Fondness for sugars of enteric viruses confronts them with human glycans genetic diversity

Jacques Le Pendu[1], Nathalie Ruvoën-Clouet[1,2]

[1]CRCINA, Inserm, Université d'Angers, Université de Nantes, Nantes, France

[2]Oniris, Ecole Nationale Vétérinaire, Agroalimentaire et de l'Alimentation, Nantes, France

Corresponding author: Jacques Le Pendu

CRCINA, IRS2, 22 Boulevard Benoni-Goullin, 44200, Nantes, France

E-mail: Jacques.le-pendu@inserm.fr

Phone: 33 228 08 02 70

Acknowledgements

The work of the authors' laboratory on this topic was supported by Inserm (Institut National de la Santé et de la Recherche Médicale), the Région des Pays de la Loire and ANR (Agence Nationale de la Recherche) through several grants. The authors are most grateful to the past and present members of the laboratory Séverine Marionneau, Maha Zakhour, Jézabel Rocher, Béatrice Le Moullac-Vaidye, Laure Barbé, Adrien Breiman, Amira Khachou and Tasnuva Ahmed for their outstanding contributions to the research presented in this review.

Abstract

Together, norovirus and rotavirus are responsible for the majority of gastroenteritis cases worldwide, leading to a large number of children deaths in low-income countries. Both attach to glycans of the histo-blood group antigen type (HBGAs) widely expressed in the digestive tract of vertebrates, albeit with interspecies differences. In humans, their synthesis is performed by glycosyltransferases encoded by the highly polymorphic *ABO*, *FUT2* and *FUT3* genes that are under long-term balanced selection. The combination of functional and null or weak alleles at these loci provides a diversity of glycan structures that define the ABO, secretor and Lewis phenotypes. At the initial stage of infection norovirus and rotavirus attach to these glycans although distinct strains of each virus present different specificities for individual glycans, hence exhibiting preferences for different human phenotypes. Absence or low expression of the recognized glycan motifs due to the genetic polymorphism is associated with resistance to the disease, showing that the HBGA polymorphisms provide a population-based innate protection. Epidemiologically dominant strains of either norovirus or rotavirus display specificity for glycan motifs present in large fractions of the population, which may differ between geographical areas in accordance with the frequency of the *ABO*, *FUT2*, *FUT3* gene polymorphisms. Evidence for virus adaptation to these geographical differences is amounting, indicative of a host-pathogen co-evolution and suggesting that enteric pathogens such as norovirus and rotavirus are likely driving forces behind the balanced HBGA polymorphisms.

Diarrheal disease affects millions of people of all ages worldwide. It is a leading cause of mortality in children younger than 5 years, second only to pneumonia (Liu et al. 2012a). Although a large array of pathogens is involved, enteric viruses are the principal agents of the disease, the two main viruses responsible for this burden being rotavirus (RV) and norovirus (NoV) (Banyai et al. 2018). Genetic adaptation to these viruses is to be expected given their massive extant impact on human populations, akin to other pathogens that imposed strong selection pressures during human evolution (Quintana-Murci 2019).

Norovirus is the leading cause of acute gastroenteritis when considered across all ages worldwide, with an estimate of 685 million cases amounting to 18% of all cases (Kirk et al. 2015). Although NoV gastroenteritis is generally rather mild, it is estimated to cause the death of 70,000 to 200,000 young children/year, the vast majority of them in low-income countries while there is no vaccine commercially available as yet (Banyai et al. 2018). NoVs are a group of genetically highly diverse viruses that belong to the *Caliciviridae* family. They are non-enveloped icosahedral viruses with a single stranded RNA genome, currently classified in ten genogroups, of which three infect humans: GI, GII and GIV, with GIV being comparatively rare. Owing to the virus diversity, genogroups are further subdivided into genotypes (Chhabra et al. 2019). Of these and for the past 20 years, the GII.4 genotype has been overwhelmingly dominant worldwide, claiming as many as 70-80% of NoV-related gastroenteritis cases. Every few years, increases in NoV outbreaks are recorded due to the emergence of antigenically distinct GII.4 variants (de Graaf et al. 2016). More recently, non-GII.4 norovirus genotypes such as the previously rare GII.17 genotype emerged and predominated in various parts of Asia (Ao et al. 2017; Kwok et al. 2017; Niendorf et al. 2017).

RV is the most common cause of diarrheal mortality in children under 5 years of age globally. Prior the availability of a vaccine, in 2008, RV caused an estimated 440.000 children deaths yearly, primarily occurring in low-income countries. Following introduction of vaccines in numerous countries, this number diminished substantially, with the most recent estimate counting 129.000 deaths/year (Tate et al. 2016; Troeger et al. 2018). Unfortunately, and for reasons that are still unclear, the vaccines prove less efficient in several low-income countries (Glass et al. 2014). RVs are double-stranded segmented and non-enveloped RNA viruses about 70 nm in size. They belong to the *Reoviridae* family. Based on the genetic variability of the VP6 proteins that constitute the viral capsid intermediate layer, at least 8 different species (A to H) of rotaviruses are defined (Matthijnssens et al. 2012). The vast majority of human

cases are caused by group A rotaviruses (RVAs). Within RVAs, the two outer capsid proteins VP7 and VP4 are used for further classification of strains, defining the G- and P- genotypes, respectively. The majority of cases is due to viruses carrying G1P[8], G2P[4], G3P[8], G4P[8], G12P[8] and G9P[8], although other G and P combinations are being found in increasing numbers (Desselberger 2014). Notably, recent reports describe increasing rotavirus genotype diversity and circulation of non-vaccine genotypes in post-vaccine era (Roczo-Farkas et al. 2018; Tanaka et al. 2017), though immune-escaped strains have yet to be identified.

