Dissecting the contribution of single nucleotide polymorphisms in CCR9 and CCL25 genomic regions to the celiac disease phenotype

Airaksinen, Laura

2021

Airaksinen , L , Cerqueira , J X M , Huhtala , H , Saavalainen , P , Yohannes , D A , Mäki , M , Kurppa , K , Kilpeläinen , E , Shcherban , A , Palotie , A , Kaukinen , K & Lindfors , K 2021 , ' Dissecting the contribution of single nucleotide polymorphisms in CCR9 and CCL25 genomic regions to the celiac disease phenotype ' , Journal of translational autoimmunity , vol. 4 , 100128 . https://doi.org/10.1016/j.jtauto.2021.100128

http://hdl.handle.net/10138/338767
https://doi.org/10.1016/j.jtauto.2021.100128

Journal of Translational Autoimmunity 4 (2021) 100128

Contents lists available at ScienceDirect

Journal of Translational Autoimmunity

journal homepage: www.sciencedirect.com/journal/journal-of-translational-autoimmunity

Dissecting the contribution of single nucleotide polymorphisms in *CCR9* and *CCL25* genomic regions to the celiac disease phenotype

Laura Airaksinen [a,1], Juliana XM. Cerqueira [a,b,1], Heini Huhtala [c], Päivi Saavalainen [d], Dawit A. Yohannes [d], Markku Mäki [a,e], Kalle Kurppa [a,e,f], Elina Kilpeläinen [g], Anastasia Shcherban [g], Aarno Palotie [g,h], Katri Kaukinen [a,i], Katri Lindfors [a,*]

[a] *Celiac Disease Research Center, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland*
[b] *Faculty of Nutrition and Food Sciences, University of Porto, Porto, Portugal*
[c] *Faculty of Social Sciences, Tampere University, Tampere, Finland*
[d] *Translational Immunology Research Program, and Department of Medical and Clinical Genetics, University of Helsinki, Helsinki, Finland*
[e] *Tampere Center for Child, Adolescent, and Maternal Health Research, Tampere University, and Department of Pediatrics, Tampere University Hospital, Tampere, Finland*
[f] *University Consortium of Seinäjoki, Seinäjoki, Finland*
[g] *Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland*
[h] *Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Analytic and Translational Genetics Unit, Department of Medicine, and the Department of Neurology, Massachusetts General Hospital, Boston, MA, USA*
[i] *Department of Internal Medicine, Tampere University Hospital, Tampere, Finland*

ARTICLE INFO

Keywords:
Celiac disease
Clinical picture
Chemokine receptor
Genetic variation
Genetic association

ABSTRACT

Purpose and objectives: Given their role in homing immune cells to the intestine, CC motif chemokine receptor 9 (CCR9) and its specific ligand CC motif chemokine ligand 25 (CCL25) are interesting candidate genes for celiac disease. These genes are located in regions previously shown to be associated with or linked to celiac disease, but no investigations on their association with various celiac disease phenotypes have so far been conducted. Here we studied such associations of both genotyped and imputed single nucleotide polymorphisms (SNPs) with either regulatory function or exonic location of the *CCR9* and *CCL25* loci.

Results: Exploiting a carefully phenotyped cohort of 625 celiac disease patients and 1817 non-celiac controls, we identified that multiple SNPs with predicted regulatory function (RegulomeDB score $\leq$3a and/or eQTL effect) located between 100 kB upstream and downstream of *CCR9* and *CCL25* are associated with celiac disease and/or selected phenotypes. Of the genotyped SNPs in the *CCR9* loci, rs213360 with an eQTL effect on *CCR9* expression in blood was associated with celiac disease and all investigated phenotypes except high HLA risk. Rs1545985 with an eQTL on *CCR9* expression and rs7652331 and rs12493471, both with RegulomeDB score $\leq$3a, were all associated with gastrointestinal symptoms and malabsorption and the latter additionally with anemia. The genotyped *CCL25* SNPs rs952444 and rs882951, with RegulomeDB scores 1d and 1f respectively and eQTL effect on *CCL25* expression in small intestine, were associated with gastrointestinal symptoms and malabsorption. The *CCL25* SNP rs2303165 identified in sequencing followed by imputation was associated with partial villous atrophy. However, the association did not pass the permutation based multiple testing correction ($P_{EMP2} > 0.05$).

Conclusions: We conclude that SNPs in the region of *CCR9* and *CCL25* with predicted functional effect or exonic localization likely contribute only modestly to various celiac disease phenotypes.

Abbreviations: CCR9, CC motif chemokine receptor 9; CCL25, CC motif chemokine ligand 25; SNP, single nucleotide polymorphism; HLA, human leukocyte antigen; TG2, transglutaminase 2; FUMA, Functional Mapping and Annotation of GWAS; GWAS, genome-wide association study; eQTL, expression quantitative trait loci; QC, quality control; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; OR, odds ratio; CI, confidence interval; PBMC, peripheral blood mononuclear cell.

* Corresponding author. Celiac Disease Research Center, Faculty of Medicine and Health Technology, Tampere University, FIN-33014, Finland.
E-mail address: katri.lindfors@tuni.fi (K. Lindfors).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jtauto.2021.100128
Received 12 August 2021; Received in revised form 2 October 2021; Accepted 13 October 2021
Available online 14 October 2021

1. Introduction

Celiac disease is an immune-mediated chronic condition where oral tolerance to dietary gluten has been lost. The multifaceted disease can present with varying gastrointestinal and/or extraintestinal signs and symptoms. In addition, patients may be completely asymptomatic [1]. A prerequisite for the development of the disease is the presence of human leukocyte antigen (HLA) molecules HLA-DQ2 or -DQ8 and individuals homozygous for HLA-DQ2 encoding alleles are at a particularly high risk [2,3]. Celiac disease, with its autoimmune nature, is characterized by an IgA class autoantibody response against endogenous enzyme transglutaminase 2 (TG2) measurable in serum. In patients with celiac disease, the ingestion of gluten leads to small-bowel mucosal villous atrophy and crypt hyperplasia of varying severity. Moreover, a profound mucosal inflammation, characterized by increased density of intraepithelial lymphocytes and infiltration of both T cells and plasma cells in the lamina propria is usually present [1]. The inflammatory cells arrive at the small bowel from secondary lymphoid tissue in a process involving the gut-homing chemokine CC motif chemokine ligand 25 (CCL25) and its specific receptor CC motif chemokine receptor 9 (CCR9) [4].

In addition to being involved in homing immune cells to the intestine, the main organ affected by celiac disease, the *CCR9-CCL25* axis is implicated in celiac disease by several other studies. Firstly, CCR9-positive dendritic cells play an essential role in maintaining gut homeostasis and tolerance by regulating the phenotype and function of both innate and adaptive immune cells [5]. Secondly, there is data to suggest that the number of CCR9-positive type 2 conventional dendritic cells is increased in circulation in untreated celiac disease [6] while the number of CCR9-expressing T cells is diminished in the small intestines of patients [7]. Thirdly, a clinical trial to test the efficacy of an oral CCR9 inhibitor as an alternative form of treatment for celiac disease has been performed but not yet published [8]. Lastly, the genes coding for CCL25 and CCR9 are located in chromosomal regions previously linked to or shown to be associated with celiac disease [9–11]. However, in earlier genetic association studies the coverage of single nucleotide polymorphism (SNPs) at *CCR9* and *CCL25* loci, particularly those in exonic regions, has been sparse. Moreover, to the best of our knowledge no studies focusing on genomic variants at these loci on different celiac disease phenotypes exist.

For the above mentioned reasons, we considered *CCR9* and *CCL25* to be potential positional and functional candidate genes in celiac disease. Consequently, we identified SNPs in *CCR9* and *CCL25* genomic regions that likely have functional effects or are located in the exons and tested their association with celiac disease and selected disease phenotypes.

