

# Project background

# Multi-omics of Celiac in Down syndrome

- Celiac disease (CD) is highly influenced by genetics.
- Most (but not all!) genetic risk for CD is conferred by variation in the HLA-DQ region (chr6, genes HLA-DQA1 and HLA-DQB1).
- Sharp et al., 2019 developed a genetic risk score (GRS) for Celiac disease that accurately predicts CD using a combination of *both* HLA-DQ genotypes *and* non-HLA-DQ SNPs.

Received: 10 December 2019 | First decision: 5 January 2020 | Accepted: 10 May 2020  
DOI: 10.1111/apt.15826

AP<sub>3</sub>T Alimentary Pharmacology & Therapeutics WILEY

## A single nucleotide polymorphism genetic risk score to aid diagnosis of coeliac disease: a pilot study in clinical care

Seth A. Sharp<sup>1</sup>  | Samuel E. Jones<sup>1</sup>  | Robert A. Kimmitt<sup>2</sup> | Michael N. Weedon<sup>1</sup>  |  
Anne M. Halpin<sup>3</sup> | Andrew R. Wood<sup>1</sup>  | Robin N. Beaumont<sup>1</sup>  | Seema King<sup>4</sup> |  
David A. van Heel<sup>5</sup>  | Patricia M. Campbell<sup>3</sup> | William A. Hagopian<sup>6</sup>  |  
Justine M. Turner<sup>4</sup> | Richard A Oram<sup>1,2</sup> 

<sup>1</sup>Institute of Biomedical and Clinical Science,  
University of Exeter Medical School, Exeter, UK  
<sup>2</sup>Royal Devon & Exeter, NHS Foundation  
Trust, Exeter, UK

<sup>3</sup>Division of Nephrology and Transplant  
Immunology, University of Alberta,  
Edmonton, AB, Canada

<sup>4</sup>Faculty of Medicine and Dentistry, Pediatrics,  
University of Alberta, Edmonton, AB, Canada

<sup>5</sup>Blizard Institute, Barts and the London  
School of Medicine and Dentistry, Queen  
Mary University of London, London, UK

<sup>6</sup>Pacific Northwest Research Institute,  
Seattle, WA, USA

### Correspondence

Richard A. Oram, RILD Level 3, University  
of Exeter Medical School, Royal Devon and  
Exeter NHS Foundation Trust, Barrack Road,  
Exeter, EX2 5DW UK.  
Email: r.oram@exeter.ac.uk

### Funding information

This research was supported by funding from  
the Women and Children's Health Research  
Institute, University of Alberta. R.A.O. is  
supported by a Diabetes UK Harry Keen  
Fellowship (16/0005529). S.S. is supported by  
a Diabetes UK PhD studentship (17/0005757).  
W.A.H. is supported by Grant U01DK063829.  
M.N.W. is supported by the Wellcome Trust  
Institutional Support Fund (WT097835MF).  
The funding was provided by the Wellcome  
Trust (084743 to D.A.v.H.), by grants from the  
Coeliac Disease Consortium and an Innovative  
Cluster approved by the Netherlands  
Genomics Initiative. This research utilised  
resources supported by National Institutes of  
Health grant U01-DK062418.

### Summary

**Background:** Single nucleotide polymorphism-based genetic risk scores (GRS) model genetic risk as a continuum and can discriminate coeliac disease but have not been validated in clinic. Human leukocyte antigen (HLA) DQ gene testing is available in clinic but does not include non-HLA attributed risk and is limited by discrete risk stratification.

**Aims:** To accurately characterise both HLA and non-HLA coeliac disease genetic risk as a single nucleotide polymorphism-based GRS and evaluate diagnostic utility.

**Methods:** We developed a 42 single nucleotide polymorphism coeliac disease GRS from a European case-control study (12 041 cases vs 12 228 controls) using HLA-DQ imputation and published genome-wide association studies. We validated the GRS in UK Biobank (1237 cases) and developed direct genotyping assays. We tested the coeliac disease GRS in a pilot clinical cohort of 128 children presenting with suspected coeliac disease.

**Results:** The GRS was more discriminative of coeliac disease than HLA-DQ stratification in UK Biobank (receiver operating characteristic area under the curve [ROC-AUC] = 0.88 [95% CIs: 0.87-0.89] vs 0.82 [95% CIs: 0.80-0.83]). We demonstrated similar discrimination in the pilot clinical cohort (114 cases vs 40 controls, ROC-AUC = 0.84 [95% CIs: 0.76-0.91]). As a rule-out test, no children with coeliac disease in the clinical cohort had a GRS below 38th population centile.

**Conclusions:** A single nucleotide polymorphism-based GRS may offer more effective and cost-efficient testing of coeliac disease genetic risk in comparison to HLA-DQ stratification. As a comparatively inexpensive test it could facilitate non-invasive coeliac disease diagnosis but needs detailed assessment in the context of other diagnostic tests and against current diagnostic algorithms.

# Multi-omics of Celiac in Down syndrome

- At the Linda Crnic Institute for Down Syndrome, we applied the Sharp et al. GRS to a cohort of individuals with **Trisomy 21 +/- Celiac disease**.
- We evaluated...
  - How does the GRS perform in individuals with Down syndrome (DS), as compared to the typical population?
  - How do the individual components in the GRS relate to CD in DS?
  - Can we improve the performance of this CD GRS in DS with the addition of protein biomarkers?

medRxiv preprint doi: <https://doi.org/10.1101/2022.09.27.22280436>; this version posted September 29, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

**Title:** Multi-Omics Assessment of Genetic Risk for Celiac Disease in Down Syndrome

**Authors:** Marisa Gallant Stahl<sup>1,2</sup> MD MSCS, Jessica R Shaw<sup>2</sup>, Neetha Paul Eduthan<sup>2</sup>, Angela L Rachubinski<sup>2</sup>, Keith P Smith<sup>2</sup>, Belinda Enriquez Estrada<sup>2</sup>, Matthew D Galbraith<sup>2,3</sup>, Ronald J Sokol<sup>1</sup> MD, Sameer Chavan<sup>4</sup>, Laura Ann Leaton<sup>4,5</sup>, Katherine M Kichula<sup>4,5</sup>, Paul J Norman<sup>4,5</sup> PhD, Jill M Norris<sup>6</sup> PhD MPH, Edwin Liu<sup>1</sup> MD & Joaquin M Espinosa<sup>2,3</sup> PhD

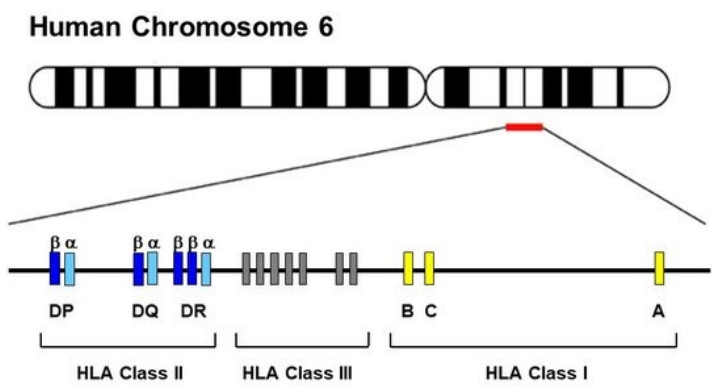
**Affiliations:**

1. Digestive Health Institute, Children's Hospital Colorado, Department of Pediatrics, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States.
2. Linda Crnic Institute for Down Syndrome, University of Colorado, Aurora, CO
3. Department of Pharmacology, University of Colorado Anschutz Medical Campus.
4. Department of Biomedical Informatics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
5. Department of Immunology and Microbiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
6. Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, United States



