Tampere University

LAURA AIRAKSINEN

Contribution of Relatedness and Genetic Factors to the Clinical Picture of Coeliac Disease

Tampere University Dissertations 728

Tampere University Dissertations 728

LAURA AIRAKSINEN

Contribution of Relatedness and Genetic Factors to the Clinical Picture of Coeliac Disease

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the auditorium F114
of the Arvo building, Arvo Ylpön katu 34, Tampere,
on 13 January 2023, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Medicine and Health Technology
Finland

Responsible supervisor and Custos	Professor Katri Lindfors Tampere University Finland	
Pre-examiners	Docent Laura Merras-Salmio University of Helsinki Finland	Docent Åsa Torinsson Naluai University of Gothenburg Sweden
Opponent	Associate Professor Jernej Dolinšek University of Maribor Slovenia	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-2712-5 (print)
ISBN 978-952-03-2713-2 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2713-2

Carbon dioxide emissions from printing Tampere University dissertations have been compensated.

PunaMusta Oy – Yliopistopaino
Joensuu 2022

"Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more, so that we may fear less."

Marie Curie

To my family

ACKNOWLEDGEMENTS

The work for this dissertation was carried out at the Celiac Disease Research Center (CeliRes), Faculty of Medicine and Health Technology, Tampere University. I am very grateful to each and every person who has helped, advised, taught, supported, challenged and mentored me during my PhD journey throughout the years. I am also thankful for the financial support from the Finnish Coeliac Disease Society, the Doctoral Programme in Medicine and Life Sciences, University of Tampere Foundation, the Foundation for Pediatric Research, the Finnish Cultural Foundation, the Paulo Foundation, the City of Tampere and the Orion Research Foundation.

I want to express my deepest gratitude towards my supervisor, Professor Katri Lindfors. Your effectiveness, practical attitude and optimism have all served as a true inspiration for me. You have always had time, patience and willingness to provide scientific guidance when needed. You have encouragingly supported the growth of my independence as a researcher, by showing trust and giving freedom. In addition to your intelligence and experience as a scientist, I greatly appreciate your sense of humour. It has made me feel comfortable from day one.

I am most thankful to Professor Katri Kaukinen, Director of CeliRes, for providing facilities and resources for this dissertation. Your infectious excitement for science sets an example to us researchers. You always seem to solve any problem we may face and have an admirable ability to dissolve complex issues in research into more intelligible ones. I am also grateful for the clinical discussions we have had. They have been very valuable for me as a non-clinician. I also wish to express my gratitude to Professor Kalle Kurppa. I appreciate the engagement dedication to research you possess. Your precise and insightful comments as a co-author were of great value. Thank you also for being patient in explaining to me the clinical details and aspects throughout our common study projects. Your readiness always to help and to be consulted has meant a lot to me.

I am profoundly grateful to Emeritus Professor Markku Mäki for opening the doors to the world of coeliac disease research for me. Your inspiring enthusiasm, colossal knowledge and approachable personality had a profound impact on me. So did the whole Coeliac Disease Study Group with its active and international research

atmosphere. And suddenly, a research career, previously appearing somewhat unattractive to me, began to seem the only desirable option. I always felt that you are truly proud of your researchers and are on their side, irrespective of whether they were undergraduate students or experienced research fellows. Thank you, Markku.

Warm thanks are due to my co-authors for their contribution to the thesis articles: Lauri Myllymäki, Professor Teea Salmi, Dawit Yohannes, Elina Kilpeläinen, Anastasia Shcherban and Professor Aarno Palotie. I would especially like to thank Päivi Saavalainen, whom I consider as my godmother in genetics and sequence analyses. Moreover, I have been privileged to have a godmother in statistics too, Heini Huhtala, whose patience and excellence in emphasizing the importance of careful statistical consideration behind every creditable study is highly appreciated. Pilvi Laurikka deserves my warmest thanks for her thorough guidance in the jungle of patient data and clinical parameters as well as for being such a great companion for discussions touching both work and family life. Last but by no means least, I want to thank Juliana Cerqueira, my "partner-in-crime", whose hard-working attitude and perseverance are beyond compare. Juliana, I will be eternally grateful to you for your scientific help and partnership. Even more importantly, however, I cherish our friendship. Thank you for always being there for me, regardless of your geographical location.

I want to acknowledge Professor Johanna Schleutker and Professor Olli Silvennoinen for their valuable comments and advice as members of the thesis committee during the first years of my PhD project. I would also like to cordially thank my pre-examiners, Docent Laura Merras-Salmio and Docent Åsa Torinsson Naluai for their kind words and invaluable comments and amendments which helped me to improve my doctoral thesis. In addition, I wish to warmly thank Virginia Mattila for astonishingly prompt and neat language revision for both article and thesis manuscripts.

From the bottom of my heart, I want to thank all former and present members of CeliRes. I venture to say, that our Keliporukka is something quite unique as a work community. I want to thank especially the backbone of Kelilab throughout the years: Kaija Laurila, Marja-Terttu Oksanen, Anne Heimonen, Mervi Himanka, Soili Peltomäki, Jorma Kulmala and Nina Koivisto. I wish to cordially thank Heidi Kontro for her motherly guidance and sincere help in so many practical things, big and small. My profound gratitude is due to Professor Ilma Korponay-Szabo, who served as my "Hungarian supervisor" during the secondment in Debrezen in 2011. Nagyon köszönöm, Ilma! From the home office to the conferences abroad (and everything in-between) I have always been accompanied by such lovely people. I wish to extend

my heartfelt thanks especially to Kati Juuti-Uusitalo, Esko Kemppainen, Helka Kaunisto, Atte Kukkurainen, Keijo Viiri, Juha Taavela, Saana Kauma, Anna Eurén, Camilla Pasternack, Valma Fuchs, Linnea Aitokari and Sara Koskimaa. Laura Kivelä, thank you for being such a Laura! I wish we will share mutual connection, laughter and singing in parts in future too. Marleena Repo (a.k.a. Martsu), your sarcastic sense of humour has made my day many times. I have always felt certain affinity with you, maybe partly due to your "biotechnologist" mindset. Rakel Nurmi, heartfelt thanks for your valuable peer support during the thesis submission process. It has been great to get to know you better and be able to relate to you because of compatible phases in family life too. Alina Popp, thank you for being such a great support for me, especially when things didn't go quite as planned. I highly appreciate your professionalism, heart and soul. Tiina Rauhavirta, Suvi Kalliokoski and Minna Hietikko, my dear Tinde, Suikka and Minde, it's impossible to find words to describe how important it has been for me to have you as colleagues-turned-into-friends. Without your peer support, care, symphathy and loyalty I wouldn't be at this point now. We have experienced so much together and our funny memories and jokes still make me laugh. Thank you for everything!

I have been blessed with wonderful, loving family and friends whose support, care and trust have been revitalizing and kept me going through all these years. Elina Ojala, thank you for being my long-standing soulmate, inspiration and peer support in many different roles in life we have. Sini Savonen, thank you for being such a loyal friend, always genuinely willing to offer your help and support. Rami Kivistö, thank you for sitting in my dustbin during our common Master's thesis journey. That started a friendship which I hope will never fade. Kaisa Teittinen, thank you for prestigious friendship as well as for being my important mentor in academia and outside of it. Thank you Maiju Kupi, Anu Haapanen, Satu Haapanen and Hanne Piekkola-Fabrin for unforgettable moments and precious discussions off-duty. You are pure gold.

Kaikista syvimmät ja rakkaimmat kiitokseni kuuluvat pienelle, mutta sitäkin tiiviimmälle perheelleni: nonnalle, isopapalle, äidille, Laurille, Adalle, Noelle, Terhille ja Jukalle. Olette aina uskoneet, tukeneet ja rakastaneet ehdoitta sekä auttaneet muistamaan ne työtä tärkeämmät asiat elämässä. Juha, mitkään sanat eivät riitä kertomaan kiitollisuuteni, arvostukseni ja rakkauteni määrää. Tämä(kään) projekti ei olisi mitenkään ollut mahdollista ilman sinua. Rakkaat Aapo ja Elmo, olette kaikkeni.

Tampere, December 2022 Laura

ABSTRACT

Coeliac disease, a systemic autoimmune disorder induced by dietary gluten, is widespread globally, but in Finland even particularly prevalent although still heavily underdiagnosed. One reason contributing to this is the remarkably multifaceted clinical picture of coeliac disease. The originally typical gastrointestinal symptoms are currently being increasingly replaced by a milder presentation with extraintestinal symptoms and even totally asymptomatic presentation is no longer abnormal. Nevertheless, coeliac disease patients do not differ from each other only in terms of symptoms. There is a wide individual variation concerning the onset, severity and progression of symptoms as well as response to dietary treatment i.e., gluten-free diet (GFD). The reasons and contributary factors underlying this wide clinical heterogeneity remain obscure.

For many people nowadays, suspicion of coeliac disease is closely connected to known familial history of the disease. These familial cases are in the majority, but not every patient has affected relatives and such patients are considered sporadic. Where coeliac disease heredity is concerned, HLA-DQ2.5 is the best-known genetic component. However, numerous loci outside the HLA region have also been associated with the disease.

The studies presented in this dissertation focused in investigating whether there are genetic and/or phenotypic differences between familial and sporadic coeliac disease and whether the dose of HLA-DQ2.5 or the presence of genetic variants outside HLA region, including the ones within candidate genes *CCR9* and *CCL25*, contributes to the clinical picture of the disease.

The findings of this dissertation managed to describe sporadic coeliac disease as an independent entity with a distinct HLA-DQ genotype even though not many significant phenotypic differences were observed between familial and sporadic coeliac disease. Nevertheless, the sporadic cases had more severe clinical phenotype at diagnosis as well as poorer overall health even after dietary treatment. Moreover, HLA-DQ2.5-negative coeliac disease patients were observed to present with classical phenotype at diagnosis as well as with persistent symptoms after dietary treatment more often than patients heterozygous or homozygous for high-risk HLA-DQ2.5. These findings will hopefully encourage physicians to pay special attention

to both these patient groups. Four distinct non-HLA variants were associated with increased risk for familial coeliac disease, but the associations need to be confirmed in future studies. The contribution of HLA-DQ2.5 dose as well as non-HLA single nucleotide polymorphisms (SNPs) within *CCR9* and *CCL25* to the clinical picture of coeliac disease was found to be only modest.

TIIVISTELMÄ

Keliakia on yleinen autoimmuunisairaus, jossa tärkein altistava ympäristötekijä on ravinnon sisältämä gluteeni. Keliakiaa esiintyy maailmanlaajuisesti, mutta sen esiintyvyys Suomessa on erityisen korkea, vaikkakin laajalti alidiagnosoitu. Osasyy tälle on keliakian kirjava taudinkuva. Ruuansulatuskanavan oireet muodostivat pitkään keliakialle tyypillisen oirekuvan, mutta se on vähitellen korvautunut lievemmillä, usein ruuansulatuskanavan ulkopuolisilla oireilla. Nykyään keliakia voi ilmentyä myös täysin oireettomana. Oireiden lisäksi niiden puhkeamisessa, voimakkuudessa sekä kehittymisessä on paljon yksilöllisiä eroja. Myös keliakiapotilaiden vaste gluteenittomaan ruokavaliohoitoon vaihtelee. Syitä tai tekijöitä tämän potilaskohtaisesti vaihtelevan taudinkuvan taustalla ei kuitenkaan vielä tunneta.

Nykyisin moni keliaakikko on osannut aiemmin epäillä kohdallaan keliakiaa ja hakeutunut lääkäriin, koska sairaus on ollut tuttu keliakiaa sairastavien sukulaisten tai perheenjäsenten kautta. Nämä ns. familiaaliset keliaakikot ovat enemmistö, mutta keliakiaa esiintyy myös ns. sporadisesti, jolloin potilaalla ei ole tiedettyä sukutaustaa keliakialle. Tunnetuin keliakialle altistava perintötekijä on HLA-DQ2.5-haplotyyppi, mutta muidenkin, HLA-alueen ulkopuolisten, geenialueiden on todettu assosioituvan keliakiaan.

Tässä väitöskirjatutkimuksessa selvitettiin familiaalisen ja sporadisen keliakian eroavaisuuksia perintötekijöiden ja taudinkuvan suhteen. Lisäksi tutkittiin, onko HLA-DQ2.5-kuormalla (eli sillä, onko potilas HLA-DQ2.5-negatiivinen, -heterotsygootti vai -homotsygootti) ja HLA:n ulkopuolisilla geneettisillä varianteilla, mukaan lukien variantit kandidaattigeeneissä *CCR9 ja CCL25*, vaikutusta keliakian kliiniseen taudinkuvaan.

Tämän väitöskirjan tulokset paljastivat, että sporadinen keliakia omana erillisenä tautimuotonaan eroaa HLA-DQ-genotyypiltään familiaalisesta tautimuodosta, vaikka useita eroavaisuuksia kliinisen taudinkuvan suhteen ei löydettykään. Sporadisten potilaiden taudinkuva diagnoosihetkellä oli kuitenkin vakavampi ja yleinen terveydentila seurantahetkellä, gluteenittomalla dieetillä, huonompi kuin familiaalisilla potilailla. Tutkimukset myös osoittivat, että HLA-DQ2.5-negatiivisia potilaita leimasi HLA-DQ2.5-heterotsygootteja ja -homotsygootteja useammin

klassinen oirekuva diagnoosihetkellä sekä pitkittyneet oireet seurantahetkellä. Nämä tutkimuslöydökset toivottavasti rohkaisevat lääkäreitä sekä muuta hoitohenkilökuntaa kiinnittämään erityistä huomiota edellä mainittuihin potilasryhmiin. Lisäksi väitöstutkimuksen geneettisten analyysien tulokset osoittivat neljän HLA-alueen ulkopuolisen variantin assosioituvan familiaaliseen keliakiaan. Näiden assosiaatioiden merkitys tulee kuitenkin vahvistaa jatkotutkimuksilla. HLA-DQ2.5-kuorman tai *CCR9*- ja *CCL25*-kandidaattigeenialueiden varianttien vaikutus keliakian taudinkuvaan oli tämän väitöskirjatutkimuksen perusteella vähäinen.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

AGA	anti-gliadin antibody
APC	antigen-presenting cell
BCR	B cell receptor
CI	confidence interval
CCL25	c-c motif chemokine ligand 25
CCR9	c-c motif chemokine receptor 9
DNA	deoxyribonucleic acid
DGP	deamidated gliadin peptide
ELISA	enzyme-linked immunosorbent assay
eQTL	expression quantitative trait loci
EmA	endomysial antibody
FDR	false discovery rate or first-degree relative
FUMA	Functional Mapping and Annotation
GERD	gastroesophageal reflux disease
GFD	gluten-free diet
GSRS	Gastrointestinal Symptom Rating Scale
GWAS	genome-wide association study
HLA	human leukocyte antigen
IBD	inflammatory bowel disease
IBS	irritable bowel syndrome
IEL	intraepithelial lymphocyte
IFN	interferon
Ig	immunoglobulin
IL	interleukin
LD	linkage disequilibrium
MAF	minor allele frequency
OR	odds ratio
PBMC	peripheral blood mononuclear cell
PCR	polymerase chain reaction
PGWB	Psychological General Well-Being Questionnaire

QC	quality control
RCD	refractory coeliac disease
SF-36	Short Form 36 Questionnaire
SNP	single nucleotide polymorphism
TCR	T cell receptor
TG2	transglutaminase 2
TG3	transglutaminase 3
T1D	type 1 diabetes mellitus
Vh:CrD	villous height – crypt depth ratio

ORIGINAL PUBLICATIONS

This dissertation is based on the following original peer-reviewed publications, referred to in the text as Studies I-III.

Publication I Airaksinen L, Myllymäki L, Kaukinen K, Saavalainen P, Huhtala H, Lindfors K, Kurppa K (2021). Differences between familial and sporadic celiac disease. *Digestive Diseases and Sciences*, *66*(6), 1981–1988. https://doi.org/10.1007/s10620-020-06490-1

Publication II Airaksinen L, Laurikka P, Huhtala H, Kurppa K, Salmi T, Saavalainen P, Kaukinen K, Lindfors K (2020). Influence of HLA-DQ2.5 Dose on Clinical Picture of Unrelated Celiac Disease Patients. *Nutrients*, *12*(12), 3775. https://doi.org/10.3390/nu12123775

Publication III Airaksinen L*, Cerqueira XM. J*, Huhtala H, Saavalainen P, Yohannes DA, 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, 14*(4), 100128. https://doi.org/10.1016/j.jtauto.2021.100128

* Shared first authorship.

AUTHOR'S CONTRIBUTIONS

The author's contribution in the original publication is listed below.

Publication I L. Airaksinen conceptualized, designed and performed all genetic association analyses and related statistical analyses. L. Airaksinen analysed, curated and interpreted genetic data. L. Airaksinen had a principal role in integrating data from both genetic and clinical analyses, in writing and revising the manuscript as well as in producing the tables and figures.

Publication II L. Airaksinen participated in study conceptualization and design and performed all analyses, analysed data and made a major contribution to data interpretation. L. Airaksinen had a principal role in writing, reviewing and editing the manuscript as well as in creating all tables and figures.

Publication III L. Airaksinen performed all the laboratory work related to sequencing project (i.e., sample preparation, DNA extraction, primer design and optimization and PCR) as well as sequence analyses, association analyses and interpretation of sequenced data. J. Cerqueira performed the association analyses, functional annotation and statistical analyses of genotyped and imputed data. L. Airaksinen and J. Cerqueira participated equally in data interpretation of the association analyses, in writing and revising the manuscript and in the production of tables and figures.

1 INTRODUCTION

Coeliac disease is one of the most common life-long gastrointestinal diseases globally (Catassi et al., 2022). It is a systemic autoimmune condition driven by the ingestion of dietary gluten present in wheat, rye and barley, in genetically susceptible individuals. Certain variants of human leukocyte antigen (HLA) complex, HLA-DQ2 or HLA-DQ8, account for the strongest genetic association with coeliac disease but there are also additional predisposing variants outside the HLA region contributing hereditary susceptibility to the disease. (Withoff et al., 2016). Family members of coeliac disease patients are at increased risk of developing the disease (Singh et al., 2015), but sporadic cases are also common. Primarily, coeliac disease affects the small intestine, where, for still incompletely understood reasons, both innate and adaptive immune systems become activated. Immune cascade is triggered by harmful gliadin peptides originating from gluten and modified by transglutaminase 2 (TG2) enzyme. (Dieterich et al., 1997; Levescot et al., 2022; Molberg et al., 1998). The activation of immune response eventually results in the development of chronic enteropathy characterized by flattening of the villi (villous atrophy) and elongation of the crypts (crypt hyperplasia) of the small intestinal epithelium. The only effective treatment for the disease is life-long avoidance of gluten-containing food products i.e., gluten-free diet (GFD) that gradually normalizes the mucosal architecture and alleviates the symptoms (Lindfors et al., 2019).

Characteristic of coeliac disease is its protean clinical phenotype. Symptoms are categorized as gastrointestinal or extraintestinal, depending on the site of manifestation, and some patients may be even completely asymptomatic (Durazzo et al., 2022; Reale et al., 2018; Tapsas et al., 2016). Moreover, onset, severity and progression of symptoms as well as response to GFD vary widely between patients. Further, several comorbidities associated with coeliac disease, together with the challenges and social secondary effects of maintaining strict GFD add to the individual burden of coeliac disease (Spijkerman et al., 2016; Uche-Anya & Lebwohl, 2021; White et al., 2016).

This dissertation investigated differences in both clinical phenotype and predisposing genotype at diagnosis as well as well-being and treatment compliance during follow-up between familial and sporadic coeliac disease patients. Furthermore, since the reasons behind the substantial individual variation in the clinical presentation of coeliac disease are unknown, this dissertation focuses on examining whether the dose of major HLA-DQ risk haplotype or the non-HLA genetic variants could establish an explanation for this phenotypic heterogeneity.

2 REVIEW OF THE LITERATURE

2.1 Epidemiology of coeliac disease

Coeliac disease was in the past described as a rare paediatric disorder affecting western Europeans. Through advanced recognition and diagnostics during the past four decades it has become evident that coeliac disease occurs globally and affects people of all ages (Lindfors et al., 2019), incidence rates being highest among people <5 and >50 years of age (West et al., 2014). Even though coeliac disease is still heavily under-diagnosed, its true prevalence can be estimated based on coeliac disease specific seropositivity, that is positivity for TG2 and endomysium autoantibodies, among general population.

The pooled global seroprevalence of coeliac disease is estimated to be 1.4% (Singh et al., 2018), which suggests that the disease is a significant worldwide public health problem. However, coeliac disease seroprevalence varies widely both intercontinentally and internationally, and even from region to region within the same country. It remains obscure whether these differences in coeliac disease prevalence are simply linked to known risk factors, HLA haplotype and gluten intake (Lebwohl et al., 2018). The highest prevalence of coeliac disease (5.6%) has been reported in children in Saharawi population (Catassi et al., 1999) whereas very low prevalences (<0.1%) have been reported, for example, in Japan and Vietnam (Fukunaga et al., 2018; Zanell et al., 2016). Nevertheless, there are many regions where the prevalence of coeliac disease has not yet been thoroughly studied and is thus still unknown. Such areas include, for example, Sub-Saharan Africa and East Asia. In Europe, countries with the highest prevalence of coeliac disease include Sweden, Finland and Italy while the lowest prevalences are found, for instance, in Switzerland, Estonia and Poland (Makharia et al., 2022). Like autoimmune diseases in general, coeliac disease presents with a greater prevalence among women, the female-to-male ratio being almost 2:1 (Megiorni et al., 2008). Additionally, familial aggregation of coeliac disease is common, which is due to the hereditary nature of the disease (genetics described in more detail in sections 2.5.2 and 2.5.5). However, sporadic coeliac disease also occurs.

The global prevalence of coeliac disease seems to be constantly increasing. In the Western world, the incidence of coeliac disease has been reported to have increased tenfold in the past forty years (King et al., 2020). The phenomenon could be explained by improved disease awareness and better availability of serological tests but it has also been reported that, for unknown reasons, the total prevalence of coeliac disease has truly increased, which has been shown in cross-sectional studies evaluating coeliac disease specific seropositivity in Northern American and Finnish populations, for example (Catassi et al., 2010; Lohi et al., 2007; Rubio-Tapia et al., 2009). Interestingly, according to more recent studies in some high prevalence countries, such as Sweden and Finland, the incidence of coeliac disease seems to have stabilized or even decreased (Kivelä et al., 2015; Namatovu et al., 2014; Virta et al., 2017). suggesting that the disease has reached its peak incidence among those populations.

2.2 Diagnostics of coeliac disease

2.2.1 Small-bowel mucosal biopsy

Long, finger-like villi and short crypts are characteristic of normal, healthy small bowel mucosa whereas in untreated coeliac disease this structure gradually starts to become flat. Since the development of methods for obtaining peroral biopsies in 1960s, the cornerstone of coeliac disease diagnosis has been the histologic examination and the determination of small intestinal mucosal morphology (Crosby et al., 1957; Shiner, 1956). In Tampere, Finland, the diagnostic status of gradual structural deterioration of the mucosa is classified in routine clinical practise as partial, subtotal or total villous atrophy (PVA, SVA or TVA respectively). An overt villous atrophy, crypt hyperplasia and more than 30 intraepithelial lymphocytes (IELs) per 100 enterocytes is typical for all of these stages. PVA indicates a minor to moderate amount of villous blunting and SVA that there are truncated villous remnants. In the last stage, TVA, the mucosa is totally flat with no visible villi (Dickson et al., 2006). The clinical classification is always made on the basis of the pathology report and has been used systematically in clinical practice in Tampere from the 1960s (Kröger et al., 2020). Other classification criteria also exist, of which the Marsh-Oberhuber and Corazza-Villanacci criteria are the most used (G. R.

Corazza & Villanacci, 2005; Marsh, 1992; Oberhuber et al., 1999). Specimens with PVA, SVA and TVA correspond respectively to IIIa, IIIb and IIIc (Marsh-Oberhuber) or B1 (PVA and SVA) and B2 (TVA)(Corazza-Villanacci). An additional, more quantitative way of determining the structural alterations in small intestinal biopsy is to measure the ratio of villous heights and crypt depths (Vh:CrD). This method gives precise numerical values, a Vh:CrD of less than 2.0 being generally accepted as indicative of coeliac disease lesion (Kuitunen et al., 1982; Taavela et al., 2013). Furthermore, more rigorous methods for evaluating the histopathologic status of biopsy specimens have recently been developed. These include, for instance, digital quantitation of IEL counts and Vh:CrD (Popp et al., 2021; Taavela et al., 2019) as well as virtual imaging of biopsy specimens utilizing X-ray microtomography (Virta et al., 2020).

2.2.2 Serology

In the 1970s, the first coeliac disease specific autoantibodies targeting reticulin fibres of connective tissue (anti-reticulin antibodies, ARAs) were discovered and became widely used for diagnostic purposes (Seah et al., 1971). ARAs were measured by indirect immunofluorescence using rodent tissues as antigens. Even though they had an excellent specificity (90-100%), their sensitivity was relatively good only in children (Mäki et al., 1984; Seah et al., 1973). Soon after, autoantibodies targeting gliadin in wheat gluten (anti-gliadin antibodies, AGAs) were found and measured using enzyme-linked immunosorbent assay (ELISA). This method was also commonly used until it became evident that AGAs showed substantial variation in terms of sensitivity and specificity and were not actually coeliac disease specific antibodies, but present in serum of patients with other gastrointestinal diseases as well (Hill et al., 1991; Lindberg et al., 1985; Mäki et al., 1991). The first autoantibody that has remained in diagnostic usage until these days is endomysial antibody (EmA), discovered in the 1980s (Chorzelski et al., 1984). It is measured by indirect immunofluorescence assay often using human umbilical cord as an antigen (Ladinser et al., 1994). Although the method is relatively laborious, it has high sensitivity and specificity (Schyum & Rumessen, 2013) which has led to its position as a gold standard for coeliac disease antibody detection (S. Husby et al., 2012). After the recognition of TG2 as the predominant autoantigen of coeliac disease in late 1990s (Dieterich et al., 1997), practical and economical ELISA tests were soon developed to detect antibodies against it. TG2 testing shows better sensitivity (90–100% vs. 83–

100%) but lower specificity (90–100% vs. 95–100%) than EmA testing (Giersiepen et al., 2012) and currently it is used widely in clinical practice (Schyum & Rumessen, 2013). It is noteworthy that both EmA and TG2 antibody tests are most reliable when measuring immuneglobulin A (IgA) class autoantibodies, but in the case of IgA-deficient patients, immunoglobulin G (IgG) class autoantibody determination is useful. There are also autoantibodies generated against TG2-deamidated gliadin peptides (DGPs) derived from gluten, but these are not useful as a screening marker and their diagnostic role is unclear since they afford low positive predictive value among TG2 autoantibody negative individuals (Gould et al., 2019).

2.2.3 Current diagnostic procedures

A clinician should nowadays suspect coeliac disease if a patient has indicative symptoms or belongs to a risk group due to associated disorder(s) (to be introduced later in section 2.3.4). Additionally, coeliac disease should be immediately suspected if a person has a family history of coeliac disease, since the risk of developing the disease in his/her lifetime is then higher, approximately 10–15% (Catassi et al., 2022). The investigations generally begin with laboratory tests to determine the levels of serum autoantibodies, notably TG2 antibodies. Since modern serological tests have excellent accuracy (Fuchs et al., 2019; Giersiepen et al., 2012; Lewis & Scott, 2006; Werkstetter et al., 2017), TG2 antibody values ≥10 times the upper limit of normal detection and positive EmA are nowadays considered sufficient criteria for coeliac disease diagnosis in children in Europe with no need for small bowel mucosal biopsy (Husby et al., 2020). In Finland, the same criterion is also applied to adults (Coeliac Disease. Current Care Guidelines, 2018). If antibody levels are not high enough, patients with positive serology are referred to endoscopy and the diagnosis is confirmed with the determination of injury in the duodenal architecture in one representative mucosal biopsy from the anatomical bulb and four from the distal duodenum (Husby et al., 2020; Ludvigsson et al., 2014).

There are several challenges and pitfalls that physicians need to be aware of when suspecting and further diagnosing coeliac disease. First of all, the awareness of a significantly heterogenous phenotype related to coeliac disease is crucial for detecting both patients with family history of the disease and patients without known coeliac disease cases in their family. In Finland, the key to this has been the concerted effort to raise awareness of coeliac disease among both the general public and general medical practitioners (Burki, 2019). Nevertheless, coeliac disease is still a

highly under-diagnosed disorder, leaving as any as two-thirds of affected individuals unrecognized for long (Laurikka et al., 2018). In other words, diagnostic delay in coeliac disease is unacceptably long (Gray & Papanicolas, 2010; Norström et al., 2011). When suspicion of coeliac disease has arisen, patients are often found rapidly by a serology-based approach. However, the risk in antibody testing in the case of adults is that seronegative patients remain undiagnosed. Possible reasons for ostensible seronegativity are, for instance, self-initiated GFD, short period of gluten consumption in infants and selective IgA deficiency (Husby et al., 2020). The latter can be resolved by testing total IgAs and if an individual turns out to be IgA-deficient, coeliac disease specific autoantibodies can be measured in IgG class (Chow et al., 2012; Collin et al., 1992; Savilahti et al., 1971). However, IgG class autoantibodies are less specific and thus cannot serve as the sole basis for a coeliac disease diagnosis and a confirmatory small bowel mucosal biopsy with histological evaluation is needed (Cataldo et al., 2000).

2.3 Clinical picture of coeliac disease

2.3.1 History and general aspects of clinical picture

The very first account of coeliac disease was written as early as in the second century A.D. by the Greek physician Arataeus of Cappadocia, naming the disease as "koiliakos" after the Greek word "koelia" which stands for abdomen (Adams, 1856). However, it took over 17 centuries before gluten was identified as an environmental factor for coeliac disease, intestinal mucosal lesion was detected as a sign for the disease and, finally, advanced instrumentation enabled biopsies to be obtained. Indeed, coeliac disease was for a long time described simply as chronic enteropathy with gastrointestinal symptoms and signs of malabsorption, as also described in the first publication of consensus definition and diagnostic criteria for coeliac disease in 1970 (Meeuwisse, 1970). When clinicians started to observe that coeliac disease could manifest outside the gastrointestinal track and moreover that some patients might even be asymptomatic, the clinical presentation of coeliac disease was categorized as "classical" (presence of mild to severe gastrointestinal symptoms) or "non-classical" (absence of gastrointestinal symptoms). Along with growing knowledge of the systemic nature and individual symptom-wise variation of coeliac disease, the terminology describing the clinical picture expanded and for years, there was a lack

of consensus on the use of terms related to coeliac disease. Thus, a decade ago, the descriptors of coeliac disease used in the literature were evaluated, the clinical presentations of the disease were collectively identified and the outcome has subsequently been referred as the "Oslo definitions" (Ludvigsson et al., 2013).

Symptoms associated with coeliac disease can be divided into gastrointestinal and extraintestinal according to the site of manifestation and a patient may experience only one or both types of symptoms. Interestingly, the most common clinical presentation of coeliac disease is reported to have changed over time, from the severe malabsorption symptoms of paediatric patients to a milder clinical picture with primarily extraintestinal symptoms affecting people of all ages (Volta et al., 2014). In parallel, asymptomatic patients are more common today than before, presumably due to screening programmes and active case-finding among coeliac disease risk groups (Agardh et al., 2015; McGowan et al., 2009). However, it is controversial whether "asymptomatic coeliac disease" truly exists since very often asymptomatic patients report impaired quality of life prior to diagnosis and after initiation of GFD their previously unrecognized symptoms are alleviated. This means that they only recognize the symptoms retrospectively (Rosén et al., 2014; Ukkola et al., 2011). Thus, such patients should rather be classified as suffering from subclinical coeliac disease, which means that their disease is below the threshold of clinical detection and they experience signs and symptoms that are not sufficient to initiate coeliac disease suspicion in routine clinical practice. Additionally, patients have sometimes adapted to the vague symptoms that have persisted for many years and do not report them or even seek medical help (Mahadev et al., 2016). Therefore, determining the symptoms of coeliac disease is still no easy task. Furthermore, it is noteworthy that there is wide individual variation in many symptom-related parameters of coeliac disease, such as its existence *per se*, onset, duration (both before and after treatment) and severity. Of note, the association between the degree of mucosal damage and the severity of symptoms remains obscure. There are several studies showing that patients without villous atrophy may suffer from symptoms or even complications and may subsequently develop mucosal atrophy (Kurppa et al., 2012; Volta et al., 2016) Additionally, those patients often have increased density of IELs in their small intestinal mucosa already before the appearance of the symptoms (Salmi et al., 2006). By contrast, a recent study by Kröger et al. demonstrated that advanced villous atrophy at diagnosis in childhood is associated with more severe clinical characteristics but not in the long-term perspective after GFD treatment (Kröger et al., 2020). All in all, the factors determining the manner of clinical presentation are poorly understood.

2.3.2 Gastrointestinal manifestations

Gastrointestinal symptoms are the most canonical signs of coeliac disease. Originally, malabsorptive gastrointestinal symptoms such as chronic diarrhoea, steatorrhoea, weight loss and failure to thrive were the typical complaints and still refer to "classical coeliac disease" in the Oslo definitions (Ludvigsson et al., 2013). Malabsorption is often defined as weight loss and/or presence of characteristic laboratory abnormalities, such as anaemia, hypoalbuminemia, low folate or low vitamin B12. The classical phenotype has, however, become less frequent over time, nowadays seen in 13–50% of coeliac disease patients (Dominguez Castro et al., 2017; Spijkerman et al., 2016; Volta et al., 2014). At the same time, many other, usually milder gastrointestinal symptoms have been recognised (Husby et al., 2020; Kivelä et al., 2015; Tapsas et al., 2016). These symptoms include nausea, vomiting, abdominal pain, bloating, heartburn, loose stool, flatulence and constipation. Symptoms are, however, always subjective and expressed verbally to a clinician. Severity of symptoms varies significantly across patients and needs to be discovered and defined. Generally, severity of symptoms depends upon the patient's self-reported experience but could for research purposes be categorized for example as "none", "mild", "moderate" and "severe". The presence and severity of gastrointestinal symptoms can be evaluated using structured and validated questionnaires, for instance the Gastrointestinal Symptom Rating Scale (GSRS) measuring self-reported symptoms with 15 selected questions (Svedlund et al., 1988).

2.3.3 Extraintestinal manifestations

Reflecting its systemic nature, coeliac disease also presents with various extraintestinal signs and symptoms affecting almost any organ. Actually, extraintestinal manifestations account for a substantial proportion of the clinical phenotypes of coeliac disease. Gastrointestinal manifestation of malabsorption is generally paralleled by deficiencies of certain vitamins (A, D, E and K) and micronutrients (vitamin B12, iron, folic acid and calcium) which may be reflected in extraintestinal complaints such as weight loss, failure to thrive, low bone mineral density and anaemia (Visakorpi & Mäki, 1994). Actually, anaemia is one of the most common, and sometimes even the sole extraintestinal presentation of untreated coeliac disease (Martín-Masot et al., 2019) and has been associated with clinically more severe forms of the disease (Berry et al., 2018; Singh et al., 2014). Another well-established extraintestinal manifestation of coeliac disease is *dermatitis herpetiformis,* a

cutaneous condition that presents with a pruritic blistering rash and granular IgA precipitates against transglutaminase 3 enzyme (TG3) in the papillary dermis (Marks et al., 1966; Reunala, 2001). Other extraintestinal manifestations include dental enamel defects (Aine et al., 1990), aphtous ulcerations (Ferguson et al., 1980), migraine (Gabrielli et al., 2003), arthritis (Lubrano et al., 1996), abnormal liver values (Äärelä et al., 2016; Farre et al., 2002) and even infertility (Collin et al., 1996). Psychiatric and neurological ailments reportedly associated with coeliac disease include depression, anxiety and most severely gluten ataxia (Briani et al., 2005).

2.3.4 Co-morbidities

Coeliac disease has been recognized to coexist with a spectrum of other disorders, mostly conditions of autoimmune origin. These include type 1 diabetes mellitus (T1D)(M. Mäki et al., 1984), autoimmune thyroidal diseases (Roy et al., 2016) and Sjögren's syndrome (Iltanen et al., 1999). In fact, the estimated prevalence of any other autoimmune disorders among coeliac disease patients may reach up to 31% (Viljamaa et al., 2009). Further examples of disorders associated with coeliac disease are selective IgA deficiency (McGowan et al., 2008) as well as chromosomal disorders such as Down's and Turner's syndromes (Gale et al., 1997; Marild et al., 2016). The risk for malignancies was previously suggested to be higher and likewise the risk of mortality among coeliac disease patients than among general population. Later on, however, the general risk of cancer as well as all-cause mortality has been reported to be comparable to those in normal population. Nevertheless, coeliac disease patients still have an elevated risk of non-Hodgkin lymphoma (Holmes et al., 1989; Koskinen et al., 2022; Tio et al., 2012).

2.4 Management of coeliac disease

2.4.1 Dietary treatment

Long ago, when the environmental driver of coeliac disease was still unknown, various dietary treatment options were experimented with to alleviate the symptoms of coeliac disease patients. The first person to suspect that wheat-containing products might contain the triggering agent, in the 1930s, was the Dutch clinician Dicke. Later, he became even more convinced of his suspicion since the clinical

condition of his young patients improved when bread was in short supply during the Second World War (Dicke et al., 1953). Hence, Dicke is considered a pioneer of GFD and indeed, in the early 1960s, avoidance of gluten was confirmed to be an effective treatment for coeliac disease (Collins & Isselbacher, 1964; Van Berge-Henegouwen & Mulder, 1993). A strict life-long GFD is the only accepted and curative treatment for coeliac disease and practically means total removal from the diet of wheat, rye, barley as well as products with added gluten. In practice, absolute avoidance of gluten is impossible. In Europe for example, foodstuffs labelled "gluten-free" may contain trace amounts of gluten, but the amount of it must not exceed 20 mg/kg (European Commission, 2014). Since only a gluten amount starting from 30-100 mg per day has been reported to injure the small-bowel mucosal structure (Catassi et al., 2007; Collin et al., 2004), that should be tolerated by the vast majority of coeliac disease patients. Nevertheless, gluten tolerance is individual and a universally valid safety margin for gluten ingestion is very challenging to set (Lähdeaho et al., 2011). Purified oats are widely reported to be safe for the majority of coeliac disease patients and their long-term usage may even be beneficial for patients in terms of improved quality of life (Aaltonen et al., 2017; Holm et al., 2006; Pinto-Sánchez et al., 2017). There is still variation in national guidelines and in some countries consumption of oats on GFD is not recommended.

Even though the responsiveness to dietary treatment also varies individually, for the majority of coeliac disease patients GFD is recuperative, alleviating rapidly both gastrointestinal and extraintestinal symptoms and lowering serum autoantibody titers within weeks to months (Haines et al., 2008; Murray et al., 2004). In the case of small intestinal mucosa, the recovery time is significantly longer, taking usually more than a year. Nevertheless, approximately 90% of patients achieve full histological remission after five years (Hære et al., 2016; Newnham et al., 2016). Patients with *dermatitis herpetiformis* are often prescribed treated with an additional anti-inflammatory dapsone medication for the first couple of years after diagnosis to control the rash (Reunala, 2001). In addition, GFD has been demonstrated to eliminate the elevated risk of malignancies and mortality associated with untreated coeliac disease (West et al., 2004). Nevertheless, there is also a small subset of patients who do not respond to GFD regardless of their good adherence to the diet. Such non-responsiveness is generally due to co-existing diseases, for example irritable bowel syndrome (IBS) or gastroesophageal reflux disease (GERD) (Barratt et al., 2011; Volta et al., 2014). The reason for non-responsiveness may be, very seldom though, small-bowel lymphoma or refractory coeliac disease (RCD). RCD is a severe and unresponsive form of the disease among adult patients with a high risk

of complications (Dewar et al., 2012; Ilus et al., 2014). In the case of RCD, immunomodulators, biologic agents and even chemotherapy for the most severe cases have been tried to manage the disease (Hujoel & Murray, 2020).

2.4.2 Alternative treatment options

GFD is considered burdensome for several reasons. For example, strict adherence to the diet is laborious, gluten-free products are costly and the diet is socially restrictive. Moreover, patients experiencing only mild symptoms or who are asymptomatic may regret being diagnosed with coeliac disease. This may lead to poor compliance with the diet and further increase the risk of complications (White et al., 2016). For all these reasons, coeliac disease patients are interested in alternative treatment options to supplement or replace GFD (Tennyson et al., 2013; Ukkola et al., 2012). Novel treatment options have been sought for years, targeting different points and key players in coeliac disease pathogenesis (Kivelä et al., 2021). One of the candidates is CCX282-B, an orally-administered antagonist of C-C motif chemokine receptor 9 (CCR9) whose efficacy and safety have been investigated in a randomized double-blind placebo-controlled phase 2 study (ClinicalTrials.gov identifier: NCT00540657). It was hypothesized that CCX282-B could mitigate the harmful effects of gluten ingestion in patients with coeliac disease by regulating migration and activation of inflammatory cells in the small intestine. However, the results of this clinical trial remain unpublished. In addition to CCX282-B, other candidate drugs have already undergone phase 2–3 clinical studies with published results. Such therapies include gluten-degrading oral enzymes, cytokine therapies and induction of tolerance to gluten ingestion (Verdu & Schuppan, 2021). Among the most promising of these is the inhibition of TG2 with ZED1227, a selective oral TG2 inhibitor, which in the preliminary trial attenuated gluten-induced duodenal mucosal damage (Schuppan et al., 2021).

2.4.3 Follow-up

Follow-up is important and recommended part of coeliac disease treatment since it provides support for patients adapting to the demanding life-long diet and subsequently promotes alleviation of symptoms and recuperation of small-bowel mucosa. In addition, it enables prevention and early detection of possible complications and co-morbidities (Haines et al., 2008; S. Husby et al., 2012; Valitutti

et al., 2017). Nevertheless, there is no consensus on specific follow-up practices. Some studies propose annual follow-up appointments with a physician and/or dietitian, which has also been supported by patients (Bebb et al., 2006; Haines et al., 2008). In terms of follow-up biopsy, according to some studies, it could be recommended for adult patients after one year on GFD (Rubio-Tapia et al., 2010; Sharkey et al., 2013) but this recommendation has been challenged, especially in the case of asymptomatic patients with no other characteristics suggesting an increased risk of complications (Ludvigsson et al., 2014). In Finland, a clinical and serological check-up for adult patients is recommended one year after diagnosis at the latest. If response to diet is good, follow-up may take place less frequently, every 2–3 years, for example. In the case of children and patients with *dermatitis herpetiformis* the need for follow-up frequency is evaluated case-by-case. A routine follow-up biopsy is not recommended unless the response to diet is poor or the patient was seronegative or presented with severe clinical findings at diagnosis. In case of paediatric patients, a follow-up biopsy is taken very seldom (Coeliac Disease. Current Care Guidelines, 2018). It is noteworthy to mention here that RCD is extremely rare in children. In fact, more criticism about the actual need and time frame of follow-up biopsy has been published supporting a more personalized course of action in coeliac disease follow-up in general (Pekki et al., 2015; Pekki et al., 2017; Rubio-Tapia et al., 2013).

2.5 Pathogenesis of coeliac disease

2.5.1 Gluten and other environmental factors

The most important identified environmental factor in coeliac disease is gluten, or, to be precise, gluten-derived peptides resistant to proteolysis of digestive enzymes that trigger disease pathogenesis (Shan et al., 2002). Gluten forms the main protein content of wheat, rye and barley grains. In wheat, gluten can be further divided into prolamins gliadins and glutenins, the former being the major source of immunogenic peptides for coeliac disease patients. The corresponding harmful prolamins are called secalins in rye and hordeins in barley. (Lindfors et al., 2019; Wieser, 2007).

Additional environmental factors besides gluten contributing to coeliac disease pathogenesis have been proposed and studied, but their association with disease development remains generally somewhat inconclusive. Nevertheless, there is, for example, some evidence on the role of intestinal microbiota imbalance and viral

infections in the onset of coeliac disease (Kemppainen et al., 2017; Verdu & Caminero, 2017; Wacklin et al., 2014), but their clear causality remains obscure. Interestingly, however, a high gluten intake has been suggested to increase the risk of coeliac disease among at-risk children. Moreover, it has been even demonstrated to have a cumulative effect in the development of coeliac disease together with enterovirus infections during the first two years of life (Andrén Aronsson et al., 2019; Lindfors et al., 2020).

