

## COELIAC DISEASE

## Concordance, disease progression, and heritability of coeliac disease in Italian twins

L Nisticò, C Fagnani, I Coto, S Percopo, R Cotichini, M G Limongelli, F Paparo, S D'Alfonso, M Giordano, C Sferlazzas, G Magazzù, P Momigliano-Richiardi, L Greco, M A Stazi



Gut 2006;55:803–808. doi: 10.1136/gut.2005.083964

See end of article for authors' affiliations

Correspondence to:  
Dr L Nisticò, Genetic Epidemiology Unit, CNESPS, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Roma, Italia; lnistico@iss.it

Revised version received 24 November 2005  
Accepted for publication 28 November 2005  
Published online first 24 January 2006

**Background and aims:** We adopted the twin method to disentangle the genetic and environmental components of susceptibility to coeliac disease (CD). We estimated disease concordance rate by zygosity and HLA genotypes, discordance times, progression rates to disease, and heritability.

**Methods:** We crosslinked the Italian Twin Registry with the membership lists of the Italian Coeliac Disease Association and recruited 23 monozygotic (MZ) and 50 dizygotic (DZ) twin pairs with at least one affected member. Zygosity was assigned by DNA fingerprinting, and HLA-DQ and DR alleles were genotyped. Disease status was ascertained by antiendomysial, anti-human tissue transglutaminase antibodies, and bowel biopsy.

**Results:** Concordance was significantly higher in MZ (83.3% probandwise, 71.4% pairwise) than in DZ (16.7% probandwise, 9.1% pairwise) pairs. Concordance was not affected by sex or HLA genotype of the co-twin and being MZ was significantly associated with the occurrence of CD (Cox adjusted hazard ratio 14.3 (95% confidence interval 4.0–50.3)). In 90% of concordant pairs the discordance time was  $\leq 2$  years. MZ and DZ co-twins had 70% and 9% cumulative probability of having symptomatic or silent forms of CD, respectively, within five years. Under ACE (additive genetic, common, and unshared environmental factors) models, with CD population prevalences of 1/91 and 1/1000, heritability estimates were 87% and 57%, respectively.

**Conclusion:** MZ pairs have a high probability of being concordant, regardless of sex or HLA genotype. Most of the affected co-twins receive a diagnosis within two years. A remarkable proportion of phenotypic variance is due to genetic factors.

Coeliac disease (CD) is an immune mediated intolerance triggered by ingestion of gluten containing grains in susceptible individuals. It occurs with inflammation of the upper small intestine, which may lead to malabsorption of important nutrients, such as iron, folic acid, and calcium. Several symptoms may be related to untreated CD. The clinical presentation of these symptoms varies according to age, and a great number of patients present with the condition in a silent form; as a consequence, the population prevalence of CD is probably underestimated. Indeed, a large scale screening in the Italian paediatric population revealed the high prevalence of CD (1:91), with two thirds of cases being clinically silent.<sup>1</sup>

From an aetiological standpoint, CD is a multifactorial disease, involving both genetic and environmental factors. Apart from the HLA region, already known to bear one or more risk loci for CD, at least two additional susceptibility loci map onto chromosome 5q31-33 and 2q33.<sup>2-3</sup>

Despite a number of assumptions that have been investigated extensively in the past and might not be met in specific situations, the twin method constitutes a powerful tool to disentangle the genetic and environmental causes of family resemblance for a given trait.<sup>4</sup> Over the past years, the potential of twin studies has enormously increased with the establishment of population based registries of data on twins that represent some of the best resources for genetic epidemiological research.<sup>5</sup> Matching one such registry with disease records makes it feasible to collect twin pairs in which at least one member is affected; this results in samples that are largely representative of the general twin population, and also ensures a gain in terms of statistical power.

We adopted the twin approach to estimate: (1) CD concordance rate by zygosity and HLA status, (2) discordance times, (3) progression rates to disease, and (4) heritability.

## SUBJECTS AND METHODS

### Twin recruitment

Twins were identified according to a procedure already described in a previous paper.<sup>6</sup> Briefly, membership lists (6998 records of the Italian Coeliac Disease Association (AIC)) were matched with the Italian Twin Registry that includes approximately 650 000 potential twin pairs born before 31 December 1996.<sup>7-8</sup> AIC lists were from Southern (Campania, Basilicata, Puglia, Calabria, Sicilia) and Northern (Piemonte) Italian regions. Eighty one twin pairs were identified, giving a ratio of 2.3 twins/100 individuals: 58 pairs were already known,<sup>6</sup> and 23 were identified de novo (13 from Piemonte and 10 from Southern Italy). Eight pairs (four opposite sex, three female, and one male) refused to participate and 73 pairs entered our study. In this paper, the twin in a pair who first received a diagnosis of CD is referred to as the proband or the index twin; co-twin usually indicates the second twin. The study was approved by the ethics committee of the University of Naples and by the AIC. All twin pairs or their parents gave informed consent.