# Primer on HLA genetics of Celiac disease

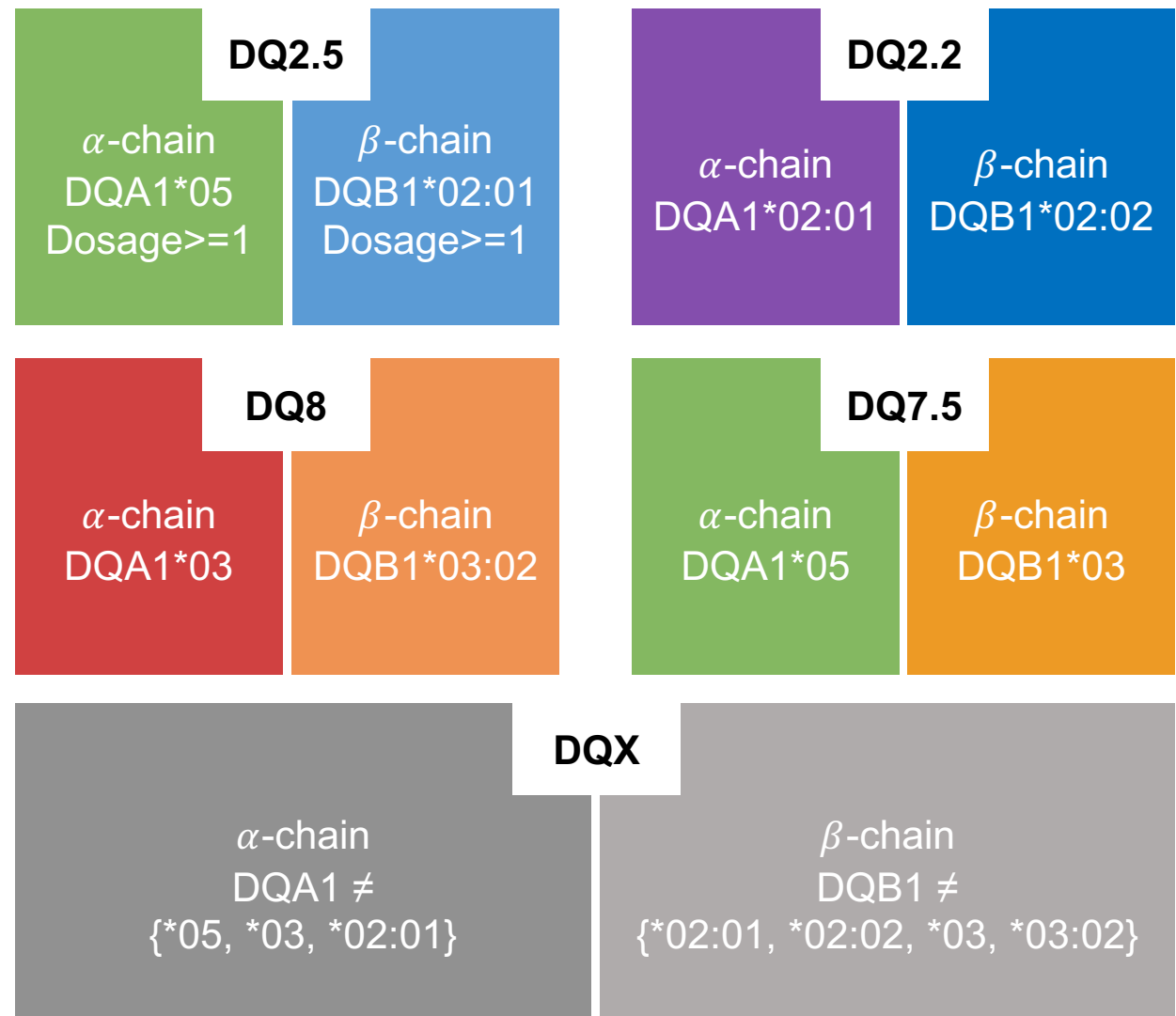
# Pairs of HLA-DQ alleles form HLA-DQ heterodimers



## Alleles encoding CD-associated heterodimers

Heterodimer	DQA1* allele	DQB1* allele
DQ2.5	05	0201
DQ2.2	0201	0202
DQ8	03	0302
DQ7.5	05	03

Subset of many possible HLA alleles at DQA1 and DQB1

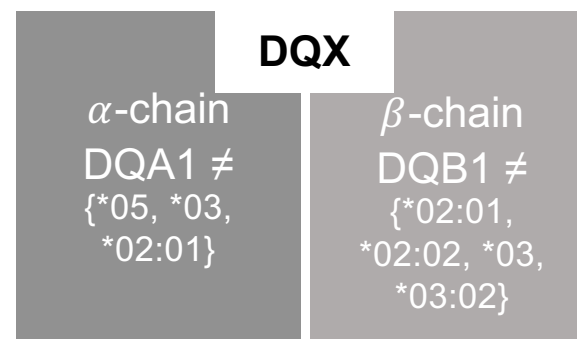
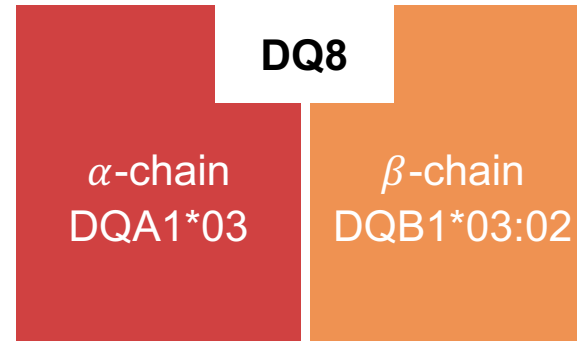
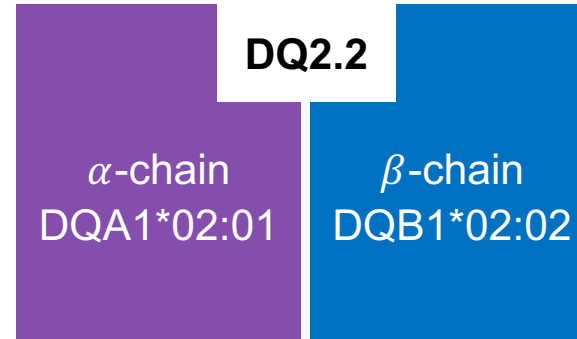
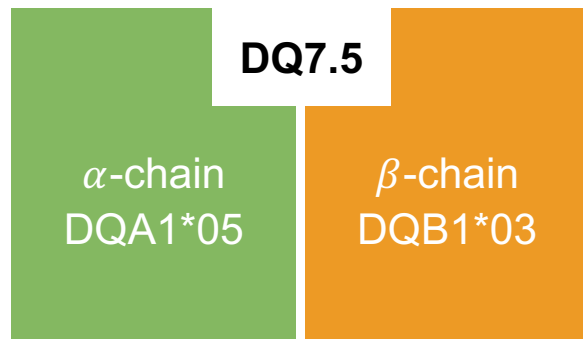
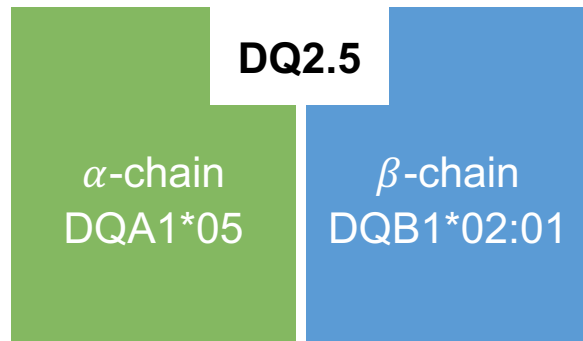
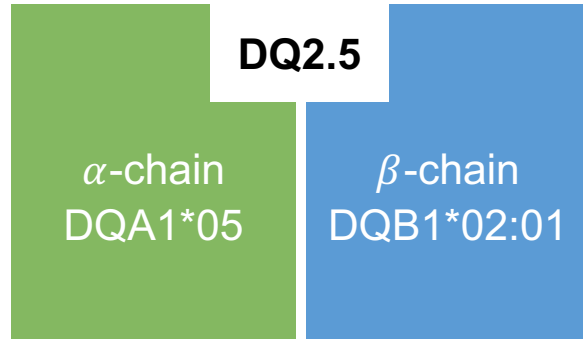


