Prevalence of Adult Celiac Disease in India: Regional Variations and Associations

B.S. Ramakrishna, DM, PhD[1], Govind K. Makharia, DM[2], Kamal Chetri, DM[3], Sangitanjan Dutta, DM[4], Prashant Mathur, MD[5], Vineet Ahuja, DM[2], Ritvik Amarchand, MBBS, PhD[6], Ramadass Balamurugan, PhD[1], Sudipta D. Chowdhury, DM[1], Dolly Daniel, MD[7], Anup Das, MD[4], Gemlyn George, MBBS[1], Siddhartha Datta Gupta, MD[8], Anand Krishnan, MD[6], Jasmin H. Prasad, MD[9], Gurvinder Kaur, PhD[10], Srinivasan Pugazhendhi, PhD[1], Anna Pulimood, MD, PhD[1], Kartik Ramakrishna, MBBS[1] and Anil K. Verma, MSc[1]

OBJECTIVES: **Although celiac disease (CeD) affects 1% of people in the northern part of India, it is believed to be uncommon in the southern and northeastern parts because of significant differences in dietary pattern and ethnicity. We estimated the prevalence of CeD in these three populations. In a subset, we also investigated differences in the prevalence of HLA-DQ 2/8 allelotype and dietary grain consumption.**

METHODS: **A total of 23,331 healthy adults were sampled from three regions of India—northern (*n*=6207), northeastern (*n*=8149), and southern (*n*=8973)—and screened for CeD using IgA anti-tissue transglutaminase antibody. Positive tests were reconfirmed using a second ELISA. CeD was diagnosed if the second test was positive and these participants were further investigated. A subsample of participants was tested for HLA-DQ2/-DQ8 and underwent detailed dietary evaluation.**

RESULTS: **Age-adjusted prevalence of celiac autoantibodies was 1.23% in northern, 0.87% in northeastern, and 0.10% in southern India (*P*<0.0001). Prevalence of CeD and latent CeD, respectively, was 8.53/1,000 and 3.70/1,000 in northern, 4.66/1,000 and 3.92/1,000 in northeastern, and 0.11/1,000 and 1.22/1,000 in the southern part. The population prevalence of genes determining HLA-DQ2 and/or -DQ8 expression was 38.1% in northern, 31.4% in northeastern, and 36.4% in southern India. Mean daily wheat intake was highest in northern (455 g) compared with northeastern (37 g) or southern part (25 g), whereas daily rice intake showed an inverse pattern.**

CONCLUSIONS: **CeD and latent CeD were most prevalent in northern India and were the least in southern India. The prevalence correlated with wheat intake and did not reflect differences in the genetic background.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

Am J Gastroenterol 2016; 111:115–123; doi:10.1038/ajg.2015.398; published online 5 January 2016

INTRODUCTION

Celiac disease (CeD), caused by hypersensitivity to gluten present in wheat and related grains, affects ~1% of the population of Europe and North America (1–5). There are major geographic differences in the prevalence of CeD, with the disease being very frequent in the Saharawi people of Africa and very uncommon in Japan and South East Asia (6–9). CeD has been recognized in northern India, primarily in children, since the 1960s (10,11). Active case finding has previously determined that the prevalence of CeD in children was ~1 in 300 school children and 1 in 100 adults in northern India (12,13). Reports of CeD from southern India are scarce (14).

There are two requisites for the development of CeD in a population: a pool of individuals who are capable of expressing the human leukocyte antigen (HLA)-DQ antigens 2 or 8 (1,15), and consumption of wheat, which is the major gluten-containing grain

[1]Department of Gastrointestinal Sciences, Christian Medical College, Vellore, India; [2]Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences, New Delhi, India; [3]Department of Gastroenterology, International Hospital, Guwahati, India; [4]Department of Medicine, Guwahati Medical College, Guwahati, India; [5]Indian Council of Medical Research, New Delhi, India; [6]Department of Community Medicine, All India Institute of Medical Sciences, New Delhi, India; [7]Department of Transfusion Medicine, Christian Medical College, Vellore, India; [8]Department of Pathology, All India Institute of Medical Sciences, New Delhi, India; [9]Department of Community Health, Christian Medical College, Vellore, India; [10]Department of Transfusion Medicine and Immunohematology, All India Institute of Medical Sciences, New Delhi, India. **Correspondence:** B.S. Ramakrishna, DM, PhD, Department of Gastrointestinal Sciences, Christian Medical College, Vellore, India. E-mail: wurama@hotmail.com

Received 9 June 2015; accepted 10 November 2015

associated with the development of CeD (1,16). Geographic differences in CeD prevalence are explained on the basis of differences in HLA-DQ background and in wheat consumption. The allelotypes determining HLA-DQ2 or HLA-DQ8 are present in about 15–30% of the population in different parts of India (17–19). There are differences in wheat consumption between different regions of India, with this grain being predominantly consumed in northern India.

We hypothesized that CeD would be very low in southern India compared with northern India, likely related to differences in dietary grain consumption pattern and HLA-DQ2/HLA-DQ8 prevalence between these populations. These studies aimed to determine the true prevalence of CeD in northern and southern India, represented by Haryana and Tamil Nadu, respectively. As northeastern India includes a large proportion of people with an Indo-Chinese genetic background, a representative population from Assam was included for comparison.

The objectives of the study were to estimate the true prevalence of IgA antibodies for tissue transglutaminase, a sensitive and specific marker of latent and clinical CeD, in three populations in southern, northern, and northeastern parts of India, determine whether prevalence of the HLA-DQ 2/8 genotype differed among these three populations, and determine dietary grain consumption patterns in these three populations.

