Interplay between human leukocyte antigen genes and the microbial colonization process of the newborn intestine

De Palma. G.[1], Capilla, A[2], Nadal, I.[1], Nova, E.[3], Pozo, T.[3], Varea, V.[4], Polanco, I.[5,] Castillejo, G.[6], López, A.[7], Garrote, J.A.[8], Calvo C.[8], García-Novo M.D[9], Cilleruelo, M.L.[10], Ribes-Koninckx, C.[7], Palau, F.[2] and Y. Sanz[1]*

[1]Ecofisicología Microbiana. Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Valencia, Spain.
[2]Instituto de Biomedicina de Valencia (CSIC). CIBER de Enfermedades Raras (CIBERER), Valencia, Spain.
[3]Departmento de Metabolismo y Nutrición. Instituto del Frío- ICTAN (CSIC), Madrid. Spain.
[4]Gastroenterología, Nutrición y Hepatología Pediátrica. Hospital Universitario Sant Joan de Deu. Barcelona, Spain.
[5]Servicio de Gastroenterologia y Nutrición Pediátrica, Hospital Universitario La Paz, Madrid, Spain.
[6]Unidad de Gastroenterología Pediátrica, Hospital Universitario Sant Joan de Reus, Tarragona, Spain.
[7]Servcio de Gastroenterologia Pediátrica, Hospital Universitario La Fe, Valencia, Spain.
[8]Unidades de Investigación y Pediatría, Hospital Clínico Universitario, Valladolid, Spain.
9Unidad de Gastroenterologia, Hospital Universitario Infantil Niño Jesús, Madrid, Spain.
[10]Servicio de Pediatría, Hospital Severo Ochoa, Leganés, Madrid

*Corresponding author: Yolanda Sanz
                            Instituto de Agroquímica y Tecnología de Alimentos (CSIC)
                            PO Box 73, 46100 Burjassot, Valencia
                            Tel: 34 963900022
                            Fax: 34 963636301
                            E-mail: yolsanz@iata.csic.es

1

Abstract

Coeliac disease (CD) development involves genetic (HLA-DQ2/DQ8) and environmental factors. Herein, the influence of the HLA-DQ genotype on the gut colonization process of breast-fed children was determined. A cohort of 20 newborns, with at least one first-degree relative with CD, were classified according to their HLA-DQ genotype into high, intermediate and low genetic risk groups, showing 24-28%, 7-8% and less than 1% probability to develop CD, respectively. Faecal microbiota was analysed at 7 days, 1 and 4 months of children's age by fluorescence *in situ* hybridization. When considering all data, *Bacteroides-Prevotella* group proportions were higher (P<0.05) in the high than in the intermediate and low genetic risk groups. Gram-negative bacteria, *E. coli*, *Streptococcus-Lactococcus*, *E. rectale-C. coccoides,* sulphate-reducing bacteria, *C. lituseburense* and *C. histolyticum* group proportions were also significantly higher (P<0.05) in the high than in the low genetic risk group. Correlations between these bacterial groups and the genetic risk were also detected (P<0.05). In addition, the number and type of CD relative seemed to influence (P<0.050) these bacterial proportions in children at CD risk. At 4 months of age, similar relationships were established between the high genetic risk to develop CD and the proportions of *Streptococcus-Lactococcus* (P<0.05), *E. rectale-C. coccoides* (P<0.05), *C. lituseburense* (P<0.05), *C. histolyticum* (P<0.05), *Bacteroides-Prevotella* (P<0.10) groups and total Gram-negative bacteria (P<0.05). The results suggest a relationship between HLA-DQ genes and the gut microbial colonization process that could lead to a change in the way this disorder is investigated.

Introduction

Coeliac disease (CD) is the commonest immune-mediated enteropathy triggered by the ingestion of cereal gluten proteins in genetically susceptible individuals. CD is a chronic inflammatory disorder of the small intestinal mucosa caused by an abnormal Th1 immune response to ingested gluten (Rodrigo, 2006). The typical manifestation of CD is the malabsorption syndrome with chronic diarrhoea, weight loss, abdominal distension and impaired growth (Fasano and Catassi, 2005). CD can develop at any age, but it often manifests in early childhood between 6 and 24 months of age after the introduction of gluten in the diet (Fasano and Catassi, 2005).

CD is strongly associated with the human leukocyte antigen (HLA) genes of the major histocompatibility complex (MHC) genomic region. Approximately 95 % of the patients are HLA-DQ2 or HLA-DQ8 positives; however, only a small percentage (20-50%) of those bearing high risk HLA haplotypes develop CD (Mearin et al., 2005). Studies with twins showed that in 25% of the cases one twin of the pair did not develop CD, supporting a role for environmental factors in the presentation of this disorder (Greco et al., 2002). In particular, the infant feeding pattern is thought to be involved in CD risk (Branski et al., 2006). Breast-feeding seems to exert a protective effect against CD development (Branski et al., 2006; Ivarsson et al., 2002), which may be related to its influence on the microbial colonization process of the newborn intestine (Sanz et al., 2008). The incidence of infections in early life, which may trigger a Th1 pro-inflammatory response and disrupt the gut microbial balance, has also been related to an increased risk of CD development (Ivarsson et al., 2003; Stene et al., 2006; Sanz et al., 2008).