**Abbreviations:** CD, coeliac disease; MZ, monozygotic, DZ, dizygotic; AIC, Italian Coeliac Disease Association; anti-tTG, anti-human tissue transglutaminase antibodies; ACE, additive genetic, common, and unshared environmental factors

### HLA typing

Peripheral blood was drawn from 73 twin pairs and 128 parents, using EDTA as anticoagulant. Genomic DNA was purified from leucocytes of all subjects using the salting out procedure.<sup>9</sup> HLA-DRB1 and -DQB1 loci were typed in all twins and their parents, when available, using commercial kits (Dynal Oxoid, Italy). HLA chromosomes identical by descent were unambiguously assigned in 41 pairs.

### Zygosity test

Zygosity was provisionally assigned by a standardised questionnaire and then confirmed using DNA fingerprinting only in same sex twin pairs having identical HLA genotypes. We used the AmpFLSTR Profiler Plus PCR Amplification Kit (Applied Biosystems, Foster City, California, USA) that amplifies, at the same time, nine short tandem repeats all localised on different chromosomes. Probability of identity by chance is approximately  $10^{-4}$ . Polymerase chain reaction products were separated by capillary electrophoresis on an automatic sequencer ABI Prism 310 (Applied Biosystems).

### Disease status

Seemingly unaffected co-twins were visited and screened for CD: serum samples were examined for antiendomysial and anti-human tissue transglutaminase (anti-tTG) antibodies. IgA endomysial antibodies were detected by an indirect immunofluorescence method using cryostat sections of human umbilical cord as antigen<sup>10</sup>; sera were tested at 1:5 dilution. Serum IgA anti-tTG levels were measured by an enzyme linked immunosorbent assay method.<sup>11</sup> All samples were tested in duplicate and values were expressed as percentage of a reference pool of sera, obtained from untreated coeliac patients. All unaffected co-twins were tested for IgA levels to exclude IgA deficiency. In all cases of positive screening, a diagnosis of CD was confirmed by a small bowel biopsy showing typical pathological changes. All affected subjects fulfilled the ESPGHAN (European Society of Paediatric Gastroenterology, Hepatology, and Nutrition) diagnostic criteria.

### Statistical analysis

#### Concordance

We estimated CD concordance by zygosity and sex, using probandwise ( $P_C$ ) and pairwise ( $P_P$ ) concordance rates under incomplete ascertainment.<sup>12</sup> For concordant affected pairs, a distinction was made between “doubly” (D) ascertained, for which both twins were in the disease records, and “singly” (S) ascertained, where only one twin was in the records and the second twin was found to be affected on further examination.  $P_C$  is defined as the probability that one twin in a pair is affected, given that his/her co-twin is affected, and can be estimated as:

$$P_C = (2D+S)/(2D+S+d)$$

where d is the number of discordant pairs.

$P_P$  is the probability that both twins in a pair are affected, given that at least one is affected, and can be expressed as:

$$P_P = (2D+S)/(2D+S+2d).$$

### Survival

We performed survival analyses using the Kaplan-Meier method on 72 pairs to describe progression of CD in co-twins. One MZ concordant pair was excluded because both twins were screened shortly after CD was diagnosed in another sibling. We considered two survival models: one in which the terminating event was defined as CD diagnosis as a consequence of appearance of symptoms or positive screening and biopsy in silent co-twins, and another where the event was defined as CD diagnosis as a consequence of symptom appearance only, while positively screened twins were incorporated as censored observations. Moreover, we adopted the Cox regression model to evaluate the impact of co-twin sex, HLA genotype group, and zygosity on CD concordance.