METHODS

The population prevalence of CeD was estimated in each of the three sites by sampling selected rural and urban areas in Haryana (close to New Delhi), Assam, and Tamil Nadu, respectively, with a rural:urban sampling ratio approximating the population distribution in this area during the 2001 Census of India (20).

Study areas

The northern site comprised 29 villages and 27 urban wards in Ballabgarh block in Faridabad district of Haryana state. The northeastern site comprised 29 urban wards of Guwahati city and 42 villages in the Kamrup and Sonitpur districts of Assam state. The southern site comprised 123 villages within a 25 km radius around the city and 18 wards within the city of Vellore in Tamil Nadu state.

Selection and recruitment of participants

The villages and urban blocks in the selected areas were listed and the population enumerated from local panchayat or municipal records. The number of households to be enrolled was calculated on the basis of the target sample size. The study team visited the target areas and held focus group discussions with community leaders and people's groups. The scope and purpose of the study was explained and the study procedures explained. Promotional material concerning the study was developed in the local language and posters and banners displayed prominently to give publicity, while pamphlets were distributed to the selected households. Eligible households included those with at least one male and one female adult. Two unrelated individuals over the age of 18 years were targeted for recruitment from each selected household. Genetically related first-degree relatives of an index participant were excluded. Households were selected as every "n"th household and respondents selected using the Kish grid (21). Selected households were visited by the field team assigned to the area, and two unrelated adults in each household were requested to participate and informed consent obtained. Participant demographic and other details were collected using predesigned forms. As the socioeconomic status of an individual may potentially influence the development of CeD, the socioeconomic status of participants was assessed and scored using the Kuppuswami method (22).

Height and weight were recorded. Participants were reviewed by the physician member of the team. The individual's ID was confirmed, and 8 ml of blood was drawn into labeled Vacutainer tubes for serum and EDTA-treated blood. Supervisory visits and random checks were employed to check the enrollment of participants, blood collection, and filling of questionnaire. Meetings were conducted and the data reviewed at regular intervals.

Laboratory processing and further testing

Blood samples were transported to the laboratories on ice in cold boxes. Serum was separated from blood by centrifugation on the same day and stored in labeled 1.5 ml centrifuge tubes in aliquots. One aliquot was used for estimation of anti-human tissue transglutaminase IgA antibody (anti-tTG Ab), whereas a duplicate aliquot was stored at −80 °C. DNA was isolated from whole blood by means of the salting out procedure, quantitated using fluorescence (Qubit DS DNA BR Assay Kit, on a Qubit 2.0 Fluorometer, Invitrogen, Bangalore, India) or spectrometry (NanoDrop 2000, Thermo Scientific, Wilmington, DE), and stored at −20 °C until analysis. A sample log was maintained.

Testing for anti-tTG Ab

All blood samples were screened for anti-tTG Ab using the Aeskulisa CELICHEK tTg-A New Generation kits (Catalog No. 3503, Aesku Diagnostics Gmbh, Wendelsheim, Germany). A single lot (Lot No. 11430) was used for all three study sites, and testing was carried out according to the manufacturer's instructions. Test values below 12 U/ml were considered negative, 12–18 U/ml indeterminate, and values >18 U/ml were considered positive. Recombinant human tTG incorporating gliadin peptide moieties capture anti-tTG in this assay, making it highly sensitive for the diagnosis of CeD (23). All samples that were positive in the first assay were tested with a second enzyme-linked immunosorbent assay (ELISA; QUANTA Lite h-tTg IgA ELISA, Catalog No. 708760, Inova Diagnostics, San Diego, CA) that used purified human erythrocyte tTG to capture and measure anti-tTG Ab. The latter is considered very specific for the diagnosis of CeD, particularly in the setting of a high pre-test probability of the sample being from an individual with CeD (23). In this test, values <20 U/ml were considered negative, values 20–30 U/ml were considered equivocal, and those >30 U/ml were considered positive.

Quality standardization

Laboratories used quality control for the automated tests. ELISA for anti-tTG Ab had quality controls in each kit, which performed

as per expectation. Quality control for the anti-tTG Ab testing between centers was achieved by exchanging 200 samples of serum between each center for testing using the Aesku screening ELISA kit. The concordance between the centers ranged from 79.5 to 94%. The observed unweighted kappa statistic for all three centers combined was 0.7824 (95% confidence interval (CI) 0.7324–0.8324).

Counseling and further evaluation of anti-tTG Ab-positive individuals

Individuals testing positive for the antibody in both ELISAs were counseled about the probability of having CeD and invited to undergo further testing. Those who consented were brought to hospital, clinically evaluated in detail, and interviewed by a dietician, using 24 h recall and food frequency questionnaires of commonly used foods ingested over the previous 3 months. Food quantities were measured according to a standard set of cups and spoons, and the daily intake of the relevant nutrients was calculated from standard tables for the composition of Indian foods (24). The dietary intake of wheat, rice, and other grains was quantified. Individuals who provided written, informed consent underwent gastroduodenoscopy as per usual clinical protocol, and biopsies were obtained from the second and third part of the duodenum, using non-spiked biopsy forceps, transferred to 10% buffered formalin and sent to the laboratory where the biopsies were oriented, processed to paraffin, and 4-micron hematoxylin and eosin-stained sections were examined by an experienced gastrointestinal pathologist. Mucosal changes were graded as per the modified Marsh-Oberhuber classification for CeD (25).

HLA testing for the genes expressing DQ2 and DQ8

The HLA-DQ locus was typed at low resolution for the presence of DQB1*02 and DQB1*0302 using polymerase chain reaction (PCR) sequence-specific oligonucleotide typing (Lifecodes HLA-DQB Typing Kit for use with Luminex, Product # 628610-50, GenProbe Transplant Diagnostics, Stamford, CT) following the manufacturer's instructions. The samples from the Haryana site were HLA typed at the All India Institute of Medical Sciences, New Delhi, whereas the samples from the other two sites were HLA typed at the Christian Medical College, Vellore. In cases where the internal controls did not work satisfactorily, PCR and sequencing was used to identify the genotype. In such cases, exon 2 of the DQB1 gene on chromosome 6 was amplified using the following primers (26).