The microbiota rapidly colonizes the gut after birth and this process is influenced by a number of factors, including the heredity of the mother, the immediate living environment, the feeding practices, microbial infections, and the host's genetics (Bjorkstén, 2006). The

early stage of colonization is characterized by the presence of higher levels of facultative anaerobes (*Enterobacteriaceae*, *Lactobacillus* and *Streptococcus*) than anaerobic bacteria (e.g. *Bifidobacterium*, *Bacteroides*, *Clostridium* and *Eubacterium*), but these proportions are reversed within one week after birth. The *Bifidobacterium* genus is predominant in breast-fed infants, reaching up to 90% of the total faecal microbiota, whereas in formula-fed infants the microbiota is more heterogeneous (Salminen and Isolauri, 2006). The intestinal microbiota plays a pivotal role in the morphology and physiology of the gut, providing the basis for a healthy gut (Salminen and Isolauri, 2006).

The gut microbiota serves several functions including regulation of gut permeability and immunity to develop adequate responses to microbial and dietary antigens and avoid overreactions (Conte et al., 2006; Sanz et al., 2007). It was recently demonstrated that the neonatal gut microbiota plays an important role in the development of atopy. In particular, a lower ratio of bifidobacteria to clostridia in the gut microbiota at neonatal age was shown to precede the development of atopy (Kalliomäki et al., 2001a). A recent study also demonstrated that probiotic supplementation to mothers before delivery and to their infants could prevent atopic disease in children (Kalliomäki et al., 2001b). Differences in the composition of the gut microbiota have been detected between healthy controls and CD patients recently (Nadal et al., 2007; Sánchez et al., 2008; Collado et al., 2009). Nevertheless, the knowledge of the role played by the gut microbiota in CD is limited and there is no information of its possible contribution to the disease risk based on prospective studies.

The objective of this preliminary study was to determine whether the gut colonization pattern of newborns and breast-fed children could be influenced by the genetic risk to develop CD based on their HLA-DQ status, in order to get new insights into the influence of host-microbe interactions in this disorder.

Materials and Methods

Subjects and samplings

This study included 20 healthy infants with at least one first-degree relative suffering from CD. Faecal samples were collected from every subject at 7 days, 1 month and 4 months of age and kept immediately at -20ºC until analysed. DNA for HLA-DQ genotyping was also obtained from oral mucosa cells by scraping the inner side of the children's cheek with sterile swabs (Copan innovation, Sarstedt, Germany) at 4 months of age. Data on clinical and dietary history of every subject were recorded including type of delivery, time of gestation, weight, size and type of feeding (Table 1). Only subjects under exclusive breast-milk feeding practices were included in the study. The study was conducted in accordance with the ethical rules of the Helsinki Declaration (Hong Kong revision, September 1989), following the EEC Good Clinical Practice guidelines (document 111/3976/88 of July 1990) and current Spanish law which regulates clinical research in humans (Royal Decree 561/1993 regarding clinical trials). The study was approved by the local Ethic Committees and written informed consent was obtained from the parents of children included in the study.

Analysis of the faecal microbiota by fluorescence in situ hybridization (FISH) and flow cytometry (FCM)

Faecal samples were diluted 1 to 10 in phosphate-buffered saline solution (PBS 130 mM sodium chloride, 10 mM sodium phosphate, pH between 7.2 and 7.4) and homogenized in a Lab Blender 400 stomacher (Seward Medical London, UK) for 5 min. Cells were fixed overnight with 4% paraformaldehyde. Then, bacteria were washed twice with PBS and bacterial pellets were suspended in 1:1 chilled ethanol-PBS and stored at -80ºC until analyzed (Nadal et al., 2007).

The universal and group-specific probes used for detection of the main bacterial groups colonizing the human gut are shown in Table 2 (MOLBIOL, Berlin, Germany). The group-specific probes were labelled at the 5'-end with fluorescein isothiocyanate (FITC); showing green fluorescence. EUB 338 probe, targeting conserved sequences within the bacterial domain, was labelled at the 5'end with the indocyanine dye Cy3 showing red fluorescence, and was used as positive control of total bacteria (Amann et al., 1990). NON EUB 338 probe was labelled with FITC and used as negative control to eliminate the background fluorescence (Wallner et al., 1993). Aliquots of 45 µl fixed bacterial samples were incubated with 5 µl of each fluorescent probe (50ng/µl) in hybridisation solution (10 mM Tris–HCl, 0.9 M NaCl, pH 8.0, and 0.1% [w/v] sodium dodecyl sulphate) at appropriate temperature (45-50 ºC) overnight. Hybridization conditions were those described by other authors (Table 2), except for *E. coli* and *Staphylococcus*, which were detected by hybridisation at 50 ºC, and for *Lactobacillus*, which were detected by hybridisation at 45 ºC. Afterwards, bacterial cells were incubated with 500 µl washing solution (10 mM Tris–HCl, 0.9 M NaCl, pH 8.0) at the corresponding hybridization temperature for 30 min to remove non specific binding of the probes. Hybridised cells were finally pelleted by centrifugation (12 000 *g* for 5 min) and suspended in 500 µl of PBS for flow cytometry detection. Bacterial groups were enumerated by combining each FITC-labelled group-specific probe with the EUB 338-Cy3 probe, and expressed as a ratio of cells hybridising with the FITC-labelled specific probe to cells hybridising with the EUB 338-Cy3 probe This proportion was corrected by subtracting the background fluorescence obtained with the negative control probe NON EUB 338 (Sokol et al., 2006; Nadal et al., 2007).

Flow cytometry detections were performed using an EPICS® XL-MCL flow cytometer (Beckman Coulter, Florida, USA) as previously described (Nadal et al., 2007). This instrument is equipped with two light scatter detectors that measure forward (FSC) and side

scatter (SSC) and fluorescence detectors that detect appropriately filtered light at green (FL1, 525 nm) and red-orange (FL3, 620 nm) wavelengths. The event rate was kept at the lowest setting (200-300 events per second) to avoid cell overlap. A total of 15 000 events were recorded in a list mode file and analyzed with the System II V.3 software (Beckman Coulter).

