HLA-related genetic risk for coeliac disease.

Mathieu Bourgey, Giuseppe Calcagno, Nada Tinto, Daniela Gennarelli, Patricia Margaritte-Jeannin, Luigi Greco, Maria Giovanna Limongelli, Oscar Esposito, Caterina Marano, Riccardo Troncone, et al.

HLA-RELATED GENETIC RISK FOR COELIAC DISEASE

Correspondence to:
Prof. Luigi Greco
Department of Pediatrics, University of Naples "Federico II"
Via Pansini 5,
80131 Naples, Italy
Tel. +39.081.7463275
Fax. +39.081.5469811
E-mail address: ydongre@unina.it

Mathieu Bourgey, Giuseppe Calcagno, Nadia Tinto, Daniela Gennarelli, Patricia Margaritte-Jeannin, Luigi Greco, Maria Giovanna Limongelli, Oscar Esposito, Caterina Marano, Riccardo Troncone, Antonella Spampanato, Françoise Clerget-Darpoux, Lucia Sacchetti.

Authors' affiliations:

Luigi Greco, Maria Giovanna Limongelli , Oscar Esposito, Caterina Marano, Riccardo Troncone:
Department of Pediatrics *and European Laboratory for the Investigation of Food-Induced Diseases*, University of Naples "Federico II", Naples, Italy.

Giuseppe Calcagno:
SPeS Dipartiment, University of Molise Campobasso, Italy.

Lucia Sacchetti, Giuseppe Calcagno, Nadia Tinto, Daniela Gennarelli, Antonella Spampanato:
Department of Biochemistry and Medical Biotechnology, University of Naples "Federico II" and CEINGE advanced biotechnology, Naples, Italy.

Mathieu Bourgey, Patricia Margaritte-Jeannin, Francoise Clerget-Darpoux :
INSERM U535, Villejuif, F-94817 France
Université Paris-Sud, IFR 69, UMR-S535Villejuif, F-94817 France

Key words: Coeliac Disease, HLA, first degree relatives, recurrence risk, *family screening, DQ-typing*

Word count: 3537

ABSTRACT

Background: Several studies have shown a higher prevalence of Coeliac Disease in sibs of coeliac patients (risk 8-12%).
Aim and method: Aim of the study is to evaluate the risk for a sib of a patient to develop CD. A cohort of 188 Italian families was ascertained through a symptomatic CD patient. In addition, an Italian trio data set (127 patients and their 2 parents) also genotyped for HLA-DQ was available for this study. All members were diagnosed for CD status and their HLA-DQ genotyping was performed.
Results: The overall risk for a sib of a CD patient to develop the disease is estimated at 10% in this sample. When taking into account the HLA-DQ information of the patient, of parents and of the sib, the estimate of the risk varies from 1 to 29%. We found that 40% of sibs of a coeliac proband have a negligible risk (lower than 1%), 30% have a risk lower than 10 % and greater than 1%; finally we have 1/3rd of the families in which the risk is high or very high (above 25%).
Conclusion: These results allow a more accurate information to be given to parents having an affected child about the actual risk for another child. It is now possible, before the birth of a child, to provide to the parents an estimate of his order risk for coeliac disease. Finally it will be possible to propose a particular follow-up for a baby at high risk.

INTRODUCTION

Coeliac Disease (CD) is an immune-mediated enteropathy caused by a permanent sensitivity to gluten in genetically susceptible individual. [1] CD has a strong genetic association with human leukocyte antigen (HLA). Most CD patients (90-95%) express the heterodimer DQA1*05/DQB1*02. [2, 3] Among the DQA1*05/DQB1*02 heterodimer carriers, the risk of disease is greater in individuals homozygous for DQB1*02. [4, 5]

Using very large samples of patients from different European countries, Margaritte-Jeannin et al [6] recently showed that the genetic risk for an individual to develop a symptomatic form can be stratified in five classes according his HLA-DQ genotype. DQ2 carriers can be divided in 3 groups : with two copies of DQB1*02 (group G1), with one copy of DQB1*02 acting in *trans* with DQA1*05 (group G2) or with one copy of DQB1*02 acting in *cis* with DQA1*05 (group G3). DQ2 non-carriers are divided in 2 groups: having two copies of a DQB1*02, DQ8 or one copy of each (group G4) or with other DQ genotypes (group G5). In all populations, G1 is the highest risk group whereas the relative risks for the other genotypes vary from one population to another. In the Italian population the relative risks are 0.68, 0.23, 0.10 and 0.02 for individuals belonging to G2, G3, G4 and G5 respectively.

Several studies have shown a higher prevalence of CD in sibs of coeliac patients compared to the general population, with estimates of risk falling between 8 and 12%. [7-13] However, the meaning of these estimates are unclear since the CD phenotype is most often not precisely defined and may or may not include different forms of the disease. We can distinguish three different forms of coeliac disease according to the ESPGHAN criteria [14-16]: the latent (or potential) form defined only by the presence of specific antibodies; the silent form defined by the presence of specific antibodies and the observation of a villous atrophy of the small intestine; and the symptomatic form defined by the presence of specific antibodies, the observation of a villous atrophy and the observation of clinical symptoms.