Forward: 5′-GCCGGTGATTCCCCGCAGAGGATTTCG-3′;
Reverse: 5′-GGAGGGGCGACGACGCTCACCTC-3′.

The PCR reaction mix contained 1×Taq DNA Polymerase Master Mix Red with 0.2 units/μl Ampliqon Taq DNA polymerase, NH_4^+ buffer, 0.4 mM dNTPs, 1.5 mM $MgCl_2$, 200 nM of each primer, and 1.6 M betaine. PCR was undertaken in a gradient thermal cycler (Master Cycler Gradient, Eppendorf, Hamburg, Germany) with initial denaturation at 95 °C for 5 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 65 °C for 30 s, initial extension at 72 °C for 30 s, and final extension at 72 °C for 10 min. PCR products were purified (QiagenPure Link PCR Purification Kit, Invitrogen, Carlsbad, CA) and bidirectionally sequenced using BigDye Terminator chemistry in a 3730xl DNA Analyzer (Life Technologies Corporation, Grand Island, NY). Sequences were trimmed, checked for quality control using FinchTV version 1.5 (http://www.geospiza.com/ftvdlinfo.html), and genotyped using the IMGT/HLA sequence alignment tool (http://www.ebi.ac.uk/ipd/imgt/hla/align.html).

Sample size

The sample size calculation was guided by published population CeD prevalence of 1.44% (95% CI 1.22–1.69) in northern India (13). For the northern state, a population prevalence of 1% with relative accuracy of 25% was assumed, leading to a sample of 6,336 individuals; assuming 10% refusal, the target recruitment was 6,970 individuals. For the presumed low prevalence states in southern and northeastern India, assumed prevalence of 0.2%, a relative accuracy of 50%, and refusal rate of 10% were used to calculate a target sample size of 8,782 in each site. Sample size calculations for HLA-DQ typing were based on earlier reported HLA-DQ*B1 0201* gene or allele frequencies (17,18). In order to detect a 50% difference in allele frequency between sites with 90% power and alpha error of 0.05, the calculated sample size was 120 participants per site.

Definitions

CeD was defined as the presence of anti-tTG Ab with positive Aesku ELISA with titer >18 U/ml with subsequent positive Inova ELISA >30 U/l. Participants who were antibody positive were categorized as classical (gastrointestinal symptoms plus Marsh grade ≥2 change), non-classical (extraintestinal symptoms plus Marsh grade ≥2 change), asymptomatic (no symptoms, Marsh grade ≥2 change), or latent (no symptoms, no mucosal change) (27). Categories 1–3 have histological or clinical abnormalities, whereas category 4 does not have either.

Statistics

Crude prevalence of CeD was defined as the number of participants with CeD as a proportion of the total population tested. Adjusted prevalence of CeD was defined as prevalence adjusted for age distribution. The 95% CIs were calculated and presented for prevalence data. Bivariate analysis of dichotomous variables was performed after entering the data in SAS. Pearson's χ^2 test was applied and *P*-values calculated. Pearson's χ^2 test was then conducted on each pair of groups and the *P*-value adjusted for multiple comparisons using the Bonferroni correction. In the case of continuous variables, analysis of variance was used to compare multiple groups with *post hoc* comparison of two groups using the Bonferroni test. Associations between CeD prevalence and sociodemographic variables were computed and shown as odds ratios with 95% CIs.

Ethics

The study protocol was approved by the institutional ethics committees of the All India Institute of Medical Sciences, New Delhi, the Christian Medical College, Vellore, and the Guwahati Medical College, Guwahati.

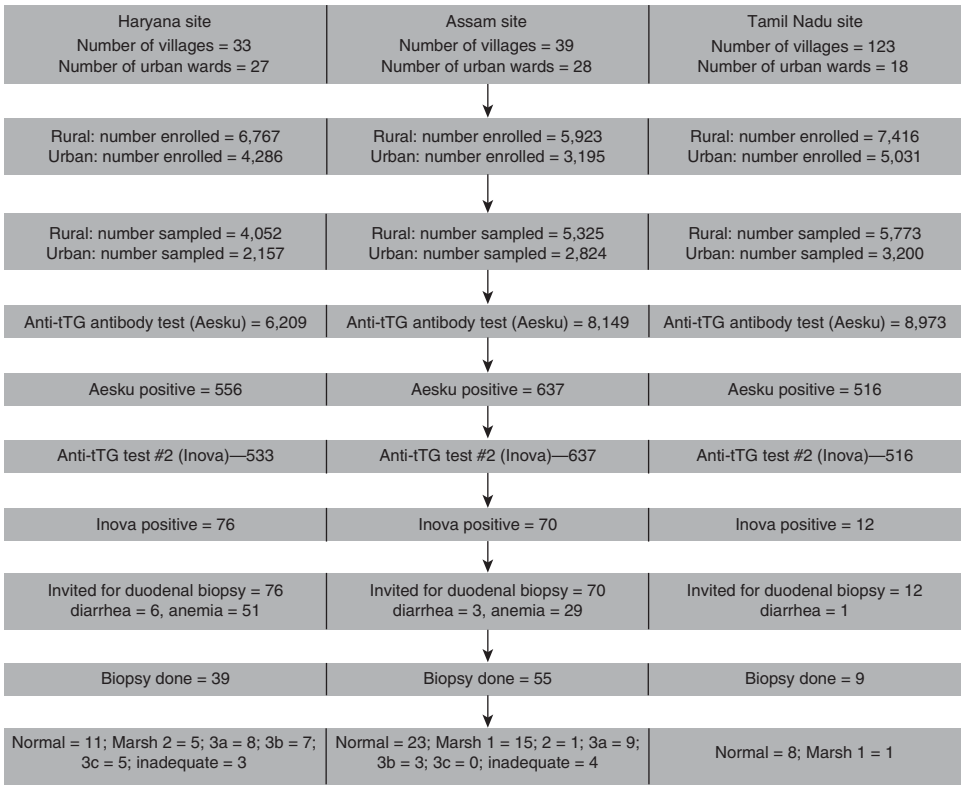

Figure 1. Flow of participants through the study at the three study sites.