### Heritability

We estimated genetic and environmental components of variance in CD using structural equation modelling and the software Mx.<sup>13</sup> We considered an ACE model incorporating parameters for additive genetic (A), common (shared) environmental (C), and individual specific (unshared) environmental (E) components of variance.<sup>4</sup> Additive genetic factors are completely correlated in MZ twins who are genetically identical, and correlate 0.5 in DZ twins, who share on average 50% of their segregating genes. Common environmental factors are shared completely by the twins regardless of zygosity, while unshared environmental influences act separately on each twin and are therefore responsible for less than perfect resemblance between MZ twins. Under these assumptions, the expectations for total variance (V) and the covariance within twin pairs are given by<sup>4</sup>:

$$V = A+C+E$$

$$\text{Cov}(MZ) = A+C$$

$$\text{Cov}(DZ) = 0.5 \times A+C.$$

Heritability ( $h^2$ ) is the proportion of the total variance that is attributable to the genetic variance; in the formula

$$h^2 = A/V.$$

The A, C, and E parameters were estimated by maximum likelihood, under a “liability threshold” model<sup>14</sup>; based on this model, there exists an underlying (continuous) liability to CD, normally distributed in the population, with a threshold such that individuals above the threshold are affected and those below are not.

This study relied on incomplete ascertainment of twins.<sup>12</sup> Thus the likelihood estimation had to be corrected by taking into account the ascertainment probability ( $0 < \pi < 1$ ) (that is, the probability that an affected twin is in the disease records) and by fixing the threshold of liability at the value corresponding to the population prevalence of the disease. Ascertainment probability can be approximated as  $\pi = 2D/(2D+S)$  where D and S are numbers of concordant pairs doubly and singly ascertained, as described in the concordance section. The appropriate script is available on the Mx website (<http://www.vcu.edu/mx/examples.html>).

**Table 1** Age at diagnosis and at enrolment of the twins

| No of pairs               | Monozygotic twins   |                    |                | Dizygotic twins    |                     |                |
|---------------------------|---------------------|--------------------|----------------|--------------------|---------------------|----------------|
|                           | Concordant (n = 17) | Discordant (n = 6) | Total (n = 23) | Concordant (n = 5) | Discordant (n = 45) | Total (n = 50) |
| Mean age at diagnosis (y) |                     |                    |                |                    |                     |                |
| Index twin                | 10.3                | 6.9                | 9.4            | 4.6                | 9.4                 | 8.9            |
| Co-twin                   | 11.3                | –                  | –              | 12.1               | –                   | –              |
| Mean age at enrolment (y) | 21                  | 20.5               | 20.8           | 22.4               | 17.8                | 18.3           |

**Table 2** Concordance by zygosity and sex in twin pairs

|                 | Concordant | Discordant | Total | Concordance (%) |           |          |           |
|-----------------|------------|------------|-------|-----------------|-----------|----------|-----------|
|                 |            |            |       | Probandwise     | 95% CI    | Pairwise | 95% CI    |
| MZ male         | 5          | 1          | 6     | 90              | 70.6–100  | 81.8     | 49.7–100  |
| MZ female       | 12         | 5          | 17    | 80.8            | 64.4–97.1 | 67.7     | 44.7–90.7 |
| All MZ          | 17         | 6          | 23    | 83.3*           | 70.3–96.4 | 71.4†    | 52.3–90.6 |
| DZ male         | 0          | 12         | 12    | 0               |           | 0        |           |
| DZ female       | 1          | 14         | 15    | 12.5            | 0–34.7    | 6.7      | 0–19.3    |
| DZ opposite sex | 4          | 19         | 23    | 26.9            | 5.2–48.7  | 15.6     | 1–30.1    |
| All DZ          | 5          | 45         | 50    | 16.7*           | 3.6–29.8  | 9.1†     | 1.3–16.9  |
| Total           | 22         | 51         | 73    |                 |           |          |           |

MZ, monozygotic, DZ, dizygotic; 95% CI, 95% confidence interval.

Test for difference between MZ and DZ twins:  $\chi^2 = 49.98$ ,  $p = 1.55 \times 10^{-12}$ ;  $\dagger \chi^2 = 40.77$ ,  $p = 1.71 \times 10^{-10}$ .

## RESULTS

### Recruitment and zygosity

We enrolled 73 twin pairs. Twenty three pairs were MZ (six males, 17 females) and 50 were DZ (12 males, 15 females, and 23 opposite sex). The MZ/DZ same sex/DZ opposite sex ratio was 0.9/1.1/0.9, and was not significantly different from that expected (1/1/1). Female/male ratio of the index twins was 2.2. Age at enrolment was similar in MZ and DZ twins (*t* test,  $p = 0.30$ ) (table 1).

