Digestive and Liver Disease 46 (2014) 1057–1063

Contents lists available at ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Progress Report

Co-localization of gluten consumption and HLA-DQ2 and -DQ8 genotypes, a clue to the history of celiac disease

Elena Lionetti[a,*], Carlo Catassi[b,c]

[a] Department of Pediatrics, University of Catania, Catania, Italy
[b] Department of Pediatrics, Università Politecnica delle Marche, Ancona, Italy
[c] Division of Pediatric Gastroenterology and Nutrition and Center for Celiac Research, Massachusetts General Hospital for Children, and Celiac Program – Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Article history:
Received 1 April 2014
Accepted 1 August 2014
Available online 4 September 2014

Keywords:
Celiac disease
Evolution
HLA
Wheat

ABSTRACT

Celiac disease is an immune-mediated disorder triggered by gluten in genetically susceptible persons. Despite its detrimental effects on human health, it has not disappeared over time. The current evolutionary theory is that celiac disease is more common in areas reached later by agricultural revolution than in countries that started consumption of wheat earlier, due to negative selection caused by celiac disease.

We reviewed data on worldwide prevalence of celiac disease, wheat consumption, and frequencies of HLA-celiac-disease-predisposing-genotypes to investigate their mutual relationship. Studies assessing prevalence of celiac disease were identified through a MEDLINE search. Wheat consumption and frequencies of HLA-DQ2-DQ8 were obtained from Food and Agriculture Organization of the United Nations and allelefrequencies.net database. Correlations between celiac disease, wheat consumption, and HLA were analyzed by linear regression. We observed a significant correlation between wheat consumption and HLA DQ2 ($p = 0.01$) and the sum of DQ2 and DQ8 ($p = 0.01$) frequencies. Wheat consumption and HLA-DQ2 tend to co-localize in different continents. The correlation between the prevalence of celiac disease and either DQ2 and/or DQ8, or the product of DQ2 + DQ8*wheat consumption was not statistically significant.

Co-localization of gluten consumption and HLA-celiac-disease-predisposing-genotypes can be explained by positive selection of HLA-DQ2 genes in wheat-consuming areas, and "demic diffusion" of Middle East farmers into Europe.

1. Introduction

Celiac disease (CD) is a systemic immune-mediated disorder triggered by dietary gluten in genetically susceptible persons, resulting from the interplay between environmental and genetic factors. The major component of the genetic predisposition, explaining about 40% of it, depends on some HLA-related class II genes: most CD patients are HLA-DQ2 positive (90%), while half of the remaining 10% are HLA-DQ8 positive. As concerns the environmental factors, the ingestion of gliadin and glutenin fractions of wheat gluten and similar alcohol-soluble proteins in other grains (rye and barley) triggers the development of the celiac enteropathy, as well as other disease manifestations [1–3].

In the past, CD was considered a rare disorder mostly affecting individuals of European origin. In 1978, Simoons formulated a theory on the evolutionary history of the disease. He postulated that the spreading of wheat consumption from the so-called "Fertile Crescent" in the Middle East along the Mediterranean countries to northern Europe exerted a negative selective pressure on genes predisposing to CD. The higher frequency of the HLA-B8 allele (at that time the only gene known to be associated with CD) and the resulting higher frequency of CD in north-western Europe were attributed to a lack of exposure to cereals until recent times [4].

The Simoons' theory did not survive the new era of CD genetics and epidemiology. The HLA-B8 was spuriously associated with CD, basically due to linkage disequilibrium with the HLA-DQ2, the "true" CD-predisposing genotype. At the same time, serological screening studies showed that (a) CD is one the commonest

* Corresponding author at: Department of Pediatrics, University of Catania, Via S. Sofia 78, 90124 Catania, Italy. Tel.: +39 095 378 17 82; fax: +39 095 378 23 85; mobile: +39 320 11 25 476.
E-mail address: elenalionetti@inwind.it (E. Lionetti).

http://dx.doi.org/10.1016/j.dld.2014.08.002
1590-8658/

lifelong disorders, affecting around 1% of the population worldwide and (b) CD is a common disorder in both north-western and south-eastern Europe despite some inter-country variability [5,6]. Despite its negative effects on human health, the CD phenotype has not disappeared over time, but is even increasing nowadays in areas with high level of both gluten consumption and CD-predisposing HLA-genotypes (so-called "evolutionary paradox of CD") [7–9].

The analysis of (a) the geographical distribution of genes that predispose to CD, (b) wheat consumption and (c) CD prevalence, can help to understand the evolutionary history of this multi-factorial disorder. To this aim, in the present study we reviewed the data on worldwide prevalence of CD, wheat consumption, and frequencies of HLA DQ2 and DQ8 haplotypes. We combined the three geographical maps to investigate their mutual relationship.

2. Methods

2.1. Protocol

Before review we developed a protocol, including eligibility criteria, search strategies, criteria for study selection, methods for extracting related data, and methods for assessing study quality and statistical methodology.

2.2. Information sources, eligibility criteria, search, and study selection

The current worldwide distribution of CD, wheat consumption and frequency of HLA DQ2 and DQ8 haplotypes were the focus of our search.