HLA-DQ Genotyping

DNA was extracted from oral mucosa cells by scraping the inner side of the children's cheek with sterile swabs (Copan innovation, Sarstedt, Germany). The cotton was cut and incubated in DLB buffer (100mM Tris HCL pH 7.4, 10mM EDTA pH 8, 10 mM NaCl per litre of distilled water); 10% SDS and 10 μl proteinase K (20 mg/ml) at 65ºC for 1 h and then, a standard phenol-chloroform DNA extraction method was carried out. Extracted DNA was stored in TE buffer (10ml Tris HCl pH 8, 200 μl EDTA 0.5M pH 8 per litre of distilled water) at -20ºC until use for genotyping. The DNA concentration, around 50-100 μg/ml, was quantified by using the NanoDrop$^{©}$ Spectrophotometer. Low-resolution HLA-DQB1 typing was performed by PCR-SSP (Polymerase Chain Reaction-Sequence Specific Primers) analysis. Each PCR reaction was performed on about 60–90 ng of extracted DNA, 0·5 U of BIOTOOLS DNA polymerase (BIOTOOLS B&M S.A, Spain), 1 × PCR Master Mix (Dynal *AllSet*$^{+TM}$ SSP or *Olerup* SSP^{TM}) containing nucleotides (200 μmol L^{-1} each), PCR buffer (50 mmol L^{-1} KCl, 1·5 mmol L^{-1} MgCl$_2$, 10 mmol L^{-1} Tris–HCl pH 8·3, 0·001% w/v gelatine), 5% glycerol and 100 μg mL^{-1} cresol red, 0·25 μmol L^{-1} of each allele- or group-specific primer pair and 0·1 μmol L^{-1} of internal positive control primer pair matching a segment of the human growth hormone gene in a final volume of 10 μL. The PCR program was as follows: an initial denaturation step at 94 °C for 2 min, followed by 10 two-temperature cycles (94 °C for 10 s and 65 °C for 60 s) and finally by 20 three-temperature cycles (94 °C for 10 s, 61 °C for 50 s and 72 °C for 30 s). Detection of amplified alleles was

carried out on a 2% agarose gel after electrophoresis and ethidium bromide staining. Although the allele DQB1*02 is determinant to be DQ2, HLA-DQA1 alleles were genotyped in a stepwise fashion for a high resolution typing to sharpen the risk classification of each individual.

5

Statistical analysis

Statistical analysis was done using the SPSS 15.0 software (SPSS Inc, Chicago, IL, USA). To determine whether the microbiological data could be adequately modelled by a normal distribution, Kolmogorov-Smirnov test was run and the distribution of the data resulted non normal at 90-99% of confidence. Therefore, microbiological data are expressed as medians of bacterial proportions with interquartile ranges (IQR) and the differences were determined by applying the Mann–Whitney U and Wilcoxon W tests for non-parametric data. The presence of correlations between factors was tested by Spearman non parametric correlation coefficient. A mixed model analysis was applied to study the effect of the genetic risk on bacterial counts, and the interactions among all factors with the time sampling as a repeated measure. Significant differences were established at *P* value less than 0.10 and 0.05.

Results

Subjects and HLA genotypes

The cohort of 20 newborns included in this study were born at term and they had uncomplicated perinatal period. They showed a mean gestational age of 39.16 ± 1.52 weeks, and a mean weight and size of 3321.94 ± 491.26 g and 49.28 ± 2.02 cm, respectively (Table 1). Most newborns were naturally delivered (17 out of 20) and all were under exclusive breast-milk feeding. This cohort of children was analysed for the HLA class II DQA1 and DQB1 genotypes. The children were classified in three main CD risk groups, according to

their DQ haplotype, loosely based on the criterion of Bourgey et al. (2007) and considering the HLA distribution of the Eastern Spanish population (Capilla et al., 2007). The first one included those individuals carrying the DQ2 haplotype in both cis (DQA1*0501-DQB1*0201 in homozygosis) or trans conformation (DQA1*0201-DQB1*0202 with DQA1*0505-DQB1*0301 in heterozygosis). The second group included those individuals carrying the DQA1*0501-DQB1*0201 haplotype along with any other haplotype. The third group included those individuals with other common genotypes not associated with CD. Of the total number of children, 28.6% were in the first group with the highest probability (24-28% of developing CD (the high risk group), 33.3% were in the second group with a probability of 7-8% of developing CD (the intermediate risk group), and 38.1% were in the third group with less than 1% probability of developing CD (the low risk group). These probabilities are theoretical and have been estimated on basis of the genotype of each newborn, without considering the genotype of the affected relative.

Microbial colonization process in breast-fed children during the first 4 months of life

The dynamics of the microbial colonization process from 7 days till 4 months of life was determined (Table 3). In general, most of the analysed bacterial group proportions tended to increase with age. *Bacteroides-Prevotella* proportions significantly increase from 7 days to 1 month of age (*P*=0.017), and sulphate-reducing bacteria proportions increased from 7 days to 4 months (*P*=0.009) and from 1 to 4 months (*P*=0.021). *Staphylococcus* proportions also increased from 1 to 4 months of age (*P*=0.086). Statistically significant correlations were also found between age and bacterial counts for sulphate-reducing bacteria (R=0.356, P=0.005) and Gram-negative bacteria (R=0.258, P=0.046).