The aim of the present study is to evaluate in the Italian population, the risk for a sib of a symptomatic patient to develop any of the 3 forms of CD and to provide to the parents having a CD affected child the risk for a future baby with the best possible precision. We estimate this risk according to familial and genetic information. Finally, we will be able in some situations to predict the risk for a future baby before his birth and in other situation it will be necessary to precise it by genotyping the baby after his birth. Thus, a special attention to wean and to follow the sibs with higher risks will be made as soon as possible.

MATERIALS AND METHODS

Family samples

A cohort of 188 nuclear families was ascertained through a symptomatical CD patient (the proband) having at least one sib. For the vast majority (184/188) of probands we had families with both parents available for typing. All probands were confirmed symptomatic coeliac, diagnosed according to the revised ESPGHAN protocol, with positive AntiEndomisum and or Anti Transglutaminase antibodies and above Marsh 1 lesions at the small intestinal biopsy. [17] A total of 798 individuals were sampled: 614 first degree relatives of 184 coeliac probands. The cohort is

composed of 246 sibs and 368 parents (184 fathers, 184 mothers). They were enrolled in a follow up program of 3 years form January 2001 to December 2003 to check their coeliac status.

Two fasting blood samples were collected in the morning from all 798 individuals. Written informed consent was obtained from each participant. The research was approved by the Ethic Committee of the School of the Medicine, University of Naples "Federico II" and was in accordance with the principles of Helsinki II declaration.

They were all screened for Antiendomisial and anti-Transglutaminase antibodies, and Class II HLA (DQ-DR) status molecular staging was performed.

After the first stage of analysis , in order to have the HLA-DQ distribution from a larger sample of proband patients and increase the robustness of the estimate, we added in the finalstage to the sample an Italian trio data set (127 probands and their 2 parents) which was part of the European samples studied by Margaritte-Jeannin et al [6] .

Disease status

Serological test

Serum samples obtained after centrifugation were used for antibodies evaluation. AntiEndomisum IgA antibodies (EMA) were detected by indirect immunofluorescence microscopy on section of *Rhesus* monkey esophagus (Eurospital Trieste Italia).

Anti tissue Transglutaminase IgA antibodies (tTG) were analyzed by enzyme linked immunosorbent assay (ELISA) using human recombinant tTG as antigen (DIA Medix Corp.-Ivax Diagnostics, Florida USA). Total serums IgA were evaluated by nephelometric assay (BN ProSpec System Behring, Marburg Germany).

Small intestinal biopsy

EMA and tTG positive individuals underwent a small intestinal biopsy, 18 by a Watson capsule and 9 by an endoscopy. The mucosal injury was graded according to Marsh [17] : T0: normal mucosal; T1: more than 40% increased intraepithelial lymphocytes infiltration; T2: T1 with crypt hyperplasia; T3a: T2 with partial villous atrophy; T3b: T2 with subtotal villous atrophy; T3c: T2 with total villous atrophy.

Genotyping

DNA extraction

Genomic DNA was extracted from the EDTA + blood sample by a commercially available Kit (Nucleon BACC 2, Amersham Biosciences Europe, Milan, Italy).

HLA DQ typing

The HLA-DQ typing for coeliac susceptibility was performed by a three step-procedure. The first step employs a single SSP-PCR home-based method previously described to detect the HLA heterodimer DQA1*0501-DQB1*0201 presence. [18] The second step uses an SSO-PCR based method to type the HLA DQB1* locus (Dynal Biotech LTD UK) to confirm the presence/absence of the DQB1*02 allele and/or to verify the presence of the other DQB1*03 risk allele. The possible DQB1*03 detection by SSO is followed by a SSP-PCR technique (Dynal Biotech LTD UK) aimed to resolve the DQB1*03 locus evidencing the DQB1*0302 allele presence/absence.

HLA DR typing

The HLA-DRB1 typing was performed by SSO-PCR based method (Dynal Biotech LTD UK) to determine the phased DR-DQ genotypes of all individuals.

Statistical analysis

Notation

Consistently with Margaritte-Jeannin et al, [6] let consider 5 DQA1-DQB1 haplotypes and 5 DQ genotypic groups. The 5 haplotypes are:

H1: DQA1*05-DQB1*02 denoted: $\alpha_5\beta_2$

H2: None(DQA1*05)-DQB1*02 denoted: $\overline{\alpha_5}\beta_2$

H3: DQA1*05-none(DQB1*02) denoted: $\alpha_5\overline{\beta_2}$

H4: DQA1*301-DQB1*302 denoted: $\alpha_3\beta_3$

H5: other haplotypes denoted: $\alpha\beta$

Among the 5 genotypic groups, 3 correspond to the DQ2 heterodimer carriers:

Group 1 (G1): H1/H1 and H1/H2 (With double dose of β_2)

Group 2 (G2): H2/H3 (heterodimer encode in *trans*)

Group 3 (G3): H1/H3, H1/H4 and H1/H5 (heterodimer encode in *cis*)

The 2 other genotypic groups are:

Group 4 (G4): H2/H2, H2/H4 and H4/H4

Group 5 (G5): other genotypes

Statistical methods

- **Estimation of the empirical risk for a sib of a proband**

In a family with an affected child (denoted hereafter proband), the risk R for another child to be affected may be estimated by the ratio of the number of affected sibs over the total number of sibs.