### Disease status

Seventeen pairs were known to be concordant before entering the study. Five co-twins (three female MZ, one male MZ, and one male DZ) were diagnosed during the study because they were positive to autoantibody screening and to intestinal biopsy. Four were clinically silent and one MZ co-twin had symptoms. Age at diagnosis varied greatly (range 0.5–57 years) although 50% of all affected twins were diagnosed before three years of age. Mean age at diagnosis was similar in MZ and DZ index and second twins (table 1) but was lower in male than in female twins (7.3 *v* 10.8 years; *t* test,  $p = 0.17$ ).

When patients or their parents were asked about symptoms that led to the diagnosis, the most frequent answers were: diarrhoea (51%), vomiting (39%), weight loss (29%), anaemia (22%), and abdominal distension (19%). Twelve twins claimed to be symptom less and to have been positively screened for CD because of affected index twins (10 co-twins), affected mother, or non-twin sister (two index twins).

### Concordance estimates

Overall, 17/23 MZ and 5/50 DZ twin pairs were concordant for CD: probandwise ( $P_c$ ) and pairwise ( $P_p$ ) concordance was

significantly different between MZ and DZ twins, with point estimates of 83% and 71% in MZ twins, and 17% and 9% in DZ twins, respectively (table 2). Concordance by sex did not significantly differ in MZ twins. In DZ, 1/5 concordant pairs was female and 4/5 were opposite sex. None of the 12 DZ male pairs was concordant (table 2). In 15/19 discordant opposite sex pairs, the affected twins were females.

Tables 3 and 4 show disease concordance according to HLA genotypes in MZ and DZ twin pairs. It was recently shown that HLA-CD association is better described by a risk hierarchy of DR-DQ genotypes rather than by DQ2-DQ8 molecules.<sup>15</sup> Accordingly, we stratified twin pairs into four genotype groups with decreasing risk of CD: the highest risk is due to DR3/3 and DR3/7 genotypes (group 1, G1); DR5/7 confers one third less risk (G2); the relative risks of DR3/X heterozygous (G3) and of DR7/7, DR4/4, and DR4/7 genotypes (G4) are estimated to be approximately one quarter of the G1 risk; finally, the fifth group (G5) includes genotypes (DR4/Y, DR5/X, DR7/Z) with very low risk of CD (1/50th of G1).

All MZ twin pairs but one belonged to the G1–G4 risk categories (table 3). Discordant MZ co-twins were DR3/7 ( $n = 2$ ), DR5/7 ( $n = 3$ ), and DR3/5 ( $n = 1$ ). In DZ pairs, 48/50 index twins carried G1–G4 genotypes (table 4) and 2/50 were DR1/7 and DR4/13 (G5).

In MZ, the highest proportion of concordant pairs was observed in the G3–G4 genotype group (eight concordant, one discordant). In DZ, the same proportions of concordant pairs were observed in the groups, with both twins having high risk G1 genotypes (one concordant and five discordant) or G3–G4 genotypes (two concordant, 10 discordant).

In each HLA stratum, disease concordance was higher in MZ pairs compared with DZ sibs sharing identical HLA chromosomes (data not shown). Indeed, in the Cox regression model, being MZ was the only factor significantly associated with the occurrence of CD in the co-twin (adjusted hazard ratio by sex and genotype group in the co-twin = 14.3 (95% confidence interval 4.0–50.3)).

### Survival analysis

There were eight disease concordant pairs (seven MZ, one DZ) in which symptoms appeared almost simultaneously (0–2 months) in both twins, and one DZ pair with a discordance time of seven months. In two MZ pairs, diagnosed before the introduction of autoantibody screening, symptom appearance in the co-twins was delayed by two years. In 10 concordant pairs the second twins were clinically silent or paucisymptomatic. Of these, seven co-twins (five MZ, two DZ) were screened for autoantibodies and diagnosed as coeliac within one year of the diagnosis of their siblings, one MZ was diagnosed within two years, and in two co-twins (one MZ and one DZ) coeliac status was ascertained after 10 and 37 years. Thus in most (19/21) of our concordant pairs the

**Table 3** Concordance by HLA-DR genotype risk group in monozygotic pairs

|                                    | Concordant | Discordant | Total |
|------------------------------------|------------|------------|-------|
| Risk genotype group†               |            |            |       |
| G1 (DR3/3, DR3/7)                  | 5          | 2          | 7     |
| G2 (DR5/7)                         | 3          | 3          | 6     |
| G3–G4 (DR3/X, DR4/4, DR4/7, DR7/7) | 8          | 1          | 9     |
| G5 (DR4/Y, DR5/X, DR7/Z)           | 1          | 0          | 1     |
| Total                              | 17         | 6          | 23    |

†HLA grouping is according to relative risks for coeliac disease estimated in the Italian population according to Margaritte-Jeannin and colleagues.<sup>15</sup> DR3 haplotype carries DQB1\*0201 and DQA1\*0501, DR7 carries DQB1\*0202, DR5 carries DQA1\*0501, and DR4 carries DQB1\*0302.