2. Materials and methods

2.1. Patients and controls

The study was conducted at Tampere University and Tampere University Hospital. Altogether 1048 celiac disease patients with biopsy-proven small bowel mucosal damage were recruited with the assistance of national and local celiac societies and by nationwide media announcements. The patient information at diagnosis was collected from medical records and from supplementary interviews by a physician or a study nurse with expertise in celiac disease. In the case of children, the guardian was interviewed. The structured interviews included questions on celiac disease diagnosis, symptoms at the time of diagnosis and in childhood, and associated medical conditions. Whole blood samples were drawn for genetic analysis. In order to avoid bias (inflation of type 1 error) caused by the inclusion of several subjects from the same family, the present study considered only one randomly selected celiac case with full genotype available per family, resulting in 625 cases. The median age of the patients was 41 (range 0.5–79) years. Patients' demographic data as well as their clinical characteristics in terms of selected celiac disease-associated phenotypes (gastrointestinal symptoms, malabsorption, anemia, severity of small bowel mucosal damage, HLA risk categories, and celiac disease antibodies) are presented in Table 1. DNA samples from 144 HLA-DQ2-positive biopsy-proven celiac disease patients and 144 non-celiac controls outside the genotyped celiac disease patient cohort were subjected to sequencing. The study design, patient recruitment, and collection of patient record data were approved by the Regional Ethics Committee of Tampere University Hospital. All participants, or in the case of children their legal guardians, gave written informed consent.

As non-celiac controls, altogether 1817 subjects with information on gender (1032 males, 785 females) and HLA-genotype (Low HLA risk, N = 939; Intermediate HLA risk, N = 855; High HLA risk, N = 23) from the population representative cohorts FINRISK and Health 2000 [12] were included in the study. Ethics committee approvals were available from the National Public Health Institute's Ethics committee and the Ethics committee in Epidemiology and Public Health in the Hospital District of Helsinki and Uusimaa.

2.2. Association analysis of the genotyped CCR9 and CCL25 SNPs

Genotypes for all subjects had been produced by Illumina 610-Quad BeadChip array (Illumina Inc., San Diego, CA, USA) [9]. First, variants spanning between 100 kB upstream and downstream of *CCR9* and *CCL25* genes were identified resulting in 105 SNPs. Thereafter, FUMA (Functional Mapping and Annotation of GWAS) platform [13] was used to annotate SNPs fulfilling our functional annotation criteria. Publicly available genome-wide association study (GWAS) summary statistic results in which our cohort has been included [9] and a pre-defined list with the 105 SNPs of interest were uploaded. The RegulomeDB 2.0 was used to identify all the SNPs with known and predicted regulatory elements and to assign them a score ranging from 1a to 7 in RegulomeDB ranking [14]. SNPs with RegulomeDB score between 1a and 3a likely to affect the gene expression, were selected for our study. In addition, the expression quantitative trait loci (eQTL) mapping using public eQTL data was used to select SNPs having significant eQTL effects (FDR <0.05) on the expression of *CCR9* gene in blood (GTEx whole blood [15], Blood eQTL [16], BIOS QTL [17], and eQTLGen) and on the expression of *CCL25* gene in the intestine (GTEx data for small intestine terminal ileum, colon sigmoid and colon transverse [15]; and of the "CEDAR" study [18], terminal ileum). Altogether 41 SNPs with RegulomeDB ranking from 1a to 3a, and/or tissue-specific eQTL effects were

Table 1
Demographic data and selected celiac disease phenotypes of 625 celiac disease patients at diagnosis.

	N	%
Females	489	78
Gastrointestinal symptoms[a]	526	84
Malabsorption[b]	267	43
Anaemia	157	25
Small bowel mucosal damage[c]		
Total or subtotal villous atrophy	361	66
Partial villous atropthy	185	34
HLA risk[d]		
High	98	16
Intermediate/low	527	84
Celiac disease autoantibodies[e]		
Positive	279	94
Negative	19	6

[a] Diarrhea, abdominal pain, flatulence, heartburn, nausea, vomiting.
[b] Anemia, vitamin and micronutrient deficiencies.
[c] Small bowel mucosal morphology data was available from 546 patients.
[d] High risk (DQ2.5/DQ2.5; DQ2.5/DQ2.2), intermediate risk (DQ2.5/X, DQ2.2/D2.2, DQ2.2/X, DQ8/DQ8, DQ8/X), low risk (DQ7/X, DQ7/DQ7).
[e] Autoantibody data (endomysial antibodies and/or tissue transglutaminase antibodies) was available from 298 patients.

identified. After applying quality control (QC) filtering for missing genotype rate <5% and missing genotype rate differences between the cases and controls (<3%), and minor allele frequency (MAF >5%), 348 cases and all the 1817 controls remained in the analysis, and 39 SNPs passed QC. All markers were in Hardy-Weinberg equilibrium (HWE) ($P > 1 \times 10^{-6}$) in the controls. Allelic associations of the 39 genotyped SNPs with celiac disease and the selected disease phenotypes were tested as described in Section 2.5.

2.3. Sequencing exonic regions of CCR9 and CCL25

DNA was extracted from whole blood samples or from leukocyte enriched buffy coats. Extractions were performed using FlexiGene DNA kit (Qiagen, Hilden, Germany). All coding exons and exon-intron boundaries of both *CCR9* and *CCL25* were amplified with PCR using primer pairs represented in Supplementary Table 1. The PCR reactions (20 μL) contained 10 mM Tris-HCl (pH 8.8 at 25 °C), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100 in 10X Optimized DyNAzyme Buffer (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 mM dNTPs, 1.0 μM of each primer, 80 ng genomic DNA and 1.0 U DyNAzyme II DNA Polymerase (Thermo Fisher Scientific). PCR conditions included initial denaturation at 95 °C for 5 min, 40 cycles of 95 °C for 30 s, 55–60 °C (depending on the PCR product) for 40 s and 72 °C for 1 min. The final extension was performed at 72 °C for 5 min. Successful PCR amplification was confirmed by analyzing PCR products with agarose gel electrophoresis and by UV–Vis spectrophotometry (NanoDrop, Thermo Fisher Scientific). Sequencing was performed by Macrogen Europe B.V. (Amsterdam, the Netherlands) exploiting ABI 3730 (Applied Biosystems, Thermo Fisher Scientific). The sequences were analyzed for SNPs using Sequencher 4.10.1 software (Gene Codes Corporation, Ann Arbor, MI, USA). Sequence data was analyzed further by calculating allele frequencies for found SNPs in the study material.

2.4. Imputation of the identified exonic SNPs and their association analysis with celiac disease and its phenotypes

As SNPs identified by sequencing were not available in the Illumina 610-Quad BeadChip array, and thus not genotyped, those having MAF ≥5% were selected to be phased and imputed using a Finnish population-specific panel of 3775 high-coverage (25–30 ×) whole-genome sequences (SISu v3). Sample-wise, variant-wise, and post-imputation QC was applied as previously described [19]. Phasing of genotyped data was performed with Eagle 2.3.5 (https://data.broadinstitute.org/alkesgroup/Eagle/) and imputation was carried out with Beagle 4.1 (version 08Jun17.d8b, https://faculty.washington.edu/browning/beagle/b4_1.html) as described in the following protocol: dx.doi.org/10.17504/protocols.io.nmndc5e. In the post-imputation QC, SNPs with good imputation quality metrics (INFO score ≥0.8) were included. The same QC filtering criteria applied in the association analysis of genotyped SNPs (section 2.2) were applied to the imputed genetic data. Out of the sequenced exonic variants, nine SNPs were found in the Finnish reference panel, resisted QC filtering, and were thus included in the association analyses.

2.5. Statistical analyses

Association analyses of genotyped SNPs were performed using PLINK v1.07 (https://zzz.bwh.harvard.edu/plink/). Results are presented as odds ratios (OR) with 95% confidence intervals (95% CI). In these analyses, associations were adjusted for multiple testing and small sample size groups by using 10^4 permutation analysis. The generated empirical P_{EMP2} value ≤ 0.05 (uncorrected P value ≤ 0.001) was assumed to be statistically significant. In the imputation analysis, the genotypes' probability dosages were handled by BCFtools (https://samtools.github.io/bcftools/) and the association analysis performed using PLINK 2.0 (https://www.cog-genomics.org/plink/2.0/). The associations of the post-imputation genotype probabilities of the nine SNPs with celiac disease and selected celiac disease phenotypes were tested using the frequentist likelihood score method implemented in SNPTEST v2.5.2. Associations reaching our permutation threshold described above were considered to be statistically significant [19,20].