RESULTS

Participant flow

Participant recruitment in the northern site was from October 2011 to August 2013, in the northeastern site from January 2012 to August 2013, and in the southern site from April 2011 to October 2012. Participant flow in these sites is shown in **Figure 1**. The demographic characteristics of the participants in relation to the presence of a positive screening and confirmatory test for anti-tTG Ab are shown in **Table 1** and **Supplementary Table S1** online.

Prevalence of CeD

The number of individuals positive for anti-tTG Ab at each step of the study is shown in **Figure 1**. In order to understand the specificity of the second assay (Inova), we also tested 204 samples, which were negative in the first (Aesku) ELISA, using the Inova ELISA. Only 1 of these 204 Aesku-negative samples tested positive in the Inova assay, confirming the specificity of the ELISA kit from Inova. The crude and age-adjusted prevalences of CeD in each of the three regions are shown in **Supplementary Table S2**. The age-adjusted prevalence was 1.23/100 (95% CI 0.98–1.54) in the Haryana site, 0.87/100 (95% CI 0.67–1.10) in the Assam site, and 0.10/100 (95% CI 0.06–0.18) in the Tamil Nadu site ($P<0.0001$). Pairwise comparisons showed that the age-standardized prevalence of CeD in the southern site was highly significantly ($P<0.0001$) lower than the prevalence in the other two sites, but the difference in prevalence between Haryana and Assam was not statistically significant.

Rural–urban differences

There were no statistically significant rural–urban differences in the prevalence of CeD, either by site-specific analysis or by analysis of the clubbed prevalence across the three study sites. Ninety six of 15,150 (0.63%) rural participants had CeD compared with 62 of 8,181 (0.75%) urban participants ($P=0.4339$). Socioeconomic status was not significantly associated with CeD prevalence.

Age distribution of the participants and prevalence by age group

Supplementary Table S3 shows the age distribution of the tested individuals by study site, and the CeD prevalence in each age interval. No statistically significant difference was noted in the prevalence of CeD across the different age groups. CeD was significantly more frequent in females compared with males.

HLA genotyping in a subsample of the general (anti-tTG antibody negative) adult population

HLA-DQ2 and/or DQ8 genotypes were found in approximately one-third of the population in each site—38.1% in Haryana, 31.4% in Assam, and 36.4% in Tamil Nadu (**Table 2**). HLA-DQ2 allele frequency was similar between Haryana (34/220), Assam (75/452), and Tamil Nadu (70/494; $P=0.5861$). HLA-DQ8 allele frequency was significantly different between Haryana (12/220), Assam (10/452), and Tamil Nadu (49/494; $P<0.001$).

Table 1. Demographic variables of the populations surveyed in the three areas in relation to a positive antibody test for anti-tissue transglutaminase

Risk variables	Celiac serology				Total	*P* value
	Negative		Positive			
	n	%	*N*	%		
Centre						
Haryana	6,133	98.8	76	1.2	6,209	—
Assam	8,079	99.1	70	0.9	8,149	<0.001
Tamil Nadu	8,961	99.9	12	0.1	8,973	—
Sex						
Male	10,728	99.5	58	0.5	10,776	0.02
Female	12,455	99.2	100	0.8	12,555	—
Age (years)						
18–20	1,562	99.4	9	0.6	1,571	—
21–30	6,069	99.2	50	0.8	6,119	0.48
31–40	6,454	99.3	46	0.7	6,500	—
41–50	4,914	99.5	26	0.5	4,940	—
51–60	2,481	99.4	14	0.6	2,495	—
>60	1,697	99.2	13	0.8	1,710	—
Area						
Urban	8,119	99.2	62	0.8	8,181	0.27
Rural	15,054	99.4	96	0.6	15,150	—
SES score						
25–16	4,251	99.5	23	0.5	4,274	—
15–11	8,362	99.4	54	0.6	8,416	0.40
10–5	8,815	99.3	65	0.7	8,880	—
<5	1,709	99.1	15	0.9	1,724	—

SES, socioeconomic status.
SES score was calculated using Kuppuswami method (22).
Those who were positive in both screening and confirmatory tests were considered positive.

Grain consumption in the anti-tTG antibody-negative general population

The daily intake of wheat and rice is shown in **Table 3**. Wheat intake was uniformly high in the Haryana cohort and much lower in the Assam and Tamil Nadu cohorts. Energy, protein, and fat intake were also higher in the Haryana cohort compared with the others.

Duodenal biopsy changes in participants positive for CeD

Six of 76 antibody-positive participants in the Haryana site and none of the celiac-positive participants in the other two study sites had diarrhea. On the other hand, a large number of participants positive for antibody had anemia. Nearly a third of antibody-positive individuals declined further investigation. Duodenal biopsy findings showed a gradient from Haryana to Assam to Tamil Nadu with 69.4, 54.9, and 12.5%, respectively, in the three sites showing changes consistent with CeD (**Table 4**). Thus, the presence of biopsy changes in serology-positive patients correlated positively with the population prevalence of CeD in the concerned geographic region (P=0.0063, χ^2 for trend).

