lau, F. Y., Chan, P. K. S., Ng, M. H. L., Sung, J. J. Y., et Wong, R. S. M. (2005). ABO blood group and susceptibility to severe acute respiratory syndrome. *JAMA* 293, 1450-1.
- Cheshenko, N., et Herold, B. C. (2002). Glycoprotein B plays a predominant role in mediating herpes simplex virus type 2 attachment and is required for entry and cell-to-cell spread. *J. Gen. Virol* 83, 2247-55.
- Chiba, S., Sakuma, Y., Kogasaka, R., Akihara, M., Horino, K., Nakao, T., et Fukui, S. (1979). An outbreak of gastroenteritis associated with calicivirus in an infant home. *J. Med. Virol* 4, 249-54.
- Choi, J., Hutson, A. M., Estes, M. K., et Prasad, B. V. V. (2008). Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc Natl Acad Sci U S A* 105, 9175-80.

- Chung, C. S., Hsiao, J. C., Chang, Y. S., et Chang, W. (1998). A27L protein mediates vaccinia virus interaction with cell surface heparan sulfate. *J. Virol* 72, 1577-85.
- Clausen, H., Lavery, S. B., Nudelman, E., Tsuchiya, S., et Hakomori, S. (1985). Repetitive A epitope (type 3 chain A) defined by blood group A1-specific monoclonal antibody TH-1: chemical basis of qualitative A1 and A2 distinction. *Proc. Natl. Acad. Sci. U.S.A* 82, 1199-203.
- Clausen, H., Watanabe, K., Kannagi, R., Lavery, S. B., Nudelman, E., Arao-Tomono, Y., et Hakomori, S. (1984). Blood group A glycolipid (Ax) with globo-series structure which is specific for blood group A1 erythrocytes: one of the chemical bases for A1 and A2 distinction. *Biochem. Biophys. Res. Commun* 124, 523-9.
- Clausen, H., White, T., Takio, K., Titani, K., Stroud, M., Holmes, E., Karkov, J., Thim, L., et Hakomori, S. (1990). Isolation to homogeneity and partial characterization of a histo-blood group A defined Fuc alpha 1----2Gal alpha 1----3-N-acetylgalactosaminyltransferase from human lung tissue. *J Biol Chem* 265, 1139-45.
- Cohney, S., Mouhtouris, E., McKenzie, I. F., et Sandrin, M. S. (1996). Molecular cloning of the gene coding for pig alpha1-->2fucosyltransferase. *Immunogenetics* 44, 76-9.
- Coutinho, P. M., Deleury, E., Davies, G. J., et Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* 328, 307-17.
- Crow, J. F. (1993). Felix Bernstein and the first human marker locus. *Genetics* 133, 4-7.
- Cubitt, W. D., et Barrett, A. D. (1985). Propagation and preliminary characterization of a chicken candidate calicivirus. *J. Gen. Virol* 66 ( Pt 7), 1431-8.

## D

- D'Adamo, P., et Whitney, C. A. (1996). *Eat Right 4 Your Type: The Individualized Diet Solution to Staying Healthy, Living Longer & Achieving Your Ideal Weight* (GP Putnam's Sons).
- Dandekar, A. A., et Perlman, S. (2006). Severe Acute Respiratory Syndrome (SARS). *ENCYCLOPEDIA OF LIFE SCIENCES*.
- Davis, C. W., Nguyen, H., Hanna, S. L., Sánchez, M. D., Doms, R. W., et Pierson, T. C. (2006). West Nile virus discriminates between DC-SIGN and DC-SIGNR for cellular attachment and infection. *J. Virol* 80, 1290-301.
- dbRBC - Blood Group Antigen Gene Mutation Database Alleles of the ABO Blood Group System. Available at: [http://www.ncbi.nlm.nih.gov/projects/gv/mhc/xslcgi.cgi?cmd=bgmutsystems\\_alleles&system=abo](http://www.ncbi.nlm.nih.gov/projects/gv/mhc/xslcgi.cgi?cmd=bgmutsystems_alleles&system=abo).
- dbRBC - Blood Group Antigen Gene Mutation Database Alleles of the H/h Blood Group System. Available at: [http://www.ncbi.nlm.nih.gov/projects/gv/mhc/xslcgi.cgi?cmd=bgmutsystems\\_alleles&system=hh](http://www.ncbi.nlm.nih.gov/projects/gv/mhc/xslcgi.cgi?cmd=bgmutsystems_alleles&system=hh).
- Decastello, A., et Sturli, A. (1902). Ueber die Isoagglutinine im Serum gesunder und kranker Menschen. *Münch Med Wochenschr* 49, 1090-5.

- Dehecchi, M. C., Tamanini, A., Bonizzato, A., et Cabrini, G. (2000). Heparan sulfate glycosaminoglycans are involved in adenovirus type 5 and 2-host cell interactions. *Virology* 268, 382-90.
- Di Martino, B., Marsilio, F., et Roy, P. (2007). Assembly of feline calicivirus-like particle and its immunogenicity. *Vet. Microbiol* 120, 173-8.
- Domino, S. E., Hiraiwa, N., et Lowe, J. B. (1997). Molecular cloning, chromosomal assignment and tissue-specific expression of a murine alpha(1,2)fucosyltransferase expressed in thymic and epididymal epithelial cells. *Biochem J* 327 ( Pt 1), 105-15.
- Domino, S. E., Zhang, L., et Lowe, J. B. (2001). Molecular cloning, genomic mapping, and expression of two secretor blood group alpha (1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract. *J Biol Chem* 276, 23748-56.
- von Dungern, E., et Hirschfeld, L. (1910). Über Vererbung gruppenspezifischer Strukturen des Blutes. *Z. Immunitätsforsch Exp. Ther.* 6, 284-92.
- von Dungern, E., et Hirschfeld, L. (1911). Über Vererbung gruppenspezifischer Strukturen des Blutes. *Mol. Genet. Genomics* 5, 196-7.
- Dürrbach, A., Baple, E., Preece, A. F., Charpentier, B., et Gustafsson, K. (2007). Virus recognition by specific natural antibodies and complement results in MHC I cross-presentation. *Eur J Immunol* 37, 1254-65.

## E

## F

- Fankhauser, R. L., Noel, J. S., Monroe, S. S., Ando, T., et Glass, R. I. (1998). Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. *J. Infect. Dis* 178, 1571-8.
- Farkas, T., Jiang, X., Guerrero, M. L., Zhong, W., Wilton, N., Berke, T., Matson, D. O., Pickering, L. K., et Ruiz-Palacios, G. (2000). Prevalence and genetic diversity of human caliciviruses (HuCVs) in Mexican children. *J. Med. Virol* 62, 217-23.
- Fastier, L. B. (1957). A new feline virus isolated in tissue culture. *Am. J. Vet. Res* 18, 382-9.
- Feldman, S. A., Hendry, R. M., et Beeler, J. A. (1999). Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J. Virol* 73, 6610-7.
- Fernandez-Mateos, P. et coll. (1998). Point mutations and deletion responsible for the Bombay H null and the Reunion H weak blood groups. *Vox Sang* 75, 37-46.

- Feyzi, E., Trybala, E., Bergström, T., Lindahl, U., et Spillmann, D. (1997). Structural requirement of heparan sulfate for interaction with herpes simplex virus type 1 virions and isolated glycoprotein C. *J. Biol. Chem* 272, 24850-7.
- Fischer, C., Jock, B., et Vogel, F. (1998). Interplay between humans and infective agents: a population genetic study. *Hum Genet* 102, 415-22.
- Forrester, N. L., Moss, S. R., Turner, S. L., Schirrmeyer, H., et Gould, E. A. (2008). Recombination in rabbit haemorrhagic disease virus: Possible impact on evolution and epidemiology. *Virology* 376, 390-6.
- Fouchier, R. A. M., Munster, V., Wallensten, A., Bestebroer, T. M., Herfst, S., Smith, D., Rimmelzwaan, G. F., Olsen, B., et Osterhaus, A. D. M. E. (2005). Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. *J. Virol* 79, 2814-22.
- Francki, R. I. B., Fauquet, C. M., Knudson, D. L., et Brown, F. (1991). Classification and nomenclature of viruses : fifth report of the International Committee on Taxonomy of Viruses. *Archives of virology. Supplementum vol. suppl. 2*, 450.
- Fukuda, M. N., et Levery, S. B. (1983). Glycolipids of fetal, newborn, and adult erythrocytes: glycolipid pattern and structural study of H3-glycolipid from newborn erythrocytes. *Biochemistry* 22, 5034-40.