A common feature of human NoVs and RVAs is that they bind to a set of glycan motifs called histo-blood group antigens (HBGAs), which include ABH and Lewis antigens, through the external domain of their protruding capsid protein: the P domain of the unique capsid protein for NoVs and the VP8* domain of the VP4 capsid protein for RVAs. This review does not attempt to provide a comprehensive view of the topic since most of the relevant work has recently been reviewed (Nordgren and Svensson 2019; Ramani et al. 2016; Ruvöen-Clouet et al. 2013; Schroten et al. 2016; Singh et al. 2015b; Tan and Jiang 2014). It rather aims at discussing the potential consequences of this shared virus property of binding to polymorphic glycans. We will argue that these enteric viruses likely constitute a major factor driving evolution of glycan polymorphism, although other factors such as pathogenic bacteria, parasitic diseases, the microbiota and nutritional factors could also be involved (Cooling 2015).

Common genetic polymorphisms of glycans in the human small intestine

The mucosal surface of the small intestine is lined with a thick layer of glycans, mainly O-glycans of the mucin type, that forms a barrier likely restricting virus access to the enterocyte cell plasma membrane. Upon entrance into the intestine, enteric viruses will first encounter these complex sugars that they will use for attachment in an initial phase of the infection process. Unlike proteins, glycans are not direct products of the genes encoding regions. Their synthesis requires a large set of proteins that includes glycosyltransferases. These enzymes are involved in glycan biosynthesis by adding monosaccharides from phosphorylated donor molecules to acceptors that can be lipids, proteins, or other saccharides. Based on sequence similarities, these enzymes have been classified into more than 100 families (Cantarel et al. 2009). Several human glycosyltransferases belonging to the GT6, GT10 and GT11 families

present common genetic polymorphisms that lead to a loss of function, or to a severely decreased catalytic activity. They contribute to synthesis of the so-called HBGAs. The frequency of these mutant alleles varies geographically. As a result, combinations of these genes polymorphisms generate glycan diversity both in terms of expressed glycan motifs and of their distribution across human populations. In addition, variations also exist between mammalian species. HBGAs substitute a wide variety of protein-bound glycans (*N*-glycans and *O*-glycans) as well as lipid-bound glycans (glycosphingolipids) found at the surface of epithelial cells of various tissues in all mammals and secreted free or in complex forms in biological fluids such as saliva and milk. In the human small intestinal mucosa, the target of enteric viruses, the glycosyltransferases genes involved in the synthesis of HBGAs are the *ABO* gene of the GT6 family, the *FUT2* or secretor gene of the GT11 family and the *FUT3* or Lewis gene of the GT10 family (Fig. 1a). As depicted, the ABO, secretor and Lewis phenotypes of each individual depend on the combined polymorphisms at the three loci *ABO*, *FUT2* and *FUT3*. The H antigen is synthesized by addition of a fucose residue in α1,2 linkage and forms the precursor of the A and B antigens, as well as of the Lewis b and Lewis y antigens. In populations of European ancestry, about 20% individuals are devoid of α1,2-linked fucose in the small intestinal surface epithelium due to homozygosity of a null *FUT2* allele. Addition of a fucose residue in α1,4 or α1,3 linkage generates the Lewis antigens. Null *FUT3* alleles are responsible for the Lewis negative phenotype, represented by below 10% of individuals in populations of European ancestry (Race and Sanger 1975). In addition to their genetic polymorphism, several members of these important glycosyltransferases gene families correspond to pseudogenes in humans while they remain functional in other species (Abrantes et al. 2009; Nystrom et al. 2015; Turcot-Dubois et al. 2007). A prime example is given by the *GGTA1* gene of the GT6 family. It encodes an α1,3galactosyltransferase that catalyzes the transfer of a galactose residue onto an N-acetyllactosamine (type 2 precursor) to generate the so-called alphaGal antigen (Fig. 1b). The gene is functional in all mammals with the exception of Apes. When functional, its tissue expression pattern varies in a species-dependent manner. Thus, the corresponding alphaGal antigen is strongly expressed in bovine small intestine epithelial cells, but not in the porcine corresponding cells whilst it is completely lacking in all human tissues (Macher and Galili 2008; Zakhour et al. 2009).

HBGA-binding specificity of NoVs and RVAs is a driver of their epidemiology

Initial studies on noroviruses showed that the capsid protein of the prototype Norwalk strain (GI.1) attached to HBGAs containing the α1,2-linked fucose dependent upon functional *FUT2* alleles, such as Lewis b, H and A antigens, but not B antigen and that volunteers of the nonsecretor phenotype were fully resistant to infection by that strain (Lindesmith et al. 2003; Marionneau et al. 2002). Consistent with the virus glycan specificity, B blood group individuals proved partially protected (Hutson et al. 2002). Later studies indicated that most other strains also attached to HBGAs, albeit with varying specificities (Huang et al. 2005; Huang et al. 2003; Ruvöen-Clouet et al. 2013; Tan and Jiang 2011). Based on their glycan specificities, human noroviruses have been classified into subgroups according to their dependency upon the ABO, the secretor or the Lewis phenotypes (Le Pendu et al. 2006; Tan and Jiang 2014). Structural analyses described the interactions between the virus capsid protein protruding domain (P-domain) and various HBGA oligosaccharides. Several distinct binding sites and modes of binding across the norovirus strains diversity have thus been uncovered, illustrating the adaptation of noroviruses to the human gut glycan diversity (Bu et al. 2008; Cao et al. 2007; Chen et al. 2011; Choi et al. 2008; Hansman et al. 2011; Koromyslova et al. 2015; Kubota et al. 2012; Qian et al. 2018; Shanker et al. 2014; Singh et al. 2015a).

A large number of epidemiological studies comprehensively reviewed recently (Nordgren and Svensson 2019) reported an overall strong protection effect of the *FUT2* mutant alleles in the homozygous state. However, the effect was not found in all studies. This is likely explained by the diversity of strains involved in outbreaks. Indeed, the glycan-binding site of some strains favors recognition of the α1,4-linked fucose added by the *FUT3* enzyme with little influence of the α1,2-linked fucose, allowing recognition of both the secretor and nonsecretor phenotypes (Huang et al. 2005). Most likely, these strains are dependent on the presence of functional *FUT3* alleles and spare Lewis negative individuals as demonstrated through a study conducted in Burkina Faso where the frequency of Lewis negative individuals is high (Nordgren et al. 2013). An impact of the ABO phenotypes has also been reported, again with disparities across studies, likely explained by variations in the strains patterns of glycan recognition (Ruvöen-Clouet et al. 2013).