Classification and projected prevalence of CeD in the three sites

Of the fully investigated participants 5 of 36 in Haryana had classical CeD compared with none in the 2 other sites. Asymptomatic CeD was noted in 20/36 participants in Haryana, 28/51 in Assam, and 1/9 in Tamil Nadu. Latent CeD was present in 11/36 participants in Haryana, 23/51 participants in Assam, and 8/9 participants in Tamil Nadu. Extrapolating from the biopsy data to the total population, the proportion of actual CeD (autoantibody positive, mucosal change present) to latent CeD (autoantibody positive, mucosal change absent) was 70 and 30% in Haryana, 55 and 45% in Assam, and 89 and 11% in Tamil Nadu. On this basis, the estimated prevalence of actual CeD and latent CeD is shown in **Table 5**.

Table 2. Prevalence of HLA-DQ 2/8 genotypes of anti-tTG-negative participants

Variable	Haryana	Assam	Tamil Nadu	*P* value
Number of participants tested	110	226	247	—
DQ2 homozygous, *n*	2	11	11	—
DQ2 heterozygous, *n*	29	52	48	—
DQ8 homozygous, *n*	1	2	18	—
DQ8 heterozygous, *n*	9	5	13	—
DQ2/DQ8 double heterozygous, *n*	1	1	0	—
Genotype DQ2/DQ8, *n* (%)	42 (38.19)	71 (31.42)	90 (36.43)	0.4659

anti-tTG, anti-tissue transglutaminase; HLA, human leukocyte antigen.
The last row shows the total number (%) of individuals in each site positive for either DQ2 or DQ8 or both genotypes.

Table 3. Twenty-four-hour dietary intakes of selected macronutrients and micronutrients of participants (negative for Aesku testing) in each of the three regions

Variable	Haryana	Assam	Tamil Nadu
Number of participants tested	101	72	128
Energy (kcal)	2,250 (467)	1,646 (375)	1,673 (427)
Protein (g)	74.2 (17.8)	40.7 (11.0)	36.8 (7.6)
Carbohydrate (g)	367.5 (81.2)	331.6 (81.0)	307.1 (72.6)
Fat (g)	46.6 (12.5)	16.0 (2.1)	25.6 (5.9)
Wheat intake (g)	455.6 (114.5)	37.8 (11.5)	25.7 (26.8)
Rice intake (g)	75.1 (39.2)	326.1 (130.6)	167.5 (77.2)
Other grain (g)	0	0	6.8 (9.2)

Values shown are mean (s.d.).

Table 4. Duodenal biopsy findings in participants with serological evidence for celiac disease (positive in the Aesku test and positive in the Inova test)

	Anti-tTG positive (Aesku and Inova)		
	Haryana	Assam	Tamil Nadu
Number positive for celiac disease	76	70	12
Number who underwent biopsy	36	51	9
Marsh grade 0 (IEL<40/100), normal C:V ratio	11	23	8
Marsh grade I** (>40/100)	—	15	1
Marsh grade II**	5	1	—
Marsh grade IIIa**	8	9	—
Marsh grade IIIb**	7	3	—
Marsh grade IIIc**	5	0	—

anti-tTG, anti-tissue transglutaminase; IEL, intraepithelial lymphocyte.
In the modified Marsh grading, biopsies were considered to show histological evidence of celiac disease if IEL were >40 IEL/100 epithelial cells (**). The grades of I–III in the Marsh classification represent the absence or presence of villous atrophy.

HLA-DQ determinants in participants with CeD

Among participants with a diagnosis of CeD and latent CeD in the three sites, HLA-DQ2 and/or DQ8 determinants of CeD were present in 75.8% in Haryana, in 39.1% in Assam, and in 28.5% in Tamil Nadu (**Table 6**). Among those with biopsy changes the relevant percentages were 83.3%, 44.4%, and nil, respectively, in the three sites.

DISCUSSION

We found significant geographic variations within India in CeD prevalence among healthy adults in this community-based prevalence study. There was a gradient of CeD within the country, with prevalence being highest in the northern state, less in the northeastern state, and least in the southern state. The proportion of adults in these parts having a HLA-DQ background permissive for CeD was similar in the three areas surveyed. However, there were significant differences in wheat consumption between the three areas.

Although CeD prevalence in the Asia-Pacific region is variable (8), ranging from 1:50 to 1:500 in Australia, Iran, Israel, New Zealand, Syria and Turkey, it is extremely rare in Japan and in China. CeD in India was first described in northern Indian children in the 1960s (10,11). A community-based study from the Punjab a decade ago estimated CeD prevalence in children of 1 in 310 (12). Using a symptom-based screen, the prevalence of CeD in adults in northern India was estimated at 1.04% (13). There are virtually no reports of CeD from southern India. The present study used a screening blood test to identify individuals with CeD in the community and established that, although the disease is indeed common in northern India, it is quite uncommon in southern India.

Antibodies to tissue transglutaminase are used in both screening and diagnosis of CeD. Here we tested for IgA anti-tTG Ab using a very sensitive ELISA in which the capture antigen was recombinant human tTG complexed with deamidated gliadin peptides (23). Those samples that were negative for anti-tTG Ab in this assay were likely to be truly negative for CeD. Positives in this test were screened further by a second ELISA, using human erythrocyte tTG as the capture antigen, which was very specific for CeD (23). Hence, samples testing positive in the second sequential ELISA were probably true positives. The use of two ELISAs in tandem or in sequence has previously been validated for the diagnosis of CeD (28,29). These serological tests measured IgA antibody and are likely to be negative in individuals with selective IgA deficiency. However, the reported prevalence of selective IgA deficiency in Indians is very low (30), and we did not test for IgG anti-tTG as this would not significantly alter the estimate of population prevalence of CeD.