## Pairs of HLA-DQ heterodimers form HLA-DQ genotypes

For example:

## Chromosome 6, paternal

## Chromosome 6, maternal



HLA-DQ genotype	
	DQ2.5/DQ2.5
	DQ2.5/DQ2.2
	DQ7.5/DQ2.2
	DQ2.5/DQ8
	DQ2.5/X
	DQ7.5/DQ2.5
	DQ8/DQ8
	DQ2.2/DQ8
	DQ2.2/DQ2.2
	DQ7.5/DQ8
	DQ8/X
	DQ2.2/X
	DQ7.5/X
	DQ7.5/DQ7.5
	X/X

# Analysis overview

# GRS analysis overview

## Key outputs

GRS

**Logistic**

**regression:** Celiac (0/1) ~

GRS

- ROC curve
- AUC for comparison to Sharp et al., 2019.

Additive genetic models

**Logistic**

**regression:** Celiac (0/1) ~

HLA-DQ **genotype** dosage  
(0/1)

+

PC1

+

...

+

PC5

- OR barplots
- Frequency stacked barplots
- Volcano plot
- Comparison to Sharp et al., 2019

**Logistic**

**regression:** Celiac (0/1) ~

GRS risk **allele** dosage  
(0/1/2)

+

PC1

+

...

+

PC5

- OR barplots
- Volcano plot
- Comparison to Sharp et al., 2019

**Logistic**

**regression:** Celiac (0/1) ~

HLA-DQ **allele group**  
dosage (0/1/2)

+

PC1

+

...

+

PC5

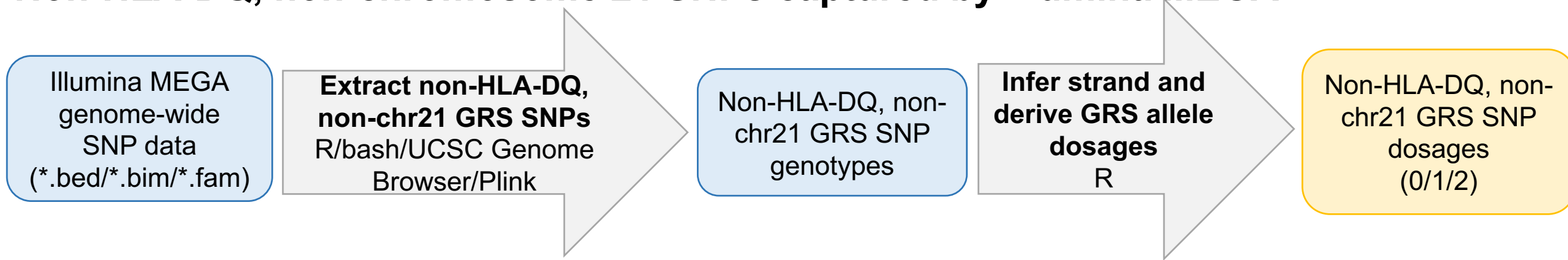
- OR barplots
- Volcano plot
- Comparison to Sharp et al., 2019



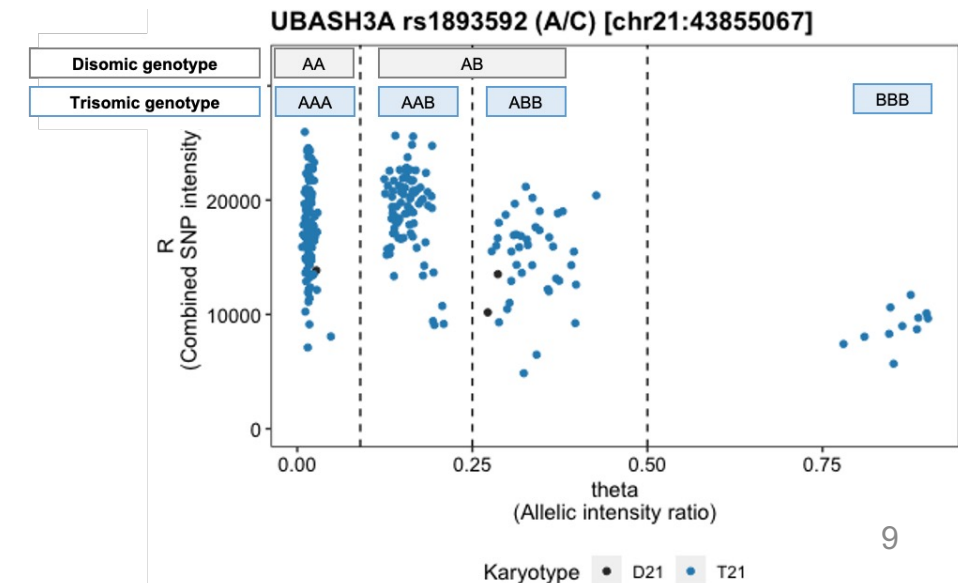
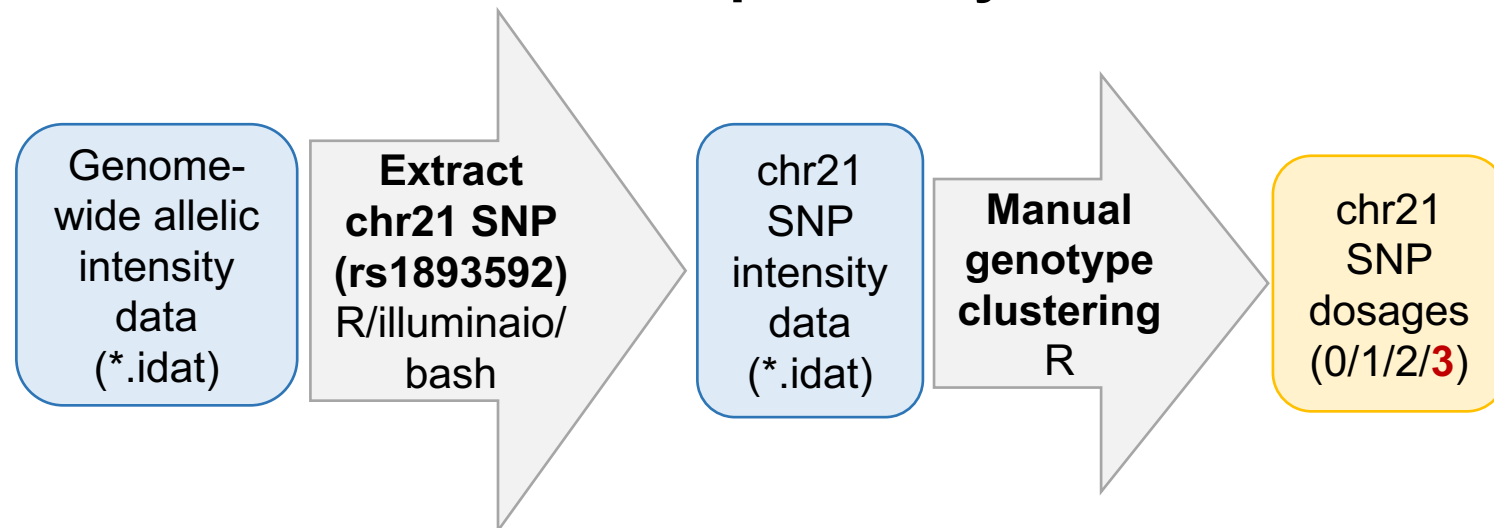
# Analysis data pipeline