An important aspect concerns the GII.4 strains that have emerged as globally dominant strains in the past 20 years and the more recently emerged GII.17 strains as dominant epidemiologically in some parts of Asia (Chan et al. 2015; de Graaf et al. 2016). Among noroviruses, GII.4 strains stand out not only because of their prevalence but also because of

their evolution with periodic replacement by new immune escape variants, akin to influenza virus (de Graaf et al. 2016; Mallory et al. 2019). The reasons behind their large epidemiological dominance remain unclear, but their HBGA-binding characteristics could represent one of the underlying driving forces. GII.4 strains appear to recognize all secretor individuals, regardless of their ABO and Lewis status. In addition, an increased relative affinity for HBGAs paralleled the appearance of epidemiological dominance of GII.4 variants (de Rougemont et al. 2011). These characteristics confer GII.4 strains a broad spectrum of susceptible hosts worldwide and might contribute to facilitate transmission (Ruvöen-Clouet et al. 2013; Tan and Jiang 2011). Quite similarly, the recently emerged GII.17 strains also show much increased and broad spectrum HBGA binding ability in comparison with their older variants, which could contribute to their strong epidemiological impact (Chan et al. 2015; Qian et al. 2018; Zhang et al. 2015). Unlike GII.4 strains, these emergent GII.17 strains failed to become dominant outside Asian countries where the major *FUT2* mutant alleles confer a weak, albeit not null, secretor phenotype that potentially expands the susceptible fraction of the population (Ferrer-Admetlla et al. 2009; Pang et al. 2001). The presence of true nonsecretors due to nonsense mutation in the *FUT2* mutant alleles of African and European populations might contribute to restrict their transmission outside these Asian countries. Alternatively, the more limited spread of emergent GII.17 strains in comparison with pandemic GII.4 strains might be related to their lower relative affinity for HBGAs.

Earlier work on animal rotaviruses indicated these viruses bind to sialic acid (Lopez and Arias 2004). However, more recently it was shown that similar to noroviruses, human rotaviruses attach to HBGAs in a strain-dependent manner with distinct P genotypes using different binding sites (Hu et al. 2018; Jiang et al. 2017; Ramani et al. 2016; Tan and Jiang 2014). Thus, the VP8* attachment protein domain from strains of the P[8] and P[4] genotypes specifically binds to the Lewis b antigen (Barbé et al. 2018; Huang et al. 2012; Liu et al. 2012b; Sun et al. 2016a). Epidemiological studies reported a strong association between RVA gastroenteritis and the secretor phenotype, consistent with the requirement of a functional *FUT2* allele for synthesis of the Lewis b antigen in the small intestine (Imbert-Marcille et al. 2013; Kambhapati et al. 2015; Nordgren et al. 2014; Sun et al. 2016b; Van Trang et al. 2014; Yang et al. 2017). Studies performed in Burkina Faso and China additionally reported a much lower risk among children presenting a Lewis negative phenotype, also consistent with the requirement of the FUT3 enzyme in the synthesis of the Lewis b structure (Nordgren et al. 2014; Yang et al. 2017) and in accordance with a recent structural study showing that both the

a1,2 and a1,4-linked fucose residues are involved in the P[8] VP8* binding site (Xu et al. 2019). Nonetheless, two studies conducted in Tunisia and Bangladesh, failed to detect any effect of the secretor phenotype for P[8] rotavirus infection (Ayouni et al. 2015; Lee et al. 2018). These seemingly contradictory observations may be explained by the co-circulation of classical P[8] strains (P[8]-1-3) with emerging P[8]-4 strains (Zeller et al. 2015) that present a distinct glycan recognition pattern in these countries where high frequencies of nonsecretor and Lewis negative individuals are present (authors manuscript in preparation). The study conducted in Burkina Faso reported an association between the Lewis negative phenotype and infection by P[6] strains, whereas no association with the secretor phenotype was detected for these strains. This could also be explained by their glycan-binding specificity since they recognize a motif of the HBGA type 1 precursor and that addition of the α1,4-linked fucose impairs binding (Barbé et al. 2018). As a result, Lewis positive individuals are less well recognized, irrespective of their secretor phenotype (Barbé et al. 2018; Liu et al. 2016). Interestingly, the frequency of the Lewis negative phenotype is much higher in Burkina Faso than in most regions of the world (≈ 30% vs <10%), likely explaining why these P[6] strains preferentially circulate in some geographical areas of Africa and Asia.

As depicted on Fig. 2, this could have a strong impact on the efficacy of vaccines in such countries since the two major presently available vaccines have a P[8] subtype, their VP8* attaching to the Lewis b epitope, akin to circulating P[8] strains. Accordingly, low vaccine take was observed among nonsecretor children (Bucardo et al. 2018; Kazi et al. 2017; Lee et al. 2018).

RV strains with P[9], P[14] and P[25] genotypes bind to the A blood group antigen and *in vitro* studies demonstrated that this antigen serves as a functional receptor for this group of strains that circulate in domestic animals while not frequently encountered in humans (Hu et al. 2012; Liu et al. 2012b; Matthijnssens et al. 2009). Likewise, a group C rotavirus strain mainly involved in animal infection and in limited human family-based outbreaks was recently shown to exclusively recognize the blood group A antigen (Sun et al. 2018). Considering that the frequency of A antigen expression at the intestinal level in the population does not exceed 30%, it is not surprising that such strains do not dominate the human epidemiology, but rather cause sporadic outbreaks. The shared presence of the A antigen in the human intestine with that of most other mammals, including pigs and cows, likely contributes to cross-species transmission of these strains.

Inversely, species-specific expression of HBGAs may represent a species barrier. Thus, both bovine specific GIII norovirus strains and the bovine specific P[5] rotavirus strains attach specifically to the alphaGal HBGA (Fig. 1b) that is not expressed in humans but present in the bovine small intestinal mucosa (Alfajaro et al. 2019; Zakhour et al. 2009). Its recognition by bovine-specific strains of both norovirus and rotavirus constitutes an interesting case of convergent host-species adaptation.