2.5.2 Role of HLA region

Alleles for HLA-DQ2.5 and HLA-DQ8, the strongest and best-characterized allotypes in coeliac disease, are located in HLA region. The HLA region is part of the Major Histocompatibility Complex (MHC), highly polymorphic and the most gene-dense region of the human genome occupying a 4 Mb segment on the short arm of chromosome 6 (6p21, precisely). MHC is divided into three regions: classes I, II and III, consisting of altogether over 200 protein-coding genes of which approximately 40% are immune-related (Douillard et al., 2021). The class II region includes *HLA-DQ* locus that is involved in coeliac disease. The locus contains genes *HLA-DQA1* and *HLA-DQB1* encoding respectively α and β subunits, which form a heterodimeric HLA-DQ molecules expressed on the surface of antigen presenting cells (APCs) (Figure 1). These heterodimers have an essential role in the initiation of adaptive immune response by binding foreign peptides to be presented to CD4+ T cells (Figure 2). An abundance of genetic polymorphisms found in HLA-DQ region affects the peptide binding and presentation repertoire of the expressed HLA-DQ molecules by altering the amino acids in their peptide-binding groove (van Heel et al., 2005).

In coeliac disease, the most permissive combination of alleles consists of *HLA-DQA1*05* and *HLA-DQB1*02* that encode heterodimer HLA-DQ2.5 which is present in more than 90% of coeliac disease patients (Sollid et al., 1989; Sollid, 2017). However, the presence of this HLA-haplotype within the patient population may differ in different populations. HLA-DQ8 heterodimer on the other hand is encoded by the alleles *DQA1*03* and *HLA-DQB1*03:02* and is present in approximately 20% of patients. Coeliac disease patients negative for both HLA-DQ2.5 and HLA-DQ8 (approx. 5% of all patients) encode mainly HLA-DQ2.2, another variant of DQ2, where α chain is encoded by allele *DQA1*02* (instead of *DQA1*05* as in HLA-DQ2.5) (Almeida et al., 2016; Liu et al., 2014; Mubarak et al., 2013)(Figure 1).

The HLA-DQ2.5 molecule is usually encoded on the same haplotype (*cis*), which means that alleles encoding α and β subunits are located in the same chromosome and thus inherited from one parent. Conversely, and more rarely, subunits are encoded by alleles residing on different chromosomes (*trans*) and thus inherited one from each parent (Figure 1). In addition to substantial polymorphism, there is a strong linkage disequilibrium (LD) in the HLA locus. It means that variants some distance apart form haplotypes occurring together significantly more often than would be randomly expected. The first such haplotype reported was A1-B8-DR3 extended HLA haplotype (also known as 8.1 ancestral haplotype or AH8.1) that is highly conserved and associated with multiple autoimmune diseases such as type 1 diabetes and autoimmune thyroid disease, in addition to coeliac disease (Price et al., 1999; Sollid, 2017; van Heel et al., 2005).

A patient is at risk of developing coeliac disease when possessing any of the HLA-DQ2.5, HLA-DQ8 or HLA-DQ2.2 heterodimers (Figure 1). There is, however, a very small group of patients negative for the aforementioned heterodimers. These patients almost exclusively encode a very low-risk heterodimer HLA-DQ7.5 which is commonly denoted as HLA-DQ7. The risk of coeliac disease is greater in the case of a person encoding two predisposing heterodimers and particularly high (fivefold risk compared to HLA-DQ2.5 heterozygotes) in the case of HLA-DQ2.5 homozygote (Margaritte-Jeannin et al., 2004; Megiorni et al., 2009; Ploski et al., 1993). Nonetheless, studies on the so-called gene dose effect of HLA-DQ2.5 on disease severity and clinical outcomes have produced inconsistent results, the study settings being rarely comparable (Bajor et al., 2019).

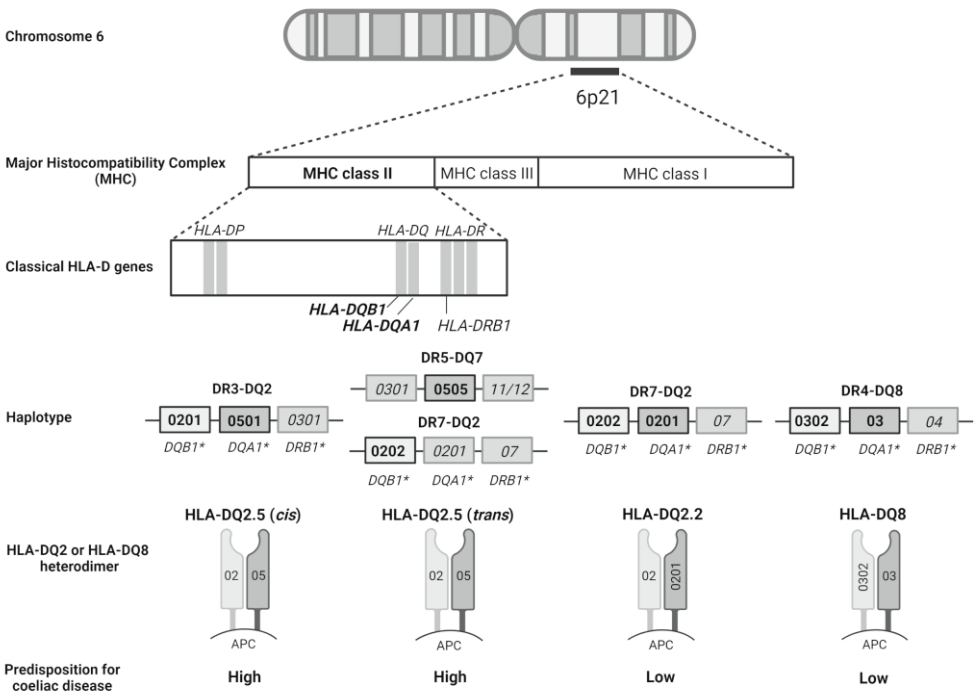

Figure 1. Genetic background of coeliac disease associated HLA-DQ molecules. Alleles from HLA-DQ and HLA-DR genes residing in Major Histocompatibility Complex (MHC) in chromosome 6 determine haplotypes coding for corresponding HLA-DQ2.5, HLA-DQ2.2 and HLA-DQ8 molecules involved in coeliac disease pathogenesis. The level of predisposition for each formed HLA-DQ heterodimer is also shown. Figure created with Biorender.com.

2.5.3 Adaptive immune mechanisms

Immunopathogenesis of coeliac disease involves both innate and adaptive immune responses. Adaptive immune activities of coeliac disease pathogenesis can be regarded to initiate when immunogenic gluten peptides, originating from incomplete digestion in the gut lumen, enter the small intestinal lamina propria by crossing the epithelial layer via paracellular or transcellular pathways (Heyman et al., 2012; Lindfors et al., 2019)(Figure 2). Gliadin peptides are deamidated by ubiquitously expressed enzyme TG2, and this deamidation has been thought to occur in the lamina propria. However, a recent study by Iversen et al. proposes that the site of action for deamidation is instead the gut lumen (Iversen et al., 2020). Either way, glutamine residues at key positions in gliadin peptides are replaced by glutamic acid making gliadin peptide negatively charged (Fleckenstein et al., 2002). This modification further increases the affinity of gliadin peptide to the antigen-binding groove of HLA-DQ2 or -DQ8 on the surface of APCs (Kim et al., 2004; Molberg et al., 1998), which may, for example be dendritic cells. APCs acquire a proinflammatory phenotype and migrate to the mesenteric lymph nodes (inductive site), where they contribute to T cell activation by presenting gliadin peptides to CD4+ T cells. Naïve CD4+ T cells have travelled to the mesenteric lymph node from the thymus (differentiation site) and after encountering APCs through their T cell receptors (TCRs), differentiate into gluten-specific CD4+ T cells. These inflammatory effector cells migrate to the lamina propria (effector site) creating an inflammatory milieu by secreting inflammatory cytokines such as interferon gamma (IFN-γ) and interleukin 21 (IL-21)(Abadie et al., 2011; Catassi et al., 2022). In addition, gluten-specific CD4+ T cells provide help to gluten- and TG2-specific B cells by activating them. These B cell populations have arisen by recognizing gliadin peptides and TG2-gliadin complexes as antigens via their B cell receptors (BCRs). They have also been proposed to act as APCs in coeliac disease pathogenesis. Activated B cells further differentiate into plasma cells that secrete autoantibodies against gliadin and TG2 (Abadie et al., 2011; Lindfors et al., 2019; Verdu & Schuppan, 2021). TG2-abs especially have been suggested to have an active role in coeliac disease pathogenesis, for example, by maintaining inflammation and provoking mucosal damage (Kalliokoski et al., 2017; Korponay-Szabó et al., 2004; Rauhavirta et al., 2019), even though contrary results have also been reported (Lindstad et al., 2022).

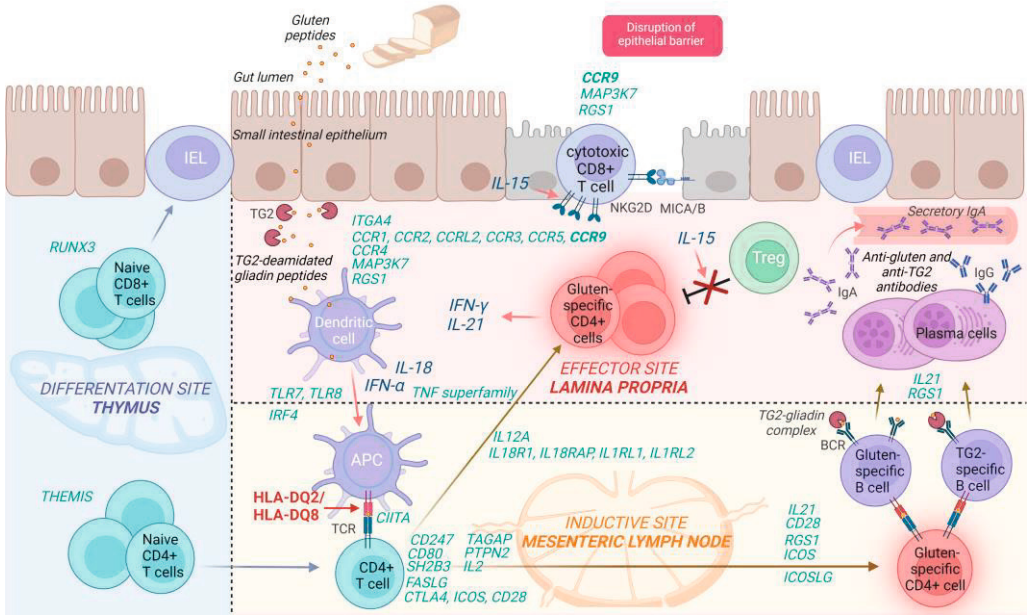

Figure 2. Model of immunopathogenesis of coeliac disease and associated non-HLA candidate genes. The figure is divided into the main three anatomical regions harbouring the key immunological pathways related to both innate and adaptive immunity. The thymus acts as a T cell differentiation site for both IELs of innate immunity and CD4+ T cells of adaptive immunity. The mesenteric lymph node serves as an inductive site where polarization of CD4+ T cells takes place and the lamina propria acts as an effector site for both innate and adaptive immune responses. Non-HLA candidate genes involved in immunological pathways are denoted in green and their full names are listed in Appendix 1. BCR, B cell receptor; IEL, intraepithelial lymphocyte; IFN-α, interferon alpha; IFN-γ, interferon gamma; IgA, immunoglobulin A; IgG, immunoglobulin G; IL-15, interleukin 15; IL-18, interleukin 18; IL-21, interleukin 21; MICA/B, MHC class I polypeptide–related sequence A/B; NKG2D, natural killer group 2D receptor; TCR, T cell receptor; TG2, tissue transglutaminase 2. Figure adapted from Abadie et al., 2011 and created with BioRender.com.

2.5.4 Innate immune mechanisms

Innate immune responses are suggested to be triggered by certain gluten peptides, enteric infections or α-amylase-trypsin inhibitors (ATIs) originating from wheat (Lindfors et al., 2019). A characteristic feature of mucosal inflammation in coeliac disease is an increasing count of IELs in the small intestinal epithelium, usually exceeding 25/100 enterocytes, whereas in the case of healthy mucosa the count is approximately 10/100 enterocytes (Corazza et al., 2007; Olivares-Villagómez & Van Kaer, 2018; Walker et al., 2010). IELs are a heterogenous T cell population that can be divided into two major groups, αβTCR+ and γδTCR+ cells. In homeostasis, IELs are responsible for maintaining epithelial integrity by eliminating infected cells and promoting epithelial repair activities (Sheridan & Lefrançois, 2010). In coeliac disease, by contrast, these cells have become dysregulated. IELs take part in innate immune responses as cytotoxic CD8+ T cells which have been stimulated by CD4+ T cells, key players in adaptive immunity, through IFN-γ and IL-21 signalling (Jabri & Sollid, 2017)(Figure 2). Cytotoxic IELs destroy the enterocytes of the intestinal barrier thus promoting epithelial permeability and the development of villous atrophy. These cytotoxic and apoptotic responses are mediated through several distinct receptor interactions such as that between killer activation receptor NKG2D expressed by IELs and stress-induced self-antigen MICA expressed by epithelial cells (Abadie et al., 2011; Du Pré & Sollid, 2015).

Another important mediator of innate immunity in coeliac disease immunopathogenesis is interleukin 15 (IL-15). In coeliac disease, IL-15 has been observed to be over-expressed both in the epithelial layer and in the lamina propria. It contributes to licensing IELs to kill epithelial cells by upregulating NKG2D expression and inhibits the function of regulatory T cells (Tregs) which further provokes proinflammatory actions of CD4+ effector T cells, mucosal injury and loss of oral tolerance. (Jabri & Abadie, 2015; Mention et al., 2003; Zorro et al., 2020). Additionally, the secretion of interferon alpha (IFN-α) and interleukin 18 (IL-18) by dendritic cells is also regarded as a hallmark of innate immunity in coeliac disease pathogenesis (Di Sabatino et al., 2007; Salvati et al., 2002).

2.5.5 Non-HLA regions

The genome-wide association studies (GWAS) conducted on samples from different populations geographically remote from each other started the new era of genetic research on complex diseases. Coeliac disease was among the first of these (Van Heel et al., 2007). The first rounds of GWAS and follow-up studies identified 43 new loci associated with coeliac disease development, also confirming the role of the HLA region (Dubois et al., 2010; Garner et al., 2009; Hunt et al., 2008; Trynka et al., 2011; Van Heel et al., 2007). To date, the number of loci outside the HLA region reportedly associated with coeliac disease is verging on 50 and the number of candidate genes residing in these regions is nearly twice that number (Table 1) (Appendix 1). The associated non-HLA regions comprise LD blocks generally consisting of multiple genes. Thus, identified coeliac disease associated SNPs tag risk haplotypes, rather than being causal variants associated with the disease themselves (Abadie et al., 2011). Nevertheless, more than half of the coeliac disease associated genetic variants affect the expression levels of at least one gene in their proximity. Such SNPs are termed *cis* expression quantitative trait loci (*cis* eQTL) SNPs (Westra & Franke, 2014). Even though it is challenging to precisely pinpoint which genes are most likely causal in terms of coeliac disease pathogenesis, the substantial number of reported SNPs reside in proximity to genes involved in immune functions (Gnodi et al., 2022). Moreover, the great majority of loci are also associated with other autoimmune diseases as well, as shown in Table 1.

Some examples of candidate genes harbouring coeliac disease associated SNPs and involved in immunity and/or immunopathological pathways related to coeliac disease are marked in green in Figure 2. *RUNX3* and *THEMIS* are involved in the thymic differentiation of CD8+ T cells and CD4+T cells respectively. *TLR7*, *TLR8* and *IRF4* are involved in viral recognition facilitating dendritic cells in the lamina propria in acquiring proinflammatory phenotype. Class II major histocompatibility complex transactivator (*CIITA*) is an example of a non-HLA gene that takes part in the process where APCs present gliadin peptides to CD4+ T cells. There are several genes (e.g., *CD247, CD80, SH2B3, FASLG, CTLA4, TAGAP, PTPN2, ICOS* and *IL18RAP*) that promote T cell activation and differentiation into gluten-specific CD4+ T cells. Some genes, such as *IL21*, which has potent regulatory effects on numerous cell types, are implicated in multiple distinct pathways at several levels, inducing cell division and proliferation. Some genes are involved in cell migration (e.g. *CCR9*, *MAP3K7* and *RGS1*) and others in the regulation of tumour necrosis

factor (TNF)-dependent pathways (TNF superfamily including *TNFAIP3*, *TNFSF4* and *TNFRSF14*) (Abadie et al., 2011; Withoff et al., 2016).

While the HLA locus contributes over 40% of coeliac disease heritability, individual non-HLA loci associated with the disease have a fairly slight impact on disease heritability. Nevertheless, their combined effect accounts for an additional 20% of genetic susceptibility to coeliac disease and may be utilized in determining genetic risk score (GRS) (Abraham et al., 2014; van der Graaf et al., 2021). However, 40% of coeliac disease heritability remains unknown. While there is a remarkable concordance between general findings from GWAS and current immunological models, the impact of SNPs identified on the immunological and clinical phenotypes of coeliac disease remains to be investigated.

Table 1. Non-HLA regions associated with coeliac disease and possibly other autoimmune and/or inflammatory diseases

Chromosomal region	Candidate gene(s) in the region[a]	Association with other autoimmune and/or inflammatory disease[b]	Reference
1p13.2	*RSBN1, PTPN22, PHTF1, AP4B1-AS1, DCLRE1B, MAGI3, AP4B1*	RA, T1D, SLE, JIA, CrD, VIT, AAV, GD, AD, MG	Tizaoui et al. 2019
1p31.3	*NFIA*	AS, PBC, PS, UC *	Coleman et al. 2016
1p36.11	*RUNX3*	AS, PS	Apel et al. 2013
1p36.23	*PARK7, TNFRSF9*	PS	Tsoi et al. 2012
1p36.32	*TNFRSF14, MMEL1, PRXL2B, TTC34*	RA	Coenen et al. 2009
1q24.2	*CD247*	RA, SSc	Dexiu et al. 2022
1q24.3	*FASLG, TNFSF18, TNFSF4*	SJO, SSc, SLE	Teruel & Alarcón-Riquelme 2016
1q25.3	*NCF2, SMG7*	SLE	Graham et al. 2011
1q31.2	*RGS1*	T1D, MS	Esposito et al. 2010, Smyth et al. 2008
1q32.1	*INAVA*	IBD	Rivas et al. 2011
1q41	*DUSP10*		Östensson et al. 2013
2p14	*PLEK*		Sharma et al. 2016
2p15–16.1	*PUS10*	CrD	Festen et al. 2011
2p16.1	*REL*	RA	Gregersen et al. 2009
2q12.1	*IL18R1, IL18RAP*	T1D, CrD, UC	Smyth et al. 2008, Zhernakova et al. 2008
2q31.3	*ITGA4, UBE2E3*	AS *	Garner et al. 2009
2q32.2–32.3	*STAT4*	RA, PBC, SJO, JIA, SLE	Lessard et al. 2013, Mells et al. 2011, Prahalad et al. 2009, Shancui-Zheng et al. 2022, Zhernakova et al. 2011
2q33.2	*CD28, CTLA4, ICOS*	RA, SLE, T1D, GD, PBC, AD	Barreto et al. 2004, Barrett et al. 2009, Blomhoff et al. 2004, Pawlak-Adamska et al. 2017, Qiu et al. 2017, Raychaudhuri et al. 2009
3p14.1	*FRMD4B*		Garner et al. 2014
3p21.31	*CCR1, CCR2, CCRL2, CCR3, CCR5, CCR9, LTF*	T1D	Smyth et al. 2008
3p22.3	*CCR4, GLB1*	MS *	Dubois et al. 2010
3q13.33	*CD80, POGLUT1, ARHGAP31*	PBC	Mells et al. 2011
3q25.33	*IL12A, SCHIP1*	MS, PBC, SJO, SLE, SSc	Esposito et al. 2010, Lessard et al. 2012, Lessard et al. 2013, Qiu et al. 2017
3q28	*LPP*	RA, VIT	Coenen et al. 2009, Jin et al. 2010
4q27	*BLTP1, ADAD1, IL2, IL21*	RA, T1D, UC	Barrett et al. 2009, Festen et al. 2009, Zhernakova et al. 2007
6p25.3	*IRF4*	PS, RA *	Trynka et al. 2011
6q15	*BACH2, MAP3K7*	T1D, AD, RA, VIT	Barrett et al. 2009, Eriksson et al. 2021, Jin et al. 2012, McAllister et al. 2013

Chromosomal region	Candidate gene(s) in the region[a]	Association with other autoimmune and/or inflammatory disease[b]	Reference
6q22.33	*PTPRK, THEMIS*	CD, IBD, MS, T1D *	Dubois et al. 2010
6q23.3	*TNFAIP3, OLIG3*	RA, SLE, PS, JIA	Graham et al. 2008, Plenge et al. 2007, Prahalad et al. 2009, Tsoi et al. 2012
6q25.3	*TAGAP*	T1D, RA, MS, CrD, PS	Eyre et al. 2010, Festen et al. 2011, Patsopoulos et al. 2011, Tsoi et al. 2012
7p14.1	*ELMO1*	JIA	Nikopensius et al. 2021
8q24.21	*PVT1*	CrD IBD MS RA *	Dubois et al. 2010
10p15.1	*PFKFB3, PRKCQ*	RA	Coenen et al. 2009
10q22.3	*ZMIZ1*	CrD, PS, MS	Ellinghaus et al. 2012, Patsopoulos et al. 2011
11q23.1	*POU2AF1, COLCA1,*	PBC	Nakamura et al. 2012
11q23.3	*TREH, DDX6*	RA, SJO, SLE, VIT	Lessard et al. 2013, Zhang et al. 2014, Zhao et al. 2017, Zhernakova et al. 2011
11q24.3	*ETS1*	MS, SLE, AS, RA	Chen et al. 2015, Lill et al. 2015, Shan et al. 2014, Yang et al. 2010
12q24.12	*SH2B3, ATXN2*	RA, T1D, PBC, PSC, JIA, VIT	Hinks et al. 2013, Jin et al. 2012, Liu et al. 2012, Liu et al. 2013, Smyth et al. 2008, Zhernakova et al. 2011
14q24.1	*ZFP36L1*	MS, RA	Beecham et al. 2013, Kwon et al. 2020
15q24.1–24.2	*CLK3, CSK*	SLE *	Trynka et al. 2011
16p13.13	*CIITA, SOCS1, CLEC16A, PRM1, PRM2*	T1D, MS, RA, SLE	Booth et al. 2008, Bronson et al. 2011, Lima et al. 2021
18p11.21	*PTPN2*	RA, T1D, CrD, JIA	Barrett et al. 2009, Cobb et al. 2013, Festen et al. 2011, Thompson et al. 2010
20q13.12	*ZNF335*	CrD, IBD, MS, RA *	Coleman et al. 2016
21q22.12	*RUNX1*	RA	Gutierrez-Achury et al. 2016
21q22.3	*ICOSLG, UBASH3A*	AA, CrD, T1D, UC	Barrett et al. 2008, Barrett et al. 2009, Conteduca et al. 2014, Wang et al. 2010
22q11.21	*UBE2L3, YDJC*	CrD, PS, RA, SLE	Ellinghaus et al. 2012, Orozco et al. 2011
Xp22.2	*TLR7, TLR8*	SLE, T1D, VIT	Cooper et al. 2009, Shen et al. 2010, Traks et al. 2015
Xq28	*HCFC1, TMEM187, IRAK1*	RA, SLE	Han et al. 2013, Kaufman et al. 2013

[a]Please see the entire list of genes in Appendix 1.

[b]Associations are found through GWAS or candidate gene approaches unless marked with * when the region has shared genetics with coeliac disease according to www.immunobase.org.

AA, alopecia areata; AAV, anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis; AD, Addison's disease; AS, ankylosing spondylitis; CrD, Crohn's disease; GD, Grave's disease; IBD, inflammatory bowel disease; JIA, juvenile idiopathic arthritis; MG, myasthenia gravis; MS, multiple sclerosis; PBC, primary biliary cholangitis; PS, psoriasis; PSC, primary sclerosing cholangitis; RA, rheumatoid arthritis; SJO, Sjögren's syndrome; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; T1D, type 1 diabetes; UC, ulcerative colitis; VIT, vitiligo.

2.5.6 Candidate genes *CCR9* and *CCL25*

The pathogenesis of coeliac disease comprises activation and recruitment of inflammatory cells. These cells are attracted to the small intestine from secondary lymphoid tissue through chemotaxis. This chemotactic process involves gut-homing chemokine CC motif chemokine ligand 25 (CCL25, formerly termed thymus-expressed chemokine, TECK) and its specific receptor CC motif chemokine receptor 9 (CCR9) (Papadakis et al., 2000; Zaballos et al., 1999). CCL25 is predominantly expressed by thymic and small intestinal epithelial cells and serves as a chemoattractant for CD4+ and CD8+ T lymphocytes that express CCR9 on their cell surface (Kunkel et al., 2000; Wurbel et al., 2000)(Figure 2). CCR9 is a G protein-coupled receptor belonging to the β chemokine receptor family. Gene *CCR9* is located in 3p21.31, the locus that was found to be associated with coeliac disease in a follow-up study performed after the first GWAS and also subsequently replicated (Hunt et al., 2008; Sharma et al., 2016) . Gene *CCL25* maps to locus 19p13.2, which is relatively proximal to locus 19p13.11, with *MYO9B* as a candidate gene that has been reported to show a strong linkage to coeliac disease even though the association observed has been population-specific (Koskinen et al., 2008; Monsuur et al., 2005; Van Belzen et al., 2003). Nevertheless, the coverage of SNPs at *CCR9* and *CCL25* loci in earlier studies focusing coeliac disease associations has been limited.

In addition to playing a central role in homing immune cells to the small intestine, the main organ affected by coeliac disease, there is functional evidence of the involvement of the CCR9-CCL25 axis in coeliac disease by several other means. Firstly, CCR9-positive dendritic cells are key players in maintaining gut homeostasis and tolerance by regulating the phenotype and function of both innate and adaptive immune cells (Pathak & Lal, 2020). Secondly, the number of circulating CCR9-positive type 2 conventional dendritic cells has been reported to be increased in untreated coeliac disease while the quantity of intestinal CCR9-positive T cells is diminished. (Escudero-Hernández et al., 2020; Olaussen et al., 2007). Interestingly, an oral CCR9 inhibitor has served as a drug candidate in a clinical trial aiming to develop an alternative treatment for coeliac disease (ClinicalTrials.gov identifier: NCT00540657).

3 AIMS OF THE STUDY

The aim of this dissertation was to investigate the phenotypic and genotypic differences between familial and sporadic coeliac disease and to examine the contribution of various genetic factors predisposing to coeliac disease to the clinical phenotype of the disease.

The specific aims in Studies I–III were:

I. To compare familial and sporadic coeliac disease in terms of the patients' HLA and non-HLA genetics, also clinical, histological and serological presentation at diagnosis and physical and psychological well-being and treatment compliance on GFD.

II. To investigate whether HLA-DQ2.5 allele dose is associated with different clinical parameters at the time of coeliac disease diagnosis and with patients' response to GFD.

III. To identify SNPs in *CCR9* and *CCL25* genomic regions having predicted functional effect and/or possessing exonic location and to test their association with coeliac disease and its distinct clinical phenotypes.

4 MATERIALS AND METHODS

4.1 Ethical considerations

All study protocols in this dissertation were in accordance with the 1964 Helsinki Declaration and its later amendments. The study design, patient recruitment and collection of patient record data were approved by the Regional Ethics Committee of Tampere University Hospital with ethical approval numbers R05183, E98012 and R03060. Additionally, for the healthy controls in Study III, approvals (558/E3/2001 for FINRISK and 407/E3/2000 for Health 2000) were granted by the Ethics Committee of the National Institute for Health and Welfare and the Ethics Committee in Epidemiology and Public Health in the Hospital District of Helsinki and Uusimaa. Written informed consent was obtained from all individual participants (or in the case of children from their legal guardians) included in the studies. This dissertation does not include any studies with animals.

4.2 Study participants

4.2.1 Patients (Studies I–III)

Data from over 1,134 Finnish biopsy-proven coeliac disease patients, both children and adults, was utilized in all three independent studies of this dissertation. The distribution of study participants and their number in each of the independent studies is shown in a flowchart in Figure 3.

A large and well-defined dataset collected as part of the coeliac disease family study R05183 in 2006-2010 made up a major part of the patient data utilized in the studies of this dissertation. Some of these patients were already HLA-DQ-genotyped when participating in earlier studies and thus the respective study codes for these are E98012 or R03060. In addition to the coeliac disease index cases, R05183 included over 3,000 family members. Participants contributing to this dataset were recruited by a nation-wide search with the help of the national and local coeliac disease

societies and by means of announcements in the media. In Study I, previously undiagnosed relatives found to have positive coeliac disease specific serology in the screening were referred to gastrointestinal endoscopy. If, based on biopsy, these individuals were confirmed as new cases, they were also considered affected family members. Moreover, relatives who refused the endoscopy but who had positive serum EmA and TG2 autoantibody values were also regarded as affected family members for the purposes of this study. This assumption was based on the evidence that seropositivity for EmA and TG2 affords excellent specificity for coeliac disease (Coeliac Disease. Current Care Guidelines, 2018; Kurppa et al., 2012). Patients having relatives with inconclusive serology and no biopsy were excluded from further analyses, likewise patients with obscure family history, non-coeliac gluten sensitivity or only self-reported coeliac disease. Finally, Study I consisted of 1,064 coeliac disease patients, who were divided into "familial cases" ($n = 761$) with one or more affected relatives and "sporadic cases" ($n = 303$) with no relatives diagnosed with coeliac disease (Fig. 3).

In Study II, only one patient from each family was randomly included in order to avoid false positive findings due to trait correlation between genetically related individuals. The final study cohort included 605 coeliac disease patients (567 from dataset R05183, 34 from R03060 and 4 from E98012) who were further divided into four different age groups: 0–6 years, 7–20 years, 21–65 years and >65 years, resulting in 124 children (<21 years of age) and 476 adults.

In Study III, in order to avoid bias (inflation of type 1 error), only one randomly selected coeliac disease patient per family with full genotype available was included in the final study cohort, resulting in 625 cases. Additional DNA samples of 144 HLA-DQ2-positive biopsy-proven coeliac disease patients from dataset R05183 were utilized in the sequencing project of *CCR9* and *CCL25* exonic regions.

In all three studies of the dissertation, missing data on individual clinical parameter(s) and/or incomplete study questionnaires were not considered as exclusion criteria.

4.2.2 Healthy controls (Study III)

A total of 1,817 individuals with information on gender (1,032 males, 785 females) and HLA-DQ-genotype (low HLA-DQ risk, *n* = 939; intermediate HLA-DQ risk, *n* = 855; high HLA-DQ risk, *n* = 23) from the population representative cohorts of FINRISK and Health 2000 (Borodulin et al., 2018) were included in the study. For the sequencing project of *CCR9* and *CCL25* exonic regions, DNA from 144 non-celiac blood donor controls was included.

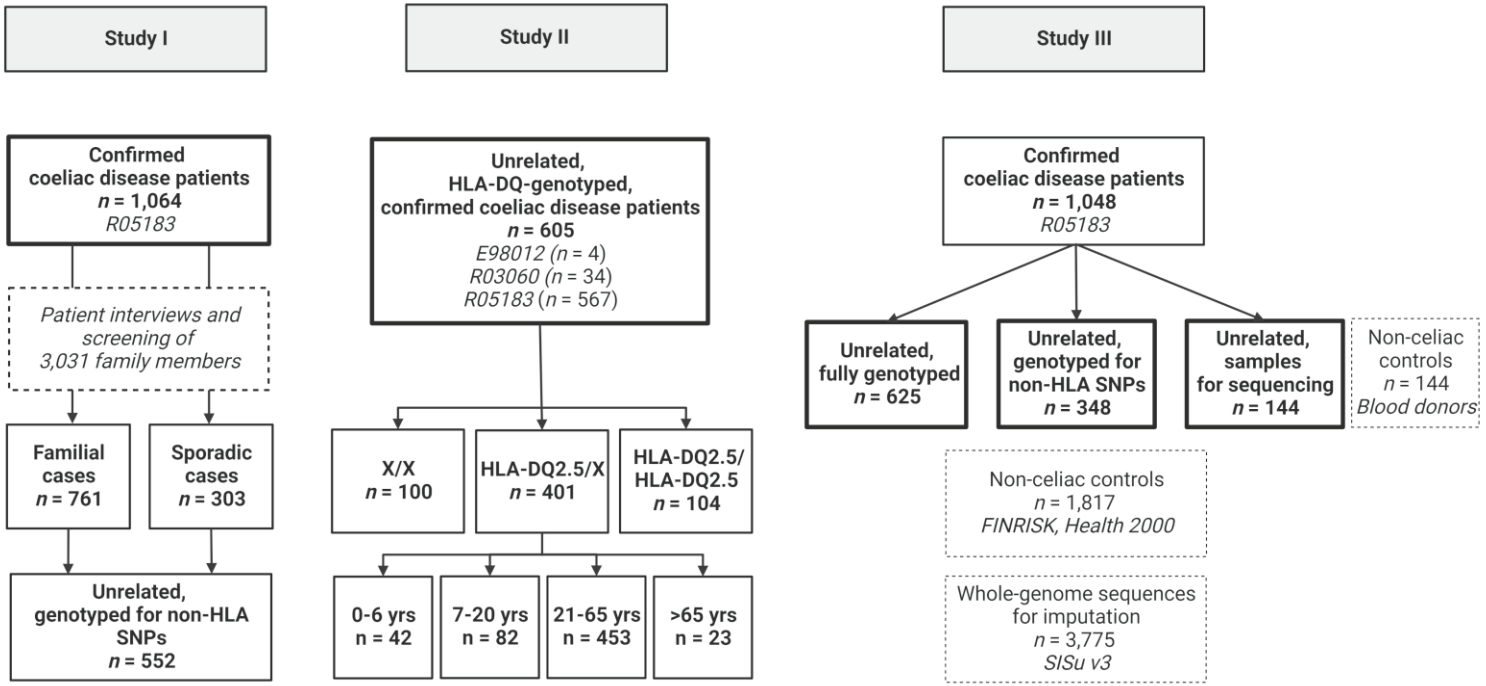

Figure 3. Distribution of study participants and samples in Studies I–III.

4.3 Methods

4.3.1 Clinical data

4.3.1.1 Data collection (Studies I–III)

To obtain information on the study patients' clinical status at the time of data collection, duration and strictness of GFD as well as the use of purified oats in the diet, structured interviews were conducted by a physician or a study nurse experienced in coeliac disease. Demographics, all relevant medical data at diagnosis (including clinical and histological data, coeliac disease serology and concomitant chronic medical conditions), presence of symptoms during childhood and family history of coeliac disease were also recorded and confirmed from patient records as available. In the case of paediatric patients, the guardian was interviewed. For further analyses of coeliac disease-associated serology and genetics, study participants were also subjected to blood sampling.

4.3.1.2 Questionnaires (Study I)

In Study I, adult participants also completed three structured and validated questionnaires eliciting prevailing symptoms and quality of life. These questionnaires were the Gastrointestinal Symptom Rating Scale (GSRS)(Appendix 2), the Psychological General Well-Being Questionnaire (PGWB)(Appendix 3), and Short Form 36 (SF-36, i.e. RAND SF 36-Item Health Survey 1.0)(Appendix 4), which are described in the following paragraphs.

GSRS measures self-reported symptoms with 15 selected questions each scored on a Likert scale from 1 to 7 points, higher scores denoting more severe gastrointestinal symptoms. Total score indicates an average of the 15 individual scores. Additionally, scores from relevant questions can be used to calculate separate average sub-scores (diarrhoea, indigestion, constipation, abdominal pain and reflux) (Svedlund et al., 1988).

The PGWB is used to evaluate quality of life and well-being. Through its 22 questions, patient's physical and psychological health is covered in terms of anxiety,

depression, well-being, self-control, general health and vitality. Scoring is from 1 to 6 points for each question and higher values indicate better self-reported quality of life and well-being. The total score is formed as a sum of each question and each sub-score as a sum of the relevant sub-category questions (Dimenäs et al., 1996; Dupuy, 1984).

SF-36 (or RAND-36) is also used to evaluate quality of life and general health. The 36 items of the questionnaire are divided into eight sub-categories consisting of physical functioning, physical role limitations, emotional role limitations, vitality, mental health, social functioning, bodily pain and general health. Scoring is from 0 to 100 points, higher scores denoting better result. The sub-category scores indicate averages of the relevant items. Physical functioning refers to an individual´s capacity to undertake daily activities such as doing dishes and cleaning, while physical role limitations indicate health issues preventing the person, e.g., from going to work or school (McHorney et al., 1994).

4.3.1.3 Clinical characteristics and their definitions (Studies I–III)

The clinical presentation was categorized into "gastrointestinal symptoms" and/or "extraintestinal symptoms". Gastrointestinal symptoms included malabsorption, diarrhoea, abdominal pain, constipation, flatulence, loose stool, vomiting, mouth ulcers, heartburn and/or bloating. Malabsorption was defined as weight loss and presence of characteristic laboratory abnormalities, such as anaemia, hypoalbuminemia, low folate or low vitamin B12. Extraintestinal symptoms included *dermatitis herpetiformis*, arthralgia, fatigue, neurological problems and/or infertility.

In Studies I–II, severity of symptoms was categorized into "no symptoms", "mild", "moderate" or "severe" symptoms based on patient report. Adherence to GFD was categorized as either "strict" or "dietary (occasional or frequent) lapses".

In Study II, diagnostic delay was categorized as 0 (screen-detected patients), <1, 1–5, 5–10 or >10 years. "First-degree relative (FDR)" referred to sibling, mother, father or offspring. "Relative" referred to any relative in a family.

4.3.1.4 Histology (Studies I–III)

The results of histological analysis of the small-bowel mucosal biopsies at the time of diagnosis were collected from the clinical pathology reports as the diagnoses were set in the healthcare system. In addition, if available, data on repeat biopsy during

GFD were recorded. According to local practice, at the time of the collection of the study dataset, a minimum of four duodenal biopsies was taken from the patients during endoscopy. Severity of small intestinal mucosal damage was evaluated from several well-orientated biopsy specimens and the degree of villous atrophy was classified as "partial", "subtotal" or "total".

4.3.1.5 Serology (Studies I–III)

The results of the coeliac disease serology at the time of diagnosis were collected from the medical records. The dataset included patients whose diagnoses had been set before the introduction of TG2 autoantibodies for diagnostic testing. Thus, in Study I, only EmA titers were considered in the analysis. In Study II, a patient was considered to be positive for coeliac disease specific autoantibodies if he/she was positive for ARA and/or EmA and/or TG2 autoantibodies at diagnosis. In Study III, patients who were positive for EmA and/or TG2 autoantibodies were regarded as positive at diagnosis. For follow-up data, both EmA and TG2 autoantibody titers were determined from serum samples drawn at the time of dataset collection (Studies I–II).

ARA and EmA were analysed using indirect immunofluorescence with rat liver, kidney or stomach tissue (ARA) (Eade et al., 1977) and human umbilical cord (EmA) (Sulkanen et al., 1998) as an antigen. Titers 1:≥5 were considered positive and positive samples were further diluted to 1:50, 1:100, 1:200, 1:500, 1:1000, 1:2000 and 1:4000, until negative. TG2 autoantibody values were tested exploiting enzyme-linked immunosorbent assay (QUANTA Lite h-tTG IgA, INOVA Diagnostics, San Diego, CA, USA), the cut-off for positivity being >30 U/L.

4.3.1.6 Statistics (Studies I–II)

Statistical analyses were performed with SPSS Statistics version 23 (IBM Corp, New York, NY, USA). P value < 0.05 was considered significant across all analyses.

In Study I, continuous variables were presented as medians with range or with lower (25th percentile) and upper (75th percentile) quartiles, or as number of subjects, and tested for statistical significance with Mann–Whitney U test. Binominal and categorical variables were presented as percentages and tested with Chi-square test.

In Study II, variables were presented as percentages and tested with Chi-square test or Fisher's Exact test, as appropriate. Statistical analyses were performed for all patients together and for younger patients (<21 years, n = 124) and adults (n = 476) separately.

4.3.2 Genetic analyses

4.3.2.1 Genotyping for HLA-DQ variants (Studies I–III)

The genotypes corresponding to coeliac disease associated HLA-DQ variants, HLA-DQ2.5, HLA-DQ8, and HLA-DQ2.2, were determined using commercial HLA typing kits (Olerup SSP low-resolution kit, Olerup SSP AB, Saltsjöbaden, Sweden or DELFIA® Celiac Disease Hybridization Assay Kit, PerkinElmer Life and Analytical Sciences, Wallac Oy, Turku, Finland) or the TaqMan chemistry based genotyping of the HLA tagging SNPs (Koskinen et al., 2009; Monsuur et al., 2008).

In Study II, patients carrying alleles *HLA-DQB1*0201* and *HLA-DQA1*0501* in *cis*-configuration were considered positive for HLA-DQ2.5. The subjects were divided into three groups according to whether they had zero, one or two copies of HLA-DQ2.5.

In Study III, patients carrying HLA-DQ2.5/HLA-DQ2.5 or HLA-DQ2.5/ HLA-DQ2.2 were considered to be high risk, patients carrying DQ2.5/X, DQ2.2/D2.2, DQ2.2/X, DQ8/DQ8, or DQ8/X, to be intermediate risk and patients carrying DQ7/X or DQ7/DQ7 to be low risk HLA category.

4.3.2.2 Genotyping for non-HLA SNPs (Studies I and III)

A subset of the study participants (n = 552 in Study I and n = 348 in Study III) were genotyped with Illumina 610-Quad BeadChip array (Illumina Inc., San Diego, CA, USA in the context of the European GWAS (Dubois et al., 2010). None of these patients were related to each other.

In Study I, 39 non-HLA SNPs previously associated with coeliac disease risk (Trynka et al., 2011) were included in the study, whereas in Study III the included SNPs were selected according to their proximal location to genes *CCR9* and *CCL25* and further filtered as described in the next paragraph. Storage of genotypes as well

as quality checks and filtering were performed with BC Genome platform, version 4.0 (BC Platforms Espoo, Finland).

4.3.2.3 Selection and filtering of genotyped *CCR9* and *CCL25* SNPs (Study III)

In Study III, the variants spanning 100 kB upstream and downstream of *CCR9* and *CCL25* genes were identified, resulting in 105 SNPs. Functional Mapping and Annotation (FUMA) platform of GWAS (Watanabe et al., 2017) was then utilized annotating SNPs fulfilling the set functional annotation criteria. The study cohort was included in the publicly available GWAS summary of statistic results (Dubois et al., 2010), which was uploaded with a pre-defined list of the 105 SNPs of interest. 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 (Boyle et al., 2012). SNPs having RegulomeDB score between 1a and 3a were selected for the study since those scores indicate the likely effect on the gene expression. In addition, eQTL mapping was performed utilizing public eQTL data, which further led to the selection of SNPs having significant eQTL effects (false discovery rate [FDR] being <0.05) on the expression of *CCR9* gene in blood (GTEx whole blood (Aguet et al., 2020), Blood eQTL (Westra et al., 2013), BIOS QTL (Zhernakova et al., 2017), and eQTLGen) and on the expression of *CCL25* gene in the intestine (GTEx data for small intestine terminal ileum, colon sigmoid and colon transverse (Aguet et al., 2020); and of the "CEDAR" study (Momozawa et al., 2018), terminal ileum). Finally, 41 SNPs with small enough RegulomeDB score (from 1a to 3a), and/or tissue-specific eQTL effect were identified. After applying quality control (QC) filtering for missing genotype rate <5%, missing genotype rate differences between the cases and controls <3%, and minor allele frequency (MAF) >5%, 348 cases and all 1817 controls remained in the analysis, and 39 SNPs passed QC.