- **Estimation of control DQA1-DQB1 haplotype frequencies (Hi)**

Since families of the two samples have been ascertained through a single affected, the parental haplotypes untransmitted to the affected child are representative of the general population (Affected Family Based Controls) and provide unbiased estimates of H_i frequencies. [19] These frequencies will be estimated from 622 untransmitted parental haplotypes (311 families).

- **Estimation of the genotypic group risks**

Given the H_i frequencies estimated in the population, we may calculate the risk to be affected (F_{ij}) for an individual with genotype H_iH_j.

$$Fij = P(Aff / H_iH_j) = \frac{P(H_iH_j / Aff) * P(Aff)}{P(H_iH_j)}$$

Where $P(H_iH_j)$ is the probability for an individual in the population to have the genotype H_iH_j. The computation of this probability is based on control haplotype frequencies assuming Hardy Weinberg proportions. P(Aff) is estimated by the probability to be affected in the population, *i.e* the disease prevalence. $P(H_iH_j/Aff)$ is estimated by the proportion of each genotype H_iH_j in the 311 affected probands.

- **Estimation of the familial correlation not due to HLA**

Given R, H_i frequencies and F_{ij}, it is possible to estimate λ the familial correlation not due to HLA assuming that λ is a multiplicative factor independent of the HLA genotypes H_iH_j (see appendix).

- **Estimation of risks for a sib of a proband**

Assuming no recombination in the DR-DQ haplotypes and based on the Mendelian law of segregation we can compute the risk R_{ij} for a sib to be affected according to the known genotype H_iH_j of the affected proband. If parents are HLA typed, we may have a more precise estimate risk R_{ikjl} for a sib to be affected according to the parents genotypes. In some situations in which the parent typing gives a large range of risk for their future baby, it may be proposed to type the baby after the birth to refine his risk according to his own genotype. Thus, we refine the estimate in computing the risk K_{ij} for a sib to be affected according to his own genotype H_iH_j (see appendix). Based on these estimates of risk and on the probability that parents of a proband have a given genotype, it is possible to provide the expected distribution of the genetic groups in sibs.

RESULTS

- **Estimation of the risk for a sib of a proband**

Table 1 gives the distribution of the three forms of CD among the additional cases. The recurrence risks (R), which correspond to the risk to develop a CD (latent + silent + symptomatic) for a sib of proband is estimated by:

R= 24/246 = 9.8% [6.1 ; 13.4]

This estimation is consistent with the value of ~10% provided by the literature. [7-13]

Table 1 Distribution of the CD forms in first degree relatives

	Parents (n=368)	Sibs (n=246)
Latent : no symptoms MI-MII	1	5
Silent : no symptoms , MIII	8	12
Symptomatic MII-MIII	8	7
Total	17	24

- **Control DQA1-DQB1 haplotype frequencies estimation**

The frequency of each control DQA1-DQB1 haplotype is presented in the table 2. For each haplotype the corresponding DR allele is reported.

Table 2 Control haplotype frequencies

Notation	DQA1-DQB1 haplotype	DR allele	Frequency
H1	DQA1*05 - DQB1*0201	DR3	0.11
H2	DQA1*02 - DQB1*0202	DR7	0.12
H3	DQA1*05 - DQB1*03	DR5	0.30
H4	DQA1*0301 – DQB1*0302	DR4	0.04
H5	Others	DRX	0.43

- **Distribution of H_iH_j in probands**

Table 3 gives the observed number and frequencies of the genotype H_iH_j for the 311 probands and the expected frequencies in the general population.

Table 3 Number and frequencies of H_iH_j in 311 probands and the expected frequencies in the general population

DQ genotypes	Corresponding DR genotypes	Genotypic group	Probands	Expected in control population
H1/H1	DR3/DR3	G1	74 (24%)	4%
H1/H2	DR3/DR7			
H2/H3	DR5/DR7	G2	117 (38%)	7%
H1/H3	DR3/DR5	G3	79 (25%)	17%
H1/H4	DR3/DR4			
H1/H5	DR3/DRX*			
H2/H2	DR7/DR7	G4	13 (4%)	3%
H2/H4	DR7/DR4			
H4/H4	DR4/DR4			
Others	Others	G5	28 (9%)	69%

*X = different from 3, 4, 5, 7

- **Risk for an individual to develop the disease according to his genotypic group**

The risk for an individual with genotype H_iH_j in the Italian population is given in Table 4. For each group the corresponding haplotype combinations and the corresponding DR genotypes (based on typing) are also shown.