X = not DR3 or 7; Y = not DR3, 4, or 7; Z = not DR3, 4, 5, or 7.

**Table 4** Concordance in dizygotic pairs by HLA-DR genotype risk group in index and co-twin

| Risk genotype group in the index twin† | Risk genotype group in the co-twin‡ |     |     |     |       |     |     |     | Total |
|--|-------------------------------------|-----|-----|-----|-------|-----|-----|-----|-------|
|  | G1                                  |     | G2  |     | G3-G4 |     | G5  |     |       |
|  | Con                                 | Dis | Con | Dis | Con   | Dis | Con | Dis |       |
| G1 (DR3/3, DR3/7)                      | 1                                   | 5   | 0   | 1   | 1     | 6   | 0   | 3   | 17    |
| G2 (DR5/7)                             | 0                                   | 1   | 0   | 1   | 0     | 0   | 0   | 7   | 9     |
| G3-G4 (DR3/X, DR4/4, DR4/7, DR7/7)     | 1                                   | 3   | 0   | 2   | 2     | 10  | 0   | 4‡  | 22    |
| G5 (DR4/Y, DR5/X, DR7/Z)               | 0                                   | 0   | 0   | 0   | 0     | 1   | 0   | 1   | 2     |
| Total                                  | 2                                   | 9   | 0   | 4   | 3     | 17  | 0   | 15  | 50    |

Con, number of concordant co-twins; Dis, number of discordant co-twins.

†HLA grouping is according to relative risks for coeliac disease estimated in the Italian population according to Margaritte-Jeannin and colleagues.<sup>15</sup> DR3 haplotype carries DQB1\*0201 and DQA1\*0501, DR7 carries DQB1\*0202, DR5 carries DQA1\*0501, and DR4 carries DQB1\*0302.

‡One co-twin is DR13/13.

X=not DR3 or 7; Y=not DR3, 4, or 7; Z=not DR3, 4, 5, or 7.

disease discordance time was  $\leq 2$  years, either as a consequence of symptom appearance or because disease signs were actively searched for in clinically silent co-twins.

The cumulative probability of being diagnosed with CD was significantly higher for MZ compared with DZ co-twins in both survival models (symptomatic CD, log rank test:  $p = 1.6 \times 10^{-5}$ ; symptomatic or silent CD, log rank test:  $p = 3.5 \times 10^{-9}$ ).

In MZ co-twins, the probability of having symptomatic forms of CD within one year after the diagnosis in their identical siblings was 36%, much higher than that observed in DZ co-twins (4%). The five year cumulative probability further increased for MZ (50%) while it did not change for DZ co-twins (fig 1A). Including silent forms, progression rates for MZ and DZ co-twins were 50% and 8% within one year, and 70% and 9% within five years (fig 1B). In our discordant pairs, the shortest follow up times were 4.6 years for MZ and 1.8 years for DZ co-twins.

### Heritability

In the present study we had 13 and four doubly ascertained pairs among MZ and DZ twins, respectively, and four and one singly ascertained pairs among MZ and DZ twins, respectively; this resulted in an ascertainment probability of  $\pi = (2 \times 17) / (2 \times 17 + 5) = 34/39 = 0.87$ .

Given the so-called "iceberg" structure of CD,<sup>16</sup> which underlines the importance of the subclinical component, we fitted the same ACE model assuming, for the population prevalence, values of 1/1000 (threshold 3.09) from clinical diagnosis data and 1/91 (threshold 2.29) from screening

data.<sup>1</sup> Each threshold is the value such that the proportion of the standard normal distribution of liability above the value exactly matches the observed proportion of affected individuals in the population, namely CD prevalence.