4.3.2.4 Sequencing exonic regions of *CCR9* and *CCL25* (Study III)

Whole blood samples or leukocyte enriched buffy coats from an additional 144 HLA-DQ2-positive biopsy-proven coeliac disease patients and 144 non-celiac controls were used for DNA extraction. Extractions were performed using FlexiGene DNA kit (Qiagen, Hilden, Germany). PCR was used to amplify all coding exons and exon-intron boundaries of both *CCR9* and *CCL25*. The designed primer

pairs used in the PCR reactions are presented in Table 2. 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 analysing 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 sequence analysis and SNP identification were performed utilizing Sequencher 4.10.1 software (Gene Codes Corporation, Ann Arbor, MI, USA). Sequence data were analysed further by calculating allele frequencies for the SNPs found in the study material.

Table 2. Primer pairs used in PCR amplification of *CCR9* and *CCL25* exonic regions (from original publication III, Supplementary Table 1)

Gene	Exon	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
CCR9[a]	1	GATCCTCCAATTTCCCCATT	CTCGCCCTACAAGAAACCTG	588
	2	CCATCTTTTTGGCATTTGGT	GCATTTGCAAAACCCTCTGT	306
	3a	CTGCCATCAGACAGGACCTT	AGAGAGCAGCTGCCAATACC	582
	3b	CAGCGTGGACAGGTACATTG	AACTGTGGAAGAAGGCGATG	512
	3c	CCACCAACATTGACATCTGC	GCATGAGAACTGAAGCCACA	599
	3d	AGGAGCCCTTGGATTTTCTC	CTGCGAACCCAGCTGTTAT	490
	3e	CGCTTATTCCTTGGTATGGTG	CCGATTGAGATAGCTCACTACCTT	599
CCL25[b]	1	TTACCACTTCCCTCCACGAC	GATGGGATTGGGGAGACTTC	250
	2&3	AGTATCTGGGCGGCTGTGT	GATGCCCAGCAGAGGTCTT	400
	4	AGCTCAAGCACATGGGAGACTCA	TGGACAGCCACCGAGGACCC	411
	5	GGCCTCTCTCATTGCTTCTG	AAAGATCGCATTGGCTTCAC	564

[a]Exon 3 of gene *CCR9* was split into 5 overlapping fragments (primers 3a-3e).
[b]Exons 2 and 3 of gene *CCL25* were covered using one primer pair (2&3).
bp, base pair

4.3.2.5 Imputation of the identified *CCR9* and *CCL25* SNPs (Study III)

Since Illumina 610-Quad BeadChip array did not include SNPs identified by sequencing and thus they were not previously genotyped, those having MAF ≥5% were selected to be phased and imputed using a Finnish population-specific panel of 3,775 high-coverage (25–30 ×) whole-genome sequences (SISu v3). Sample-wise, variant-wise and post-imputation QC was applied as previously described (Cerqueira et al., 2021). Phasing of genotyped data was carried out with Eagle v2.3.5 (Loh et al., 2016) and imputation was performed with Beagle 4.1 (version 08Jun17.d8b)(Browning & Browning, 2016). After the post-imputation QC, SNPs with good enough imputation quality metrics (INFO score ≥0.8) were included in the association analyses.

4.3.2.6 Association analyses (Studies I and III)

PLINK v1.07 (Purcell et al., 2007) was used for performing association analyses of genotyped non-HLA SNPs while PLINK 2.0 (Chang et al., 2015) was used in case of imputed data.

In Study I, 37 SNPs passed the QC filters (Hardy–Weinberg Equilibrium test, $P \leq 0.05$) and were tested for association with familiar/sporadic coeliac disease.

In Study III, all 39 selected and QC filtered genotyped markers were in Hardy-Weinberg equilibrium ($P > 1 \times 10-6$) in the controls. SNPs were tested for allelic associations with coeliac disease and the selected disease phenotypes. Selected disease phenotypes included gastrointestinal symptoms, malabsorption, anaemia, severity of small bowel mucosal damage, HLA risk category and presence of coeliac disease autoantibodies. The same QC filtering criteria as applied to the association analysis of genotyped SNPs were also applied to the imputed genetic data. Nine exonic SNPs resulting from the sequencing approach were found in the Finnish reference panel and further resisted QC filtering. Thus, these SNPs were included in the association analyses.

4.3.2.7 Statistics of genetic analyses (Studies I–III)

Results were presented as odds ratios (OR) with 95% confidence intervals (95% CI). Generally, P value < 0.05 was considered significant.

In Study III, in the case of association analyses, adjustment for multiple testing and small sample size groups was performed using 10^4 permutation analysis. The generated empirical P_{EMP2} value ≤ 0.05 (uncorrected P value ≤ 0.001) was considered statistically significant. In the case of imputation analysis, BCFtools (Danecek et al., 2021) was utilized to handle the probability dosages of the genotypes and PLINK 2.0 (Chang et al., 2015) was used to perform the association analysis. The associations of the post-imputation genotype probabilities of the 9 SNPs with coeliac disease and selected coeliac disease phenotypes were tested using the frequentist likelihood score method implemented in SNPTEST v2.5.2 (Marchini et al., 2007). Associations were considered statistically significant if they reached the permutation threshold described above.

5 SUMMARY OF RESULTS

5.1 Demographic description of subjects (Studies I–III)

In Study I, there were 761 familial and 303 sporadic cases. At the time of coeliac disease diagnosis, the median age of the familial cases was 39 (range 0–81) years, and that of sporadic cases was 41 (range 1-79) years (P = 0.010) The proportion of females did not differ statistically significantly, being 72.8% among familial cases and 75.6% among sporadic cases. At follow-up evaluation in Study I, the median age of the familial cases was 50 (range 2–89) years and that of the sporadic cases was 52 (range 6–84) years. The difference did not reach statistical significance.

Study II included 605 cases consisting of 104 HLA-DQ2.5 homozygous, 401 HLA-DQ2.5 heterozygous and 100 HLA-DQ2.5 negative cases. There was no significant difference in the distribution of age groups between HLA-DQ2.5 dose groups, most of the patients (73.1–76.1%) belonging to the age group 21–65 years in each dose group at the time of diagnosis. The proportion of females was 78.0% in the HLA-DQ2.5-negative group, 73.3% in the HLA-DQ2.5-heterozygous group and 76.0% in HLA-DQ2.5-homozygous group.

In Study III, the median age of the patients at the time of diagnosis in the main cohort consisting of 625 cases was 41 (range 0.5–79) years. The percentage of females was 78.2.

5.2 Differences in genetic predisposition and clinical picture between familial and sporadic coeliac disease (Study I)

Of the familial cases, 39% had one and 61% had two or more affected relatives. In the case of 92% of familial patients, the affected family member(s) was/were first- or second-degree relative(s). Affected relative(s) were more often maternal (64%) than paternal (31%). Moreover, in 5% of familial cases, relatives from both mother's and father's side of the family were affected.

In terms of genetic predisposition, the overall distribution of HLA-DQ genotypes differed significantly (P = 0.001) among the study groups. Patients positive for HLA-DQ2.5 (either heterozygotes or homozygotes) were in the majority in both study groups, being 88.8% of familial cases and 78.0% of sporadic cases, respectively. The proportion of HLA-DQ2.5 homozygotes was higher among familial cases (20.9% vs. 11.7%) whereas patients carrying HLA-DQ2.2/HLA-DQ2.2 or HLA-DQ2.2/HLA-DQX, HLA-DQ8/HLA-DQ8 and HLA-DQX/HLA-DQX haplotypes were more common among sporadic cases (7.7% vs. 2.1, 2.7% vs. 0.9%, 6.3% vs. 2.4% respectively) (Figure 4) (Table 4 in original publication I).

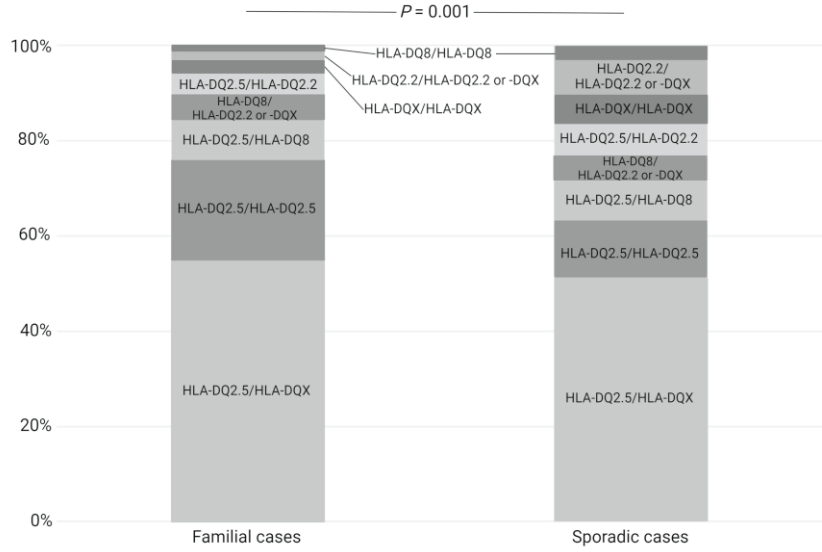

Figure 4. Distribution of coeliac disease predisposing HLA-DQ genotypes in familial and sporadic coeliac disease patients. "HLA-DQX" defines any other HLA-DQ haplotype than listed here.

Three out of 37 coeliac disease associated SNPs tested outside the HLA region (rs3748816, rs2816316 and rs2762051) were associated with increased risk for familial coeliac disease (OR 1.39, 95% CI 1.03–1.90, $P = 0.034$; OR 1.75, 95% CI 1.10–2.79, $P = 0.017$ and OR 1.48, 95% CI 1.03–2.13, $P = 0.035$ respectively) and one SNP (rs10903122) with decreased risk (OR 0.71, 95% CI 0.53–0.96, $P = 0.026$) respectively (Supplementary Table 2 in original publication I).

The main basis for the diagnosis differed significantly between groups ($P < 0.001$), familial cases being more often screen-detected than sporadic cases (26.3% vs. 2.3% respectively). Additionally, familial cases were more often seropositive for coeliac disease specific autoantibodies at diagnosis than were sporadic cases (88.6% vs. 81.8%, $P = 0.040$). Sporadic cases had significantly more often gastrointestinal presentation, *dermatitis herpetiformis* and severe symptoms than did familial cases (68.5% vs. 49.1%, 21.3% vs. 13.5%, 65.3% vs. 46.9% respectively) (Table 3). In terms of childhood diagnoses, malabsorption, or severity of small-bowel mucosal damage, no significant differences between the study groups were found (Table 1 in original publication I).

At follow-up, the study groups had similarly recovered mucosal morphology after one year on a GFD (full recovery of the villi in 59.3% of familial cases and in 60.3% of sporadic cases, $P = 0.956$). Familial cases had been longer on GFD (8 [range 4–15] years vs. 7 [range 3–13] years respectively; $P = 0.005$). Familial cases reported overall symptoms significantly less often than sporadic cases ($P = 0.004$) but were substantially more often seropositive for EmA ($P < 0.001$). With borderline significance, a greater proportion of sporadic cases had regular follow-up compared to familial cases (36.0% vs. 29.4% respectively, $P = 0.052$) (Table 3). The study groups were comparable in adherence to and management of GFD, usage of gluten-free oats and in frequency of seropositivity for TG2 autoantibodies. Familial and sporadic patients did not differ in terms of gastrointestinal symptoms measured by GSRS, but familial cases had better median general health scores measured by PGWB (14 vs. 13, $P = 0.028$) as well as higher SF-36 total, physical role functioning, vitality and mental health scores (81 vs. 78, $P = 0.011$; 100 vs. 75, $P = 0.027$; 73 vs. 70, $P = 0.015$ and 84 vs. 80, $P = 0.040$ respectively). Regarding comorbidities, there were no differences between the study groups in frequency of fractures, but familial cases were more often totally free from other conditions (21.7% vs. 14.4%, $P = 0.007$) and less often had neurological (9.9% vs. 15.3%, $P = 0.013$) and dermatological (9.4% vs. 16.8%, $P = 0.001$) diseases (Supplementary Table 1 in original publication I).

Table 3. Summary of clinical, serological and histological characteristics of 1,064 patients with familial or sporadic coeliac disease. Data were available on >90% of the patients except for the designated parameters.

	Familial *n* = 761		Sporadic *n* = 303		
	n	%	*n*	%	*P*
Main reason for the diagnosis					<0.001*
Screening	200	26.3	7	2.3	
Extraintestinal symptoms[a]	187	24.6	88	29.1	
Gastrointestinal symptoms[b]	374	49.1	207	68.5	
Other common symptoms					
Malabsorption	270	35.7	107	35.5	0.971
Dermatitis herpetiformis	102	13.5	64	21.3	0.002
Severity of symptoms at diagnosis[c]					<0.001*
No symptoms	60	10.2	3	1.4	
Mild or moderate	251	42.8	71	33.3	
Severe	275	46.9	139	65.3	
Seropositivity at diagnosis[d,e]	358	88.6	112	81.8	0.040
Severity of villous atrophy at diagnosis[f]					0.147*
Partial	192	32.9	97	39.8	
Subtotal	233	39.9	91	37.3	
Total	159	27.2	56	23.0	
Current symptoms on GFD[g]	152	21.1	85	29.5	0.004
Follow-up serology[h]					
EmA positivity	108	15.2	18	6.6	<0.001
TG2 autoantibody positivity	182	24.0	57	18.9	0.077
Regular follow-up	189	29.4	96	36.0	0.052

[a]e.g. *dermatitis herpetiformis*, arthralgia, rash, swelling, fatigue.

[b]e.g. diarrhoea, constipation, abdominal pain, flatulence, loose stools, mouth ulcers.

[d]EmA or ARA positivity.

Data was available on >90 % of the patients except for designated parameters, where [c]*n* = 586 (familial cases) and *n* = 213 (sporadic cases); [e]404 and 137; and [f]584 and 244 respectively.

[g]Any type of gastrointestinal and/or extraintestinal symptoms.

[h]Only samples taken ≥2 years after diagnosis were included.

ARA, anti-reticulin antibody; EmA, endomysial antibody; GFD, gluten-free diet; TG2, transglutaminase 2.

*Calculated across all three variables.

5.3 Impact of HLA-DQ2.5 dose on the clinical picture of coeliac disease (Study II)

The proportion of patients presenting with classical symptoms, i.e. diarrhoea and malabsorption, was lowest among heterozygotes and highest among HLA-DQ2.5-negative patients (P = 0.007) (Table 4). In terms of other types (abdominal or extraintestinal) or severity of symptoms (Table 4), presence of symptoms in childhood, diagnostic delay, presence of diarrhoea, anaemia or *dermatitis herpetiformis*, no significant differences between the groups were observed. The percentage of patients positive for coeliac disease specific autoantibodies was smallest among HLA-DQ2.5-negative subjects and greatest among HLA-DQ2.5-homozygotes (P = 0.021). Severity of mucosal damage was comparable among the study groups (Table 4). Patients homozygous for HLA-DQ2.5 most often had any relative (P = 0.011) or an FDR (P = 0.031) with coeliac disease, whereas HLA-DQ2.5-negative patients had such relatives the least often (Table 4).

During follow-up (median follow-up time 13 years, range <1–47 years), the HLA-heterozygotes managed to maintain a strict GFD most often and DQ2.5-homozygotes least often (Table 4). HLA-DQ2.5-negative patients reported current symptoms most often whereas HLA-DQ2.5-heterozygous patients did so least often (P = 0.002). At follow-up, there was no significant difference between the groups in terms of antibody positivity and mucosal recovery (Table 4), neither in terms of concomitant autoimmune diseases nor malignancy (Table 2 in original publication II).

The results of a separate analysis with adult patients only were parallel to those of the whole cohort. HLA-DQ2.5 dose groups differed significantly in terms of the presence of classical symptoms and coeliac disease specific autoantibody positivity at diagnosis (P = 0.006 and P = 0.043 respectively). Moreover, adult HLA-DQ2.5-homozygotes had more often either any relative or an FDR with coeliac disease than did HLA-DQ2.5-heterozygotes or HLA-DQ2.5-negative subjects (P = 0.001 for any relative and P = 0.003 for FDR). At follow-up, significant differences between adult study groups were observed only in adherence to the GFD and in self-reported prevailing symptoms (P = 0.016 and P = 0.002 respectively). In case of analyses with paediatric patients only, no significant differences were found in any of the parameters studied.

Table 4. Summary of clinical, serological and histological characteristics of 605 coeliac disease patients negative (X/X), heterozygous (DQ2.5/X) or homozygous (DQ2.5/ DQ2.5) for HLA-DQ2.5. Data were available on >90% of the patients except for the designated parameters.

	X/X n = 100 n (%)	DQ2.5/X n = 401 n (%)	DQ2.5/DQ2.5 n = 104 n (%)	*P*
At diagnosis				
Relatives with coeliac disease	38 (42.7)	204 (54.3)	65 (64.4)	0.011
FDRs with coeliac disease	28 (31.5)	166 (45.5)	48 (48.5)	0.031
Classical symptoms[a]	30 (30.0)	67 (16.7)	22 (21.2)	0.007
Abdominal symptoms[b]	86 (86.9)	322 (81.3)	85 (81.7)	0.426
Extraintestinal manifestations[c]	44 (44.4)	196 (49.5)	49 (47.1)	0.646
Dermatitis herpetiformis	10 (11.0)	63 (17.1)	19 (19.2)	0.270
Severity of symptoms[d]				0.484*
No symptoms	4 (5.0)	17 (5.9)	3 (4.0)	
Mild	19 (23.8)	74 (25.9)	25 (33.3)	
Moderate	12 (15.0)	35 (12.2)	4 (5.3)	
Severe	45 (56.3)	160 (55.9)	43 (57.3)	
Seropositivity[e,f]	49 (77.8)	217 (86.8)	56 (94.9)	0.021
Severity of villous atrophy[g]				0.546*
Normal morphology	2 (2.4)	10 (3.0)	2 (2.4)	
PVA	23 (27.7)	111 (33.8)	21 (25.3)	
SVA or TVA	58 (69.9)	207 (63.1)	60 (72.3)	
During follow-up				
Adherence to GFD				0.025*
Strict	91 (93.8)	376 (97.4)	92 (91.1)	
Dietary lapses	5 (5.2)	9 (2.3)	8 (7.9)	
No GFD	1 (1.0)	1 (0.3)	1 (1.0)	
Current symptoms on GFD[h,i]	26 (44.8)	59 (22.9)	20 (32.3)	0.002
Seropositivity[e,j]	4 (28.6)	10 (23.8)	1 (9.1)	0.535
Severity of villous atrophy[k]				0.108*
Normal morphology	35 (72.9)	104 (55.6)	23 (46.9)	
PVA	12 (25.0)	71 (38.0)	22 (44.9)	
SVA or TVA	1 (2.1)	12 (6.4)	4 (8.2)	

[a]Presence of diarrhoea and malabsorption.

[b]Abdominal pain, diarrhoea, loose stool, heartburn, flatulence, constipation and/or bloating.

[c]Any symptom(s) presenting outside of the gastrointestinal track.

Data were available on >90 % of the patients except for designated parameters where numbers of cases in different HLA-DQ2.5 dose groups were [d]80 (X/X), 286 (DQ2.5/X), 75 (DQ2.5/DQ2.5); [f]63, 250, 59; [g]83, 328, 83; [i]58, 258, 62; [j]14, 42, 11 and [k]48, 187, 49 respectively.

[e]TG2 autoantibody and/or EmA and/or ARA positivity.

[h]Any type of gastrointestinal and/or extraintestinal symptoms.

FDR, first-degree relative; GFD, gluten-free diet; PVA, partial villous atrophy; SVA, subtotal villous atrophy, TVA, total villous atrophy.

*Calculated across all variables.

5.4 Contribution of genetic variation of *CCR9* and *CCL25* in the clinical picture of coeliac disease (Study III)

5.4.1 Associations of genotyped functionally annotated *CCR9* and *CCL25* SNPs with coeliac disease and its different phenotypes

In *CCR9* region, SNP rs2133660 had an eQTL effect on the expression of *CCR9* in whole blood and peripheral blood mononuclear cells (PBMCs) and was associated with coeliac disease (OR = 1.20, 95% CI = 1.02–1.42, $P = 0.031$) even though the association did not resist correction for multiple testing ($P_{EMP2} = 0.558$). In addition, the same SNP similarly showed nominal associations ($P_{EMP2} > 0.05$) with malabsorption (OR = 1.45, 95% CI = 1.14–1.83, $P = 0.002$), gastrointestinal symptoms (OR = 1.29, 95% CI = 1.08–1.54, $P = 0.005$), anaemia (OR = 1.35, 95% CI = 1.01–1.81, $P = 0.041$), PVA (OR = 1.35, 95% CI = 1.01–1.79, $P = 0.039$) and negative coeliac disease serology (OR = 3.48, 95% CI = 1.38–8.74, $P = 0.005$) (Table 5). Furthermore, two SNPs with RegulomeDB scores ≤ 3a (rs7652331 and rs12493471) and one with eQTL effect on *CCR9* expression (rs1545985) were nominally associated with gastrointestinal symptoms (OR = 1.24, 95% CI = 1.03–1.47, $P = 0.012$; OR = 1.23, 95% CI = 1.03–1.47, $P = 0.022$; and OR = 1.25, 95% CI = 1.05–1.49, $P = 0.012$ respectively) and malabsorption (OR =1.35, 95% CI = 1.06–1.70, $P = 0.013$; OR =1.39, 95% CI = 1.10–1.75, $P = 0.006$ and OR =1.37, 95% CI = 1.09–1.74, $P = 0.008$ respectively). Additionally, rs12493471 showed a nominal association with anaemia (OR = 1.34, 95% CI = 1.00–1.79, $P = 0.047$). Other associations of the *CCR9* SNPs examined are shown in Table 5 and Figure 1 in original publication III. None of the studied *CCR9* SNPs nominally associated with coeliac disease and/or its distinct phenotypes showed either RegulomeDB score ≤ 3a or eQTL effect.

CCL25 SNPs rs882951, rs952444, rs12983784, and rs11667975 (all with RegulomeDB score ≤ 3a) showed nominal association ($P_{EMP2} > 0.05$) with more than one phenotype. Of these, rs882951 and rs952444 were both associated with the presence of gastrointestinal symptoms and malabsorption (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) also having the lowest RegulomeDB scores (1d and 1f respectively) as well as significant eQTL effects on *CCL25* expression in the small intestine. Rs12983784 was associated with coeliac disease (OR = 1.20, 95% CI = 1.01–1.44, P = 0.041) and coeliac disease specific seropositivity (OR = 1.43, 95% CI = 1.12–1.82, $P = 0.004$).

Rs11667975 was associated with TVA/SVA (OR = 0.65, 95% CI = 0.48–0.89, P = 0.006), high HLA risk (OR = 4.20, 95% CI = 1.21–14.53, P = 0.015) and seropositivity (OR = 1.36, 95% CI = 1.04–1.79, P = 0.025), as presented in Table 5. Additional nominal associations of *CCL25* SNPs are presented in Table 5 and in Figure 2 in original publication III. None of the associated *CCL25* SNPs showed only eQTL effect without RegulomeDB score ≤ 3a.

Table 5. Association of SNPs in the genomic regions of candidate genes *CCR9* and *CCL25* with coeliac disease and its distinct phenotypes.

Phenotype	***Candidate gene***	***Variant[a]***	***OR [95% CI]***	***P***	P_{EMP2}
Coeliac disease	*CCR9*				
		rs2133660[e]	1.20 [1.02–1.42]	0.031	0.558
	CCL25				
		rs12983784*	1.20 [1.01–1.44]	0.041	0.659
Gastrointestinal symptoms	*CCR9*				
		rs2133660[e]	1.29 [1.08–1.54]	0.005	0.145
		rs1545985[e]	1.25 [1.05–1.49]	0.012	0.289
		rs7652331*	1.24 [1.03–1.47]	0.012	0.416
		rs12493471*	1.23 [1.03–1.47]	0.022	0.439
	CCL25				
		rs952444*[e]	1.20 [1.01–1.43]	0.034	0.624
		rs882951*[e]	1.20 [1.01–1.43]	0.034	0.624
Malabsorption					
All	*CCR9*				
		rs2133660[e]	1.45 [1.14–1.83]	0.002	0.064
		rs12493471*	1.39 [1.10–1.75]	0.006	0.174
		rs1545985[e]	1.37 [1.09–1.74]	0.008	0.203
		rs7652331*	1.35 [1.06–1.70]	0.013	0.318
		rs2191031*	1.34 [1.03–1.74]	0.028	0.526
	CCL25				
		rs952444*[e]	1.32 [1.04–1.66]	0.021	0.425
		rs882951*[e]	1.32 [1.04–1.66]	0.021	0.425
Anaemia	*CCR9*				
		rs2191031*	1.51 [1.10–2.06]	0.010	0.261
		rs13069079[e]	1.53 [1.07–2.20]	0.020	0.416
		rs17214952[e]	1.48 [1.03–2.13]	0.035	0.608
		rs2133660[e]	1.35 [1.01-–1.81]	0.041	0.662
		rs12493471*	1.34 [1.00–1.79]	0.047	0.707
	CCL25				
		rs4804061*	0.51 [0.29–0.90]	0.018	0.383
Small bowel mucosal damage					
TVA/SVA	*CCL25*				
		rs1129763*	0.65 [0.48–0.89]	0.006	0.168
		rs11667975*	1.31 [1.03–1.68]	0.027	0.513
PVA	*CCR9*				
		rs2133660[e]	1.35 [1.01–1.79]	0.039	0.645
	CCL25				
		rs4804249*	0.66 [0.44–1.00]	0.046	0.693

Phenotype	*Candidate gene*	*Variant*[a]	*OR [95% CI]*	*P*	P_{EMP2}
HLA risk[b]					
High	*CCR9*				
		rs4535265[e]	2.42 [1.11–5.28]	0.025	0.536
	CCL25				
		rs11667975*	4.20 [1.21–14.53]	0.015	0.401
Coeliac disease autoantibodies[c]					
Seropositivity	*CCL25*				
		rs12983784*	1.43 [1.12–1.82]	0.004	0.114
		rs7257948*	1.38 [1.10—1.73]	0.006	0.167
		rs11667975*	1.36 [1.04–1.79]	0.025	0.506
Seronegativity	*CCR9*				
		rs2133660[e]	3.48 [1.38–8.74]	0.005	0.152
		rs13069079[e]	3.37 [1.34–8.48]	0.006	0.196
		rs17214952[e]	3.35 [1.33–8.44]	0.006	0.203
		rs10490770[e]	3.10 [1.03–9.34]	0.034	0.586
		rs12639224[e]	2.39 [0.99–5.76]	0.045	0.672
	CCL25				
		rs2287936*	2.81 [1.07–7.35]	0.028	0.485

[a]SNPs between 100 kB upstream and downstream of the candidate genes *CCR9* or *CCL25*

*SNPs with RegulomeDB scores ≤ 3a. Scores ranging from 1a to 3a indicate the strongest probability of a SNP having a regulatory effect on gene expression (Boyle et al., 2012).

[e]SNPs with significant (FDR < 0.05) cell- and tissue-specific eQTL effects on *CCR9* or *CCL25* genes in public whole blood, PBMC and small intestine data (Cerqueira et al., 2021; Momozawa et al., 2018; Zhernakova et al., 2017).

[b]Cases and controls categorized as having low HLA risk were excluded from the analysis but individuals with intermediate HLA risk were included.

[c]EmA and/or TG2 autoantibodies.

CI, confidence interval; EmA, endomysial antibody; OR, odds ratio; P_{EMP2}, empirical *P* value at 10.000 permutation threshold; PVA, partial villous atrophy; TVA/SVA, total or subtotal villous atrophy; TG2, transglutaminase 2.

5.4.2 Associations of sequenced and imputed *CCR9* and *CCL25* SNPs with coeliac disease and its distinct phenotypes

Two variants in the exonic region of *CCR9* and eight in those of *CCL25* with MAF ≥5% were identified as a result of the sequencing approach. One of the variants found in *CCR9* was previously unreported and converted threonine to alanine at position 100 of CCR9 protein sequence. However, this variant could not be imputed and thus association analyses were carried out with the remaining nine SNPs. No associations with coeliac disease were found, but the synonymous variant rs2303165 among *CCL25* SNPs was nominally associated ($P_{EMP2} > 0.05$) with PVA (OR = 0.39, 95% CI = 0.15–1.00, $P = 0.038$). Nevertheless, the SNP had no eQTL effect on *CCL25* expression in the small intestine.

6 DISCUSSION

6.1 Differences in the clinical picture between familial and sporadic coeliac disease

In terms of general clinical phenotype at diagnosis, familial and sporadic cases did not differ remarkably. However, the main reason for diagnosis among familial cases over tenfold more often than among sporadic cases was screening (Table 3). Naturally, this is most likely due to more active case-finding among risk groups, also recommended in clinical practice, to which family members of coeliac disease patients also belong. In parallel, sporadic cases were more often symptomatic at the time of coeliac disease diagnosis suffering particularly gastrointestinal symptoms significantly more often than did familial cases (Table 3). Furthermore, sporadic cases suffered more often from severe symptoms whereas familial cases were completely asymptomatic seven times more often than sporadic cases. This finding seems natural, since sporadic cases more likely seek medical advice spontaneously instead of being screened, which is more probable for familial cases. Generally, the familial form of coeliac disease seems to dominate, which may also be seen in the numbers of participating study patients representing distinct forms of the disease in Study I (761 familial cases vs. 303 sporadic cases). However, the more pronounced willingness of familial patients to participate in such studies as an explanatory factor cannot be completely excluded. A participating study patient was categorized as a familial case if they had any affected relatives, even distant ones. Yet only 8% of familial cases had third degree or more remote relative(s) as the closest affected family member(s). Such relatives are e.g. great-grandparents, great uncles/aunts and first cousins. Our findings are parallel to and corroborate the constantly growing number of studies reporting a gradual shift in the classical, gastrointestinal presentation of coeliac disease towards a milder clinical phenotype (Pedretti et al., 2021; Rampertab et al., 2006). In addition, there is evidence of a high frequency of undetected coeliac disease among the family members of coeliac disease patients (Agardh et al., 2015; Nellikkal et al., 2019) which is also in line with our findings concerning previously undetected family members. Interestingly, however, especially in view of the above-mentioned observations, the study groups were comparable in

terms of severity of mucosal damage at diagnosis, the majority of patients having either partial or subtotal villous atrophy (Table 3). On the other hand, our results concur with those of several studies reporting only a weak correlation between clinical phenotype and severity of villous atrophy (Brar et al., 2007; Kivelä et al., 2017).

At follow-up, not many statistically significant differences were found between the familial and sporadic study groups. In fact, there were only two major observations with clear discrepancies between the groups. Firstly, a larger proportion of sporadic cases reported prevailing overall symptoms (29.5% vs. 21.1%) (Table 3) accompanied by poorer general health, physical role functioning, vitality and mental health (Table 3 in original publication I). One might speculate that since sporadic cases had more often symptoms at diagnosis and furthermore, their GFD has not lasted as long as that of the familial cases (7 vs. 8 years, respectively), it would be plausible that their symptoms also persisted during follow-up. Nonetheless, there is strong evidence that the symptoms of coeliac disease typically diminish quite rapidly after the initiation of dietary treatment (Haines et al., 2008; Pekki et al., 2015; Zarkadas et al., 2006), which rather invalidates the previous speculation as an explanation. Interestingly however, in terms of co-morbidities, sporadic patients had more often neurological and dermatological disorders, which may explain the greater experience of overall symptoms if the symptoms were mistakenly thought to be caused by coeliac disease. Another alternative explanation is that severe symptoms may actually predispose to persistent symptoms being experienced even during a strict GFD (Paarlahti et al., 2013). It is also to be noted that sporadic patients lack peer support from family members, which may have a negative effect by even exacerbating the experience of symptoms as reported in terms of IBS and IBD (Halpert, 2018; Plevinsky et al., 2016). Moreover, the possibility of some other, undiagnosed disease(s) causing symptoms in the case of sporadic study patients cannot be entirely excluded. One can also speculate whether an analysis comparing the first-diagnosed individual in each family to sporadic cases would yield a different result regarding differences in symptoms and comorbidities between the two study groups. This is based on an assumption that in the case of later diagnosed family members, coeliac disease could be readily suspected before any more severe symptoms or comorbidities may begin to develop.

The second major difference between the study groups at follow-up was that the familial cases showed more pronounced EmA-positivity during GFD. In this case, higher frequency of seropositivity already at the time of diagnosis could explain the elevated autoantibody values persisting at follow-up, since, contrary to symptoms,

normalization of autoantibody levels may take several years, even though generally they decrease in weeks to months (Gidrewicz et al., 2017) although self-reported adherence to the diet was equally strict between familial and sporadic patients. However, as the data were based on patients' self-reporting, it cannot be comprehensively excluded that familial cases actually had dietary lapses. It is noteworthy that equally strict adherence and the capability to manage GFD between familial and sporadic cases was a somewhat surprising finding, since, in the case of familial patients, one could expect easier adaptation to and maintenance of the diet because of the family history of coeliac disease. In any case, Finnish coeliac disease patients generally commit conscientiously to GFD even when diagnosed by risk group screening (Kivelä et al., 2018).

6.2 Contribution of different HLA genotypes to the clinical picture of coeliac disease

When the distribution of HLA-DQ haplotypes was compared between familial and sporadic coeliac disease patients, the difference was clearly statistically significant (Table 4 in original publication I). The proportion of patients carrying one of the lower risk HLA-DQ haplotypes (HLA-DQ2.2/HLA-DQ2.2 or HLA-DQ2.2/HLA-DQX, HLA-DQ8/HLA-DQ8 and HLA-DQX/HLA-DQX) was roughly three times higher in the sporadic group than in the familial group. The opposite was observed with regard to the high-risk haplotype HLA-DQ2.5/HLA-DQ2.5: the percentage of these homozygotes was almost double among familial cases compared to sporadic cases (Figure 4). These results are understandable since it is known that risk alleles predisposing to coeliac disease accumulate in families through numerous affected relatives. Moreover, the results from Study II are in line with the aforementioned finding of HLA-DQ2.5-homozygosis among familial cases: HLA-DQ2.5-homozygotes in Study II were over-represented in having either any relative(s) or FDRs (Table 4). Actually, these differences were observed to follow HLA-DQ2.5 dose dependency (Figure I in original publication II). What is noteworthy is that the more pronounced HLA-DQ2.5-homozygocity among familial patients in Study I or HLA-DQ2.5-homozygocity *per se,* in Study II, did not reflect more severe and/or classical phenotype as it did in the recent meta-analysis by Bajor et al. (2019).

When investigating the dose effect of HLA-DQ2.5, the most prominent risk haplotype for coeliac disease, the effect was observed only in terms of coeliac disease specific serology at the time of diagnosis and the aforementioned existence of relatives. It is conceivable that HLA-DQ2.5 correlates with positive serology in a dose-dependent manner, since in the case of HLA-DQ2.5-homozygous APCs, the stimulation of gluten-specific T cells is more pronounced. This further leads to induced activation and proliferation of the T cells, which in turn take part in the induction of an anti-TG2 antibody response (du Pré et al., 2020). In earlier studies, HLA-DQ2.5 dose has been additionally associated with the presence of classical symptoms, especially among children (Bajor et al., 2019; Martínez-Ojinaga et al., 2019), but this was not observed in the present study. Unexpectedly, instead of HLA-DQ2.5-heterozygotes, the study patients negative for HLA-DQ2.5 were those presenting with classical phenotype most often. Moreover, the same patient group reported persistent symptoms at follow-up more often than did heterozygous or homozygous patients (Table 4.) Such findings contradicting earlier results may to

some extent be explained by the fact that 50% of the studies included in the meta-analysis by Bajor et al. involved only paediatric patients, whereas in the present study, the cohort consisted of both adults and children, the latter being in the minority. The small number of paediatric patients in our cohort may actually explain why significant differences in terms of the parameters studied were not found among this particular group of patients. It is plausible that additional determinants other than HLA-DQ genotype affect the disease phenotype. In fact, this assumption has gained support from a recent study showing that siblings with similar HLA-DQ genotype actually had low concordance in terms of coeliac disease phenotype (Kauma et al., 2019). Such additional factors may include both environmental factors and non-HLA genetic variants, the latter also being studied in this dissertation. Moreover, there is a probability that the factors determining clinical phenotype may even differ between children and adults, which could further account for the inconsistent findings between this dissertation and those of studies published earlier. Additionally, the number of study patients in earlier studies has been significantly lower, less than 150 patients in 16 out of 24 studies (Bajor et al., 2019), whereas the number of patients recruited for the present cohort was more than 600. This is likely to improve the reliability of our results.

At follow-up, no HLA-DQ2.5 dose dependency was observed concerning any of the parameters studied related to the clinical picture of coeliac disease (Table 2 in original publication II). Nevertheless, the groups differed statistically significantly in terms of adherence to GFD and presence of current symptoms. HLA-DQ2.5-heterozygotes were most often committed to strict adherence to the diet whereas the proportion of patients with similar adherence was lowest among homozygotes. In spite of statistical significance, HLA-DQ genotype most likely does not affect adherence to the diet in any way. Thus, it is noteworthy that there was only one person in each group who reported not following the GFD, which indicates particularly high overall adherence to GFD, as was the case also in Study I, discussed in the preceding section (6.1). Patients negative for HLA-DQ2.5 were observed to suffer most from the prevailing symptoms. This could not, however, be explained by poorer adherence to GFD, since their adherence to the diet was excellent. Moreover, neither divergent serology nor small bowel mucosal morphology accounted for more pronounced experience of prevailing symptoms since for both parameters there were no statistically significant differences between HLA-DQ2.5 dose groups. Another possible factor predisposing to persisting symptoms even during GFD has been proposed to be a long diagnostic delay (Paarlahti et al., 2013). In this case, however, this explanation also needs to be rejected since there were no

differences in the time limits within which the separate HLA-DQ2.5 dose groups received their diagnoses (Table 1 in original publication II). Alternative explanations for persisting symptoms could include small intestinal bacterial overgrowth, constant low-grade inflammation despite conscientious GFD adherence or altered intestinal microbiome (Rubio-Tapia et al., 2010; Tursi et al., 2003; Wacklin et al., 2014). In fact, it has been shown that the HLA-DQ genotype does have an effect on the intestinal microbiota and therefore this conception would deserve further investigations (Olivares et al., 2015). Interestingly, HLA-DQ2.5-negative patients were found to be symptom-wise very similar to the sporadic patients of Study I, both decidedly often having gastrointestinal phenotype at diagnosis and still experiencing symptoms at follow-up. Thus, these results concerning both HLA-DQ2.5-negative patients and sporadic coeliac disease patients add weight to considering lowering the threshold of coeliac disease suspicion for these patient groups in primary health care as well as paying special attention to them after diagnosis in clinical care in order to prevent continuous health problems and poorer quality of life.

6.3 Contribution of non-HLA variants to the clinical picture of coeliac disease

To investigate whether genetic variants outside the HLA region are associated with familial or sporadic coeliac disease, 37 SNPs previously associated with coeliac disease and genotyped in our cohort were tested for these associations. Three SNPs, rs3748816, rs2816316 and rs2762051, were associated with increased risk for familial coeliac disease. Rs3748816 maps to locus 1p36.32, which harbours e.g. candidate gene *TNFRSF14*. TNFRSF14 belongs to the TNF superfamily and functions in signal transduction pathways that activate inflammatory and inhibitory T cell immune responses (B. S. Kwon et al., 1997). The same locus has additionally been associated with rheumatoid arthritis (RA)(Coenen et al., 2009). Rs2816316 resides in locus 1q31.2, which houses candidate gene *RGS1* encoding a protein of the same name. RGS1 attenuates the signalling activity of G-proteins (Watson et al., 1996) inhibiting chemokine-induced cell migration and integrin-mediated adhesion in lymphocytes (Bowman et al., 1998). RGS1 has previously been associated with T1D and multiple sclerosis (MS) (Esposito et al., 2010; Smyth et al., 2008). The third SNP found to associate with increased risk for familial coeliac disease was intron variant rs2762051 within *DLEU1* located in 13q14.2. *DLEU1* is a long non-coding RNA (lncRNA) which, in addition to being associated with coeliac disease, has been shown

to be abnormally elevated in numerous malignancies. Additionally, it has been reported to be involved in various activities of cancerous cells, including the procession of the proliferation, invasion and inhibition of apoptosis (Song et al., 2020).

Furthermore, one SNP, rs10903122, was found to be associated with decreased risk for familial coeliac disease. This variant maps to locus 1p36.11 within gene *RUNX3,* which has also been reported to be associated with ankylosing spondylitis (AS) and psoriasis (Apel et al., 2013). RUNX3 is a transcription factor known to control blood cell development during different stages of cell lineage specification (Korinfskaya et al., 2021). All four associations with familiar coeliac disease were only modestly statistically significant (*P* range 0.017–0.035) without adjustment for small sample size. Nevertheless, since they point to essential immunological pathways, their possible role in familial coeliac disease should be further studied and confirmed by statistically more stringent means in a larger study cohort.

In the case of non-HLA variants within exonic regions of candidate genes *CCR9* and *CCL25*, several nominal associations with coeliac disease or its various phenotypes were observed (Table 5). In the *CCR9* region, none of the associated SNPs had both RegulomeDB score ≤ 3a and eQTL effect on *CCR9* expression. Nevertheless, rs12493471 was associated with gastrointestinal symptoms, malabsorption and anaemia, accompanied in each case by RegulomeDB score 3a. Furthermore, the same variant has previously been associated with coeliac disease (Sharma et al., 2016), but the association was not replicated in the present study. Rs7652331, associated with gastrointestinal symptoms and malabsorption, had an even lower RegulomeDB score, 1f. Hence it is rather interesting in terms of functionality, since SNPs scoring from 1a to 1f are, by definition, predicted to affect different DNA regulatory elements as well as to have an eQTL effect. However, rs7652331 was not observed to have any eQTL effect on *CCR9* expression, which further strengthens the assumption that most likely the probable effect of the SNP on phenotypes was due to other genes than *CCR9*. Another interesting variant was rs2133660. This SNP was associated with coeliac disease and several phenotypes (gastrointestinal symptoms, malabsorption, anaemia, PVA and negative serology), but instead of having a notable RegulomeDB score, it had a significant eQTL effect on *CCR9* expression. Thus, there is a possibility that the associations with the aforementioned phenotypes more likely reflect the effect of some other SNP(s) residing in proximity to the given locus and causing induced expression of *CCR9,* seen as a significant eQTL effect. It is worth mentioning that none of the imputed exonic variants showed any association with potential or proven functional effect.

In terms of genetic variants located in the exonic region of *CCL25*, two SNPs, rs952444 and rs882951, were associated with gastrointestinal symptoms and malabsorption with the highest likelihood (RegulomeDB scores 1d and 1f respectively) of having functional, transcription related effects on gene expression, as well as significant eQTL effects on the expression of *CCL25* in the small intestine. However, as far as we know, no evidence for altered CCL25 expression in the small intestine in different coeliac disease phenotypes has been presented. This could mean that the SNPs exert their effects via other mechanisms than those involving *CCL25* expression. Another SNP within *CCL25* region, rs2303165, that was found in the sequencing approach and later imputed, was observed to be associated with PVA, but without eQTL effects in the small intestine. This observation supports the assumption that the variant could affect the phenotype without the involvement of *CCL25*.

Nominal associations (in Study III) signify that the associations identified did not pass multiple testing correction set at 10,000 permutations. This may appear a rather stringent permutation threshold, but it was chosen because the empirically adjusted P values (P_{EMP2}) provided by such a threshold afford convincing evidence of association (Kunert-Graf et al., 2021). Due to the strict permutation threshold as well as the fairly small sample size, the statistical power was somewhat limited for detecting strong associations, and confirmatory studies with larger, but carefully phenotyped cohorts of coeliac disease patients are needed in future. As already discussed in the context of Studies I and II, it must not be forgotten that coeliac disease phenotypes most probably have additional non-genetic determinants such as the amount of gluten consumed, infections and microbes (Andrén Aronsson et al., 2019; Størdal et al., 2021).

6.4 Strengths and limitations

The main strength, that commonly applies to all the studies of this dissertation, was the opportunity to utilize very carefully phenotyped and well-defined patient material. Moreover, thorough genotype data of these phenotyped patients provided additional value by enabling investigations of the research questions between genotype and phenotype of coeliac disease that has not previously been investigated. Nevertheless, it is indisputable that retrospective study settings by their nature may appear as a limitation. Additionally, there was a variation in sample size between the parameters studied since not all the clinical data was available on each participating subject.