## G

- Gagneux, P., et Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9, 747-55.
- Gardai, S. J., Xiao, Y., Dickinson, M., Nick, J. A., Voelker, D. R., Greene, K. E., et Henson, P. M. (2003). By binding SIRPalpha or calreticulin/CD91, lung collectins act as dual function surveillance molecules to suppress or enhance inflammation. *Cell* 115, 13-23.
- Gavier-Widén, D., et Mörner, T. (1991). Epidemiology and diagnosis of the European brown hare syndrome in Scandinavian countries: a review. *Rev. - Off. Int. Epizoot* 10, 453-8.
- Geijtenbeek, T. B. et coll. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100, 587-97.
- Gentsch, J. R., et Pacitti, A. F. (1987). Differential interaction of reovirus type 3 with sialylated receptor components on animal cells. *Virology* 161, 245-8.
- Giroglou, T., Florin, L., Schäfer, F., Streeck, R. E., et Sapp, M. (2001). Human papillomavirus infection requires cell surface heparan sulfate. *J. Virol* 75, 1565-70.
- Graham, R. L., Sims, A. C., Brockway, S. M., Baric, R. S., et Denison, M. R. (2005). The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J. Virol* 79, 13399-411.
- Gramberg, T. et coll. (2005). LSEctin interacts with filovirus glycoproteins and the spike protein of SARS coronavirus. *Virology* 340, 224-36.

- Green, C. (1989). The ABO, Lewis and related blood group antigens; a review of structure and biosynthesis. *FEMS Microbiol Immunol* 1, 321-30.
- Green, K. Y. et coll. (2000). Taxonomy of the caliciviruses. *J. Infect. Dis* 181 Suppl 2, S322-30.
- Guan, Y. et coll. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302, 276-8.
- Guo, M., Qian, Y., Chang, K. O., et Saif, L. J. (2001). Expression and self-assembly in baculovirus of porcine enteric calicivirus capsids into virus-like particles and their use in an enzyme-linked immunosorbent assay for antibody detection in swine. *J. Clin. Microbiol* 39, 1487-93.
- Gustafsson, K., Dürrbach, A., Seymour, R. M., et Pomiankowski, A. (2005). The Role of ABO Histo-Blood Group Antigens in Viral Infections. *Trends Glycosci. Glycotechnol.* 17, 285-94.

## H

- Hammer, C., et Hingerle, M. (1992). Development of preformed natural antibodies in gnotobiotic dogs and pigs, impact of food antigens on antibody specificity. *Transplant. Proc* 24, 707-9.
- Hamming, I., Timens, W., Bulthuis, M. L. C., Lely, A. T., Navis, G. J., et van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol* 203, 631-7.
- Hansman, G. S., Matsubara, N., Oka, T., Ogawa, S., Natori, K., Takeda, N., et Katayama, K. (2005). Deletion analysis of the sapovirus VP1 gene for the assembly of virus-like particles. *Arch Virol* 150, 2529-38.
- Hansman, G. S., Oka, T., Katayama, K., et Takeda, N. (2007). Human sapoviruses: genetic diversity, recombination, and classification. *Rev. Med. Virol* 17, 133-41.
- Harduin-Lepers, A., Mollicone, R., Delannoy, P., et Oriol, R. (2005). The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology* 15, 805-17.
- Harrington, P. R., Lindesmith, L., Yount, B., Moe, C. L., et Baric, R. S. (2002). Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J. Virol* 76, 12335-43.
- Hart, G. W., Haltiwanger, R. S., Holt, G. D., et Kelly, W. G. (1989). Glycosylation in the nucleus and cytoplasm. *Annu Rev Biochem* 58, 841-74.
- Hartmann, G. (1941). Group antigens in human organs: a discussion of the "secreter" non-secreter" phenomenon (Ed. Einar Munksgaard).
- Hartshorn, K. L., Crouch, E. C., White, M. R., Eggleton, P., Tauber, A. I., Chang, D., et Sastry, K. (1994). Evidence for a protective role of pulmonary surfactant protein D (SP-D) against influenza A viruses. *J. Clin. Invest* 94, 311-9.

- Hartshorn, K. L., Reid, K. B., White, M. R., Jensenius, J. C., Morris, S. M., Tauber, A. I., et Crouch, E. (1996). Neutrophil deactivation by influenza A viruses: mechanisms of protection after viral opsonization with collectins and hemagglutination-inhibiting antibodies. *Blood* 87, 3450-61.
- Hawgood, S., Brown, C., Edmondson, J., Stumbaugh, A., Allen, L., Goerke, J., Clark, H., et Poulain, F. (2004). Pulmonary collectins modulate strain-specific influenza A virus infection and host responses. *J. Virol* 78, 8565-72.
- He, R. et coll. (2004). Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus. *Virus Res* 105, 121-5.
- He, Y., Lu, H., Siddiqui, P., Zhou, Y., et Jiang, S. (2005). Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. *J. Immunol* 174, 4908-15.
- He, Y., Zhou, Y., Wu, H., Luo, B., Chen, J., Li, W., et Jiang, S. (2004). Identification of immunodominant sites on the spike protein of severe acute respiratory syndrome (SARS) coronavirus: implication for developing SARS diagnostics and vaccines. *J. Immunol* 173, 4050-7.
- Herold, B. C., Gerber, S. I., Belval, B. J., Siston, A. M., et Shulman, N. (1996). Differences in the susceptibility of herpes simplex virus types 1 and 2 to modified heparin compounds suggest serotype differences in viral entry. *J. Virol* 70, 3461-9.
- Herold, B. C., Gerber, S. I., Polonsky, T., Belval, B. J., Shaklee, P. N., et Holme, K. (1995). Identification of structural features of heparin required for inhibition of herpes simplex virus type 1 binding. *Virology* 206, 1108-16.
- Herold, B. C., WuDunn, D., Soltys, N., et Spear, P. G. (1991). Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity. *J. Virol* 65, 1090-8.
- Hirabayashi, J., Arata, Y., et Kasai, K. (2001). Glycome project: concept, strategy and preliminary application to *Caenorhabditis elegans*. *Proteomics* 1, 295-303.
- Hitoshi, S., Kusunoki, S., Kanazawa, I., et Tsuji, S. (1996). Molecular cloning and expression of a third type of rabbit GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase. *J Biol Chem* 271, 16975-81.
- Hitoshi, S., Kusunoki, S., Kanazawa, I., et Tsuji, S. (1995). Molecular cloning and expression of two types of rabbit beta-galactoside alpha 1,2-fucosyltransferase. *J Biol Chem* 270, 8844-50.
- Ho, Y., Lin, P., Liu, C. Y. Y., Lee, S., et Chao, Y. (2004). Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochem. Biophys. Res. Commun* 318, 833-8.
- Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., et Pöhlmann, S. (2005). Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. U.S.A* 102, 7988-93.

- Holgerson, J., Breimer, M. E., et Samuelsson, B. E. (1992). Basic biochemistry of cell surface carbohydrates and aspects of the tissue distribution of histo-blood group ABH and related glycosphingolipids. *APMIS Suppl* 27, 18-27.
- Holmes, K. V. (2005). Structural biology. Adaptation of SARS coronavirus to humans. *Science* 309, 1822-3.
- Hong, P. W., Nguyen, S., Young, S., Su, S. V., et Lee, B. (2007). Identification of the optimal DC-SIGN binding site on human immunodeficiency virus type 1 gp120. *J. Virol* 81, 8325-36.
- Hounsell, E. F., Davies, M. J., et Renouf, D. V. (1996). O-linked protein glycosylation structure and function. *Glycoconj J* 13, 19-26.
- Hsiao, J. C., Chung, C. S., et Chang, W. (1999). Vaccinia virus envelope D8L protein binds to cell surface chondroitin sulfate and mediates the adsorption of intracellular mature virions to cells. *J. Virol* 73, 8750-61.
- Hsieh, P. et coll. (2005). Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J. Virol* 79, 13848-55.
- Huang, C., Ito, N., Tseng, C. K., et Makino, S. (2006). Severe acute respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. *J. Virol* 80, 7287-94.
- Huang, P. et coll. (2003). Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. *J Infect Dis* 188, 19-31.
- Huang, P., Farkas, T., Zhong, W., Tan, M., Thornton, S., Morrow, A. L., et Jiang, X. (2005). Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *J Virol* 79, 6714-22.
- Huskens, D., Van Laethem, K., Vermeire, K., Balzarini, J., et Schols, D. (2007). Resistance of HIV-1 to the broadly HIV-1-neutralizing, anti-carbohydrate antibody 2G12. *Virology* 360, 294-304.
- Hutson, A. M., Airaud, F., LePendu, J., Estes, M. K., et Atmar, R. L. (2005). Norwalk virus infection associates with secretor status genotyped from sera. *J Med Virol* 77, 116-20.
- Hutson, A. M., Atmar, R. L., Graham, D. Y., et Estes, M. K. (2002). Norwalk virus infection and disease is associated with ABO histo-blood group type. *J Infect Dis* 185, 1335-7.
- Hutson, A. M., Atmar, R. L., Marcus, D. M., et Estes, M. K. (2003). Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. *J. Virol* 77, 405-15.
- Hutson, A. M., Atmar, R. L., et Estes, M. K. (2004). Norovirus disease: changing epidemiology and host susceptibility factors. *Trends in Microbiology* 12, 279-87.

- Ikeda, K., Sannoh, T., Kawasaki, N., Kawasaki, T., et Yamashina, I. (1987). Serum lectin with known structure activates complement through the classical pathway. *J. Biol. Chem* 262, 7451-4.
- Ip, W. K. E. et coll. (2005). Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. *J. Infect. Dis* 191, 1697-704.
- Ito, N., Mossel, E. C., Narayanan, K., Popov, V. L., Huang, C., Inoue, T., Peters, C. J., et Makino, S. (2005). Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein. *J. Virol* 79, 3182-6.
- Iwamoto, S. et coll. (2002). Rat encodes the paralogous gene equivalent of the human histo-blood group ABO gene. Association with antigen expression by overexpression of human ABO transferase. *J. Biol. Chem* 277, 46463-9.