## GRS component: Non-HLA-DQ SNPs

### Non-HLA-DQ, non-chromosome 21 SNPs captured by Illumina MEGA



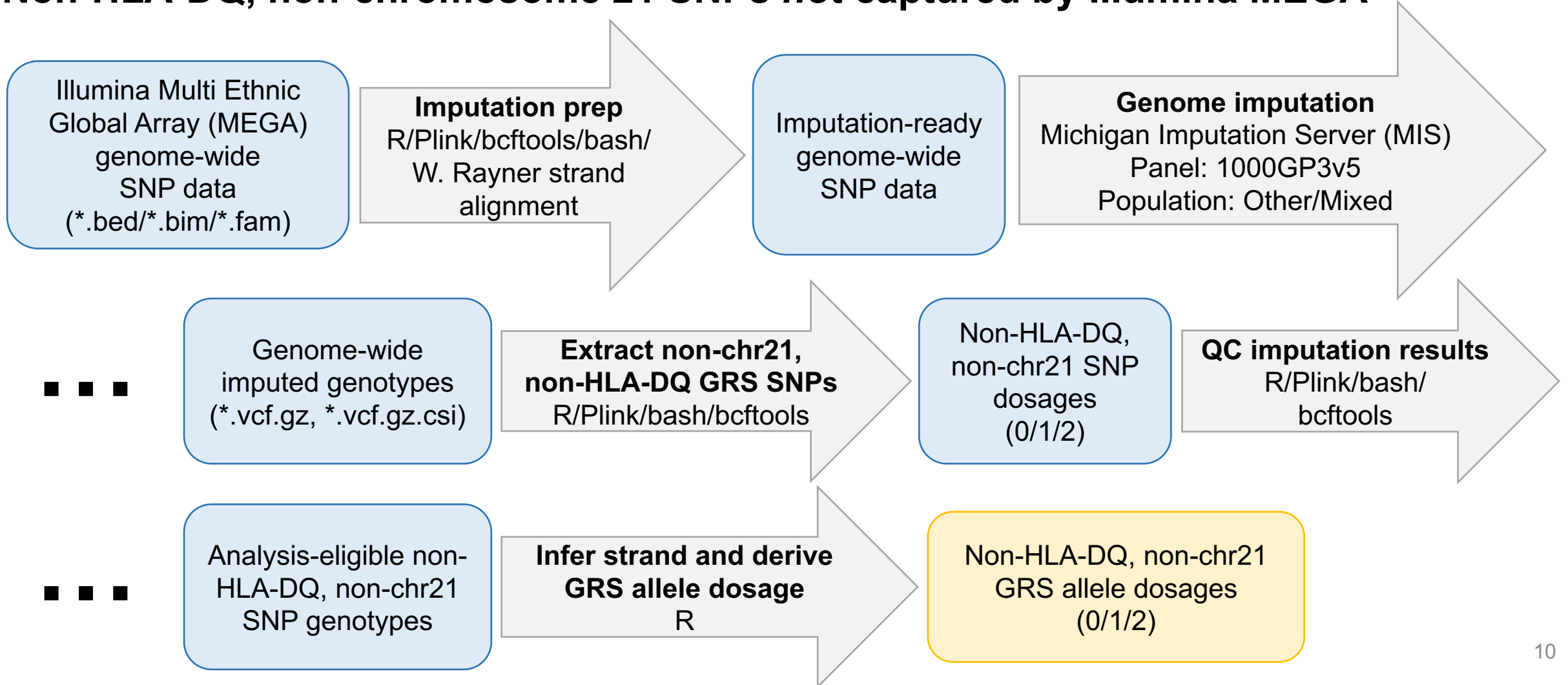
### Chromosome 21 SNP captured by Illumina MEGA



# Analysis data pipeline

## GRS component: Non-HLA-DQ SNPs (*continued*)

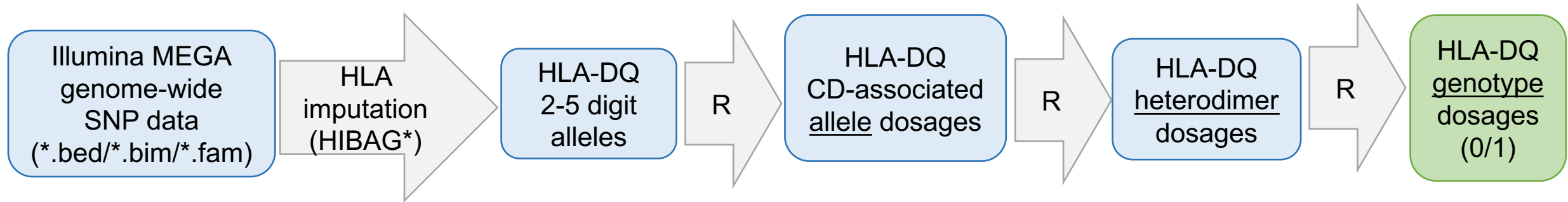
### Non-HLA-DQ, non-chromosome 21 SNPs *not* captured by Illumina MEGA



# Analysis data pipeline

GRS component: HLA-DQ genotypes

HLA-DQ genotypes *not* captured by Illumina MEGA



\*Performed by Paul Norman et al.