Studies on the prevalence of CD were identified by searching electronic databases and scanning reference lists of articles, and by consultation with experts in the field. The search was applied to the Medline database using PubMed by combining search terms for prevalence (prevalence, epidemiology) with keywords for CD (or coeliac disease, coeliac, gluten intolerance). As far as the diagnostic algorithm, we included only studies defining CD as either (a) CD positive serology plus typical small intestinal biopsy findings (Marsh 3a-3c) or (b) CD positive serology (high level of IgA anti-transglutaminase and/or antiendomysial antibodies). All types of study design (e.g. cross-sectional, cohort, case–control, and case series), except case reports, were considered for inclusion in this review. Search results were limited to studies based on the screening of unselected (low-risk) child and adult samples. Studies on healthy blood donors were excluded because they differed from the general population in social, health-related, and gender variables. No publication date or publication status restrictions were imposed. Eligibility assessment of studies was performed independently in an unblinded standardized manner by the two authors. Disagreements between reviewers were resolved by consensus. All studies described in this review were published between 1950 (start of Medline) and August 2012.

Worldwide distribution data of wheat consumption levels were obtained from the Food and Agriculture Organization of the United Nations (FAO) database (http://www.fao.org). The duration of wheat culture was taken from the work of Ammerman and Cavalli-Sforza [10]. According to the duration of wheat consumption we divided the Middle-East and European countries into four groups from 1 (gluten-containing cereals introduced around 10,000 years ago) to 4 (gluten-containing cereals introduced 6000–4000 years ago). The frequency of HLA-DQ2 and HLA-DQ8 in different countries were obtained by the allelefrequencies.net database (http://www.allelefrequencies.net) [11].

2.3. Data collection process

We developed a data extraction sheet (based on the Cochrane Consumers and Communication Review Group's data extraction template) [12], pilot tested it on three randomly selected included studies, and refined it accordingly. One author (E.L.) extracted the data from included studies, and the second author (C.C.) checked the extracted data.

2.4. Data items

Regarding the prevalence of CD, from each included study information was extracted on (1) study design; (2) characteristics of participants, including number, age, setting of enrolment, method of diagnosis of CD; and (3) prevalence of CD.

To evaluate the worldwide distribution of genes predisposing to CD we extracted the frequency of the allele DQB1*02 (DQ2), and the frequency of the haplotype DQA1*0301/DQB1*0302 (DQ8) from the allelefrequencies.net database. DQB1*02 positivity is the hallmark of the different CD-predisposing DQ2 haplo/genotypes, that is DQA1*05/DQB1*02 in *trans* or *cis* position (DQ2.5) and DQA1*02/DQB1*02 (DQ2.2).

To evaluate the worldwide wheat consumption we extracted data on wheat availability in the year 2009 for each country (expressed in grams per capita/day) from the FAO database.

2.5. Risk of bias

To ascertain the validity of the eligible studies, the study design, the size and representativeness of the study population (i.e. the presence of selection bias), the validity of outcomes (risk of confounding or bias), and the quality of the statistical analysis were taken into account. We assessed the methodological quality of included studies in accordance with the guidelines of the Cochrane Consumers and Communication Review Group, adapted for the current review concerning observational non-intervention studies. In all cases, the two authors independently assessed the quality of the studies included, with any disagreements resolved by discussion and consensus. Where necessary, study authors were contacted for additional information or for clarification of the study methods.

2.6. Statistical analysis

We calculated the prevalence of CD, the levels of wheat consumption, and the frequencies of HLA-DQ2 and HLA-DQ8 for different regions of the world. In case of multiple studies in the same country we calculated the weighted average. For each parameter we constructed a geographical map.

The correlations were studied by Spearman's correlation test and associations were further studied by multivariate linear regression analysis.

All differences were considered to be statistically significant at a 5% probability level.

Data were analyzed with SPSS 13.0 software (SPSS, Chicago, IL).

3. Results

3.1. Study selection and characteristics (CD prevalence)

Supplemental Figure 1 shows the flow diagram of study selection of studies on the prevalence of CD. Table 1 summarizes the studies reporting on the prevalence of CD [13–43]. Thirty-one publications were included, containing 38 CD prevalence estimates, of which 23 in children and 15 in adults, across 24 countries. The included studies involved 122,858 participants. None but one of the articles reported a power calculation to determine the population sample size.

Table 1
Summary of included studies evaluating the prevalence of celiac disease in different countries.