3. Results

3.1. Association of genotyped functionally annotated CCR9 and CCL25 SNPs with celiac disease and its distinct phenotypes

Of the genotyped *CCR9* SNPs, rs2133660 with an eQTL effect on the expression of CCR9 in whole blood and peripheral blood mononuclear cells (PBMCs) was associated with the presence of malabsorption (OR = 1.45, 95% CI = 1.14–1.83, P = 0.002) although this did not resist correction for multiple testing (P_{EMP2} = 0.064) (Fig. 1, Supplementary Table 2). The same SNP had nominal associations ($P_{EMP2} > 0.05$) with celiac disease (OR = 1.20, 95% CI = 1.02–1.42, P = 0.031), with the presence of gastrointestinal symptoms (OR = 1.29, 95% CI = 1.08–1.54, P = 0.005), anemia (OR = 1.35, 95% CI = 1.01–1.81, P = 0.041), partial villous atrophy (OR = 1.35, 95% CI = 1.01–1.79, P = 0.039), and with negative celiac disease autoantibodies (OR = 3.48, 95% CI = 1.38–8.74, P = 0.005) (Fig. 1). Further, rs12493471 and rs7652331 with RegulomeDB scores ≤ 3a and rs1545985 with an eQTL effect on CCR9 expression were all associated with the presence of gastrointestinal symptoms and malabsorption and rs12493471 also with anemia. Four SNPs (rs12983784, rs952444, rs882951 and rs11667975) in the *CCL25* gene region were nominally associated ($P_{EMP2} > 0.05$) with more than one phenotype, and all had a RegulomeDB score ≤ 3a (Fig. 2). Of these, rs882951 and rs952444, with the lowest RegulomeDB scores (1d and 1f respectively), had significant eQTL effects on CCL25 expression in small intestine and were both associated with the presence of gastrointestinal symptoms and malabsorption (for both SNPs OR = 1.20, 95% CI = 1.01–1.43, P = 0.034 and OR = 1.32, 95% CI = 1.04–1.66, P = 0.021 respectively). Rs12983784 was associated with celiac disease (OR = 1.20, 95% CI = 1.01–1.44, P = 0.041) and positive serum autoantibodies (OR = 1.43, 95% CI = 1.12–1.82, P = 0.004). Rs1129763 and rs11667975 were further associated with the presence of total/subtotal villous atrophy (OR = 0.65, 95% CI = 0.48–0.89, P = 0.006 and OR = 1.31, 95% CI = 1.03–1.68, P = 0.027, respectively), rs11667975 with high HLA risk (OR = 4.20, 95% CI = 1.21–14.53, P = 0.015), and positive serum autoantibodies (OR = 1.36, 95% CI = 1.04–1.79, P = 0.025). The detailed results are presented in Figs. 1 and 2 and Supplementary Table 2.

3.2. Association analysis using sequenced and imputed data

By sequencing the exons of *CCR9* and *CCL25*, we identified two variants in *CCR9* and eight in *CCL25* with MAF ≥5% (Supplementary Table 3). One of the variants found in *CCR9* was previously unreported (referred to as CCR9ex3snp7 in Supplementary Table 3) and changed threonine to alanine at position 100. However, this variant could not be imputed and therefore association analyses were carried out with the remaining nine SNPs. None of the SNPs were associated with celiac disease, but of the *CCL25* SNPs, the synonymous variant rs2303165 was nominally associated ($P_{EMP2} > 0.05$) with partial villous atrophy (OR = 0.39, 95% CI = 0.15–1.00, P = 0.038) (Fig. 2), but the SNP had no eQTL effect on CCL25 expression in the intestine.

4. Discussion

Our study identified several nominal associations for SNPs in the genomic regions of *CCR9* and *CCL25* genes with celiac disease or distinct disease phenotypes. As regards *CCR9*, associations were detected with several genotyped SNPs with potential or proven functional effect but not with any of the exonic imputed ones. None of associated *CCR9* SNPs

Phenotypes	Variant	Regulome score	Significant eQTLs to CCR9	OR [95% CI]	P value	P_{EMP2}
Celiac disease	rs2133660	4	Whole Blood/PBMC	1.20 [1.02-1.42]	0.031	0.558
Gastrointestinal symptoms	rs2133660	4	Whole Blood/PBMC	1.29 [1.08-1.54]	0.005	0.145
	rs1545985	5	Whole Blood/PBMC	1.25 [1.05-1.49]	0.012	0.289
	rs7652331	1f		1.24 [1.03-1.47]	0.012	0.416
	rs12493471	3a		1.23 [1.03-1.47]	0.022	0.439
Malabsorption	rs2133660	4	Whole Blood/PBMC	1.45 [1.14-1.83]	0.002	0.064
	rs12493471	3a		1.39 [1.10-1.75]	0.006	0.174
	rs1545985	5	Whole Blood/PBMC	1.37 [1.09-1.74]	0.008	0.203
	rs7652331	1f		1.35 [1.06-1.70]	0.013	0.318
	rs2191031	3a		1.34 [1.03-1.74]	0.028	0.526
Anemia	rs2191031	3a		1.51 [1.10-2.06]	0.010	0.261
	rs13069079	5	Whole Blood/PBMC	1.53 [1.07-2.20]	0.020	0.416
	rs17214952	7	Whole Blood/PBMC	1.48 [1.03-2.13]	0.035	0.608
	rs2133660	4	Whole Blood/PBMC	1.35 [1.01-1.81]	0.041	0.662
	rs12493471	3a		1.34 [1.00-1.79]	0.047	0.707
Small bowel mucosal damage						
PVA	rs2133660	4	Whole Blood/PBMC	1.35 [1.01-1.79]	0.039	0.645
HLA risk						
High	rs4535265	5	Whole Blood/PBMC	2.42 [1.11-5.28]	0.025	0.536
CeD-antibodies						
Negative CeD-autoantibodies	rs2133660	4	Whole Blood/PBMC	3.48 [1.38-8.74]	0.005	0.152
	rs13069079	5	Whole Blood/PBMC	3.37 [1.34-8.48]	0.006	0.196
	rs17214952	7	Whole Blood/PBMC	3.35 [1.33-8.44]	0.006	0.203
	rs10490770	6	Whole Blood/PBMC	3.10 [1.03-9.34]	0.034	0.586
	rs12639224	5	Whole Blood/PBMC	2.39 [0.99-5.76]	0.045	0.672

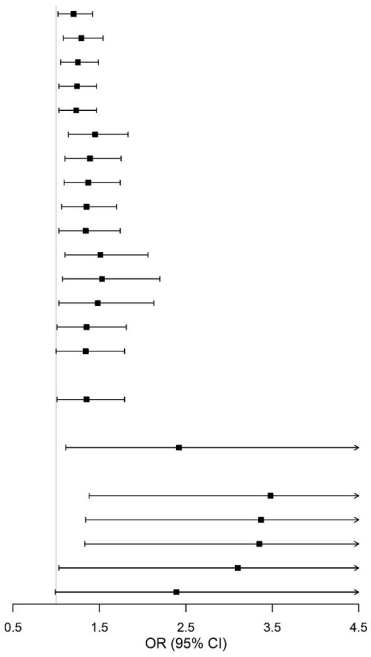

Fig. 1. Forest plot representing association of SNPs in the genomic region of the candidate gene *CCR9* with celiac disease and its different phenotypes. eQTL; expression quantitative trait loci, OR; odds ratio, CI; confidence interval, P_{EMP2}; empirical P value at 10,000 permutation threshold, PBMC; peripheral blood mononuclear cell, PVA; partial villous atrophy, CeD; celiac disease.