Table 4: Genotypic group risks in the population

DQ genotypes	Corresponding DR genotypes	Genotypic group	F_{ij}
H1/H1	DR3/DR3	G1	0.21
H1/H2	DR3/DR7		
H2/H3	DR5/DR7	G2	0.17
H1/H3	DR3/DR5	G3	0.06
H1/H4	DR3/DR4		
H1/H5	DR3/DRX*		
H2/H2	DR7/DR7	G4	0.05
H2/H4	DR7/DR4		
H4/H4	DR4/DR4		
Others	Others	G5	0.006

*X = different from 3, 4, 5, 7

- **The familial correlation λ not due to the transmission of the HLA haplotypes Hi**

Given R (0.098), H_i frequencies and F_{ij} we can compute λ, the familial correlation not due to the transmission of Hi haplotypes:

$\lambda = 1.4$

- **Risk for a sib of a proband according the DQ genotype of the proband**

Table 5 gives for each possible DQ genotypes of a proband the risk to be affected for a sib of this proband. Results are classed by genotypic group.

Table 5 Risks for a sib of a proband according the DQ genotype of the proband

proband group	G1	G1	G2	G3	G3	G3	G4	G4	G4	G5	G5	G5	G5	G5	G5
proband genotype	H_1H_1	H_1H_2	H_2H_3	H_1H_3	H_1H_4	H_1H_5	H_2H_2	H_2H_4	H_4H_4	H_2H_5	H_3H_3	H_3H_4	H_3H_5	H_4H_5	H_5H_5
R_{ij}	0.14	0.14	0.11	0.07	0.07	0.06	0.09	0.06	0.04	0.04	0.03	0.03	0.03	0.02	0.02

Blue is for $R_{ij} < 0.05$, green for $0.5 \leq R_{ij} \leq 0.1$ and red for $R_{ij} > 0.1$

- **Risk for a sib of a proband according the DQ genotype of parents**

As shown in table 5, according to the HLA DQ of the proband, the risk for a future baby is quite different. HLA typing may also be proposed to the parents if they want to be better informed. Figure 1 gives what may be concluded with the parental typing. In colored situations, the risk for the baby may be negligible (blue), moderate (green) or high (red). It represents 30% of cases. In contrast, there are situations (uncolored boxes) in which parental typing may lead to very variable risk for the baby (from negligible to high). In such a case, typing the baby will be encouraged.

For example, a mother H2H2 (DR7/DR7) who already has a child with CD has a 29% risk to have another one if the father is H1H1 (DR3/DR3), but only a 7% of risk if the father is H2H2 (DR7/DR7).

Similarly a father H2H5 (DR7/DRX) who already has a child with CD has a very small (2%) risk to have another CD child if the mother is H2H5 (DR7/DRX), but a higher (12%) risk if the mother is H3H3 (DR5/DR5).

Figure 1 also show that the decision of genotyping a new born baby depends more on the possible variability of the risk for the child than on the mean expectation obtained from the parent genotypes. For example, if parents are H1H1/H1H1, H2H4/H2H4 or H3H5/H4H5 there is no doubt on the risk for the child and the child genotyping is unnecessary. If parents are H2H4/H1H3 a new child has a mean risk of 15% but his risk may vary from 1 to 29%. In this situation the child genotyping is necessary if we want to refine the risk.

- **Risk for a sib of a proband according to his own DQ genotype**

Figure 2 gives the probability (in italics) for a sib of a proband to belong to Gi and the corresponding risk (inside the bar), showing that it is expected that approximately 40% of sibs will belong to G5 and consequently will have a negligible risk (lower than 1%). However about 30% will belong to G1 or G2 and will be predicted a risk higher than 20%.

CONCLUSION AND DISCUSSION

Genetic counseling in families with a proband affected by a multifactorial disease is generally inappropriate, due to the large uncertainty around the empirical risk of recurrence. In coeliac disease, prenatal genetic counseling is neither required nor suggested, but those who work with coeliac families are often asked about the recurrence risk for a possible future child. We answer that the risk of recurrence is approximately 10%, but do not give any more accurate information. For many parents a 10% risk of recurrence is intolerably high, and they feel discouraged to plan further pregnancies.

In this study we show that a sib of a coeliac proband has an average recurrence risk of 10%, but this average has to be broken down according to the HLA DQ information on the proband. According to the HLA DQ of the proband the risk estimate for the sib ranges from 2% to 14%. It is possible to provide better information to parents by performing their own HLA typing. In 30% of cases, the HLA parental typing will be sufficient to give an accurate estimate of their baby's risk. In other cases, for example when parents are H2H4 and H1H3, it may be advisable to genotype, after his birth, the baby in order to precise his risk. Broadly, it is expected that approximately 40% of sibs of a coeliac proband will have a negligible risk (lower than 1% to develop one form of the disease). Consequently about 40% of the families of a coeliac proband will receive a very reassuring message by this procedure. Moreover, 30% of sibs are expected to have a risk lower than 10% and greater than 1%, so it is possible also with them to share a positive attitude about their recurrence risk.