Country and reference	Setting	Population	Diagnosis*	Prevalence: *n* (%)
Algeria [13]	Healthy population	Children (2–15 yrs)	3	56/989 (5.6)
Argentine [14]	Screening pre-nuptial	Adults	2	12/2000 (0.6)
Australia [15]	Healthy population	Adults	2	12/3011 (0.4)
Brazil [16]	Care center	Children (1–14 yrs)	1	11/2034 (0.5)
Burkina Faso [17]	Healthy population	Adults	3	0/600 (0)
Egypt [18]	Healthy population	Children (7 mths–18 yrs)	2	8/1500 (0.5)
Finland [19]	School	Children (7–16 yrs)	2	37/3654 (1.0)
Finland [20]	Healthy population	Adults	1	113/4846 (2.4)
Germany [20]	Healthy population	Adults	1	8/3098 (0.2)
India [21]	School	Children (3–17 yrs)	1	14/4347 (0.3)
India [22]	Care center	Children (6–12 mths)	1	4/400 (1.0)
Iran [23]	Healthy population	Adults	1	27/2799 (1.0)
Iran [24]	Healthy population	Adults	1	7/1440 (0.5)
Iran [25]	School	Children (13 yrs)	1	3/634 (0.5)
Ireland [26]	Healthy population	Adults	1	15/1823 (0.8)
Italy [27]	School	Children	1	30/3188 (0.9)
Italy [20]	School	Children (10–19 yrs)	1	31/2645 (1.1)
Italy [20]	Healthy population	Adults	1	32/4781 (0.7)
Libya [28]	School	Children (5–17 yrs)	1	19/2920 (0.8)
Netherlands [29]	Healthy population	Children (2–4 yrs)	2	31/6127 (0.5)
New Zealand [30]	Healthy population	Adults	2	12/1064 (1.2)
Portugal [31]	School	Children (15 yrs)	1	4/536 (0.7)
Russia [32]	School	Children (6–14 yrs)	1	4/1988 (0.2)
Spain [33]	Children from birth	Children (3 yrs)	1	7/830 (0.8)
Spain [34]	Care center	Children (1–14 yrs)	1	11/780 (1.4)
Spain [34]	Care center	Adults	1	10/3450 (0.3)
Sweden [35]	Healthy population	Adults	2	10/1894 (0.5)
Sweden [36]	Care center	Children (2.5 yrs)	1	9/690 (1.3)
Sweden [37]	School	Children (12 yrs)	1	212/7274 (2.9)
Tunisia [38]	School	Children (6–12 yrs)	3	42/6286 (0.6)
Turkey [39]	School	Children (6 mths–17 yrs)	1	7/1263 (0.6)
Turkey [40]	School	Children (6–17 yrs)	3	215/20190 (1.0)
United Kingdom [41]	General practice	Adults	1	85/7550 (1.1)
United Kingdom [42]	Healthy population	Children (7 yrs)	3	54/5470 (1.0)
United Kingdom [20]	Healthy population	Children (12–15 yrs)	1	17/1975 (0.9)
United Kingdom [20]	Healthy population	Adults	1	69/4656 (1.5)
USA [43]	School	Children (6–18 yrs)	1	4/1281 (0.3)
USA [43]	Healthy population	Adults	1	27/2845 (0.9)
Total				1269/122858 (1.0)

* Diagnosis of celiac disease was defined as: 1 = serology positive for celiac disease (anti-transglutaminase antibodies) and histology compatible with celiac disease; 2 = serology positive for celiac disease (anti-endomisium antibodies antibodies) and histology compatible with celiac disaese; 3 = anti-endomisium antibodies positives.

3.2. Geo-epidemiology of CD

Overall, the worldwide prevalence of CD among low-risk (general population) adults and children ranged from 0 to 5.6%, with a mean prevalence ratio of 1% (Supplemental Figure S2). In Europe the overall prevalence of CD was 1%, with large variations between countries [20]. Similar rates have been reported from the US population (0.9%) [42], from many South American countries [14,16], and from developed countries mostly populated by individuals of European origin, e.g., Australia (0.4%) and New Zealand (1.1%) [15,30]. In areas of the developing world, rates overlapping European figures, especially in North Africa (i.e., 0.5% in Egypt, 0.8% in Libya, and 0.6% in Tunisia) [17,27,37], Middle East (i.e., 0.8% in Iran and 1% in Turkey) [23–25,39,40], and India (i.e., 0.4%) [21,22] were found. The Saharawi population of Arab-Berber origin living in Algeria had the highest prevalence of CD (5.6%) among all world populations [13]. In Burkina Faso, wheat consumption is almost negligible, and the prevalence of CD was null [17]. No CD prevalence studies were found from China or other Far Eastern countries.

3.3. Wheat consumption

Data on wheat consumption were available for 150 countries. In general, wheat consumption ranged between 21 and 564 g per capita/day (mean = 260) (Supplemental Figure S3).

3.4. HLA-DQ2 and DQ8 frequency

Data on HLA-DQ were available for 28 (DQ2) and 15 (DQ8) countries, respectively. Frequencies of HLA DQ2 and DQ8 ranged between 0 and 28% (mean = 14) and between 1 and 9% (mean = 4), respectively (Supplemental Figures S4 and S5).

3.5. CD prevalence, wheat consumption and HLA-DQ2 and–DQ8 frequency

The prevalence of CD and the frequency of DQ2 or DQ8 or DQ2 plus DQ8 were not significantly correlated (Rho = 0.4, $p = 0.07$; Rho = 0.3, $p = 0.4$). Likewise, there was no significant correlation between the nationwide prevalence of CD and the level of wheat consumption (Rho = 0.4, $p = 0.06$), and between the prevalence of CD and the product of DQ2 + DQ8 frequency * wheat consumption (Rho = 0.3, $p = 0.3$). On the other hand, a highly significant correlation between wheat consumption and the frequency of either HLA DQ2 (Fig. 1) or the sum of DQ2 and DQ8 (Fig. 2) was found. There was a trend of positive correlation between wheat consumption and HLA-DQ8 frequency, although this did not reach statistical significance (Rho = 0.5, $p = 0.06$). A significant correlation between the frequency of HLA DQ2 and the duration of wheat consumption according to the pattern of wheat culture spreading was found (Fig. 3). In the same four areas, the duration of wheat