Phenotypes	Variant	Regulome score	Significant eQTLs to CCL25	OR [95% CI]	P value	P_{EMP2}
Celiac disease	rs12983784	1f		1.20 [1.01-1.44]	0.041	0.659
Gastrointestinal symptoms	rs952444	1d	SI Terminal Ileum/Ileum	1.20 [1.01-1.43]	0.034	0.624
	rs882951	1f	SI Terminal Ileum/Ileum	1.20 [1.01-1.43]	0.034	0.624
Malabsorption	rs952444	1d	SI Terminal Ileum/Ileum	1.32 [1.04-1.66]	0.021	0.425
	rs882951	1f	SI Terminal Ileum/Ileum	1.32 [1.04-1.66]	0.021	0.425
Anemia	rs4804061	2c		0.51 [0.29-0.90]	0.018	0.383
Small bowel mucosal damage						
TVA/SVA	rs1129763	2b		0.65 [0.48-0.89]	0.006	0.168
	rs11667975	3a		1.31 [1.03-1.68]	0.027	0.513
PVA	rs4804249	2b		0.66 [0.44-1.00]	0.046	0.693
	rs2303165*	4		0.39 [0.15-1.00]	0.038	
HLA risk						
High	rs11667975	3a		4.20 [1.21-14.53]	0.015	0.401
CeD-antibodies						
Positive CeD-autoantibodies	rs12983784	1f		1.43 [1.12-1.82]	0.004	0.114
	rs7257948	3a		1.38 [1.10-1.73]	0.006	0.167
	rs11667975	3a		1.36 [1.04-1.79]	0.025	0.506
Negative CeD-autoantibodies	rs2287936	3a		2.81 [1.07-7.35]	0.028	0.485

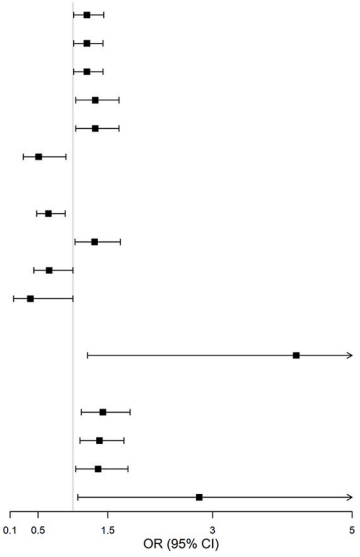

Fig. 2. Forest plot of representing association of SNPs in the genomic region of the candidate gene *CCL25* with celiac disease and its different phenotypes. SNPs identified in sequencing analysis and imputed are indicated by an asterisk. eQTL; expression quantitative trait loci, OR; odds ratio, CI; confidence interval, P_{EMP2}; empirical P value at 10,000 permutation threshold, PBMC; peripheral blood mononuclear cell, TVA/SVA; total or subtotal villous atrophy, PVA; partial villous atrophy, CeD; celiac disease.

had both RegulomeDB score less than 3a and an eQTL effect on CCR9 expression. Of the SNPs with RegulomeDB score less than 3a, rs12493471, associated with gastrointestinal symptoms, malabsorption, and anemia in the present study has previously been associated with celiac disease [21]. SNPs with RegulomeDB scores ranging from 1a to 1f are by definition predicted to affect different DNA regulatory elements and have an eQTL effect based on the Encyclopedia of DNA Elements (ENCODE) project, Gene Expression Omnibus data, and published literature [14,22] and are thus functionally particularly interesting. Rs7652331 at the *CCR9* locus associated with gastrointestinal symptoms and malabsorption falls into this category with RegulomeDB score 1f. However, as the SNP did not have an eQTL on CCR9 expression, it is likely that the effect of the SNP on the phenotype is due to other genes than *CCR9*. Of the SNPs with an eQTL on *CCR9*, rs2133660 mapping to intronic region of *FYCO1* gene downstream of *CCR9* was associated with malabsorption along with an association with celiac disease and several

other phenotypes. However, as the RegulomeDB score of rs2133660 was over 3a, it is possible that the phenotypes associated with this variant rather reflect the effect of a proxy SNP causing increased CCR9 expression in lymphocytes. Moreover, rs2133660 also had eQTL effects on other nearby genes, many of them belonging to different chemokine receptor families. Thus, the effect of the SNP on different phenotypes is hardly attributable solely to increased CCR9 expression.

As regards the SNPs in the genomic region of *CCL25*, genotyped SNPs rs952444 and rs882951 both showed nominal associations with the presence of gastrointestinal symptoms and malabsorption. These variants had the highest likelihood (RegulomeDB score $\leq$ 1f) of being connected to functional transcriptional effects in gene expression and they both had eQTL effects on the expression of CCL25 in the small intestine [15,18]. Although as far as we know no studies have addressed CCL25 expression in the small intestine in different celiac disease phenotypes, the lack of evidence for altered CCL25 expression in celiac disease [23–25] would suggest that the SNPs exert their effects through other mechanisms than CCL25 expression. Sequencing *CCL25* exons followed by imputation revealed that rs2303165 located in *CCL25* gene region was associated with partial villous atrophy, without eQTL *CCL25* effects in intestine. This exonic *CCL25* SNPs could exert its effects on the phenotypes by yet to be determined mechanisms that do not directly involve *CCL25*.

The associations identified in our study did not pass the 10,000 permutation based multiple testing correction. Although there is a debate regarding the number of permutations required, 10,000 permutations provide empirically adjusted P-values with strong evidence of association [26,27]. As the permutation threshold used in the present study is quite stringent and because of our small sample size, we had limited statistical power to detect strong associations. Thus, our findings need to be confirmed in further studies with larger cohorts of carefully phenotyped celiac disease patients. In any case, *CCR9* and *CCL25* variants likely make only a minor contribution to the generation of celiac disease phenotypes addressed in the present study. We have previously reported associations of SNPs outside the *CCR9* and *CCL25* genomic regions with the same phenotypes that were resistant to correction for multiple comparisons. One of these was the association of rs13010713 in integrin subunit alpha 4 gene (*ITGA4*) with the presence of gastrointestinal symptoms, total or subtotal villous atrophy, and intermediate HLA risk [19]. *ITGA4* codes for the α_4 subunit of heterodimeric integrin molecules involved in adhesion and the pairing of α4 subunit with β_7 subunit promotes homing of T cells to intestinal sites [28]. Interestingly, expression of CCR9 and $\alpha_4\beta_7$ on T cells and their subsequent localization to the gut is required for induction of oral tolerance, at least in mice [29]. Thus, due to this solid firm connection between *CCR9* and $\alpha_4\beta_7$, variants found in the integrin locus may possess changes of functionality in the gut-homing pathway which are reflected in celiac disease phenotypes.

Undoubtedly, a given celiac disease phenotype likely also has nongenetic determinants. These may include environmental factors such as microbes, infections or the amount of gluten consumed by a patient, all of which have been associated with the development of celiac disease [30,31]. Moreover, delay in the diagnosis of celiac disease may allow the disease to progress to a more severe form, thus affecting some of the phenotypes at the time of diagnosis. Data showing that at least diarrhea, abdominal pain, and malabsorption are associated with long diagnostic delay lends credence to this hypothesis [32].

The main strength of the present study is the carefully phenotyped cohort of celiac disease patients. In addition, the exploitation of the imputed data allowed us to study the exonic SNPs that were not present in the Illumina 610-Quad BeadChip array previously used to study celiac disease associations. As a limitation, the sequencing approach was performed using Sanger sequencing, which has its shortcomings in accuracy compared to Next-Generation Sequencing (NGS) techniques with more comprehensive coverage and higher sensitivity to detect low-frequency variants. To overcome this, we used MAF $\geq$5% as our cut-off. However, this may have resulted in missing interesting variants, including rs12721497 in *CCR9* with MAF = 0.01, according to the 1000 Genomes Project, associated with acute and chronic stage graft versus host disease [33]. In addition, unfortunately eQTL data, blood or small bowel mucosal samples from celiac disease patients were not available which precluded us from addressing eQTL effects particularly in patients or studying the effects of the SNPs on immunological changes in celiac disease more generally. Moreover, the number of individuals in our study cohort was rather small, particularly for genotype-phenotype association analyses.

5. Conclusions

We conclude that SNPS in the region of *CCR9* and *CCL25* having probable functional effect or being located in exons are weakly associated with various celiac disease phenotypes. Our results thus suggest that, regardless of the importance of *CCR9* and *CCL25* in maintaining gut homeostasis and tolerance, variation within these genomic regions likely makes a minor contribution to the phenotype of celiac disease.