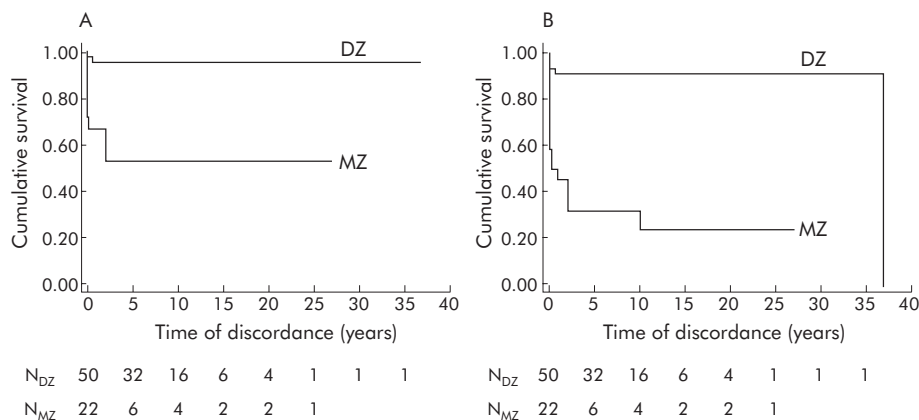
Estimates of genetic and environmental variance components are shown in table 5.

Under the ACE model with population prevalence for CD of 1/1000, 57% of the variation in liability to CD was explained by additive genetic factors, while 42% and 1% were the contributions of common and unique environmental factors, respectively. When we fitted the same ACE model with a population prevalence of 1/91, the heritability estimate became 87%, and the relative weight of common environmental factors decreased to 12%; the contribution of the unshared environmental component of variance remained at the 1% level.

### DISCUSSION

As expected, concordance in MZ twins was significantly higher compared with DZ pairs and was very close to previously reported values in smaller sample sizes.<sup>6</sup> Moreover, given that the number of singly ascertained pairs was relatively small, concordance estimates were similar under both complete and incomplete ascertainment conditions. In DZ, the proportion of affected co-twins ( $5/50 = 10\%$ ) was in line with CD prevalences of 4% to 12% in first degree relatives reported in several studies.<sup>17</sup>

Concordance rates by sex were not significantly different within MZ and DZ groups. In DZ pairs, the highest point estimates in opposite sex twins were not due to a longer



**Figure 1** Kaplan-Meier survival curves in co-twins of coeliac probands by zygosity. The terminating event is: (A) coeliac disease (CD) diagnosis as a consequence of symptom appearance only; (B) CD diagnosis as a consequence of either appearance of symptoms or positive screening.  $N_{DZ}$  and  $N_{MZ}$  are the numbers of dizygotic (DZ) and monozygotic (MZ) co-twins entering each time interval.



**Table 5** Estimates of additive genetic (A), shared environmental (C), and unshared environmental (E) variance components in coeliac disease liability, using two different values for population prevalence

|                        | Variance component |                  |                  |
|------------------------|--------------------|------------------|------------------|
|                        | A (95% CI)         | C (95% CI)       | E (95% CI)       |
| Prevalence (threshold) |                    |                  |                  |
| 1/1000 (3.09)          | 0.57 (0.32–0.93)   | 0.42 (0.06–0.67) | 0.01 (0.00–0.03) |
| 1/91 (2.29)            | 0.87 (0.49–1.00)   | 0.12 (0.00–0.49) | 0.01 (0.00–0.05) |

95% CI, 95% confidence interval.

mean follow up time as it was similar to that of female pairs and even shorter than that of male pairs.

Six of 23 MZ pairs were disease discordant after a period of 4.5–27 years of follow up. Discordance in MZ twins is usually attributed to differential exposure to environmental risk factors; in the case of CD, precautionary gluten elimination from the diet of unaffected co-twins cannot be ruled out. An alternative explanation for the differences between genetically identical individuals involves epigenetic modifications such as DNA methylation and histone acetylation, occurring after twin separation, that may control expression and silencing of disease genes.<sup>18</sup>

In both MZ and DZ pairs, we observed that proportions of affected co-twins were not significantly different between those with G1 or G3–G4 HLA genotypes. This may indicate either random fluctuation or that risk hierarchy shown in the general population cannot be applied to this specific setting where high susceptibility HLA genotypes may have been picked up.<sup>15</sup>

In our co-twins, the only significant susceptibility factor to CD was sharing the whole genetic background with the affected twin. Indeed, at least two additional loci on chromosomes 2 and 5 are associated with or linked to CD in the Italian population.<sup>2, 3</sup>

In our concordant pairs, most of the co-twins received a diagnosis of CD shortly after their siblings, with median discordance time of one month for both MZ and DZ twins; as median follow up times in discordant pairs were 10.5 and 7.7 years for MZ and DZ, respectively, we do not expect any dramatic variation in concordance rates as a consequence of an even longer follow up.

