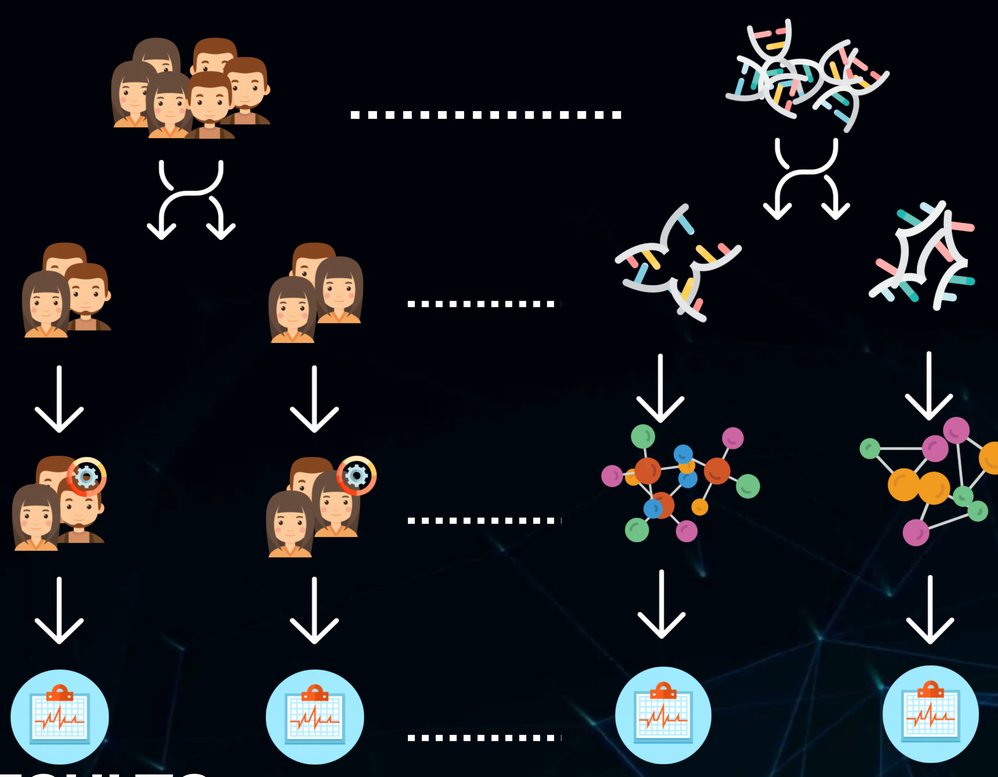
Celiac Disease Triggers

Using Mendelian Randomization to infer causality

Matthijs Knigge[1], Serena Sanna[2]

1. *Hanze University of Groningen, Institute for Life Science & Technology, Bioinformatics, University of Medical Center Groningen, Department of Genetics, Groningen, The Netherlands.*
2. *University of Groningen, University of Medical Center Groningen, Department of Genetics, Groningen, The Netherlands.*