# Analysis data pipeline

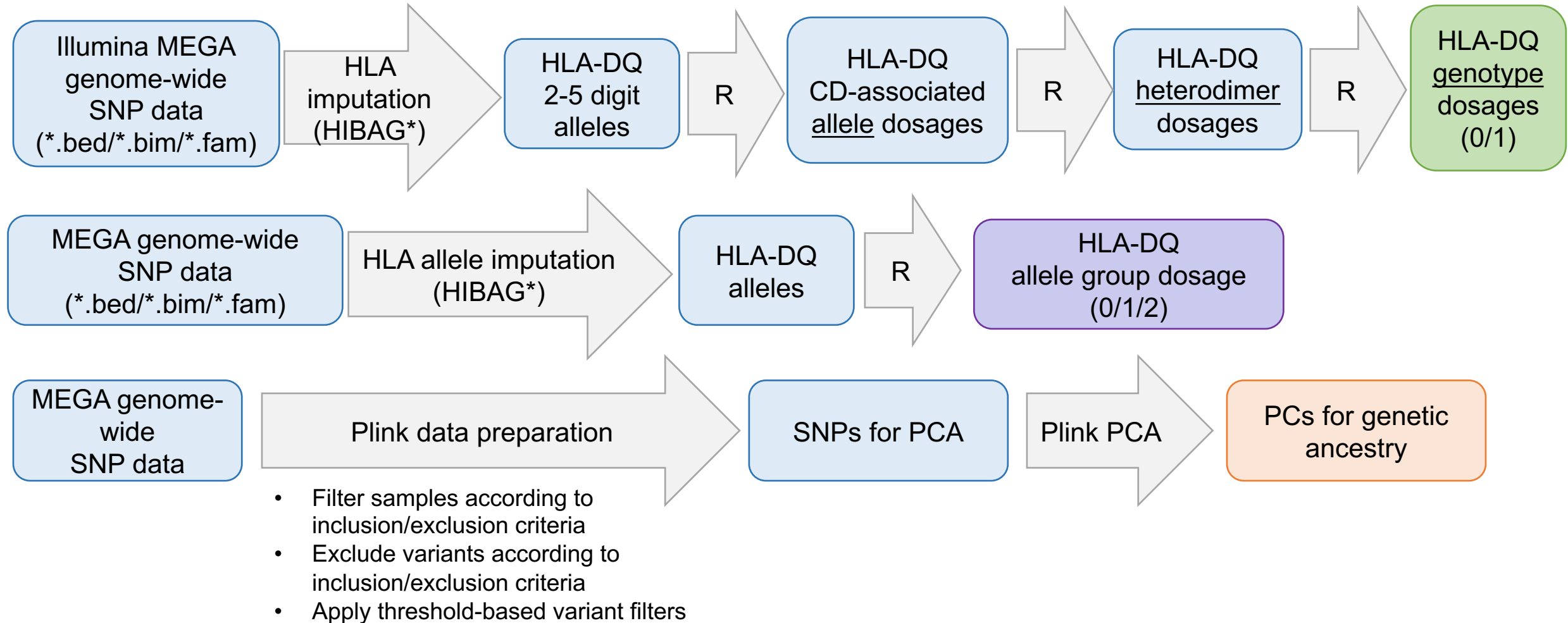
## Celiac GRS calculation

- Weight for each HLA-DQ genotype given in Sharp et al., 2019, Table S2
- Weights for each non-HLA-DQ SNP genotype given in Sharp et al., 2019 Table S3
- Score allele dosage extracted from MEGA genotype data (dosage 0/1/2), extracted from genome-wide imputation results (dosage 0/1/2), or derived after manual genotype clustering to account for trisomy 21 (dosage 0/1/2/3)

$$\begin{aligned}
 \text{GRS} = & \text{4.97} \times \begin{array}{c} \text{DQ2.5/DQ2.5} \\ \text{dosage} \\ (0/1) \end{array} + \text{4.96} \times \begin{array}{c} \text{DQ2.5/DQ2.2} \\ \text{dosage} \\ (0/1) \end{array} + \dots + \text{0.0} \times \begin{array}{c} \text{X/X} \\ \text{dosage} \\ (0/1) \end{array} + \\
 & \text{0.425} \times \begin{array}{c} \text{rs4143332 (C)} \\ \text{dosage} \\ (0/1/2) \end{array} + \text{0.372} \times \begin{array}{c} \text{rs3128927 (T)} \\ \text{dosage} \\ (0/1/2) \end{array} + \dots + \\
 & \text{-0.128} \times \begin{array}{c} \text{rs1893592 (C)} \\ \text{dosage} \\ (0/1/2/3) \end{array} + \dots + \\
 & \text{-0.083} \times \begin{array}{c} \text{rs10167650 (G)} \\ \text{dosage} \\ (0/1/2) \end{array}
 \end{aligned}$$

# Analysis data pipeline

## Additional data for variant-level analyses



\*Performed by Paul Norman et al.

# Selected results/figures

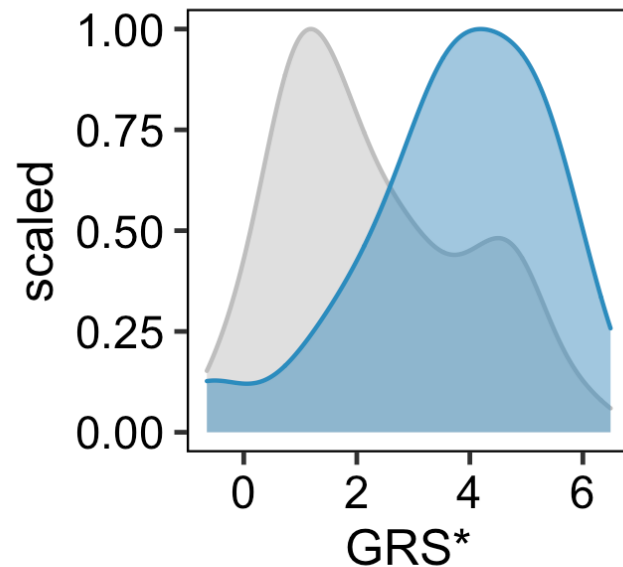
# GRS performance in DS

(GRS version: Sharp et al., 2022)

## Celiac GRS\* in DS +/- CD

OR = 1.62 (1.22, 2.22)