The major strength of Study I was its novelty in terms of study design, since there are no earlier studies comparing familial and sporadic forms of coeliac disease. Even though coeliac disease has a strong familial nature and its disease penetrance has by birth cohort studies been reported to be even deeper than generally supposed (Andrén Aronsson et al., 2019), it is equally important to acknowledge the sporadic form of the disease and investigate whether there are clinical characteristics that distinguish it from the familial form. The possibility of bias due to undetected coeliac disease among the relatives of apparently sporadic cases was substantially reduced by screening their undiagnosed family members. Nevertheless, it was impossible to obtain complete information about the family history or recruit all the family members of index patients for the study. Thus, the possibility of impaired detection of familial cases cannot be comprehensively excluded. Another issue that can be regarded as a limitation in Study I is the fact that the degree of kinship to affected family members was unevenly distributed among familial cases, since over 90% of index patients had first- or second-degree relative(s) affected, more distant relatedness being notably in the minority.

For Studies I–II, study type, i.e., follow-up study, can be regarded as a strength since it enabled long-term (>2 years) observations as well as the augmentation of study parameters to also cover patients' lives after diagnosis during the treatment period. Nevertheless, the actual studies were conducted retrospectively, which may appear as a limitation. In the case of Study II, the number of study patients was substantially higher than in earlier studies on the dose of HLA-DQ2.5 (Bajor et al., 2019). However, the small number of paediatric patients precluded reliable analyses in terms of HLA.DQ2.5 dose effect in this particular patient group.

In terms of characterization of patients' phenotype in all three studies (I–III), hidden confounding factors possibly affecting patients' experiences of symptoms

and quality of life could not be comprehensively excluded. Such factors might include e.g. sporadic autoimmunity of a close family member or latent, undiagnosed morbidity of a study patient. As for genotypic characterization, Study II considered HLA-DQ2.5 alleles only in *cis* configuration and the group of HLA-DQ2.5-negative patients was genetically rather heterogenous, including both patients with risk genotype (albeit low) for coeliac disease, e.g. HLA-DQ8 and/or HLA-DQ2.2, and patients without any major predisposing genotype. Despite clinically large patient cohorts, the number of study patients for genotype-phenotype association analyses in Studies I and III was rather modest.

In Study III, the major strength was the opportunity to utilize imputed data enabling the investigation of exonic SNPs of genes *CCR9* and *CCL25* for the first time for coeliac disease associations, since coverage of the variants within these regions in previous association studies has been sparse. Nevertheless, the sequencing approach was implemented utilizing Sanger sequencing, a method that is inferior in accuracy and coverage to Next-Generation Sequencing techniques, which might be seen as a weakness of the study. As for functional annotation analyses, unfortunately eQTL data were not available on coeliac disease patients, which prevented further studies addressing eQTL effects in patients or investigating the possible effects of the SNPs on immunological changes of coeliac disease on a larger scale.

7 CONCLUSIONS

The aims of this dissertation were to compare familial and sporadic coeliac disease in terms of phenotypic and genotypic differences and to examine the contribution of different genetic factors predisposing to coeliac disease both within the HLA region and outside of it, to the clinical phenotype of the disease. The following conclusions can be drawn based on the results obtained from Studies I-III of this dissertation:

Familial coeliac disease appeared milder in terms of clinical presentation than did sporadic coeliac disease, even though familial cases more often possessed high-risk HLA-DQ genotypes associated with coeliac disease. Sporadic coeliac disease cases were found to have more severe clinical phenotype at diagnosis as well as poorer overall health even after dietary treatment. Four distinct non-HLA SNPs, rs3748816, rs2816316, rs2762051 and rs10903122, were associated with increased risk for familial coeliac disease. Confirmatory studies establishing the role of these variants in familial coeliac disease could be of interest in future.

The contribution of HLA-DQ2.5 dose to the clinical picture of coeliac disease was modest. It was only observed for coeliac disease specific seropositivity at diagnosis and for the presence of relatives with coeliac disease. HLA-DQ2.5-negative coeliac disease patients were observed to present with classical phenotype at diagnosis as well as with prevailing symptoms after dietary treatment more often than patients with higher-risk genotypes. Taking this, and similar observations on the sporadic coeliac disease patients in Study I into consideration, special attention should be paid to these patient groups in clinics, possibly providing them with intensified support.

Non-HLA SNPs within coeliac disease candidate genes *CCR9* and *CCL25* and with probable functional effect, were nominally associated with clinical phenotypes of coeliac disease. The results suggest that, despite the importance of these genes in both small intestinal homeostasis and immunology, their genetic variation likely makes a minor contribution to the clinical picture of coeliac disease.

REFERENCES

Aaltonen, K., Laurikka, P., Huhtala, H., Mäki, M., Kaukinen, K., & Kurppa, K. (2017). The Long-Term Consumption of Oats in Celiac Disease Patients Is Safe: A Large Cross-Sectional Study. *Nutrients*, *9*(6). https://doi.org/10.3390/NU9060611

Abadie, V., Sollid, L. M., Barreiro, L. B., & Jabri, B. (2011). Integration of Genetic and Immunological Insights into a Model of Celiac Disease Pathogenesis. *Http://Dx.Doi.Org/10.1146/Annurev-Immunol-040210-092915*, *29*, 493–525. https://doi.org/10.1146/ANNUREV-IMMUNOL-040210-092915

Abraham, G., Tye-Din, J. A., Bhalala, O. G., Kowalczyk, A., Zobel, J., & Inouye, M. (2014). Accurate and robust genomic prediction of celiac disease using statistical learning. *PLoS Genetics*, *10*(2). https://doi.org/10.1371/JOURNAL.PGEN.1004137

Adams, F. (1856). Extant Works of Aretaeus the Cappadocian. *The Sydenham Society, London*.

Agardh, D., Lee, H. S., Kurppa, K., Simelld, V., Aronsson, C. A., Jörneus, O., Hummel, M., Liu, E., & Koletzko, S. (2015). Clinical features of celiac disease: A prospective birth cohort. *Pediatrics*, *135*(4), 627–634. https://doi.org/10.1542/peds.2014-3675

Aguet, F., Barbeira, A. N., Bonazzola, R., Brown, A., Castel, S. E., Jo, B., Kasela, S., Kim-Hellmuth, S., Liang, Y., Oliva, M., Flynn, E. D., Parsana, P., Fresard, L., Gamazon, E. R., Hamel, A. R., He, Y., Hormozdiari, F., Mohammadi, P., Muñoz-Aguirre, M., … Volpi, S. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, *369*(6509), 1318–1330. https://doi.org/10.1126/SCIENCE.AAZ1776

Aine, L., Mäki, M., Collin, P., & Keyriläinen, O. (1990). Dental enamel defects in celiac disease. *Journal of Oral Pathology & Medicine*, *19*(6), 241–245. https://doi.org/10.1111/J.1600-0714.1990.TB00834.X

Almeida, L. M., Gandolfi, L., Pratesi, R., Uenishi, R. H., de Almeida, F. C., Selleski, N., & Nóbrega, Y. K. de M. (2016). Presence of DQ2.2 Associated with DQ2.5 Increases the Risk for Celiac Disease. *Autoimmune Diseases*, *2016*, 5409653. https://doi.org/10.1155/2016/5409653

Andrén Aronsson, C., Lee, H. S., Hård Af Segerstad, E. M., Uusitalo, U., Yang, J., Koletzko, S., Liu, E., Kurppa, K., Bingley, P. J., Toppari, J., Ziegler, A. G., She, J. X., Hagopian, W. A., Rewers, M., Akolkar, B., Krischer, J. P., Virtanen, S. M., Norris, J. M., & Agardh, D. (2019). 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 - Journal of the American Medical Association*, *322*(6), 514–523. https://doi.org/10.1001/jama.2019.10329

Apel, M., Uebe, S., Bowes, J., Giardina, E., Korendowych, E., Juneblad, K., Pasutto, F., Ekici, A. B., McManus, R., Ho, P., Bruce, I. N., Ryan, A. W., Behrens, F., Böhm, B., Traupe, H., Lohmann, J., Gieger, C., Wichmann, H. E., Padyukov, L., … Hüffmeier, U. (2013). Variants in RUNX3 contribute to susceptibility to psoriatic arthritis, exhibiting further common ground with ankylosing spondylitis. *Arthritis and Rheumatism*, *65*(5), 1224–1231. https://doi.org/10.1002/ART.37885

Bajor, J., Szakács, Z., Farkas, N., Hegyi, P., Illés, A., Solymár, M., Pétervári, E., Balaskó, M., Pár, G., Sarlós, P., Szűcs, Á., Czimmer, J., Szemes, K., Huszár, O., Varjú, P., & Vincze, Á. (2019). Classical celiac disease is more frequent with a double dose of HLA-DQB102: A systematic review with meta-analysis. *PLoS ONE*, *14*(2). https://doi.org/10.1371/journal.pone.0212329

Barratt, S. M., Leeds, J. S., Robinson, K., Shah, P. J., Lobo, A. J., McAlindon, M. E., & Sanders, D. S. (2011). Reflux and irritable bowel syndrome are negative predictors of quality of life in coeliac disease and inflammatory bowel disease. *European Journal of Gastroenterology & Hepatology*, *23*(2), 159–165. https://doi.org/10.1097/MEG.0B013E328342A547

Barreto, M., Santos, E., Ferreira, R., Fesel, C., Fontes, M. F., Pereira, C., Martins, B., Andreia, R., Viana, J. F., Crespo, F., Vasconcelos, C., Ferreira, C., & Vicente, A. M. (2004). Evidence for CTLA4 as a susceptibility gene for systemic lupus erythematosus. *European Journal of Human Genetics : EJHG*, *12*(8), 620–626. https://doi.org/10.1038/SJ.EJHG.5201214

Barrett, J. C., Clayton, D. G., Concannon, P., Akolkar, B., Cooper, J. D., Erlich, H. A., Julier, C., Morahan, G., Nerup, J., Nierras, C., Plagnol, V., Pociot, F., Schuilenburg, H., Smyth, D. J., Stevens, H., Todd, J. A., Walker, N. M., Rich, S. S., Baskerville, T., … Zhang, Q. (2009). Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature Genetics*, *41*(6), 703–707. https://doi.org/10.1038/NG.381

Barrett, Jeffrey C., Hansoul, S., Nicolae, D. L., Cho, J. H., Duerr, R. H., Rioux, J. D., Brant, S. R., Silverberg, M. S., Taylor, K. D., Barmada, M. M., Bitton, A., Dassopoulos, T., Datta, L. W., Green, T., Griffiths, A. M., Kistner, E. O., Murtha, M. T., Regueiro, M. D., Rotter, J. I., … Daly, M. J. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genetics*, *40*(8), 955–962. https://doi.org/10.1038/NG.175

Bebb, J. R., Lawson, A., Knight, T., & Long, R. G. (2006). Long-term follow-up of coeliac disease – what do coeliac patients want? *Alimentary Pharmacology & Therapeutics*, *23*(6), 827–831. https://doi.org/10.1111/J.1365-2036.2006.02824.X

Beecham, A. H., Patsopoulos, N. A., Xifara, D. K., Davis, M. F., Kemppinen, A., Cotsapas, C., Shah, T. S., Spencer, C., Booth, D., Goris, A., Oturai, A., Saarela, J., Fontaine, B., Hemmer, B., Martin, C., Zipp, F., D'Alfonso, S., Martinelli-Boneschi, F., Taylor, B., … McCauley, J. L. (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature Genetics*, *45*(11), 1353. https://doi.org/10.1038/NG.2770

Berry, N., Basha, J., Varma, N., Varma, S., Prasad, K. K., Vaiphei, K., Dhaka, N., Sinha, S. K., & Kochhar, R. (2018). Anemia in celiac disease is multifactorial in etiology: A prospective study from India. *JGH Open : An Open Access Journal of Gastroenterology and Hepatology*, *2*(5), 196–200. https://doi.org/10.1002/JGH3.12073

Blomhoff, A., Lie, B. A., Myhre, A. G., Kemp, E. H., Weetman, A. P., Akselsen, H. E., Huseby, E. S., & Undlien, D. E. (2004). Polymorphisms in the cytotoxic T lymphocyte antigen-4 gene region confer susceptibility to Addison's disease. *The Journal of Clinical Endocrinology and Metabolism*, *89*(7), 3474–3476. https://doi.org/10.1210/JC.2003-031854

Booth, D. R., Heard, R. N., Stewart, G. J., Goris, A., Dobosi, R., Dubois, B., Lorentzen, Å. R., Celius, E. G., Harbo, H. F., Spurkland, A., Olsson, T., Kockum, I., Link, J., Hillert, J., Ban, M., Baker, A., Sawcer, S., Compston, A., Mihalova, T., … Haines, J. (2008). The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes & Immunity, 10*(1), 11–14. https://doi.org/10.1038/gene.2008.83

Borodulin, K., Tolonen, H., Jousilahti, P., Jula, A., Juolevi, A., Koskinen, S., Kuulasmaa, K., Laatikainen, T., Männistö, S., Peltonen, M., Perola, M., Puska, P., Salomaa, V., Sundvall, J., Virtanen, S. M., & Vartiainen, E. (2018). Cohort Profile: The National FINRISK Study. *International Journal of Epidemiology*, *47*(3), 696-696i. https://doi.org/10.1093/ije/dyx239

Bowman, E. P., Campbell, J. J., Druey, K. M., Scheschonka, A., Kehrl, J. H., & Butcher, E. C. (1998). Regulation of chemotactic and proadhesive responses to chemoattractant receptors by RGS (regulator of G-protein signaling) family members. *The Journal of Biological Chemistry*, *273*(43), 28040–28048. https://doi.org/10.1074/JBC.273.43.28040

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., Karczewski, K. J., Park, J., Hitz, B. C., Weng, S., Cherry, J. M., & Snyder, M. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*, *22*(9), 1790–1797. https://doi.org/10.1101/gr.137323.112

Brar, P., Kwon, G. Y., Egbuna, I. I., Holleran, S., Ramakrishnan, R., Bhagat, G., & Green, P. H. R. (2007). Lack of correlation of degree of villous atrophy with severity of clinical presentation of coeliac disease. *Digestive and Liver Disease*, *39*(1), 26–29. https://doi.org/10.1016/J.DLD.2006.07.014

Briani, C., Zara, G., Toffanin, E., Ruggero, S., Ferrarini, A., De Lazzari, F., Luca, M., Faggian, D., Grassivaro, F., Ermani, M., Pezzani, R., Giometto, B., & D'Odorico, A. (2005). Neurological complications of celiac disease and autoimmune mechanisms: preliminary data of a prospective study in adult patients. *Annals of the New York Academy of Sciences*, *1051*, 148–155. https://doi.org/10.1196/ANNALS.1361.056

Bronson, P. G., Goldstein, B. A., Ramsay, P. P., Beckman, K. B., Noble, J. A., Lane, J. A., Seldin, M. F., Kelly, J. A., Harley, J. B., Moser, K. L., Gaffney, P. M., Behrens, T. W., Criswell, L. A., & Barcellos, L. F. (2011). The rs4774 CIITA missense variant is associated with risk of systemic lupus erythematosus. *Genes and Immunity*, *12*(8), 667–671. https://doi.org/10.1038/GENE.2011.36

Browning, B. L., & Browning, S. R. (2016). Genotype Imputation with Millions of Reference Samples. *The American Journal of Human Genetics*, *98*(1), 116–126. https://doi.org/10.1016/J.AJHG.2015.11.020

Burki, T. K. (2019). Coeliac disease in Finland: what can we learn? *The Lancet Gastroenterology & Hepatology*, *4*(12), 909–910. https://doi.org/10.1016/S2468-1253(19)30354-1

Cataldo, F., Lio, D., Marino, V., Picarelli, A., Ventura, A., & Corazza, G. R. (2000). IgG(1) antiendomysium and IgG antitissue transglutaminase (anti-tTG) antibodies in coeliac patients with selective IgA deficiency. Working Groups on Celiac Disease of SIGEP and Club del Tenue. *Gut, 47(3),* 366–369. https://doi.org/10.1136/gut.47.3.366

Catassi, C., Fabiani, E., Iacono, G., D'Agate, C., Francavilla, R., Biagi, F., Volta, U., Accomando, S., Picarelli, A., De Vitis, I., Pianelli, G., Gesuita, R., Carle, F., Mandolesi, A., Bearzi, I., & Fasano, A. (2007). A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *The American Journal of Clinical Nutrition*, *85*(1), 160–166. https://doi.org/10.1093/AJCN/85.1.160

Catassi, C., Kryszak, D., Bhatti, B., Sturgeon, C., Helzlsouer, K., Clipp, S. L., Gelfond, D., Puppa, E., Sferruzza, A., & Fasano, A. (2010). Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Annals of Medicine*, *42*(7), 530–538. https://doi.org/10.3109/07853890.2010.514285

Catassi, C., Rätsch, I. M., Gandolfi, L., Pratesi, R., Fabiani, E., El Asmar, R., Frijia, M., Bearzi, I., & Vizzoni, L. (1999). Why is coeliac disease endemic in the people of the Sahara? *The Lancet*, *354*(9179), 647–648. https://doi.org/10.1016/S0140-6736(99)02609-4

Catassi, C., Verdu, E. F., Bai, J. C., & Lionetti, E. (2022). Coeliac disease. *Lancet*, *399*(10344), 2413–2426. https://doi.org/10.1016/S0140-6736(22)00794-2

Cerqueira, J. X. M., Saavalainen, P., Kurppa, K., Laurikka, P., Huhtala, H., Nykter, M., L. E. Koskinen, L., Yohannes, D. A., Kilpeläinen, E., Shcherban, A., Palotie, A., Kaukinen, K., & Lindfors, K. (2021). Independent and cumulative coeliac disease-susceptibility loci are associated with distinct disease phenotypes. *Journal of Human Genetics*, 1–11. https://doi.org/10.1038/s10038-020-00888-5

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Chen, L., Huang, Z., Yang, B., Cai, B., Su, Z., & Wang, L. (2015). Association of E26 Transformation Specific Sequence 1 Variants with Rheumatoid Arthritis in Chinese Han Population. *PloS One*, *10*(8). https://doi.org/10.1371/JOURNAL.PONE.0134875

Chorzelski, T. P., Beutner, E. H., Sulej, J., Tchorzewska, H., Jablonska, S., Kumar, V., & Kapuscinska, A. (1984). IgA anti-endomysium antibody. A new immunological marker of dermatitis herpetiformis and coeliac disease. *The British Journal of Dermatology*, *111*(4), 395–402. https://doi.org/10.1111/J.1365-2133.1984.TB06601.X

Chow, M. A., Lebwohl, B., Reilly, N. R., & Green, P. H. R. (2012). Immunoglobulin A deficiency in celiac disease. *Journal of Clinical Gastroenterology*, *46*(10), 850–854. https://doi.org/10.1097/MCG.0B013E31824B2277

Cobb, J. E., Plant, D., Flynn, E., Tadjeddine, M., Dieudé, P., Cornélis, F., Ärlestig, L., Dahlqvist, S. R., Goulielmos, G., Boumpas, D. T., Sidiropoulos, P., Krintel, S. B., Ørnbjerg, L. M., Hetland, M. L., Klareskog, L., Haeupl, T., Filer, A., Buckley, C. D., Raza, K., … Worthington, J. (2013). Identification of the tyrosine-protein phosphatase non-receptor type 2 as a rheumatoid arthritis susceptibility locus in europeans. *PloS One*, *8*(6). https://doi.org/10.1371/JOURNAL.PONE.0066456

Coeliac Disease. Current Care Guidelines. *Working group set up by the Finnish Medical Society Duodecim and the Finnish Society of Gastroenterology*. The Finnish Medical Society Duodecim (2018). https://www.kaypahoito.fi

Coenen, M. J. H., Trynka, G., Heskamp, S., Franke, B., van Diemen, C. C., Smolonska, J., van Leeuwen, M., Brouwer, E., Boezen, M. H., Postma, D. S., Platteel, M., Zanen, P., Lammers, J. W. W. J., Groen, H. J. M., Mali, W. P. T. M., Mulder, C. J., Tack, G. J., Verbeek, W. H. M., Wolters, V. M., … Zhernakova, A. (2009). Common and different genetic background for rheumatoid arthritis and coeliac disease. *Human Molecular Genetics*, *18*(21), 4195–4203. https://doi.org/10.1093/HMG/DDP365

Coleman, C., Quinn, E. M., Ryan, A. W., Conroy, J., Trimble, V., Mahmud, N., Kennedy, N., Corvin, A. P., Morris, D. W., Donohoe, G., O'Morain, C., MacMathuna, P., Byrnes, V., Kiat, C., Trynka, G., Wijmenga, C., Kelleher, D., Ennis, S., Anney, R. J. L., & McManus, R. (2016). Common polygenic variation in coeliac disease and confirmation of ZNF335 and NIFA as disease susceptibility loci. *European Journal of Human Genetics : EJHG*, *24*(2), 291–297. https://doi.org/10.1038/EJHG.2015.87

Collin, P., Mäki, M., Keyriläinen, O., Hällström, O., Reunala, T., & Pasternack, A. (1992). Selective IgA deficiency and coeliac disease. *Scandinavian Journal of Gastroenterology*, *27*(5), 367–371. https://doi.org/10.3109/00365529209000089

Collin, P., Thorell, L., Kaukinen, K., & Mäki, M. (2004). The safe threshold for gluten contamination in gluten-free products. Can trace amounts be accepted in the treatment of coeliac disease? *Alimentary Pharmacology & Therapeutics*, *19*(12), 1277–1283. https://doi.org/10.1111/J.1365-2036.2004.01961.X

Collin, P., Vilska, S., Heinonen, P. K., Hällström, O., & Pikkarainen, P. (1996). Infertility and coeliac disease. *Gut*, *39*(3), 382–384. https://doi.org/10.1136/GUT.39.3.382

Collins, J. R., & Isselbacher, K. J. (1964). Treatment of adult celiac disease (nontropical sprue). *The New England Journal of Medicine*, *271*(22), 1153–1156. https://doi.org/10.1056/NEJM196411262712208

Conteduca, G., Rossi, A., Megiorni, F., Parodi, A., Ferrera, F., Tardito, S., Battaglia, F., Kalli, F., Negrini, S., Pizzuti, A., Rizza, E., Indiveri, F., Fenoglio, D., & Filaci, G. (2014). Single nucleotide polymorphisms in the promoter regions of Foxp3 and ICOSLG genes are associated with Alopecia areata. *Clinical and Experimental Medicine*, *14*(1), 91–97. https://doi.org/10.1007/S10238-012-0224-3

Cooper, J. D., Walker, N. M., Smyth, D. J., Downes, K., Healy, B. C., & Todd, J. A. (2009). Follow-up of 1715 SNPs from the Wellcome Trust Case Control Consortium genome-wide association study in type I diabetes families. *Genes and Immunity*, *10,*(Suppl 1). https://doi.org/10.1038/GENE.2009.97

Corazza, G. R., & Villanacci, V. (2005). Coeliac disease. *Journal of Clinical Pathology*, *58*(6), 573–574. https://doi.org/10.1136/JCP.2004.023978

Corazza, Gino Roberto, Villanacci, V., Zambelli, C., Milione, M., Luinetti, O., Vindigni, C., Chioda, C., Albarello, L., Bartolini, D., & Donato, F. (2007). Comparison of the interobserver reproducibility with different histologic criteria used in celiac disease. *Clinical Gastroenterology and Hepatology : The Official Clinical Practice Journal of the American Gastroenterological Association*, *5*(7), 838–843. https://doi.org/10.1016/J.CGH.2007.03.019

Crosby, W. H., Army, U. S., & Kugler, H. W. (1957). Intraluminal biopsy of the small intestine; the intestinal biopsy capsule. *The American Journal of Digestive Diseases*, *2*(5), 236–241. https://doi.org/10.1007/BF02231100

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/GIGASCIENCE/GIAB008

Dewar, D. H., Donnelly, S. C., McLaughlin, S. D., Johnson, M. W., Ellis, H. J., & Ciclitira, P. J. (2012). Celiac disease: management of persistent symptoms in patients on a gluten-free diet. *World Journal of Gastroenterology*, *18*(12), 1348–1356. https://doi.org/10.3748/WJG.V18.I12.1348

Dexiu, C., Xianying, L., Yingchun, H., & Jiafu, L. (2022). Advances in CD247. *Scandinavian Journal of Immunology*, *96*(1), e13170. https://doi.org/10.1111/SJI.13170

Di Sabatino, A., Pickard, K. M., Gordon, J. N., Salvati, V., Mazzarella, G., Beattie, R. M., Vossenkaemper, A., Rovedatti, L., Leakey, N. A. B., Croft, N. M., Troncone, R., Corazza, G. R., Stagg, A. J., Monteleone, G., & MacDonald, T. T. (2007). Evidence for the Role of Interferon-alfa Production by Dendritic Cells in the Th1 Response in Celiac Disease. *Gastroenterology*, *133*(4), 1175–1187. https://doi.org/10.1053/J.GASTRO.2007.08.018

Dicke, W. K., Weijers, H. A., & Kamer, J. H. v. D. (1953). Coeliac disease. II. The presence in wheat of a factor having a deleterious effect in cases of coeliac disease. *Acta Paediatrica*, *42*(1), 34–42. https://doi.org/10.1111/J.1651-2227.1953.TB05563.X

Dickson, B. C., Streutker, C. J., & Chetty, R. (2006). Coeliac disease: an update for pathologists. *Journal of Clinical Pathology*, *59*(10), 1008. https://doi.org/10.1136/JCP.2005.035345

Dieterich, W., Ehnis, T., Bauer, M., Donner, P., Volta, U., Riecken, E. O., & Schuppan, D. (1997). Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nature Medicine*, *3*(7), 797–801. https://doi.org/10.1038/nm0797-797

Dimenäs, E., Carlsson, G., Glise, H., Israelsson, B., & Wiklund, I. (1996). Relevance of norm values as part of the documentation of quality of life instruments for use in upper gastrointestinal disease. *Scandinavian Journal of Gastroenterology. Supplement*, *221*, 8–13. http://www.ncbi.nlm.nih.gov/pubmed/9110389

Dominguez Castro, P., Harkin, G., Hussey, M., Christopher, B., Kiat, C., Liong Chin, J., Trimble, V., McNamara, D., MacMathuna, P., Egan, B., Ryan, B., Kevans, D., Farrell, R., Byrnes, V., Mahmud, N., & McManus, R. (2017). Changes in Presentation of Celiac Disease in Ireland From the 1960s to 2015. *Clinical Gastroenterology and Hepatology : The Official Clinical*

Practice Journal of the American Gastroenterological Association, *15*(6), 864-871.e3. https://doi.org/10.1016/J.CGH.2016.11.018

Douillard, V., Castelli, E. C., Mack, S. J., Hollenbach, J. A., Gourraud, P. A., Vince, N., & Limou, S. (2021). Approaching Genetics Through the MHC Lens: Tools and Methods for HLA Research. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/FGENE.2021.774916

Du Pré, M. F., Blazevski, J., Dewan, A. E., Stamnaes, J., Kanduri, C., Sandve, G. K., Johannesen, M. K., Lindstad, C. B., Hnida, K., Fugger, L., Melino, G., Qiao, S.-W., & Sollid, L. M. (2020). B cell tolerance and antibody production to the celiac disease autoantigen transglutaminase 2. *Journal of Experimental Medicine*, *217*(2). https://doi.org/10.1084/jem.20190860

Du Pré, M. F., & Sollid, L. M. (2015). T-cell and B-cell immunity in celiac disease. *Best Practice & Research Clinical Gastroenterology*, *29*(3), 413–423. https://doi.org/10.1016/J.BPG.2015.04.001

Dubois, P. C. A., Trynka, G., Franke, L., Hunt, K. A., Romanos, J., Curtotti, A., Zhernakova, A., Heap, G. A. R., Ádány, R., Aromaa, A., Bardella, M. T., Van Den Berg, L. H., Bockett, N. A., De La Concha, E. G., Dema, B., Fehrmann, R. S. N., Fernández-Arquero, M., Fiatal, S., Grandone, E., … Van Heel, D. A. (2010). Multiple common variants for celiac disease influencing immune gene expression. *Nature Genetics*, *42*(4), 295–302. https://doi.org/10.1038/ng.543

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Dupuy, H. J. (1984). The Psychological General Well-being (PGWB) Index. In Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies. In E. J. Wenger NK, Mattson ME, Furburg CD (Ed.), *The Psychological General Well-being (PGWB) Index. In Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*. Le Jacq Publishing.

Durazzo, M., Ferro, A., Brascugli, I., Mattivi, S., Fagoonee, S., & Pellicano, R. (2022). Extra-Intestinal Manifestations of Celiac Disease: What Should We Know in 2022? *Journal of Clinical Medicine*, *11*(1). https://doi.org/10.3390/JCM11010258

Eade, O. E., Lloyd, R. S., Lang, C., & Wright, R. (1977). IgA and IgG reticulin antibodies in coeliac and non-coeliac patients. *Gut*, *18*(12), 991–993. https://doi.org/10.1136/gut.18.12.991

Ellinghaus, D., Ellinghaus, E., Nair, R. P., Stuart, P. E., Esko, T., Metspalu, A., Debrus, S., Raelson, J. V., Tejasvi, T., Belouchi, M., West, S. L., Barker, J. N., Kõks, S., Kingo, K., Balschun, T., Palmieri, O., Annese, V., Gieger, C., Wichmann, H. E., … Franke, A. (2012). Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *American Journal of Human Genetics*, *90*(4), 636–647. https://doi.org/10.1016/J.AJHG.2012.02.020

Eriksson, D., Røyrvik, E. C., Aranda-Guillén, M., Berger, A. H., Landegren, N., Artaza, H., Hallgren, Å., Grytaas, M. A., Ström, S., Bratland, E., Botusan, I. R., Oftedal, B. E., Breivik, L., Vaudel, M., Helgeland, Ø., Falorni, A., Jørgensen, A. P., Hulting, A. L., Svartberg, J., … Husebye, E. S. (2021). GWAS for autoimmune Addison's disease identifies multiple risk loci and highlights AIRE in disease susceptibility. *Nature Communications*, *12*(1). https://doi.org/10.1038/S41467-021-21015-8

Escudero-Hernández, C., Martín, Á., de Pedro Andrés, R., Fernández-Salazar, L., Garrote, J. A., Bernardo, D., & Arranz, E. (2020). Circulating Dendritic Cells from Celiac Disease Patients Display a Gut-Homing Profile and are Differentially Modulated by Different

Gliadin-Derived Peptides. *Molecular Nutrition and Food Research*, *64*(6). https://doi.org/10.1002/mnfr.201900989

Esposito, F., Patsopoulos, N. A., Cepok, S., Kockum, I., Leppä, V., Booth, D. R., Heard, R. N., Stewart, G. J., Cox, M., Scott, R. J., Lechner-Scott, J., Goris, A., Dobosi, R., Dubois, B., Rioux, J. D., Oturai, A. B., Søndergaard, H. B., Sellebjerg, F., Sørensen, P. S., … De Jager, P. L. (2010). IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. *Genes and Immunity*, *11*(5), 397–405. https://doi.org/10.1038/GENE.2010.28

European Commission. (2014). *Commission implementing regulation (EU) No 828/2014 of 30 July 2014 on the requirements for the provision of information to consumers on the absence or reduced presence of gluten in food text with EEA relevance. Brussels.* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32014R0828

Eyre, S., Hinks, A., Bowes, J., Flynn, E., Martin, P., Wilson, A. G., Morgan, A. W., Emery, P., Steer, S., Hocking, L. J., Reid, D. M., Harrison, P., Wordsworth, P., Thomson, W., Worthington, J., & Barton, A. (2010). Overlapping genetic susceptibility variants between three autoimmune disorders: rheumatoid arthritis, type 1 diabetes and coeliac disease. *Arthritis Research & Therapy*, *12*(5). https://doi.org/10.1186/AR3139

Farre, C., Esteve, M., Curcoy, A., Cabré, E., Arranz, E., Amat, L. L., & Garcia-Tornel, S. (2002). Hypertransaminasemia in pediatric celiac disease patients and its prevalence as a diagnostic clue. *The American Journal of Gastroenterology*, *97*(12), 3176–3181. https://doi.org/10.1111/J.1572-0241.2002.07127.X

Ferguson, M. M., Wray, D., Carmichael, H. A., Russell, R. I., & Lee, F. D. (1980). Coeliac disease associated with recurrent aphthae. *Gut*, *21*(3), 223–226. https://doi.org/10.1136/GUT.21.3.223

Festen, E. A.M., Goyette, P., Scott, R., Annese, V., Zhernakova, A., Lian, J., Lefèbvre, C., Brant, S. R., Cho, J. H., Silverberg, M. S., Taylor, K. D., De Jong, D. J., Stokkers, P. C., Mcgovern, D., Palmieri, O., Achkar, J. P., Xavier, R. J., Daly, M. J., Duerr, R. H., … Rioux, J. D. (2009). Genetic variants in the region harbouring IL2/IL21 associated with ulcerative colitis. *Gut*, *58*(6), 799–804. https://doi.org/10.1136/GUT.2008.166918

Festen, Eleonora A.M., Goyette, P., Green, T., Boucher, G., Beauchamp, C., Trynka, G., Dubois, P. C., Lagacé, C., Stokkers, P. C. F., Hommes, D. W., Barisani, D., Palmieri, O., Annese, V., van Heel, D. A., Weersma, R. K., Daly, M. J., Wijmenga, C., & Rioux, J. D. (2011). A Meta-Analysis of Genome-Wide Association Scans Identifies IL18RAP, PTPN2, TAGAP, and PUS10 As Shared Risk Loci for Crohn's Disease and Celiac Disease. *PLoS Genetics*, *7*(1). https://doi.org/10.1371/JOURNAL.PGEN.1001283

Fleckenstein, B., Molberg, Ø., Qiao, S. W., Schmid, D. G., Von Mülbe, F. Der, Elgstøen, K., Jung, G., & Sollid, L. M. (2002). Gliadin T cell epitope selection by tissue transglutaminase in celiac disease. Role of enzyme specificity and pH influence on the transamidation versus deamidation process. *The Journal of Biological Chemistry*, *277*(37), 34109–34116. https://doi.org/10.1074/JBC.M204521200

Fuchs, V., Kurppa, K., Huhtala, H., Laurila, K., Mäki, M., Collin, P., Salmi, T., Luostarinen, L., Saavalainen, P., & Kaukinen, K. (2019). Serology-based criteria for adult coeliac disease have excellent accuracy across the range of pre-test probabilities. *Alimentary Pharmacology & Therapeutics*, *49*(3), 277–284. https://doi.org/10.1111/apt.15109

Fukunaga, M., Ishimura, N., Fukuyama, C., Izumi, D., Ishikawa, N., Araki, A., Oka, A., Mishiro, T., Ishihara, S., Maruyama, R., Adachi, K., & Kinoshita, Y. (2018). Celiac disease in non-clinical populations of Japan. *Journal of Gastroenterology*, *53*(2), 208–214. https://doi.org/10.1007/S00535-017-1339-9/FIGURES/2

Gabrielli, M., Cremonini, F., Fiore, G., Addolorato, G., Padalino, C., Candelli, M., De Leo, M. E., Santarelli, L., Giacovazzo, M., Gasbarrini, G., Pola, P., & Gasbarrini, A. (2003). Association between migraine and Celiac disease: results from a preliminary case-control and therapeutic study. *The American Journal of Gastroenterology*, *98*(3), 625–629. https://doi.org/10.1111/J.1572-0241.2003.07300.X

Gale, L., Wimalaratna, H., Brotodiharjo, A., & Duggan, J. M. (1997). Down's syndrome is strongly associated with coeliac disease. *Gut*, *40*(4), 492–496. https://doi.org/10.1136/GUT.40.4.492

Garner, C., Ahn, R., Ding, Y. C., Steele, L., Stoven, S., Green, P. H., Fasano, A., Murray, J. A., & Neuhausen, S. L. (2014). Genome-Wide Association Study of Celiac Disease in North America Confirms FRMD4B as New Celiac Locus. *PLoS ONE*, *9*(7). https://doi.org/10.1371/JOURNAL.PONE.0101428

Garner, C. P., Murray, J. A., Ding, Y. C., Tien, Z., van Heel, D. A., & Neuhausen, S. L. (2009). Replication of celiac disease UK genome-wide association study results in a US population. *Human Molecular Genetics*, *18*(21), 4219–4225. https://doi.org/10.1093/HMG/DDP364

Gidrewicz, D., Trevenen, C. L., Lyon, M., & Decker Butzner, J. (2017). Normalization Time of Celiac Serology in Children on a Gluten-free Diet. *Journal of Pediatric Gastroenterology and Nutrition*, *64*(3), 362–367. https://doi.org/10.1097/MPG.0000000000001270

Giersiepen, K., Lelgemann, M., Stuhldreher, N., Ronfani, L., Husby, S., Koletzko, S., & Korponay-Szabó, I. R. (2012). Accuracy of diagnostic antibody tests for coeliac disease in children: Summary of an evidence report. *Journal of Pediatric Gastroenterology and Nutrition*, *54*(2), 229–241. https://doi.org/10.1097/MPG.0B013E318216F2E5

Gnodi, E., Meneveri, R., & Barisani, D. (2022). Celiac disease: From genetics to epigenetics. *World Journal of Gastroenterology*, *28*(4), 449. https://doi.org/10.3748/WJG.V28.I4.449

Gould, M. J., Brill, H., Marcon, M. A., Munn, N. J., & Walsh, C. M. (2019). In Screening for Celiac Disease, Deamidated Gliadin Rarely Predicts Disease When Tissue Transglutaminase Is Normal. *Journal of Pediatric Gastroenterology and Nutrition*, *68*(1), 20–25. https://doi.org/10.1097/MPG.0000000000002109

Graham, D. S., Morris, D. L., Bhangale, T. R., Criswell, L. A., Syvänen, A. C., Rönnblom, L., Behrens, T. W., Graham, R. R., & Vyse, T. J. (2011). Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. *PLoS Genetics*, *7*(10). https://doi.org/10.1371/JOURNAL.PGEN.1002341

Graham, R. R., Cotsapas, C., Davies, L., Hackett, R., Lessard, C. J., Leon, J. M., Burtt, N. P., Guiducci, C., Parkin, M., Gates, C., Plenge, R. M., Behrens, T. W., Wither, J. E., Rioux, J. D., Fortin, P. R., Cunninghame Graham, D., Wong, A. K., Vyse, T. J., Daly, M. J., … Gaffney, P. M. (2008). Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nature Genetics*, *40*(9), 1059–1061. https://doi.org/10.1038/NG.200

Gray, A. M., & Papanicolas, I. N. (2010). Impact of symptoms on quality of life before and after diagnosis of coeliac disease: Results from a UK population survey. *BMC Health Services Research*, *10*. https://doi.org/10.1186/1472-6963-10-105

Gregersen, P. K., Amos, C. I., Lee, A. T., Lu, Y., Remmers, E. F., Kastner, D. L., Seldin, M. F., Criswell, L. A., Plenge, R. M., Holers, V. M., Mikuls, T. R., Sokka, T., Moreland, L. W., Bridges, S. L., Xie, G., Begovich, A. B., & Siminovitch, K. A. (2009). REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nature Genetics*, *41*(7), 820–823. https://doi.org/10.1038/NG.395

Gutierrez-Achury, J., Zorro, M. M., Ricaño-Ponce, I., Zhernakova, D. V., Diogo, D., Raychaudhuri, S., Franke, L., Trynka, G., Wijmenga, C., & Zhernakova, A. (2016). Functional implications of disease-specific variants in loci jointly associated with coeliac disease and rheumatoid arthritis. *Human Molecular Genetics*, *25*(1), 180–190. https://doi.org/10.1093/HMG/DDV455

Hære, P., Høie, O., Schulz, T., Schönhardt, I., Raki, M., & Lundin, K. E. A. (2016). Long-term mucosal recovery and healing in celiac disease is the rule – not the exception. *Scandinavian Journal of Gastroenterology* *51*(12), 1439–1446. https://doi.org/10.1080/00365521.2016.1218540

Haines, M. L., Anderson, R. P., & Gibson, P. R. (2008). Systematic review: the evidence base for long-term management of coeliac disease. *Alimentary Pharmacology & Therapeutics*, *28*(9), 1042–1066. https://doi.org/10.1111/J.1365-2036.2008.03820.X

Halpert, A. (2018). Irritable Bowel Syndrome: Patient-Provider Interaction and Patient Education. *Journal of Clinical Medicine*, *7*(1), 3. https://doi.org/10.3390/jcm7010003

Han, T. U., Cho, S. K., Kim, T., Joo, Y. Bin, Bae, S. C., & Kang, C. (2013). Association of an activity-enhancing variant of IRAK1 and an MECP2-IRAK1 haplotype with increased susceptibility to rheumatoid arthritis. *Arthritis and Rheumatism*, *65*(3), 590–598. https://doi.org/10.1002/ART.37804

Heyman, M., Abed, J., Lebreton, C., & Cerf-Bensussan, N. (2012). Intestinal permeability in coeliac disease: insight into mechanisms and relevance to pathogenesis. *Gut*, *61*(9), 1355–1364. https://doi.org/10.1136/GUTJNL-2011-300327

Hill, P., Thompson, S., & Holmes, G. (1991). IgA anti-gliadin antibodies in adult celiac disease. *Clinical Chemistry*, *37*(5), 647–650.