We will be left with about 1/3rd of the families in which the risk of recurrence is high or very high (above 20%). In these cases we do propose to share the information with the family in order to set up a plan to deal with this risk:

1. Breast feeding should be strongly supported, [20] although it is well know that it does not prevent the diseases, but it affects only the phenotype, by delaying the onset of symptoms. [21]

2. Gluten containing foods should be introduced at weaning according to the usual practices adopted for infants from unaffected families, since there is no evidence that time of gluten introduction may ever affect the incidence of the disease. It is also desirable to umask the disease as soon as possible and not delaying to later ages, when risk of complications increase (Ventura e Greco Gastroenterology). It has been suggested that there is a 'window' of time of gluten introduction (4 to 6 months) when the child is 'protected' by developing the disease, but again, this is likely to affect only the phenotype of this heavily genetic determined disease. [22] Twin studies have indeed recently suggested that unshared environment has little or no effect in the onset of coeliac disease. [23] Genetic factors do play the major role. Gluten avoidance, which is environmental, may stop completely the pathogenesis of the disease, but this is not the real life. Who, in our communities, will indeed accept to go, for life, on a gluten free diet because of the genetic risk of developing the disease at the present estimates (never higher than 30%) ?

3. The easy availability of anti-transglutaminase antibodies test, also at point-of-care, does give the possibility to suspect the onset of the disease, by measuring the antibodies titer, much in advance of clinical symptoms. [24] A secondary prevention may be put in place for subjects with a significant risk estimate: the vast majority of infants who will eventually develop the disease might be diagnosed before the disease become clinically manifest.

In conclusion for these infants at high risk of recurrence in coeliac families, the doctors should encourage breast feeding, ordinary weaning with gluten and an accurate follow up after

gluten introduction. A great amount of suffering, anxieties and health care resources consumption would really be prevented.

Over and above the intention of the authors, ethical issue will arise from the interpretation of the genetic risk tables. To reduce this risk we avoided to assign to each cross of Figure 1 an exact point estimate, but people in the red area may feel the burden of the increased risk estimate.

We are not discovering new markers: we are just improving the rough risk estimate which is actually in the daily use. The majority of families may feel reassured by this information. But a few in the red area might feel depressed. We do not have jet an exact estimate of the effects of this information in the families: we have just started a four years European multicentric prospective study to gain this knowledge. Our anectodical professional experience suggests that family do prefere to have the best estimate of their own risk of recurrence. We do hope that not one child birth is prevented but an enhancement of childbirths is promoted.

Giving this solid base of risk estimation provided by HLA genotyping, it will be possible to improve the risk estimate by adding other genetic information. Currently, other susceptibility genes for CD have been mapped either by linkage or by association studies, [25-28] but this information cannot be used in the risk estimation until susceptibility variants are clearly identified. Our hope is that some of the predisposing variants will be shortly identified: a multivariate combination of these information will then really improve the risk estimate, although it is possible that the information will not immediately provide better estimates because of the complexity of the genetic contribution of each locus.

ACKNOWLEDGEMENTS

We would like to thank the patients and their families.
This work was supported by:

- FRM (Fondation de la Recherche Médicale) for the doctoral funding of M. Bourgey;
- ELFID (European Laboratory for the Investigation of Food-Induced Diseases);
- CEINGE, Regione Campania Convenzione- Del G. R. 27/12/2002 N. 6276;
- Regione Campania (progetto finalizzato Ricerca Sanitaria, D.G 10 del 21/01/05, sistema B linea di ricerca 6);
- MIUR (Italian Ministero dell'Istruzione, dell'Università e della Ricerca).

Competing interests: none declared.

REFERENCES

1 Green PH, Jabri B. Coeliac disease. *Lancet* 2003;**362**:383-91.

2 Sollid LM, Markussen G, EK J, *et al*. Evidence for a primary association of coeliac disease to a particular HLA-DQ α/β heterodimer. *J Exp Med* 1989;**169**:345-350.

3 Sollid LM, Thorsby E. The primary association of coeliac disease to a given HLA-DQ α/β heterodimer explains the divergent HLA-DR association observed in various Caucasian population. *Tissue Antigens* 1990;**36**:136-7.

4 Plosky R, Ek J, Thorsby E, Sollid LM. On the HLA-DQ associated susceptibility in Coeliac Disease: a possible gene dosage effect of DQB1*0201. *Tissue Antigens* 1993;**41**:173-7.

5 Clerget-Darpoux F, Bouguerra F, Kastally R, *et al*. High risk genotypes for Coeliac Disease. *C R Acad Sci Paris III* 1994;**317**:931-6.

6 Margaritte-Jeannin P, Babron MC, Bourgey M, *et al*. HLA-DQ relatives risk for coeliac disease in European population: a study of the European Genetic cluster on Coeliac Disease. *Tissue Antigen* 2004;**63**:562-7.

7 Houlston RS, Ford D. Genetics of Coeliac Disease. *Q J Med* 1996;**89**:737-43.

8 Bonamico M, Mariani P, Mazzilli MC, *et al*. Frequency and clinical pattern of coeliac disease among siblings of coeliac children. *J Pediatr Gastroenterol Nutr* 1996;**23**:159-63.