All microbiological data were also analysed as a function of the genetic risk to develop CD (Figure 1). Total bacterial proportions were significantly higher in the high and intermediate

genetic risk group than in the low genetic risk group (P=0.010 and P=0.030, respectively). Total Gram-negative bacterial proportions were higher in the high genetic risk group than in the intermediate and low genetic risk groups, showing significant differences between the high and the low genetic risk groups (*P*=0.009). In addition, significantly higher proportions

5 of *Bacteroides-Prevotella* group were found in children of the high genetic risk group in comparison with those of the intermediate and low genetic risk groups (P=0.015 and P=0.042, respectively). Similar trends were found for *E. coli* (P=0.011), *Streptococcus-Lactococcus* (P=0.003), *E. rectale-C. coccoides* (P=0.004), sulphate-reducing bacteria (P=0.022), *C. lituseburense* (P=0.002) and *C. histolyticum* (P=0.020) groups, showing

10 proportions significantly higher in the high than in the low genetic risk group. Significant correlations were also found between the genetic risk group of the infants and specific bacterial group proportions by applying Spearman correlation coefficient (Figure1). Higher proportions of total Gram-negative bacteria (R=-0.364, P=0.004), *E. coli* (R=-0.352, P=0.007), *Bacteroides-Prevotella* (R=-0.265, P=0.046), *Staphylococcus* (R=-0.294,

15 P=0.022), *Streptococcus-Lactococcus* (R=-0.386, P=0.002), sulphate-reducing bacteria (R=-0.311, P=0.016), *E. rectale-C. coccoides* (R=-0.379, P=0.003), *C. histolyticum* (R=-0.317, P=0.014) and *C. lituseburense* (R=-0.358, P=0.005) groups were correlated with increased genetic risk (Figure 1).

When considering the effect of the diseased relatives of the children under study, higher

20 proportions of *E. coli* and *C. lituseburense* groups were observed when the affected relative was either the mother or the father than when the affected relative was one siblings (*P*=0.033 and *P*=0.016, respectively). Differences were also observed in bacterial proportions between those children, who have one of the parents plus one sibling with CD, and those who only have one relative (the mother or the father) with CD. In particular, *E. coli* (*P*=0.042), *E.*

25 *rectale-C. coccoides* (*P*=0.019), *C. lituseburense* (*P*=0.028)*, Lactobacillus* (*P*=0.040),

Streptococcus-Lactococcus (*P*=0.040), Gram-positive bacteria (*P*=0.019) and total bacteria (*P*=0.028) were higher when the affected relatives were one of the parents together with a sibling. Statistically significant correlations were found between the CD relative being the mother and increased proportions of *E. coli* (R=0.275, P=0.048), *E. rectale-C. coccoides* (R=0.304, P=0.025), *C. lituseburense* (R=0.315, P=0.019)*, Lactobacillus, Streptococcus-Lactococcus* (R=0.272, P=0.047) and Gram positive bacteria (R=0.323, P=0.016).

When data were adjusted by age, total bacteria (P=0.016), *Streptococcus-Lactococcus* (*P*=0.010), *E. rectale-C. coccoides* (*P*=0.093), *C. lituseburense* (*P*=0.015), *C. histolyticum* (P=0.015) and *Bacteroides-Prevotella* (*P*=0.100) group proportions were higher in the high genetic risk group than in the low genetic risk group at 4 months of age. Similar correlations were found between higher proportions of these bacterial groups and increased genetic risk of CD development (Figure 2). At 7 days and 1 month of age similar trends were found but statistical significance was not always reached probably due to the limited number of studied cases. Interactions between age, genetic risk and bacterial proportions were not significant.

Discussion

This is the first follow-up study of the intestinal colonization process of the microbiota in children at family risk of developing CD. A cohort of first-degree relatives of CD patients was analyzed for their HLA class II DQA1 and DQB1 genotype. These two loci are transmitted in haplotypes and encode the alpha and beta-chain that forms the DQ heterodimer, a cell surface receptor of antigen presenting cells involved in gliadin peptide presentation in CD patients. In this study, three groups of different genetic risk were defined to establish a possible relationship between the HLA-DQ genotype and the gut microbial colonization process. This classification was based on the recent stratification of CD risk into five classes according to the HLA-DQ genotype in European population, which was

comparable to the Spanish one (Margaritte-Jeannin et al*.,* 2004 and Bourgey et al., 2007, Capilla et al., 2007, Donat et al., 2009). In this work, we have considered the need to change the initial five group classification into three main groups given our moderate sample size and the absence of patient's genotype. On the other hand, the DQ8 haplotype (DQA1*0301-

5   DQB1*0302) was not considered as a coeliac risk haplotype by itself in Spanish population, as occurs in other European countries. DQ8 haplotype is considered as a risk haplotype only in the case that it is transmitted in heterozygosis together with the DQA1*0501-DQB1*0201 or DQA1*0201-DQB1*0202 haplotypes as well as in homozygosis with itself, so newborns carrying this genotype are included in the second risk group.

10   Overall, the dynamics of the colonization process of the whole cohort of children showed similar patterns as those reported in previous studies for general population (Fanaro et al., 2003; Adlerberth et al., 2007; 2006). In this study, most bacterial populations tended to increase with age, particularly the strict anaerobic groups *Bacteroides-Prevotella* and sulphate-reducing bacteria and the facultative group *Staphylococcus*. In previous reports,

15   *Bacteroides* counts also increased as the colonization process progressed, while levels of facultative anaerobes dropped within the first week of life (Fanaro et al., 2003). Nevertheless, in this study *Staphylococcus* levels increased in the studied period as reported in other recent cohort studies on European and Swedish infants (Adlerberth et al., 2007; 2006). Human milk that is considered a continuous source of commensal bacteria to the infant gut, including

20   *Staphylococcus* species, could partly explain the abundance and origin of this bacterial group in the faeces of the studied babies under breast-feeding practices (Delgado et al., 2008). *Bifidobacterium* genus dominated the microbiota of the studied population over the entire period, which is characteristic of breast-fed babies, but its proportions did not increase with age in agreement with previous studies after 3 months of life (Fanaro et al., 2003).