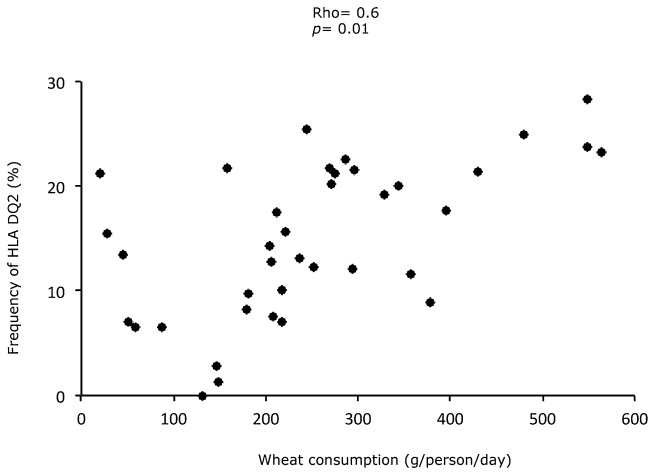

Fig. 1. Correlation between the level of wheat consumption and the frequency of HLA-DQ2 in different countries of the world.

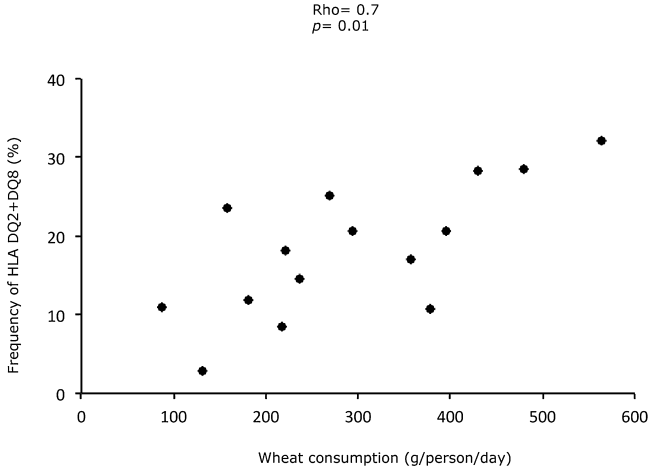

Fig. 2. Correlation between the level of wheat consumption and the sum of the frequencies of HLA DQ2 and DQ8 in different countries around the world.

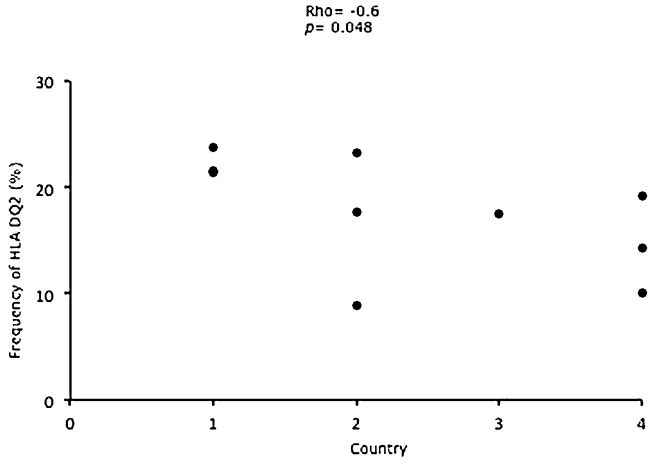

Fig. 3. Correlation between the frequency of HLA DQ2 and the duration of wheat consumption in four different areas of Middle East and Europe. Countries are grouped according to the duration of wheat consumption from 1 (gluten-containing cereals introduced around 10,000 years ago) to 4 (gluten-containing cereals introduced 6000–4000 years ago). 1: Turkey, Iran, Israel; 2: Tunisia, Italy, Greece; 3: Germany; 4: Sweden, Finland, Ireland.

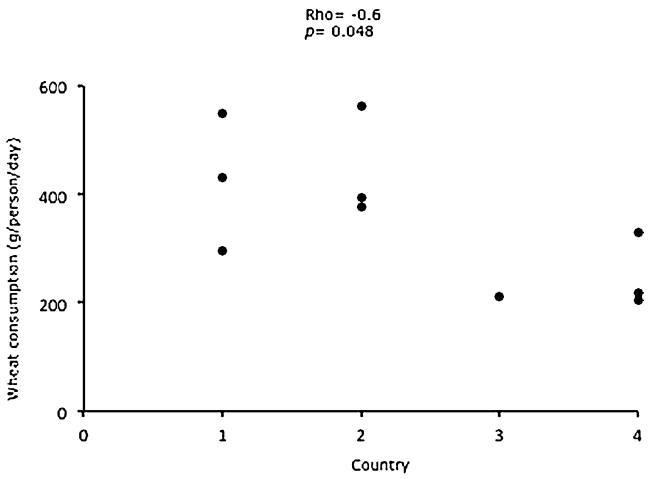

Fig. 4. Correlation between the current level of wheat consumption and the duration of wheat consumption in four different areas of Middle East and Europe. Countries are grouped according to the duration of wheat consumption from 1 (gluten-containing cereals introduced around 10,000 years ago) to 4 (gluten-containing cereals introduced 6000–4000 years ago). 1: Turkey, Iran, Israel; 2: Tunisia, Italy, Greece; 3: Germany; 4: Sweden, Finland, Ireland.

consumption was significantly correlated with the level of current wheat consumption (Fig. 4).

