ONLINE METHODS

Subjects. Written informed consent was obtained from all subjects, with Ethics Committee / Institutional Review Board approval. All individuals, except the Indian population sample, are of European ancestry. DNA samples were from blood, lymphoblastoid cell lines or saliva.

Affected celiac individuals were diagnosed according to standard clinical criteria, compatible serology and in all cases small intestinal biopsy - most cases were diagnosed using the revised ESPGHAN criteria as a minimum requirement[34]. More specific requirements were: UK cases[3-5] (hospital outpatients, n=1145) required Marsh III stage intestinal biopsy (HLA-DQ2.5cis tag SNP rs2187668 MAF=0.4699); UK cases [4,5] (Coeliac UK members, n=6583), self-reported diagnosis by intestinal biopsy (note the rs2187668 MAF=0.4803 was similar to UK hospital cases, versus combined UK controls MAF 0.1419); Italy (Milan) [5,35] and Polish[5] required Marsh III stage intestinal biopsy and positive endomysial/tissue transglutaminase antibodies; Spain (CEGEC)[36] required at least Marsh II stage intestinal biopsy; Netherlands cases[5] required Marsh III stage intestinal biopsy, or Marsh II stage intestinal biopsy with compatible HLA-DQ type; India (Punjab) cases required Marsh III stage intestinal biopsy and strongly positive tissue transglutaminase antibodies; Italy (Naples, Rome) required abnormal intestinal biopsy and positive tissue transglutaminase antibodies[37].

The UK 1958 Birth Cohort and UK Blood Services-Common Controls are unselected population controls. Polish controls and Italian (Naples) controls excluded celiac serology positive samples. Spain (Madrid) controls were unselected blood donors and hospital employees. Spain (CEGEC), Italy (Rome), Indian (Punjab) controls were unselected blood donors. Italian controls (Milan) were unselected healthy individuals. Netherlands controls were unselected blood donors and population controls.

SNP selection: All 1000 Genomes Project low-coverage whole genome sequencing pilot CEU variants within 0.1cM of the lead SNP for each disease and region were selected. The Sept 2009 release comprising 60 CEU individuals was used (~5x mean read-depth whole genome sequencing) selecting markers called in at least two of the Broad Institute / Sanger Institute / University of Michigan algorithms. Additional genomic region re-sequencing content was submitted for Immunochip at specific loci from coeliac disease, Crohn's disease and type 1 diabetes cases and controls (**Supplementary Note**).

Genotyping. Samples were genotyped using the Immunochip as per Illumina's protocols (at labs in London, Hinxton, Groningen and Charlottesville). NCBI build 36 (hg18) mapping was used (Illumina manifest file Immuno_BeadChip_11419691_B.bpm).

Data Quality Control. Very low call rate samples and variants were first excluded (and samples repeated). The Illumina GenomeStudio GenTrain2.0 algorithm was used to cluster an initial 2,000 UK samples. Subsequently with additional sample data (case and control data were analysed together) clusters were re-adjusted or excluded (manual or automated) for variants with low quality statistics (call rate<99.5%, low GenCall score, many high-intensity no-calls). This method was superior to the GenoSNP or Illuminus clustering algorithms (not shown). A cluster set based on 172,242 autosome/X-chromosome variants (available on request) was then applied to all samples. Samples were excluded for call rate <99.5% across 172,242 markers. We then removed 15,657 non-polymorphic markers (i.e. only one of three expected genotype clouds observed) which reflect a combination of ethnic-specific variants, allele-specific assay failure, as well as substantial false-positive rates in early next-generation sequencing SNP calling algorithms.

Samples were excluded for incompatible recorded gender and genotype inferred gender, duplicates and first/second degree relatives. Potential ethnic outliers were identified by multi-dimensional scaling plots of samples merged with HapMap3 data, the subset of SNPs common to HapMap3 and Immunochip accurately identified the different HapMap3 population samples. We considered the white European and Indian collections separately.

Stepwise conditional logistic regression is sensitive to missing data and subtle genotyping error, and we therefore desired an ultra-high quality dataset. Markers were excluded from all sample collections for deviation from Hardy-Weinberg equilibrium in controls (P<0.0001) and/or differential missingness in genotype no-calls between cases and controls (P<0.001), in any of the seven collections. Finally we required a per-SNP call rate of >99.95% (a maximum of 12 no-call genotypes from 24,269 samples per autosomal marker), generating a dataset of 139,553 markers (of which all but 372 indels are SNPs).

We visually inspected intensity plot genotype clouds for all markers described in **Table 2** (and further potential loci with $P<1.9\times10^{-6}$), and confirmed all to be high quality.

Genotype data has been deposited at the European Genome-phenome Archive (EGA, http://www.ebi.ac.uk/ega/), which is hosted by the EBI, under accession number EGAS00000000053.

Statistical Analysis. Analyses were performed using PLINK v1.07[38], using logistic regression tests with gender as a covariate, and collection membership (**Table 1**) as a factorized covariate. Stepwise conditional logistic regression was performed, in order of markers with the smallest P value. Graphs were plotted in R, and using a modified version of LocusZoom[39].

We permuted affection status for the dataset at each region (**Online Methods**), to establish locus wide statistical significance thresholds for defining independently associated SNPs. For each locus, given by the linkage disequilibrium

boundaries (**Table 2**) we calculated the 5th percentile based on the nominal *P*-value distribution for 1000 permutations controlling for multiple marker testing. This approach proved slightly more stringent than a per-locus Bonferroni correction for independent (using an estimate for "independence" of pairwise $r^2<0.05$) variants (**Supplementary Table 3**). We estimated our dataset contained 26,146 completely uncorrelated variants (using pairwise $r^2<0.05$ and a sliding 1000 SNP window).

The fraction of additive variance was calculated using a liability threshold model[40] assuming a population prevalence of 1%. Effect sizes and control allele frequencies were estimated from the UK dataset. Genetic variance was calculated assuming 50% heritability.