The study of the effects of the HLA genotype on the colonization process revealed an association between higher proportions of *Bacteroides-Prevotella* group and the high genetic risk group, compared with the intermediate and low genetic risk groups. Total Gram-negative bacteria, *E. coli*, *Streptococcus-Lactococcus*, *E. rectale-C. coccoides, C. lituseburense* and *C. histolyticum* group proportions followed a similar trend when comparing the high versus the low genetic risk groups. In addition, higher levels of all these bacterial groups correlated with higher genetic risk. An increased number of affected coeliac relatives as well as the mother being the affected relative seemed to be associated with higher proportions of bacterial groups (*E. coli*, *E. rectale-C. coccoides*, *C. lituseburense* and *Streptococcus-Lactococcus*) related to increased CD risk in children. These results suggest that the features of the affected relative's microbiota could partially be transferred to the newborns with higher probability in both cases. Similar associations between increased *Streptococcus-Lactococcus, E. rectale- C. coccoides, C. lituseburense, C. histolyticum* and *Bacteroides-Prevotella* proportions and increased genetic risk of CD development were detected at 4 months. These preliminary results suggest that HLA class II molecules of the MHC could be a genetic host factor modulating the gut microbial colonization process and thereby disease risk. In murine models, it has been suggested that the MHC is one of the genetic factors exerting a pronounced influence on composition of the faecal microbiota using diverse mouse strains congenic for MHC (Toivanen et al., 2001). In human subjects, the only evidence of the influence of the genetic background in the structure of the intestinal microbiota is based on comparisons between twins and non-genetic related individuals, showing that subjects with a common genetic background had the highest similarities in their gut microbiota (Stewart et al., 2005; Palmer et al., 2007; Dicksved et al., 2008). MHC genes are responsible for encoding cell-surface antigen-presenting proteins that have a vital role in the immunological dialogue that occurs between T cells and other cells involved in autoimmune disorders. Many

autoimmune conditions have close genetic linkages to particular human HLA class II genes, such as CD, type-1 diabetes and rheumatoid arthritis. CD and rheumatoid arthritis have also been associated with imbalances in the intestinal microbiota particularly in newly diagnosed patients (Nadal et al., 2007; Vaahtovuo et al., 2008; Thompson-Chagoyán et al., 2007; Wen et al., 2008). MHC genes together with non-MHC autosomal genes have also been associated with susceptibility to infectious agents such as *Helicobacter felis*, *Listeria monocytogenes* and *Salmonella enterica* serovar *Typhimurium* (Plant and Glynn, 1976; Gervais et al, 1984; Mohammadi et al., 1996). In particular, HLA-DQ genes have been shown to influence the host response against *Helicobacter pylori* and the HLA-DQB1 genotype, which is also associated with CD and type-1 diabetes, was related to development of atrophic gastritis in *H. pylori*-infected patients (Sakai et al., 1999). In addition, higher incidence of infections, which can promote a Th1 protective immune response, has also been associated with increased risk of developing CD in early life (Ivarsson et al., 2003, Sanz et al., 2008).

Although the knowledge of the mechanisms by which MHC and particularly HLA-DQ molecules could influence the gut microbiota composition will require further investigations, these molecules could be involved in immune activation leading to tolerance or elimination of the microbe. Intestinal dendritic cells act as sentinels, sampling the mucosal surface for dietary or microbial antigens, thereby activating antigen presentation and specific T-cell responses that could favour tolerance to specific bacterial groups or facilitate their immune elimination modifying their relative proportions (Winkler et al., 2007).

Given that the major source of microbial stimuli in infancy is the gut microbiota, its distinctive characteristics in each infant may differently influence host physiology and immune responses, leading to either tolerance or a biased Th1 or Th2 cytokine response and thereby to disease (Hessle et al., 2000; Kalliomaki et al., 2001a; Adlerberth et al., 2007). The development of atopy has been shown to be preceded by deviations in gut microbiota,

characterized by lower prevalence of Gram-positive anaerobes and a higher ratio of clostridia to bifidobacteria at neonatal age (Penders et al., 2007; Kalliomaki et al., 2001a). In addition, studies in 1-year-old healthy Swedish and Estonian children, which represent two population groups with high and low incidence of CD and allergy, respectively, also showed increased

5   levels of *Clostridium* in Swedish group (Sepp et al., 1997). In addition, *Bacteroides* and *Clostridium* were those groups present in many Swedish infants, while not in Estonian infants (Sepp et al., 1997).

In summary, the obtained results suggest a relationship between the HLA-DQ genes and the colonization process of the gut microbiota at early life, which could play a role in the risk to

10   develop CD. Further work is in progress to confirm these preliminary observations that could lead to a significant change in the way autoimmune disorders and their therapies are investigated.

REFERENCES

Adleberth, I., Lindberg, E., Aberg, N., Hesselmar, B., Saalman, R., Strannegard, I.L., and Wold, A.E. (2006). Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? Pediatr. Res. 59, 96-101.

Adlerberth, I., Strachan, D.P., Matricardi, P., Ahrné, S., Orfei, L., Aberg, N., Perkin, M.R., Tripodi, S., Hesselmar, B., Saalman, R., Coates, A.R., Bonanno, C.L., Panetta, V., and Wold, A.E. (2007). Gut microbiota and development of atopic eczema in 3 european birth cohorts. J. Allergy Clin. Immunol. 120, 343-50.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R. and Stahl, D.A. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. 56, 1919-25.

Björkstén, B. (2006). The gut microbiota: a complex ecosystem. Clin. Experimental Allergy. 36, 1215-1217.