P<0.0014

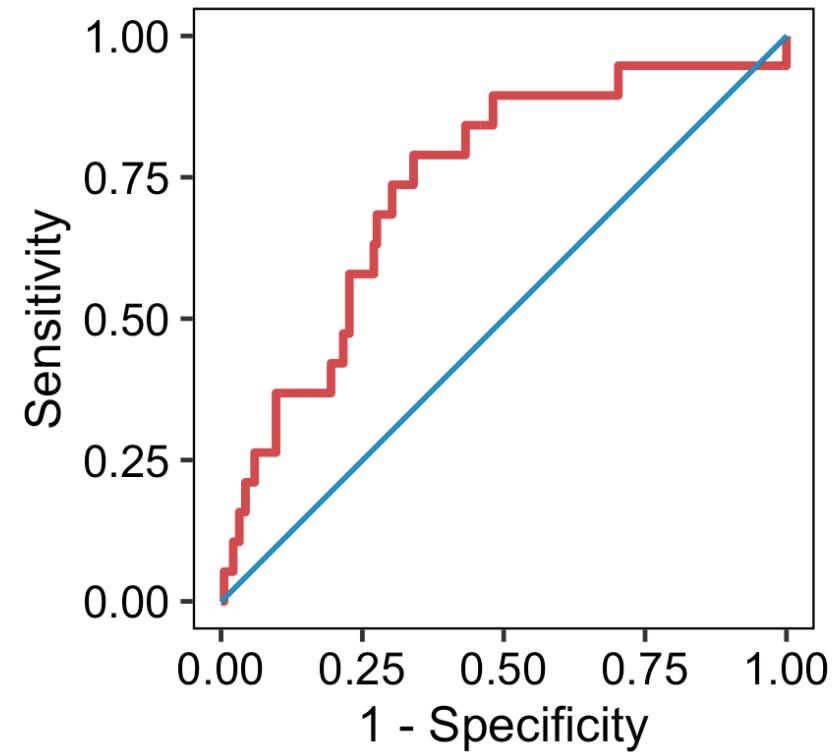


DS without CD DS with CD

\*Sharp et al., 2022

## Predictive accuracy of GRS\* in DS

Area under the curve: 0.7354

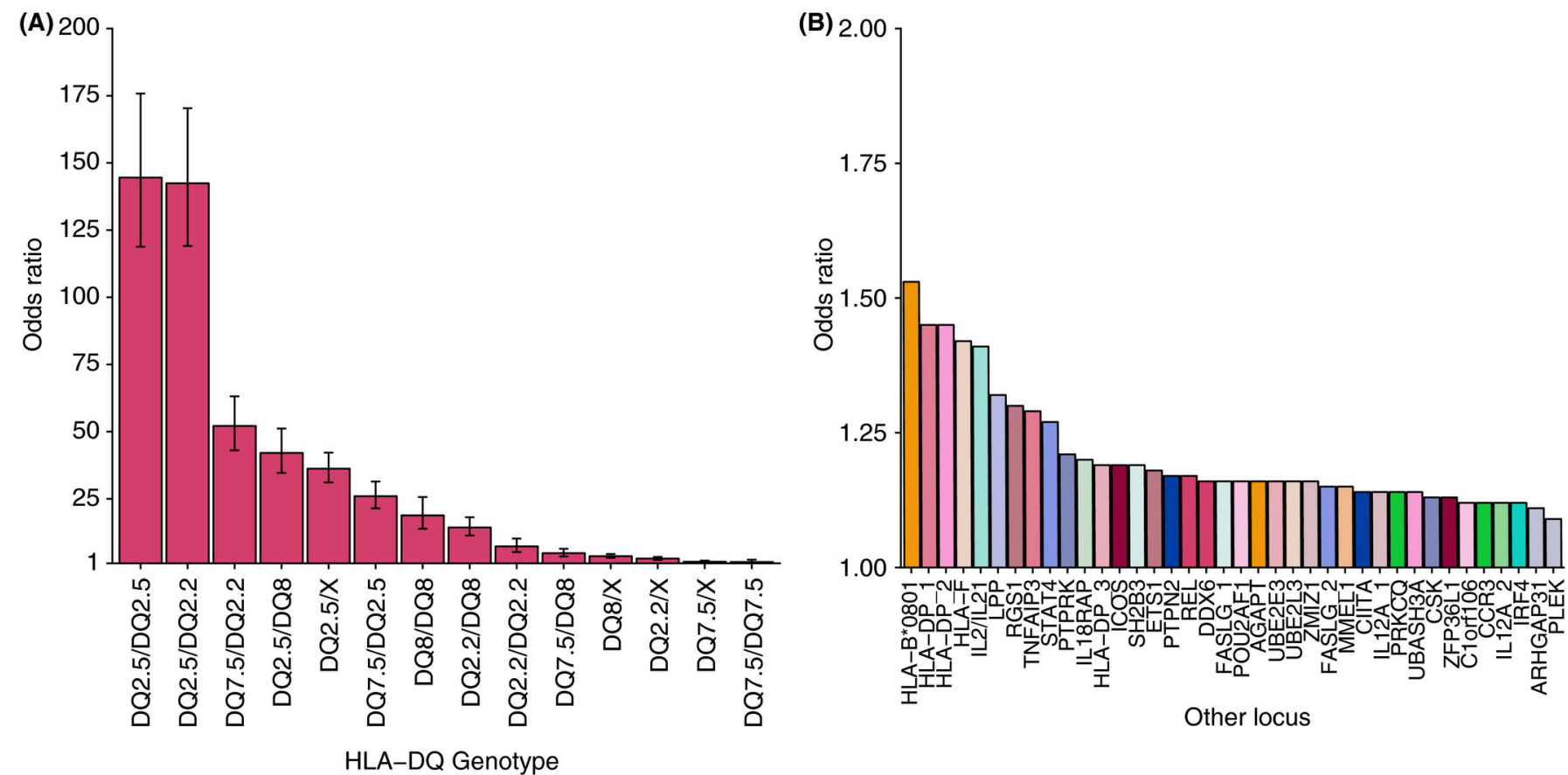


Logistic: Celiac ~ GRS\*

\*Sharp et al., 2022

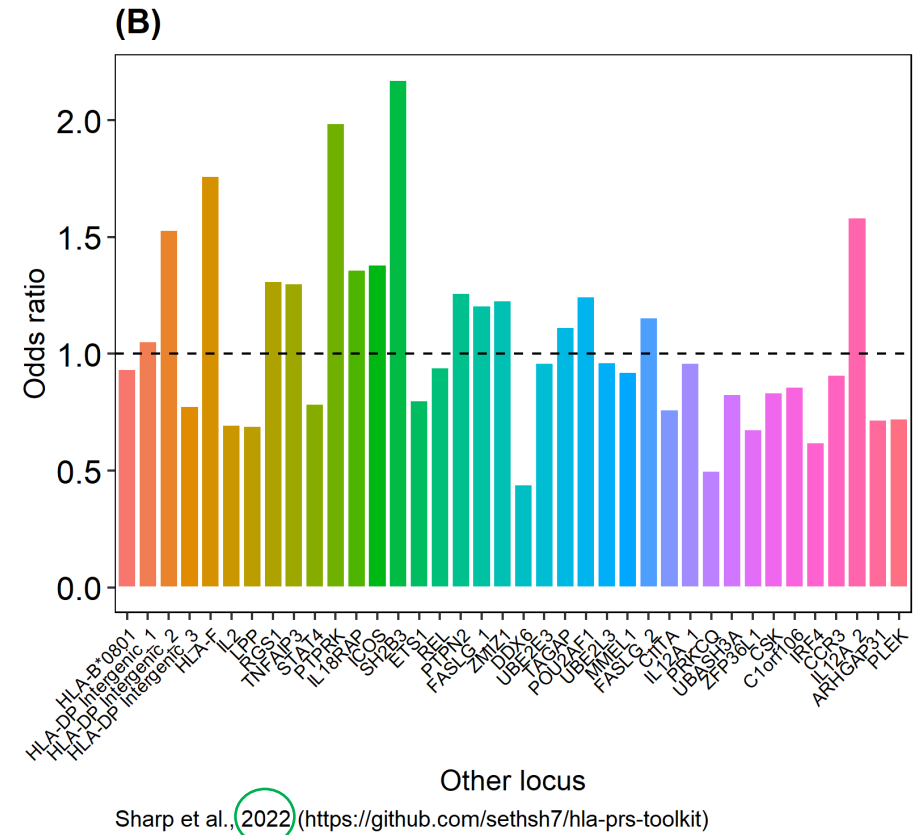
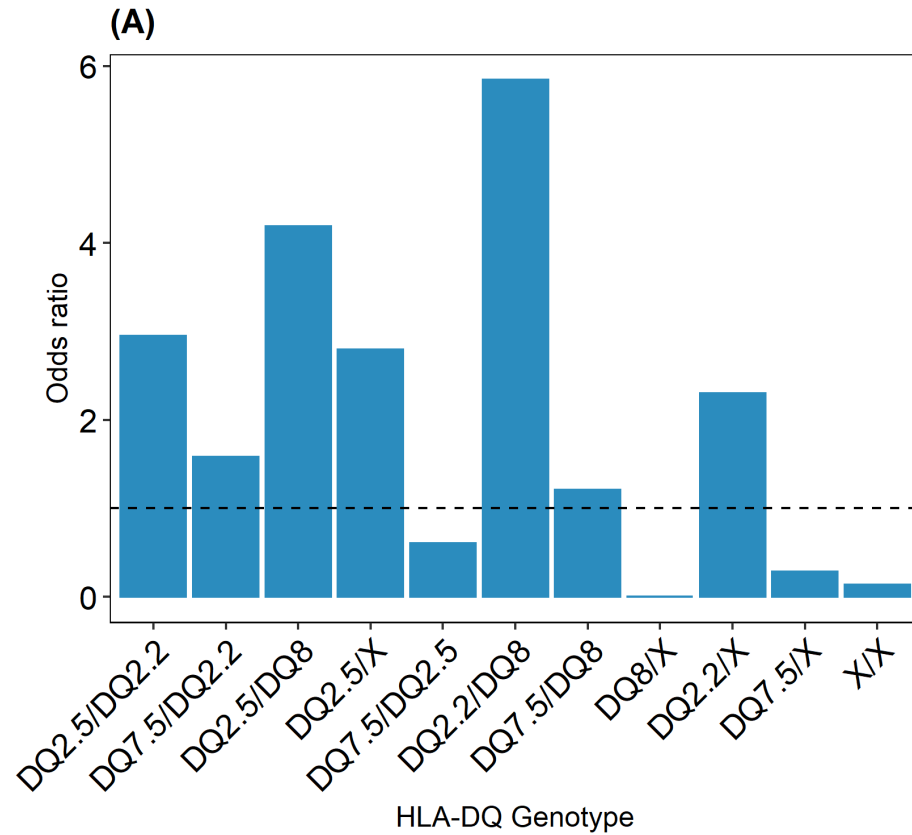
# Figure 1 of Sharp et al., 2019