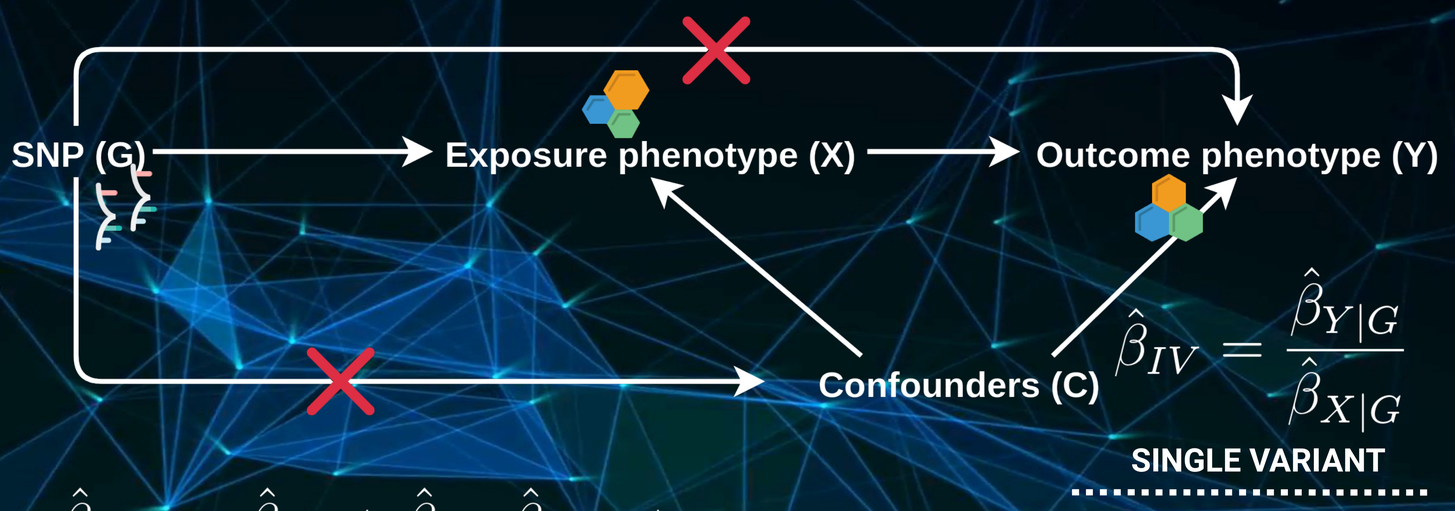


INTRODUCTION

Celiac disorder affects up to two percent of the European population, and it is hereditary. The inheritance model of celiac disease is oligogenic.[1] In this model, a genetic variant in the Major Histocompatibility Complex (MHC) region, the HLA-DQ2 haplotype, accounts for 40% of the inheritable risk, but it is not sufficient enough to develop the disease. 30% of the Europeans are in fact carriers of the variant, but only 3% express celiac disease. Other 42 non-HLA variants have been identified, but still remains little understanding on what triggers the disease. We aimed to systematically assess multiple clinical parameters and molecular mechanisms to identify causal and protective factors, that can explain what is triggering disease manifestation in HLA-DQ2 carriers.[2]

MATERIALS AND METHODS

We investigated over 500.000 clinical parameters and molecular phenotypes, for which summary statistics from genome-wide association studies (GWAS) were available. We used the concept of Mendelian Randomization[3] (MR) approach, and applied two-samples MR methods: Inverse Variance Weighted[4] (IVW) and MR-Egger method[5] to infer causality links by combining summary statistics from GWAS. For our analyses, we discarded variants located in the MHC region to avoid over-estimation of causality for factors that also influenced by the same or other HLA haplotypes.



$$\hat{\beta}_{Y|G} = \hat{\beta}_{0E} + \hat{\beta}_{1E}\ \hat{\beta}_{X|G} + \epsilon_G;$$

$$\epsilon_G \sim \mathcal{N}(0, \sigma^2_{\epsilon_G}\sigma^2_{\hat{\beta}_{Y|X}})$$