Bourgey, M., Calcagno, G., Tinto, N., Pennarelli, D., Margaritte-Jeannin, P., Greco, L., Limongelli, M.G., Esposito, O., Marano, C., Troncone, R., Spampanato, A., Clerget-Darpoux, F., and Sacchetti, L. (2007). HLA related genetic risk for coeliac disease. Gut. 56, 1054-59.

Branski, D., Fasano, A., and Troncone, R. (2006). Latest developments in the pathogenesis and treatment of celiac disease. J. Pediatr. 149, 295-300.

Bullock, N.R., Booth, J.C., and Gibson, G.R. (2004). Comparative composition of bacteria in the human intestinal microflora during remission and active ulcerative colitis. Curr. Issues. Intest. Microbiol. 5, 59-64.

Capilla, A., Donat, E., Planelles, D., Espinós, C., Ribes-Koninckx, C., and Palau, F. (2007). Genetic analyses of celiac disease in a Spanish population confirm association with CELIAC3 but not with CELIAC4.Tissue Antigens. 70, 324-9.

Collado, M.C., Donat, E., Ribes-Koninckx, C., Calabuig, M., and Sanz, Y. (2009). Specific duodenal and faecal bacterial groups are associated with paediatric celiac disease. J. Clin. Pathol. 62, 264-269.

Conte, M.P., Schippa, S., Zamboni, I., Penta, M., Chiarini, F., Seganti, L., Osborn, J., Falconieri, P., Borrelli, O., and Cucchiara, S. (2006). Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease. Gut. 55, 1760-1767.

Delgado, S., Arroyo, R., Martín, R., and Rodríguez, JM. (2008). PCR-DGGE assessment of the bacterial diversity of breast milk in women with lactational infectious mastitis. BMC Infect. Dis. 8, 51.

Dicksved, J., Halfvarson, J., Rosenquist, M., Järnerot, G., Tysk, C., Apajalahti, J., Engstrand, L., and Jansson, J.K. (2008). Molecular analysis of the gut microbiota of identical twins with Crohn's disease. ISME J. 2, 716-27.

Donat, E., Planelles, D., Capilla-Villanueva, A., Montoro, J.A., Palau, F., and Ribes-Koninckx, C. (2009). Allelic distribution and the effect of haplotype combination for HLA type II loci in the celiac disease population of the Valencian community (Spain). Tissue Antigens. 73, 255-61.

Fanaro, S., Chierici, R., Guerrini, P., and Vigi, V. (2003). Intestinal microflora in early infancy: composition and develoment. Acta. Paediatr. Suppl. 91, 48-55.

Fasano, A., and Catassi, C. (2005). Coeliac disease in children Best Practice & Res. Clin Gastroenterol. 19, 3: 467-78.

Franks, A.H., Harmsen, H.J., Raangs, G.C., Jansen, G.J., Schut, F., and Welling, G.W. (1998). Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl. Environ. Microbiol. 64, 3336-3345.

5   Gervais, F., Stevenson, M., and Skamene, E. (1984). Genetic control of resistance to *Listeria monocytogenes*: regulation of leukocyte inflammatory responses by the Hc locus. J. Immunol. 132, 2078-83.

Greco, L., Romino, R., Coto, I., Di Cosmo, N., Percopo, S., Maglio, M., Paparo, F., Gasperi, V., Limongelli, M.G., Cotichini, R., D'Agate, C., Tinto, N., Sacchetti, L., Tosi, R.,

10   and Stazi, M.A. (2002). The first large population based twin study of coeliac disease. Gut. 50, 624-8.

Harmsen, J.H.M., Elfferich, P., Schut, F., and Welling, G.W. (1999). A 16S rRNA Targeted probe for detection of lactobacilli and enterococci in faecal samples by fluorescent in situ hybridisation. Microbiol. Ecol. Health. Dis. 11, 3-12.

15   Hessle, C., Andersson, B., and Wold, A. (2000). Gram-positive bacteria are potent inducers of monocytic interleuchin-12 (IL-12) while gram-negative bacteria preferentially stimulate IL-10 production. Infection and Immunity. 68, 3581-3586.

Ivarsson, A., Hernell, O., Stenlund, H., and Persson, L.A. (2002). Breast-feeding protect against celiac disease. Am. J. Clin. Nutr. 75, 914-21.

20   Ivarsson, A., Hernell, O., Nyström, L., and Persson, LA. (2003). Children born in the summer have increased risk for coeliac disease. J. Epidemiol. Community. Health. 57, 36-39.

Kalliomäki, M., Kirjavainen, P., Eerola, E., Kero, P., Salminen, S., and Isolauri., E. (2001a). Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. J. Allergy Clin. Immunol. 107, 129-34.

Kalliomäki, M., Salminen, S., Arvilommi, H., Kero, P., Koskinen, P., and Isolauri, E. (2001b). Probiotics in primary prevention f atopic disease: a randomized placebo-controlled trial. Lancet. 357, 1076-79.

Kempf, V.A.J., Trebesius, K., and Autenrieth, I.B. (2000). Fluorescent in situ Hybridization allows rapid identification of microorganisms in blood cultures J. Clin. Microbiol. 38, 830-838.

Langendijk, P.S., Schut, F., Jansen, G.J., Raangs, G.C., Kamphuis, G.R., Wilkinson, M.H., and Welling, G.W. (1995). Quantitative fluorescence in situ hybridization of *Bifidobacterium* spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. Appl. Environ. Microbiol. 61, 3069-3075.

Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.H. (1996). Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. Microbiol. 142, 1097-1106.