Funding

This study was supported by the Academy of Finland, the Sigrid Juselius Foundation, the Päivikki and Sakari Sohlberg Foundation, the Paulo Foundation, the Foundation for Pediatric Research, and the Competitive State Research Financing of the Expert Area of Tampere University Hospital. Funding sources were not involved in the design of the study, in the collection, analysis or interpretation of data, or in the decision to submit the article for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtauto.2021.100128.

Credit author statement

LA, JXMC, MM, KKu, KKa and KL: Conceptualization. LA, JXMC, PS, KKu, EK, AS, AP, KKa and KL: Data curation. LA, JXMC, HH, PS, DAY, EK, AS, AP and KL: Formal analysis. LA, JXMC, MM, KKu, KKa and KL: Funding acquisition. LA, JXMC, HH, PS, DAY and KL: Investigation. LA, JXMC and KL: Writing – original draft. Resources: MM and KKa. Supervision: KL. All the authors performed Writing – review & editing.

References

[1] K. Lindfors, C. Ciacci, K. Kurppa, K.E.A. Lundin, G.K. Makharia, M.L. Mearin, J. A. Murray, E.F. Verdu, K. Kaukinen, Coeliac disease, Nat. Rev. Dis. Prim. 5 (2019) 1–18, https://doi.org/10.1038/s41572-018-0054-z.

[2] P. Margaritte-Jeannin, M.C. Babron, M. Bourgey, A.S. Louka, F. Clot, S. Percopo, I. Coto, J.P. Hugot, H. Ascher, L.M. Sollid, L. Greco, F. Clerget-Darpoux, HLA-DQ relative risks for coeliac disease in European populations: a study of the European Genetics Cluster on Coeliac Disease, Tissue Antigens 63 (2004) 562–567, https://doi.org/10.1111/j.0001-2815.2004.00237.x.

[3] M.M. Pietzak, T.C. Schofield, M.J. McGinniss, R.M. Nakamura, Stratifying risk for celiac disease in a large at-risk United States population by using HLA alleles, Clin. Gastroenterol. Hepatol. 7 (2009) 966–971, https://doi.org/10.1016/j.cgh.2009.05.028.

[4] K.A. Papadakis, J. Prehn, V. Nelson, L. Cheng, S.W. Binder, P.D. Ponath, D. P. Andrew, S.R. Targan, The role of thymus-expressed chemokine and its receptor CCR9 on lymphocytes in the regional specialization of the mucosal immune system, J. Immunol. 165 (2000) 5069–5076, https://doi.org/10.4049/jimmunol.165.9.5069.