Conclusions

A common feature of human noroviruses and rotaviruses binding to HBGAs is that the combined *ABO*, *FUT2* and *FUT3* genes polymorphisms allow protection of a substantial fraction of individuals from the disease caused by any given strain. This corresponds to a population or herd innate protection that may have emerged from a host-pathogen co-evolution process as documented in the case of RHDV and European rabbits (Le Pendu et al. 2014; Ruvöen-Clouet et al. 2013).

Overall, epidemiologically dominant strains possess HBGA-binding characteristics that allow them to recognize a broad host-spectrum, facilitating their transmission. In contrast strains that bind to glycans expressed in a more limited fraction of the population show a globally lower epidemiological impact. Moreover, regional differences in the epidemiology of these viruses can be accounted for, at least in part, by the variable frequencies of HBGA polymorphisms across human populations. This may contribute, among other factors, to explain the variable vaccines efficacies observed across countries since available rotavirus vaccines are live attenuated viruses.

The *ABO* gene comprises a large number of alleles, making it one of the most polymorphic human genes. Consistent with a role in host-microbes interactions, several studies indicated that it underwent a long-lived balancing selection that contributed to maintain the two major functional alleles along with the silent alleles (Calafell et al. 2008; Segurel et al. 2012; Villanea et al. 2015). Likewise, a long history of balancing selection has been detected for *FUT2* alleles (Ferrer-Admetlla et al. 2009; Silva et al. 2010). Maintenance of their polymorphisms over long evolutionary times indicates strong selective pressure by pathogenic agents. Interestingly, signs of long-term adaptation in the human genome were recently shown to have primarily occurred in response to RNA virus selection pressure (Azevedo et al. 2015; Enard and Petrov 2018). Since ABH antigens are expressed in the gut of all mammalian

species, whilst their additional presence on the vascular endothelium and red blood cells is restricted to Old World Monkeys and Apes, respectively, the balancing selection at the *ABO* locus was suggested to have originated in response to co-evolution with gut pathogens (Segurel et al. 2013). Thus, considering the high burden they impose on human populations and their strong impact on young children, noroviruses and rotaviruses arguably constitute major drivers of the balanced polymorphisms at the *ABO*, *FUT2* and possibly *FUT3* loci.

Conflict of interest statement:

On behalf of both authors, the corresponding author states that there is no conflict of interest.

References

Abrantes J, Posada D, Guillon P, Esteves PJ, Le Pendu J (2009) Widespread gene conversion of alpha-2-fucosyltransferase genes in mammals. J Mol Evol 69: 22-31.

Alfajaro MM, Kim J-Y, Barbé L, Cho EH, Park J-G, Soliman M, Baek Y-B, Kang M-I, Kim SH, Kim G-J, Park S-I, Le Pendu J, Cho KO (2019) Dual recognition of sialic acid and aGal epitopes by the VP8* domains of the bovine rotavirus G6P[5] WC3 and of its monoreassortant G4P[5] RotaTeq vaccines strains. J Virol 93: e00941-19.

Ao Y, Wang J, Ling H, He Y, Dong X, Wang X, Peng J, Zhang H, Jin M, Duan Z (2017) Norovirus GII.P16/GII.2-Associated Gastroenteritis, China, 2016. Emerg Infect Dis 23: 1172-1175.

Ayouni S, Sdiri-Loulizi K, de Rougemont A, Estienney M, Ambert-Balay K, Aho S, Hamami S, Aouni M, Neji-Guediche M, Pothier P, Belliot G (2015) Rotavirus P[8] Infections in Persons with Secretor and Nonsecretor Phenotypes, Tunisia. Emerg Infect Dis 21: 2055-8.

Azevedo L, Serrano C, Amorim A, Cooper DN (2015) Trans-species polymorphism in humans and the great apes is generally maintained by balancing selection that modulates the host immune response. Human Genomics 9. doi: 10.1186/s40246-015-0043-1

Banyai K, Estes MK, Martella V, Parashar UD (2018) Viral gastroenteritis. Lancet 392: 175-186.

Barbé L, Le Moullac-Vaidye B, Echasserieau K, Bernardeau K, Carton T, Bovin N, Nordgren J, Svensson L, Ruvoën-Clouet N, Le Pendu J (2018) Histo-blood group antigen-binding specificities of human rotaviruses are associated with gastroenteritis but not with in vitro infection. Scientific Reports 8: 12961.

Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS (2008) Structural basis for the receptor binding specificity of Norwalk virus. J Virol 82: 5340-5347.

Bucardo F, Nordgren J, Reyes Y, Gonzalez F, Sharma S, Svensson L (2018) The Lewis A phenotype is a restriction factor for Rotateq and Rotarix vaccine-take in Nicaraguan children. Sci Rep 8: 1502.

Calafell F, Roubinet F, Ramirez-Soriano A, Saitou N, Bertanpetit J, Blancher A (2008) Evolutionary dynamics of the human ABO gene. Hum Genet 124: 123-135.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucl Acid Res 37: D233-D238.

Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, Zhang X, Jiang X, Li X, Rao Z (2007) Structural basis for the recognition of blood group trisaccharides by norovirus. J Virol 81: 5949-5957.

Chan MC, Lee N, Hung TN, Kwok K, Cheung K, Tin EK, Lai RW, Nelson EA, Leung TF, Chan PK (2015) Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. Nat Commun 6: 10061.

Chen Y, Tan M, Xia M, Hao N, Zhang XC, Huang P, Jiang X, Li X, Rao Z (2011) Crystallography of a Lewis-binding norovirus, elucidation of strain-specificity to the polymorphic human histo-blood group antigens. PLoS Pathog 7: e1002152.

Chhabra P, de Graaf M, Parra F, Chan MC, Green KY, Martella V, Wang Q, White PA, Katayama K, Vennema H, Koopmans MPG, Vinjé J (2019) Updated classification of norovirus genogroups and genotypes. J Gen Virol: 1393-1406.

Choi J-M, Hutson AM, Estes MK, Prasad BV (2008) Atomic resolution structural characterisation of recognition of histo-blood group antigens by Norwalk virus. Proc Natl Acad Sci 105: 9175-9180.