A single nucleotide polymorphism genetic risk score to aid diagnosis of coeliac disease: a pilot study in clinical care

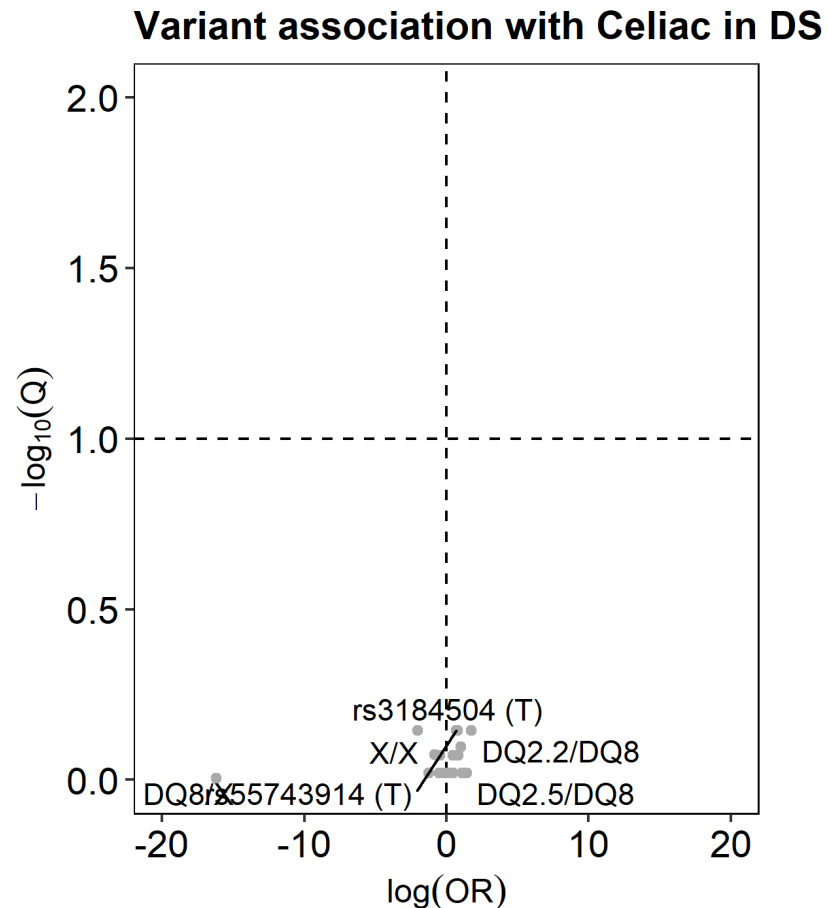




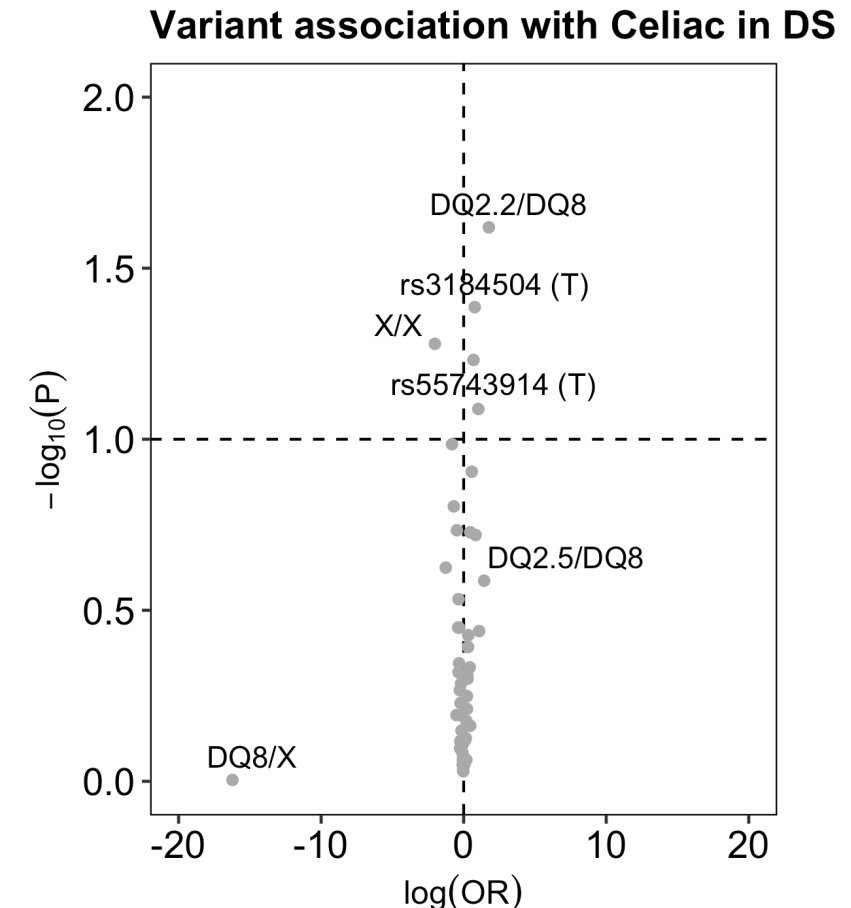
# Odds ratios observed in the HTP Revised GRS (Sharp et al., 2022)



# Results of single-variant logistic regression models for CD (0/1) versus dosage of score allele (GRS version: Sharp et al., 2022)



+ PC1 + PC2 + PC3 + PC4 + PC5, family = binomial(link = "logit"), data = .)



+ PC1 + PC2 + PC3 + PC4 + PC5, family = binomial(link = "logit"), data = .)