Hinks, A., Cobb, J., Marion, M. C., Prahalad, S., Sudman, M., Bowes, J., Martin, P., Comeau, M. E., Sajuthi, S., Andrews, R., Brown, M., Chen, W. M., Concannon, P., Deloukas, P., Edkins, S., Eyre, S., Gaffney, P. M., Guthery, S. L., Guthridge, J. M., … Thompson, S. D. (2013). Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nature Genetics*, *45*(6), 664. https://doi.org/10.1038/NG.2614

Holm, K., Mäki, M., Vuolteenaho, N., Mustalahti, K., Ashorn, M., Ruuska, T., & Kaukinen, K. (2006). Oats in the treatment of childhood coeliac disease: a 2-year controlled trial and a long-term clinical follow-up study. *Alimentary Pharmacology & Therapeutics*, *23*(10), 1463–1472. https://doi.org/10.1111/J.1365-2036.2006.02908.X

Holmes, G. K. T., Prior, P., Lane, M. R., Pope, D., & Allan, R. N. (1989). Malignancy in coeliac disease--effect of a gluten free diet. *Gut*, *30*(3), 333–338. https://doi.org/10.1136/GUT.30.3.333

Hujoel, I. A., & Murray, J. A. (2020). Refractory Celiac Disease. In *Current Gastroenterology Reports* *22*(4), 18. https://doi.org/10.1007/s11894-020-0756-8

Hunt, K. A., Zhernakova, A., Turner, G., Heap, G. A. R., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L. C., Ryan, A. W., Panesar, D., Gwilliam, R., Takeuchi, F., McLaren, W. M., Holmes, G. K. T., Howdle, P. D., Walters, J. R. F., Sanders, D. S., Playford, R. J., Trynka, G., … Van Heel, D. A. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. *Nature Genetics*, *40*(4), 395–402. https://doi.org/10.1038/ng.102

Husby, S., Koletzko, S., Korponay-Szabó, I. R., Mearin, M. L., Phillips, A., Shamir, R., Troncone, R., Giersiepen, K., Branski, D., Catassi, C., Lelgeman, M., Mäki, M., Ribes-Koninckx, C., Ventura, A., & Zimmer, K. P. (2012). European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *Journal of Pediatric Gastroenterology and Nutrition*, *54*(1), 136–160. https://doi.org/10.1097/MPG.0B013E31821A23D0

Husby, Steffen, Koletzko, S., Korponay-Szabó, I., Kurppa, K., Mearin, M. L., Ribes-Koninckx, C., Shamir, R., Troncone, R., Auricchio, R., Castillejo, G., Christensen, R., Dolinsek, J., Gillett, P., Hróbjartsson, A., Koltai, T., Maki, M., Nielsen, S. M., Popp, A., Størdal, K., … Wessels, M. (2020). European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. *Journal of Pediatric Gastroenterology and Nutrition*, *70*(1), 141–156. https://doi.org/10.1097/MPG.0000000000002497

Iltanen, S., Collin, P., Korpela, M., Holm, K., Partanen, J., Polvi, A., & Mäki, M. (1999). Celiac disease and markers of celiac disease latency in patients with primary Sjögren's syndrome. *The American Journal of Gastroenterology*, *94*(4), 1042–1046. https://doi.org/10.1111/J.1572-0241.1999.01011.X

Ilus, T., Kaukinen, K., Virta, L. J., Huhtala, H., Mäki, M., Kurppa, K., Heikkinen, M., Heikura, M., Hirsi, E., Jantunen, K., Moilanen, V., Nielsen, C., Puhto, M., Pölkki, H., Vihriälä, I., & Collin, P. (2014). Refractory coeliac disease in a country with a high prevalence of clinically-diagnosed coeliac disease. *Alimentary Pharmacology & Therapeutics*, *39*(4), 418–425. https://doi.org/10.1111/APT.12606

Iversen, R., Amundsen, S. F., Kleppa, L., du Pré, M. F., Stamnaes, J., & Sollid, L. M. (2020). Evidence That Pathogenic Transglutaminase 2 in Celiac Disease Derives From Enterocytes. *Gastroenterology*, *159*(2), 788–790. https://doi.org/10.1053/J.GASTRO.2020.04.018

Jabri, B., & Abadie, V. (2015). IL-15 functions as a danger signal to regulate tissue-resident T cells and tissue destruction. *Nature Reviews. Immunology*, *15*(12), 771–783. https://doi.org/10.1038/NRI3919

Jabri, B., & Sollid, L. M. (2017). T cells in celiac disease. *Journal of Immunology)*, *198*(8), 3005. https://doi.org/10.4049/JIMMUNOL.1601693

Jin, Y., Birlea, S. A., Fain, P. R., Ferrara, T. M., Ben, S., Riccardi, S. L., Cole, J. B., Gowan, K., Holland, P. J., Bennett, D. C., Luiten, R. M., Wolkerstorfer, A., Van Der Veen, J. P. W., Hartmann, A., Eichner, S., Schuler, G., Van Geel, N., Lambert, J., Kemp, E. H., … Spritz, R. A. (2012). Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nature Genetics*, *44*(6), 676. https://doi.org/10.1038/NG.2272

Jin, Y., Birlea, S. A., Fain, P. R., Gowan, K., Riccardi, S. L., Holland, P. J., Mailloux, C. M., Sufit, A. J. D., Hutton, S. M., Amadi-Myers, A., Bennett, D. C., Wallace, M. R., McCormack, W. T., Kemp, E. H., Gawkrodger, D. J., Weetman, A. P., Picardo, M., Leone, G., Taïeb, A., … Spritz, R. A. (2010). Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *The New England Journal of Medicine*, *362*(18), 1686–1697. https://doi.org/10.1056/NEJMOA0908547

Kalliokoski, S., Piqueras, V. O., Frías, R., Sulic, A. M., Määttä, J. A. E., Kähkönen, N., Viiri, K., Huhtala, H., Pasternack, A., Laurila, K., Sblattero, D., Korponay-Szabó, I. R., Mäki, M., Caja, S., Kaukinen, K., & Lindfors, K. (2017). Transglutaminase 2-specific coeliac disease autoantibodies induce morphological changes and signs of inflammation in the small-bowel mucosa of mice. *Amino Acids*, *49*(3), 529–540. https://doi.org/10.1007/S00726-016-2306-0

Karell, K., Louka, A. S., Moodie, S. J., Ascher, H., Clot, F., Greco, L., Ciclitira, P. J., Sollid, L. M., Partanen, J., European Genetics Cluster on Celiac Disease (2003). HLA types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. *Human Immunology, 64(*4*)*, 469–477. https://doi.org/10.1016/s0198-8859(03)00027-2.

Kaufman, K. M., Zhao, J., Kelly, J. A., Hughes, T., Adler, A., Sanchez, E., Ojwang, J. O., Langefeld, C. D., Ziegler, J. T., Williams, A. H., Comeau, M. E., Marion, M. C., Glenn, S. B., Cantor, R. M., Grossman, J. M., Hahn, B. H., Song, Y. W., Yu, C. Y., James, J. A., … Tsao, B. P. (2013). Fine mapping of Xq28: both MECP2 and IRAK1 contribute to risk for systemic lupus erythematosus in multiple ancestral groups. *Annals of the Rheumatic Diseases*, *72*(3), 437–444. https://doi.org/10.1136/ANNRHEUMDIS-2012-201851

Kauma, S., Kaukinen, K., Huhtala, H., Kivelä, L., Pekki, H., Salmi, T., Saavalainen, P., Lindfors, K., Kurppa, K., Kauma, S., Kaukinen, K., Huhtala, H., Kivelä, L., Pekki, H., Salmi, T., Saavalainen, P., Lindfors, K., & Kurppa, K. (2019). The Phenotype of Celiac Disease Has Low Concordance between Siblings, Despite a Similar Distribution of HLA Haplotypes. *Nutrients*, *11*(2), 479. https://doi.org/10.3390/nu11020479

Kemppainen, K. M., Lynch, K. F., Liu, E., Lönnrot, M., Simell, V., Briese, T., Koletzko, S., Hagopian, W., Rewers, M., She, J. X., Simell, O., Toppari, J., Ziegler, A. G., Akolkar, B., Krischer, J. P., Lernmark, Å., Hyöty, H., Triplett, E. W., & Agardh, D. (2017). Factors That Increase Risk of Celiac Disease Autoimmunity After a Gastrointestinal Infection in Early Life. *Clinical Gastroenterology and Hepatology, 15*(5), 694-702.e5. https://doi.org/10.1016/J.CGH.2016.10.033

Kim, C. Y., Quarsten, H., Bergseng, E., Khosla, C., & Sollid, L. M. (2004). Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4175–4179. https://doi.org/10.1073/PNAS.0306885101

King, J. A., Jeong, J., Underwood, F. E., Quan, J., Panaccione, N., Windsor, J. W., Coward, S., Debruyn, J., Ronksley, P. E., Shaheen, A. A., Quan, H., Godley, J., Veldhuyzen Van Zanten, S., Lebwohl, B., Ng, S. C., Ludvigsson, J. F., & Kaplan, G. G. (2020). Incidence of Celiac Disease Is Increasing over Time: A Systematic Review and Meta-analysis. *American Journal of Gastroenterology*, *115*(4), 507–525. https://doi.org/10.14309/AJG.0000000000000523

Kivelä, L., Caminero, A., Leffler, D. A., Pinto-Sanchez, M. I., Tye-Din, J. A., & Lindfors, K. (2021). Current and emerging therapies for coeliac disease. *Nature Reviews. Gastroenterology & Hepatology*, *18*(3), 181–195. https://doi.org/10.1038/S41575-020-00378-1

Kivelä, L., Kaukinen, K., Huhtala, H., Lähdeaho, M.-L., Mäki, M., & Kurppa, K. (2017). At-Risk Screened Children with Celiac Disease are Comparable in Disease Severity and Dietary Adherence to Those Found because of Clinical Suspicion: A Large Cohort Study. *The Journal of Pediatrics*, *183*, 115-121.e2. https://doi.org/10.1016/J.JPEDS.2016.12.077

Kivelä, L., Kaukinen, K., Lähdeaho, M. L., Huhtala, H., Ashorn, M., Ruuska, T., Hiltunen, P., Visakorpi, J., Mäki, M., & Kurppa, K. (2015). Presentation of Celiac Disease in Finnish Children Is No Longer Changing: A 50-Year Perspective. *Journal of Pediatrics*, *167*(5), 1109-1115.e1. https://doi.org/10.1016/j.jpeds.2015.07.057

Kivelä, L., Popp, A., Arvola, T., Huhtala, H., Kaukinen, K., & Kurppa, K. (2018). Long-term health and treatment outcomes in adult coeliac disease patients diagnosed by screening in childhood. *United European Gastroenterology Journal*, *6*(7), 1022–1031. https://doi.org/10.1177/2050640618778386

Korinfskaya, S., Parameswaran, S., Weirauch, M. T., & Barski, A. (2021). Runx Transcription Factors in T Cells—What Is Beyond Thymic Development? *Frontiers in Immunology*, *12*. https://doi.org/10.3389/FIMMU.2021.701924

Korponay-Szabó, I. R., Halttunen, T., Szalai, Z., Laurila, K., Király, R., Kovács, J. B., Fésüs, L., & Mki, M. (2004). In vivo targeting of intestinal and extraintestinal transglutaminase 2 by coeliac autoantibodies. *Gut*, *53*(5), 641–648. https://doi.org/10.1136/GUT.2003.024836

Koskinen, I., Hervonen, K., Huhtala, H., Pasternack, C., Salmi, T., Reunala, T., Collin, P., & Kaukinen, K. (2022). Mortality and causes of death in different celiac disease phenotypes

during long-term follow-up. *Digestive and Liver Disease*. https://doi.org/10.1016/J.DLD.2022.04.016

Koskinen, L. L. E., Korponay-Szabo, I. R., Viiri, K., Juuti-Uusitalo, K., Kaukinen, K., Lindfors, K., Mustalahti, K., Kurppa, K., Ádány, R., Pocsai, Z., Széles, G., Einarsdottir, E., Wijmenga, C., Mäki, M., Partanen, J., Kere, J., & Saavalainen, P. (2008). Myosin IXB gene region and gluten intolerance: linkage to coeliac disease and a putative dermatitis herpetiformis association. *Journal of Medical Genetics*, *45*(4), 222–227. https://doi.org/10.1136/JMG.2007.053991

Koskinen, L., Romanos, J., Kaukinen, K., Mustalahti, K., Korponay-Szabo, I., Barisani, D., Bardella, M. T., Ziberna, F., Vatta, S., Széles, G., Pocsai, Z., Karell, K., Haimila, K., Ádány, R., Not, T., Ventura, A., Mäki, M., Partanen, J., Wijmenga, C., & Saavalainen, P. (2009). Cost-effective HLA typing with tagging SNPs predicts celiac disease risk haplotypes in the Finnish, Hungarian, and Italian populations. *Immunogenetics*, *61*(4), 247–256. https://doi.org/10.1007/s00251-009-0361-3

Kröger, S., Kurppa, K., Repo, M., Huhtala, H., Kaukinen, K., Lindfors, K., Arvola, T., & Kivelä, L. (2020). Severity of Villous Atrophy at Diagnosis in Childhood Does Not Predict Long-term Outcomes in Celiac Disease. *Journal of Pediatric Gastroenterology and Nutrition*, *71*(1), 71–77. https://doi.org/10.1097/MPG.0000000000002675

Kuitunen, P., Kosnai, I., & Savilahti, E. (1982). Morphometric study of the jejunal mucosa in various childhood enteropathies with special reference to intraepithelial lymphocytes. *Journal of Pediatric Gastroenterology and Nutrition*, *1*(4), 525–531. https://doi.org/10.1097/00005176-198212000-00012

Kunert-Graf, J. M., Sakhanenko, N. A., & Galas, D. J. (2021). Optimized permutation testing for information theoretic measures of multi-gene interactions. *BMC Bioinformatics*, *22*(1), 180. https://doi.org/10.1186/s12859-021-04107-6

Kunkel, E. J., Campbell, J. J., Haraldsen, G., Pan, J., Boisvert, J., Roberts, A. I., Ebert, E. C., Vierra, M. A., Goodman, S. B., Genovese, M. C., Wardlaw, A. J., Greenberg, H. B., Parker, C. M., Butcher, E. C., Andrew, D. P., & Agace, W. W. (2000). Lymphocyte CC chemokine receptor 9 and epithelial thymus-expressed chemokine (TECK) expression distinguish the small intestinal immune compartment: Epithelial expression of tissue-specific chemokines as an organizing principle in regional immunity. *The Journal of Experimental Medicine*, *192*(5), 761–767. https://doi.org/10.1084/JEM.192.5.761

Kurppa, K., Räsänen, T., Collin, P., Iltanen, S., Huhtala, H., Ashorn, M., Saavalainen, P., Haimila, K., Partanen, J., Mäki, M., & Kaukinen, K. (2012). Endomysial antibodies predict celiac disease irrespective of the titers or clinical presentation. *World Journal of Gastroenterology : WJG*, *18*(20), 2511. https://doi.org/10.3748/WJG.V18.I20.2511

Kurppa, K., Salminiemi, J., Ukkola, A., Saavalainen, P., Löytynoja, K., Laurila, K., Collin, P., Mäki, M., & Kaukinen, K. (2012). Utility of the new ESPGHAN criteria for the diagnosis of celiac disease in at-risk groups. *Journal of Pediatric Gastroenterology and Nutrition*, *54*(3), 387–391. https://doi.org/10.1097/MPG.0B013E3182407C6B

Kwon, B. S., Tan, K. B., Ni, J., Kwi-Ok-Oh, Lee, Z. H., Kim, K. K., Kim, Y. J., Wang, S., Gentz, R., Yu, G. L., Harrop, J., Lyn, S. D., Silverman, C., Porter, T. G., Truneh, A., & Young, P. R. (1997). A newly identified member of the tumor necrosis factor receptor superfamily with a wide tissue distribution and involvement in lymphocyte activation. *The Journal of Biological Chemistry*, *272*(22), 14272–14276. https://doi.org/10.1074/JBC.272.22.14272

Kwon, Y. C., Lim, J., Bang, S. Y., Ha, E., Hwang, M. Y., Yoon, K., Choe, J. Y., Yoo, D. H., Lee, S. S., Lee, J., Chung, W. T., Kim, T. H., Sung, Y. K., Shim, S. C., Choi, C. B., Jun, J. B., Kang, Y. M., Shin, J. M., Lee, Y. K., … Bae, S. C. (2020). Genome-wide association study in a Korean population identifies six novel susceptibility loci for rheumatoid arthritis.

Annals of the Rheumatic Diseases, *79*(11), 1438–1445. https://doi.org/10.1136/ANNRHEUMDIS-2020-217663

Ladinser, B., Rossipal, E., & Pittschieler, K. (1994). Endomysium antibodies in coeliac disease: an improved method. *Gut*, *35*(6), 776–778. https://doi.org/10.1136/GUT.35.6.776

Laurikka, P., Nurminen, S., Kivelä, L., & Kurppa, K. (2018). Extraintestinal Manifestations of Celiac Disease: Early Detection for Better Long-Term Outcomes. *Nutrients*, *10*(8). https://doi.org/10.3390/NU10081015

Lebwohl, B., Sanders, D. S., & Green, P. H. R. (2018). Coeliac disease. *The Lancet*, *391*(10115), 70–81. https://doi.org/10.1016/S0140-6736(17)31796-8

Lessard, C. J., Adrianto, I., Ice, J. A., Wiley, G. B., Kelly, J. A., Glenn, S. B., Adler, A. J., Li, H., Rasmussen, A., Williams, A. H., Ziegler, J., Comeau, M. E., Marion, M., Wakeland, B. E., Liang, C., Ramos, P. S., Grundahl, K. M., Gallant, C. J., Alarcón, G. S., … Moser, K. L. (2012). Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. *American Journal of Human Genetics*, *90*(4), 648–660. https://doi.org/10.1016/J.AJHG.2012.02.023

Lessard, C. J., Li, H., Adrianto, I., Ice, J. A., Rasmussen, A., Grundahl, K. M., Kelly, J. A., Dozmorov, M. G., Miceli-Richard, C., Bowman, S., Lester, S., Eriksson, P., Eloranta, M. L., Brun, J. G., Gøransson, L. G., Harboe, E., Guthridge, J. M., Kaufman, K. M., Kvarnström, M., … Sivils, K. L. (2013). Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. *Nature Genetics*, *45*(11), 1284–1294. https://doi.org/10.1038/NG.2792

Levescot, A., Malamut, G., & Cerf-Bensussan, N. (2022). Immunopathogenesis and environmental triggers in coeliac disease. *Gut*. https://doi.org/10.1136/GUTJNL-2021-326257

Lewis, N. R., & Scott, B. B. (2006). Systematic review: the use of serology to exclude or diagnose coeliac disease (a comparison of the endomysial and tissue transglutaminase antibody tests). *Alimentary Pharmacology & Therapeutics*, *24*(1), 47–54. https://doi.org/10.1111/J.1365-2036.2006.02967.X

Lill, C. M., Luessi, F., Alcina, A., Sokolova, E. A., Ugidos, N., de la Hera, B., Guillot-Noël, L., Malhotra, S., Reinthaler, E., Schjeide, B. M. M., Mescheriakova, J. Y., Mashychev, A., Wohlers, I., Akkad, D. A., Aktas, O., Alloza, I., Antigüedad, A., Arroyo, R., Astobiza, I., … Bertram, L. (2015). Genome-wide significant association with seven novel multiple sclerosis risk loci. *Journal of Medical Genetics*, *52*(12), 848–855. https://doi.org/10.1136/JMEDGENET-2015-103442

Lima, S. C., Gomes da Silva, I. I. F., Nascimento, D. de Q., de Moura, R. R., Mesquita, M. da S., Asano, N. M. J., Fernandes, G. V., Valente, L. M., Rushansky, E., Mariano, M. H. Q. de A., Xavier, R. M., Chies, J. A. B., Crovella, S., & Sandrin-Garcia, P. (2021). CIITA gene polymorphism (rs3087456) in systemic lupus erythematosus and rheumatoid arthritis: A population-based cohort study. *International Journal of Immunogenetics*, *48*(5), 429–434. https://doi.org/10.1111/IJI.12548

Lindberg, T., Nilsson, L., Borulf, S., Cavell, B., Fällström, S. P., Jansson, U., Stenhammar, L., & Stintzing, G. (1985). Serum IgA and IgG gliadin antibodies and small intestinal mucosal damage in children. *Journal of Pediatric Gastroenterology and Nutrition*, *4*(6), 917–922. https://doi.org/10.1097/00005176-198512000-00011

Lindfors, K., Ciacci, C., Kurppa, K., Lundin, K. E. A., Makharia, G. K., Mearin, M. L., Murray, J. A., Verdu, E. F., & Kaukinen, K. (2019). Coeliac disease. *Nature Reviews Disease Primers, 5*(1), 1–18. https://doi.org/10.1038/s41572-018-0054-z

Lindfors, K., Lin, J., Lee, H. S., Hyöty, H., Nykter, M., Kurppa, K., Liu, E., Koletzko, S., Rewers, M., Hagopian, W., Toppari, J., Ziegler, A. G., Akolkar, B., Krischer, J. P., Petrosino, J. F., Lloyd, R. E., & Agardh, D. (2020). Metagenomics of the faecal virome indicate a cumulative effect of enterovirus and gluten amount on the risk of coeliac disease autoimmunity in genetically at risk children: the TEDDY study. *Gut*, *69*(8), 1416–1422. https://doi.org/10.1136/GUTJNL-2019-319809

Lindstad, C. B., du Pré, M. F., Stamnaes, J., & Sollid, L. M. (2022). Injection of prototypic celiac anti-transglutaminase 2 antibodies in mice does not cause enteropathy. *PloS One*, *17*(4). https://doi.org/10.1371/JOURNAL.PONE.0266543

Liu, E., Lee, H.-S., Aronsson, C. A., Hagopian, W. A., Koletzko, S., Rewers, M. J., Eisenbarth, G. S., Bingley, P. J., Bonifacio, E., Simell, V., & Agardh, D. (2014). Risk of Pediatric Celiac Disease According to HLA Haplotype and Country. *New England Journal of Medicine*, *371*(1), 42–49. https://doi.org/10.1056/nejmoa1313977

Liu, J. Z., Almarri, M. A., Gaffney, D. J., Mells, G. F., Jostins, L., Cordell, H. J., Ducker, S. J., Day, D. B., Heneghan, M. A., Neuberger, J. M., Donaldson, P. T., Bathgate, A. J., Burroughs, A., Davies, M. H., Jones, D. E., Alexander, G. J., Barrett, J. C., Sandford, R. N., & Anderson, C. A. (2012). Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nature Genetics*, *44*(10), 1137–1141. https://doi.org/10.1038/NG.2395

Liu, J. Z., Hov, J. R., Folseraas, T., Ellinghaus, E., Rushbrook, S. M., Doncheva, N. T., Andreassen, O. A., Weersma, R. K., Weismüller, T. J., Eksteen, B., Invernizzi, P., Hirschfield, G. M., Gotthardt, D. N., Pares, A., Ellinghaus, D., Shah, T., Juran, B. D., Milkiewicz, P., Rust, C., … Karlsen, T. H. (2013). Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nature Genetics*, *45*(6), 670–675. https://doi.org/10.1038/NG.2616

Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., Reshef, Y. A., Finucane, H. K., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., & Price, A. L. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nature Genetics*, *48*(11), 1443–1448. https://doi.org/10.1038/NG.3679

Lohi, S., Mustalahti, K., Kaukinen, K., Laurila, K., Collin, P., Rissanen, H., Lohi, O., Bravi, E., Gasparin, M., Reunanen, A., & Mäki, M. (2007). Increasing prevalence of coeliac disease over time. *Alimentary Pharmacology & Therapeutics*, *26*(9), 1217–1225. https://doi.org/10.1111/J.1365-2036.2007.03502.X

Lubrano, E., Ciacci, C., Ames, P. R. J., Mazzacca, G., Oriente, P., & Scarpa, R. (1996). The arthritis of coeliac disease: prevalence and pattern in 200 adult patients. *British Journal of Rheumatology*, *35*(12), 1314–1318. https://doi.org/10.1093/RHEUMATOLOGY/35.12.1314

Ludvigsson, J. F., Bai, J. C., Biagi, F., Card, T. R., Ciacci, C., Ciclitira, P. J., Green, P. H. R., Hadjivassiliou, M., Holdoway, A., Van Heel, D. A., Kaukinen, K., Leffler, D. A., Leonard, J. N., Lundin, K. E. A., McGough, N., Davidson, M., Murray, J. A., Swift, G. L., Walker, M. M., … Sanders, D. S. (2014). Diagnosis and management of adult coeliac disease: guidelines from the British Society of Gastroenterology. *Gut*, *63*(8), 1210–1228. https://doi.org/10.1136/GUTJNL-2013-306578

Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H. R., Hadjivassiliou, M., Kaukinen, K., Kelly, C. P., Leonard, J. N., Lundin, K. E. A., Murray, J. A., Sanders, D. S., Walker, M. M., Zingone, F., & Ciacci, C. (2013). The Oslo definitions for coeliac disease and related terms. *Gut*, *62*(1), 43–52. https://doi.org/10.1136/gutjnl-2011-301346

Lähdeaho, M. L., Mäki, M., Laurila, K., Huhtala, H., & Kaukinen, K. (2011). Small- bowel mucosal changes and antibody responses after low- and moderate-dose gluten challenge in celiac disease. *BMC Gastroenterology*, *11*. https://doi.org/10.1186/1471-230X-11-129

Mahadev, S. H., Gardner, R., Lewis, S. K., Lebwohl, B., & Green, P. H. (2016). Quality of Life in Screen-detected Celiac Disease Patients in the United States. *Journal of Clinical Gastroenterology*, *50*(5), 393–397. https://doi.org/10.1097/MCG.0000000000000433

Makharia, G. K., Singh, P., Catassi, C., Sanders, D. S., Leffler, D., Ali, R. A. R., & Bai, J. C. (2022). The global burden of coeliac disease: opportunities and challenges. *Nature Reviews Gastroenterology & Hepatology 2022*, *19*(5), 313-327. https://doi.org/10.1038/s41575-021-00552-z

Marchini, J., Howie, B., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genetics*, *39*(7), 906–913. https://doi.org/10.1038/ng2088

Margaritte-Jeannin, P., Babron, M. C., Bourgey, M., Louka, A. S., Clot, F., Percopo, S., Coto, I., Hugot, J. P., Ascher, H., Sollid, L. M., Greco, L., & Clerget-Darpoux, F. (2004). HLA-DQ relative risks for coeliac disease in European populations: A study of the European Genetics Cluster on Coeliac Disease. *Tissue Antigens*, *63*(6), 562–567. https://doi.org/10.1111/j.0001-2815.2004.00237.x

Marild, K., Størdal, K., Hagman, A., & Ludvigsson, J. F. (2016). Turner Syndrome and Celiac Disease: A Case-Control Study. *Pediatrics*, *137*(2). https://doi.org/10.1542/PEDS.2015-2232

Marks, J., Shuster, S., & Watson, A. J. (1966). Small-bowel Changes in Dermatitis Herpetiformis. *The Lancet*, *288*(7476), 1280–1282. https://doi.org/10.1016/S0140-6736(66)91692-8

Marsh, M. N. (1992). Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology*, *102*(1), 330–354.https://doi.org/10.1016/0016-5085(92)91819-P

Martín-Masot, R., Nestares, M. T., Diaz-Castro, J., López-Aliaga, I., Alférez, M. J. M., Moreno-Fernandez, J., & Maldonado, J. (2019). Multifactorial Etiology of Anemia in Celiac Disease and Effect of Gluten-Free Diet: A Comprehensive Review. *Nutrients*, *11*(11). https://doi.org/10.3390/NU11112557

Martínez-Ojinaga, E., Fernández-Prieto, M., Molina, M., Polanco, I., Urcelay, E., & Núñez, C. (2019). Influence of HLA on clinical and analytical features of pediatric celiac disease. *BMC Gastroenterology*, *19*(1). https://doi.org/10.1186/s12876-019-1014-0

McAllister, K., Yarwood, A., Bowes, J., Orozco, G., Viatte, S., Diogo, D., Hocking, L. J., Steer, S., Wordsworth, P., Wilson, A. G., Morgan, A. W., Kremer, J. M., Pappas, D., Gregersen, P., Klareskog, L., Plenge, R., Barton, A., Greenberg, J., Worthington, J., & Eyre, S. (2013). Identification of BACH2 and RAD51B as rheumatoid arthritis susceptibility loci in a meta-analysis of genome-wide data. *Arthritis and Rheumatism*, *65*(12), 3058–3062. https://doi.org/10.1002/ART.38183

McGowan, K. E., Castiglione, D. A., & Butzner, J. D. (2009). The Changing Face of Childhood Celiac Disease in North America: Impact of Serological Testing. *Pediatrics*, *124*(6), 1572–1578. https://doi.org/10.1542/PEDS.2008-2373

McGowan, K. E., Lyon, M. E., & Butzner, J. D. (2008). Celiac Disease and IgA Deficiency: Complications of Serological Testing Approaches Encountered in the Clinic. *Clinical Chemistry*, *54*(7), 1203–1209. https://doi.org/10.1373/CLINCHEM.2008.103606

McHorney, C. A., Ware, J. E., Rachel Lu, J. F., & Sherbourne, C. D. (1994). The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, *32*(1), 40–66. https://doi.org/10.1097/00005650-199401000-00004

Meeuwisse, G. W. (1970). Round table discussion. Diagnostic criteria in coeliac disease. *Acta Paediatrica*, *59*(4), 461–464.

Megiorni, F., Mora, B., Bonamico, M., Barbato, M., Montuori, M., Viola, F., Trabace, S., & Mazzilli, M. C. (2008). HLA-DQ and susceptibility to celiac disease: evidence for gender differences and parent-of-origin effects. *The American Journal of Gastroenterology*, *103*(4), 997–1003. https://doi.org/10.1111/J.1572-0241.2007.01716.X

Megiorni, F., Mora, B., Bonamico, M., Barbato, M., Nenna, R., Maiella, G., Lulli, P., & Mazzilli, M. C. (2009). HLA-DQ and risk gradient for celiac disease. *Human Immunology*, *70*(1), 55–59. https://doi.org/10.1016/J.HUMIMM.2008.10.018

Mells, G. F., Floyd, J. A. B., Morley, K. I., Cordell, H. J., Franklin, C. S., Shin, S. Y., Heneghan, M. A., Neuberger, J. M., Donaldson, P. T., Day, D. B., Ducker, S. J., Muriithi, A. W., Wheater, E. F., Hammond, C. J., Dawwas, M. F., Jones, D. E., Peltonen, L., Alexander, G. J., Sandford, R. N., & Anderson, C. A. (2011). Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nature Genetics*, *43*(4), 329–332. https://doi.org/10.1038/NG.789

Mention, J. J., Ahmed, M. Ben, Bègue, B., Barbe, U., Verkarre, V., Asnafi, V., Colombel, J. F., Cugnenc, P. H., Ruemmele, F. M., McIntyre, E., Brousse, N., Cellier, C., & Cerf-Bensussan, N. (2003). Interleukin 15: a key to disrupted intraepithelial lymphocyte homeostasis and lymphomagenesis in celiac disease. *Gastroenterology*, *125*(3), 730–745. https://doi.org/10.1016/S0016-5085(03)01047-3

Molberg, Ø., Mcadam, S. N., Körner, R., Quarsten, H., Kristiansen, C., Madsen, L., Fugger, L., Scott, H., Norën, O., Roepstorff, P., Lundin, K. E. A., Sjöström, H., & Sollid, L. M. (1998). Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nature Medicine*, *4*(6), 713–717. https://doi.org/10.1038/NM0698-713

Momozawa, Y., Dmitrieva, J., Théâtre, E., Deffontaine, V., Rahmouni, S., Charloteaux, B., Crins, F., Docampo, E., Elansary, M., Gori, A. S., Lecut, C., Mariman, R., Mni, M., Oury, C., Altukhov, I., Alexeev, D., Aulchenko, Y., Amininejad, L., Bouma, G., … Zhao, Z. Z. (2018). IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04365-8

Monsuur, A. J., Bakker, P. I. W. D., Alizadeh, B. Z., Zhernakova, A., Bevova, M. R., Strengman, E., Franke, L., Slot, R. V. t., Belzen, M. J. V., Lavrijsen, I. C. M., Diosdado, B., Daly, M. J., Mulder, C. J. J., Mearin, M. L., Meijer, J. W. R., Meijer, G. A., Oort, E. Van, Wapenaar, M. C., Koeleman, B. P. C., & Wijmenga, C. (2005). Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. *Nature Genetics*, *37*(12), 1341–1344. https://doi.org/10.1038/NG1680

Monsuur, A. J., de Bakker, P. I. W., Zhernakova, A., Pinto, D., Verduijn, W., Romanos, J., Auricchio, R., Lopez, A., van Heel, D. A., Crusius, J. B. A., & Wijmenga, C. (2008). Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. *PloS One*, *3*(5), e2270. https://doi.org/10.1371/journal.pone.0002270

Mubarak, A., Spierings, E., Wolters, V., Van Hoogstraten, I., Kneepkens, C. M. F., & Houwen, R. (2013). Human leukocyte antigen DQ2.2 and celiac disease. *Journal of Pediatric Gastroenterology and Nutrition*, *56*(4), 428–430. https://doi.org/10.1097/MPG.0B013E31827913F9

Murray, J. A., Watson, T., Clearman, B., & Mitros, F. (2004). Effect of a gluten-free diet on gastrointestinal symptoms in celiac disease. *The American Journal of Clinical Nutrition*, *79*(4), 669–673. https://doi.org/10.1093/AJCN/79.4.669

Mäki, M., Hällstrom, O., Huupponen, T., Vesikari, T., & Visakorpi, J. K. (1984). Increased prevalence of coeliac disease in diabetes. *Archives of Disease in Childhood*, *59*(8), 739–742. https://doi.org/10.1136/ADC.59.8.739

Mäki, M., Holm, K., Hallstrom, O., Collin, P., Viander, M., Savilahti, E., Lipsanen, V., & Koskimies, S. (1991). Serological markers and HLA genes among healthy first-degree relatives of patients with coeliac disease. *Lancet*, *338*(8779), 1350–1353. https://doi.org/10.1016/0140-6736(91)92234-S

Mäki, Markku, Hällström, O., Vesikari, T., & Visakorpi, J. K. (1984). Evaluation of a serum IgA-class reticulin antibody test for the detection of childhood celiac disease. *The Journal of Pediatrics*, *105*(6), 901–905. https://doi.org/10.1016/S0022-3476(84)80074-8

Nakamura, M., Nishida, N., Kawashima, M., Aiba, Y., Tanaka, A., Yasunami, M., Nakamura, H., Komori, A., Nakamuta, M., Zeniya, M., Hashimoto, E., Ohira, H., Yamamoto, K., Onji, M., Kaneko, S., Honda, M., Yamagiwa, S., Nakao, K., Ichida, T., … Ishibashi, H. (2012). Genome-wide Association Study Identifies TNFSF15 and POU2AF1 as Susceptibility Loci for Primary Biliary Cirrhosis in the Japanese Population. *American Journal of Human Genetics*, *91*(4), 721. https://doi.org/10.1016/J.AJHG.2012.08.010

Namatovu, F., Sandström, O., Olsson, C., Lindkvist, M., & Ivarsson, A. (2014). Celiac disease risk varies between birth cohorts, generating hypotheses about causality: evidence from 36 years of population-based follow-up. *BMC Gastroenterology*, *14*(1). https://doi.org/10.1186/1471-230X-14-59

Nellikkal, S. S., Hafed, Y., Larson, J. J., Murray, J. A., & Absah, I. (2019). High Prevalence of Celiac Disease Among Screened First-Degree Relatives. *Mayo Clinic Proceedings*, *94*(9), 1807–1813. https://doi.org/10.1016/J.MAYOCP.2019.03.027

Newnham, E. D., Shepherd, S. J., Strauss, B. J., Hosking, P., & Gibson, P. R. (2016). Adherence to the gluten-free diet can achieve the therapeutic goals in almost all patients with coeliac disease: A 5-year longitudinal study from diagnosis. *Journal of Gastroenterology and Hepatology*, *31*(2), 342–349. https://doi.org/10.1111/JGH.13060

Nikopensius, T., Niibo, P., Haller, T., Jagomägi, T., Voog-Oras, Ü., Tõnisson, N., Metspalu, A., Saag, M., & Pruunsild, C. (2021). Association analysis of juvenile idiopathic arthritis genetic susceptibility factors in Estonian patients. *Clinical Rheumatology*, *40*(10), 4157–4165. https://doi.org/10.1007/S10067-021-05756-X

Norström, F., Lindholm, L., Sandström, O., Nordyke, K., & Ivarsson, A. (2011). Delay to celiac disease diagnosis and its implications for health-related quality of life. *BMC Gastroenterology*, *11*, 118. https://doi.org/10.1186/1471-230X-11-118

Oberhuber, G., Granditsch, G., & Vogelsang, H. (1999). The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *European Journal of Gastroenterology & Hepatology*, *11*(10), 1185–1194. https://doi.org/10.1097/00042737-199910000-00019

Olaussen, R. W., Karlsson, M. R., Lundin, K. E. A., Jahnsen, J., Brandtzaeg, P., & Farstad, I. N. (2007). Reduced Chemokine Receptor 9 on Intraepithelial Lymphocytes in Celiac Disease Suggests Persistent Epithelial Activation. *Gastroenterology*, *132*(7), 2371–2382. https://doi.org/10.1053/j.gastro.2007.04.023

Olivares-Villagómez, D., & Van Kaer, L. (2018). Intestinal Intraepithelial Lymphocytes: Sentinels of the Mucosal Barrier. *Trends in Immunology*, *39*(4), 264–275. https://doi.org/10.1016/J.IT.2017.11.003

Olivares, M., Neef, A., Castillejo, G., De Palma, G., Varea, V., Capilla, A., Palau, F., Nova, E., Marcos, A., Polanco, I., Ribes-Koninckx, C., Ortigosa, L., Izquierdo, L., & Sanz, Y. (2015). The HLA-DQ2 genotype selects for early intestinal microbiota composition in infants at high risk of developing coeliac disease. *Gut*, *64*(3), 406–417. https://doi.org/10.1136/GUTJNL-2014-306931

Orozco, G., Eyre, S., Hinks, A., Bowes, J., Morgan, A. W., Wilson, A. G., Wordsworth, P., Steer, S., Hocking, L., Thomson, W., Worthington, J., Barton, A., Gibbons, L. J., Flynn, E., Martin, P., Ke, X., Donn, R., Martin, S., Robinson, J. I., … Harrison, P. (2011). Study of the common genetic background for rheumatoid arthritis and systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, *70*(3), 463–468. https://doi.org/10.1136/ARD.2010.137174

Paarlahti, P., Kurppa, K., Ukkola, A., Collin, P., Huhtala, H., Mäki, M., & Kaukinen, K. (2013). Predictors of persistent symptoms and reduced quality of life in treated coeliac disease patients: A large cross-sectional study. *BMC Gastroenterology*, *13*(1). https://doi.org/10.1186/1471-230X-13-75

Papadakis, K. A., Prehn, J., Nelson, V., Cheng, L., Binder, S. W., Ponath, P. D., Andrew, D. P., & Targan, S. R. (2000). The Role of Thymus-Expressed Chemokine and Its Receptor CCR9 on Lymphocytes in the Regional Specialization of the Mucosal Immune System. *The Journal of Immunology*, *165*(9), 5069–5076. https://doi.org/10.4049/jimmunol.165.9.5069

Pathak, M., & Lal, G. (2020). The Regulatory Function of CCR9+ Dendritic Cells in Inflammation and Autoimmunity. *Frontiers in Immunology, 11*. https://doi.org/10.3389/fimmu.2020.536326

Patsopoulos, N. A., Esposito, F., Reischl, J., Lehr, S., Bauer, D., Heubach, J., Sandbrink, R., Pohl, C., Edan, G., Kappos, L., Miller, D., Montalbán, J., Polman, C. H., Freedman, M. S., Hartung, H. P., Arnason, B. G. W., Comi, G., Cook, S., Filippi, M., … de Bakker, P. I. W. (2011). Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Annals of Neurology*, *70*(6), 897–912. https://doi.org/10.1002/ANA.22609

Pawlak-Adamska, E., Frydecka, I., Bolanowski, M., Tomkiewicz, A., Jonkisz, A., Karabon, L., Partyka, A., Nowak, O., Szalinski, M., & Daroszewski, J. (2017). CD28/CTLA-4/ICOS haplotypes confers susceptibility to Graves' disease and modulates clinical phenotype of disease. *Endocrine*, *55*(1), 186–199. https://doi.org/10.1007/S12020-016-1096-1

Pedretti, M., Sbravati, F., Allegri, D., Labriola, F., Lombardo, V., Spisni, E., Zarbo, C., & Alvisi, P. (2021). Is the clinical pattern of pediatric celiac disease changing? A thirty-years real-life experience of an Italian center. *Italian Journal of Pediatrics*, *47*(1). https://doi.org/10.1186/S13052-021-01183-5

Pekki, H., Kurppa, K., Mäki, M., Huhtala, H., Laurila, K., Ilus, T., & Kaukinen, K. (2017). Performing routine follow-up biopsy 1 year after diagnosis does not affect long-term outcomes in coeliac disease. *Alimentary Pharmacology & Therapeutics*, *45*(11), 1459–1468. https://doi.org/10.1111/APT.14048

Pekki, Henna, Kurppa, K., Mäki, M., Huhtala, H., Sievänen, H., Laurila, K., Collin, P., & Kaukinen, K. (2015). Predictors and significance of incomplete mucosal recovery in celiac disease after 1 year on a gluten-free diet. *American Journal of Gastroenterology*, *110*(7), 1078–1085. https://doi.org/10.1038/ajg.2015.155

Pinto-Sánchez, M. I., Causada-Calo, N., Bercik, P., Ford, A. C., Murray, J. A., Armstrong, D., Semrad, C., Kupfer, S. S., Alaedini, A., Moayyedi, P., Leffler, D. A., Verdú, E. F., & Green, P. (2017). Safety of Adding Oats to a Gluten-Free Diet for Patients With Celiac Disease: Systematic Review and Meta-analysis of Clinical and Observational Studies. *Gastroenterology*, *153*(2), 395-409.e3. https://doi.org/10.1053/J.GASTRO.2017.04.009

Plenge, R. M., Cotsapas, C., Davies, L., Price, A. L., De Bakker, P. I. W., Maller, J., Pe'er, I., Burtt, N. P., Blumenstiel, B., DeFelice, M., Parkin, M., Barry, R., Winslow, W., Healy, C., Graham, R. R., Neale, B. M., Izmailova, E., Roubenoff, R., Parker, A. N., … Altshuler, D. (2007). Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nature Genetics*, *39*(12), 1477–1482. https://doi.org/10.1038/NG.2007.27

Plevinsky, J. M., Greenley, R. N., & Fishman, L. N. (2016). Self-management in patients with inflammatory bowel disease: Strategies, outcomes, and integration into clinical care. *Clinical and Experimental Gastroenterology*, *9*, 259–267. https://doi.org/10.2147/CEG.S106302

Ploski, R., Ek, J., Thorsby, E., & Sollid, L. M. (1993). On the HLA-DQ(α1*0501, β1*0201)-associated susceptibility in celiac disease: A possible gene dosage effect of *DQB1*0201*. *Tissue Antigens*, *41*(4), 173–177. https://doi.org/10.1111/j.1399-0039.1993.tb01998.x

Popp, A., Taavela, J., Graziano, P., Parente, P., Covelli, C., Lamacchia, C., Andriulli, A., Mäki, M., & Isola, J. (2021). A New Intraepithelial γδ T-Lymphocyte Marker for Celiac Disease Classification in Formalin-Fixed Paraffin-Embedded (FFPE) Duodenal Biopsies. *Digestive Diseases and Sciences*, *66*(10), 3352–3358. https://doi.org/10.1007/S10620-020-06680-X

Prahalad, S., Hansen, S., Whiting, A., Guthery, S. L., Clifford, B., McNally, B., Zeft, A. S., Bohnsack, J. F., & Jorde, L. B. (2009). Variants in TNFAIP3, STAT4, and C12orf30 loci associated with multiple autoimmune diseases are also associated with juvenile idiopathic arthritis. *Arthritis and Rheumatism*, *60*(7), 2124–2130. https://doi.org/10.1002/ART.24618

Price, P., Witt, C., Allcock, R., Sayer, D., Garlepp, M., Kok, C. C., French, M., Mallal, S., & Christiansen, F. (1999). The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunological Reviews*, *167*, 257–274. https://doi.org/10.1111/J.1600-065X.1999.TB01398.X

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Qiu, F., Tang, R., Zuo, X., Shi, X., Wei, Y., Zheng, X., Dai, Y., Gong, Y., Wang, L., Xu, P., Zhu, X., Wu, J., Han, C., Gao, Y., Zhang, K., Jiang, Y., Zhou, J., Shao, Y., Hu, Z., … Ma, X. (2017). A genome-wide association study identifies six novel risk loci for primary biliary cholangitis. *Nature Communications*, *8*(1). https://doi.org/10.1038/NCOMMS14828

Rampertab, S. D., Pooran, N., Brar, P., Singh, P., & Green, P. H. R. (2006). Trends in the Presentation of Celiac Disease. *The American Journal of Medicine*, *119*(4), 355.e9-355.e14. https://doi.org/10.1016/j.amjmed.2005.08.044

Rauhavirta, T., Hietikko, M., Salmi, T., & Lindfors, K. (2019). Transglutaminase 2 and Transglutaminase 2 Autoantibodies in Celiac Disease: a Review. *Clinical Reviews in Allergy & Immunology*, *57*(1), 23–38. https://doi.org/10.1007/S12016-016-8557-4

Raychaudhuri, S., Thomson, B. P., Remmers, E. F., Eyre, S., Hinks, A., Guiducci, C., Catanese, J. J., Xie, G., Stahl, E. A., Chen, R., Alfredsson, L., Amos, C. I., Ardlie, K. G., Barton, A., Bowes, J., Burtt, N. P., Chang, M., Coblyn, J., Costenbader, K. H., … Plenge, R. M. (2009). Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nature Genetics*, *41*(12), 1313–1318. https://doi.org/10.1038/NG.479

Reale, A., Trevisan, M., Alvisi, G., & Palù, G. (2018). The silent enemy: Celiac disease goes viral. *Journal of Cellular Physiology*, *233*(4), 2693–2694. https://doi.org/10.1002/JCP.26155

Reunala, T. L. (2001). Dermatitis herpetiformis. *Clinics in Dermatology*, *19*(6), 728–736. https://doi.org/10.1016/S0738-081X(00)00184-X

Rivas, M. A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C. K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N., Fennell, T., Kirby, A., Latiano, A., Goyette, P., Green, T., Halfvarson, J., Haritunians, T., Korn, J. M., Kuruvilla, F., … Daly, M. J. (2011). Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nature Genetics*, *43*(11), 1066. https://doi.org/10.1038/NG.952

Rosén, A., Sandström, O., Carlsson, A., Högberg, L., Olén, O., Stenlund, H., & Ivarsson, A. (2014). Usefulness of Symptoms to Screen for Celiac Disease. *Pediatrics*, *133*(2), 211–218. https://doi.org/10.1542/PEDS.2012-3765