A detailed analysis of within-continent distribution showed that gluten and HLA-DQ2 co-segregated not only in Europe/Middle East, but also in other Continents. In India wheat consumption is higher in Northern-Western states such as Punjab, Haryana, Delhi, Rajasthan, Uttar Pradesh, Bihar and Madhya Pradesh (so-called "celiac belt"). Likewise, the DQ2 allele frequency is higher in North (31.9%) than in South India (12.78% for Piramalai Kallars and 9% for Yadavas) [44]. In Africa the frequency of DQ2 is higher in the Northern populations consuming a gluten-rich diet, such as Saharawi (39%) [45], Libya (34%) [46], Algeria (28.3%) [11], Tunisia (23.4%) [11], Morocco (25%) [11], than sub-Saharan populations showing low frequency of DQ2 (Rwanda 15.5%, Tanzania 13.5%, Cameroon 7%) [11] and low level of wheat consumption. The frequency of the DQB1*0201 allele is higher in the northern Chinese than in the southern Chinese populations, and the consumption of wheat in the northern area is higher than in the southern area, where rice is the staple diet [47].

The multivariable linear regression analysis did not show any significant effect of predisposing genes and wheat consumption on the prevalence of CD.

4. Discussion

Our study shows that CD is a common condition in different continents, including many developing areas, with a mean prevalence of 0.9% worldwide. We did not find a significant correlation between the prevalence of CD and predisposing HLA genes, a finding that can be explained, at least in part, by a certain degree of data heterogeneity. For example, the Australian and New Zealand HLA-DQ allele frequencies derived from individuals of Aboriginal and Maori origin, while studies in the same countries on CD prevalence were performed in individuals of prominently European origin. Although a trend of positive correlation between the prevalence of CD and the level of wheat consumption was identified, this also did not reach statistical significance, probably due to non-linear correlation between the level of gluten intake and the risk of CD development. A dose–effect relationship between gluten intake and development of CD has been previously shown, but only at very low gluten intakes (ranging between 50 and 500 mg of daily gluten) [48]. Although the worldwide variation of wheat

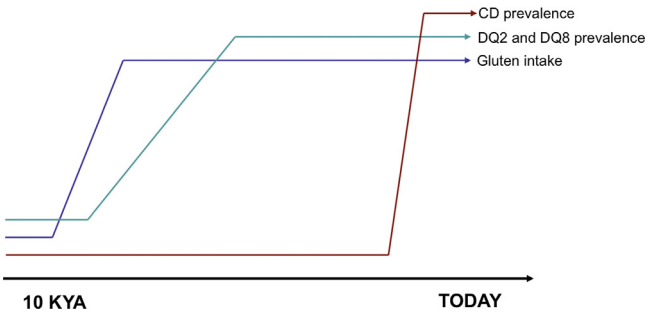

Fig. 5. The chain of events eventually leading to the current high prevalence of celiac disease: (1) increasing gluten intake in the early era of the agriculture revolution in Middle East and then Europe; (2) increased prevalence of HLA-DQ2 due to demic population diffusion and/or positive selection caused by adverse effects of wheat consumption on health, e.g. dental caries; (3) in modern time, e.g. the last centuries, other environmental changes triggered the surge of celiac disease in the general population worldwide (length, height and slant of the lines are arbitrary). KYA, kiloyear (thousand years); CD, Celiac disease.

consumption is high, the current intake is well above 500 mg per day in most countries, therefore minimizing the effect that the level of gluten intake may have on the prevalence of CD.

The analysis of correlation between wheat consumption, HLA-DQ2 and DQ8 frequency and CD prevalence had already been performed by Abadie and coworkers in their recent review paper [49]. However our work took this analysis a step forward and showed a previously unreported strong correlation between the geographical distribution of HLA CD-predisposing genotypes and wheat consumption. The paradoxical distribution of these two CD-causative factors is best exemplified by findings in the Europe/Middle East area. Domestication of gluten-containing cereals (particularly ancient varieties of wheat and barley) began approximately 10,000 years ago in Neolithic settlements in the north-eastern (Turkey, Iran, and Iraq) and in the south-western (Palestine, Syria, and Lebanon) regions of the so-called "Fertile Crescent" area. Cultivation of wheat and barley slowly spread northwest ward across Europe to reach Western countries, as Britain, "only" 4000 years ago [10]. The agriculture revolution paved the way to new diseases, such as CD. Since treatment with the gluten-free diet was not available in the ancient time, the CD geno/phenotype exerted a negative impact on human reproductive fitness and survival over thousand years. As a consequence of a prolonged negative selective pressure, CD and its predisposing genotypes should be less common in Middle East than in European countries (as originally theorized by Simoons) [4]. On the contrary, nowadays CD is equally common in UK and Turkey, and the frequency of HLA-DQ2 (the major CD-predisposing haplotype) is even higher in Turkey and Iran, i.e. countries consuming more wheat and for a longer duration of time, than in Finland and Ireland, i.e. countries consuming less gluten and for a shorter duration of time. This "evolutionary celiac paradox" is strengthened by the co-localization of gluten consumption and HLA-DQ2 we found in other areas of the world, e.g. Africa and India. To explain these counterintuitive findings, we suggest that the diffusion of HLA CD-predisposing genotypes and the diffusion of CD occurred in different times (Fig. 5). The currently high prevalence of CD is most likely a recent event, explained by recent changes in environmental conditions, e.g. related to the quantity and quality of ingested gluten, dough fermentation, human intestinal microbioma and pattern of infant nutrition [7,50].