[5] M. Pathak, G. Lal, The regulatory function of CCR9+ dendritic cells in inflammation and autoimmunity, Front. Immunol. 11 (2020), https://doi.org/10.3389/fimmu.2020.536326.
[6] C. Escudero-Hernández, Á. Martín, R. de Pedro Andrés, L. Fernández-Salazar, J. A. Garrote, D. Bernardo, E. Arranz, Circulating dendritic cells from celiac disease patients display a gut-homing profile and are differentially modulated by different gliadin-derived peptides, Mol. Nutr. Food Res. 64 (2020), https://doi.org/10.1002/mnfr.201900989.
[7] R.W. Olaussen, M.R. Karlsson, K.E.A. Lundin, J. Jahnsen, P. Brandtzaeg, I. N. Farstad, Reduced chemokine receptor 9 on intraepithelial lymphocytes in celiac disease suggests persistent epithelial activation, Gastroenterology 132 (2007) 2371–2382, https://doi.org/10.1053/j.gastro.2007.04.023.
[8] A Phase II Study of CCX282-B in Patients With Celiac Disease - Full Text View - ClinicalTrials.gov, (n.d.). https://www.clinicaltrials.gov/ct2/show/NCT00540657 (accessed April 26, 2021).
[9] P.C.A. Dubois, G. Trynka, L. Franke, K.A. Hunt, J. Romanos, A. Curtotti, A. Zhernakova, G.A.R. Heap, R. Ádány, A. Aromaa, M.T. Bardella, L.H. van den Berg, N.A. Bockett, E.G. de la Concha, B. Dema, R.S.N. Fehrmann, M. Fernández-Arquero, S. Fiatal, E. Grandone, P.M. Green, H.J.M. Groen, R. Gwilliam, R.H. J. Houwen, S.E. Hunt, K. Kaukinen, D. Kelleher, I. Korponay-Szabo, K. Kurppa, P. MacMathuna, M. Mäki, M.C. Mazzilli, O.T. McCann, M.L. Mearin, C.A. Mein, M. M. Mirza, V. Mistry, B. Mora, K.I. Morley, C.J. Mulder, J.A. Murray, C. Núñez, E. Oosterom, R.A. Ophoff, I. Polanco, L. Peltonen, M. Platteel, A. Rybak, V. Salomaa, J.J. Schweizer, M.P. Sperandeo, G.J. Tack, G. Turner, J.H. Veldink, W. H.M. Verbeek, R.K. Weersma, V.M. Wolters, E. Urcelay, B. Cukrowska, L. Greco, S. L. Neuhausen, R. McManus, D. Barisani, P. Deloukas, J.C. Barrett, P. Saavalainen, C. Wijmenga, D.A. van Heel, Multiple common variants for celiac disease influencing immune gene expression, Nat. Genet. 42 (2010) 295–302, https://doi.org/10.1038/ng.543.
[10] G. Trynka, K.A. Hunt, N.A. Bockett, J. Romanos, V. Mistry, A. Szperl, S.F. Bakker, M.T. Bardella, L. Bhaw-Rosun, G. Castillejo, E.G. de la Concha, R.C. de Almeida, K.-R.M. Dias, C.C. van Diemen, P.C.A. Dubois, R.H. Duerr, S. Edkins, L. Franke, K. Fransen, J. Gutierrez, G.A.R. Heap, B. Hrdlickova, S. Hunt, L.P. Izurieta, V. Izzo, L.A.B. Joosten, C. Langford, M.C. Mazzilli, C.A. Mein, V. Midah, M. Mitrovic, B. Mora, M. Morelli, S. Nutland, C. Núñez, S. Onengut-Gumuscu, K. Pearce, M. Platteel, I. Polanco, S. Potter, C. Ribes-Koninckx, I. Ricaño-Ponce, S.S. Rich, A. Rybak, J.L. Santiago, S. Senapati, A. Sood, H. Szajewska, R. Troncone, J. Varadé, C. Wallace, V.M. Wolters, A. Zhernakova, B.K. Thelma, B. Cukrowska, E. Urcelay, J.R. Bilbao, M.L. Mearin, D. Barisani, J.C. Barrett, V. Plagnol, P. Deloukas, C. Wijmenga, D.A. van Heel, P. Deloukas, C. Wijmenga, D.A. van Heel, Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease, Nat. Genet. 43 (2011) 1193–1201, https://doi.org/10.1038/ng.998.
[11] A. Sharma, X. Liu, D. Hadley, W. Hagopian, E. Liu, W.-M. Chen, S. Onengut-Gumuscu, V. Simell, M. Rewers, A.-G. Ziegler, A. Lernmark, O. Simell, J. Toppari, J.P. Krischer, B. Akolkar, S.S. Rich, D. Agardh, J.-X. She, Identification of non-HLA genes associated with celiac disease and country-specific differences in a large, international pediatric cohort, PLoS One 11 (2016), e0152476, https://doi.org/10.1371/journal.pone.0152476.
[12] K. Borodulin, H. Tolonen, P. Jousilahti, A. Jula, A. Juolevi, S. Koskinen, K. Kuulasmaa, T. Laatikainen, S. Männistö, M. Peltonen, M. Perola, P. Puska, V. Salomaa, J. Sundvall, S.M. Virtanen, E. Vartiainen, Cohort profile: the national FINRISK study, Int. J. Epidemiol. 47 (2018), https://doi.org/10.1093/ije/dyx239, 696-696i.
[13] K. Watanabe, E. Taskesen, A. Van Bochoven, D. Posthuma, Functional mapping and annotation of genetic associations with FUMA, Nat. Commun. 8 (2017) 1–11, https://doi.org/10.1038/s41467-017-01261-5.
[14] A.P. Boyle, E.L. Hong, M. Hariharan, Y. Cheng, M.A. Schaub, M. Kasowski, K. J. Karczewski, J. Park, B.C. Hitz, S. Weng, J.M. Cherry, M. Snyder, Annotation of functional variation in personal genomes using RegulomeDB, Genome Res. 22 (2012) 1790–1797, https://doi.org/10.1101/gr.137323.112.
[15] F. Aguet, A.N. Barbeira, R. Bonazzola, A. Brown, S.E. Castel, B. Jo, S. Kasela, S. Kim-Hellmuth, Y. Liang, M. Oliva, E.D. Flynn, P. Parsana, L. Fresard, E. R. Gamazon, A.R. Hamel, Y. He, F. Hormozdiari, P. Mohammadi, M. Muñoz-Aguirre, Y.S. Park, A. Saha, A.V. Segrè, B.J. Strober, X. Wen, V. Wucher, K. G. Ardlie, A. Battle, C.D. Brown, N. Cox, S. Das, E.T. Dermitzakis, B.E. Engelhardt, D. Garrido-Martín, N.R. Gay, G.A. Getz, R. Guigó, R.E. Handsaker, P.J. Hoffman, H. K. Im, S. Kashin, A. Kwong, T. Lappalainen, X. Li, D.G. MacArthur, S. B. Montgomery, J.M. Rouhana, M. Stephens, B.E. Stranger, E. Todres, A. Viñuela, G. Wang, Y. Zou, S. Anand, S. Gabriel, A. Graubert, K. Hadley, K.H. Huang, S. R. Meier, J.L. Nedzel, D.T. Nguyen, B. Balliu, D.F. Conrad, D.J. Cotter, O. M. deGoede, J. Einson, E. Eskin, T.Y. Eulalio, N.M. Ferraro, M.J. Gloudemans, L. Hou, M. Kellis, X. Li, S. Mangul, D.C. Nachun, A.B. Nobel, Y. Park, A.S. Rao, F. Reverter, C. Sabatti, A.D. Skol, N.A. Teran, F. Wright, P.G. Ferreira, G. Li, M. Melé, E. Yeger-Lotem, M.E. Barcus, D. Bradbury, T. Krubit, J.A. McLean, L. Qi, K. Robinson, N.V. Roche, A.M. Smith, L. Sobin, D.E. Tabor, A. Undale, J. Bridge, L. E. Brigham, B.A. Foster, B.M. Gillard, R. Hasz, M. Hunter, C. Johns, M. Johnson, E. Karasik, G. Kopen, W.F. Leinweber, A. McDonald, M.T. Moser, K. Myer, K. D. Ramsey, B. Roe, S. Shad, J.A. Thomas, G. Walters, M. Washington, J. Wheeler, S. D. Jewell, D.C. Rohrer, D.R. Valley, D.A. Davis, D.C. Mash, P.A. Branton, L. Sobin, L.K. Barker, H.M. Gardiner, M. Mosavel, L.A. Siminoff, P. Flicek, M. Haeussler, T. Juettemann, W.J. Kent, C.M. Lee, C.C. Powell, K.R. Rosenbloom, M. Ruffier, D. Sheppard, K. Taylor, S.J. Trevanion, D.R. Zerbino, N.S. Abell, J. Akey, L. Chen, K. Demanelis, J.A. Doherty, A.P. Feinberg, K.D. Hansen, P.F. Hickey, L. Hou, F. Jasmine, L. Jiang, R. Kaul, M. Kellis, M.G. Kibriya, J.B. Li, Q. Li, S. Lin, S. E. Linder, B.L. Pierce, L.F. Rizzardi, K.S. Smith, M. Snyder, J. Stamatoyannopoulos, H. Tang, M. Wang, P.A. Branton, L.J. Carithers, P. Guan, S.E. Koester, A.R. Little, H. M. Moore, C.R. Nierras, A.K. Rao, J.B. Vaught, S. Volpi, The GTEx Consortium atlas of genetic regulatory effects across human tissues, Science 369 (80) (2020) 1318–1330, https://doi.org/10.1126/SCIENCE.AAZ1776.
[16] H.J. Westra, M.J. Peters, T. Esko, H. Yaghootkar, C. Schurmann, J. Kettunen, M. W. Christiansen, B.P. Fairfax, K. Schramm, J.E. Powell, A. Zhernakova, D. V. Zhernakova, J.H. Veldink, L.H. Van Den Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P.A.C. Hoen, E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A.B. Singleton, D.G. Hernandez, M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Völker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey, S.A. Gharib, D.A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder, M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R.C. Jansen, P.M. Visscher, J.C. Knight, B.M. Psaty, S. Ripatti, A. Teumer, T.M. Frayling, A. Metspalu, J.B.J. Van Meurs, L. Franke, Systematic identification of trans eQTLs as putative drivers of known disease associations, Nat. Genet. 45 (2013) 1238–1243, https://doi.org/10.1038/ng.2756.
[17] D.V. Zhernakova, P. Deelen, M. Vermaat, M. Van Iterson, M. Van Galen, W. Arindrarto, P. Van't Hof, H. Mei, F. Van Dijk, H.J. Westra, M.J. Bonder, J. Van Rooij, M. Verkerk, P.M. Jhamai, M. Moed, S.M. Kielbasa, J. Bot, I. Nooren, R. Pool, J. Van Dongen, J.J. Hottenga, C.D.A. Stehouwer, C.J.H. Van Der Kallen, C. G. Schalkwijk, A. Zhernakova, Y. Li, E.F. Tigchelaar, N. De Klein, M. Beekman, J. Deelen, D. Van Heemst, L.H. Van Den Berg, A. Hofman, A.G. Uitterlinden, M.M. J. Van Greevenbroek, J.H. Veldink, D.I. Boomsma, C.M. Van Duijn, C. Wijmenga, P. E. Slagboom, M.A. Swertz, A. Isaacs, J.B.J. Van Meurs, R. Jansen, B.T. Heijmans, P. A.C. Hoen't, L. Franke, Identification of context-dependent expression quantitative trait loci in whole blood, Nat. Genet. 49 (2017) 139–145, https://doi.org/10.1038/ng.3737.
[18] Y. Momozawa, J. Dmitrieva, E. Théâtre, V. Deffontaine, S. Rahmouni, B. Charloteaux, F. Crins, E. Docampo, M. Elansary, A.S. Gori, C. Lecut, R. Mariman, M. Mni, C. Oury, I. Altukhov, D. Alexeev, Y. Aulchenko, L. Amininejad, G. Bouma, F. Hoentjen, M. Löwenberg, B. Oldenburg, M.J. Pierik, A.E. Vander Meulen-De Jong, C.J. Van Der Woude, M.C. Visschedijk, M. Lathrop, J.P. Hugot, R. K. Weersma, M. De Vos, D. Franchimont, S. Vermeire, M. Kubo, E. Louis, M. Georges, C. Abraham, J.P. Achkar, T. Ahmad, A.N. Ananthakrishnan, V. Andersen, C.A. Anderson, J.M. Andrews, V. Annese, G. Aumais, L. Baidoo, R. N. Baldassano, P.A. Bampton, M. Barclay, J.C. Barrett, T.M. Bayless, J. Bethge, A. Bitton, G. Boucher, S. Brand, B. Brandt, S.R. Brant, C. Büning, A. Chew, J.H. Cho, I. Cleynen, A. Cohain, A. Croft, M.J. Daly, M. D'Amato, S. Danese, D. De Jong, G. Denapiene, L.A. Denson, K.L. Devaney, O. Dewit, R. D'Inca, M. Dubinsky, R. H. Duerr, C. Edwards, D. Ellinghaus, J. Essers, L.R. Ferguson, E.A. Festen, P. Fleshner, T. Florin, A. Franke, K. Fransen, R. Gearry, C. Gieger, J. Glas, P. Goyette, T. Green, A.M. Griffiths, S.L. Guthery, H. Hakonarson, J. Halfvarson, K. Hanigan, T. Haritunians, A. Hart, C. Hawkey, N.K. Hayward, M. Hedl, P. Henderson, X. Hu, H. Huang, K.Y. Hui, M. Imielinski, A. Ippoliti, L. Jonaitis, L. Jostins, T.H. Karlsen, N.A. Kennedy, M.A. Khan, G. Kiudelis, K. Krishnaprasad, S. Kugathasan, L. Kupcinskas, A. Latiano, D. Laukens, I.C. Lawrance, J.C. Lee, C. W. Lees, M. Leja, J. Van Limbergen, P. Lionetti, J.Z. Liu, G. Mahy, J. Mansfield, D. Massey, C.G. Mathew, D.P.B. McGovern, R. Milgrom, M. Mitrovic, G. W. Montgomery, C. Mowat, W. Newman, A. Ng, S.C. Ng, S.M.E. Ng, S. Nikolaus, K. Ning, M. Nöthen, I. Oikonomou, O. Palmieri, M. Parkes, A. Phillips, C. Y. Ponsioen, U. Potocnik, N.J. Prescott, D.D. Proctor, G. Radford-Smith, J.F. Rahier, S. Raychaudhuri, M. Regueiro, F. Rieder, J.D. Rioux, S. Ripke, R. Roberts, R. K. Russell, J.D. Sanderson, M. Sans, J. Satsangi, E.E. Schadt, S. Schreiber, D. Schulte, L.P. Schumm, R. Scott, M. Seielstad, Y. Sharma, M.S. Silverberg, L. A. Simms, J. Skieceviciene, S.L. Spain, A.H. Steinhart, J.M. Stempak, L. Stronati, J. Sventoraityte, S.R. Targan, K.M. Taylor, A. Ter Velde, L. Torkvist, M. Tremelling, S. Van Sommeren, E. Vasiliauskas, H.W. Verspaget, T. Walters, K. Wang, M. H. Wang, Z. Wei, D. Whiteman, C. Wijmenga, D.C. Wilson, J. Winkelmann, R. J. Xavier, B. Zhang, C.K. Zhang, H. Zhang, W. Zhang, H. Zhao, Z.Z. Zhao, IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes, Nat. Commun. 9 (2018), https://doi.org/10.1038/s41467-018-04365-8.
[19] J.X.M. Cerqueira, P. Saavalainen, K. Kurppa, P. Laurikka, H. Huhtala, M. Nykter, L. L.E. Koskinen, D.A. Yohannes, E. Kilpeläinen, A. Shcherban, A. Palotie, K. Kaukinen, K. Lindfors, Independent and cumulative coeliac disease-susceptibility loci are associated with distinct disease phenotypes, J. Hum. Genet. (2021) 1–11, https://doi.org/10.1038/s10038-020-00888-5.
[20] J. Marchini, B. Howie, S. Myers, G. McVean, P. Donnelly, A new multipoint method for genome-wide association studies by imputation of genotypes, Nat. Genet. 39 (2007) 906–913, https://doi.org/10.1038/ng2088.
[21] A. Sharma, X. Liu, D. Hadley, W. Hagopian, E. Liu, W.M. Chen, S. Onengut-Gumuscu, V. Simell, M. Rewers, A.G. Ziegler, Å. Lernmark, O. Simell, J. Toppari, J. P. Krischer, B. Akolkar, S.S. Rich, D. Agardh, J.X. She, Identification of non-HLA genes associated with celiac disease and country-specific differences in a large, international pediatric cohort, PLoS One 11 (2016), https://doi.org/10.1371/journal.pone.0152476.
[22] I. Dunham, A. Kundaje, S.F. Aldred, P.J. Collins, C.A. Davis, F. Doyle, C.B. Epstein, S. Frietze, J. Harrow, R. Kaul, J. Khatun, B.R. Lajoie, S.G. Landt, B.K. Lee, F. Pauli, K.R. Rosenbloom, P. Sabo, A. Safi, A. Sanyal, N. Shoresh, J.M. Simon, L. Song, N. D. Trinklein, R.C. Altshuler, E. Birney, J.B. Brown, C. Cheng, S. Djebali, X. Dong, J. Ernst, T.S. Furey, M. Gerstein, B. Giardine, M. Greven, R.C. Hardison, R.S. Harris, J. Herrero, M.M. Hoffman, S. Iyer, M. Kellis, P. Kheradpour, T. Lassmann, Q. Li, X. Lin, G.K. Marinov, A. Merkel, A. Mortazavi, S.C.J. Parker, T.E. Reddy, J. Rozowsky, F. Schlesinger, R.E. Thurman, J. Wang, L.D. Ward, T.W. Whitfield, S. P. Wilder, W. Wu, H.S. Xi, K.Y. Yip, J. Zhuang, B.E. Bernstein, E.D. Green, C. Gunter, M. Snyder, M.J. Pazin, R.F. Lowdon, L.A.L. Dillon, L.B. Adams, C.