9 Farrè C, Humbert P, Vilar P, *et al*. Serological markers and HLA-DQ2 haplotype among first-degree relatives of coeliac patient. *Dig Dis Sci* 1999;**44**:2344-9.

10 Korponay-Szabo I, Kovacs J, Lorincz M, *et al*. Families with multiple cases of gluten-sensitive enteropathy. *Z Gastroenterol* 1998;**36**:553-8.

11 Trier JS. Coeliac Sprue. *N Engl J Med* 1991;**325**(24):1709-19.

12 Hogberg L, Falth-Magnusson K, Grodzinsky E, Stenhammar L. Familial prevalence of Coeliac Disease: a twenty year follow-up study. *Scand J Gastroenterol* 2003;**38**:61-5.

13 Pittshieler K, Gentili L, Niederhofer H. Onset of coeliac disease: a prospective longitudinal study. *Acta Paediatr* 2003;**92**:1149-52

14 Walker-Smith JA, Guandalini S, Schmitz J, *et al*. Revised criteria for diagnosis of Coeliac Disease. Report of the Working Group of the European Society of Paediatric Gastroenterology and Nutricion. *Arch Dis Child* 1990;**65**:909-11.

15 Paparo F, Petrone E, Tosco A, *et al*. Clinical, HLA, and small bowel immunohistochemical features of children with positive serum antiendomysium antibodies and architecturally normal small intestinal mucosa. *Am J Gastroenterol* 2005;**100**:2294-98.

16 Fasano A. Clinical presentation of coeliac disease in the pediatric population. *Gastroenterology* 2005;**128**:S68-S73.

17 Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for standardized report scheme for pathologists. *Eur J Gastroenterol Hepatol* 1999;**11**:1185-94.

18 Sacchetti L, Tinto N. Calcagno G, *et al*. Multiplex PCR typing of the three most frequent HLA alleles in coeliac disease. *Clin Chim Acta* 2001;**310**(2):205-7.

19 Thomson G. Mapping disease genes: family Based Association studies. *Am J Hum Genet* 1995; **57**:487-498.

20 Ivarsson A, Hernell O, Stenlund H, Persson A. Breast-feeding protects against coeliac disease. *Am J Clin Nutr* 2002;**75**: 914-21.

21 Greco L, Auricchio S, Mayer M, Grimaldi M. Case control study on nutritional risk factors in coeliac disease. *J Pediat Gastroenterol Nutr* 1988;7:395-399.

22 Norris JM, Barriga K, Hoffenberg EJ, *et al*. Risk of coeliac disease autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. *JAMA* 2005;293:2310-2.

23 Nistico L, Fagnani C, Coto I, *et al*. Concordance, disease progression, and heritability of coeliac disease in italian twins. *Gut* 2006 Jun;55(6):803-8.

24 Korponay-Szabo I R, Raivio T, Laurila K, *et al*. Coeliac disease case finding and diet monitoring by point-of-care testing. *Aliment Pharmacol Ther* 2005;**22**(8):729-737.

25 Greco L, Babron MC, Corazza GR, *et al*. Existence of a genetic risk factor on chromosome 5q in Italian coeliac disease families. *Ann Hum Genet* 2001 Jan; **65**(Pt 1):35-41.

26 Amundsen SS, Monsuur AJ, Wapenaar MC, *et al*. Association analysis of MYO9B gene polymorphisms with coeliac disease in a Swedish/Norwegian cohort. *Hum Immunol* 2006 Apr-May; **67**(4-5):341-5.

27 Naluai AT, Nilsson S, Samuelsson L, *et al*. The CTLA-4/CD28 gene region on chromosome 2q/33 confers susceptibility to coeliac disease in a way possibly distinct from that of type 1 diabetes and other chronic infiammatory disorders. *Tissue Antigens* 2000;**56**:350-355.

28 King AL, Moodie SJ, Fraser JS, *et al*. CTLA-4/CD28 gene region is associated with genetic susceptibility to coeliac disease in UK families. *J Med. Genet* 2002;**39**:51-54.

APPENDIX

In this section, we detail formulae used in the method section.

We use the following notations:

S_{aff}: to be an affected sib

P_{aff}: the proband is affected

P_{ij}: the proband has a genotype H_iH_j

C_{ikjl}: parents have genotypes H_iH_k and H_jH_l

Aff: to be an affected individual

λ: familial correlation not due to HLA assuming that it is a multiplicative factor independent of the HLA DQ genotypes