Cooling L (2015) Blood groups in infection and host susceptibility. Clin Microbiol Rev 28: 801-870.

de Graaf M, van Beek J, Koopmans MP (2016) Human norovirus transmission and evolution in a changing world. Nat Rev Microbiol 14: 421-33.

de Rougemont A, Ruvoën-Clouet N, Simon B, Estienney M, Elie-Caille C, Aho S, Pothier P, Le Pendu J, Boireau W, Belliot G (2011) Qualitative and quantitative analysis of the binding of GII.4 norovirus variants onto human blood group antigens. J Virol 85: 4057-4070.

Desselberger U (2014) Rotaviruses. Virus Research 190C: 75-96.

Enard D, Petrov DA (2018) Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. Cell 175: 360-371.

Ferrer-Admetlla A, Sikora M, Laayouni H, Esteve A, Roubinet F, Blancher A, Calafell F, Bertanpetit J, Calafell F (2009) A natural history of FUT2 polymorphism in humans. Mol Biol Evol 26: 1993-2003.

Glass RI, Parashar U, Patel M, Gentsch J, Jiang B (2014) Rotavirus vaccines: successes and challenges. The Journal of infection 68 Suppl 1: 18.

Hansman G, Biertumpfel C, Georgiev I, McLellan J, Chen L, Zhou T, Katayama K, Kwong P (2011) Crystal structures of GII.10 and GII.12 norovirus protruding domains in complex with histo-blood group antigens reveal details for a potential site of vulnerability. J Virol 85: 6687-7388.

Hu L, Crawford SE, Czako R, Cortes-Penfield NW, Smith DF, Le Pendu J, Estes MK, Prasad BV (2012) Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature 485: 256-9.

Hu L, Sankaran B, Laucirica DR, Patil K, Salmen W, Ferreon ACM, Tsoi PS, Lasanajak Y, Smith DF, Ramani S, Atmar RL, Estes MK, Ferreon JC, Prasad BVV (2018) Glycan recognition in globally dominant human rotaviruses. Nat Commun 9: 2631.

Huang P, Farkas T, Zhong W, Tan M, Thornton SA, Morrow AL, 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: 6714-6722.

Huang P, Xia M, Tan M, Zhong W, Wei C, Wang L, Morrow A, Jiang X (2012) Spike protein VP8* of human rotavirus recognizes histo-blood group antigens in a type-specific manner. J Virol 86: 4833-4843.

Huang PW, Farkas T, Marionneau S, Zhong WM, Ruvoën-clouet N, Morrow A, Pickering LK, Newburg DS, Le Pendu J, 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 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.

Imbert-Marcille B-M, Barbé L, Dupé M, Le Moullac-Vaidye B, Besse B, Peltier C, Ruvoën-Clouet N, Le Pendu J (2013) A FUT2 gene common polymorphism determines resistance to rotavirus A of the P[8] genotype. J Infect Dis 209: 1227-1230.

Jiang X, Liu Y, Tan M (2017) Histo-blood group antigens as receptors for rotaivurs, new understanding on rotavirus epidemiology and vaccine strategy. Emerging Microbes and Infections 6: e22.

Kambhapati A, Payne DC, Costantini V, Lopman BA (2015) Host genetic susceptibility to enteric viruses: a systematic review and metaanalysis. Clin Infect Dis 62: 11-18.

Kazi AM, Cortese MM, Yu Y, Lopman B, Morrow AL, Fleming JA, McNeal MM, Steele AD, Parashar UD, Zaidi AKM, Ali A (2017) Secretor and Salivary ABO Blood Group Antigen Status Predict Rotavirus Vaccine Take in Infants. J Infect Dis 215: 786-789.

Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleesschauwer B, Dopfer D, Fazil A, Fischer-Walker CL, Hald T, Hall AJ, Keddy KH, Lake RJ, Lanata CF, Torgerson PR, Havelaar AH, Angulo FJ (2015) World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLoS Med 12: e1001921.

Koromyslova AD, Leuthold MM, Bowler MW, Hansman GS (2015) The sweet quartet: Binding of fucose to the norovirus capsid. Virology 483: 203-208.

Kubota T, Kumagai A, Ito H, Furukawa S, Someya Y, Takeda N, Ishii K, Wakita T, Narimatsu H, Shirato H (2012) Structural basis for the recognition of Lewis antigens by genogroup I norovirus. J Virol 86: 11138-50.

Kwok K, Niendorf S, Lee N, Hung TN, Chan LY, Jacobsen S, Nelson EAS, Leung TF, Lai RWM, Chan PKS, Chan MCW (2017) Increased Detection of Emergent Recombinant Norovirus GII.P16-GII.2 Strains in Young Adults, Hong Kong, China, 2016-2017. Emerg Infect Dis 23: 1852-1855. doi: 10.3201/eid2311.170561

Le Pendu J, Nystrom K, Ruvoen-Clouet N (2014) Host-pathogen co-evolution and glycan interactions. Curr Opin Virol 7: 88-94.

Le Pendu J, Ruvöen-Clouet N, Kindberg E, Svensson L (2006) Mendelian resistance to human norovirus infections. Sem Immunol 18: 375-386.

Lee B, Dickson DM, deCamp AC, Ross Colgate E, Diehl SA, Uddin MI, Sharmin S, Islam S, Bhuiyan TR, Alam M, Nayak U, Mychaleckyj JC, Taniuchi M, Petri WA, Jr., Haque R, Qadri F, Kirkpatrick BD (2018) Histo-Blood Group Antigen Phenotype Determines Susceptibility to Genotype-Specific Rotavirus Infections and Impacts Measures of Rotavirus Vaccine Efficacy. J Infect Dis 217: 1399-1407.

Lindesmith L, Moe CL, Marionneau S, Ruvoën-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.

Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, Rudan I, Campbell H, Cibulskis R, Li M, Mathers C, Black RE (2012a) Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. . Lancet 379: 2151-2161.

Liu Y, Huang P, Tan M, Biesiada J, Meller J, Castello AA, Jiang B, Jiang X (2012b) Rotavirus VP8*: phylogeny, host range, and interaction with histo-blood group antigens. J Virol 86: 9899-910.