On the other hand, the parallel geographical distribution of gluten intake and HLA-DQ2 and DQ8 genotypes could be explained by different mechanisms, which are not mutually exclusive:

(a) a founder effect reflecting higher prevalence of DQ2/DQ8 in the early farmers that 10,000 years ago spread from Middle East and gradually colonized Europe determining a South/East–North/West genetic gradient (so called "demic diffusion" model of Cavalli-Sforza and Ammerman) between Middle East and Europe [10] (local effect). This slow "wave of advance" of a population of early farmers with a high frequency of DQ2 and DQ8 could lead to progressive dilution of their genetic background by mixing with the local European populations;
(b) a positive selection of HLA CD-predisposing genotypes pushed by protection against deleterious conditions driven by the agricultural revolution (worldwide effect). An extensive search of the literature showed that only one disease is partially "protected" by an HLA-DQ, CD-predisposing haplotype (DQ2), that is dental caries. A recent Brazilian study showed that adolescents positive for HLA-DQ2 allele were less likely to have dental caries than their counterparts who were negative for this allele (OR = 0.33, CI 0.16–0.66) [51]. Nothing was found about disease protection induced by the HLA-DQ8 haplotype.

Dental caries is the typical Neolithic disease that dramatically increased after wheat carbohydrates acquired a primary role in human diet [52]. Interestingly, also CD patients show a reduced prevalence of dental caries when compared to controls, despite high incidence of dental enamel hypoplasia [53]. *Streptococcus mutans*, the principal agent responsible for oral fermentation of sugars leading to dental caries, is less abundant in the intestinal mucosa of CD children, regardless of inflammatory status [54]. The consumption of wheat is positively correlated to the development of dental caries whereas rice, the typical gluten-free cereal, shows less cariogenic potential [55]. In the context of the ancestral wheat-consuming society, when cure and prevention of dental caries were basically non-existent, some sort of genetically determined resistance to dental caries could well drive a positive selection of HLA-DQ2. It remains to be elucidated how DQ2 positivity protects from dental caries development, e.g. by accelerating the clearance of "sticky" (and possibly cariogenic) gluten peptides from the mouth and teeth surface. The DQ2-dental caries protection hypothesis needs to find confirmation by further population studies.

Our (b) explanation is more generalizable than (a), and could explain the high gluten/high DQ2 co-localization in different, apparently unrelated, areas of the world such as Europe, Africa and India. Thus, the highest prevalence of DQ2 among wheat consuming populations could be another example of evolutionary adaptation driven by the dietary changes imposed by the agricultural revolution, in analogy to lactase persistence in milk drinking populations [56] and glucose-6-phosphate deficiency (G6PD) in Mediterranean populations consuming broad beans [57]. In this context CD could simply represent the modern "dark side" of a genetic structure (DQ2 positivity) conferring resistance to a negative consequence of a wheat-based diet.

We are aware of the limitations of our geo-epidemiological analysis: (a) current wheat consumption does not perfectly reflect historical wheat consumption (although both historical consideration and our data suggest that a strong correlation does indeed exist); (b) FAO data reflect individual availability of wheat (including unconsumed leftover) instead of individual wheat intake; (c) heterogeneity of HLA data does not allow to differentiate high-risk (e.g. double copy of the DQB1*02 gene) from lower risk genotypes; (d) CD prevalence data are based on different diagnostic methods and different age groups; (e) regional differences in the consumption of CD-triggering grains different from wheat (rye and barley) have not been investigated; (f) the possible evolutionary role of non-HLA predisposing genes has not been evaluated. At least 39 non-HLA CD predisposing genes have been identified, most of them influencing the inflammatory and immune pathways.

Recent studies showed consistent signs of positive selection for CD-associated non HLA alleles, i.e. IL12A, IL18RAP, and SH2B3 [58].

In conclusion, we suggest that the current high prevalence of CD is the last link in a chain of events started about 10,000 years ago after wheat domestication and diffusion from the Middle East. To explain the so-called "evolutionary celiac paradox" of co-localization of gluten consumption and HLA CD predisposing genotypes, we suggest a primary role for positive selection of the major CD-predisposing haplotype, that is HLA-DQ2, in wheat-consuming populations. The force driving HLA-DQ2 selection could have been protection against dental caries, a disease causing a tremendous negative impact on the health and reproductive fitness of ancient populations. The "DQ2-dental caries protection" hypothesis needs to find confirmation by further population studies.

Conflict of interest

Carlo Catassi served as consultant for Menarini diagnostics s.r.l., for Dr Shaer, and for Heinz Company. Elena Lionetti served as consultant for Heinz Company.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.dld.2014.08.002.

