Biobank Norway
Nordic Biobank Network
NORDIC Biobank Conference
NSHG-PM

Nordic Conference on Future Health 2024

NTNU
HUNT4
HELSEUNDERSØKELSEN I NORD-TRØNDELAG
Faculty of Medicine and Health Sciences
Department of Public Health and Nursing

A genome-wide association study of celiac disease identifies novel risk variants in the 5p15.33 locus: results from a population-based screening of adults, the HUNT4 study.

Mohammad Sayeef Alam[1,2,✉], Rebecka Hjort[1,2], Laurent F. Thomas[1,3,4], Ben Brumpton[1,2], Kristian Hveem[1], Knut E. A. Lundin[5,6], Iris H. Jonkers[7], Ludvig M. Sollid[6,8], and Eivind Ness-Jensen[1,2,9,10]

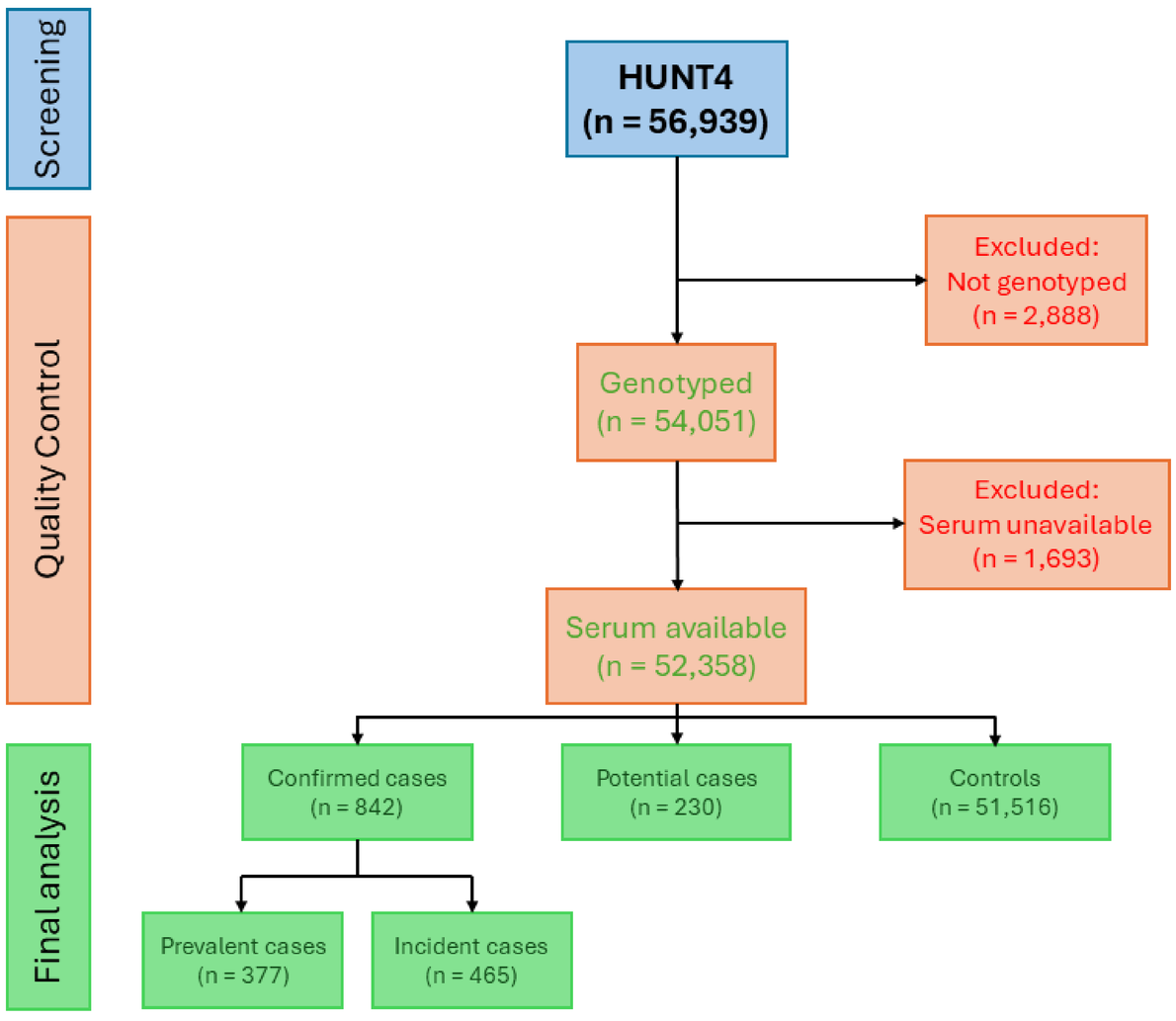
1 HUNT Center for Molecular and Clinical Epidemiology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
2 HUNT Research Centre, Norwegian University of Science and Technology (NTNU), Levanger, Norway
3 Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
4 BioCore - Bioinformatics Core Facility, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
5 Department of Gastroenterology, Oslo University Hospital - Rikshospitalet, Oslo, Norway
6 Norwegian Coeliac Disease Research Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
7 Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
8 Department of Immunology, Oslo University Hospital - Rikshospitalet, Oslo, Norway
9 Department of Medicine, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway
10 Department of Molecular Medicine and Surgery, Karolinska Institutet & Karolinska University Hospital, Stockholm, Sweden