Liu Y, Ramelot TA, Huang P, Liu Y, Li Z, Feizi T, Zhong W, Wu F-T, Tan M, Kennedy MA, Jiang X (2016) Glycan specificity of P[19] rotavirus and comparison with those of related P genotypes. J Virol 90: 9983-9996.

Lopez S, Arias CF (2004) Multistep entry of rotavirus into cells: a Versaillesque dance. Trends Microbiol 12: 271-8.

Macher BA, Galili U (2008) The Galα1,3Galβ1,4GlcNAc-R (α-Gal) epitope: A carbohydrate of unique evolution and clinical relevance. Biochem Biophys Acta 1780: 75-88.

Mallory ML, Lindesmith LC, Graham RL, Baric RS (2019) GII.4 Human Norovirus: Surveying the Antigenic Landscape. Viruses 11: E177.

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 histo-blood group antigens on gastro-duodenal epithelial cells of secretor individuals. Gastroenterology 122: 1967-1977.

Matthijnssens J, Otto PH, Ciarlet M, Desselberger U, Van Ranst M, Johne R (2012) VP6-sequence-based cutoff values as a criterion for rotavirus species demarcation. Arch Virol 157: 1177-82.

Matthijnssens J, Potgieter CA, Ciarlet M, Parreno V, Martella V, Banyai K, Garaicoechea L, Palombo EA, Novo L, Zeller M, Arista S, Gerna G, Rahman M, Van Ranst M (2009) Are human P[14] rotavirus strains the result of interspecies transmissions from sheep or other ungulates that belong to the mammalian order Artiodactyla? J Virol 83: 2917-29.

Niendorf S, Jacobsen S, Faber M, Eis-Hubinger AM, Hofmann J, Zimmermann O, Hohne M, Bock CT (2017) Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016. Euro Surveill 22.

Nordgren J, Nitiema LW, Ouermi D, Simpore J, Svensson L (2013) Host genetic factors affect susceptibility to norovirus infections in Burkina Faso. PLoS One 8: e69557.

Nordgren J, Sharma SB, bucardo F, Nasir W, Günaydin G, Ouermi D, Nitiema LW, Becker-Dreps S, Simpore J, Hammarström L, Larson G, Svensson L (2014) Both Lewis and Secretor Status Mediate Susceptibility to Rotavirus Infections in a Rotavirus Genotype Dependent Manner. Clin Infect Dis 59: 1567-1573.

Nordgren J, Svensson L (2019) Genetic susceptibility to human norovirus infection: an update. Viruses 11: E226.

Nystrom K, Abrantes J, Lopes AM, Le Moullac-Vaidye B, Marchandeau S, Rocher J, Ruvoen-Clouet N, Esteves PJ, Le Pendu J (2015) Neofunctionalization of the Sec1 alpha1,2fucosyltransferase paralogue in leporids contributes to glycan polymorphism and resistance to rabbit hemorrhagic disease virus. PLoS Pathog 11: e1004759.

Pang H, Koda Y, Soejima M, Fujitani N, Ogaki T, Saito A, Kawasaki T, Kimura H (2001) Polymorphism of the human ABO-Secretor locus (FUT2) in four populations in Asia: indication of distinct Asian subpopulations. Ann Hum Genet 65: 429-437.

Patnaik SK, Helmberg W, Blumenfeld OO (2012) BGMUT: NCBI dbRBC database of allelic variations of genes encoding antigens of blood group systems. Nucleic Acids Res 40: D1023-9.

Qian Y, Song M, Jiang X, Xia M, Meller J, Tan M, Chen Y, Li X, Rao Z (2018) Structural Adaptations of Norovirus GII.17/13/21 Lineage through Two Distinct Evolutionary Paths. J Virol 93: e-01655-18.

Quintana-Murci L (2019) Human Immunology through the Lens of Evolutionary Genetics. Cell 177: 184-199.

Race RR, Sanger R (1975) Blood groups in man. Blackwell Scientific Publications, Oxford

Ramani S, Hu L, Venkataram Prasad BV, Estes MK (2016) Diversity in rotavirus-host glycan interactions: a "sweet spectrum". Cell Mol Gastroenterol Hepatol 12: 263-273.

Roczo-Farkas S, Kirkwood CD, Cowley D, Barnes GL, Bishop RF, Bogdanovic-Sakran N, Boniface K, Donato CM, Bines JE (2018) The Impact of Rotavirus Vaccines on Genotype Diversity: A Comprehensive Analysis of 2 Decades of Australian Surveillance Data. J Infect Dis 218: 546-554.

Ruvöen-Clouet N, Belliot G, Le Pendu J (2013) Noroviruses and histo-blood groups: the impact of common host genetic polymorphisms on virus transmission and evolution. Rev Med Virol 23: 355-366.

Schroten H, Hanish FG, Hansman GS (2016) Human norovirus interactions with histo-blood group antigens and human milk oligosaccharides. J Virol 90: 5855-5859.

Segurel L, Gao Z, Przeworski M (2013) Ancestry runs deeper than blood: The evolutionary history of ABO points to cryptic variation of functional importance. Bioessays 35: 862-867.

Segurel L, Thompson EE, Flutre T, Lovstad J, Venkat A, Margulis SW, Moyse J, Ross S, Gamble K, Sella G, Ober C, Przeworski M (2012) The ABO blood group is a trans-species polymorphism in primates. Proc Natl Acad Sci 109: 18493-18498.

Shanker S, Czako R, Sankaran B, Atmar RL, Estes MK, Prasad BV (2014) Structural analysis of determinants of histo-blood group antigen binding specificity in genogroup I noroviruses. J Virol 88: 6168-6180.

Silva LM, Carvalho AS, Guillon P, Seixas S, Azevedo M, Almeida R, Ruvoën-Clouet N, Reis CA, Le Pendu J, Rocha J, David L (2010) Infection-associated FUT2 (fucosyltransferase 2) genetic variation and impact on functionality assessed by in vivo studies. Glycoconj J 27: 61-68.