References

[1] Fasano A, Catassi C. Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum. Gastroenterology 2001;120:636–51.
[2] Lionetti E, Catassi C. New clues in celiac disease epidemiology, pathogenesis, clinical manifestations, and treatment. International Reviews of Immunology 2011;30:219–31.
[3] Fasano A, Catassi C. Clinical practice. Celiac disease. New England Journal of Medicine 2012;367:2419–26.
[4] Simoons FJ. Coeliac disease as a geographic problem. In: Walcher DN, Kretchmer N, editors. Food, nutrition and evolution. New York: Masson; 1981. p. 179–99.
[5] Catassi C. Editorial Celiac disease in Turkey: lessons from the fertile crescent. American Journal of Gastroenterology 2011;106:1518–20.
[6] Barada K, Abu Daya H, Rostami K, et al. Celiac disease in the developing world. Gastrointestinal Endoscopy Clinics of North America 2012;22:773–96.
[7] Catassi C, Kryszak D, Bhatti B, et al. Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. Annals of Medicine 2010;42:530–8.
[8] Aaron L. The last two millennias echo-catastrophes are the driving forces for the potential genetic advantage mechanisms in celiac disease. Medical Hypotheses 2011;77:773–6.
[9] Sams A, Hawks J. Patterns of population differentiation and natural selection on the celiac disease background risk network. PLOS ONE 2013;8:e70564.
[10] Ammerman AJ, Cavalli-Sforza LL. The Neolithic transition and the genetics of populations in Europe. Princeton: Princeton University Press; 1984.
[11] Gonzalez-Galarza FF, Christmas S, Middleton D, et al. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. Nucleic Acids Research 2011;39:D913–9.
[12] Cochrane Consumers and Communication Review Group. Data Extraction Template for Cochrane Reviews. http://www.latrobe.edu.au/chcp/assets/downloads/DET_2011.doc
[13] Catassi C, Rätsch IM, Gandolfi L, et al. Why is coeliac disease endemic in the people of Sahara? Lancet 1999;354:647–8.
[14] Gomez JC, Selvaggio GS, Viola M, et al. Prevalence of CD in Argentina: screening of an adult population in the La Plata area. American Journal of Gastroenterology 2001;96:2700–4.
[15] Hovell CJ, Collett JA, Vautier G, et al. High prevalence of coeliac disease in a population-based study from Western Australia: a case for screening? Medical Journal of Australia 2001;175:247–50.
[16] Pratesi R, Gandolfi L, Garcia SG, et al. Prevalence of coeliac disease: unexplained age-related variation in the same population. Scandinavian Journal of Gastroenterology 2003;38:747–50.
[17] Cataldo F, Lio D, Simpore J, et al. Consumption of wheat foodstuffs not a risk for celiac disease occurrence in Burkina Faso. Journal of Pediatric Gastroenterology and Nutrition 2002;35:233–4.
[18] Abu-Zekry M, Kryszak D, Diab M, et al. Prevalence of CD in Egyptian children disputes the east-west agriculture-dependent spread of the disease. Journal of Pediatric Gastroenterology and Nutrition 2008;47:136–40.
[19] Mäki M, Mustalahti K, Kokkonen J, et al. Prevalence of Celiac disease among children in Finland. New England Journal of Medicine 2003;348:2517–24.
[20] Mustalahti K, Catassi C, Reunanen A, et al. The prevalence of CD in Europe: results of a centralized, international mass-screening project. Annals of Medicine 2010;42:587–95.
[21] Sood A, Midha V, Sood N, et al. Prevalence of CD among school children in Punjab, North India. Journal of Gastroenterology and Hepatology 2006;21:1622–5.
[22] Bhattacharya M, Dubey AP, Mathur NB. Prevalence of celiac disease in north Indian children. Indian Pediatrics 2009;46:415–7.
[23] Akbari MR, Mohammadkhani A, Fakheri H, et al. Screening of the adult population in Iran for coeliac disease: comparison of the tissue-transglutaminase antibody and anti-endomysial antibody tests. European Journal of Gastroenterology and Hepatology 2006;18:1181–6.
[24] Saberi-Firouzi M, Omrani GR, Nejabat M, et al. Prevalence of celiac disease in Shiraz, southern Iran. Saudi Journal of Gastroenterology 2008;14:135–8.
[25] Farahmand F, Mir-Nasseri MM, Shahraki T, et al. Prevalence of occult celiac disease in healthy Iranian school age children. Archives of Iranian Medicine 2012;15:342–5.
[26] Johnston SD, Watson RG, McMillan SA, et al. Prevalence of coeliac disease in Northern Ireland. Lancet 1997;350:1370.
[27] Tommasini A, Not T, Kiren V, et al. Mass screening for coeliac disease using antihuman transglutaminase antibody assay. Archives of Disease in Childhood 2004;89:512–5.
[28] Alarida K, Harown J, Ahmaida A, et al. Coeliac disease in Libyan children: a screening study based on the rapid determination of anti-transglutaminase antibodies. Digestive and Liver Disease 2011;43:688–91.
[29] Csizmadia CG, Mearin ML, von Blomberg BM, et al. An iceberg of childhood coeliac disease in the Netherlands. Lancet 1999;353:813–4.