R is the probability to have an affected sib of an individual P conditionally that P is affected (P is a proband):

$$R = P(S_{Aff} / P_{Aff}) = \frac{P(S_{Aff}, P_{Aff})}{P(P_{Aff})} = \frac{\sum_{i=1}^{5}\sum_{j=1}^{5}\left[P(S_{Aff}, P_{Aff}, P_{ij})\right]}{P(P_{Aff})}$$

$$R = \frac{\sum_{i=1}^{5}\sum_{j=1}^{5}\left[\sum_{k=1}^{5}\sum_{l=1}^{5}\left[\lambda * P(Aff, C_{ikjl}, P_{Aff}, P_{ij})\right]\right]}{P(P_{Aff})}$$

$$R = \frac{P(P_{Aff}) * \sum_{i=1}^{5}\sum_{j=1}^{5}\left[P(P_{ij} / P_{Aff}) * \lambda * \sum_{k=1}^{5}\sum_{l=1}^{5}\left[P(C_{ikjl} / P_{ij}) * P(Aff / C_{ikjl})\right]\right]}{P(P_{Aff})}$$

$$R = \lambda * \sum_{i=1}^{5}\sum_{j=1}^{5}\left[\frac{F_{ij} * P(H_iH_j)}{P(Aff)} * \sum_{k=1}^{5}\sum_{l=1}^{5}\left[P(H_k) * P(H_l) * \frac{1}{4} * (F_{ij} + F_{il} + F_{kj} + F_{kl})\right]\right]$$

As we can estimate R, F_{ij} and the allelic frequencies of H_i haplotypes on the sample, it is possible to estimate λ the familial correlation not due to HLA haplotypes:

$$\lambda = \frac{R}{\sum_{i=1}^{5}\sum_{j=1}^{5}\left[\frac{F_{ij} * P(H_iH_j)}{P(Aff)} * \sum_{k=1}^{5}\sum_{l=1}^{5}\left[P(H_k) * P(H_l) * \frac{1}{4} * (F_{ij} + F_{il} + F_{kj} + F_{kl})\right]\right]}$$

With:

$$P(Aff) = \sum_{i=1}^{5}\sum_{j=1}^{5} F_{ij} * P(H_iH_j)$$

Given λ, F_{ij} and the allelic frequencies of H_i haplotypes it is possible to calculated R_{ij} the probability for a sib of proband to be affected conditionally that the proband has a genotype H_iH_j:

$$R_{ij} = P(S_{Aff} / P_{Aff}, P_{ij}) = \frac{P(S_{Aff}, P_{Aff}, P_{ij})}{P(P_{Aff}, P_{ij})} = \frac{\lambda * \sum_{k=1}^{5} \sum_{l=1}^{5} P(Aff, C_{ikjl}, P_{Aff}, P_{ij})}{P(P_{Aff}, P_{ij})}$$

$$R_{ij} = \frac{P(P_{Aff}) * P(P_{ij} / P_{Aff}) * \lambda * \sum_{k=1}^{5} \sum_{l=1}^{5} \left[P(C_{ikjl} / P_{ij}) * P(Aff / C_{ikjl}) \right]}{P(P_{Aff}) * P(P_{ij} / P_{Aff})}$$

$$R_{ij} = \lambda * \sum_{k=1}^{5} \sum_{l=1}^{5} \left[P(C_{ikjl} / P_{ij}) * P(Aff / C_{ikjl}) \right]$$

$$R_{ij} = \lambda * \sum_{k=1}^{5} \sum_{l=1}^{5} \left[P(H_k) * P(H_l) * \frac{1}{4} * (F_{ij} + F_{il} + F_{kj} + F_{kl}) \right]$$

When the genotypes of parents are given, it is possible to refine this risk by calculated the R_{ikjl} risk which is the probability that a sib of a proband is affected conditionally that parents have genotypes H_iH_k and H_jH_l:

$$R_{ikjl} = P(S_{Aff} / C_{ikjl}) = \lambda * P(Aff / C_{ikjl}) = \lambda * \frac{1}{4} * (F_{ij} * F_{il} * F_{kj} * F_{kl})$$