$$\hat{\beta}_{IV} = \frac{\hat{\beta}_{Y|G}}{\hat{\beta}_{X|G}}$$

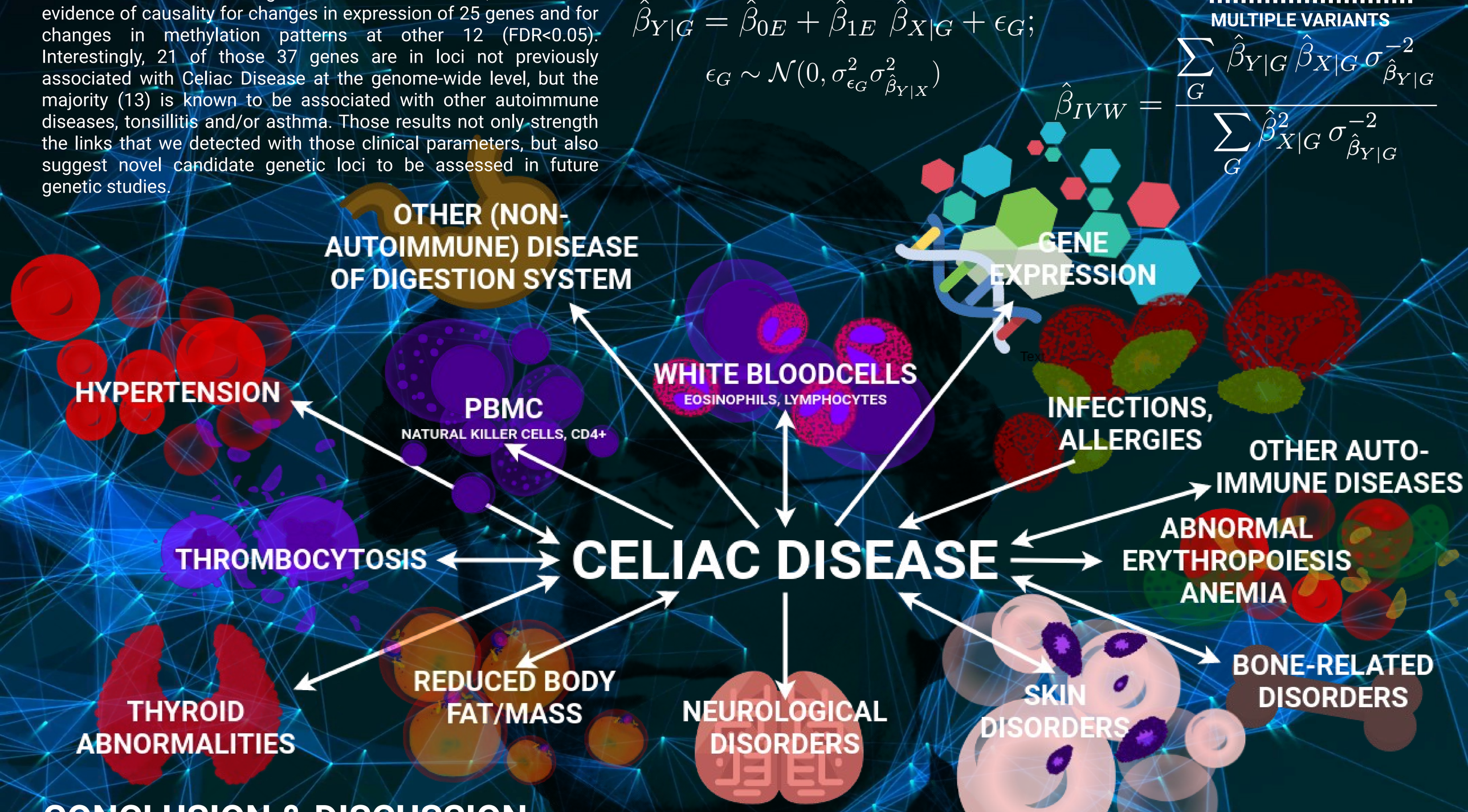
SINGLE VARIANT

MULTIPLE VARIANTS

$$\hat{\beta}_{IVW} = \frac{\sum_G \hat{\beta}_{Y|G}\,\hat{\beta}_{X|G}\,\sigma^{-2}_{\hat{\beta}_{Y|G}}}{\sum_G \hat{\beta}^2_{X|G}\,\sigma^{-2}_{\hat{\beta}_{Y|G}}}$$

RESULTS

We identified 1133 significant (FDR < 0.05) clinical parameters that cause or protect for celiac disease, and others that show feedback mechanisms . For example autoimmune diseases such as Type 1 Diabetes (T1D), have been largely described to co-occur with celiac disease.[6,7] We also confirmed a causative effect of eosinphils (and protective of lymphocytes) in increasing the risk of celiac disease, which was recently described using the same approach [8]. Here we also observed that there is feedback mechanisms, with the disease also affecting, in turn, the number of eosinophils and lymphoctyes. Finally, in addition to parameters that are supported by previous epidemiological studies, we found an interesting causative role for infection diseases and allergies. At the molecular level, we observed evidence of causality for changes in expression of 25 genes and for changes in methylation patterns at other 12 (FDR<0.05). Interestingly, 21 of those 37 genes are in loci not previously associated with Celiac Disease at the genome-wide level, but the majority (13) is known to be associated with other autoimmune diseases, tonsillitis and/or asthma. Those results not only strength the links that we detected with those clinical parameters, but also suggest novel candidate genetic loci to be assessed in future genetic studies.



CONCLUSION & DISCUSSION

We have carried out a systematic application of Mendelian Randomization to hundred of thousands of phenotypes, making this study the largest of its kind. We have identified causal and protective factors for celiac disease, recapitulating knowledge from epidemiological studies and also highlighting new key players. We expect that application to other diseases can bring new insights in the understanding of pathophysiology as we well as of biological mechanism, especially as more genome-wide association studies on -omics measurements will be available in the future.

REFERENCES. [1]Robert Di Niro, et al. Nature Medicine (2012); [2]Gujral, N. et al. World journal of Gastroenterology (2012); [3]David M. Evans, et al. Human Genetics (2014); [4]Debbie A. Lawlor. Journal of Epidemiology (2016); [5]Stephan Burgess, et al. Eur J Epidemiol (2017); [6]William Hagopian, et al. Pediatrics (2017); [7]Festen EAM, et al. PloS Genetics (2011); [8]William J. Astle, et al. Cell (2016).

Hanze University of Applied Sciences Groningen

umcg

https://bitbucket.org/MatthijsKnigge/mendelianRandomization

ma.knigge@st.hanze.nl. m.knigge@umcg.nl