It is well known that CD is frequently paucisymptomatic or silent. In our study, approximately half of the MZ and DZ affected co-twins were positive to antibody screening and had a flat gut mucosa despite the scarcity or absence of symptoms. Thus if CD diagnoses due to symptoms appearance only had been considered, disease incidence would have been largely underestimated in MZ and DZ non-index twins. The five year cumulative incidence ratios of MZ relative to DZ co-twins were 12.5 (0.50/0.04) when silent co-twins were excluded and 7.8 (0.70/0.09) if asymptomatic second twins were included.

Our data showed a substantial heritability for CD. Using the ACE models with population prevalences for CD of 1/1000 and 1/91, heritability estimates were 57% and 87%, respectively, indicating that approximately 60–90% of the variance in liability to the disease had a genetic origin. However, because of the limited power of this study, which is reflected in the large confidence intervals, we have to be cautious in interpreting the point estimates.

Moreover, it could be worth reminding the reader that these estimates give us only a general understanding of the genetic influence on the phenotype—that is, they simply suggest a genetic role in determining interindividual phenotypic differences but they tell us nothing about which genes are directly implicated in the emergence of the disease. Also,

they provide no information on whether it is possible to modify the prognosis by environmental interventions, such as lifestyle changes. With respect to this issue, elimination of gluten from the diet is a typical example of environmental intervention that, in the case of CD, can result in total recovery of gut function and correction of most other consequences, despite considerable heritability.

It is clear that the approach we used strongly depends on fixing a priori the level of the population prevalence; this is required to correct the likelihood function for the ascertainment procedure when selected samples of twins, including only pairs with at least one affected member, are considered. In our case, if we had solely relied on the prevalence estimated from clinical data, we would have underestimated the genetic penetrance and overestimated the shared environmental influences.

Based on our data, common environment could not be ruled out completely as a contributing aetiological factor. Common environment refers to any shared environmental factor that contributes to the resemblance of members of a twin pair, regardless of zygosity. It may include biological events such as exposure to infectious agents, dietary characteristics, and other intrauterine factors that may influence the similarity of CD patterns within pairs. Moreover, it is clear that the environments of members of twin pairs are most alike in utero and during the first postnatal period, while they tend to diverge over time. Therefore, possible effects of common environment on liability to CD was seen both in the short discordance time and in the young age at diagnosis observed in our sample for the majority of concordant pairs.

To our knowledge, no previous investigations have estimated, using the twin model, the heritability of CD in Italy or in other countries. Rather, family studies have been extensively used for unravelling genetic predisposition to CD.<sup>19, 20</sup> In principle, they consist of comparing the risk for relatives of affected individuals with that of the general population, and unlike the twin method they suffer from the disadvantage of not providing information on the extent to which the observed familial resemblance has a genetic basis or is attributable to shared environmental exposure.

The heritability point estimate under the high prevalence scenario was in line with values found in twin studies for other HLA mediated diseases, such as type 1 diabetes (88%),<sup>21</sup> Graves' disease (79%), and psoriasis (80%)<sup>22, 23</sup>; this could reflect shared pathogenetic mechanisms that may partially explain the comorbidity of these conditions.<sup>24–26</sup>

## ACKNOWLEDGEMENTS

We warmly thank the twins, their families, and the Associazione Italiana Celiachia (AIC).

We are grateful to the following clinical investigators: Giuseppina Oderda (Dipartimento Scienze Mediche, Università del Piemonte Orientale), Cristiana Barbera (Clinica Pediatrica, Università di Torino), Carla Sategna-Guidetti (Dipartimento Medicina Interna, Università di Torino), and Giuseppe Di Pasquale and Riccardo Scoglio (Dipartimento Scienze Pediatriche, Mediche e Chirurgiche,

Università di Messina). We also acknowledge the skilful assistance of Cristina D'Ippolito in the management of the Italian Twin Registry database.

This work was supported by ELFID (European Laboratory for Food Induced Disease), the Italian Ministero della Salute (project 2000, No 0AB/F to MAS) and the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR-PRIN No MM06187812 to SD, GM, and LG).

CF was supported by the Genomewin project (European Union Contract No QL2G-CT-2002-01254).

#### Authors' affiliations

**L Nisticò, C Fagnani, R Cotichini, M A Stazi**, Genetic Epidemiology Unit, National Centre of Epidemiology, Surveillance and Health Promotion (CNESPS), Istituto Superiore di Sanità, Rome, Italy

**I Coto, S Percopo, M G Limongelli, F Paparo, I Greco**, Department of Paediatrics, University of Naples Federico II, Naples and European Laboratory of Food Induced Disease (ELFID), Naples, Italy

**S D'Alfonso, M Giordano, P Momigliano-Richiardi**, Department of Medical Sciences, Eastern Piedmont University "A Avogadro" and Interdisciplinary Research Center for Autoimmune Diseases (IRCAD), Novara, Italy

**C Sferlazzas, G Magazzù**, Department of Paediatric Sciences, University of Messina, Messina, Italy

Conflict of interest: None declared.