## J

- Janský, J. (1907). Haematologické studie u psychotiku. *Sborn Klin Praha* 1907, 85-139.
- Javaud, C., Dupuy, F., Maftah, A., Julien, R., et Petit, J. (2003). The fucosyltransferase gene family: an amazing summary of the underlying mechanisms of gene evolution. *Genetica* 118, 157-70.
- Jeffers, S. A. et coll. (2004). CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc. Natl. Acad. Sci. U.S.A* 101, 15748-53.
- Jiang, X., Huang, P., Zhong, W., Tan, M., Farkas, T., Morrow, A. L., Newburg, D. S., Ruiz-Palacios, G. M., et Pickering, L. K. (2004). Human milk contains elements that block binding of noroviruses to human histo-blood group antigens in saliva. *J. Infect. Dis* 190, 1850-9.
- Jiang, X., Wang, M., Graham, D. Y., et Estes, M. K. (1992). Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *J. Virol* 66, 6527-32.
- Jiang, X., Zhong, W., Kaplan, M., Pickering, L. K., et Matson, D. O. (1999). Expression and characterization of Sapporo-like human calicivirus capsid proteins in baculovirus. *J. Virol. Methods* 78, 81-91.
- Joziase, D. H. (1992). Mammalian glycosyltransferases: genomic organization and protein structure. *Glycobiology* 2, 271-7.
- Joziase, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. J., et Shaper, N. L. (1989). Bovine alpha 1-3-galactosyltransferase: isolation and characterization of a cDNA clone. Identification of homologous sequences in human genomic DNA. *J Biol Chem* 264, 14290-7.
- Julenius, K., Mølgaard, A., Gupta, R., et Brunak, S. (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15, 153-64.

## K

- Kadoi, K., Kiryu, M., Iwabuchi, M., Kamata, H., Yukawa, M., et Inaba, Y. (1997). A strain of calicivirus isolated from lions with vesicular lesions on tongue and snout. *New Microbiol* 20, 141-8.
- Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J., et Chiorini, J. A. (2001). Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J. Virol* 75, 6884-93.
- Kamitani, W., Narayanan, K., Huang, C., Lokugamage, K., Ikegami, T., Ito, N., Kubo, H., et Makino, S. (2006). Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc. Natl. Acad. Sci. U.S.A* 103, 12885-90.
- Kapikian, A. Z., Wyatt, R. G., Dolin, R., Thornhill, T. S., Kalica, A. R., et Chanock, R. M. (1972). Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *J. Virol* 10, 1075-81.
- Kari, B., et Gehrz, R. (1993). Structure, composition and heparin binding properties of a human cytomegalovirus glycoprotein complex designated gC-II. *J. Gen. Virol* 74 ( Pt 2), 255-64.
- Kasturi, L., Chen, H., et Shakin-Eshleman, S. H. (1997). Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors. *Biochem J* 323 ( Pt 2), 415-9.
- Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G., et Lowe, J. B. (1995). Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J Biol Chem* 270, 4640-9.
- Kermarrec, N., Roubinet, F., Apoil, P. A., et Blancher, A. (1999). Comparison of allele O sequences of the human and non-human primate ABO system. *Immunogenetics* 49, 517-26.
- King, M. (1994). Blood group antigens on human erythrocytes-distribution, structure and possible functions. *Biochim Biophys Acta* 1197, 15-44.
- Kishore, U. et coll. (2006). Surfactant proteins SP-A and SP-D: structure, function and receptors. *Mol. Immunol* 43, 1293-315.
- Klenk, H., Wagner, R., Heuer, D., et Wolff, T. (2002). Importance of hemagglutinin glycosylation for the biological functions of influenza virus. *Virus Res* 82, 73-5.
- Koda, Y., Soejima, M., Johnson, P. H., Smart, E., et Kimura, H. (2000). An Alu-mediated large deletion of the FUT2 gene in individuals with the ABO-Bombay phenotype. *Hum Genet* 106, 80-5.
- Koda, Y., Soejima, M., Johnson, P. H., Smart, E., et Kimura, H. (1997). Missense mutation of FUT1 and deletion of FUT2 are responsible for Indian Bombay phenotype of ABO blood group system. *Biochem Biophys Res Commun* 238, 21-5.
- Koda, Y., Soejima, M., et Kimura, H. (1997). Structure and expression of H-type GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase gene (FUT1). Two

- transcription start sites and alternative splicing generate several forms of FUT1 mRNA. *J Biol Chem* 272, 7501-5.
- Koda, Y., Soejima, M., et Kimura, H. (2001). The polymorphisms of fucosyltransferases. *Leg Med (Tokyo)* 3, 2-14.
- Koda, Y., Tachida, H., Soejima, M., Takenaka, O., et Kimura, H. (2000). Ancient origin of the null allele se(428) of the human ABO-secretor locus (FUT2). *J Mol Evol* 50, 243-8.
- Kominato, Y., McNeill, P. D., Yamamoto, M., Russell, M., Hakomori, S., et Yamamoto, F. (1992). Animal histo-blood group ABO genes. *Biochem Biophys Res Commun* 189, 154-64.
- Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., et Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. *Br. Med. Bull* 58, 19-42.
- Kritschewski, I. L., et Schwarzmann, L. A. (1927). Die Gruppenspezifische Differenzierung der Menschlichen Organe. *Klin Wochenschr* 6, 2081-6.
- Krokhin, O. et coll. (2003). Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. *Mol Cell Proteomics* 2, 346-56.
- Krusat, T., et Streckert, H. J. (1997). Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch. Virol* 142, 1247-54.
- Kuba, K. et coll. (2005). A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat. Med* 11, 875-9.
- Kumazaki, T., et Yoshida, A. (1984). Biochemical evidence that secretor gene, Se, is a structural gene encoding a specific fucosyltransferase. *Proc Natl Acad Sci U S A* 81, 4193-7.
- Künkel, F., et Herrler, G. (1993). Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* 195, 195-202.
- Kwong, P. D. et coll. (2002). HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 420, 678-82.

## L

- Lai, M. M., et Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv. Virus Res* 48, 1-100.
- Land, A., et Braakman, I. (2001). Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. *Biochimie* 83, 783-90.
- Landsteiner, K. (1901). Über Agglutinationserscheinungen normalen menschlichen Blutes. *Wien Klin Wochenschr* 14, 1132-4.
- Larsen, R. D., Ernst, L. K., Nair, R. P., et Lowe, J. B. (1990). Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-

- L-fucosyltransferase cDNA that can form the H blood group antigen. *Proc Natl Acad Sci U S A* 87, 6674-8.
- Larsen, R. D., Rivera-Marrero, C. A., Ernst, L. K., Cummings, R. D., et Lowe, J. B. (1990). Frameshift and nonsense mutations in a human genomic sequence homologous to a murine UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-galactosyltransferase cDNA. *J. Biol. Chem* 265, 7055-61.
- Lau, S. K. P. et coll. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc Natl Acad Sci U S A* 102, 14040-5.
- Le Gall-Reculé, G., Zwingelstein, F., Laurent, S., de Boisséson, C., Portejoie, Y., et Rasschaert, D. (2003). Phylogenetic analysis of rabbit haemorrhagic disease virus in France between 1993 and 2000, and the characterisation of RHDV antigenic variants. *Arch Virol* 148, 65-81.
- Le Pendu, J. (1989). A hypothesis on the dual significance of ABH, Lewis and related antigens. *J Immunogenet* 16, 53-61.
- Le Pendu, J. (2004). Histo-blood group antigen and human milk oligosaccharides: genetic polymorphism and risk of infectious diseases. *Adv. Exp. Med. Biol* 554, 135-43.
- Le Pendu, J., Cartron, J. P., Lemieux, R. U., et Oriol, R. (1985). The presence of at least two different H-blood-group-related beta-D-gal alpha-2-L-fucosyltransferases in human serum and the genetics of blood group H substances. *Am J Hum Genet* 37, 749-60.
- Le Pendu, J. et coll. (1986). Monoclonal antibodies specific for type 3 and type 4 chain-based blood group determinants: Relationship to the A1 and A2 subgroups. *Glycoconjugate J* 3, 255-71.
- Le Pendu, J., Ruvoën-Clouet, N., Kindberg, E., et Svensson, L. (2006). Mendelian resistance to human norovirus infections. *Semin Immunol* 18, 375-86.
- Lee, B. et coll. (2001). cis Expression of DC-SIGN allows for more efficient entry of human and simian immunodeficiency viruses via CD4 and a coreceptor. *J. Virol* 75, 12028-38.
- Lee, E., Hall, R. A., et Lobigs, M. (2004). Common E protein determinants for attenuation of glycosaminoglycan-binding variants of Japanese encephalitis and West Nile viruses. *J. Virol* 78, 8271-80.
- Lenghaus, C., Studdert, M., et Gavier-Widén, D. (2001). Calicivirus infections. Dans *Infectious Diseases of Wild Mammals* (Elizabeth S. Williams, Ian K. Barker).
- LeVine, A. M., Whitsett, J. A., Hartshorn, K. L., Crouch, E. C., et Korfhagen, T. R. (2001). Surfactant protein D enhances clearance of influenza A virus from the lung in vivo. *J. Immunol* 167, 5868-73.
- Li, F., Li, W., Farzan, M., et Harrison, S. C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309, 1864-8.
- Li, W., Greenough, T. C., Moore, M. J., Vasilieva, N., Somasundaran, M., Sullivan, J. L., Farzan, M., et Choe, H. (2004). Efficient replication of severe acute respiratory syndrome coronavirus in mouse cells is limited by murine angiotensin-converting enzyme 2. *J. Virol* 78, 11429-33.