Margaritte-Jeannin, P., Babron, M.C., Bourgey, M., Louka, A.S., Clot, F., Percopo, S., Coto, I., Hugot, J.P., Ascher, H., Sollid, L.M., Greco, L., and Clerget-Darpoux, F. (2004). HLA-DQ relative risks for coeliac disease in European populations: a study of the European genetics cluster on coeliac disease. Tissue Antigens. 63, 562-567.

Mearin, M.L., Ivarsson, A., and Dickey, W. (2005). Coeliac disease: is it time for mass screening? Best. Pract. Res. Clin. Gastroenterol. 19, 441-52.

Mohammadi, M., Redline, R., Nedrud, J., and Czinn, S. (1996). Role of the host in pathogenesis of Helicobacter-associated gastritis: H. felis infection of inbred and congenic mouse strains. Infect. Immun. 64, 238-45.

Nadal, I., Donant, E., Ribes-Koninckx, C., Calabuig, M., and Sanz, Y. (2007). Imbalance in the composition of the duodenal microbiota of children with coeliac disease. J. Med. Microbiol. 56, 1669-74.

Palmer, C., Bik, E.M., Digiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the Human Infant Intestinal Microbiota. PLoS Biol. 5, e177.

Penders, J., Thijs, C., van den Brandt, P.A., Kummeling, I., Snijders, B., Stelma, F., Adams, H., van Ree, R., and Stobberingh, E.E. (2007). Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study. Gut .56, 661-667.

Plant, J., and Glynn, A.A. (1976). Genetics of resistance to infection with Salmonella typhimurium in mice. J. Infect. Dis. 133, 72-8.

Poulsen, L.K., Lan, F., Kristensen, C.S, Hobolth, P., Molin, S., and Krogfelt, K.A. (1994). Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. Infect. Immun. 62, 5191-5194.

Rodrigo, L. (2006). Celiac disease World J. Gastroenterol. 12, 41: 6585-6593.

Sakai, T., Aoyama, N., Satonaka, K., Shigeta, S., Yoshida, H., Shinoda, Y., Shirasaka, D., Miyamoto, M., Nose, Y., and Kasuga, M. (1999). HLA-DQB1 locus and the development of atrophic gastritis with *Helicobacter pylori* infection. J. Gastroenterol. 34, S11, 24-7.

Salminen, S., and Isolauri, E. (2006). Intestinal colonization, Microbiota, and probiotics. J. Pediatr. 149, S115-S120.

Sánchez, E., Nadal, I., Donat, E., Ribes-Koninckx, C., Calabuig, M., and Sanz, Y. (2008). Reduced diversity and increased virulence-gene carriage in intestinal enterobacteria of coeliac children. BMC Gastroenterol. 8, 50.

Sanz, Y., Nadal, I., and Sánchez, E. (2007). Probiotics as drugs against human gastrointestinal infections. Recent Patents Anti-Infect. Drug Disc. 2, 148-56.

Sanz, Y., Sánchez, E., De Palma, G., Medina, M., Marcos, A., and Nova, E. (2008). Indigenous gut microbiota, probiotics, and coeliac disease. In: Child Nutrition & Physiology, L.T. Overton & M R. Ewente, ed. (Nova Science Publishers, Inc, NY, USA) pp. 210-224.

Sepp, E., Julge, K., Vasar, M., Naaber, P., Björksten, B., and Mikelsaar, M. (1997). Intestinal microflora of Estonian and Swedish infants. Acta Paediatr. 86, 956-61.

Sokol, H., Seksik, P., Rigottier-Gois, L., Lay, C., Lepage, P., Podglajen, I., Marteau, P. and Dore, J. (2006). Specificities of the fecal microbiota in inflammatory bowel disease. Inflamm. Bowel Dis. 12, 106-11.

Stene, L.C., Honeyman, M.C., Hoffenberg, E.J., Haas, J.E., Sokol, R.J., Emery, L., Taki, I., Norris, J.M., Erlich, H.A., Eisenbarth, G.S. and Rewers, M. (2006). Rotavirus infection frequency and risk of coeliac disease autoimmunity in early childhood: a longitudinal study. Am. J. Gastroenterol. 101, 2333-40.

Stewart, J.A., Chadwick, V.S., and Murray, A. (2005). Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. J. Med. Microbiol. 54, 1239-42.

Thompson-Chagoyán, O.C., Maldonado, J., and Gil, A. (2007). Colonization and impact of disease and other factors on intestinal microbiota. Dig. Dis. Sci. 52, 2069-77.

Toivanen, P., Vaahtovuo, J., and Eerola, E. (2001). Influence of major histocompatibility complex on bacterial composition of fecal flora. Infect. Immun. 69, 2372-7.

Vaahtovuo, J., Toivanen, P., and Eerola, E. (2001). Study of murine faecal microflora by cellular fatty acid analysis; effect of age and mouse strain. Antonie Van Leeuwenhoek. 80, 35-42.

Wallner, G., Amann, R., and Beisker, W. (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry. 14, 136-43.

Wen, L., Ley, R.E., Volchkov, P.Y., Stranges, P.B., Avanesyan, L., Stonebraker, A.C., Hu, C., Wong, F.S., Szot, G.L., Bluestone, J.A., Gordon, J.I., and Chervonsky, A.V.(2008). Innate immunity and intestinal microbiota in the development of Type 1 diabetes. Nature. 455, 1109-13.

Winkler, P., Ghadimi, D., Schrezenmeir, J., Kraehenbuhl, J.P. (2007). Molecular and cellular basis of microflora-host interactions. J. Nutr. 137, 756S-72S.