Table 5. Estimated prevalence of actual (symptomatic and asymptomatic) and latent CeD in the three sites

Site	Symptomatic CeD	Asymptomatic CeD	Latent CeD
Haryana	1.77/1,000 (0.88–3.17/1,000)	6.76/1,000 (4.88–9.14/1,000)	3.70/1,000 (2.34–5.56/1,000)
Assam	0	4.66/1,000 (3.31–6.40/1,000)	3.92/1,000 (2.68–5.54/1,000)
Tamil Nadu	0	0.11/1,000	1.22/1,000 (0.6–2.1/1,000)

CeD, celiac disease.

Values shown are the prevalence (95% confidence intervals) per 1,000 population. Prevalences were extrapolated to the population based on the proportion of non-responding participants who refused further assessment after antibody testing.

Table 6. Classification of celiac serology-positive participants according to further investigation

	Haryana	Assam	Tamil Nadu
Number positive for celiac disease	76	70	12
Number who underwent biopsy	36	51	9
Biopsy and HLA done	29	45	7
HLA-DQB1 allele	—	—	—
Biopsy abnormal			
DQB1*02 present	14	10	0
DQB1*0302 present	1	1	0
Both absent	3	15	1
Biopsy normal			
DQB1*02 present	7	4	2
DQB1*0302 present	0	2	0
Both absent	4	13	4

HLA, human leukocyte antigen.

It is notable that all the reports on CeD in India originate from the northern states of Delhi, Haryana, Punjab, and Uttar Pradesh (31–35). In the present study, symptomatic CeD was prevalent only in the northern Indian cohort, being absent in the other two cohorts. The findings support the prevalent notion that CeD is rare or does not exist in southern or northeastern India. Asymptomatic CeD was noted in both northern and northeastern cohorts, but its prevalence in the southern Indian cohort was very much lower than in the other two cohorts. Latent CeD was noted in all three cohorts, with the southern cohort exhibiting a significantly lower prevalence of latent CeD than the other two. These differences merit further exploration.

The development of CeD requires a permissive genetic background related to the expression of HLA-DQ antigens. Almost all patients with CeD have a specific HLA type genetically defined as *HLA-DQB1*02* (serologically defined as HLA-DQ2) or *HLA-DQB1*0302* (which is serologically defined as HLA-DQ8). In Europe and North America, a third of the population belongs to these two HLA-DQ types, and the prevalence of CeD in countries in these regions approximates or exceeds 1% (1,5). The absence of these two genotypes in a population is associated with the absence of CeD in that population, as noted in Japan and South East Asia (8,15). Previous studies from India suggested that the allele prevalence of *HLA-DQB1*02* in northern India ranged from 16.3 to 31.9%, whereas that of *HLA-DQB1*0302* ranged from 0 to 5.4% (17,19). Studies in selected ethnic groups in southern India revealed *HLA-DQB1*02* allele frequencies of 8.9–14.3% and *HLA-DQB1*0302* allele frequencies of 6.6–10.5% (18). In the present study, the allele frequencies of *HLA-DQB1*02* and *HLA-DQB1*0302* were broadly within the same range. The permissive HLA-DQ 2/8 genotypes were present in about a third each of the three regional populations tested.

Ninety three percent or more of northern Indian children with CeD were previously shown to have the *HLA-DQB1*02* allelotype (36–38). Of the participants with CeD and histological changes in the present study, 77.73% in the northern state, 38.4% in the northeastern state, and 0% in the southern state had the *HLA-DQB1*02* allelotype. Of those participants with CeD without histological changes, 63.63% in the northern state, 21.05% in the northeastern state, and 33.33% in the southern state had the *HLA-DQB1*02* allelotype. It was interesting that, whereas the HLA-DQ 2/8 distribution was homogenous in the general population across the areas, it was tremendously different when CeD cases were considered. No doubt other explanations apply here, and this requires further investigation. In a proportion of patients, CeD has been shown to occur in association with half a DQ2 molecule and not the entire DQB1*02-DQA1*05 (39). The Tamil Nadu population had a higher prevalence of DQ8 alleles, but only a small number of CeD cases were identified here.

Wheat intake in northern India was very high, and was lower in the northeastern and southern populations. Thus, although the quantity of wheat intake could be related to regional CeD prevalence, the relationship was somewhat less obvious in the northeastern population. There was a gradient of wheat intake between the three regions, and conversely there was a gradient in the opposite direction of symptomatic CeD in the three regions. It is possible that genetic makeup, involving either HLA-DQ or other genes, interacts with the quantum of wheat intake to determine CeD prevalence in the three regions.

It was interesting that in the northern site, 83.3% of individuals with abnormal biopsies were HLA-DQ2 or -DQ8 positive, compared with 44.4% in the northeastern region and 0% in the southern region. Similarly, in those with normal biopsies and CeD by serology, HLA-DQ2 or -DQ8 was present in 63.6% in the northern region, in 31.5% in the northeastern region, and in 33.3% in the southern region. Thus, a number of the participants expressing antibody responses consistent with CeD in the northeastern and southern regions did not have the HLA-DQ antigen types conventionally associated with CeD. It is possible that these represent false-positive results of the serology, even though we used two quite different serological tests to make the diagnosis of

CeD. It is possible that the latter participants may have a condition other than CeD that led to antibody generation. We would like to speculate also that perhaps non-DQ2, non-DQ8 HLA types lead to antibody generation alone without leading to intestinal mucosal damage—i.e., to latent CeD.