J. Kelly, J. Zhang, J.R. Wexler, P.J. Good, E.A. Feingold, G.E. Crawford, J. Dekker, L. Elnitski, P.J. Farnham, M.C. Giddings, T.R. Gingeras, R. Guigó, T.J. Hubbard, W. J. Kent, J.D. Lieb, E.H. Margulies, R.M. Myers, J.A. Stamatoyannopoulos, S. A. Tenenbaum, Z. Weng, K.P. White, B. Wold, Y. Yu, J. Wrobel, B.A. Risk, H. P. Gunawardena, H.C. Kuiper, C.W. Maier, L. Xie, X. Chen, T.S. Mikkelsen, S. Gillespie, A. Goren, O. Ram, X. Zhang, L. Wang, R. Issner, M.J. Coyne, T. Durham, M. Ku, T. Truong, M.L. Eaton, A. Dobin, A. Tanzer, J. Lagarde, W. Lin, C. Xue, B.A. Williams, C. Zaleski, M. Röder, F. Kokocinski, R.F. Abdelhamid, T. Alioto, I. Antoshechkin, M.T. Baer, P. Batut, I. Bell, K. Bell, S. Chakrabortty, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R. Duttagupta, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M.J. Fullwood, H. Gao, D. Gonzalez, A. Gordon, C. Howald, S. Jha, R. Johnson, P. Kapranov, B. King, C. Kingswood, G. Li, O.J. Luo, E. Park, J.B. Preall, K. Presaud, P. Ribeca, D. Robyr, X. Ruan, M. Sammeth, K.S. Sandhu, L. Schaeffer, L.H. See, A. Shahab, J. Skancke, A.M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, Y. Hayashizaki, A. Reymond, S.E. Antonarakis, G.J. Hannon, Y. Ruan, P. Carninci, C.A. Sloan, K. Learned, V.S. Malladi, M.C. Wong, G.P. Barber, M.S. Cline, T. R. Dreszer, S.G. Heitner, D. Karolchik, V.M. Kirkup, L.R. Meyer, J.C. Long, M. Maddren, B.J. Raney, L.L. Grasfeder, P.G. Giresi, A. Battenhouse, N.C. Sheffield, K.A. Showers, D. London, A.A. Bhinge, C. Shestak, M.R. Schaner, S.K. Kim, Z. Z. Zhang, P.A. Mieczkowski, J.O. Mieczkowska, Z. Liu, R.M. McDaniell, Y. Ni, N. U. Rashid, M.J. Kim, S. Adar, Z. Zhang, T. Wang, D. Winter, D. Keefe, V.R. Iyer, M. Zheng, P. Wang, J. Gertz, J. Vielmetter, E.C. Partridge, K.E. Varley, C. Gasper, A. Bansal, S. Pepke, P. Jain, H. Amrhein, K.M. Bowling, M. Anaya, M.K. Cross, M. A. Muratet, K.M. Newberry, K. McCue, A.S. Nesmith, K.I. Fisher-Aylor, B. Pusey, G. DeSalvo, S.L. Parker, S. Balasubramanian, N.S. Davis, S.K. Meadows, T. Eggleston, J.S. Newberry, S.E. Levy, D.M. Absher, W.H. Wong, M.J. Blow, A. Visel, L.A. Pennachio, H.M. Petrykowska, A. Abyzov, B. Aken, D. Barrell, G. Barson, A. Berry, A. Bignell, V. Boychenko, G. Bussotti, C. Davidson, G. Despacio-Reyes, M. Diekhans, I. Ezkurdia, A. Frankish, J. Gilbert, J.M. Gonzalez, E. Griffiths, R. Harte, D.A. Hendrix, T. Hunt, I. Jungreis, M. Kay, E. Khurana, J. Leng, M.F. Lin, J. Loveland, Z. Lu, D. Manthravadi, M. Mariotti, J. Mudge, G. Mukherjee, C. Notredame, B. Pei, J.M. Rodriguez, G. Saunders, A. Sboner, S. Searle, C. Sisu, C. Snow, C. Steward, E. Tapanari, M.L. Tress, M.J. Van Baren, S. Washietl, L. Wilming, A. Zadissa, Z. Zhang, M. Brent, D. Haussler, A. Valencia, N. Addleman, R.P. Alexander, R.K. Auerbach, S. Balasubramanian, K. Bettinger, N. Bhardwaj, A.P. Boyle, A.R. Cao, P. Cayting, A. Charos, Y. Cheng, C. Eastman, G. Euskirchen, J.D. Fleming, F. Grubert, L. Habegger, M. Hariharan, A. Harmanci, S. Iyengar, V.X. Jin, K.J. Karczewski, M. Kasowski, P. Lacroute, H. Lam, N. Lamarre-Vincent, J. Lian, M. Lindahl-Allen, R. Min, B. Miotto, H. Monahan, Z. Moqtaderi, X.J. Mu, H. O'Geen, Z. Ouyang, D. Patacsil, D. Raha, L. Ramirez, B. Reed, M. Shi, T. Slifer, H. Witt, L. Wu, X. Xu, K.K. Yan, X. Yang, K. Struhl, S. M. Weissman, L.O. Penalva, S. Karmakar, R.R. Bhanvadia, A. Choudhury, M. Domanus, L. Ma, J. Moran, A. Victorsen, T. Auer, L. Centanin, M. Eichenlaub, F. Gruhl, S. Heermann, B. Hoeckendorf, D. Inoue, T. Kellner, S. Kirchmaier, C. Mueller, R. Reinhardt, L. Schertel, S. Schneider, R. Sinn, B. Wittbrodt, J. Wittbrodt, G. Jain, G. Balasundaram, D.L. Bates, R. Byron, T.K. Canfield, M. J. Diegel, D. Dunn, A.K. Ebersol, T. Frum, K. Garg, E. Gist, R.S. Hansen, L. Boatman, E. Haugen, R. Humbert, A.K. Johnson, E.M. Johnson, T.V. Kutyavin, K. Lee, D. Lotakis, M.T. Maurano, S.J. Neph, F.V. Neri, E.D. Nguyen, H. Qu, A.P. Reynolds, V. Roach, E. Rynes, M.E. Sanchez, R.S. Sandstrom, A.O. Shafer, A.B. Stergachis, S. Thomas, B. Vernot, J. Vierstra, S. Vong, H. Wang, M.A. Weaver, Y. Yan, M. Zhang, J.M. Akey, M. Bender, M.O. Dorschner, M. Groudine, M.J. MacCoss, P. Navas, G. Stamatoyannopoulos, K. Beal, A. Brazma, P. Flicek, N. Johnson, M. Lukk, N.M. Luscombe, D. Sobral, J.M. Vaquerizas, S. Batzoglou, A. Sidow, N. Hussami, S. Kyriazopoulou-Panagiotopoulou, M.W. Libbrecht, M.A. Schaub, W. Miller, P.J. Bickel, B. Banfai, N.P. Boley, H. Huang, J.J. Li, W.S. Noble, J. A. Bilmes, O.J. Buske, A.D. Sahu, P.V. Kharchenko, P.J. Park, D. Baker, J. Taylor, L. Lochovsky, An integrated encyclopedia of DNA elements in the human genome, Nature 489 (2012) 57–74, https://doi.org/10.1038/nature11247.