5

FIGURE LEGEND

Figure 1. Composition of the faecal microbiota of a cohort of children at family risk of CD development as a function of their HLA-DQ genotype, considering all data points. Median values of specific-group probe labelled bacteria in relation to total EUB-probe labelled bacteria (% Group-specific/EUB) and interquartile ranges (IQR). Correlations between bacterial group proportions and the genetic risk (High risk, 24-280% probability; Intermediate risk, 7-8% probability; and low risk <1% probability of developing CD) were established by the Spearman's test and significant differences were defined at $P < 0.10$ or 0.05.

Figure 2. Composition of the faecal microbiota of a cohort of children at family risk of CD development as a function of their HLA-DQ genotype at 4 months of age. Median values of specific-group probe labelled bacteria in relation to total EUB-probe labelled bacteria (% Group-specific/EUB) and interquartile ranges (IQR). Correlations between bacterial group proportions and the genetic risk (High risk, 24-28% probability; Intermediate risk, 7-8% probability; and low risk <1% probability of developing CD) were established by the Spearman's test and significant differences were defined at $P < 0.10$ or 0.05.

Table 1. Characteristics of the studied cohort of newborns

	Total subjects (n=20)
Delivery	
vaginal	17/20
caesarean	3/20
Size (cm)	49.28 ± 2.02
Weight (g)	3321.94 ± 491.26
Weeks of gestation	39.16 ± 1.52

Table 2. Oligonucleotide probes and hybridization conditions used in FISH analysis of faecal bacteria

Probe	Target bacterial group	Sequence (5'-3')	Hybridization conditions	References
Eub 338	All bacteria	GCTGCCTCCCGTAGGAGT	50°C	Wallner *et al*. (1993)
Non EUB 338	Negative control	ACTCCTACGGGAGGCAGC	50°C	Wallner *et al*. (1993)
Bif 164	*Bifidobacterium*	CATCCGGCATTACCACCC	50ºC	Langendijk *et al.* (1995)
Bac 303	*Bacteroides/Prevotella*	CCAATGTGGGGGACCTT	46ºC	Manz *et al*. (1996)
Erec 0482	*Eubacterium rectale/ Clostridium coccoides*	GCTTCTTAGTCARGTACCG	50ºC	Franks *et al*. (1998)
E.coli 1513	*Escherichia coli*	CACCGTAGTGCCTCGTCATCA	50ºC	Poulsen *et al*. (1994)
Strc 493	*Streptococcus/Lactococcus*	GTTAGCCGTCCCTTTCTGG	50°C	Franks *et al*. (1998)
Lab 158	*Lactobacillus/*other lactic acid bacteria	GGTATTAGCA(T/C)CTGTTTCCA	45°C	Harmsen *et al.* (1999)
Sta	*Staphylococcus*	TCCTCCATATCTCTGCGC	50°C	Kempf *et al*.(2000)
C. His 150	*Clostridium histolyticum*	TTATGCGGTATTAATCT(C/T)CCTTT	50°C	Franks *et al*. (1998)
C. Lit 135	*Clostridium lituseburense*	GTTATCCGTGTGTACAGGG	50°C	Franks *et al*. (1998)
SRB	Sulphate reducing bacteria	TACGGATTTCACTCCT	50°C	Bullock et al. (2004)

Table 3. Microbial colonization process in children at family risk of CD development during the first 4 months of life

Bacterial group	7 days (n=20)	1 month (n=20)	4 months (n=20)	Mann –Whitney U-test P value		
	Group-specific bacteria/EUB(%) (IQR)			7 d-1 month	7d-4 months	1m-4 months
Lactobacillus group	2.09 (0.80-2.97)	1.77 (1.20-5.34)	2.81 (1.51-8.21)	0.513	0.095	0.231
Bifidobacterium	9.20 (6.22-12.20)	10.57 (6.20-25.00)	12.21 (4.17-15.43)	0.234	0.301	0.589
Streptococcus-Lactococcus	2.85 (1.64-6.01)	2.91 (1.54-7.15)	3.07 (2.02-5.04)	0.925	0.841	0.758
Staphylococcus	1.99 (1.16-6.57)	2.47 (1.24-7.00)	4.71 (2.83-7.60)	0.678	0.086	0.086
E. rectale-C. coccoides	1.95 (1.30-4.68)	2.30 (1.20-5.38)	2.22 (1.20-6.63)	0.911	0.813	0.989
E. coli	2.37 (1.36-4.63)	2.25 (1.43-3.49)	3.13 (1.42-5.72)	0.707	0.515	0.380
Bacteroides-Prevotella	2.58 (1.36-5.33)	6.15 (3.45-8.58)	4.61 (2.32-8.81)	0.017*	0.125	0.407
C. histolyticum	2.85 (1.44-6.15)	2.47 (1.47-5.43)	2.76 (1.80-5.36)	0.925	0.718	0.620
C. lituseburense	2.14 (1.05-5.18)	1.97 (1.34-4.98)	1.62 (1.13-4.76)	0.529	0.947	0.512
Sulphate-reducing bacteria	1.12 (0.76-3.99)	1.47 (1.06-2.93)	3.62 (1.38-10.11)	0.461	0.009*	0.021*
‡Sum Gram positive	24.72 (18.54-48.17)	37.57 (17.83-62.65)	32.22 (20.82-66.51)	0.547	0.327	0.718
‡Sum Gram negative	9.04 (5.16-13.46)	11.37 (7.06-17.92)	12.37 (7.33-32.38)	0.114	0.081	0.417
‡Sum Total	38.12 (24.24-55.96)	51.27 (22.64-75.63)	42.10 (28.77-53.90)	0.383	0.377	0.874

‡Total bacteria, Gram-negative bacteria and Gram-positive bacteria were calculated by adding the relative proportions of the corresponding groups related to total bacteria (%specific group/EUB), excluding overlapping probes.
**Significant differences (P<0.050) between mean values of different group proportions by using the Wilcoxon W-test