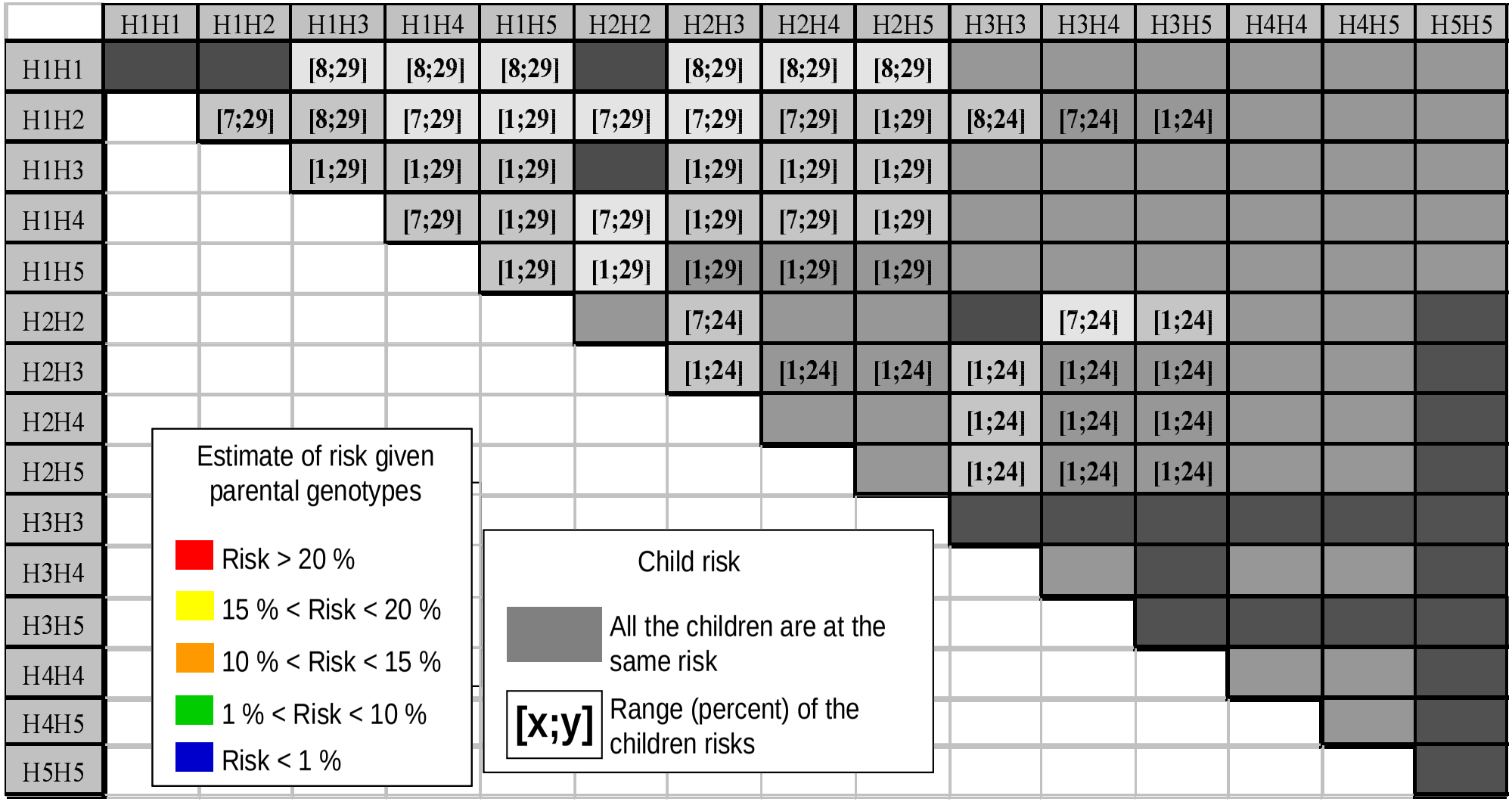

Figure 1: Risk for a sib of a proband according the DQ genotype of the parents

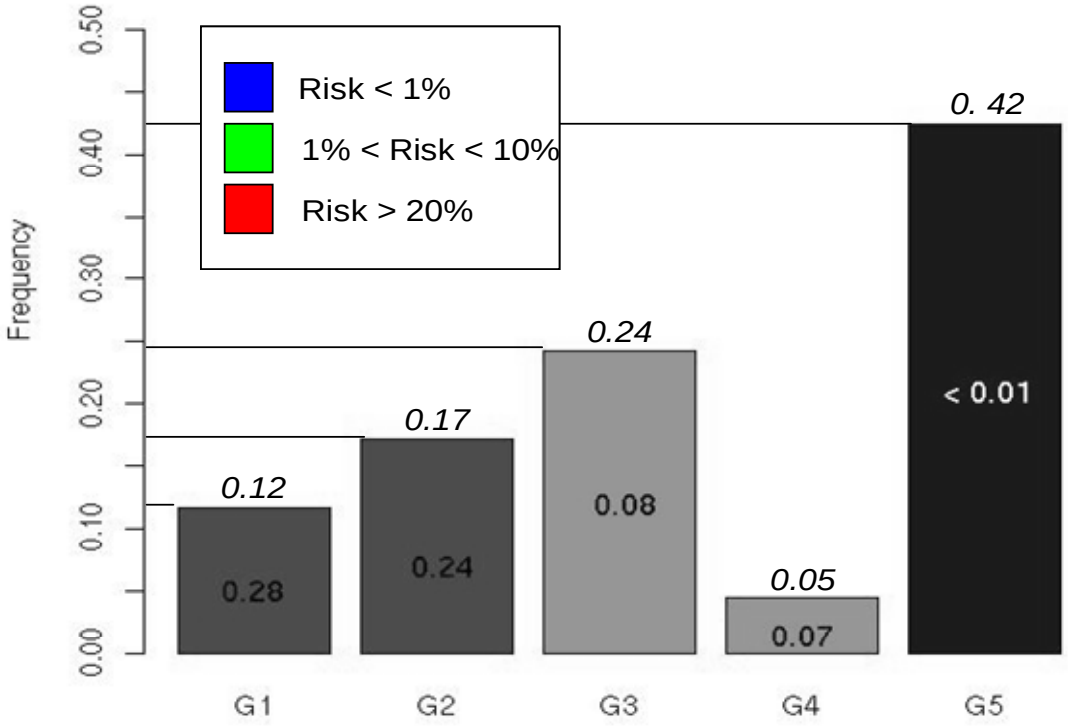

Figure 2: Probability for a sib of a proband to belong to Gi and corresponding risk.