[30] Cook HB, Burt MJ, Collett JA, et al. Adult coeliac disease: prevalence and clinical significance. Journal of Gastroenterology and Hepatology 2000;15:1032–6.
[31] Antunes H, Abreu I, Nogueiras A, et al. First determination of the prevalence of celiac disease in a Portuguese population. Acta Medica Portuguesa 2006;19:115–20.
[32] Kondrashova A, Mustalahti K, Kaukinen K, et al. Lower economic status and inferior hygienic environment may protect against celiac disease. Annals of Medicine 2008;40:223–31.
[33] Castaño L, Blarduni E, Ortiz L, et al. Prospective population screening for celiac disease: high prevalence in the first 3 years of life. Journal of Pediatric Gastroenterology and Nutrition 2004;39:80–4.
[34] Mariné M, Farre C, Alsina M, et al. The prevalence of coeliac disease is significantly higher in children compared with adults. Alimentary Pharmacology and Therapeutics 2011;33:477–86.
[35] Ivarsson A, Persson LA, Nyström L, et al. Epidemic of coeliac disease in Swedish children. Acta Paediatrica 2000;89:165–71.
[36] Carlsson AK, Axelsson IE, Borulf SK, et al. Serological screening for celiac disease in healthy 2.5-year-old children in Sweden. Pediatrics 2001;107:42–5.
[37] Myléus A, Ivarsson A, Webb C, et al. Celiac disease revealed in 3% of Swedish 12-year-olds born during an epidemic. Journal of Pediatric Gastroenterology and Nutrition 2009;49:170–6.
[38] Ben Hariz M, Kallel-Sellami M, Kallel L, et al. Prevalence of celiac disease in Tunisia: mass-screening study in school children. European Journal of Gastroenterology and Hepatology 2007;19:687–94.
[39] Ertekin V, Selimoglu MA, Kardas F, et al. Prevalence of CD in Turkish children. Journal of Clinical Gastroenterology 2005;39:689–91.
[40] Dalgic B, Sari S, Basturk B, et al. Prevalence of celiac disease in healthy Turkish school children. American Journal of Gastroenterology 2011;106:1512–7.
[41] West J, Logan RF, Hill PG, et al. Seroprevalence, correlates, and characteristics of undetected coeliac disease in England. Gut 2003;52:960–5.
[42] Bingley PJ, Williams AJ, Norcross AJ, et al. Undiagnosed coeliac disease at age seven: population based prospective birth cohort study. British Medical Journal 2004;328:322–3.
[43] Fasano A, Berti I, Gerarduzzi T, et al. Prevalence of CD in at-risk and non at-risk groups: a large, multicentre study. Archives of Internal Medicine 2003;163:286–92.
[44] Shanmugalakshmi S, Balakrishnan K, Manoharan K, et al. HLA-DRB1*, -DQB1* in Piramalai Kallars and Yadhavas, two Dravidian-speaking castes of Tamil Nadu, South India. Tissue Antigens 2003;61:451–64.
[45] Catassi C, Doloretta Macis M, Ratsch IM, et al. The distribution of DQ genes in the Saharawi population provides only a partial explanation for the high CD prevalence. Tissue Antigens 2001;58:402–6.
[46] Alarida K, Harown J, Di Pierro MR, et al. HLA-DQ2 and -DQ8 genotypes in celiac and healthy Libyan children. Digestive and Liver Disease 2010;42:425–7.
[47] Yuan J, Gao J, Li X, et al. The tip of the celiac iceberg in China: a systematic review and meta-analysis. PLOS ONE 2013;8:e81151.
[48] Catassi C, Rossini M, Rätsch IM, et al. Dose dependent effects of protracted ingestion of small amounts of gliadin in coeliac disease children: a clinical and jejunal morphometric study. Gut 1993;34:1515–9.
[49] Abadie V, Sollid LM, Barreiro LB, et al. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. Annual Review of Immunology 2011;29:493–525.
[50] Catassi C, Anderson RP, Hill ID, et al. World perspective on celiac disease. Journal of Pediatric Gastroenterology and Nutrition 2012;55:494–9.
[51] Valarini N, Maciel SM, Moura SK, et al. Association of dental caries with HLA Class II allele in Brazilian adolescents. Caries Research 2012;46:530–5.
[52] Larsen CS. The agricultural revolution as environmental catastrophe: implications for health and lifestyle in the Holocene. Quaternary International 2006;150:12–20.
[53] Pastore L, Carroccio A, Compilato D, et al. Oral manifestations of celiac disease. Journal of Clinical Gastroenterology 2008;42:224–32.

[54] Sánchez E, Donat E, Ribes-Koninckx C, et al. Duodenal-mucosal bacteria associated with celiac disease in children. Applied and Environment Microbiology 2013;79:5472–9.
[55] Sreebny LM. Cereal availability and dental caries. Community Dentistry and Oral Epidemiology 1983;11:148–55.
[56] Tishkoff SA, Reed FA, Ranciaro A, et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nature Genetics 2007;39:31–40.
[57] Mehta A, Mason PJ, Vulliamy TJ. Glucose-6-phosphate dehydrogenase deficiency. Baillière's Best Practice and Research Clinical Haematology 2000;13:21–38.
[58] Zhernakova A, Elbers CC, Ferwerda B, et al. Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. American Journal of Human Genetics 2010;86:970–7.