- Li, W. et coll. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450-4.
- Li, W. et coll. (2005). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J* 24, 1634-43.
- Li, Y., Teneberg, S., Thapa, P., Bendelac, A., Levery, S. B., et Zhou, D. (2008). Sensitive detection of isoglobo and globo series tetraglycosylceramides in human thymus by ion trap mass spectrometry. *Glycobiology* 18, 158-65.
- Lin, B., Saito, M., Sakakibara, Y., Hayashi, Y., Yanagisawa, M., et Iwamori, M. (2001). Characterization of three members of murine alpha1,2-fucosyltransferases: change in the expression of the Se gene in the intestine of mice after administration of microbes. *Arch Biochem Biophys* 388, 207-15.
- Lin, M. et coll. (2003). Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *BMC Med. Genet* 4, 9.
- Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., Stewart, P., LePendu, J., et Baric, R. (2003). Human susceptibility and resistance to Norwalk virus infection. *Nat Med* 9, 548-53.
- Liu, C. K., Wei, G., et Atwood, W. J. (1998). Infection of glial cells by the human polyomavirus JC is mediated by an N-linked glycoprotein containing terminal alpha(2-6)-linked sialic acids. *J. Virol* 72, 4643-9.
- Liu, S. J., Xue, H. P., Pu, B. Q., et Quian, N. H. (1984). A new viral disease in rabbits. *Anim Husb Vet Med* 16, 253-5.
- Lively, C., et Apanius, V. (1995). Genetic diversity in host-parasite interactions. *Ecology of Infectious Diseases in Natural Populations*, 421--49.
- Lozach, P. et coll. (2003). DC-SIGN and L-SIGN are high affinity binding receptors for hepatitis C virus glycoprotein E2. *J. Biol. Chem* 278, 20358-66.
- Lynch, M., et Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151-5.

## M

- Marcato, P. S., Benazzi, C., Vecchi, G., Galeotti, M., Della Salda, L., Sarli, G., et Lucidi, P. (1991). Clinical and pathological features of viral haemorrhagic disease of rabbits and the European brown hare syndrome. *Rev. - Off. Int. Epizoot* 10, 371-92.
- Marchandeu, S., Chaval, Y., et Le Goff, E. (2000). Prolonged decline in the abundance of wild European rabbits *Oryctolagus cuniculus* and high immunity level over three years following the arrival of rabbit haemorrhagic disease. *Wildl Biol* 6, 141-7.

- Marionneau, S., Ruvoën, N., Le Moullac-Vaidye, B., Clement, M., Cailleau-Thomas, A., Ruiz-Palacois, G., Huang, P., Jiang, X., et Le Pendu, J. (2002). Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology* 122, 1967-77.
- Marra, M. A. et coll. (2003). The Genome sequence of the SARS-associated coronavirus. *Science* 300, 1399-404.
- Martín, J., Wharton, S. A., Lin, Y. P., Takemoto, D. K., Skehel, J. J., Wiley, D. C., et Steinhauer, D. A. (1998). Studies of the binding properties of influenza hemagglutinin receptor-site mutants. *Virology* 241, 101-11.
- Martinko, J. M., Vincek, V., Klein, D., et Klein, J. (1993). Primate ABO glycosyltransferases: evidence for trans-species evolution. *Immunogenetics* 37, 274-8.
- Martinsohn, J. T., Sousa, A. B., Guethlein, L. A., et Howard, J. C. (1999). The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics* 50, 168-200.
- Matiz, K., Ursu, K., Kecskeméti, S., Bajmócy, E., et Kiss, I. (2006). Phylogenetic analysis of rabbit haemorrhagic disease virus (RHDV) strains isolated between 1988 and 2003 in eastern Hungary. *Arch Virol* 151, 1659-66.
- Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A., et Klenk, H. (2004). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc. Natl. Acad. Sci. U.S.A* 101, 4620-4.
- Meijerink, E., Fries, R., Vögeli, P., Masabanda, J., Wigger, G., Stricker, C., Neuenschwander, S., Bertschinger, H. U., et Stranzinger, G. (1997). Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci. *Mamm Genome* 8, 736-41.
- Meijerink, E., Neuenschwander, S., Dinter, A., Yerle, M., Stranzinger, G., et Vögeli, P. (2001). Isolation of a porcine UDP-GalNAc transferase cDNA mapping to the region of the blood group EAA locus on pig chromosome 1. *Anim Genet* 32, 132-8.
- Meunier, J. C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J., et Wychowski, C. (1999). Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex. *J. Gen. Virol* 80 ( Pt 4), 887-96.
- Mifsud, N. A., Watt, J. M., Condon, J. A., Haddad, A. P., et Sparrow, R. L. (2000). A novel cis-AB variant allele arising from a nucleotide substitution A796C in the B transferase gene. *Transfusion* 40, 1276-7.
- Mollicone, R., Davies, D. R., Evans, B., Dalix, A. M., et Oriol, R. (1986). Cellular expression and genetic control of ABH antigens in primary sensory neurons of marmoset, baboon and man. *J Neuroimmunol* 10, 255-69.
- Morisse, J. P., Le Gall, G., et Boilletot, E. (1991). Hepatitis of viral origin in Leporidae: introduction and aetiological hypotheses. *Rev Sci Tech* 10, 269-310.
- Morrow, A. L., et Rangel, J. M. (2004). Human milk protection against infectious diarrhea: implications for prevention and clinical care. *Seminars in Pediatric Infectious Diseases* 15, 221-8.

- Morrow, A. L., Ruiz-Palacios, G. M., et coll. (2004). Human milk oligosaccharide blood group epitopes and innate immune protection against campylobacter and calicivirus diarrhea in breastfed infants. *Adv. Exp. Med. Biol* 554, 443-6.
- Mourant, A. E. (1982). ABH secretion and natural selection. *Ann. Hum. Biol* 9, 575-7.
- Mourant, A. E., et Ernest, A. (1983). *Blood Relations: Blood Groups and Anthropology* (Oxford University Press New York, N. Y.).
- Mourant, A. E., Kopec, A. C., et Domanieewska-Sobczak, K. (1978). Blood groups and diseases, a study of associations of diseases with blood groups and other polymorphisms. (Oxford, NY: Oxford University press).
- Mourant, A. E., Kopec, A. C., et Domaniewska-Sobczak, K. (1976). *The distribution of the human blood groups, and other polymorphisms* (Oxford University Press London).

## N

- Nagai, M., Davè, V., Kaplan, B. E., et Yoshida, A. (1978). Human blood group glycosyltransferases. I. Purification of n-acetylgalactosaminyltransferase. *J Biol Chem* 253, 377-9.
- Nagesha, H. S., Wang, L. F., Hyatt, A. D., Morrissy, C. J., Lenghaus, C., et Westbury, H. A. (1995). Self-assembly, antigenicity, and immunogenicity of the rabbit haemorrhagic disease virus (Czechoslovakian strain V-351) capsid protein expressed in baculovirus. *Arch. Virol* 140, 1095-1108.
- Navaratnam, N., Findlay, J. B., Keen, J. N., et Watkins, W. M. (1990). Purification, properties and partial amino acid sequence of the blood-group-A-gene-associated alpha-3-N-acetylgalactosaminyltransferase from human gut mucosal tissue. *Biochem J* 271, 93-8.
- Nei, M., Gu, X., et Sitnikova, T. (1997). Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. U.S.A* 94, 7799-806.
- Nei, M., et Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet* 39, 121-52.
- Neil, S. J. D., McKnight, A., Gustafsson, K., et Weiss, R. A. (2005). HIV-1 incorporates ABO histo-blood group antigens that sensitize virions to complement-mediated inactivation. *Blood* 105, 4693-9.
- Neill, J. D., Meyer, R. F., et Seal, B. S. (1995). Genetic relatedness of the caliciviruses: San Miguel sea lion and vesicular exanthema of swine viruses constitute a single genotype within the Caliciviridae. *J. Virol* 69, 4484-8.
- Ng, M. H. L., Lau, K., Li, L., Cheng, S., Chan, W. Y., Hui, P. K., Zee, B., Leung, C., et Sung, J. J. Y. (2004). Association of human-leukocyte-antigen class I (B\*0703) and class II (DRB1\*0301) genotypes with susceptibility and resistance to the development of severe acute respiratory syndrome. *J. Infect. Dis* 190, 515-8.