Singh B, K, Koromyslova A, Hefele L, Gürth C, Hansman GS (2015a) Structural Evolution of the Emerging 2014-2015 GII.17 Noroviruses. J Virol 90: 2710-2715.

Singh B, K, Leuthold MM, Hansman GS (2015b) Human noroviruses' fondness for histo-blood group antigens. J Virol 89: 2024-2040.

Sun X, Guo N, Li D, Jin M, Zhou Y, Xie G, Pang L, Zhang Q, Cao Y, Duan Z (2016a) Binding specificity of P[8] VP8* proteins of rotavirus vaccine strains with histo-blood group antigens. Virology 495: 129-135.

Sun X, Guo N, Li J, Yan X, He Z, Li D, Jin M, Xie G, Pang L, Zhang Q, Liu N, Duan ZJ (2016b) Rotavirus infection and histo-blood group antigens in the children hospitalized with diarrhoea in China. Clin Microbiol Infect 22: 740.e1-3.

Sun X, Wang L, Qi J, Li D, Wang M, Cong X, Peng R, Chai W, Zhang Q, Wang H, Wen H, Gao GF, Tan M, Duan Z (2018) Human Group C Rotavirus VP8*s Recognize Type A Histo-Blood Group Antigens as Ligands. J Virol 92: e00442-18.

Tan M, Jiang X (2011) Norovirus-host interaction: Multi-selections by human histo-blood group antigens. Trends Microbiol 19: 382-388.

Tan M, Jiang X (2014) Histo-blood group antigens: a common niche for norovirus and rotavirus. Expert Rev Mol Med 16: e5.

Tanaka T, Kamiya H, Asada K, Suga S, Ido M, Umemoto M, Ouchi K, Ito H, Kuroki H, Nakano T, Taniguchi K (2017) Changes in Rotavirus Genotypes before and after Vaccine Introduction: a Multicenter, Prospective Observational Study in Three Areas of Japan. Jpn J Infect Dis 70: 448-452.

Tate JE, Burton AH, Boschi-Pinto C, Parashar UD (2016) Global, regional, and national estimates of rotavirus mortality in children <5 years of age, 2000-2013. Clin Infect Dis 62S: S96-S105.

Troeger C, Khalil IA, Rao PC, Cao S, Blacker BF, Ahmed T, Armah G, Bines JE, Brewer TG, Colombara DV, Kang G, Kirkpatrick BD, Kirkwood CD, Mwenda JM, Parashar UD, Petri WA, Jr., Riddle MS, Steele AD, Thompson RL, Walson JL, Sanders JW, Mokdad AH, Murray CJL, Hay SI, Reiner RC, Jr. (2018) Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea Among Children Younger Than 5 Years. JAMA Pediatr 172: 958-965.

Turcot-Dubois AL, Le Moullac-Vaidye B, Despiau S, Roubinet F, Bovin N, Le Pendu J, 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: 516-528.

Van Trang N, Vu HT, Le NT, Huang P, Jiang X, Anh DD (2014) Association between norovirus and rotavirus infection and histo-blood group antigen types in Vietnamese children. Journal of clinical microbiology 52: 1366-1374.

Villanea FA, Safi KN, Busch JW (2015) A General Model of Negative Frequency Dependent Selection Explains Global Patterns of Human ABO Polymorphism. PLoS One 10: e0125003.

Xu S, Liu Y, Tan M, Zhong W, Zhao D, Jiang X, Kennedy MA (2019) Molecular basis of P[6] and P[8] major human rotavirus VP8* domain interactions with histo-blood group antigens. bioRxiv http://dx.doi.org/10.1101/512301.

Yang T-A, Hou J-Y, Huang Y-C, Chen C-J (2017) Genetic susceptibility to rotavirus gastroenteritis and vaccine effectiveness in Taiwanese children. Sci Rep 7: 6412.

Zakhour M, Ruvoën-Clouet N, Charpilienne A, Langpap B, Poncet D, Peters T, Bovin N, Le Pendu J (2009) The alphaGal epitope of the histo-blood group antigen family is a ligand for bovine norovirus Newbury2 expected to prevent cross-species prevention. PLoS Pathog 5: e1000504.

Zeller M, Heylen E, Damanka S, Pietsch C, Donato C, Tamura T, Kulkarni R, Arora R, Cunliffe N, Maunula L, Potgieter C, Tamim S, Coster SD, Zhirakovskaya E, Bdour S, O'Shea H, Kirkwood CD, Seheri M, Nyaga MM, Mphahlele J, Chitambar SD, Dagan R, Armah G, Tikunova N, Van Ranst M, Matthijnssens J (2015) Emerging OP354-Like P[8] Rotaviruses Have Rapidly Dispersed from Asia to Other Continents. Mol Biol Evol 32: 2060-71.

Zhang XF, Huang Q, Long Y, Jiang X, Zhang T, Tan M, Zhang QL, Huang ZY, Li YH, Ding YQ, Hu GF, Tang S, Dai YC (2015) An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility. Sci Rep 5: 17687.