Roy, A., Laszkowska, M., Sundström, J., Lebwohl, B., Green, P. H. R., Kämpe, O., & Ludvigsson, J. F. (2016). Prevalence of Celiac Disease in Patients with Autoimmune Thyroid Disease: A Meta-Analysis. *Thyroid : Official Journal of the American Thyroid Association*, *26*(7), 880–890. https://doi.org/10.1089/THY.2016.0108

Rubio-Tapia, A., Hill, I. D., Kelly, C. P., Calderwood, A. H., & Murray, J. A. (2013). ACG clinical guidelines: diagnosis and management of celiac disease. *The American Journal of Gastroenterology*, *108*(5), 656–676. https://doi.org/10.1038/AJG.2013.79

Rubio-Tapia, A., Kyle, R. A., Kaplan, E. L., Johnson, D. R., Page, W., Erdtmann, F., Brantner, T. L., Kim, W. R., Phelps, T. K., Lahr, B. D., Zinsmeister, A. R., Melton, L. J., & Murray, J. A. (2009). Increased prevalence and mortality in undiagnosed celiac disease. *Gastroenterology*, *137*(1), 88–93. https://doi.org/10.1053/J.GASTRO.2009.03.059

Rubio-Tapia, A., Rahim, M. W., See, J. A., Lahr, B. D., Wu, T. T., & Murray, J. A. (2010). Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *American Journal of Gastroenterology*, *105*(6), 1412–1420. https://doi.org/10.1038/ajg.2010.10

Salmi, T. T., Collin, P., Korponay-Szabó, I. R., Laurila, K., Partanen, J., Huhtala, H., Király, R., Lorand, L., Reunala, T., Mäki, M., & Kaukinen, K. (2006). Endomysial antibody-negative coeliac disease: clinical characteristics and intestinal autoantibody deposits. *Gut*, *55*(12), 1746–1753. https://doi.org/10.1136/GUT.2005.071514

Salvati, V. M., MacDonald, T. T., Bajaj-Elliott, M., Borrelli, M., Staiano, A., Auricchio, S., Troncone, R., & Monteleone, G. (2002). Interleukin 18 and associated markers of T helper cell type 1 activity in coeliac disease. *Gut*, *50*(2), 186–190. https://doi.org/10.1136/GUT.50.2.186

Savilahti, E., Pelkonen, P., & Visakorpi, J. K. (1971). IgA deficiency in children. A clinical study with special reference to intestinal findings. *Archives of Disease in Childhood*, *46*(249), 665–670. https://doi.org/10.1136/ADC.46.249.665

Schuppan, D., Mäki, M., Lundin, K. E. A., Isola, J., Friesing-Sosnik, T., Taavela, J., Popp, A., Koskenpato, J., Langhorst, J., Hovde, Ø., Lähdeaho, M.-L., Fusco, S., Schumann, M., Török, H. P., Kupcinskas, J., Zopf, Y., Lohse, A. W., Scheinin, M., Kull, K., … Greinwald, R. (2021). A Randomized Trial of a Transglutaminase 2 Inhibitor for Celiac Disease. *The New England Journal of Medicine*, *385*(1), 35–45. https://doi.org/10.1056/NEJMOA2032441

Schyum, A. C., & Rumessen, J. J. (2013). Serological testing for celiac disease in adults. *United European Gastroenterology Journal*, *1*(5), 319–325. https://doi.org/10.1177/2050640613502788

Seah, P. P., Fry, L., Holborow, E. J., Rossiter, M. A., Doe, W. F., Magalhaes, A. F., & Hoffbrand, A. V. (1973). Antireticulin antibody: incidence and diagnostic significance. *Gut*, *14*(4), 311–315. https://doi.org/10.1136/GUT.14.4.311

Seah, P. P., Fry, L., Rossiter, M. A., Hopfbrand, A. V., & Holborow, E. J. (1971). Anti-reticulin antibodies in childhood coeliac disease. *Lancet*, *2*(7726), 681–682. https://doi.org/10.1016/S0140-6736(71)92248-3

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., Sollid, L. M., & Khosla, C. (2002). Structural basis for gluten intolerance in celiac sprue. *Science*, *297*(5590), 2275–2279. https://doi.org/10.1126/SCIENCE.1074129

Shan, S., Dang, J., Li, J., Yang, Z., Zhao, H., Xin, Q., Ma, X., Liu, Y., Bian, X., Gong, Y., & Liu, Q. (2014). ETS1 variants confer susceptibility to ankylosing spondylitis in Han Chinese. *Arthritis Research & Therapy*, *16*(2). https://doi.org/10.1186/AR4530

Shancui-Zheng, Jinping-Zhang, Guoyuan-Lu, Liu, L., & Zhiyong-Deng. (2022). Polymorphism in STAT4 Increase the Risk of Systemic Lupus Erythematosus: An Updated Meta-Analysis. *International Journal of Rheumatology*, *2022*. https://doi.org/10.1155/2022/5565057

Sharkey, L. M., Corbett, G., Currie, E., Lee, J., Sweeney, N., & Woodward, J. M. (2013). Optimising delivery of care in coeliac disease – comparison of the benefits of repeat biopsy and serological follow-up. *Alimentary Pharmacology & Therapeutics*, *38*(10), 1278–1291. https://doi.org/10.1111/APT.12510

Sharma, A., Liu, X., Hadley, D., Hagopian, W., Liu, E., Chen, W.-M., Onengut-Gumuscu, S., Simell, V., Rewers, M., Ziegler, A.-G., Lernmark, A., Simell, O., Toppari, J., Krischer, J. P., Akolkar, B., Rich, S. S., Agardh, D., & She, J.-X. (2016). Identification of Non-HLA Genes Associated with Celiac Disease and Country-Specific Differences in a Large, International Pediatric Cohort. *PloS One*, *11*(3), e0152476. https://doi.org/10.1371/journal.pone.0152476

Shen, N., Fu, Q., Deng, Y., Qian, X., Zhao, J., Kaufman, K. M., Wu, Y. L., Yu, C. Y., Tang, Y., Chen, J. Y., Yang, W., Wong, M., Kawasaki, A., Tsuchiya, N., Sumida, T., Kawaguchi, Y., Howe, H. S., Mok, M. Y., Bang, S. Y., … Tsao, B. P. (2010). Sex-specific association of X-linked Toll-like receptor 7 (TLR7) with male systemic lupus erythematosus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(36), 15838–15843. https://doi.org/10.1073/PNAS.1001337107

Sheridan, B. S., & Lefrançois, L. (2010). Intraepithelial lymphocytes: to serve and protect.*Current Gastroenterology Reports*, *12*(6), 513–521. https://doi.org/10.1007/S11894-010-0148-6

Shiner, M. (1956). Duodenal biopsy. *Lancet*, *270*(6906), 17–19. https://doi.org/10.1016/S0140-6736(56)91854-2

Singh, P., Arora, A., Strand, T. A., Leffler, D. A., Catassi, C., Green, P. H., Kelly, C. P., Ahuja, V., & Makharia, G. K. (2018). Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. *Clinical Gastroenterology and Hepatology*, *16*(6), 823-836.e2. https://doi.org/10.1016/J.CGH.2017.06.037

Singh, P., Arora, S., Lal, S., Strand, T. A., & Makharia, G. K. (2015). Risk of celiac disease in the first- and second-degree relatives of patients with celiac disease: A systematic review and meta-analysis. In *American Journal of Gastroenterology*, *110*(11), 1539–1548. https://doi.org/10.1038/ajg.2015.296

Singh, P., Arora, S., & Makharia, G. K. (2014). Presence of anemia in patients with celiac disease suggests more severe disease. *Indian Journal of Gastroenterology : Official Journal of the Indian Society of Gastroenterology*, *33*(2), 161–164. https://doi.org/10.1007/S12664-013-0423-1

Smyth, D. J., Plagnol, V., Walker, N. M., Cooper, J. D., Downes, K., Yang, J. H. M., Howson, J. M. M., Stevens, H., McManus, R., Wijmenga, C., Heap, G. A., Dubois, P. C., Clayton, D. G., Hunt, K. A., van Heel, D. A., & Todd, J. A. (2008). Shared and distinct genetic variants in type 1 diabetes and celiac disease. *The New England Journal of Medicine*, *359*(26), 2767–2777. https://doi.org/10.1056/NEJMOA0807917

Sollid, L. M. (2017). The roles of MHC class II genes and post-translational modification in celiac disease. *Immunogenetics, 69*(8–9), 605–616. https://doi.org/10.1007/s00251-017-0985-7

Sollid, L. M., Markussen, G., Ek, J., Gjerde, H., Vartdal, F., & Thorsby, E. (1989). Evidence for a primary association of celiac disease to a particular HLA-DQ a/ß heterodimer. *Materials and Methods Brief Definitive Report*. *169*, 345–350. https://doi.org/10.1084/jem.169.1.345

Song, C., Zhang, J., Zhao, Z., Yang, Y., Meng, D., Wang, J., Guo, C., & Yuan, C. (2020). DLEU1: A Functional Long Noncoding RNA in Tumorigenesis. *Current Pharmaceutical Design*, *26*(15), 1742–1748. https://doi.org/10.2174/1381612826666200122145305

Spijkerman, M., Tan, I. L., Kolkman, J. J., Withoff, S., Wijmenga, C., Visschedijk, M. C., & Weersma, R. K. (2016). A large variety of clinical features and concomitant disorders in celiac disease - A cohort study in the Netherlands. *Digestive and Liver Disease : Official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, *48*(5), 499–505. https://doi.org/10.1016/J.DLD.2016.01.006

Størdal, K., Kahrs, C., Tapia, G., Agardh, D., Kurppa, K., & Stene, L. C. (2021). Review article: exposure to microbes and risk of coeliac disease. *Alimentary Pharmacology and Therapeutics, 53*(1), 43–62. https://doi.org/10.1111/apt.16161

Sulkanen, S., Collin, P., Laurila, K., & Mäki, M. (1998). IgA- and IgG-class antihuman umbilical cord antibody tests in adult coeliac disease. *Scandinavian Journal of Gastroenterology*, *33*(3), 251–254. https://doi.org/10.1080/00365529850170810

Svedlund, J., Sjödin, I., & Dotevall, G. (1988). GSRS-A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Digestive Diseases and Sciences*, *33*(2), 129–134. https://doi.org/10.1007/BF01535722

Taavela, J., Koskinen, O., Huhtala, H., Lähdeaho, M.-L., Popp, A., Laurila, K., Collin, P., Kaukinen, K., Kurppa, K., & Mäki, M. (2013). Validation of Morphometric Analyses of Small-Intestinal Biopsy Readouts in Celiac Disease. *PLoS ONE*, *8*(10), e76163. https://doi.org/10.1371/journal.pone.0076163

Taavela, J., Viiri, K., Popp, A., Oittinen, M., Dotsenko, V., Peraäho, M., Staff, S., Sarin, J., Leon, F., Mäki, M., & Isola, J. (2019). Histological, immunohistochemical and mRNA gene expression responses in coeliac disease patients challenged with gluten using PAXgene fixed paraffin-embedded duodenal biopsies. *BMC Gastroenterology*, *19*(1). https://doi.org/10.1186/S12876-019-1089-7

Tapsas, D., Hollén, E., Stenhammar, L., & Fälth-Magnusson, K. (2016). The clinical presentation of coeliac disease in 1030 Swedish children: Changing features over the past four decades. *Digestive and Liver Disease : Official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, *48*(1), 16–22. https://doi.org/10.1016/J.DLD.2015.09.018

Tennyson, C. A., Simpson, S., Lebwohl, B., Lewis, S., & Green, P. H. R. (2013). Interest in medical therapy for celiac disease. *Therapeutic Advances in Gastroenterology*, *6*(5), 358–364. https://doi.org/10.1177/1756283X13492580

Teruel, M., & Alarcón-Riquelme, M. E. (2016). Genetics of systemic lupus erythematosus and Sjögren's syndrome: An update. *Current Opinion in Rheumatology*, *28*(5), 506–514. https://doi.org/10.1097/BOR.0000000000000310

Thompson, S. D., Sudman, M., Ramos, P. S., Marion, M. C., Ryan, M., Tsoras, M., Weiler, T., Wagner, M., Keddache, M., Haas, J. P., Mueller, C., Prahalad, S., Bohnsack, J., Wise, C. A., Punaro, M., Zhang, D., Rosé, C. D., Comeau, M. E., Divers, J., … Langefeld, C. D. (2010). The susceptibility loci juvenile idiopathic arthritis shares with other autoimmune diseases extend to PTPN2, COG6, and ANGPT1. *Arthritis and Rheumatism*, *62*(11), 3265–3276. https://doi.org/10.1002/ART.27688

Tio, M., Cox, M. R., & Eslick, G. D. (2012). Meta-analysis: coeliac disease and the risk of all-cause mortality, any malignancy and lymphoid malignancy. *Alimentary Pharmacology & Therapeutics*, *35*(5), 540–551. https://doi.org/10.1111/J.1365-2036.2011.04972.X

Tizaoui, K., Kim, S. H., Jeong, G. H., Kronbichler, A., Lee, K. S., Lee, K. H., & Shin, J. Il. (2019). Association of PTPN22 1858C/T Polymorphism with Autoimmune Diseases: A Systematic Review and Bayesian Approach. *Journal of Clinical Medicine*, *8*(3). https://doi.org/10.3390/JCM8030347

Traks, T., Keermann, M., Karelson, M., Rätsep, R., Reimann, E., Silm, H., Vasar, E., Kõks, S., & Kingo, K. (2015). Polymorphisms in Toll-like receptor genes are associated with vitiligo. *Frontiers in Genetics*, *6*, 278. https://doi.org/10.3389/FGENE.2015.00278

Trynka, G., Hunt, K. A., Bockett, N. A., Romanos, J., Mistry, V., Szperl, A., Bakker, S. F., Bardella, M. T., Bhaw-Rosun, L., Castillejo, G., de la Concha, E. G., de Almeida, R. C., Dias, K.-R. M., van Diemen, C. C., Dubois, P. C. A., Duerr, R. H., Edkins, S., Franke, L., Fransen, K., … Heel, D. A. van. (2011). Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nature Genetics*, *43*(12), 1193–1201. https://doi.org/10.1038/ng.998

Tsoi, L. C., Spain, S. L., Knight, J., Ellinghaus, E., Stuart, P. E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J. E., Kang, H. M., Allen, M. H., McManus, R., Novelli, G., Samuelsson, L., Schalkwijk, J., Ståhle, M., Burden, A. D., Smith, C. H., … Trembath, R. C. (2012). Identification of fifteen new psoriasis susceptibility loci highlights the role of innate immunity. *Nature Genetics*, *44*(12), 1341. https://doi.org/10.1038/NG.2467

Tursi, A., Brandimarte, G., & Giorgetti, G. M. (2003). High prevalence of small intestinal bacterial overgrowth in celiac patients with persistence of gastrointestinal symptoms after gluten withdrawal. *American Journal of Gastroenterology*, *98*(4), 839–843. https://doi.org/10.1111/j.1572-0241.2003.07379.x

Uche-Anya, E., & Lebwohl, B. (2021). Celiac disease: Clinical update. *Current Opinion in Gastroenterology*, *37*(6), 619–624. https://doi.org/10.1097/MOG.0000000000000785

Ukkola, A., Mäki, M., Kurppa, K., Collin, P., Huhtala, H., Kekkonen, L., & Kaukinen, K. (2011). Diet Improves Perception of Health and Well-being in Symptomatic, but Not Asymptomatic, Patients With Celiac Disease. *Clinical Gastroenterology and Hepatology*, *9*(2), 118-123.e1. https://doi.org/10.1016/J.CGH.2010.10.011

Ukkola, A., Mäki, M., Kurppa, K., Collin, P., Huhtala, H., Kekkonen, L., & Kaukinen, K. (2012). Patients' experiences and perceptions of living with coeliac disease - implications for optimizing care. *Journal of Gastrointestinal and Liver Diseases : JGLD*, *21*(1), 17–22. https://pubmed.ncbi.nlm.nih.gov/22457855/

Valitutti, F., Trovato, C. M., Montuori, M., & Cucchiara, S. (2017). Pediatric Celiac Disease: Follow-Up in the Spotlight. *Advances in Nutrition*, *8*(2), 356. https://doi.org/10.3945/AN.116.013292

van Belzen, M. J., Meijer, J. W. R., Sandkuijl, L. A., Bardoel, A. F. J., Mulder, C. J. J., Pearson, P. L., Houwen, R. H. J., & Wijmenga, C. (2003). A major non-HLA locus in celiac disease maps to chromosome 19. *Gastroenterology*, *125*(4), 1032–1041. https://doi.org/10.1016/S0016-5085(03)01205-8

van Berge-Henegouwen, G. P., & Mulder, C. J. J. (1993). Pioneer in the gluten free diet: Willem-Karel Dicke 1905-1962, over 50 years of gluten free diet. *Gut*, *34*(11), 1473. https://doi.org/10.1136/GUT.34.11.1473

van der Graaf, A., Zorro, M. M., Claringbould, A., Võsa, U., Aguirre-Gamboa, R., Li, C., Mooiweer, J., Ricaño-Ponce, I., Borek, Z., Koning, F., Kooy-Winkelaar, Y., Sollid, L. M., Qiao, S. W., Kumar, V., Li, Y., Franke, L., Withoff, S., Wijmenga, C., Sanna, S., & Jonkers, I. (2021). Systematic Prioritization of Candidate Genes in Disease Loci Identifies TRAFD1 as a Master Regulator of IFNγ Signaling in Celiac Disease. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/FGENE.2020.562434

van Heel, D. A., Franke, L., Hunt, K. A., Gwilliam, R., Zhernakova, A., Inouye, M., Wapenaar, M. C., Barnardo, M. C. N. M., Bethel, G., Holmes, G. K. T., Feighery, C., Jewell, D., Kelleher, D., Kumar, P., Travis, S., Walters, J. R. F., Sanders, D. S., Howdle, P., Swift, J., … Wijmenga, C. (2007). A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nature Genetics 2007 39:7*, *39*(7), 827–829. https://doi.org/10.1038/ng2058

van Heel, D. A., Hunt, K., Greco, L., & Wijmenga, C. (2005). Genetics in coeliac disease. *Best Practice & Research Clinical Gastroenterology*, *19*(3), 323–339. https://doi.org/10.1016/J.BPG.2005.01.001

Verdu, E. F., & Caminero, A. (2017). How infection can incite sensitivity to food: Microbes can trigger chronic immune disorders to dietary antigens, such as in celiac disease. *Science*, *356*(6333), 29–30. https://doi.org/10.1126/SCIENCE.AAN1500/ASSET/5ADE8C30-D709-49C9-B4A0-FF181380BC87/ASSETS/GRAPHIC/356_29_F1.JPEG

Verdu, E. F., & Schuppan, D. (2021). Co-factors, Microbes, and Immunogenetics in Celiac Disease to Guide Novel Approaches for Diagnosis and Treatment. *Gastroenterology*, *161*(5), 1395-1411.e4. https://doi.org/10.1053/J.GASTRO.2021.08.016

Viljamaa, M., Kaukinen, K., Huhtala, H., Kyrönpalo, S., Rasmussen, M., & Collin, P. (2009). Coeliac Disease, autoimmune diseases and gluten exposure. *Scandinavian Journal of Gastroenterology*, *40*(4), 437–443. https://doi.org/10.1080/00365520510012181

Virta, J., Hannula, M., Tamminen, I., Lindfors, K., Kaukinen, K., Popp, A., Taavela, J., Saavalainen, P., Hiltunen, P., Hyttinen, J., & Kurppa, K. (2020). X-ray microtomography is a novel method for accurate evaluation of small-bowel mucosal morphology and surface area. *Scientific Reports*, *10*(1). https://doi.org/10.1038/S41598-020-69487-W

Virta, L. J., Saarinen, M. M., & Kolho, K. L. (2017). Declining trend in the incidence of biopsy-verified coeliac disease in the adult population of Finland, 2005-2014. *Alimentary Pharmacology & Therapeutics*, *46*(11–12), 1085–1093. https://doi.org/10.1111/APT.14335

Visakorpi, J. K., & Mäki, M. (1994). Changing clinical features of coeliac disease. *Acta Paediatrica Supplement*, *83*(395), 10–13. https://doi.org/10.1111/J.1651-2227.1994.TB13221.X

Volta, U., Caio, G., Giancola, F., Rhoden, K. J., Ruggeri, E., Boschetti, E., Stanghellini, V., & De Giorgio, R. (2016). Features and Progression of Potential Celiac Disease in Adults. *Clinical Gastroenterology and Hepatology : The Official Clinical Practice Journal of the American Gastroenterological Association*, *14*(5), 686-693.e1. https://doi.org/10.1016/J.CGH.2015.10.024

Volta, U., Caio, G., Stanghellini, V., & De Giorgio, R. (2014). The changing clinical profile of celiac disease: a 15-year experience (1998-2012) in an Italian referral center. *BMC Gastroenterology*, *14*(1). https://doi.org/10.1186/S12876-014-0194-X

Wacklin, P., Laurikka, P., Lindfors, K., Collin, P., Salmi, T., Lähdeaho, M. L., Saavalainen, P., Mäki, M., Mättö, J., Kurppa, K., & Kaukinen, K. (2014). Altered Duodenal microbiota composition in celiac disease patients suffering from persistent symptoms on a long-term gluten-free diet. *American Journal of Gastroenterology*, *109*(12), 1933–1941. https://doi.org/10.1038/ajg.2014.355

Walker, M. M., Murray, J. A., Ronkainen, J., Aro, P., Storskrubb, T., D'Amato, M., Lahr, B., Talley, N. J., & Agreus, L. (2010). Detection of celiac disease and lymphocytic enteropathy by parallel serology and histopathology in a population-based study. *Gastroenterology*, *139*(1), 112–119. https://doi.org/10.1053/J.GASTRO.2010.04.007

Wang, K., Baldassano, R., Zhang, H., Qu, H. Q., Imielinski, M., Kugathasan, S., Annese, V., Dubinsky, M., Rotter, J. I., Russell, R. K., Bradfield, J. P., Sleiman, P. M. A., Glessner, J. T., Walters, T., Hou, C., Kim, C., Frackelton, E. C., Garris, M., Doran, J., … Hakonarson, H. (2010). Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes implicates multiple loci with opposite effects. *Human Molecular Genetics*, *19*(10), 2059. https://doi.org/10.1093/HMG/DDQ078

Watanabe, K., Taskesen, E., Van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1), 1–11. https://doi.org/10.1038/s41467-017-01261-5

Watson, N., Linder, M. E., Druey, K. M., Kehrl, J. H., & Blumer, K. J. (1996). RGS family members: GTPase-activating proteins for heterotrimeric G-protein alpha-subunits. *Nature*, *383*(6596), 172–175. https://doi.org/10.1038/383172A0

Werkstetter, K. J., Korponay-Szabó, I. R., Popp, A., Villanacci, V., Salemme, M., Heilig, G., Lillevang, S. T., Mearin, M. L., Ribes-Koninckx, C., Thomas, A., Troncone, R., Filipiak, B., Mäki, M., Gyimesi, J., Najafi, M., Dolinšek, J., Dydensborg Sander, S., Auricchio, R., Papadopoulou, A., … Eftekhar Sadat, A. T. (2017). Accuracy in Diagnosis of Celiac Disease Without Biopsies in Clinical Practice. *Gastroenterology*, *153*(4), 924–935. https://doi.org/10.1053/J.GASTRO.2017.06.002

West, J., Logan, R. F. A., Smith, C. J., Hubbard, R. B., & Card, T. R. (2004). Malignancy and mortality in people with coeliac disease: population based cohort study. *BMJ : British Medical Journal*, *329*(7468), 716. https://doi.org/10.1136/BMJ.38169.486701.7C

West, J., Fleming, K. M., Tata, L. J., Card, T. R., Crooks, C. J. (2014). Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. *American Journal of Gastroenterology*, *109*(5), 757–768. https://doi.org/10.1038/ajg.2014.55

Westra, H. J., & Franke, L. (2014). From genome to function by studying eQTLs. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1842*(10), 1896–1902. https://doi.org/10.1016/J.BBADIS.2014.04.024

Westra, H. J., Peters, M. J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M. W., Fairfax, B. P., Schramm, K., Powell, J. E., Zhernakova, A., Zhernakova, D. V., Veldink, J. H., Van Den Berg, L. H., Karjalainen, J., Withoff, S., Uitterlinden, A. G., Hofman, A., Rivadeneira, F., … Franke, L. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics*, *45*(10), 1238–1243. https://doi.org/10.1038/ng.2756

White, L. E., Bannerman, E., & Gillett, P. M. (2016). Coeliac disease and the gluten-free diet: a review of the burdens; factors associated with adherence and impact on health-related quality of life, with specific focus on adolescence. *Journal of Human Nutrition and Dietetics*, *29*(5), 593–606. https://doi.org/10.1111/JHN.12375

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology*, *24*(2), 115–119. https://doi.org/10.1016/J.FM.2006.07.004

Withoff, S., Li, Y., Jonkers, I., & Wijmenga, C. (2016). Understanding Celiac Disease by Genomics. *Trends in Genetics*, *32*(5), 295–308. https://doi.org/10.1016/J.TIG.2016.02.003

Wurbel, A., Philippe, J.-M., Nguyen, C., eve Victorero, G., Freeman, T., Wooding, P., Miazek, A., eve Mattei, M.-G., Malissen, M., Jordan, B. R., Malissen, B., Carrier, A., & Naquet, P. (2000). The chemokine TECK is expressed by thymic and intestinal epithelial cells and attracts double-and single-positive thymocytes expressing the TECK receptor CCR9. *European Journal of Immunology*, *30*(1), 262–271. https://doi.org/10.1002/1521-4141

Yang, W., Shen, N., Ye, D. Q., Liu, Q., Zhang, Y., Qian, X. X., Hirankarn, N., Ying, D., Pan, H. F., Mok, C. C., Chan, T. M., Wong, R. W. S., Lee, K. W., Mok, M. Y., Wong, S. N., Leung, A. M. H., Li, X. P., Avihingsanon, Y., Wong, C. M., … Lau, Y. L. (2010). Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS Genetics*, *6*(2). https://doi.org/10.1371/JOURNAL.PGEN.1000841

Zaballos, Á., Gutiérrez, J., Varona, R., Ardavín, C., & Márquez, G. (1999). Cutting Edge: Identification of the Orphan Chemokine Receptor GPR-9-6 as CCR9, the Receptor for the Chemokine TECK. *The Journal of Immunology*, *162*(10).

Zanell, S., De Leo, L., Le, N. N. Q., Bo, N. D., Not, T., Mai, T. T. C., Son, P. D., Hai, L. T., Malaventura, C., Vatta, S., Ziberna, F., Mazzocco, M., Volpato, S., Lan, P. T., Le-Thi-Minh, H., & Borgna-Pignatti, C. (2016). Cross-sectional study of coeliac autoimmunity in a population of Vietnamese children. *BMJ Open*, *6*(6), e011173. https://doi.org/10.1136/BMJOPEN-2016-011173

Zarkadas, M., Cranney, A., Case, S., Molloy, M., Switzer, C., Graham, I. D., Butzner, J. D., Rashid, M., Warren, R. E., & Burrows, V. (2006). The impact of a gluten-free diet on adults with coeliac disease: Results of a national survey. *Journal of Human Nutrition and Dietetics*, *19*(1), 41–49. https://doi.org/10.1111/j.1365-277X.2006.00659.x

Zhang, J., Zhang, Y., Yang, J., Zhang, L., Sun, L., Pan, H. F., Hirankarn, N., Ying, D., Zeng, S., Lee, T. L., Lau, C. S., Chan, T. M., Leung, A. M. H., Mok, C. C., Wong, S. N., Lee, K. W., Ho, M. H. K., Lee, P. P. W., Chung, B. H. Y., … Yang, W. (2014). Three SNPs in chromosome 11q23.3 are independently associated with systemic lupus erythematosus in Asians. *Human Molecular Genetics*, *23*(2), 524–533. https://doi.org/10.1093/HMG/DDT424

Zhao, S., Fang, F., Tang, X., Dou, J., Wang, W., Zheng, X., Sun, L., & Zhang, A. (2017). An in-depth analysis identifies two new independent signals in 11q23.3 associated with vitiligo in the Chinese Han population. *Journal of Dermatological Science*, *88*(1), 103–109. https://doi.org/10.1016/J.JDERMSCI.2017.05.001

Zhernakova, A., Alizadeh, B. Z., Bevova, M., Van Leeuwen, M. A., Coenen, M. J. H., Franke, B., Franke, L., Posthumus, M. D., Van Heel, D. A., Van Der Steege, G., Radstake, T. R. D. J., Barrera, P., Roep, B. O., Koeleman, B. P. C., & Wijmenga, C. (2007). Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. *American Journal of Human Genetics*, *81*(6), 1284–1288. https://doi.org/10.1086/522037

Zhernakova, A., Festen, E. M., Franke, L., Trynka, G., van Diemen, C. C., Monsuur, A. J., Bevova, M., Nijmeijer, R. M., van 't Slot, R., Heijmans, R., Boezen, H. M., van Heel, D. A., van Bodegraven, A. A., Stokkers, P. C. F., Wijmenga, C., Crusius, J. B. A., & Weersma, R. K. (2008). Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. *American Journal of Human Genetics*, *82*(5), 1202–1210. https://doi.org/10.1016/J.AJHG.2008.03.016

Zhernakova, A., Stahl, E. A., Trynka, G., Raychaudhuri, S., Festen, E. A., Franke, L., Westra, H. J., Fehrmann, R. S. N., Kurreeman, F. A. S., Thomson, B., Gupta, N., Romanos, J., McManus, R., Ryan, A. W., Turner, G., Brouwer, E., Posthumus, M. D., Remmers, E. F., Tucci, F., … Plenge, R. M. (2011). Meta-Analysis of Genome-Wide Association Studies in Celiac Disease and Rheumatoid Arthritis Identifies Fourteen Non-HLA Shared Loci. *PLoS Genetics*, *7*(2), 1002004. https://doi.org/10.1371/JOURNAL.PGEN.1002004

Zhernakova, D. V., Deelen, P., Vermaat, M., Van Iterson, M., Van Galen, M., Arindrarto, W., Van't Hof, P., Mei, H., Van Dijk, F., Westra, H. J., Bonder, M. J., Van Rooij, J., Verkerk, M., Jhamai, P. M., Moed, M., Kielbasa, S. M., Bot, J., Nooren, I., Pool, R., … Franke, L. (2017). Identification of context-dependent expression quantitative trait loci in whole blood. *Nature Genetics*, *49*(1), 139–145. https://doi.org/10.1038/ng.3737

Zorro, M. M., Aguirre-Gamboa, R., Mayassi, T., Ciszewski, C., Barisani, D., Hu, S., Weersma, R. K., Withoff, S., Li, Y., Wijmenga, C., Jabri, B., & Jonkers, I. H. (2020). Tissue alarmins and adaptive cytokine induce dynamic and distinct transcriptional responses in tissue-resident intraepithelial cytotoxic T lymphocytes. *Journal of Autoimmunity*, *108*. https://doi.org/10.1016/J.JAUT.2020.102422

Äärelä, L., Nurminen, S., Kivelä, L., Huhtala, H., Mäki, M., Viitasalo, A., Kaukinen, K., Lakka, T., & Kurppa, K. (2016). Prevalence and associated factors of abnormal liver values in children with celiac disease. *Digestive and Liver Disease : Official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, *48*(9), 1023–1029. https://doi.org/10.1016/J.DLD.2016.05.022

Östensson, M., Montén, C., Bacelis, J., Gudjonsdottir, A. H., Adamovic, S., Ek, J., Ascher, H., Pollak, E., Arnell, H., Browaldh, L., Agardh, D., Wahlström, J., Nilsson, S., & Torinsson-Naluai, Å. (2013). A possible mechanism behind autoimmune disorders discovered by genome-wide linkage and association analysis in celiac disease. *PloS One*, *8*(8). https://doi.org/10.1371/JOURNAL.PONE.0070174

APPENDIX 1: GENE NAME ABBREVIATIONS

Nomenclature and formatting of gene names and their abbreviations are denoted according to the guidelines of the Human Genome Organization Gene Nomenclature Committee (HGNC).

ADAD1	Adenosine Deaminase Domain Containing 1
AP4B1	Adaptor Related Protein Complex 4 Subunit Beta 1
AP4B1-AS1	AP4B1 Antisense RNA 1
ARHGAP31	Rho GTPase Activating Protein 31
ATXN2	Ataxin 2
BACH2	BTB Domain And CNC Homolog 2
BLTP1	Bridge-Like Lipid Transfer Protein Family Member 1
CCR("number")	C-C Motif Chemokine Receptor
CCRL2	C-C Motif Chemokine Receptor Like 2
GLB1	Galactosidase Beta 1
CIITA	Class II Major Histocompatibility Complex Transactivator
CD28	CD28 Molecule
CD80	CD80 Molecule
CD247	CD247 Molecule
CLEC16	C-Type Lectin Domain Containing 16A
CLK3	CDC Like Kinase 3
COLCA1	Colorectal Cancer Associated 1
COLCA2	Colorectal Cancer Associated 2
CSK	C-Terminal Src Kinase
CTLA4	Cytotoxic T-Lymphocyte Associated Protein 4
DCLRE1B	DNA Cross-Link Repair 1B
DDX6	DEAD-Box Helicase 6
DUSP10	Dual Specificity Phosphatase 10
ELMO1	Engulfment and Cell Motility 1
ETS1	ETS Proto-Oncogene 1, Transcription Factor
FASLG	Fas Ligand

FRMD4B	FERM Domain Containing 4B
HCFC1	Host Cell Factor C1
ICOS	Inducible T Cell Costimulator
ICOSLG	Inducible T Cell Costimulator Ligand
IL2	Interleukin 2
IL21	Interleukin 21
IL12A	Interleukin 12A
IL18R1	Interleukin 18 Receptor 1
IL18RAP	Interleukin 18 Receptor Accessory Protein
INAVA	Innate Immunity Activator
IRAK1	Interleukin 1 Receptor Associated Kinase 1
IRF4	Interferon Regulatory Factor 4
ITGA4	Integrin Subunit Alpha 4
LPP	LIM Domain Containing Preferred Translocation Partner In Lipoma
LTF	Lactotransferrin
MAGI3	Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 3
MAP3K7	Mitogen-Activated Protein Kinase Kinase Kinase 7
MMEL1	Membrane Metalloendopeptidase Like 1
NCF2	Neutrophil Cytosolic Factor 2
NFIA	Nuclear Factor I A
OLIG3	Oligodendrocyte Transcription Factor 3
PARK7	Parkinsonism Associated Deglycase
PFKFB3	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 3
PHTF1	Putative Homeodomain Transcription Factor 1
PLEK	Pleckstrin
POGLUT1	Protein O-Glucosyltransferase 1
POU2AF1	POU Class 2 Homeobox Associating Factor 1
PRKCQ	Protein Kinase C Theta
PRM1	Protamine 1
PRM2	Protamine 2
PRXL2B	Peroxiredoxin Like 2B
PTPN2	Protein Tyrosine Phosphatase Non-Receptor Type 2
PTPN22	Protein Tyrosine Phosphatase Non-Receptor Type 22
PTPRK	Protein Tyrosine Phosphatase Receptor Type K

PUS10	Pseudouridine Synthase 10
PVT1	Pvt1 Oncogene
REL	REL Proto-Oncogene, NF-KB Subunit
RGS1	Regulator of G Protein Signaling 1
RSBN1	Round Spermatid Basic Protein 1
RUNX1	RUNX Family Transcription Factor 1
RUNX3	RUNX Family Transcription Factor 3
SCHIP1	Schwannomin Interacting Protein 1
SH2B3	SH2B Adaptor Protein 3
SMG7	SMG7 Nonsense Mediated MRNA Decay Factor
SOCS1	Suppressor of Cytokine Signaling 1
STAT4	Signal Transducer And Activator Of Transcription 4
TAGAP	T Cell Activation RhoGTPase Activating Protein
THEMIS	Thymocyte Selection Associated
TLR7	Toll Like Receptor 7
TLR8	Toll Like Receptor 8
TMEM187	Transmembrane Protein 187
TNFAIP3	TNF Alpha Induced Protein 3
TNFRSF9	TNF Receptor Superfamily Member 9
TNFRSF14	TNF Receptor Superfamily Member 14
TNFSF4	TNF Superfamily Member 4
TNFSF18	TNF Superfamily Member 18
TREH	Trehalase
TTC34	Tetratricopeptide Repeat Domain 34
UBASH3A	Ubiquitin Associated and SH3 Domain Containing A
UBE2E3	Ubiquitin Conjugating Enzyme E2 E3
UBE2L3	Ubiquitin Conjugating Enzyme E2 L3
YDJC	YdjC Chitooligosaccharide Deacetylase Homolog
ZFP36L1	ZFP36 Ring Finger Protein Like 1
ZMIZ1	Zinc Finger MIZ-Type Containing 1
ZNF335	Zinc Finger Protein 335

APPENDIX 2: GSRS QUESTIONNAIRE

R_____

THE GASTROINTESTINAL SYMPTOM RATING SCALE (GSRS)

Nimi_____

Lue tämä ensin:
Tutkimus sisältää kysymyksiä voinnistasi ja tilastasi kuluneen viikon aikana.
Merkitse rastilla (X) se vaihtoehto, joka sopii parhaiten sinuun ja tilaasi.

1. Onko Sinulla ollut VATSAKIPUJA kuluneen viikon aikana? (Vatsakivuilla tarkoitetaan kaikenlaista kipua tai särkyä vatsassa.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

2. Onko Sinulla ollut NÄRÄSTYSTÄ kuluneen viikon aikana? (Närästyksellä tarkoitetaan kirvelevää tai polttavaa pahanolontunnetta rintalastan takana.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

GSRS (F)

R_____

3. Onko Sinulla ollut HAPPAMIA RÖYHTÄISYJÄ kuluneen viikon aikana? (Happamilla röyhtäisyillä tarkoitetaan äkillisiä, hapanta vatsanestettä sisältäviä röyhtäisyjä.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

4. Onko Sinua HIUKAISSUT kuluneen viikon aikana? (Hiukaisulla tarkoitetaan vatsassa olevaa hiukovaa tunnetta, johon liittyy tarve syödä aterioiden välillä.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

5. Onko Sinulla ollut PAHOINVOINTIA kuluneen viikon aikana? (Pahoinvoinnilla tarkoitetaan pahanolontunnetta, joka saattaa muuttua kuvotukseksi tai oksentamiseksi.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

GSRS (F)

R_____

6. Onko vatsasi KURISSUT kuluneen viikon aikana? (Kurinalla tarkoitetaan vatsassa tuntuvaa värinää tai ”murinaa”.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

7. Onko vatsaasi TURVOTTANUT kuluneen viikon aikana? (Turvotuksella tarkoitetaan vatsassa tuntuvaa pingotusta, johon usein liittyy tuntemuksia ilmavaivoista.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

8. Onko Sinua vaivannut RÖYHTÄILY kuluneen viikon aikana? (Röyhtäilyllä tarkoitetaan tarvetta päästää ilmaa suun kautta, minkä yhteydessä vatsassa tuntuva pingotus usein helpottuu.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

GSRS (F)

R_____

9. Onko Sinulla ollut ILMAVAIVOJA kuluneen viikon aikana? (Ilmavaivoilla tarkoitetaan tässä tarvetta päästää ilmaa, jonka yhteydessä vatsassa tuntuva pingotus usein helpottuu.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

10. Onko Sinua vaivannut UMMETUS kuluneen viikon aikana? (Ummetuksella tarkoitetaan ulostuskertojen harventumista.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

11. Onko Sinua vaivannut RIPULI kuluneen viikon aikana? (Ripulilla tarkoitetaan ulostuskertojen lisääntymistä.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

GSRS (F)

R_____

12. Onko Sinua vaivannut LÖYSÄ VATSA kuluneen viikon aikana? (Jos ulosteesi on välillä ollut kovaa ja välillä löysää, ilmoita vain, missä määrin ulosteesi löysyys on Sinua vaivannut.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

13. Onko Sinua vaivannut KOVA VATSA kuluneen viikon aikana? (Jos ulosteesi on välillä ollut kovaa ja välillä löysää, ilmoita vain, missä määrin ulosteesi kovuus on Sinua vaivannut.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

14. Onko Sinua vaivannut kuluneen viikon aikana PAKOTTAVA ULOSTAMISEN TARVE? (Pakottavalla ulostamisen tarpeella tarkoitetaan äkillistä tarvetta käydä WC:ssä. Siihen liittyy usein puutteellisen pidättämiskyvyn tunne.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

GSRS (F)

R_____

15. Onko Sinulla kuluneen viikon aikana ollut ULOSTAMISEN YHTEYDESSÄ TUNNE, ETTÄ SUOLI EI OLE TYHJENTYNYT KOKONAAN? (Tällä tarkoitetaan, että suoli ei ponnistuksista huolimatta tunnu tyhjentyneen kunnolla.)

☐ Ei minkäänlaisia vaivoja
☐ Vähäpätöisiä vaivoja
☐ Lieviä vaivoja
☐ Kohtalaisia vaivoja
☐ Melko pahoja vaivoja
☐ Pahoja vaivoja
☐ Erittäin pahoja vaivoja

16. ONKO SINULLA VIIMEISEN KUUKAUDEN AIKANA ESIINTYNYT SEURAAVIA OIREITA (rengasta sopivat vaihtoehdot)

a. kielikipuja
b. haavaumia suussa
c. luustokipuja
d. puutumista
e. muuta, mitä _____

TARKISTA, ETTÄ OLET VASTANNUT KAIKKIIN KYSYMYKSIIN, ENNEN KUIN PALAUTAT LOMAKKEEN.

KIITOS AVUSTASI!

GSRS (F)

APPENDIX 3: PGWB QUESTIONNAIRE

The R_____

PGWB INDEX

Nimi_____

Tutkimuksen tämä osa sisältää kysymyksiä siitä, miltä Teistä tuntuu ja kuinka Teillä on mennyt VIIMEKSI KULUNEEN VIIKON AIKANA. Jokaisen kysymyksen osalta rastittakaa (X) se vaihtoehto, joka parhaiten sopii Teidän kohdallenne.

1. Miltä Teistä on YLEISESTI ottaen TUNTUNUT viimeksi kuluneen viikon aikana?

☐ Mielialani on ollut erinomainen
☐ Mielialani on ollut oikein hyvä
☐ Mielialani on ollut enimmäkseen hyvä
☐ Mielialani on vaihdellut paljon
☐ Mielialani on ollut enimmäkseen huono
☐ Mielialani on ollut hyvin huono

2. Kuinka usein Teitä on VAIVANNUT JOKIN SAIRAUS, RUUMIILLINEN VAIVA, SÄRYT tai KIVUT viimeksi kuluneen viikon aikana?

☐ Joka päivä
☐ Melkein joka päivä
☐ Noin puolet ajasta
☐ Silloin tällöin, mutta vähemmän kuin puolet ajasta
☐ Harvoin
☐ Ei koskaan

PGWB (F)

R_____

3. Tunsitteko itsenne MASENTUNEEKSI viimeksi kuluneen viikon aikana?
☐ Kyllä – niin paljon, että minusta tuntui siltä, että ottaisin itseni hengiltä
☐ Kyllä – niin paljon, etten välittänyt mistään
☐ Kyllä – hyvin masentuneeksi melkein joka päivä
☐ Kyllä – melko masentuneeksi useita kertoja
☐ Kyllä – lievästi masentuneeksi silloin tällöin
☐ Ei – en ole kertaakaan tuntenut itseäni lainkaan masentuneeksi

4. Oletteko pystynyt HALLITSEMAAN KÄYTTÄYTYMISTÄNNE, AJATUKSIANNE, MIELIALOJANNE tai TUNTEITANNE viimeksi kuluneen viikon aikana?
☐ Kyllä, ehdottomasti
☐ Kyllä – useimmiten
☐ Yleensä
☐ En kovin hyvin
☐ En, ja se häiritsee minua jonkin verran
☐ En, ja se häiritsee minua kovasti

5. Onko Teitä vaivannut HERMOSTUNEISUUS tai LEVOTTOMUUS viimeksi kuluneen viikon aikana?
☐ Erittäin paljon, jopa niin, että en ole voinut tehdä työtä tai huolehtia asioista
☐ Hyvin paljon
☐ Melko paljon
☐ Jonkin verran, niin että se on vaivannut minua
☐ Vähän
☐ Ei lainkaan

PGWB (F)

R_____

6. Kuinka paljon TARMOA, PIRTEYTTÄ tai ELINVOIMAA Teillä on ollut viimeksi kuluneen viikon aikana?

☐ Hyvin täynnä tarmoa – erittäin pirteä
☐ Melko tarmokas suurimman osan ajasta
☐ Tarmokkuuteni on vaihdellut melkoisesti
☐ Yleensä vähän tarmoa tai pirteyttä
☐ Hyvin vähän elinvoimaa tai tarmoa suurimman osan ajasta
☐ Ei lainkaan tarmoa tai elinvoimaa – olen tuntenut itseni loppuun ajetuksi tai loppuunkuluneeksi

7. Olen tuntenut itseni ALAKULOISEKSI JA SYNKKÄMIELISEKSI viimeksi kuluneen viikon aikana?

☐ En kertaakaan
☐ Vähän tänä aikana
☐ Jonkin verran tänä aikana
☐ Melkoisen osan tästä ajasta
☐ Suurimman osan tästä ajasta
☐ Koko ajan

8. Oletteko yleisesti ollut KIREÄ tai tuntenut itsenne JÄNNITTYNEEKSI viimeksi kuluneen viikon aikana?

☐ Kyllä, erittäin jännittyneeksi suurimman osan ajasta tai koko ajan
☐ Kyllä, hyvin jännittyneeksi suurimman osan ajasta
☐ En ole ollut koko ajan kireä, mutta olen tuntenut itseni melko jännittyneeksi useita kertoja
☐ Olen tuntenut itseni vähän jännittyneeksi muutamia kertoja
☐ En ole yleensä tuntenut itseäni jännittyneeksi
☐ En ole lainkaan tuntenut itseäni jännittyneeksi

PGWB (F)

R_____

9. Kuinka ONNELLINEN, TYYTYVÄINEN tai MIELISSÄNNE olette ollut viimeksi kuluneen viikon aikana?

☐ Erittäin onnellinen, en olisi voinut olla tyytyväisempi tai enemmän mielissäni
☐ Hyvin onnellinen suurimman osan ajasta
☐ Yleensä tyytyväinen ja mielissäni
☐ Joskus melko onnellinen ja joskus melko onneton
☐ Yleensä tyytymätön ja onneton
☐ Hyvin tyytymätön tai onneton suurimman osan ajasta tai koko ajan

10. Oletteko tuntenut itsenne riittävän TERVEEKSI tekemään asioita, joita haluatte tehdä tai Teidän on ollut pakko tehdä viimeksi kuluneen viikon aikana?