There were several limitations in the present study. We did not include children in the population prevalence study essentially because of the difficulties surrounding consenting children in the community for obtaining blood samples for study. Extrapolating from the adult data, we can assume that the same regional differences between northern and southern parts of the country exist with respect to CeD in children, but this remains to be formally proven. It is clear that, althoguh there are numerous reports of CeD in children from northern India, similar reports from southern India are virtually absent. A second limitation of the study is that it was not possible to biopsy all participants who had positive serology. This was because of lack of participant consent during the phase of the study when the serology-positive individuals were approached for further participation. As a result, the prevalence of CeD and latent CeD is based on the assumption that distribution of intestinal mucosal abnormalities in the non-biopsied participants would be in the same proportion as in those who were biopsied.

Finally, the findings of this study have important implications for public health policy in India and other countries in this region. Latent CeD, with positive antibody and normal small bowel mucosa, showed a prevalence of approximately 3 per 1,000 in the northern and northeastern state. There is considerable controversy over the significance of latent CeD. However, carefully conducted studies, which retrospectively examined celiac autoantibody in a cohort of well-characterized healthy individuals followed up for many years, suggest that latent CeD may be associated with up to fourfold increase in mortality compared with those without latent CeD (40). CeD characterized by mucosal changes was found in ~8 per 1,000 adult population in the northern state, whereas it was in the range of 1 per 10,000 adult population in the southern state. As there was no difference in the frequency of HLA-DQ2-expressing genotypes between the three states, it seems likely that quantum of wheat consumption was the major determinant of CeD prevalence. These findings are likely generalizable to the various regions of India, based on the dietary preferences. Wheat is the staple grain used in the northern and western states of India, whereas rice is the staple grain in eastern and southern states of India. The present study showed that the genetic background for HLA-DQ2 expression was broadly similar in the regions of the country studied, and there is no reason to expect that the frequency of HLA-DQ2 expression would differ in the population of the other states. Thus, we can expect that CeD with histological change is present in nearly 1% of the population of the northern Indian states, whereas it is much less prevalent in the southern states.

ACKNOWLEDGMENTS

The assistance provided by the ICMR Task Force on Celiac Disease, including Narender Arora, Rakesh Kochhar, Meera Mathur, Sita Naik, R.M. Pandey, Amarender Puri, and H.P.S. Sachdeva, in study design and monitoring is gratefully acknowledged. We also gratefully acknowledge assistance provided by Dr L. Jeyaseelan in data analysis. We also acknowledge assistance provided on the ground by Kirankumar Baskaran, Sanjit Kumar, Manohar Madhukar, John Mechenro, Sandeep Paras, G. Priya, Surendranath, and Giriprasad Venugopal. The study sponsors appointed a committee of experts that reviewed and modified the initial study design, reviewed mid-term progress and final completion reports, and the data analysis and interpretation.

CONFLICT OF INTEREST

Guarantor of the article: B.S. Ramakrishna, DM, PhD.

Specific author contributions: Planning the study, overall coordination between the three study sites, supervision of the Tamil Nadu site, data interpretation, and drafting the manuscript: B.S. Ramakrishna; planning the study, supervision of the Haryana site, data interpretation, and drafting the manuscript: Govind K. Makharia; planning the study, supervision of the Assam site, and data interpretation: Kamal Chetri and Sangitanjan Dutta; planning and overall coordination, arranging an independent review of study design, study progress, and interpretation: Prashant Mathur; supervision of the Haryana site: Vineet Ahuja; participant recruitment and data collection at the Haryana site: Ritvik Amarchand and Anil K. Verma; project management at the Tamil Nadu site: Ramadass Balamurugan; participant recruitment and data collection at the Tamil Nadu site: Sudipta D. Chowdhury, Gemlyn George, and Kartik Ramakrishna; analysis and interpretation of the HLA typing: Dolly Daniel, Gurvinder Kaur, and Srinivasan Pugazhendhi; participant recruitment and pathology analysis and interpretation: Anup Das, Siddhartha Datta Gupta and Anna Pulimood; study design and co-supervision of the Haryana study site: Anand Krishnan; study design and co-supervision of the Tamil Nadu study site: Jasmin H. Prasad. All authors approved the final draft of the manuscript submitted.

Financial support: The study was funded by a Task Force grant (No. 5/4/3-1/TF/2011-NCD-II) from the Indian Council of Medical Research, New Delhi.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Celiac disease prevalence in northern India is close to that in Europe and North America.
- ✓ Celiac disease in adults may be classical with gastrointestinal manifestations or non-classical with non-gastrointestinal manifestations.
- ✓ Human leukocyte antigen (HLA)-DQ2 or DQ8 expression is needed for manifestation of celiac disease.

WHAT IS NEW HERE

- ✓ Compared with northern India, celiac disease is infrequent or absent in southern India.
- ✓ Celiac disease in southern India is almost always latent.
- ✓ HLA-DQ 2/8 genotypes are found in one-third of the population in India.
- ✓ HLA-DQ genotypes other than 2/8 were found more often in latent CeD.