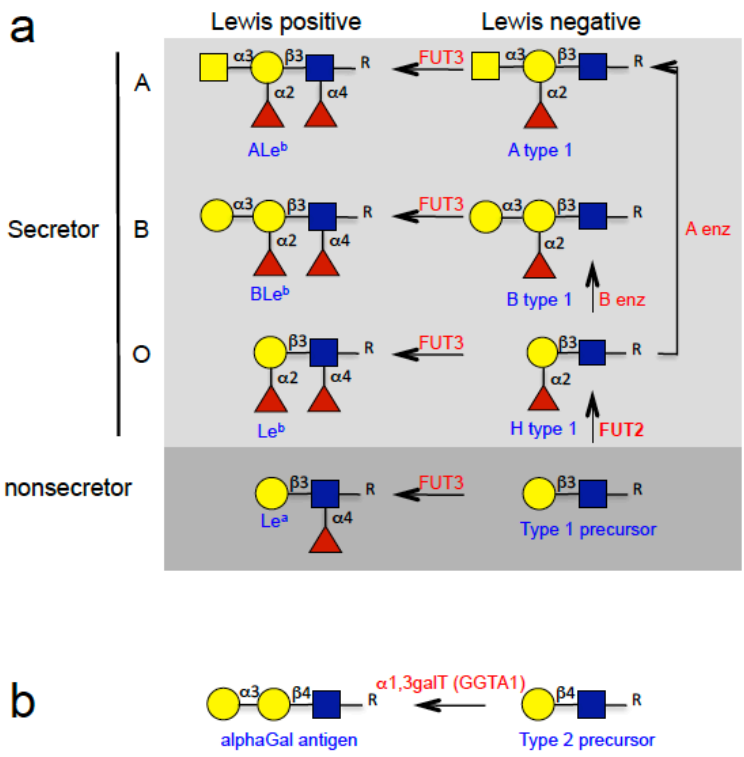

Figure 1: Biosynthesis of histo-blood group antigens in epithelial cells. **a/** synthesis of ABH and Lewis antigens starts from a precursor disaccharide (type 1 precursor) that represents the terminal portion of glycolipids, O- or N-glycans of glycoproteins. The enzymes FUT2 (in bold), FUT3, A and B (A enz, B enz) denoted in red, sequentially add monosaccharides to form the A, B, H and Lewis antigens denoted in blue. Polymorphisms of the *FUT2*, *FUT3* and *ABO* genes generate subgroups of people with distinct phenotypes with characteristic antigens expression (Secretor/nonsecretor, Lewis positive/Lewis negative, A, B and O). All allelic variants can be found in the BGMUT database (Patnaik et al. 2012). The *FUT2* G428A mutation introduces a stop codon and corresponds to a fully inactive allele found in many populations except in Asia. By Contrast, the A385T mutation found in Asia generates a protein with I>F amino acid change at position 129 which decreases its enzymatic activity. It is responsible for the so-called "secretor weak" phenotype. Of note, at each step of the biosynthesis, some untransformed structures remain (i.e. some type 1 precursor will still be available in Lewis positive individuals; or H type 1 in A or B secretors). **b**/ The alphaGal antigen is synthesized from type 2 precursor by addition of a galactose residue in α1,3 linkage. This is catalyzed by the α1,3 galactosyltransferase of the GT6 or ABO family

encoded by the *GGTA1* gene. The gene is functional in all mammal species but Apes. Owing to the lack of this epitope, humans possess circulating so-called natural anti-alphaGal antibodies, akin to natural anti-A and anti-B.

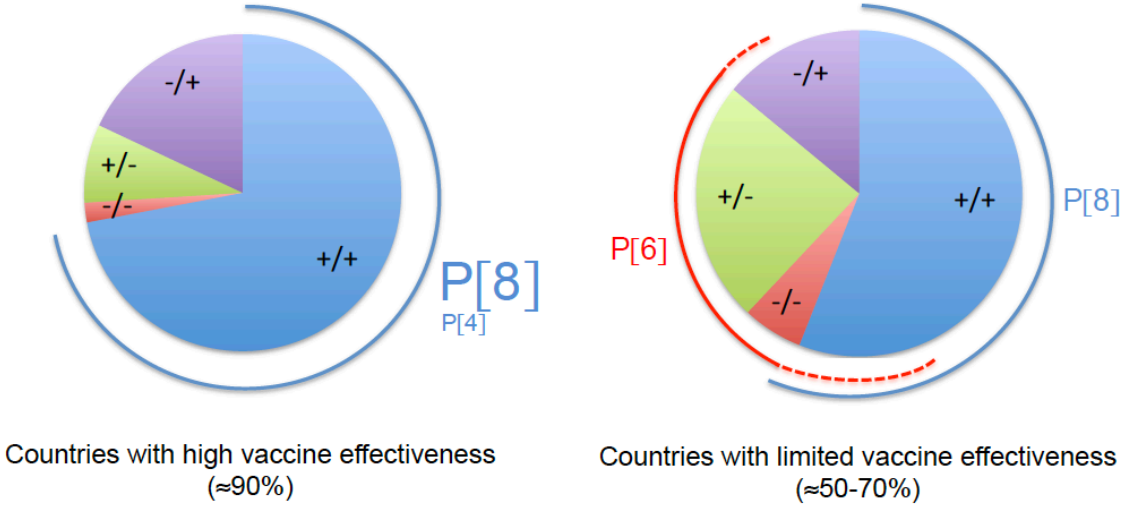

Figure 2: Impact of the frequencies of HBGA polymorphisms on rotavirus epidemiology and expected consequences on vaccine efficacy. Frequencies of HBGAs in two types of geographical areas are represented. In most developed countries (left) the combination of FUT2 and FUT3 positive individuals (Secretor/Lewis positive phenotype in blue) is high (Race and Sanger 1975), allowing circulation of strains with specificity for the difucosylated Lewis b antigen (P[8] and P[4]) that include the Rotarix and Rotateq attenuated live vaccine strains (Barbé et al. 2018; Desselberger 2014). Children with the remaining phenotypes, nonsecretor/Lewis positive (purple), Secretor/Lewis negative (green) and nonsecretor/Lewis negative (orange) are largely resistant to the disease caused by these strains (Imbert-Marcille et al. 2013; Kambhapati et al. 2015; Nordgren et al. 2014; Yang et al. 2017). These countries present a good match between genetic susceptibility to the disease, circulating strains and vaccine strains, explaining the high vaccine effectiveness. In countries where higher frequencies of either *FUT2* or *FUT3* null alleles are encountered such as Burkina Faso (Nordgren et al. 2014) (right), strains such as P[6] that favor infection of Lewis negative individuals (orange and green) and recognize glycans from Lewis positive individuals to some extent can maintain transmission (Barbé et al. 2018). In such countries there is a partial mismatch between susceptibility to the disease and that to vaccine strains, likely contributing to the observed lower vaccine effectiveness (Glass et al. 2014). Arc of circles indicate the susceptible fraction of the population to strains of the same color. Broken parts of the orange

arc indicate partial genetic susceptibility. +/+, +/-, -/+ and -/- indicate positivity or negativity for the Secretor and Lewis (FUT2/FUT3) characters, respectively.