☐ Kyllä, ehdottomasti
☐ Suurimman osan ajasta
☐ Terveysongelmat ovat merkittävästi rajoittaneet minua
☐ Olen ollut vain niin terve, että olen voinut huolehtia itsestäni
☐ Olen tarvinnut jonkin verran apua itseni huolehtimisessa
☐ Olen tarvinnut toista henkilöä auttamaan itseäni useimmissa tai kaikissa asioissa, joita minun on täytynyt tehdä

11. Oletteko tuntenut itsenne niin SURULLISEKSI, LANNISTUNEEKSI tai TOIVOTTOMAKSI, että olette miettinyt, onko millään mitään merkitystä viimeksi kuluneen viikon aikana?

☐ Erittäin paljon – niin paljon, että olen ollut valmis luovuttamaan
☐ Hyvin paljon
☐ Melko lailla
☐ Jonkin verran – sen verran, että se on vaivannut minua
☐ Vähän
☐ En lainkaan

PGWB (F)

R_____

12. Oletteko herännyt PIRTEÄNÄ ja LEVÄNNEENÄ viimeksi kuluneen viikon aikana?
☐ En kertaakaan
☐ Muutaman harvan kerran
☐ Joitakin kertoja
☐ Aika monta kertaa
☐ Useimmiten
☐ Joka kerta

13. Oletteko ollut HUOLISSANNE tai LEVOTON TERVEYDESTÄNNE viimeksi kuluneen viikon aikana?
☐ Erittäin paljon
☐ Hyvin paljon
☐ Melko paljon
☐ Jonkin verran, mutta en kovin paljon
☐ Käytännöllisesti katsoen en koskaan
☐ En lainkaan

14. Onko Teistä tuntunut siltä, että olisitte ”MENETTÄMÄSSÄ JÄRKENNE” tai KONTROLLINNE siitä, miten TOIMITTE, PUHUTTE, AJATTELETTE, TUNNETTE tai MITÄ MUISTATTE viimeksi kuluneen viikon aikana?
☐ Ei lainkaan
☐ Vain vähän
☐ Jonkin verran, mutta ei niin paljon, että olisin ollut huolissani tai levoton siitä
☐ Jonkin verran ja olen ollut vähän huolissani
☐ Jonkin verran ja olen ollut melko huolissani
☐ Kyllä, hyvin paljon ja olen ollut hyvin huolissani

PGWB (F)

R_____

15. Päivittäinen elämäni on ollut TÄYNNÄ minua KIINNOSTAVIA ASIOITA viimeksi kuluneen viikon aikana?

☐ Ei lainkaan tänä aikana
☐ Vain pienen osan tästä ajasta
☐ Joskus
☐ Melkoisen osan tästä ajasta
☐ Suurimman osan tästä ajasta
☐ Koko ajan

16. Oletteko tuntenut itsenne AKTIIVISEKSI/TARMOKKAAKSI tai TYLSÄKSI/VELTOKSI viimeksi kuluneen viikon aikana?

☐ Hyvin aktiiviseksi/tarmokkaaksi joka päivä
☐ Enimmäkseen aktiiviseksi/tarmokkaaksi – en koskaan tylsäksi/veltoksi
☐ Melko aktiiviseksi/tarmokkaaksi – harvoin tylsäksi/veltoksi
☐ Melko tylsäksi/veltoksi – harvoin aktiiviseksi/tarmokkaaksi
☐ Enimmäkseen tylsäksi/veltoksi – en koskaan aktiiviseksi/tarmokkaaksi
☐ Hyvin tylsäksi/veltoksi joka päivä

17. Oletteko ollut HUOLESTUNUT, HARMISSANNE tai AHDISTUNUT viimeksi kuluneen viikon aikana?

☐ Erittäin paljon – niin paljon, että olen tuntenut itseni melkein sairaaksi huolestuneisuudesta
☐ Hyvin paljon
☐ Melko lailla
☐ Jonkin verran – sen verran, että se on vaivannut minua
☐ Vähän
☐ En lainkaan

PGWB (F)

R_____

18. Olen tuntenut itseni TASAPAINOISEKSI ja VARMAKSI viimeksi kuluneen viikon aikana?

☐ En lainkaan tänä aikana
☐ Pienen osan tästä ajasta
☐ Joskus
☐ Huomattavan osan tästä ajasta
☐ Suurimman osan tästä ajasta
☐ Koko ajan

19. Oletteko tuntenut itsenne LEVOLLISEKSI/HUOJENTUNEEKSI vai PINGOTTUNEEKSI/KIREÄKSI viimeksi kuluneen viikon aikana?

☐ Olen tuntenut itseni levolliseksi ja huojentuneeksi koko viikon
☐ Olen tuntenut itseni levolliseksi ja huojentuneeksi suurimman osan ajasta
☐ Yleensä olen tuntenut itseni levolliseksi, mutta ajoittain olen tuntenut itseni melko pingottuneeksi
☐ Yleensä olen tuntenut itseni pingottuneeksi, mutta ajoittain olen tuntenut itseni melko levolliseksi
☐ Olen tuntenut itseni pingottuneeksi/kireäksi suurimman osan ajasta
☐ Olen tuntenut itseni hyvin pingottuneeksi/kireäksi koko ajan

20. Olen tuntenut itseni ILOISEKSI/HUOLETTOMAKSI viimeksi kuluneen viikon aikana?

☐ En lainkaan tänä aikana
☐ Pienen osan tästä ajasta
☐ Joskus
☐ Melkoisen osan tästä ajasta
☐ Suurimman osan tästä ajasta
☐ Koko ajan

PGWB (F)

R_____

21. Olen tuntenut itseni VÄSYNEEKSI ja LOPPUUN KULUNEEKSI viimeksi kuluneen viikon aikana?

☐ En lainkaan tänä aikana
☐ Pienen osan tästä ajasta
☐ Joskus
☐ Melkoisen osan tästä ajasta
☐ Suurimman osan tästä ajasta
☐ Koko ajan

22. Oletteko tuntenut itsenne ”STRESSAANTUNEEKSI”, RASITTUNEEKSI tai PAINEEN ALAISEKSI viimeksi kuluneen viikon aikana?

☐ Kyllä, melkein enemmän kuin voin sietää tai kestää
☐ Kyllä melko lailla
☐ Kyllä, jonkin verran – enemmän kuin tavallisesti
☐ Kyllä, jonkin verran – kuten tavallisesti
☐ Kyllä, vähän
☐ En lainkaan

TARKISTAKAA, ETTÄ OLETTE VASTANNUT KAIKKIIN KYSYMYKSIIN!
KIITOS HYVÄSTÄ YHTEISTYÖSTÄ.

PGWB (F)

APPENDIX 4: SF -36 QUESTIONNAIRE

Liite 3

RAND 36-ITEM HEALTH SURVEY 1.0 (RAND-36)

Suomenkielinen versio

STAKES/KTL

1. **Onko terveytenne yleisesti ottaen ...**
(ympyröikää yksi numero)

1 erinomainen
2 varsin hyvä
3 hyvä
4 tyydyttävä
5 huono

2. **Jos vertaatte nykyistä terveydentilaanne vuoden takaiseen, onko terveytenne yleisesti ottaen ...**
(ympyröikää yksi numero)

1 tällä hetkellä paljon parempi kuin vuosi sitten
2 tällä hetkellä jonkin verran parempi kuin vuosi sitten
3 suunnilleen samanlainen
4 tällä hetkellä jonkin verran huonompi kuin vuosi sitten
5 tällä hetkellä paljon huonompi kuin vuosi sitten

Seuraavassa luetellaan erilaisia päivittäisiä toimintoja. Rajoittaako terveydentilanne nykyisin suoriutumistanne seuraavista päivittäisistä toiminnoista? Jos rajoittaa, kuinka paljon?
(ympyröikää yksi numero joka riviltä)

		kyllä, rajoittaa paljon	kyllä, rajoittaa hiukan	ei rajoita lainkaan
3.	huomattavia ponnistuksia vaativat toiminnat (esimerkiksi juokseminen, raskaiden tavaroiden nostelu, rasittava urheilu)	1	2	3
4.	kohtuullisia ponnistuksia vaativat toiminnat, kuten pöydän siirtäminen, imurointi, keilailu	1	2	3
5.	ruokakassien nostaminen tai kantaminen	1	2	3
6.	nouseminen portaita useita kerroksia	1	2	3
7.	nouseminen portaita yhden kerroksen	1	2	3
8.	vartalon taivuttaminen, polvistuminen, kumartuminen	1	2	3
9.	noin kahden kilometrin matkan kävely	1	2	3
10.	noin puolen kilometrin matkan kävely	1	2	3
11.	noin 100 metrin matkan kävely	1	2	3
12.	kylpeminen tai pukeutuminen	1	2	3

Onko teillä viimeisen 4 viikon aikana ollut RUUMIILLISEN TERVEYDENTILANNE TAKIA alla mainittuja ongelmia työssänne tai muissa tavanomaisissa päivittäisissä tehtävissänne?
(ympyröikää yksi numero joka riviltä)

		kyllä	ei
13.	Vähensitte työhön tai muihin tehtäviin käyttämäänne aikaa	1	2
14.	Saitte aikaiseksi vähemmän kuin halusitte	1	2
15.	Terveydentilanne asetti teille rajoituksia joissakin työ- tai muissa tehtävissä	1	2
16.	Töistänne tai tehtävistänne suoriutuminen tuotti vaikeuksia (olette joutunut esim. ponnistelemaan tavallista enemmän)	1	2

Onko teillä viimeisen 4 viikon aikana ollut TUNNE-ELÄMÄÄN LIITTYVIEN vaikeuksien (esim. masentuneisuus tai ahdistuneisuus) takia alla mainittuja ongelmia työssänne tai muissa tavanomaisissa päivittäisissä tehtävissänne?
(ympyröikää yksi numero joka riviltä)

		Kyllä	ei
17.	Vähensitte työhön tai muihin tehtäviin käyttämäänne aikaa	1	2
18.	Saitte aikaiseksi vähemmän kuin halusitte	1	2
19.	Ette suorittanut töitänne tai muita tehtäviänne yhtä huolellisesti kuin tavallisesti	1	2

20. **MISSÄ MÄÄRIN ruumiillinen terveydentilanne tai tunne-elämän vaikeudet ovat viimeisen 4 viikon aikana häirinneet tavanomaista (sosiaalista) toimintaanne perheen, ystävien, naapureiden tai muiden ihmisten parissa?**
(ympyröikää yksi numero)

1 ei lainkaan
2 hieman
3 kohtalaisesti
4 melko paljon
5 erittäin paljon

21. **Kuinka voimakkaita ruumiillisia kipuja teillä on ollut viimeisen 4 viikon aikana?**
(ympyröikää yksi numero)

1 ei lainkaan
2 hyvin lieviä
3 lieviä
4 kohtalaisia
5 voimakkaita
6 erittäin voimakkaita

22. **Kuinka paljon kipu on häirinnyt tavanomaista työtänne (kotona tai kodin ulkopuolella) viimeisen 4 viikon aikana?**
(ympyröikää yksi numero)

1 ei lainkaan
2 hieman
3 kohtalaisesti
4 melko paljon
5 erittäin paljon

Seuraavat kysymykset koskevat sitä, miltä teistä on tuntunut viimeisen 4 viikon aikana. Merkitkää kunkin kysymyksen kohdalla se numero, joka parhaiten kuvaa tuntemuksianne.
(ympyröikää yksi numero joka riviltä)

Kuinka suuren osan ajasta olette viimeisen 4 viikon aikana ...	koko ajan	suurimman osan aikaa	huomattavan osan aikaa	jonkin aikaa	vähän aikaa	en lainkaan
23. tuntenut olevanne täynnä elinvoimaa	1	2	3	4	5	6
24. ollut hyvin hermostunut	1	2	3	4	5	6
25. tuntenut mielialanne niin matalaksi, ettei mikään ole voinut teitä piristää	1	2	3	4	5	6
26. tuntenut itsenne tyyneksi ja rauhalliseksi	1	2	3	4	5	6
27. ollut täynnä tarmoa	1	2	3	4	5	6
28. tuntenut itsenne alakuloiseksi ja apeaksi	1	2	3	4	5	6
29. tuntenut itsenne “loppuunkuluneeksi”	1	2	3	4	5	6
30. ollut onnellinen	1	2	3	4	5	6
31. tuntenut itsenne väsyneeksi	1	2	3	4	5	6

32. **Kuinka suuren osan ajasta ruumiillinen terveydentilanne tai tunne-elämän vaikeudet ovat viimeisen 4 viikon aikana häirinneet tavanomaista sosiaalista toimintaanne (ystävien, sukulaisten, muiden ihmisten tapaaminen)?**
(ympyröikää yksi numero)

1 koko ajan
2 suurimman osan aikaa
3 jonkin aikaa
4 vähän aikaa
5 ei lainkaan

Kuinka hyvin seuraavat väittämät pitävät paikkansa teidän kohdallanne?
(ympyröikää yksi numero joka riviltä)

	pitää ehdotto-masti paikkansa	pitää enimmäk-seen paikkansa	en osaa sanoa	enimmäk-seen ei pidä paikkansa	ehdotto-masti ei pidä paikkansa
33. Minusta tuntuu, että sairastun jonkin verran helpommin kuin muut ihmiset	1	2	3	4	5
34. Olen vähintään yhtä terve kuin kaikki muutkin tuntemani ihmiset	1	2	3	4	5
35. Uskon, että terveyteni tulee heikkenemään	1	2	3	4	5
36. Terveyteni on erinomainen	1	2	3	4	5

PUBLICATIONS

PUBLICATION I

Differences between familial and sporadic celiac disease

Laura Airaksinen, Lauri Myllymäki, Katri Kaukinen, Päivi Saavalainen, Heini Huhtala, Katri Lindfors, Kalle Kurppa

Digestive Diseases and Sciences, 2021, 66(6), 1981–1988
https://doi.org/10.1007/s10620-020-06490-1

Digestive Diseases and Sciences (2021) 66:1981–1988
https://doi.org/10.1007/s10620-020-06490-1

ORIGINAL ARTICLE

Differences Between Familial and Sporadic Celiac Disease

Laura Airaksinen[1] · Lauri Myllymäki[1] · Katri Kaukinen[1,2] · Päivi Saavalainen[3] · Heini Huhtala[4] · Katri Lindfors[1] · Kalle Kurppa[5,6,7]

Received: 4 March 2020 / Accepted: 11 July 2020 / Published online: 23 July 2020

Abstract

Background It is not known if genetic background, characteristics at diagnosis, physical and psychological well-being, and adherence to a gluten-free diet are comparable between patients with familial or sporadic celiac disease. These issues were investigated in a follow-up study.

Methods Altogether 1064 patients were analyzed for celiac disease-associated serology, predisposing HLA-DQ, and non-HLA genotypes. Medical data were collected from patient records and supplementary interviews. Current symptoms and quality of life were further evaluated with the Gastrointestinal Symptom Rating Scale (GSRS), the Psychological General Well-Being questionnaire (PGWB), and Short Form 36 (SF-36) questionnaires.

Results Familial and sporadic groups differed ($P<0.001$) in the reason for diagnosis and clinical presentation at diagnosis, familial patients being more often screen-detected (26% vs. 2%, $P<0.001$) and having less often gastrointestinal (49% vs. 69%) and severe symptoms (47% vs. 65%). The groups were comparable in terms of histological damage, frequency of malabsorption, comorbidities, childhood diagnoses, and short-term treatment response. At the time of the study, familial cases reported fewer symptoms (21% vs. 30%, $P=0.004$) and lower prevalence of all (78% vs. 86%, $P=0.007$), neurological (10% vs. 15%, $P=0.013$), and dermatological (9% vs. 17%, $P=0.001$) comorbidities. Dietary adherence and GSRS scores were comparable, but familial cases had better quality of life according to PGWB and SF-36. High-risk genotype HLA-DQ2.5/DQ2.5 was more frequent among familial cases, and four non-HLA SNPs were associated with familial celiac disease.

Conclusions Despite the greater proportion of high-risk genotypes, familial cases had milder symptoms at presentation than did sporadic cases. Worse experience of symptoms and poorer quality of life in sporadic disease indicate a need for intensified support.

Keywords Celiac disease · Familial · Sporadic · Symptoms · Quality of life · HLA

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10620-020-06490-1) contains supplementary material, which is available to authorized users.

✉ Kalle Kurppa
kalle.kurppa@tuni.fi

Laura Airaksinen
laura.airaksinen@tuni.fi

Lauri Myllymäki
lauri.myllymaki@tuni.fi

Katri Kaukinen
katri.kaukinen@tuni.fi

Päivi Saavalainen
paivi.saavalainen@helsinki.fi

Heini Huhtala
heini.huhtala@tuni.fi

Katri Lindfors
katri.lindfors@tuni.fi

[1] Celiac Disease Research Center, Faculty of Medicine and Health Technology, Tampere University, PO BOX 100, 33014 Tampere, Finland

[2] Department of Internal Medicine, Tampere University Hospital, Tampere, Finland

[3] Research Programs Unit, Immunobiology and Department of Medical and Clinical Genetics, University of Helsinki, Helsinki, Finland

[4] Faculty of Social Sciences, Tampere University, Tampere, Finland

[5] Centre for Child Health Research, Tampere University Hospital and Tampere University, Tampere, Finland

[6] Department of Pediatrics, Seinäjoki University Hospital, Seinäjoki, Finland

[7] University Consortium of Seinäjoki, Seinäjoki, Finland

Introduction

Celiac disease is a chronic immune-mediated disorder in which ingestion of dietary gluten typically causes inflammation and morphological damage in the small bowel mucosa. According to population-based screening studies, the true prevalence of this heavily underdiagnosed disease is approximately 1–3% [1–3]. In specific at-risk groups, such as relatives of patients and subjects with another autoimmune disease, prevalence may reach as high as 5–15% [4–8]. At the individual level, the risk of celiac disease is increased by several factors including gender, predisposing HLA-DQ genotype and, in the case of familial celiac disease, the degree of relatedness with the index patient [7–9].

HLA class II genes encoding HLA-DQ2 and DQ8 are required for the development of celiac disease. Approximately 90% of patients carry HLA-DQ2.5 (encoded by *HLA-DQA1*0501* and HLA-*DQB1*0201*; approximately 20% homozygotes) [10, 11]. The rest carry either HLA-DQ2.2 (*DQA1*0201/DQB1*0202*) or HLA-DQ8 (*DQA1*03/DQB1*0302*). In addition, more than 40 non-HLA loci may contribute to disease susceptibility [12–14]. Interestingly, the presentation of celiac disease varies widely and patients may suffer either gastrointestinal or extraintestinal symptoms, or be even completely asymptomatic [15]. In fact, the phenotype may even vary between identical twins [16], indicating a modifying effect of environmental factors. It is currently unclear whether familial risk, either in conjunction with or independently of the genotype, also affects the phenotype and treatment outcomes in celiac disease, as well as long-term coping with the gluten-free diet.

The aim of this study was to compare familial and sporadic celiac disease with regard to the clinical, histological, and serological presentation at diagnosis and physical and psychological well-being and treatment compliance after being on dietary treatment for several years. This was established by exploiting large and well-defined cohorts of patients with or without affected family members.

Materials and Methods

Patients and Study Design

The study was carried out at the Celiac Disease Research Center, Tampere University, and at Tampere University Hospital. Biopsy-proven celiac disease patients and their relatives were recruited by a nationwide search with the help of nationwide and local celiac societies and by means newspaper announcements. In order to ascertain whether the presence of family risk affects coping with a gluten-free diet, all voluntary adult study participants completed specific questionnaires eliciting symptoms and quality of life. Furthermore, they, or in the case of a child the guardian, were interviewed by a physician or a study nurse with expertise in celiac disease. All relevant medical data and diagnoses were confirmed from patient records as available. In addition, blood samples were drawn from both the patients and their relatives for further analyses of celiac disease-associated serology and genetics (Fig. 1).

Family history of celiac disease was assessed by interview and from the medical records if reported. Furthermore, previously undiagnosed relatives with positive celiac antibodies in the present screening were referred to gastrointestinal endoscopy and the new confirmed cases were considered to be affected family members. Moreover, for the purposes of this study, relatives who refused the biopsy but had positive serum endomysium (EmA) and tissue transglutaminase antibodies (tTGab) were also regarded as affected family

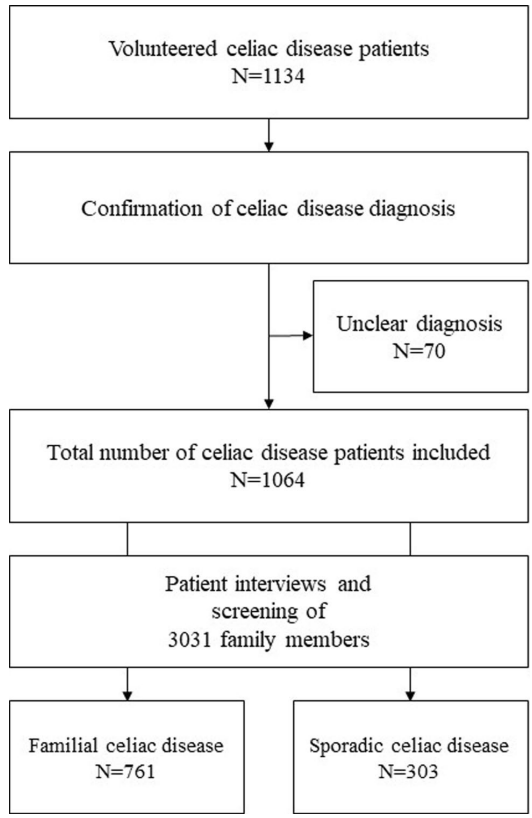

Fig. 1 Flowchart of the study

members based on the evidence that seropositivity for EmA and tTGab affords excellent specificity for celiac disease [5, 17]. Patients whose relatives had inconclusive serology and no biopsy were excluded from further analyses, as were those with unclear family history, non-celiac gluten sensitivity or only self-reported celiac disease.

The final study cohort included 1064 celiac disease patients, who were divided into "familial cases" ($n = 761$) with one or more affected relatives and "sporadic cases" ($n = 303$) with no diagnosed relatives (Fig. 1).

Clinical Data

Clinical information was gathered by patient interviews and supplemented from the patient records. In the case of children, the parents/guardians were interviewed. The data collected included demographic information, clinical presentation at diagnosis, and the main reason for suspicion of celiac disease, as well as celiac disease-associated (e.g., type 1 diabetes and autoimmune thyroidal disease) or other concomitant chronic illnesses. Moreover, data on adherence and capability to maintain a gluten-free diet, use of purified oats in the diet, and presence of any kind of (e.g., gastrointestinal and extraintestinal) recurrent self-reported symptoms and complications were recorded. Malabsorption was defined as weight loss and presence of characteristic laboratory abnormalities, such as anemia, hypoalbuminemia, low folate or low vitamin B12.

The main reason for suspecting celiac disease was further categorized into "gastrointestinal symptoms," "extraintestinal symptoms," and "screen-detected" and severity of symptoms before diagnosis as "none," "mild or moderate," and "severe" as previously defined [18]. Adherence to gluten-free diet was categorized as either "strict" or "occasional or frequent lapses" based on the dietary interview.

Serology

The results of celiac disease serology at the time of diagnosis were collected from the medical records. Only EmA titers were considered in this analysis, since some of the patients had been diagnosed before the introduction of tTGab tests. From serum samples collected at the time of the present study, tTGab values were tested by enzyme-linked immunosorbent assay (QUANTA Lite h-tTG IgA, INOVA Diagnostics, San Diego, CA; cutoff for positivity > 30 U/l) and EmA titers using indirect immunofluorescence with human umbilical cord as an antigen. Titers 1: ≥ 5 were considered positive for EmA, and positive samples were further diluted until negative to 1:50, 1:100, 1:200, 1:500, 1:1000, 1:2000, and 1:4000.

Histology

The results of histological analysis of the small-bowel mucosal biopsies were collected from the pathology reports. In our clinical practice, a minimum of four duodenal biopsies are taken upon endoscopy from each patient with suspected celiac disease and during the repeat endoscopy while on a gluten-free diet. Severity of small intestinal mucosal damage is evaluated from several representative and well-orientated biopsy specimens, and the degree of diagnostic villous atrophy is classified as partial, subtotal, or total.

Questionnaires

Three structured and validated questionnaires were used to evaluate current gastrointestinal symptoms and quality of life. This was done with adult patients only since the questionnaires are not validated in subjects under 18 years of age.

The Gastrointestinal Symptom Rating Scale (GSRS) measures self-reported symptoms with 15 selected questions [19]. Each individual question is scored on a Likert scale from 1 to 7 points, with higher scores indicating more severe gastrointestinal symptoms. Total score is calculated as an average of the 15 individual scores. In addition, five separate sub-scores, including diarrhea, indigestion, constipation, abdominal pain, and reflux, can be calculated as an average of the relevant questions.

The Psychological General Well-Being questionnaire (PGWB) was used to evaluate quality of life and well-being [20, 21]. It consists of 22 questions covering anxiety, depression, well-being, self-control, general health, and vitality. Each question is scored from 1 to 6 points, higher values indicating better self-reported quality of life and well-being. The total score is reported as a sum of each question and each sub-score as a sum of the relevant sub-category questions.

The Short Form 36 (SF-36) was also used to evaluate quality of life and health [22]. The questionnaire consists of 36 items divided into eight sub-categories including physical functioning, physical role limitations, emotional role limitations, vitality, mental health, social functioning, bodily pain, and general health. Each question is scored from 0 to 100 points, with higher scores indicating a better result. The subcategory scores are calculated as averages of the relevant items. Physical functioning refers to an individual´s capacity to undertake daily activities such as doing dishes and cleaning, while physical role limitations elicit if health issues prevent the subject, e.g., from going to work or school.

Genetic Analysis

The genotypes corresponding to disease-associated HLA variants HLA-DQ2.5, HLA-DQ8, and HLA-DQ2.2 were

determined from the patients using commercial HLA typing kits (Olerup SSP low-resolution kit, Olerup SSP AB, Saltsjöbaden, Sweden, or DELFIA® Celiac Disease Hybridization Assay Kit, PerkinElmer Life and Analytical Sciences, Wallac Oy, Turku, Finland) or the TaqMan chemistry-based genotyping of the HLA tagging SNPs as previously described [23, 24].

A further 552 patients were genotyped with Illumina 610-Quad BeadChip array for 39 non-HLA SNPs previously associated with celiac disease risk as a part of the European Genome-wide Association Study [13]. Of these, 37 SNPs passed the quality control filters (Hardy–Weinberg Equilibrium test, $P \leq 0.05$) and were tested for association with familiar/sporadic celiac disease. Genotypes were stored on and quality checks and filtering performed with BC Genome platform, version 4.0 (BC Platforms Espoo, Finland). Single marker association analyses were performed using PLINK, version 1.07 [25]. Patients with unclear genotype were excluded and, in order to avoid false positive findings due to trait correlation between genetically related individuals, only one patient from each family was included.

Statistics

Statistical analyses were performed with SPSS Statistics version 23 (IBM Corp, New York, NY, USA). Continuous variables were presented as medians with range or with lower (25th percentile) and upper (75th percentile) quartiles, or as number of subjects, and tested for statistical significance by Mann–Whitney U test. Binominal and categorical variables were presented as percentages and tested by Chi-square test. P value < 0.05 was considered significant across all analyses. Odds ratios (OR) with 95% confidence intervals (CI) were calculated for the non-HLA SNPs in both study groups.

Results

At diagnosis of celiac disease, the median age of the familial cases was 39 (range 0–81) years and of the sporadic cases 41 (range 1–79) years ($P = 0.010$). Of the familial cases, 39% had one and 61% had two or more affected relatives and 92% of all familial cases had affected first- or second-degree relative(s). Affected relative(s) were more often from mother's (64%) than father's (31%) side of the family. In 5% of familial cases, both maternal and paternal relatives were affected.

Familial cases were more often screen-detected and EmA positive and had less often gastrointestinal presentation, dermatitis herpetiformis, and severe symptoms at diagnosis (Table 1). There were no significant differences between the study groups in the prevalence of childhood diagnoses and malabsorption, or severity of small-bowel mucosal damage (Table 1). The groups also achieved comparable recovery of the mucosal morphology after 1 year on a gluten-free diet (full recovery of the villi in 59.3% and 60.3%, respectively, $P = 0.956$).

At present follow-up evaluation, the median age was 50 (range 2–89) years in the familial cases and 52 (6–84) years in the sporadic cases. The former group had been on gluten-free diet significantly longer (median 8 [range 4–15] vs. 7 [range 3–13] years, respectively; $P = 0.005$). Familial cases reported overall symptoms less often but were more often EmA positive on a gluten-free diet (Table 2). They also had less often regular follow-up with borderline significance, whereas the groups were comparable in current adherence and capability to manage a gluten-free diet, use of gluten-free oats, and frequency of tTGab positivity (Table 2). In addition, the groups did not differ in gastrointestinal

Table 1 Clinical, serological, and histological characteristics at diagnosis in 1064 patients with familial or sporadic celiac disease

	Familial n = 761		Sporadic n = 303		*P*
	N	%	N	%	
Females	554	72.8	229	75.6	0.353
Celiac disease diagnosis in childhood	143	18.8	44	14.5	0.088
Main reason for the diagnosis					**< 0.001***
Screening	200	26.3	7	2.3	
Extraintestinal symptoms[a]	187	24.6	88	29.1	
Gastrointestinal symptoms[b]	374	49.1	207	68.5	
Other common symptoms					
Malabsorption	270	35.7	107	35.5	0.971
Dermatitis herpetiformis	102	13.5	64	21.3	**0.002**
Severity of symptoms at diagnosis[c]					**< 0.001***
No symptoms	60	10.2	3	1.4	
Mild or moderate	251	42.8	71	33.3	
Severe	275	46.9	139	65.3	
Seropositivity[d] at diagnosis[e]	358	88.6	112	81.8	**0.040**
Severity of villous atrophy at diagnosis[f]					0.147
Partial	192	32.9	97	39.8	
Subtotal	233	39.9	91	37.3	
Total	159	27.2	56	23.0	

Bold values indicate statistically significant difference with P value < 0.05

[a]For example, dermatitis herpetiformis, arthralgia, rash, swelling, fatigue [18]

[b]For example, diarrhea, constipation, abdominal pain, flatulence, loose stools, mouth ulcers [18]

[d]Endomysium or antireticulin antibodies

Data were available in > 90% of the cases except [c]586 and 213; [e]404 and 137; and [f]584 and 244, respectively

*Calculated across all three variables

Table 2 Follow-up characteristics in 1064 celiac patients with familial or sporadic celiac disease

	Familial, $n=761$		Sporadic, $n=303$		P
	N	%	*N*	%	
Self-reported adherence to gluten-free diet					0.202
Strict	704	96.6	291	97.7	
Occasional or frequent lapses	29	3.4	7	2.3	
Capable to manage the diet	673	94.4	274	93.8	0.732
Use of purified oats	611	83.2	253	85.5	0.378
Current symptoms[a]	152	21.1	85	29.5	**0.004**
Follow-up serology[b]					
Positive endomysium antibodies	108	15.2	18	6.6	**< 0.001**
Positive tTGab	182	24.0	57	18.9	0.077
Regular follow-up	189	29.4	96	36.0	0.052

Bold values indicate statistically significant difference with P value < 0.05

tTGab tissue transglutaminase antibodies

[a]Any type of recurrent gastrointestinal and extraintestinal symptoms

[b]Only samples taken ≥ 2 years after diagnosis were counted

symptoms as measured by GSRS, but familial cases had better median PGWB general health score and SF-36 total, physical functioning, vitality, and mental health scores (Table 3).

Regarding concomitant chronic conditions, there were no differences between the groups in frequency of fractures, but familial cases were more often completely free from other conditions and had less often neurological and dermatological diseases (Supplementary Table 1).

Celiac disease-associated HLA haplotypes were available (one case per family) from 330 familial and 222 sporadic cases. The overall HLA-DQ distribution differed significantly between the two groups (Table 4). Homozygosity for HLA-DQ2.5 was also more common among the familial cases, while HLA-DQ2.2/DQ2.2 or HLA-DQ2.2/DQX, HLA-DQ8/DQ8, and HLA-DQX/DQX haplotypes were more common among sporadic cases (Table 4).

Of the 37 tested celiac disease-associated non-HLA SNPs, rs3748816 (OR 1.39, 95% CI 1.03–1.90; $P=0.034$), rs2816316 (OR 1.75, 95% CI 1.10–2.79; $P=0.017$), and rs2762051 (OR 1.48, CI 1.03–2.13; $P=0.035$) were associated with increased risk and rs10903122 (OR 0.71, 95% CI 0.53–0.96; $P=0.026$) with decreased risk for familial celiac disease (Supplementary Table 2).

Discussion

Patients with familial and sporadic celiac disease were found to have mostly comparable characteristics at diagnosis, except that the former were more often screen-detected and had milder symptoms. The minor differences in diagnostic approach and symptoms are probably attributable to the active screening of at-risk groups recommended in our national guidelines [17]. While there are no earlier studies with similar design, there are reports of a high frequency of undiagnosed celiac disease among family members of patients [26–29]. Altogether, there seems to be a gradual shift in the typical presentation of celiac disease toward a milder form [30, 31]. Interestingly, despite the greater proportion of asymptomatic/mildly symptomatic cases among the familial patients, the degree of histological damage was comparable between the groups. This concurs with reports showing a weak correlation between clinical presentation and severity of the mucosal lesions [28, 32–34], the ultimate reasons for which remain unclear.

The study groups were also found to have similar adherence to gluten-free diet, which is somewhat surprising as maintaining the diet could be expected to be less challenging in subjects with a family history of celiac disease. The excellent adherence in both groups is likely attributable to several factors, including the widespread availability and labeling of gluten-free products as well as the generally high awareness of the disease in Finnish food stores and restaurants, along with the former (now discontinued) governmentally granted financial reimbursement for officially diagnosed patients. Interestingly, despite equal dietary adherence, a greater proportion of sporadic patients reported having current self-perceived overall symptoms according to the interview. This experience is unlikely to be explained the minor difference in the duration of the gluten-free diet, since the symptoms generally diminish quite rapidly on treatment [35–37]. It must be mentioned that in spite of the equal self-reported dietary adherence, there was higher proportion of EmA positivity in the familial group on gluten-free diet. This may reflect their higher frequency of seropositivity already at diagnosis, since normalization of the autoantibodies may take longer than 2 years [38]. However, the possibility of

Table 3 Current symptoms and quality of life as measured by validated questionnaires in 627 adult* celiac patients with familial or sporadic celiac disease

	Familial, $n=420$		Sporadic, $n=207$		P
	Median	Q_1, Q_3	Median	Q_1, Q_3	
Gastrointestinal Symptom Rating Scale[a]					
Total score	1.9	1.5, 2.5	1.9	1.5, 2.6	0.379
Diarrhea	1.3	1.0, 2.3	1.7	1.0, 2.0	0.791
Indigestion	2.3	1.8, 3.3	2.3	1.8, 3.3	0.348
Constipation	1.7	1.0, 2.7	1.7	1.3, 2.7	0.282
Pain	2.0	1.3, 2.3	1.7	1.3, 2.5	0.875
Reflux	1.5	1.0, 2.0	1.5	1.0, 2.0	0.906
Psychological General Well-Being[b]					
Total score	107	95, 116	105	93, 115	0.171
Anxiety	25	21, 27	24	22, 27	0.688
Depression	17	15, 18	17	15, 18	0.283
Well-being	18	15, 20	17	15, 19	0.235
Self-control	16	14, 17	16	14, 17	0.707
General health	14	11, 15	13	10, 15	**0.028**
Vitality	18	16, 20	18	16, 20	0.410
Short Form 36[c,d]					
Total score	81	67, 89	78	63, 86	**0.011**
Physical functioning	95	80, 100	90	80, 98	0.126
Physical role functioning	100	50, 100	75	25, 100	**0.027**
Emotional role functioning	100	67, 100	100	67, 100	0.708
Vitality, energy	73	55, 85	70	50, 80	**0.015**
Mental health	84	72, 92	80	71, 88	**0.040**
Social role functioning	88	75, 100	88	75, 100	0.209
Bodily pain	78	58, 90	68	55, 90	0.216
General health perceptions	65	50, 80	60	40, 75	0.084

Data were available in > 90% of cases in each category except in[d] only from 376 familial cases

Bold values indicate statistically significant difference with P value < 0.05

Q_1, lower (25th percentile) quartile; Q_3, upper (75th percentile) quartile

*Children were excluded since the questionnaires are validated for adults only

Higher scores indicate either more severe symptoms[a], better well-being[b] or better functioning[c]

Table 4 Celiac disease-related human leukocyte antigen (HLA) genotypes in 552 patients with familial or sporadic celiac disease

	Familial, $n=330$		Sporadic, $n=222$		P[a]
	N	%	N	%	
HLA haplotype					**0.001**
DQ2.5/DQ2.5	69	20.9	26	11.7	
DQ2.5/DQ2.2	15	4.5	14	6.3	
DQ2.5/DQ8	27	8.2	19	8.6	
DQ2.5/DQX[b]	182	55.2	114	51.4	
DQ2.2/DQ2.2 or DQX[b]	7	2.1	17	7.7	
DQ8/DQ2.2 or DQX[b]	19	5.8	12	5.4	
DQ8/DQ8	3	0.9	6	2.7	
DQX[b]/DQX[b]	8	2.4	14	6.3	

Bold value indicates statistically significant difference with P value < 0.05

[a]Calculated between all haplotypes by Pearson Chi-square test

[b]DQX defines haplotype other than listed here

familiar cases actually having poorer dietary adherence cannot be fully excluded.

The sporadic patients had more often neurological and dermatological disorders, which could possibly explain the higher frequency of experienced symptoms as these complaints could be mistakenly attributed to celiac disease. Absence of peer support from family members with the disease might further hamper this assessment of causality and exacerbate the experience of symptoms [39, 40]. Alternatively, severe symptoms, more common among sporadic cases at diagnosis, may also predispose to persistent symptoms on a strict gluten-free diet [41], which could offer another explanation for the difference observed here. The experience of persistent symptoms, concomitant disorders, and lack of peer support may also explain the poorer quality of life as measured by PGWB and SF-36 scores in subjects with sporadic disease [40, 41]. These findings emphasize the importance of adequate guidance and support both at diagnosis and during the management of celiac disease.

There was also a significant difference in the HLA-DQ distribution between the groups. The high-risk genotype DQ2.5/DQ2.5 in particular was almost twice as frequent among familial cases, whereas the medium and low-risk genotypes [42] were, correspondingly, more common in sporadic disease. This is not surprising, since the predisposing risk alleles cluster within families with multiple affected members. In contrast to the findings of a recent meta-analysis [43], this was not reflected in a more severe and classic phenotype. However, the more active screening among familial cases complicates this issue, and further studies with larger numbers of cases are needed to confirm our findings. Besides the HLA genotypes, four SNPs were associated with familial celiac disease. Rs2762051 is located within the long non-coding RNA *DLEU1*, whereas the other three, rs3748816, rs2816316, and rs10903122, map to loci harboring genes *MMEL1/TNFRSF14*, *RGS1,* and *RUNX3*, respectively. These genes are all involved in immunological functions, and thus, the possible role of these non-HLA gene loci in familial celiac disease could be of interest in future studies.

The main strength of the present study is the carefully phenotyped cohort of patients with and without family history of celiac disease. Furthermore, a potential bias caused by undiagnosed disease among the relatives was reduced by serological screening of previously undiagnosed participants. One may criticize the fact that no biopsy was required for the diagnosis of these individuals, but this is no longer required in the Finnish diagnostic guidelines, and, in our opinion, it would be more biased to classify subjects with positive tTG and EmA as non-celiacs [5, 17]. As a limitation, it was not possible to recruit all the family members or to access comprehensive information on the family histories of all index patients, which may have impaired the detection of familial cases in the cohort. Moreover, the degree of familial relation to the index patient varied to some extent, since a minority of the familial cases had more distant than first- or second-degree relative(s) affected. Nor can it be fully excluded that even though not specifically reported here, the experienced symptoms and quality of life may in fact be attributable to confounding factors such as sporadic autoimmunity in close family members. In addition, although the study is clinically large, the groups were still small for purposes of genetic association analyses, and the systematic questionnaires used were validated only in adults.

To conclude, despite the greater proportion of high-risk genotypes among the subjects in the familial cohort, their clinical presentation was milder and other features comparable with those subjects with sporadic disease. The increased frequency of self-perceived symptoms and poorer health and quality of life scores in the questionnaires in sporadic cases underlines the need for physicians to pay special attention and possibly provide intensified support to this patient group.

Funding This study was supported by the Academy of Finland, the Finnish Medical Foundation, the Sohlberg Foundation, the Paulo Foundation, the Sigrid Juselius Foundation, the Foundation for Pediatric Research, and the Competitive State Research Financing of the Expert Area of Tampere University Hospital.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in this study were in accordance with the 1964 Helsinki Declaration and its later amendments. The study design, patient recruitment, and collection of patient record data were approved by the Regional Ethics Committee of Pirkanmaa Hospital District.

Informed consent Informed consent was obtained from all individual participants included in the study. This article does not contain any studies with animals performed by any of the authors.