REFERENCES

1. Green PHR, Cellier C. Celiac disease. N Engl J Med 2007;357:1731–43.
2. West J, Fleming KM, Tata LJ *et al.* Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. Am J Gastroenterol 2014;109:757–68.
3. Altobelli E, Paduano R, Petrocelli R *et al.* Burden of celiac disease in Europe: a review of its childhood and adulthood prevalence and incidence as of September 2014. Ann Ig 2014;26:485–98.
4. Ludvigsson JF, Bai JC, Biagi F *et al.* Diagnosis and management of adult coeliac disease: guidelines from the British Society of Gastroenterology. Gut 2014;63:1210–28.
5. Catassi C, Gatti S, Fasano A. The new epidemiology of celiac disease. J Pediatr Gastroenterol Nutr 2014;59(Suppl 1):S7–9.
6. Catassi C, Rätsch IM, Gandolfi L *et al.* Why is coeliac disease endemic in the people of the Sahara? Lancet 1999;354:647–8.
7. Catassi C, Doloretta Macis M, Rätsch IM *et al.* The distribution of DQ genes in the Saharawi population provides only a partial explanation for the high celiac disease prevalence. Tissue Antigens 2001;58:402–6.
8. Cummins AG, Roberts-Thomson IC. Prevalence of celiac disease in the Asia-Pacific region. J Gastroenterol Hepatol 2009;24:1347–51.
9. Nakazawa H, Makishima H, Ito T *et al.* Screening tests using serum tissue transglutaminase IgA may facilitate the identification of undiagnosed celiac disease among Japanese population. Int J Med Sci 2014;11:819–23.
10. Walia BN, Sidhu JK, Tandon BN *et al.* Coeliac disease in North Indian children. Br Med J 1966;2:1233–4.
11. Walia BN, Mehta S, Gupte SP. Coeliac disease. Indian Pediatr 1972;9:16–19.
12. Sood A, Midha V, Sood N *et al.* Prevalence of celiac disease among school children in Punjab, North India. J Gastroenterol Hepatol 2006;21:1622–5.
13. Makharia GK, Verma AK, Amarchand R *et al.* Prevalence of celiac disease in the northern part of India: a community based study. J Gastroenterol Hepatol 2011;26:894–900.
14. Ganesh R, Suresh N, Sathiyasekaran M. Celiac disease, still an uncommon problem in Tamilians? Indian J Gastroenterol 2009;28:189.
15. Liu E, Lee HS, Aronsson CA *et al.* Risk of pediatric celiac disease according to HLA haplotype and country. N Engl J Med 2014;371:42–9.
16. Bruins MJ. The clinical response to gluten challenge: a review of the literature. Nutrients 2013;5:4614–41.
17. Rani R, Fernandez-Vina MA, Stastny R. Association between HLA class II alleles in a north Indian population. Tissue Antigens 1998;52:37–43.
18. 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.
19. Agrawal S, Srivastava SK, Borkar M *et al.* Genetic affinities of north and northeastern populations of India: inference from HLA-based study. Tissue Antigens 2008;72:120–30.
20. Government of India, Ministry of Home Affairs. Available at: http://www.censusindia.gov.in/PopulationFinder/Population_Finder.aspx (accessed on 3 March 2015).
21. Kish L. A procedure for objective respondent selection within the household. J Am Stat Assoc 1949;22:380–7.
22. Kumar N, Shekhar C, Kumar P *et al.* Kuppuswami's socioeconomic status scale—updating for 2007. Indian J Pediatr 2007;74:1131–2.
23. Wong RC, Wilson RJ, Steele RH *et al.* A comparison of 13 guinea pig and human anti-tissue transglutaminase antibody ELISA kits. J Clin Pathol 2002;55:488–94.
24. Gopalan C, Rama Sastri BV, Balasubramanian SC. Nutritive values of Indian foods Indian Council of Medical Research: New Delhi, 2004.
25. Dickson BC, Streutker CJ, Chetty R. Coeliac disease: an update for pathologists. J Clin Pathol 2006;59:1008–16.
26. Genotyping of the HLA-DQB1 gene. Available at: https://www.niddkrepository.org/static/studies/gokind/moop/genotyping%20moop/Genotyping_HLA-DQB1.pdf (accessed on 04 March 2015).
27. Ludvigsson JF, Leffler DA, Bai JC *et al.* The Oslo definitions for coeliac disease and related terms. Gut 2013;62:43–52.
28. Aberg AK, Olcén P. Serologic screening for celiac disease in children: a comparison between established assays and tests with deamidated gliadin-derived peptides plus conjugates for both IgA and IgG antibodies. APMIS 2009;117:808–13.
29. Rozenberg O, Lerner A, Pacht A *et al.* A new algorithm for the diagnosis of celiac disease. Cell Mol Immunol 2011;8:146–9.
30. Chandran S, Khetan D, Chaudhary R *et al.* Low prevalence of IgA deficiency in north Indian population. Indian J Med Res 2006;123:653–6.
31. Yachha SK, Poddar U. Celiac disease in India. Indian J Gastroenterol 2007;26:230–7.
32. Makharia GK, Baba CS, Khadgawat R *et al.* Celiac disease: variations of presentations in adults. Indian J Gastroenterol 2007;26:162–6.
33. Agarwal N, Puri AS, Grover R. Non-diarrheal celiac disease: a report of 31 cases from northern India. Indian J Gastroenterol 2007;26:122–6.
34. Sood A, Midha V, Sood N *et al.* Adult celiac disease in northern India. Indian J Gastroenterol 2003;22:124–6.
35. Sachdev A, Srinivasan V, Maheswary S *et al.* Adult onset celiac disease in north India. Trop Gastroenterol 2002;23:117–9.
36. Agrawal S, Gupta A, Yachha SK *et al.* Association of human leucocyte-DR and DQ antigens in coeliac disease: a family study. J Gastroenterol Hepatol 2000;15:771–4.
37. Kaur G, Sarkar N, Bhatnagar S *et al.* Pediatric celiac disease in India is associated with multiple DR3-DQ2 haplotypes. Hum Immunol 2002;63:677–82.
38. Mehra NK, Kaur G, Kanga U *et al.* Immunogenetics of autoimmune diseases in Asian Indians. Ann N Y Acad Sci 2002;958:333–6.
39. Margaritte-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.
40. Rubio-Tapia A, Kyle RA, Kaplan EL *et al.* Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology 2009;137:88–93.