## O

- Ogasawara, K. et coll. (1996). Extensive polymorphism of ABO blood group gene: three major lineages of the alleles for the common ABO phenotypes. *Hum Genet* 97, 777-83.
- Ogasawara, K., Yabe, R., Uchikawa, M., Bannai, M., Nakata, K., Takenaka, M., Takahashi, Y., Juji, T., et Tokunaga, K. (1998). Different alleles cause an imbalance in A2 and A2B phenotypes of the ABO blood group. *Vox Sang* 74, 242-7.
- Ogasawara, K., Yabe, R., Uchikawa, M., Nakata, K., Watanabe, J., Takahashi, Y., et Tokunaga, K. (2001). Recombination and gene conversion-like events may contribute to ABO gene diversity causing various phenotypes. *Immunogenetics* 53, 190-9.
- Ogasawara, K. et coll. (1996). Molecular genetic analysis of variant phenotypes of the ABO blood group system. *Blood* 88, 2732-7.
- O'hUigin, C., Sato, A., et Klein, J. (1997). Evidence for convergent evolution of A and B blood group antigens in primates. *Hum. Genet* 101, 141-8.
- Oliver, S. L., Asobayire, E., Dastjerdi, A. M., et Bridger, J. C. (2006). Genomic characterization of the unclassified bovine enteric virus Newbury agent-1 (Newbury1) endorses a new genus in the family Caliciviridae. *Virology* 350, 240-50.
- Olofsson, S., et Bergström, T. (2005). Glycoconjugate glycans as viral receptors. *Ann. Med* 37, 154-72.
- Opie, S. R., Warrington, K. H., Agbandje-McKenna, M., Zolotukhin, S., et Muzyczka, N. (2003). Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J. Virol* 77, 6995-7006.
- Oriol, R. (1987). ABH and related tissue antigens. *Biochem Soc Trans* 15, 596-9.
- Oriol, R. (1990). Genetic control of the fucosylation of ABH precursor chains. Evidence for new epistatic interactions in different cells and tissues. *J Immunogenet* 17, 235-45.
- Oriol, R., Candelier, J. J., Taniguchi, S., Balanzino, L., Peters, L., Niekrasz, M., Hammer, C., et Cooper, D. K. (1999). Major carbohydrate epitopes in tissues of domestic and African wild animals of potential interest for xenotransplantation research. *Xenotransplantation* 6, 79-89.
- Oriol, R., et Dalix, A. M. (1977). Differences in the maturation of the immune response of A- and A+ rabbits. Good and poor responders respectively for the A antigen. *Immunology* 33, 91-9.
- Oriol, R., Danilovs, J., et Hawkins, B. R. (1981). A new genetic model proposing that the Se gene is a structural gene closely linked to the H gene. *Am J Hum Genet* 33, 421-31.

- Oriol, R., Le Pendu, J., et Mollicone, R. (1986). Genetics of ABO, H, Lewis, X and related antigens. *Vox Sang* 51, 161-71.
- Oriol, R., Martinez-Duncker, I., Chantret, I., Mollicone, R., et Codogno, P. (2002). Common origin and evolution of glycosyltransferases using Dol-P-monosaccharides as donor substrate. *Mol Biol Evol* 19, 1451-63.
- Oriol, R., Mollicone, R., Coullin, P., Dalix, A. M., et Candelier, J. J. (1992). Genetic regulation of the expression of ABH and Lewis antigens in tissues. *APMIS Suppl* 27, 28-38.

## P

- Pacitti, A. F., et Gentsch, J. R. (1987). Inhibition of reovirus type 3 binding to host cells by sialylated glycoproteins is mediated through the viral attachment protein. *J. Virol* 61, 1407-15.
- Patel, M. M., Widdowson, M., Glass, R. I., Akazawa, K., Vinjé, J., et Parashar, U. D. (2008). Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg Infect Dis* 14, 1224-31.
- Paul, R. W., Choi, A. H., et Lee, P. W. (1989). The alpha-anomeric form of sialic acid is the minimal receptor determinant recognized by reovirus. *Virology* 172, 382-5.
- Peiris, J. S. M., Guan, Y., et Yuen, K. Y. (2004). Severe acute respiratory syndrome. *Nat. Med* 10, S88-97.
- Peter-Katalinić, J. (2005). Methods in enzymology: O-glycosylation of proteins. *Methods Enzymol* 405, 139-71.
- Phan, T. G. et coll. (2007). Genetic heterogeneity, evolution, and recombination in noroviruses. *J. Med. Virol* 79, 1388-400.
- Piau, J. P., Labarriere, N., Dabouis, G., et Denis, M. G. (1994). Evidence for two distinct alpha(1,2)-fucosyltransferase genes differentially expressed throughout the rat colon. *Biochem J* 300 ( Pt 3), 623-6.
- Poet, S. E., Skilling, D. E., Megyesi, J. L., Gilmartin, W. G., et Smith, A. W. (1996). Detection of a non-cultivable calicivirus from the white tern (*Gygis alba rothschildi*). *J. Wildl. Dis* 32, 461-7.
- Prasad, B. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., et Estes, M. K. (1999). X-ray crystallographic structure of the Norwalk virus capsid. *Science* 286, 287-90.
- Prato, C. M., Akers, T. G., et Smith, A. W. (1974). Serological evidence of calicivirus transmission between marine and terrestrial mammals. *Nature* 249, 255-6.
- Preece, A. F., Strahan, K. M., Devitt, J., Yamamoto, F., et Gustafsson, K. (2002). Expression of ABO or related antigenic carbohydrates on viral envelopes leads to neutralization in the presence of serum containing specific natural antibodies and complement. *Blood* 99, 2477-82.

Pu, B. Q., Qian, N., et Cui, S. (1985). HA and HI tests for the detection of antibody titres to so called "haemorrhagic pneumonia" in rabbits. *Chin. J. Vet. Med.*, 16-7.

Puissant, B., Roubinet, F., Dellacasagrande, J., Massip, P., Abbal, M., Pasquier, C., Izopet, J., et Blancher, A. (2005). Decrease of Lewis frequency in HIV-infected patients: possible competition of fucosylated antigens with HIV for binding to DC-SIGN. *AIDS* 19, 627-30.

## Q

Qu, X. et coll. (2005). Identification of two critical amino acid residues of the severe acute respiratory syndrome coronavirus spike protein for its variation in zoonotic tropism transition via a double substitution strategy. *J. Biol. Chem* 280, 29588-95.

Queney, G., Ferrand, N., Marchandeu, S., Azevedo, M., Mougel, F., Branco, M., et Monnerot, M. (2000). Absence of a genetic bottleneck in a wild rabbit (*Oryctolagus cuniculus*) population exposed to a severe viral epizootic. *Mol Ecol* 9, 1253-64.

## R

Rademacher, C., Krishna, N., Palcic, M., Parra, F., et Peters, T. (2008). NMR Experiments Reveal the Molecular Basis of Receptor Recognition by a Calicivirus. *J Am Chem Soc*.

Raman, R., Raguram, S., Venkataraman, G., Paulson, J. C., et Sasisekharan, R. (2005). Glycomics: an integrated systems approach to structure-function relationships of glycans. *Nat Methods* 2, 817-24.

Reading, P. C., Morey, L. S., Crouch, E. C., et Anders, E. M. (1997). Collectin-mediated antiviral host defense of the lung: evidence from influenza virus infection of mice. *J. Virol* 71, 8204-12.

Reid, M. E., et Lomas-Francis, C. (2003). *The Blood Group Antigen FactsBook* (Academic Press).

Ren, W. et coll. (2006). Full-length genome sequences of two SARS-like coronaviruses in horseshoe bats and genetic variation analysis. *J Gen Virol* 87, 3355-9.

Rockx, B., De Wit, M., Vennema, H., Vinjé, J., De Bruin, E., Van Duynhoven, Y., et Koopmans, M. (2002). Natural history of human calicivirus infection: a prospective cohort study. *Clin. Infect. Dis* 35, 246-53.

Rogers, G. N., Pritchett, T. J., Lane, J. L., et Paulson, J. C. (1983). Differential sensitivity of human, avian, and equine influenza A viruses to a glycoprotein inhibitor of infection: selection of receptor specific variants. *Virology* 131, 394-408.

Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C. H., Meier-Ewert, H., Skehel, J. J., et Wiley, D. C. (1998). Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* 396, 92-6.

- Rota, P. A. et coll. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 300, 1394-9.
- Roubinet, F., Janvier, D., et Blancher, A. (2002). A novel cis AB allele derived from a B allele through a single point mutation. *Transfusion* 42, 239-46.
- Rouquier, S., Giorgi, D., Bergmann, A., Brandriff, B., et Lennon, G. (1994). The fucosyltransferase locus FUT1 maps distal to apolipoprotein loci E and C2 on human chromosome 19. *Cytogenet Cell Genet* 66, 70-1.
- Rouquier, S., Lowe, J. B., Kelly, R. J., Fertitta, A. L., Lennon, G. G., et Giorgi, D. (1995). Molecular cloning of a human genomic region containing the H blood group alpha(1,2)fucosyltransferase gene and two H locus-related DNA restriction fragments. Isolation of a candidate for the human Secretor blood group locus. *J Biol Chem* 270, 4632-9.
- Ruvoën-Clouet, N., Blanchard, D., André-Fontaine, G., et Ganière, J. P. (1995). Partial characterization of the human erythrocyte receptor for rabbit haemorrhagic disease virus. *Res Virol* 146, 33-41.
- Ruvoën-Clouet, N., Ganière, J. P., André-Fontaine, G., Blanchard, D., et Le Pendu, J. (2000). Binding of rabbit hemorrhagic disease virus to antigens of the ABH histo-blood group family. *J Virol* 74, 11950-4.