[23] H. Bragde, U. Jansson, M. Fredrikson, E. Grodzinsky, J. Söderman, Celiac disease biomarkers identified by transcriptome analysis of small intestinal biopsies, Cell. Mol. Life Sci. 75 (2018) 4385–4401, https://doi.org/10.1007/s00018-018-2898-5.

[24] M.M. Leonard, Y. Bai, G. Serena, K.P. Nickerson, S. Camhi, C. Sturgeon, S. Yan, M. R. Fiorentino, A. Katz, B. Nath, J. Richter, M. Sleeman, C. Gurer, A. Fasano, RNA sequencing of intestinal mucosa reveals novel pathways functionally linked to celiac disease pathogenesis, PLoS One 14 (2019), https://doi.org/10.1371/journal.pone.0215132.

[25] V. Dotsenko, M. Oittinen, J. Taavela, A. Popp, M. Peräaho, S. Staff, J. Sarin, F. Leon, J. Isola, M. Mäki, K. Viiri, Genome-wide transcriptomic analysis of intestinal mucosa in celiac disease patients on a gluten-free diet and postgluten challenge, CMGH 11 (2021) 13–32, https://doi.org/10.1016/j.jcmgh.2020.07.010.

[26] B. Han, H.M. Kang, E. Eskin, Rapid and accurate multiple testing correction and power estimation for millions of correlated markers, PLoS Genet. 5 (2009), https://doi.org/10.1371/journal.pgen.1000456.

[27] J.M. Kunert-Graf, N.A. Sakhanenko, D.J. Galas, Optimized permutation testing for information theoretic measures of multi-gene interactions, BMC Bioinf. 22 (2021) 180, https://doi.org/10.1186/s12859-021-04107-6.

[28] A. Hamann, D.P. Andrew, D. Jablonski-Westrich, B. Holzmann, E.C. Butcher, Role of alpha 4-integrins in lymphocyte homing to mucosal tissues in vivo, J. Immunol. 152 (1994).

[29] B. Cassani, E.J. Villablanca, F.J. Quintana, P.E. Love, A. Lacy-Hulbert, W.S. Blaner, T. Sparwasser, S.B. Snapper, H.L. Weiner, J.R. Mora, Gut-tropic T cells that express integrin α4β7 and CCR9 Are required for induction of oral immune tolerance in mice, Gastroenterology 141 (2011) 2109–2118, https://doi.org/10.1053/j.gastro.2011.09.015.

[30] K. Størdal, C. Kahrs, G. Tapia, D. Agardh, K. Kurppa, L.C. Stene, Review article: exposure to microbes and risk of coeliac disease, Aliment. Pharmacol. Ther. 53 (2021) 43–62, https://doi.org/10.1111/apt.16161.

[31] C. Andrén Aronsson, H.S. Lee, E.M. Hård Af Segerstad, U. Uusitalo, J. Yang, S. Koletzko, E. Liu, K. Kurppa, P.J. Bingley, J. Toppari, A.G. Ziegler, J.X. She, W. A. Hagopian, M. Rewers, B. Akolkar, J.P. Krischer, S.M. Virtanen, J.M. Norris, D. Agardh, Association of gluten intake during the first 5 years of life with incidence of celiac disease autoimmunity and celiac disease among children at increased risk, JAMA, J. Am. Med. Assoc. 322 (2019) 514–523, https://doi.org/10.1001/jama.2019.10329.

[32] V. Fuchs, K. Kurppa, H. Huhtala, P. Collin, M. Mäki, K. Kaukinen, Factors associated with long diagnostic delay in celiac disease, Scand. J. Gastroenterol. 49 (2014) 1304–1310, https://doi.org/10.3109/00365521.2014.923502.

[33] Y. Inamoto, M. Murata, A. Katsumi, Y. Kuwatsuka, A. Tsujimura, Y. Ishikawa, K. Sugimoto, M. Onizuka, S. Terakura, T. Nishida, T. Kanie, H. Taji, H. Iida, R. Suzuki, A. Abe, H. Kiyoi, T. Matsushita, K. Miyamura, Y. Kodera, T. Naoe, Donor single nucleotide polymorphism in the CCR9 gene affects the incidence of skin GVHD, Bone Marrow Transplant. 45 (2010) 363–369, https://doi.org/10.1038/bmt.2009.131.