#### REFERENCES

- Tommasini A, Not T, Kiren V, et al. Mass screening for coeliac disease using antihuman transglutaminase antibody assay. *Arch Dis Child* 2004;**89**:512-15.
- 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;**65**:35-41.
- Holopainen P, Naluai AT, Moodie S, et al. Candidate gene region 2q33 in European families with coeliac disease. *Tissue Antigens* 2004;**63**:212-22.
- Neale MC, Cardon LR. *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publisher, 1992:82-8.
- Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nature* 2002;**3**:872-82.
- Greco L, Romino R, Coto I, et al. The first large population based twin study of coeliac disease. *Gut* 2002;**50**:624-8.
- Salveti M, Ristori G, Tosi R, et al. Italian population yields world's largest twin registry. *Nat Med* 1997;**3**:1176.
- Stazi MA, Cotichini R, Patriarca V, et al. The Italian twin project: from the personal identification number to a national twin registry. *Twin Res* 2002;**5**:382-6.
- Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
- Not T, Città A, Lucchesi A, et al. Antiendomysium antibody on human umbilical cord vein tissue: an inexpensive and sensitive diagnostic tool for the screening of the coeliac disease. *Eur J Pediatr* 1997;**156**:616-18.
- Sblattero D, Berti I, Trevisiol C, et al. Human recombinant tissue transglutaminase ELISA: an innovative diagnostic assay for celiac disease. *Am J Gastroenterol* 2000;**95**:1253-7.
- Witte JS, Carlin JB, Hopper JL. Likelihood-based approach to estimating twin concordance for dichotomous traits. *Genet Epidemiol* 1999;**16**:290-304.
- Neale MC, Boker SM, Xie G, et al. *Mx: Statistical modeling*, 6th Edn., Richmond, VA: Department of Psychiatry, Virginia Commonwealth University 2002:11-151.
- Sham P. *Statistics in human genetics*. London: Arnold, 1998:221-34.
- Margarite-Jeannin P, Babron MC, Bourgey M, et al. HLA-DQ relative risks for coeliac disease in European populations: a study of the European Genetics Cluster on Coeliac Disease. *Tissue Antigens* 2004;**63**:562-7.
- Fasano A, Catassi C. Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum. *Gastroenterology* 2001;**120**:636-51.
- Dubé C, Rostom A, Sy R, et al. The prevalence of celiac disease in average-risk and at-risk Western European populations: a systematic review. *Gastroenterology* 2005;**128**(suppl 1):S57-67.
- Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005;**102**:10604-9.
- Hogberg L, Falth-Magnusson K, Grodzinsky E, et al. Familial prevalence of coeliac disease: a twenty-year follow-up study. *Scand J Gastroenterol* 2003;**38**:61-5.
- Book L, Zone JJ, Neuhausen SL. Prevalence of celiac disease among relatives of sib pairs with celiac disease in U.S. families. *Am J Gastroenterol* 2003;**98**:377-81.
- Hyttinen V, Kaprio J, Kinnunen L, et al. Genetic liability of type 1 diabetes and the onset age among 22650 young Finnish twin pairs: a nationwide follow-up study. *Diabetes* 2003;**52**:1052-5.
- Brix TH, Kyvik KO, Christensen K, et al. Evidence for a major role of heredity in Graves' disease: a population-based study of two Danish twin cohorts. *J Clin Endocrinol Metab* 2001;**86**:930-4.
- Duffy DL, Spelman LS, Martin NG. Psoriasis in Australian twins. *J Am Acad Dermatol* 1993;**29**:428-34.
- Barera G, Bonfanti R, Viscardi M, et al. Occurrence of celiac disease after onset of type 1 diabetes: a 6-year prospective longitudinal study. *Pediatrics* 2002;**109**:833-8.
- Ch'ng CL, Biswas M, Benton A, et al. Prospective screening for coeliac disease in patients with Graves' hyperthyroidism using anti-gliadin and tissue transglutaminase antibodies. *Clin Endocrinol (Oxf)* 2005;**62**:303-6.
- Ojetti V, Aguilar Sanchez J, Guerriero C, et al. High prevalence of celiac disease in psoriasis. *Am J Gastroenterol* 2003;**98**:2574-5.