## S

- Sagar, M., Wu, X., Lee, S., et Overbaugh, J. (2006). Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J. Virol* 80, 9586-98.
- Saitou, N., et Yamamoto, F. (1997). Evolution of primate ABO blood group genes and their homologous genes. *Mol Biol Evol* 14, 399-411.
- Sandhoff, K., van Echten, G., Schröder, M., Schnabel, D., et Suzuki, K. (1992). Metabolism of glycolipids: the role of glycolipid-binding proteins in the function and pathobiochemistry of lysosomes. *Biochem Soc Trans* 20, 695-9.
- Sarnesto, A., Köhlin, T., Hindsgaul, O., Thurin, J., et Blaszczyk-Thurin, M. (1992). Purification of the secretor-type beta-galactoside alpha 1->2-fucosyltransferase from human serum. *J Biol Chem* 267, 2737-44.
- Sarnesto, A., Köhlin, T., Thurin, J., et Blaszczyk-Thurin, M. (1990). Purification of H gene-encoded beta-galactoside alpha 1->2 fucosyltransferase from human serum. *J Biol Chem* 265, 15067-75.
- Saunier, K., Barreaud, J. P., Eggen, A., Oriol, R., Levéziel, H., Julien, R., et Petit, J. M. (2001). Organization of the bovine alpha 2-fucosyltransferase gene cluster suggests that the Sec1 gene might have been shaped through a nonautonomous L1-retrotransposition event within the same locus. *Mol. Biol. Evol.* 18, 2083-91.
- Sawicki, S. G., et Sawicki, D. L. (2005). Coronavirus transcription: a perspective. *Curr. Top. Microbiol. Immunol* 287, 31-55.

- Schaecher, S. R., Mackenzie, J. M., et Pekosz, A. (2007). The ORF7b protein of severe acute respiratory syndrome coronavirus (SARS-CoV) is expressed in virus-infected cells and incorporated into SARS-CoV particles. *J. Virol* 81, 718-31.
- Schaffer, F. L., Soergel, M. E., Black, J. W., Skilling, D. E., Smith, A. W., et Cubitt, W. D. (1985). Characterization of a new calicivirus isolated from feces of a dog. *Arch. Virol* 84, 181-95.
- Seal, B. S., Lutze-Wallace, C., Kreutz, L. C., Sapp, T., Dulac, G. C., et Neill, J. D. (1995). Isolation of caliciviruses from skunks that are antigenically and genotypically related to San Miguel sea lion virus. *Virus Res* 37, 1-12.
- Secchiero, P., Sun, D., De Vico, A. L., Crowley, R. W., Reitz, M. S., Zauli, G., Lusso, P., et Gallo, R. C. (1997). Role of the extracellular domain of human herpesvirus 7 glycoprotein B in virus binding to cell surface heparan sulfate proteoglycans. *J. Virol* 71, 4571-80.
- Seymour, R. M., Allan, M. J., Pomiankowski, A., et Gustafsson, K. (2004). Evolution of the human ABO polymorphism by two complementary selective pressures. *Proc Biol Sci* 271, 1065-72.
- Shafti-Keramat, S., Handisurya, A., Kriehuber, E., Meneguzzi, G., Slupetzky, K., et Kirnbauer, R. (2003). Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. *J. Virol* 77, 13125-35.
- Shao, H., He, X., Achnine, L., Blount, J. W., Dixon, R. A., et Wang, X. (2005). Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* 17, 3141-54.
- Sharon, N. (1996). Carbohydrate-lectin interactions in infectious disease. *Adv Exp Med Biol* 408, 1-8.
- Shields, B., Mills, J., Ghildyal, R., Gooley, P., et Meanger, J. (2003). Multiple heparin binding domains of respiratory syncytial virus G mediate binding to mammalian cells. *Arch. Virol* 148, 1987-2003.
- Sinibaldi, L., Viti, D., Goldoni, P., Cavallo, G., Caroni, C., et Orsi, N. (1987). Inhibition of BK virus haemagglutination by gangliosides. *J. Gen. Virol* 68 ( Pt 3), 879-83.
- Slater-Handshy, T., Droll, D. A., Fan, X., Di Bisceglie, A. M., et Chambers, T. J. (2004). HCV E2 glycoprotein: mutagenesis of N-linked glycosylation sites and its effects on E2 expression and processing. *Virology* 319, 36-48.
- Smith, A. W., et Akers, T. G. (1976). Vesicular exanthema of swine. *J. Am. Vet. Med. Assoc* 169, 700-3.
- Smith, A. W., Akers, T. G., Madin, S. H., et Vedros, N. A. (1973). San Miguel sea lion virus isolation, preliminary characterization and relationship to vesicular exanthema of swine virus. *Nature* 244, 108-10.
- Smith, A. W., Mattson, D. E., Skilling, D. E., et Schmitz, J. A. (1983). Isolation and partial characterization of a calicivirus from calves. *Am. J. Vet. Res* 44, 851-5.
- Smith, A. W., Skilling, D. E., Cherry, N., Mead, J. H., et Matson, D. O. (1998). Calicivirus emergence from ocean reservoirs: zoonotic and interspecies movements. *Emerging Infect. Dis* 4, 13-20.

- Smith, A. W., Skilling, D. E., Ensley, P. K., Benirschke, K., et Lester, T. L. (1983). Calicivirus isolation and persistence in a pygmy chimpanzee (*Pan paniscus*). *Science* 221, 79-81.
- Smith, A. W., Skilling, D. E., et Ridgway, S. (1983). Calicivirus-induced vesicular disease in cetaceans and probable interspecies transmission. *J. Am. Vet. Med. Assoc* 183, 1223-5.
- Soejima, M., Fujihara, J., Takeshita, H., et Koda, Y. (2008). Sec1-FUT2-Sec1 hybrid allele generated by interlocus gene conversion. *Transfusion* 48, 488-92.
- Speak, A. O. et coll. (2007). Implications for invariant natural killer T cell ligands due to the restricted presence of isoglobotrihexosylceramide in mammals. *Proc. Natl. Acad. Sci. U.S.A* 104, 5971-6.
- Spear, P. G. (1993). Entry of alphaherpesviruses into cells. *Semin. Virol* 4, 167-180.
- Springer, G. F., et Horton, R. E. (1969). Blood group isoantibody stimulation in man by feeding blood group-active bacteria. *J. Clin. Invest* 48, 1280-91.
- Stevens, J., Blixt, O., Paulson, J. C., et Wilson, I. A. (2006). Glycan microarray technologies: tools to survey host specificity of influenza viruses. *Nat. Rev. Microbiol* 4, 857-64.
- Surjit, M., et Lal, S. K. (2008). The SARS-CoV nucleocapsid protein: a protein with multifarious activities. *Infect. Genet. Evol* 8, 397-405.
- Suzuki, T., Portner, A., Scroggs, R. A., Uchikawa, M., Koyama, N., Matsuo, K., Suzuki, Y., et Takimoto, T. (2001). Receptor specificities of human respiroviruses. *J. Virol* 75, 4604-13.
- Suzuki, Y., Ito, T., Suzuki, T., Holland, R. E., Chambers, T. M., Kiso, M., Ishida, H., et Kawaoka, Y. (2000). Sialic acid species as a determinant of the host range of influenza A viruses. *J. Virol* 74, 11825-31.
- Suzuki, Y., Nagao, Y., Kato, H., Suzuki, T., Matsumoto, M., et Murayama, J. (1987). The hemagglutinins of the human influenza viruses A and B recognize different receptor microdomains. *Biochim. Biophys. Acta* 903, 417-24.

## T

- Takeuchi, Y., Porter, C. D., Strahan, K. M., Preece, A. F., Gustafsson, K., Cosset, F. L., Weiss, R. A., et Collins, M. K. (1996). Sensitization of cells and retroviruses to human serum by (alpha 1-3) galactosyltransferase. *Nature* 379, 85-8.
- Tan, M., Huang, P., Meller, J., Zhong, W., Farkas, T., et Jiang, X. (2003). Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. *J. Virol* 77, 12562-71.
- Tang, N. L., Chan, P. K., Hui, D. S., To, K., Zhang, W., Chan, F. K. L., Sung, J. J., et Lo, Y. M. D. (2007). Lack of support for an association between CLEC4M homozygosity and protection against SARS coronavirus infection. *Nat. Genet* 39, 691-2; author reply 694-6.