This is what our issue is:

1. **Celiac disease (CeD) is a chronic, multifactorial disease affecting approximately 1% of individuals of European ancestry, caused by an immune response to gluten, leading to enteropathy of the small intestine.**
2. **Genetic studies have identified the role of HLA-DQ heterodimers in CeD, but only 1-2% of carriers develop the disease, indicating that HLA-DQ is necessary but not sufficient for CeD development.**
3. **Previous genome-wide association studies (GWAS) have included only prevalent cases, leading to selection bias and underrepresentation of the true population scenario.**
4. **The use of immune-specific chips and emphasis on immune-specific loci in non-HLA regions during genotyping has limited the discovery of novel genetic variants.**
5. **This study aims to address these biases by conducting a large population-based screening of CeD among adults, including all prevalent and incident cases, and implementing a genotyping strategy that emphasizes all regions equally.**

To investigate it, we did this:

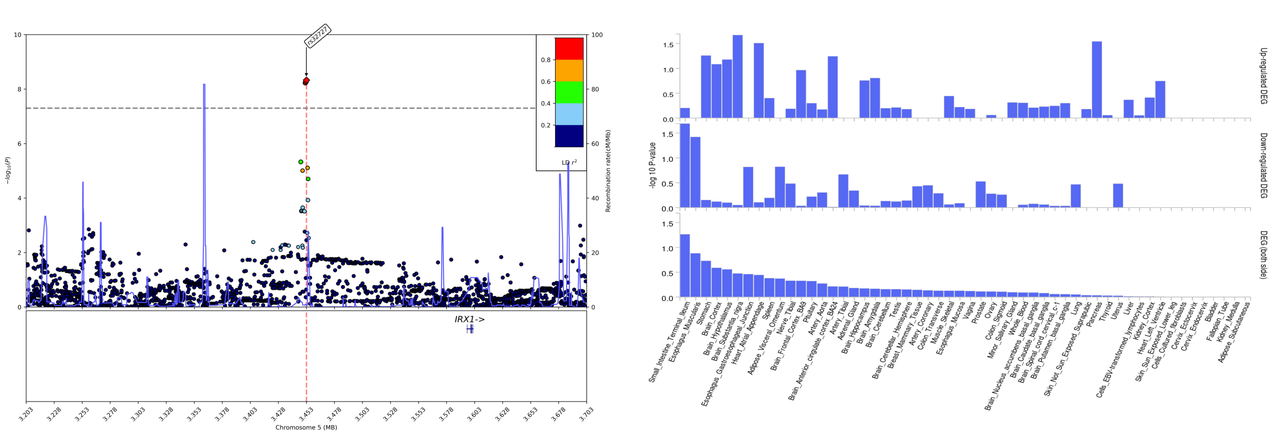
1. **Data was collected from the fourth round of the Trøndelag Health Study (HUNT4) with a 54% response rate, including 56,939 serum samples for analysis.**
2. **A comprehensive screening process was implemented, including serological testing for TG2 IgA and IgG antibodies, clinical evaluations, endoscopic examinations, and histopathological assessments to confirm CeD diagnoses.**



3. **Genotyping was performed using four different Illumina HumanCoreExome arrays, with stringent quality control measures applied, resulting in 358,964 polymorphic variants included in the study.**
4. **The SAIGE tool was used for GWAS, accounting for sample relatedness and case-control imbalance, with a focus on biopsy-confirmed CeD and including various covariates.**
5. **The genome-wide heritability of the non-HLA region was calculated using the LDSC function, and functional mapping and annotation were performed using FUMA, including tissue-specificity and pathway enrichment analysis.**

We found out that:

1. **13 novel variants from 12 novel and one known locus visualized in a Manhattan plot for all the known and novel loci with minor allele count ≥ 3. The novel loci are indicated in green and the known in blue. The significance threshold used was $P \leq 5 \times 10^{-8}$**



2. **Regional plot shows that the closest gene to the lead SNP is *IRX1*.**
3. **The genome-wide heritability using LDSC function was estimated to be around 23%.**
4. **The differentially expressed gene enrichment score is observed highest in the small intestines.**
5. **In the six of the subset validation analysis, the chr5p15.33 locus reached a significance threshold of $P \leq 5 \times 10^{-7}$ in each of them.**

Our take home message is that:

1. **This study found a new genetic association at the chr5p15.33 locus and confirmed 38 previously reported loci. Gene expression was significant in the gut, immune, and brain tissues which indicates presence of extra-gastrointestinal influence on the development of the disease.**
2. **The chr5p15.33 locus was missed in recent studies, highlighting the need for comprehensive genotyping. CeD shares genetic risk with other autoimmune diseases like rheumatoid arthritis and type 1 diabetes.**
3. **The strengths of the study are large, population-based cohort with thorough diagnosis and use of the SAIGE tool for accurate estimates.**
4. **The major limitation of the study are lack of data on children and non-European populations; small number of cases for rare variants.**
5. **Our recommendation for future studies would be to include diverse populations and age groups, use denser genotyping chips, and conduct functional pathway analyses.**
6. **To conclude this study identifies a new genetic association and confirms several known loci, highlighting shared genetic risks with other autoimmune diseases and the importance of comprehensive genotyping.**