References

1. Mäki M, Mustalahti K, Kokkonen J, et al. Prevalence of celiac disease among children in Finland. *N Engl J Med*. 2003;348:2517–2524.
2. Myléus A, Ivarsson A, Webb C, et al. Celiac disease revealed in 3% of Swedish 12-year-olds born during an epidemic. *J Pediatr Gastroenterol Nutr*. 2009;49:170–176.
3. Singh P, Arora A, Strand TA, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2018;16:e2.
4. Pittschieler K, Gentili L, Niederhofer H. Onset of coeliac disease: a prospective longitudinal study. *Acta Paediatr*. 2007;92:1149–1152.
5. Kurppa K, Salminiemi J, Ukkola A, et al. Utility of the new ESPGHAN criteria for the diagnosis of celiac disease in at: risk groups. *J Pediatr Gastroenterol Nutr*. 2012;54:387–391.
6. Oliveira A, Trindade E, Tavares M, Lima R, Terra M, Dias JA. Celiac disease in first degree relatives of celiac children. *Arq Gastroenterol*. 2012;49:204–207.
7. Singh P, Arora S, Lal S, Strand TA, Makharia GK. Risk of celiac disease in the first- and second-degree relatives of patients with celiac disease: a systematic review and meta-analysis. *Am J Gastroenterol*. 2015;110:1539–1548.
8. Wessels MMS, de Rooij N, Roovers L, Verhage J, de Vries W, Mearin ML. Towards an individual screening strategy for first-degree relatives of celiac patients. *Eur J Pediatr*. 2018;177:1585–1592.
9. Bourgey M, Calcagno G, Tinto N, et al. HLA related genetic risk for coeliac disease. *Gut*. 2007;56:1054–1059.
10. Kårhus LL, Thuesen BH, Skaaby T, Rumessen JJ, Linneberg A. The distribution of HLA DQ2 and DQ8 haplotypes and their association with health indicators in a general Danish population. *United Eur Gastroenterol J*. 2018;6:866–878.
11. Liu E, Dong F, Barón AE, et al. High incidence of celiac disease in a long-term study of adolescents with susceptibility genotypes. *Gastroenterology*. 2017;152:e1.
12. Dubois PCA, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet*. 2010;42:295–302.
13. Trynka G, Hunt KA, Bockett NA, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet*. 2011;43:1193–1201.
14. Gutierrez-Achury J, Zorro MM, Ricaño-Ponce I, et al. Functional implications of disease-specific variants in loci jointly associated with coeliac disease and rheumatoid arthritis. *Hum Mol Genet*. 2016;25:180–190.

15. Lindfors K, Ciacci C, Kurppa K, et al. Coeliac disease. *Nat Rev Dis Prim*. 2019;5:3.
16. Hervonen K, Karell K, Holopainen P, Collin P, Partanen J, Reunala T. Concordance of dermatitis herpetiformis and celiac disease in monozygous twins. *J Invest Dermatol*. 2000;115:990–993.
17. Celiac Disease. Current care guidelines, 2018. Working group set up by the Finnish Medical Society Duodecim and the Finnish Society of Gastroenterology. https://www.kaypahoito.fi. Accessed February 28, 2020.
18. Kivelä L, Kaukinen K, Lähdeaho ML, et al. Presentation of celiac disease in Finnish children is no longer changing: a 50-year perspective. *J Pediatr*. 2015;167:e1.
19. Svedlund J, Sjödin I, Dotevall G. GSRS: a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Dig Dis Sci*. 1988;33:129–134. https://doi.org/10.1007/252FBF01535722.
20. Dupuy HJ. The psychological general well-being (PGWB) index. In: Wenger NK, Mattson ME, Furburg CD EJ, eds. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: Le Jacq Publishing; 1984.
21. Dimenäs E, Carlsson G, Glise H, Israelsson B, Wiklund I. Relevance of norm values as part of the documentation of quality of life instruments for use in upper gastrointestinal disease. *Scand J Gastroenterol Suppl*. 1996;221:8–13.
22. McHorney CA, Ware JE, Rachel Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF–36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care.*. 1994;32:40–66.
23. Monsuur AJ, de Bakker PIW, Zhernakova A, et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. *PLoS ONE*. 2008;3:e2270.
24. Koskinen L, Romanos J, Kaukinen K, et al. Cost-effective HLA typing with tagging SNPs predicts celiac disease risk haplotypes in the Finnish, Hungarian, and Italian populations. *Immunogenetics*. 2009;61:247–256.
25. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
26. Rubio-Tapia A, Van Dyke CT, Lahr BD, et al. Predictors of family risk for celiac disease: a population-based study. *Clin Gastroenterol Hepatol*. 2008;6:983–987.
27. Agardh D, Lee HS, Kurppa K, et al. Clinical features of celiac disease: a prospective birth cohort. *Pediatrics*. 2015;135:627–634.
28. Kivelä L, Kaukinen K, Huhtala H, Lähdeaho M-L, Mäki M, Kurppa K. At-risk screened children with celiac disease are comparable in disease severity and dietary adherence to those found because of clinical suspicion: a large cohort study. *J Pediatr*. 2017;183:e2.
29. Nellikkal SS, Hafed Y, Larson JJ, Murray JA, Absah I. High prevalence of celiac disease among screened first-degree relatives. *Mayo Clin Proc*. 2019;94:1807–1813.
30. Brandimarte G, Tursi A, Giorgetti GM. Changing trends in clinical form of celiac disease. Which is now the main form of celiac disease in clinical practice? *Minerva Gastroenterol Dietol*. 2002;48:121–130.
31. Rampertab SD, Pooran N, Brar P, Singh P, Green PHR. Trends in the presentation of celiac disease. *Am J Med*. 2006;119:355.e9–355.e14.
32. Brar P, Kwon GY, Egbuna II, et al. Lack of correlation of degree of villous atrophy with severity of clinical presentation of coeliac disease. *Dig Liver Dis*. 2007;39:26–29.
33. Thomas HJ, Ahmad T, Rajaguru C, Barnardo M, Warren BF, Jewell DP. Contribution of histological, serological, and genetic factors to the clinical heterogeneity of adult-onset coeliac disease. *Scand J Gastroenterol*. 2009;44:1076–1083.
34. Taavela J, Koskinen O, Huhtala H, et al. Validation of morphometric analyses of small-intestinal biopsy readouts in celiac disease. *PLoS ONE.*. 2013;8:e76163.
35. Murray JA, Watson T, Clearman B, Mitros F. Effect of a gluten-free diet on gastrointestinal symptoms in celiac disease. *Am J Clin Nutr*. 2004;79:669–673.
36. Zarkadas M, Cranney A, Case S, et al. The impact of a gluten-free diet on adults with coeliac disease: results of a national survey. *J Hum Nutr Diet*. 2006;19:41–49.
37. Pekki H, Kurppa K, Mäki M, et al. Predictors and significance of incomplete mucosal recovery in celiac disease after 1 year on a gluten-free diet. *Am J Gastroenterol*. 2015;110:1078–1085.
38. Gidrewicz D, Trevenen CL, Lyon M, Butzner JD. Normalization time of celiac serology in children on a gluten-free diet. *J Pediatr Gastroenterol Nutr*. 2017;64:362–367.
39. Plevinsky JM, Greenley RN, Fishman LN. Self-management in patients with inflammatory bowel disease: strategies, outcomes, and integration into clinical care. *Clin Exp Gastroenterol*. 2016;9:259–267.
40. Halpert A. Irritable bowel syndrome: patient-provider interaction and patient education. *J Clin Med*. 2018;7:3.
41. Paarlahti P, Kurppa K, Ukkola A, et al. Predictors of persistent symptoms and reduced quality of life in treated coeliac disease patients: a large cross-sectional study. *BMC Gastroenterol*. 2013;13:75.
42. Romanos J, van Diemen CC, Nolte IM, et al. Analysis of HLA and non-HLA alleles can identify individuals at high risk for celiac disease. *Gastroenterology*. 2009;137:e3.
43. Bajor J, Szakács Z, Farkas N, et al. Classical celiac disease is more frequent with a double dose of HLA-DQB102: a systematic review with meta-analysis. *PLoS ONE*. 2019;14:e0212329.

PUBLICATION II

Influence of HLA-DQ2.5 Dose on Clinical Picture of Unrelated Celiac Disease Patients

Laura Airaksinen, Pilvi Laurikka, Heini Huhtala, Kalle Kurppa, Teea Salmi, Päivi Saavalainen, Katri Kaukinen, Katri Lindfors

Nutrients, 2020, 12(12), 3775
https://doi.org/10.3390/nu12123775

nutrients

MDPI

Article

Influence of HLA-DQ2.5 Dose on Clinical Picture of Unrelated Celiac Disease Patients

Laura Airaksinen [1], Pilvi Laurikka [1], Heini Huhtala [2], Kalle Kurppa [3,4], Teea Salmi [1,5], Päivi Saavalainen [6], Katri Kaukinen [1,7] and Katri Lindfors [1,*]

1 Celiac Disease Research Center, Faculty of Medicine and Health Technology, Tampere University, 33520 Tampere, Finland; laura.airaksinen@tuni.fi (L.A.); pilvi.laurikka@tuni.fi (P.L.); teea.salmi@tuni.fi (T.S.); katri.kaukinen@tuni.fi (K.K.)
2 Faculty of Social Sciences, Tampere University, 33520 Tampere, Finland; heini.huhtala@tuni.fi
3 Tampere Centre for Child Health Research, Tampere University Hospital and Tampere University, 33521 Tampere, Finland; kalle.kurppa@tuni.fi
4 Department of Pediatrics, Seinäjoki Central Hospital and University Consortium of Seinäjoki, 60220 Seinäjoki, Finland
5 Department of Dermatology, Tampere University Hospital, 33521 Tampere, Finland
6 Research Programs Unit, Immunobiology, and Haartman Institute, Department of Medical Genetics, University of Helsinki, 00014 Helsinki, Finland; paivi.saavalainen@helsinki.fi
7 Department of Internal Medicine, Tampere University Hospital, 33521 Tampere, Finland
* Correspondence: katri.lindfors@tuni.fi

Received: 13 November 2020; Accepted: 7 December 2020; Published: 9 December 2020

Abstract: The clinical phenotype of celiac disease varies considerably among patients and the dosage of HLA-DQ2.5 alleles has been suggested to be a contributing factor. We investigated whether HLA-DQ2.5 allele dosage is associated with distinct clinical parameters at the time of diagnosis and with patients' response to a gluten-free diet. The final cohort included 605 carefully phenotyped non-related Finnish celiac disease patients grouped as having 0, 1 or 2 copies of HLA-DQ2.5. Clinical data at the time of diagnosis and during gluten-free diet were collected systematically from medical records and supplementary interviews. An increasing HLA-DQ2.5 dose effect was detected for celiac disease antibody positivity at diagnosis ($p = 0.021$) and for the presence of any first-degree relatives with celiac disease ($p = 0.011$ and $p = 0.031$, respectively). Instead, DQ2.5-negative patients were suffering most often from classical symptoms at diagnosis ($p = 0.007$ between HLA groups). In addition, during follow-up they were most often symptomatic despite a gluten-free diet ($p = 0.002$ between groups). Our results thus suggest that increasing HLA-DQ2.5 dose only has a minor effect on the clinical picture of celiac disease. However, HLA-DQ2.5-negative patients should not be overlooked in clinical practice and particular attention should be paid to this patient group during gluten-free diet.

Keywords: celiac disease; HLA-DQ2.5; dose effect; clinical presentation; gluten-free diet

1. Introduction

Celiac disease is a chronic immune mediated condition driven by the ingestion of dietary gluten. It is characterized by small bowel mucosal damage and autoantibody response to transglutaminase 2 (TG2). The disease can be diagnosed at any age from childhood to older age and there is often a marked diagnostic delay [1,2]. Moreover, there is a substantial variation in the clinical picture of the disease, which may present with gastrointestinal and/or extraintestinal symptoms of varying severity or be completely asymptomatic [3]. Further variation to the disease phenotype is brought by several associated conditions [4]. The only efficient treatment for celiac disease is a strict gluten-free diet

(GFD), which usually results in alleviation of symptoms and normalization of small bowel mucosal morphology. Nevertheless, in a subset of patients, the symptoms and mucosal damage may persist despite a GFD [5,6]. If malabsorption and villous atrophy persist despite strict avoidance of gluten for a minimum of 12 months on GFD and after alternative causes have been excluded, the condition is termed refractory celiac disease (RCD) [7].

Family members of celiac disease patients are at an increased risk of being affected, likely due to their higher frequency of HLA-DQ2.5 or HLA-DQ8, the major determinants contributing to disease susceptibility [8]. HLA-DQ2.5 and DQ8 are heterodimeric molecules present on the surface of antigen presenting cells. In celiac disease, they bind and present deamidated gluten peptides to CD4-positive T cells, leading to the generation of an immune response [9,10]. HLA-DQ2.5 is far more common than HLA-DQ8 as it is present in more than 90% of patients [11]. The HLA-DQ2.5 heterodimer can be encoded by *HLA-DQA1*0501* and *HLA-DQB1*0201* alleles located on the same chromosome (*cis*) or by *DQA1*0505* and *DQB1*0202* on different chromosomes (*trans*). Homozygosity for HLA-DQ2.5 is associated with a particularly high risk for celiac disease [12–14]. This phenomenon has been attributed to the premise that gluten presented by antigen presenting cells in HLA-DQ2.5 homozygous individuals can induce a four-fold higher T cell response than in heterozygous individuals [15]. In addition to the disease risk, the dose of HLA-DQ2.5 has been suggested to affect the phenotype of celiac disease [16]. Earlier research on this issue has nevertheless reported conflicting findings, possibly due to rather small patient cohorts comprising a substantial portion of pediatric patients [16]. Nowadays, the majority of celiac disease diagnoses are made on adults and their clinical picture may be contributed to a larger extent to factors other than the HLA-type. Moreover, earlier research has not excluded patients originating from the same family, thereby increasing the possibility of bias caused by similar genetic background.

We investigated the HLA-DQ2.5 dose effect on various clinical parameters at the time of diagnosis and on patients' response to GFD by exploiting a large and carefully phenotyped cohort of unrelated pediatric and adult celiac disease patients.

2. Materials and Methods

2.1. Patients and Study Design

The study was conducted at Tampere University and Tampere University Hospital. Altogether, 1048 biopsy-proven celiac disease patients were recruited by a nationwide search with the help of national and local celiac societies and by media announcements. The patient information at diagnosis (demographics, clinical and histological data, celiac disease serology, presence of symptoms during childhood, and other concomitant chronic medical conditions as well as family history of celiac disease) was collected from medical records and from supplementary interviews by a physician or a study nurse. In the case of children, the guardian was interviewed. Patients were divided into four different age groups: 0–6 years, 7–20 years, 21–65 years, and >65 years. Diagnostic delay was categorized as 0 (screen-detected patients), <1, 1–5, 5–10 or >10 years. Abdominal symptoms included abdominal pain, diarrhea, loose stools, heartburn, flatulence, constipation, and/or bloating. Malabsorption was defined as weight loss and/or presence of characteristic laboratory abnormalities, such as anemia, hypoalbuminemia, low folate or low vitamin B12. Classical symptoms referred to the presence of both diarrhea and malabsorption [17]. Extraintestinal symptoms included any symptom(s) presenting outside of the gastrointestinal tract, such as dermatitis herpetiformis, infertility, joint pains, and neurological problems. Severity of symptoms was categorized into "no symptoms", "mild symptoms", "moderate symptoms" or "severe symptoms". "First-degree relative" referred to sibling, mother, father or offspring. "Relative" referred to any relative in a family. In addition, follow-up data on self-reported current symptoms, both gastrointestinal and extraintestinal, as well as adherence to GFD, were assessed by interviews. Adherence to GFD was described as either "strict

GFD", "dietary lapses" or "no GFD". In addition, histological and serological data at follow-up were assessed as described below.

In order to avoid false positive findings due to trait correlation between genetically related individuals, only one patient from each family was included (randomly). The final study cohort included 605 celiac disease patients.

The study design, patient recruitment, and collection of patient record data were approved by the Regional Ethics Committee of Tampere University Hospital. All participants gave written informed consent.

2.2. Histology

The results of histological analysis of the small-bowel mucosal biopsies at the time of diagnosis were collected from the pathology reports. In addition, if available, the degree of mucosal recovery on GFD evaluated from possible repeat biopsy was recorded. In both cases, severity of small intestinal mucosal damage was evaluated from several representative and well-orientated biopsy specimens and the degree of diagnostic villous atrophy was classified as partial, subtotal, or total, corresponding approximately to the IIIa, IIIb, and IIIc Marsh–Oberhuber classifications, respectively [18].

2.3. Serology

The results of celiac disease serology at the time of diagnosis were collected from the medical records. A patient was regarded as positive for celiac disease-specific antibodies if he or she was positive for TG2 autoantibodies [19] and/or endomysial autoantibodies (EmA) [20] and/or antireticulin autoantibodies (ARAs, measured in 1980/90s and later replaced by EmA) [21]. ARA and EmA were analyzed using indirect immunofluorescence with rat liver, kidney or stomach tissue (ARA) [21] or human umbilical cord (EmA) [22] as an antigen. Titers 1: ≥5 were considered positive. From serum samples collected at the time of the present study (follow-up data), TG2 antibodies and EmA were determined. TG2 antibodies values were tested by enzyme-linked immunosorbent assay (QUANTA Lite h-tTG IgA, INOVA Diagnostics, San Diego, CA, USA) with the cut-off for positivity being >30 U/L.

2.4. Genetic Analysis

The genotypes corresponding to disease-associated HLA variants, HLA-DQ2.5, HLA-DQ8, and HLA-DQ2.2, were determined using commercial HLA typing kits (Olerup SSP low-resolution kit, Olerup SSP AB, Saltsjöbaden, Sweden or DELFIA® Celiac Disease Hybridization Assay Kit, PerkinElmer Life and Analytical Sciences, Wallac Oy, Turku, Finland) or the TaqMan chemistry based genotyping of the HLA tagging single-nucleotide polymorphisms (SNPs) as previously described [23,24]. In this study, patients carrying alleles *HLA-DQB1*0201* and *HLA-DQA1*0501* in *cis* configuration were considered positive for HLA-DQ2.5. The subjects were divided into three groups according to whether they had zero, one or two copies of HLA-DQ2.5.

2.5. Statistics

Statistical analyses were performed with SPSS Statistics version 23 (IBM Corp, Armonk, NY, USA). Variables were presented as percentages and tested by Chi-square test or Fisher's Exact test, as appropriate. p value < 0.05 was considered significant across all analyses. Statistical analyses were performed for all patients together and for younger patients (<21 years, n = 124) and adults (n = 476) separately.

3. Results

Altogether, 100 (16.5%) celiac disease patients were negative (X/X group), 401 (66.3%) were heterozygous (DQ2.5/X group), and 104 (17.2%) were homozygous (DQ2.5/DQ2.5 group) for HLA-DQ2.5 (Table 1).

Table 1. Clinical, serological and histological characteristics of 605 celiac disease patients negative (X/X), heterozygous (DQ2.5/X) or homozygous (DQ2.5/DQ2.5) for HLA-DQ2.5 at the time of diagnosis.

	X/X ***n* = 100**	**DQ2.5/X** ***n* = 401**	**DQ2.5/DQ2.5** ***n* = 104**	***p***
	% (*n*)	**% (*n*)**	**% (*n*)**	
Females	78.0 (78)	73.3 (294)	76.0 (79)	0.589
Age group				0.996 *
0–6 years	7.1 (7)	6.8 (27)	7.7 (8)	
7–20 years	13.1 (13)	13.6 (54)	14.4 (15)	
21–65 years	75.8 (75)	76.1 (302)	73.1 (76)	
>65 years	4.0 (4)	3.5 (14)	4.8 (5)	
Symptoms in childhood	48.9 (43)	43.8 (161)	51.5 (51)	0.334
Diagnostic delay				0.908 *
0 [1]	4.3 (4)	4.9 (18)	3.0 (3)	
<1 year	20.7 (19)	23.6 (86)	25.7 (26)	
1–5 years	37.0 (34)	31.5 (115)	32.7 (33)	
5–10 years	7.6 (7)	10.4 (38)	6.9 (7)	
>10 years	30.4 (28)	29.6 (108)	31.7 (32)	
Classical symptoms [2]	30.0 (30)	16.7 (67)	21.2 (22)	**0.007**
Abdominal symptoms [3]	86.9 (86)	81.3 (322)	81.7 (85)	0.426
Diarrhea	44.7 (42)	37.1 (144)	41.7 (43)	0.337
Anemia	30.3 (30)	28.5 (113)	37.5 (39)	0.209
Extraintestinal manifestations [4]	44.4 (44)	49.5 (196)	47.1 (49)	0.646
Dermatitis herpetiformis	11.0 (10)	17.1 (63)	19.2 (19)	0.270
Severity of symptoms [5]				0.484 *
No symptoms	5.0 (4)	5.9 (17)	4.0 (3)	
Mild	23.8 (19)	25.9 (74)	33.3 (25)	
Moderate	15.0 (12)	12.2 (35)	5.3 (4)	
Severe	56.3 (45)	55.9 (160)	57.3 (43)	
Celiac disease antibody positivity [6,7]	77.8 (49)	86.8 (217)	94.9 (56)	**0.021**
Severity of mucosal damage [8]				0.546 *
Normal morphology	2.4 (2)	3.0 (10)	2.4 (2)	
Partial villous atrophy	27.7 (23)	33.8 (111)	25.3 (21)	
Subtotal/total villous atrophy	69.9 (58)	63.1 (207)	72.3 (60)	

Values in bold face indicate statistically significant difference with p value < 0.05. * Calculated across all variables. Data were available for >90% of patients except for designated parameters where number of cases in different HLA-DQ2.5 dose groups was [5] 80, 286, 75; [7] 63, 250, 59 and [8] 83, 328, 83, respectively. [1] Screen-detected patients. [2] Presence of diarrhea and malabsorption. [3] Abdominal pain, diarrhea, loose stools, heartburn, flatulence, constipation, and/or bloating. [4] Any symptom(s) presenting outside of the gastrointestinal tract, such as dermatitis herpetiformis, infertility, joint pains and neurological problems. [6] Positivity for serum TG2 autoantibodies and/or endomysial autoantibodies and/or antireticulin autoantibodies.

At the time of diagnosis, there was no significant difference in age distribution between the HLA-DQ2.5 dose groups. The proportion of patients suffering from classical symptoms was lowest among heterozygotes and highest among DQ2.5-negative patients (Table 1). No significant differences between the groups were observed in presence of symptoms in childhood, diagnostic delay or the type (abdominal, diarrhea, anemia, extraintestinal) or severity of symptoms (Table 1). The percentage of patients positive for celiac disease-specific autoantibodies was smallest in the HLA-DQ2.5-negative group and greatest in HLA-DQ2.5-homozygous group. The groups were comparable in terms of severity of mucosal damage (Table 1). DQ2.5 homozygotes most often had any relative or a first-degree relative with celiac disease, whereas DQ2.5-negative patients had such relatives the least often (Figure 1A,B).

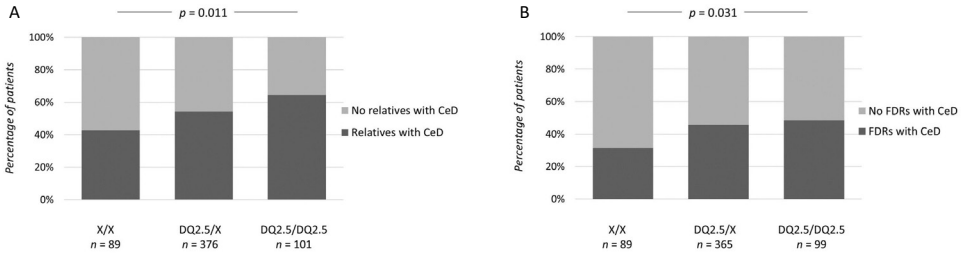

Figure 1. The percentages of patients among different HLA-DQ2.5 dose groups having (**A**) any relatives with CeD and (**B**) having first-degree relatives (FDR) with CeD. Patients are divided into different HLA-DQ2.5 dose groups based on whether they are either negative (X/X), heterozygous (DQ2.5/X) or homozygous (DQ2.5/DQ2.5) for HLA-DQ2.5.

At follow-up (median follow-up time 13 years, range <1–47 years), the DQ2.5-heterozygous patients maintained a strict GFD most often and homozygous patients least often (Table 2). The proportion of patients suffering from self-reported current symptoms was greatest among DQ2.5-negative and smallest among DQ2.5-heterozygous patients. Groups did not differ significantly in terms of antibody positivity, mucosal recovery, concomitant autoimmune diseases or malignancy at the time of the follow-up (Table 2).

Table 2. Clinical, serological and histological characteristics of 605 celiac disease patients negative (X/X), heterozygous (DQ2.5/X) or homozygous (DQ2.5/DQ2.5) for HLA-DQ2.5 at the time of the follow-up.

	X/X ***n* = 100**	**DQ2.5/X** ***n* = 401**	**DQ2.5/DQ2.5** ***n* = 104**	
	% (*n*)	**% (*n*)**	**% (*n*)**	***p***
Adherence to GFD [1]				**0.025** *
Strict GFD [1]	93.8 (91)	97.4 (376)	91.1 (92)	
Dietary lapses	5.2 (5)	2.3 (9)	7.9 (8)	
No GFD [1]	1.0 (1)	0.3 (1)	1.0 (1)	
Self-reported current symptoms [2,3]	44.8 (26)	22.9 (59)	32.3 (20)	**0.002**
Celiac disease antibody positivity [4,5]	28.6 (4)	23.8 (10)	9.1 (1)	0.535
Severity of mucosal damage [6]				0.108 *
Normal morphology	72.9 (35)	55.6 (104)	46.9 (23)	
Partial villous atrophy	25.0 (12)	38.0 (71)	44.9 (22)	
Subtotal/total villous atrophy	2.1 (1)	6.4 (12)	8.2 (4)	
Other illnesses				
Any autoimmune disease [7]	29.8 (28)	23.6 (91)	18.0 (18)	0.155
Type 1 diabetes	3.2 (3)	3.4 (13)	1.0 (1)	0.504
Thyroidal disease	17.9 (17)	14.0 (54)	9.7 (10)	0.248
Malignancy	3.2 (3)	3.9 (15)	5.8 (6)	0.581

Values in bold face indicate statistically significant difference with p value < 0.05. * Calculated across all variables. Data were available for >90% of patients except for designated parameters where numbers of cases in different HLA-DQ2.5 dose groups were [3] 58, 258, 62; [5] 14, 42, 11 and [6] 48, 187, 49, respectively. [1] GFD = gluten-free diet. [2] Any type of recurrent gastrointestinal and extraintestinal symptoms. [4] Positivity for serum TG2 and/or endomysial and/or antireticulin autoantibodies. [7] Any autoimmune disease including type 1 diabetes (DM1), thyroidal diseases, IgA nephropathy, Sjögren's syndrome, rheumatoid arthritis, sarcoidosis, psoriasis, vitiligo, and lichen planus.

The results of a subanalysis with adult patients only were parallel to those of the whole cohort. Significant differences between distinct HLA-DQ2.5 dose groups were observed in the presence of classical symptoms and celiac disease-specific antibody positivity at diagnosis (p = 0.006 and p = 0.043, respectively). Moreover, HLA-DQ2.5 homozygous adults also had more often either any relative or a first-degree relative with celiac disease than did heterozygous or HLA-DQ2.5-negative patients (p = 0.001 for any relative and p = 0.003 for first-degree relative). For adults at follow-up,

significant differences between dose groups were observed in adherence to the GFD and in self-reported current symptoms ($p = 0.016$ and $p = 0.002$, respectively). When children were analyzed separately, no significant differences were found in any of the parameters studied (data not shown).

4. Discussion

It has previously been reported that HLA-DQ2.5 dose is associated with increased risk of celiac disease [14], stronger disease-specific T cell response in vitro [15], and the presence of classical symptoms, particularly in children [16,25]. Here, we observed a dose effect for celiac autoantibody positivity at diagnosis as well as for the presence of any and first-degree relatives with the disease. However, we observed no increase in the presence of classical symptoms with an increasing HLA-DQ2.5 dose; instead, the HLA-DQ2.5-negative group presented most often with this phenotype. In addition, we found no association between HLA-DQ2.5 dose and abdominal symptoms, diarrhea, anemia, extraintestinal manifestations, severity of symptoms or mucosal morphology.

Our finding of an HLA-DQ2.5 dose effect with antibody positivity is rational since the gluten-specific T cells that proliferate and activate more robustly after stimulation in the context of the homozygous HLA-DQ2.5 antigen presenting cells participate in the induction of an anti-TG2 antibody response [26]. Moreover, such a dose effect in patients having a relative with celiac disease likely reflects the presence of the predisposing HLA-DQ2.5 within these families. When comparing our results with those of earlier studies, it is noteworthy that our cohort included mostly adult celiac disease patients, whereas earlier studies were conducted predominantly on pediatric patients [16]. It is possible that the factors affecting the disease phenotype differ between adults and children and also include other determinants besides the HLA-DQ type. The identity of such phenotype-modulating factors remains obscure but may, for instance, include non-HLA genetic variants and/or environmental factors. This assumption is supported by our previous finding that sib pairs with discordant clinical presentation had similar HLA haplotypes more often than pairs with the concordant phenotype did [27]. Interestingly, we have also observed that the diversity and composition of intestinal microbiota varies markedly between different celiac disease phenotypes, this being an interesting issue for further study [28].

HLA-DQ2.5 homozygosity has previously been observed in over 40% of patients with RCD type II (RCDII) in contrast to 20% in uncomplicated celiac disease [29]. RCDII is a severe condition with a poor prognosis [7]. Due to this and the fact that RCDII is very rare in Finland [30], our cohort did not include such cases and we were unable to address the HLA-DQ2.5 dose effect on this parameter. In any case, we found that HLA-DQ2.5 dose is not associated with poorer recovery of the intestinal damage. The investigation of HLA-DQ2.5 dose in RCDII would require a multi-center approach to achieve a sufficient number of patients for statistical power.

We found patients negative for HLA-DQ2.5 to suffer most often from classical symptoms at celiac disease diagnosis. Moreover, they most often experienced symptoms during GFD although their adherence to GFD was excellent. In addition, at follow-up, they did not differ from the other HLA-DQ2.5 groups in terms of celiac disease antibody positivity or small bowel mucosal morphology. A long diagnostic delay has been reported to predispose to persistent symptoms [31], but here patients negative for HLA-DQ2.5 received their diagnoses within the same time limits as the other HLA groups. Alternative explanations for persistent symptoms while on GFD could be altered composition of the small bowel mucosal microbiome [32], small-intestinal bacterial overgrowth [33] or continuous low-grade inflammation in spite of a strict GFD [34]. In any case, our results stress the need to pay special attention to HLA-DQ2.5-negative patients in clinics in order to prevent long-lasting health problems in this patient subset.

4.1. Strengths and Weaknesses

The main strength of our study was a large and well-defined cohort of unrelated adult and pediatric celiac disease patients enabling us to address the effect of HLA-DQ2.5 dose reliably. However, the study

was retrospective, which, given its nature may appear as a limitation. In addition, our study considered HLA-DQ2.5 alleles only in the *cis* configuration. Therefore, HLA-DQ2.5/HLA-DQ2.2 genotype was categorized as HLA-DQ2.5-heterozygous in spite of evidence to suggest that this genotype carries an equal risk for celiac disease as HLA-DQ2.5 homozygosity [35]. Moreover, the HLA-DQ2.5-negative group was heterogenous, comprising both the HLA-DQ8-positive cases as well as those without any of the major HLA types predisposing to celiac disease. Further, the small number of pediatric patients inhibited any reliable investigation of the effect of HLA-DQ2.5 dose in this particular subgroup.

4.2. Conclusions

In our cohort with a preponderance of adults, we demonstrated that the effect of HLA-DQ2.5 dose on the clinical picture of celiac disease was only modest. Patients negative for HLA-DQ2.5 were characterized by the most marked seronegativity, by the presence of classic symptoms at diagnosis, and also by symptoms persisting in spite of GFD.

Author Contributions: Conceptualization, L.A., K.K. (Kalle Kurppa), T.S., K.K. (Katri Kaukinen) and K.L.; Data curation, L.A., K.K. (Kalle Kurppa), K.K. (Katri Kaukinen) and K.L.; Formal analysis, L.A., P.L., H.H., P.S. and K.L.; Funding acquisition, L.A., K.K. (Kalle Kurppa), K.K. (Katri Kaukinen) and K.L.; Investigation, L.A., P.L., H.H., P.S. and K.L.; Supervision, K.K. (Katri Kaukinen) and K.L.; Writing—original draft, L.A. and K.L.; Writing—review and editing, L.A., P.L., H.H., K.K. (Kalle Kurppa), T.S., P.S., K.K. (Katri Kaukinen) and K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Academy of Finland, the Sigrid Juselius Foundation, the Paivikki 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.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Gray, A.M.; Papanicolas, I.N. Impact of symptoms on quality of life before and after diagnosis of coeliac disease: Results from a UK population survey. *BMC Health Serv. Res.* **2010**, *10*, 1–7. [CrossRef] [PubMed]
2. Norström, F.; Lindholm, L.; Sandström, O.; Nordyke, K.; Ivarsson, A. Delay to celiac disease diagnosis and its implications for health-related quality of life. *BMC Gastroenterol.* **2011**, *11*. [CrossRef] [PubMed]
3. Di Sabatino, A.; Corazza, G.R. Coeliac disease. *Lancet* **2009**, *373*, 1480–1493. [CrossRef]
4. Caio, G.; Volta, U.; Sapone, A.; Leffler, D.A.; De Giorgio, R.; Catassi, C.; Fasano, A. Celiac disease: A comprehensive current review. *BMC Med.* **2019**, *17*, 1–20. [CrossRef] [PubMed]
5. Midhagen, G.; Hallert, C. High rate of gastrointestinal symptoms in celiac patients living on a gluten-free diet: Controlled study. *Am. J. Gastroenterol.* **2003**, *98*, 2023–2026. [CrossRef] [PubMed]
6. Wahab, P.J.; Meijer, J.W.R.; Mulder, C.J.J. Histologic follow-up of people with celiac disease on a gluten-free diet: Slow and incomplete recovery. *Am. J. Clin. Pathol.* **2002**, *118*, 459–463. [CrossRef]
7. Hujoel, I.A.; Murray, J.A. Refractory Celiac Disease. *Curr. Gastroenterol. Rep.* **2020**, *22*, 18. [CrossRef]
8. Singh, P.; Arora, S.; Lal, S.; Strand, T.A.; Makharia, G.K. Risk of celiac disease in the first- and second-degree relatives of patients with celiac disease: A systematic review and meta-analysis. *Am. J. Gastroenterol.* **2015**, *110*, 1539–1548. [CrossRef]
9. Van De Wal, Y.; Kooy, Y.M.C.; Drijfhout, J.W.; Amons, R.; Koning, F. Peptide binding characteristics of the coeliac disease associated DQ(α1(*)0501, β1(*)0201) molecule. *Immunogenetics* **1996**, *44*, 246–253. [CrossRef] [PubMed]
10. Vartdal, F.; Johansen, B.H.; Friede, T.; Thorpe, C.J.; Stevanović, S.; Eriksen, J.E.; Sletten, K.; Thorsby, E.; Rammensee, H.G.; Sollid, L.M. The peptide binding motif of the disease associated HLA-DQ (α 1(*) 0501, β 1(*) 0201) molecule. *Eur. J. Immunol.* **1996**, *26*, 2764–2772. [CrossRef] [PubMed]
11. Sollid, L.M.; Thorsby, E. HLA susceptibility genes in celiac disease: Genetic mapping and role in pathogenesis. *Gastroenterology* **1993**, *105*, 910–922. [CrossRef]

12. Ploski, R.; Ek, J.; Thorsby, E.; Sollid, L.M. On the HLA-DQ(α1*0501, β1*0201)-associated susceptibility in celiac disease: A possible gene dosage effect of *DQB1*0201*. *Tissue Antigens* **1993**, *41*, 173–177. [CrossRef] [PubMed]
13. van Belzen, M.J.; Koeleman, B.P.C.; Crusius, J.B.A.; Meijer, J.W.R.; Bardoel, A.F.J.; Pearson, P.L.; Sandkuijl, L.A.; Houwen, R.H.J.; Wijmenga, C. Defining the contribution of the HLA region to cis DQ2-positive coeliac disease patients. *Genes Immun.* **2004**, *5*, 215–220. [CrossRef] [PubMed]
14. Margaritte-Jeannin, P.; Babron, M.C.; Bourgey, M.; Louka, A.S.; Clot, F.; Percopo, S.; Coto, I.; Hugot, J.P.; Ascher, H.; Sollid, L.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*, 562–567. [CrossRef]
15. Vader, W.; Stepniak, D.; Kooy, Y.; Mearin, L.; Thompson, A.; van Rood, J.J.; Spaenij, L.; Koning, F. The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12390–12395. [CrossRef]
16. Bajor, J.; Szakács, Z.; Farkas, N.; Hegyi, P.; Illés, A.; Solymár, M.; Pétervári, E.; Balaskó, M.; Pár, G.; Sarlós, P.; et al. Classical celiac disease is more frequent with a double dose of HLA-DQB102: A systematic review with meta-analysis. *PLoS ONE* **2019**, *14*, e0212329. [CrossRef]
17. Ludvigsson, J.F.; Leffler, D.A.; Bai, J.C.; Biagi, F.; Fasano, A.; Green, P.H.R.; Hadjivassiliou, M.; Kaukinen, K.; Kelly, C.P.; Leonard, J.N.; et al. The Oslo definitions for coeliac disease and related terms. *Gut* **2013**, *62*, 43–52. [CrossRef]
18. Dickson, B.C.; Streutker, C.J.; Chetty, R. Coeliac disease: An update for pathologists. *J. Clin. Pathol.* **2006**, *59*, 1008–1016. [CrossRef]
19. Dieterich, W.; Ehnis, T.; Bauer, M.; Donner, P.; Volta, U.; Riecken, E.O.; Schuppan, D. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat. Med.* **1997**, *3*, 797–801. [CrossRef]
20. Chorzelski, T.P.; Sulej, J.; Tchorzewska, H.; Jablonska, S.; Beutner, E.H.; Kumar, V. IgA Class Endomysium Antibodies in Dermatitis Herpetiformis and Coeliac Disease. *Ann. N. Y. Acad. Sci.* **1983**, *420*, 325–334. [CrossRef]
21. Eade, O.E.; Lloyd, R.S.; Lang, C.; Wright, R. IgA and IgG reticulin antibodies in coeliac and non-coeliac patients. *Gut* **1977**, *18*, 991–993. [CrossRef] [PubMed]
22. Sulkanen, S.; Collin, P.; Laurila, K.; Mäki, M. IgA- and IgG-class antihuman umbilical cord antibody tests in adult coeliac disease. *Scand. J. Gastroenterol.* **1998**, *33*, 251–254. [CrossRef] [PubMed]
23. Monsuur, A.J.; de Bakker, P.I.W.; Zhernakova, A.; Pinto, D.; Verduijn, W.; Romanos, J.; Auricchio, R.; Lopez, A.; van Heel, D.A.; Crusius, J.B.A.; et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. *PLoS ONE* **2008**, *3*, e2270. [CrossRef] [PubMed]
24. Koskinen, L.; Romanos, J.; Kaukinen, K.; Mustalahti, K.; Korponay-Szabo, I.; Barisani, D.; Bardella, M.T.; Ziberna, F.; Vatta, S.; Széles, G.; et al. Cost-effective HLA typing with tagging SNPs predicts celiac disease risk haplotypes in the Finnish, Hungarian, and Italian populations. *Immunogenetics* **2009**, *61*, 247–256. [CrossRef] [PubMed]
25. Martínez-Ojinaga, E.; Fernández-Prieto, M.; Molina, M.; Polanco, I.; Urcelay, E.; Núñez, C. Influence of HLA on clinical and analytical features of pediatric celiac disease. *BMC Gastroenterol.* **2019**, *19*, 91. [CrossRef] [PubMed]
26. du Pré, M.F.; Blazevski, J.; Dewan, A.E.; Stamnaes, J.; Kanduri, C.; Sandve, G.K.; Johannesen, M.K.; Lindstad, C.B.; Hnida, K.; Fugger, L.; et al. B cell tolerance and antibody production to the celiac disease autoantigen transglutaminase 2. *J. Exp. Med.* **2020**, *217*. [CrossRef]
27. Kauma, S.; Kaukinen, K.; Huhtala, H.; Kivelä, L.; Pekki, H.; Salmi, T.; Saavalainen, P.; Lindfors, K.; Kurppa, K.; Kauma, S.; et al. The Phenotype of Celiac Disease Has Low Concordance between Siblings, Despite a Similar Distribution of HLA Haplotypes. *Nutrients* **2019**, *11*, 479. [CrossRef]
28. Wacklin, P.; Kaukinen, K.; Tuovinen, E.; Collin, P.; Lindfors, K.; Partanen, J.; Mäki, M.; Mättuö, J. The duodenal microbiota composition of adult celiac disease patients is associated with the clinical manifestation of the disease. *Inflamm. Bowel Dis.* **2013**, *19*, 934–941. [CrossRef]
29. Al-Toma, A.; Goerres, M.S.; Meijer, J.W.R.; Peña, A.S.; Crusius, J.B.A.; Mulder, C.J.J. Human leukocyte antigen-DQ2 homozygosity and the development of refractory celiac disease and enteropathy-associated T-cell lymphoma. *Clin. Gastroenterol. Hepatol.* **2006**, *4*, 315–319. [CrossRef]
30. Ilus, T.; Kaukinen, K.; Virta, L.J.; Pukkala, E.; Collin, P. Incidence of malignancies in diagnosed celiac patients: A population-based estimate. *Am. J. Gastroenterol.* **2014**, *109*, 1471–1477. [CrossRef]

31. Paarlahti, P.; Kurppa, K.; Ukkola, A.; Collin, P.; Huhtala, H.; Mäki, M.; Kaukinen, K. Predictors of persistent symptoms and reduced quality of life in treated coeliac disease patients: A large cross-sectional study. *BMC Gastroenterol.* **2013**, *13*, 75. [CrossRef] [PubMed]
32. Wacklin, P.; Laurikka, P.; Lindfors, K.; Collin, P.; Salmi, T.; Lähdeaho, M.L.; Saavalainen, P.; Mäki, M.; Mättö, J.; Kurppa, K.; et al. Altered Duodenal microbiota composition in celiac disease patients suffering from persistent symptoms on a long-term gluten-free diet. *Am. J. Gastroenterol.* **2014**, *109*, 1933–1941. [CrossRef] [PubMed]
33. Tursi, A.; Brandimarte, G.; Giorgetti, G.M. High prevalence of small intestinal bacterial overgrowth in celiac patients with persistence of gastrointestinal symptoms after gluten withdrawal. *Am. J. Gastroenterol.* **2003**, *98*, 839–843. [CrossRef] [PubMed]
34. Rubio-Tapia, A.; Rahim, M.W.; See, J.A.; Lahr, B.D.; Wu, T.T.; Murray, J.A. Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *Am. J. Gastroenterol.* **2010**, *105*, 1412–1420. [CrossRef]
35. Sollid, L.M. The roles of MHC class II genes and post-translational modification in celiac disease. *Immunogenetics* **2017**, *69*, 605–616. [CrossRef]

PUBLICATION III

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

Laura Airaksinen, Juliana XM Cerqueira, Heini Huhtala, Päivi Saavalainen, Dawit A. Yohannes, Markku Mäki, Kalle Kurppa, Elina Kilpeläinen, Anastasia Shcherban, Aarno Palotie, Katri Kaukinen, Katri Lindfors

Journal of Translational Autoimmunity, 2021, 14(4), 100128
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 ≤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 ≤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

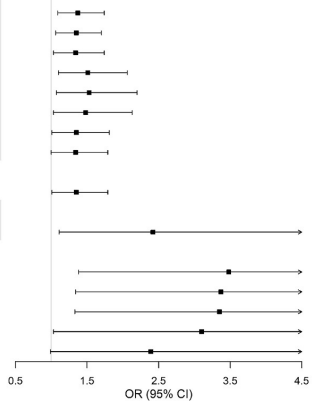

Fig. 1. Forest plot representing association of SNPs in the genomic region of the candidate gene *CCR9* with celiac disease and its different phenotypes. eQTL; expression quantitative trait loci, OR; odds ratio, CI; confidence interval, P_{EMP2}; empirical P value at 10,000 permutation threshold, PBMC;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

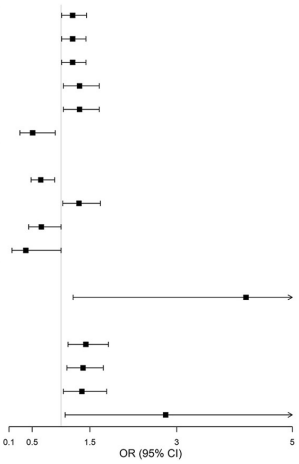

Fig. 2. Forest plot of representing association of SNPs in the genomic region of the candidate gene *CCL25* with celiac disease and its different phenotypes. SNPs identified in sequencing analysis and imputed are indicated by an asterisk. eQTL; expression quantitative trait loci, OR; odds ratio, CI; 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 $\alpha 4$ 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.