- Thielens, N. M., Tacnet-Delorme, P., et Arlaud, G. J. (2002). Interaction of C1q and mannan-binding lectin with viruses. *Immunobiology* 205, 563-74.
- Traggiai, E., Becker, S., Subbarao, K., Kolesnikova, L., Uematsu, Y., Gismondo, M. R., Murphy, B. R., Rappuoli, R., et Lanzavecchia, A. (2004). An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nat. Med* 10, 871-5.
- Traum, J. (1934). Foot-and-mouth disease: specific treatment, eradication and differential diagnosis. *Proceedings of the 12th International Veterinary Congress. New York* 2, 87.
- Trinchera, M., Fabbri, M., et Ghidoni, R. (1991). Topography of glycosyltransferases involved in the initial glycosylations of gangliosides. *J Biol Chem* 266, 20907-12.
- Tripet, B., Howard, M. W., Jobling, M., Holmes, R. K., Holmes, K. V., et Hodges, R. S. (2004). Structural characterization of the SARS-coronavirus spike S fusion protein core. *J. Biol. Chem* 279, 20836-49.
- Trkola, A. et coll. (1996). Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J. Virol* 70, 1100-8.
- Trybala, E., Liljeqvist, J. A., Svennerholm, B., et Bergström, T. (2000). Herpes simplex virus types 1 and 2 differ in their interaction with heparan sulfate. *J. Virol* 74, 9106-14.
- Trybala, E., Olofsson, S., Mårdberg, K., Svennerholm, B., Umemoto, K., Glorioso, C., et Bergström, T. (2004). Structural and functional features of the polycationic peptide required for inhibition of herpes simplex virus invasion of cells. *Antiviral Res* 62, 125-34.
- Tseng, C. K. et coll. (2007). Severe acute respiratory syndrome coronavirus infection of mice transgenic for the human Angiotensin-converting enzyme 2 virus receptor. *J. Virol* 81, 1162-73.
- Turcot, A., Blancher, A., Le Moullac-Vaidye, B., Despiau, S., Rocher, J., Roubinet, F., Szpirer, C., et Le Pendu, J. (2003). Cloning of a rat gene encoding the histo-blood group B enzyme: rats have more than one Abo gene. *Glycobiology* 13, 919-28.
- Turcot-Dubois, A., Le Moullac-Vaidye, B., Despiau, S., Roubinet, F., Bovin, N., Le Pendu, J., et Blancher, A. (2007). Long-term evolution of the CAZY glycosyltransferase 6 (ABO) gene family from fishes to mammals--a birth-and-death evolution model. *Glycobiology* 17, 516-28.
- Tyrrel, D. A. J., Almedia, J. D., Berry, D. M., Cunningham, C. H., Hamre, D., Hofstad, M. S., Malluci, L., et McIntosh, K. (1968). Coronavirus. *Nature* 220, 650.

## U

- Uncapher, C. R., DeWitt, C. M., et Colonno, R. J. (1991). The major and minor group receptor families contain all but one human rhinovirus serotype. *Virology* 180, 814-7.

## V

- Van Valen, L. (1973). A new evolutionary law. *Evol. Theory* 1, 1-30.
- Varki, A. (2006). Nothing in glycobiochemistry makes sense, except in the light of evolution. *Cell* 126, 841-5.
- Vigerust, D. J., et Shepherd, V. L. (2007). Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol* 15, 211-8.
- Vigerust, D. J., Ulett, K. B., Boyd, K. L., Madsen, J., Hawgood, S., et McCullers, J. A. (2007). N-linked glycosylation attenuates H3N2 influenza viruses. *J. Virol* 81, 8593-600.

## W

- Wagner, F. F., et Flegel, W. A. (1997). Polymorphism of the h allele and the population frequency of sporadic nonfunctional alleles. *Transfusion* 37, 284-90.
- Walters, R. W., Yi, S. M., Keshavjee, S., Brown, K. E., Welsh, M. J., Chiorini, J. A., et Zabner, J. (2001). Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J. Biol. Chem* 276, 20610-6.
- Wang, F. Z., Akula, S. M., Pramod, N. P., Zeng, L., et Chandran, B. (2001). Human herpesvirus 8 envelope glycoprotein K8.1A interaction with the target cells involves heparan sulfate. *J. Virol* 75, 7517-27.
- Wang, S., Chou, T. W., Sakhatskyy, P. V., Huang, S., Lawrence, J. M., Cao, H., Huang, X., et Lu, S. (2005). Identification of two neutralizing regions on the severe acute respiratory syndrome coronavirus spike glycoprotein produced from the mammalian expression system. *J. Virol* 79, 1906-10.
- Watkins, W., et Morgan, W. (1959). Possible genetical pathways for the biosynthesis of blood group mucopolysaccharides. *Vox Sang* 4, 97-119.
- Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., et Kawakami, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiol. Rev* 56, 152-79.
- Wei, X. et coll. (2003). Antibody neutralization and escape by HIV-1. *Nature* 422, 307-12.
- White, L. J., Ball, J. M., Hardy, M. E., Tanaka, T. N., Kitamoto, N., et Estes, M. K. (1996). Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines. *J. Virol* 70, 6589-97.
- WHO (2003). Acute respiratory syndrome China, Hong Kong Special Administrative Region of China, and Viet Nam. *WHO WER* 78, 73-4.
- WHO Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. Available at: [http://www.who.int/csr/sars/country/table2004\\_04\\_21/en/index.html](http://www.who.int/csr/sars/country/table2004_04_21/en/index.html).

- Wilson, L., McKinlay, C., Gage, P., et Ewart, G. (2004). SARS coronavirus E protein forms cation-selective ion channels. *Virology* 330, 322-31.
- Witebsky, E. (1927). Okabe. *Stschr f Immun* 52, 359.
- de Witte, L., de Vries, R. D., van der Vlist, M., Yüksel, S., Litjens, M., de Swart, R. L., et Geijtenbeek, T. B. (2008). DC-SIGN and CD150 have distinct roles in transmission of measles virus from dendritic cells to T-lymphocytes. *PLoS Pathog* 4, e1000049.
- Wolk, T., et Schreiber, M. (2006). N-Glycans in the gp120 V1/V2 domain of the HIV-1 strain NL4-3 are indispensable for viral infectivity and resistance against antibody neutralization. *Med. Microbiol. Immunol* 195, 165-72.
- Wong, S. K., Li, W., Moore, M. J., Choe, H., et Farzan, M. (2004). A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J. Biol. Chem* 279, 3197-201.
- WuDunn, D., et Spear, P. G. (1989). Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J. Virol* 63, 52-8.
- Wyde, P. R., Moylett, E. H., Chetty, S. N., Jewell, A., Bowlin, T. L., et Piedra, P. A. (2004). Comparison of the inhibition of human metapneumovirus and respiratory syncytial virus by NMSO3 in tissue culture assays. *Antiviral Res* 63, 51-9.

## X

- Xu, H., Storch, T., Yu, M., Elliott, S. P., et Haslam, D. B. (1999). Characterization of the human Forssman synthetase gene. An evolving association between glycolipid synthesis and host-microbial interactions. *J. Biol. Chem* 274, 29390-8.
- Xu, Z. J., et Chen, W. X. (1989). Viral haemorrhagic disease in rabbits: a review. *Vet. Res. Commun* 13, 205-12.

## Y

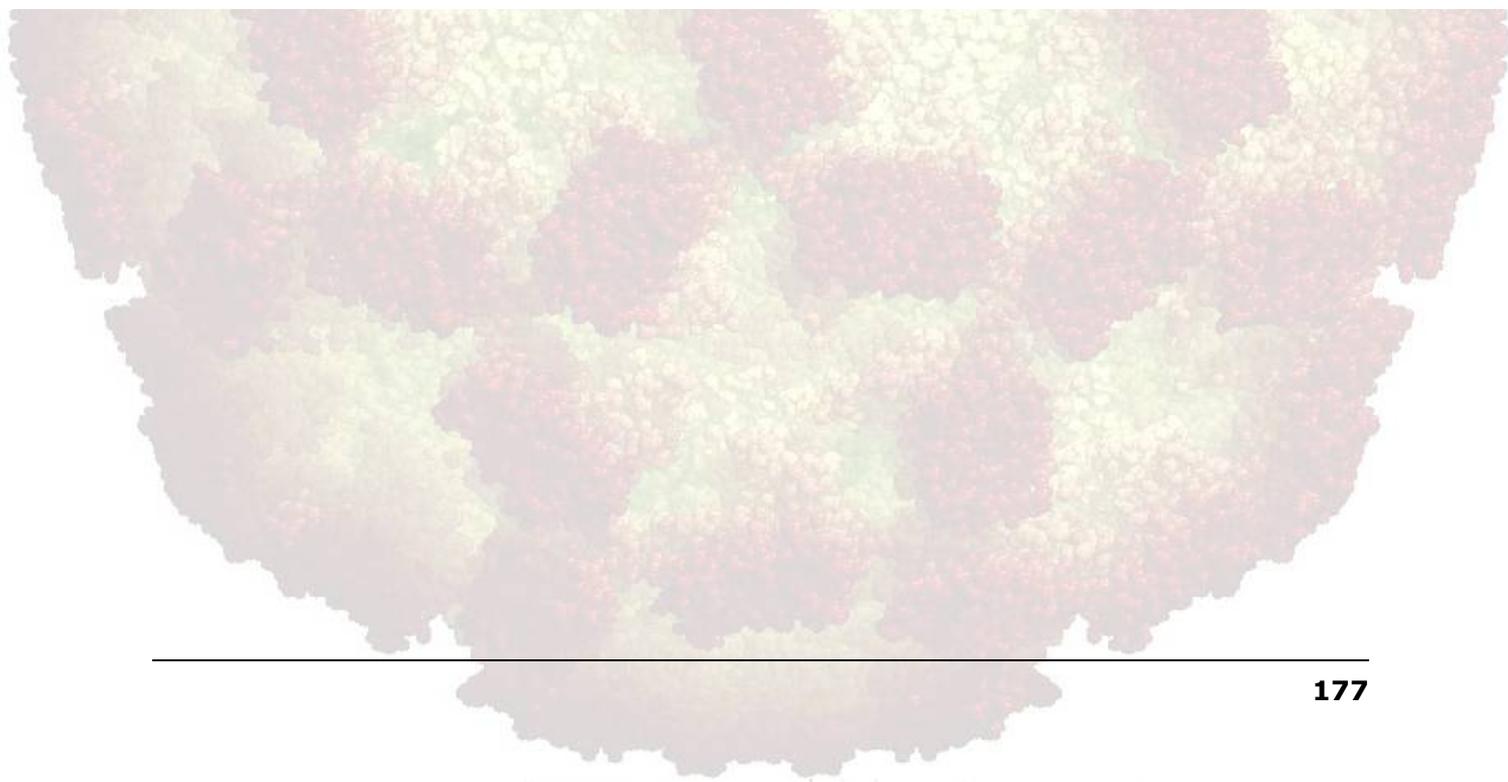
- Yamamoto, F., Clausen, H., White, T., Marken, J., et Hakomori, S. (1990). Molecular genetic basis of the histo-blood group ABO system. *Nature* 345, 229-33.
- Yamamoto, F., et Hakomori, S. (1990). Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. *J Biol Chem* 265, 19257-62.
- Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H., et Hakomori, S. (1990). Cloning and characterization of DNA complementary to human UDP-GalNAc: Fuc alpha 1->2Gal alpha 1->3GalNAc transferase (histo-blood group A transferase) mRNA. *J Biol Chem* 265, 1146-51.

- Yamamoto, F., et McNeill, P. D. (1996). Amino acid residue at codon 268 determines both activity and nucleotide-sugar donor substrate specificity of human histo-blood group A and B transferases. In vitro mutagenesis study. *J Biol Chem* 271, 10515-20.
- Yamamoto, F., McNeill, P. D., et Hakomori, S. (1992). Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem Biophys Res Commun* 187, 366-74.
- Yamamoto, F., McNeill, P. D., et Hakomori, S. (1991). Identification in human genomic DNA of the sequence homologous but not identical to either the histo-blood group ABH genes or alpha 1----3 galactosyltransferase pseudogene. *Biochem. Biophys. Res. Commun* 175, 986-94.
- Yamamoto, F., McNeill, P. D., Kominato, Y., Yamamoto, M., Hakomori, S., Ishimoto, S., Nishida, S., Shima, M., et Fujimura, Y. (1993). Molecular genetic analysis of the ABO blood group system: 2. cis-AB alleles. *Vox Sang* 64, 120-3.
- Yamamoto, F., et Yamamoto, M. (2001). Molecular genetic basis of porcine histo-blood group AO system. *Blood* 97, 3308-10.
- Yamamoto, M., Lin, X. H., Kominato, Y., Hata, Y., Noda, R., Saitou, N., et Yamamoto, F. (2001). Murine equivalent of the human histo-blood group ABO gene is a cis-AB gene and encodes a glycosyltransferase with both A and B transferase activity. *J Biol Chem* 276, 13701-8.
- Yamashita, Y., Chung, Y. S., Horie, R., Kannagi, R., et Sowa, M. (1995). Alterations in gastric mucin with malignant transformation: novel pathway for mucin synthesis. *J Natl Cancer Inst* 87, 441-6.
- Yang, Z., Huang, Y., Ganesh, L., Leung, K., Kong, W., Schwartz, O., Subbarao, K., et Nabel, G. J. (2004). pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. *J. Virol* 78, 5642-50.
- Yip, S. P. (2000). Single-tube multiplex PCR-SSCP analysis distinguishes 7 common ABO alleles and readily identifies new alleles. *Blood* 95, 1487-92.

## Z

- Zambon, M. C. (1999). Epidemiology and pathogenesis of influenza. *J. Antimicrob. Chemother* 44 Suppl B, 3-9.
- Zheng, D., Ando, T., Fankhauser, R. L., Beard, R. S., Glass, R. I., et Monroe, S. S. (2006). Norovirus classification and proposed strain nomenclature. *Virology* 346, 312-23.
- Zhi, L. et coll. (2007). Lack of support for an association between CLEC4M homozygosity and protection against SARS coronavirus infection. *Nat. Genet* 39, 692-4; author reply 694-6.
- Zhou, T. et coll. (2004). An exposed domain in the severe acute respiratory syndrome coronavirus spike protein induces neutralizing antibodies. *J. Virol* 78, 7217-26.

Zhu, Z., Gershon, M. D., Ambron, R., Gabel, C., et Gershon, A. A. (1995). Infection of cells by varicella zoster virus: inhibition of viral entry by mannose 6-phosphate and heparin. *Proc. Natl. Acad. Sci. U.S.A* 92, 3546-50.



**Les antigènes tissulaires de groupe sanguin ABH acteurs de la protection innée antivirale : exemples des Calicivirus (NV et RHDV) et du Coronavirus SRAS-CoV.**

Les antigènes tissulaires de groupe sanguin sont des motifs glycaniques présents sur les cellules épithéliales ou dans les sécrétions et également trouvés sur les érythrocytes de certains primates. Leur synthèse nécessite l'action de glycosyltransférases codées par les gènes *Fut1*, *Fut2*, *Sec1* ou *Abo*. Ils servent de ligands à diverses souches de Calicivirus du genre norovirus responsables de gastro-entérites chez l'homme, le polymorphisme aux loci *FUT2* et *ABO* contrôlant la sensibilité à ces virus. Dans ce travail nous montrons que le lait maternel contient des molécules de type mucine qui inhibent la fixation de la souche NV. L'inhibition est dépendante de la présence d'un allèle *FUT2* permettant la synthèse de glycanes  $\alpha$ 1,2fucosylés.

Un autre Calicivirus, le RHDV, responsable de la maladie hémorragique du lapin (RHD), se fixe sur un glycanne  $\alpha$ 1,2fucosylé. Nous avons observé qu'il existe des lapins sauvages présentant un déficit d'expression de ce ligand dont la fréquence dans les populations est augmentée à la suite d'épidémies de RHD. Un allèle du gène *Sec1* a pu être associé à la survie des animaux, suggérant une sélection par le virus.

La protéine d'enveloppe du Coronavirus du SRAS lui permet de se fixer sur son récepteur cellulaire et peut porter des antigènes ABH. Des anticorps anti-A du plasma de sujets O inhibent la fixation de cette protéine sur son récepteur lorsqu'elle est synthétisée dans des cellules A, expliquant la plus grande sensibilité des individus A (ou B) au SRAS. Ainsi les antigènes de groupes sanguins tissulaires pourraient contribuer à la protection innée antivirale par des mécanismes complémentaires, expliquant leur maintien au cours de l'évolution.

**Mots clés** : ABH, Norovirus, RHDV, SRAS-CoV, Coévolution hôte/pathogène

**Histo-blood group antigens ABH actors of the innate antiviral protection : examples of Caliciviruses (NV and RHDV) and of the SARS-CoV.**

Histo-blood group antigens are glycan motifs present on mammalian epithelial surfaces and in secretion. They are also found on erythrocytes from apes. Their synthesis requires glycosyltransferases encoded by the *Fut1*, *Fut2*, *Sec1* and *Abo* genes. They are ligands for various Calicivirus strains of the norovirus genus which cause gastroenteritis in humans and polymorphisms at the *FUT2* and *ABO* loci control sensitivity to these viruses. In this work, we show that maternal milk contains mucin-type molecules that inhibit NV strain binding to its  $\alpha$ 1,2fucosylated ligand. The inhibition depends upon a *FUT2* allele that allows  $\alpha$ 1,2fucosylation.

Another Calicivirus, RHDV, responsible for the rabbit hemorrhagic disease, binds to an  $\alpha$ 1,2fucosylated ligand. We observed that there exist wild rabbits presenting a deficit of expression of the ligand. Their frequency is increased following RHD outbreaks. Moreover, a *Sec1* allele has been associated with the survival of animals, suggesting a selection by the virus.

The envelope protein of the SARS Coronavirus allows binding to its cellular receptor and may carry ABH antigens. We show that natural antibodies from the plasma of O individuals inhibit attachment of this protein to its receptor when it is synthesized in A cells. This may explain the higher sensitivity of A (or B) individuals to SARS.

Overall, the results are discussed to suggest that histo-blood group antigens could contribute to innate antiviral protection by various complementary mechanisms, which could explain why they have been maintained throughout mammalian evolution.

**keywords** : ABH, Norovirus, RHDV, SRAS-CoV, host/pathogen coevolution