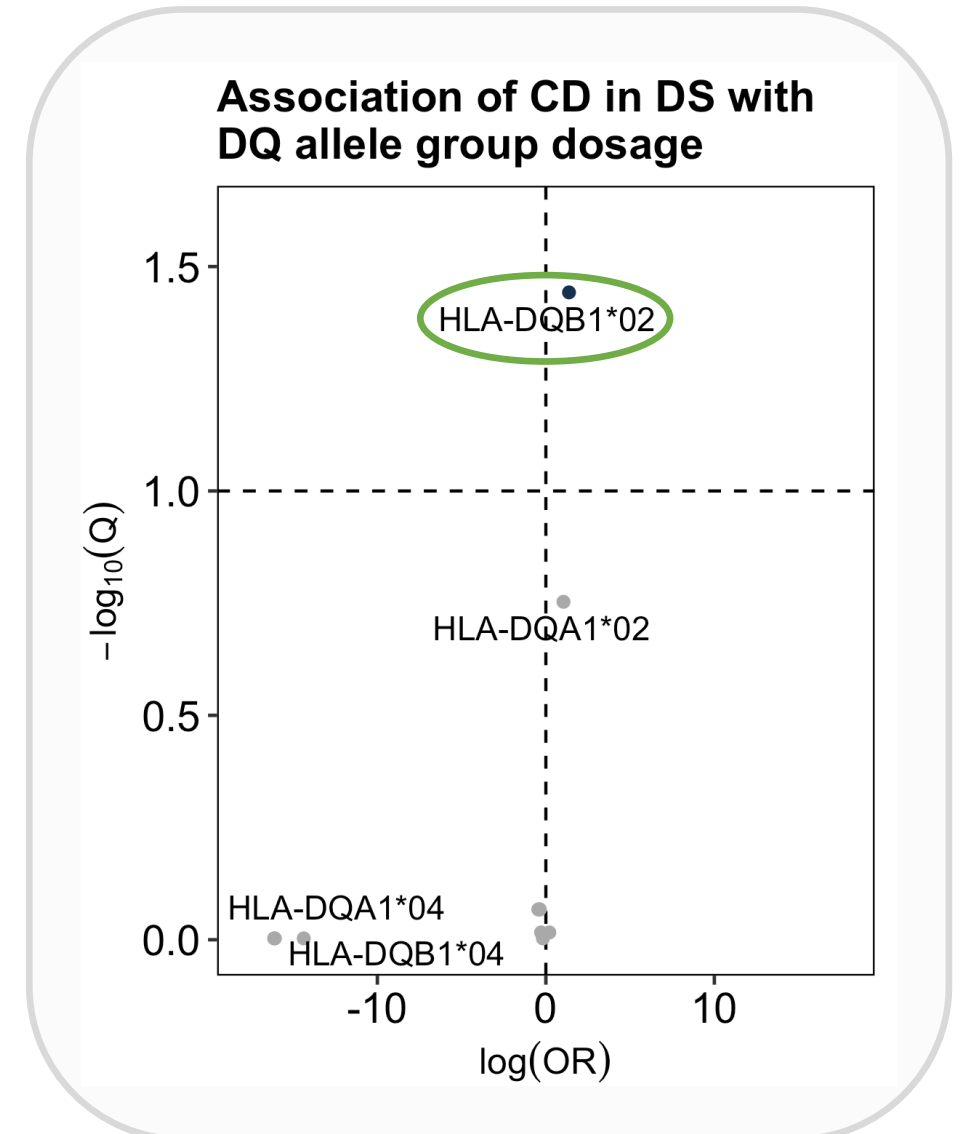
# Dosage of DQB1\*02 alleles is significantly associated with CD in DS

## Alleles encoding CD-associated heterodimers

Heterodimer	DQA1*	DQB1*
DQ2.5	05	0201
DQ2.2	0201	0202
DQ8	03	0302
DQ7.5	05	03

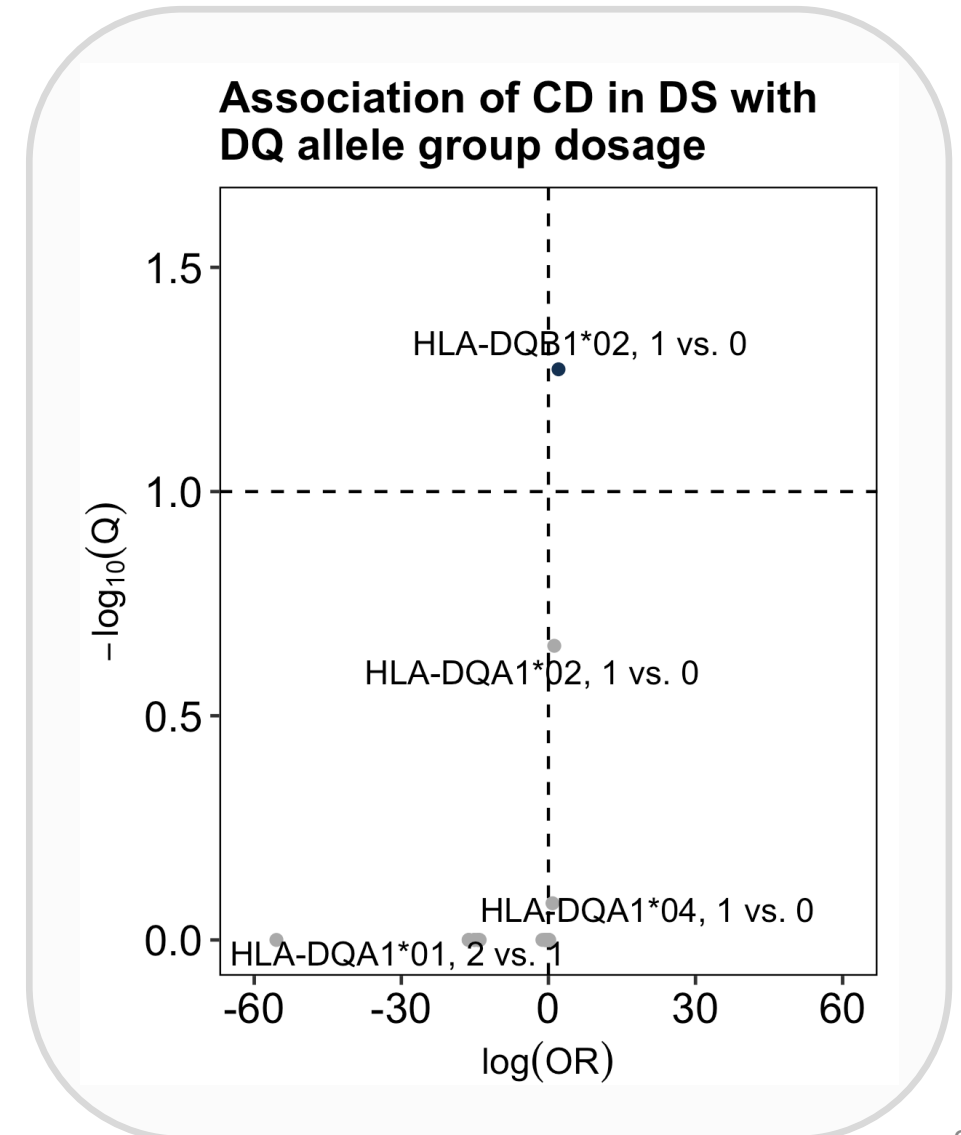
HLA-DQB1\*02

- In an **additive genetic model**, dosage of **DQB1\*02** alleles (e.g., DQB1\*02:01, DQB1\*02:02) was significantly associated with Celiac disease in Down syndrome (**OR = 3.95**, **Q<0.0361**).
- On average, each additional **DQB1\*02** allele corresponded to a **3.95-fold** increase in the odds of **CD in DS**.



# Possible dominant association of DQB1\*02 with a 7.81-fold increase in the odds of CD in DS

- To identify non-additive associations (i.e., dominant, recessive), we performed a follow-up analysis of DQ allele group dosage as a categorical predictor, rather than numeric.
- **Carrying one DQB1\*02 allele was associated with a 7.81-fold increase in the odds of CD in DS (OR = 7.81,  $Q < 0.0534$ ), as compared to carrying zero DQB1\*02 alleles.**
- By contrast, we found no significant difference in the odds of Celiac given two **DQB1\*02** alleles compared to one (OR = 0.80,  $Q < 0.9999$ ).
- **These results may indicate a dominant effect of DQB1\*02 in the pathogenesis of CD in individuals with DS (OR = 7.81,  $Q < 0.0534$ ).**



# Dominant association of DQB1\*02 with a 7.81-fold increase in the odds of Celiac disease in Down syndrome

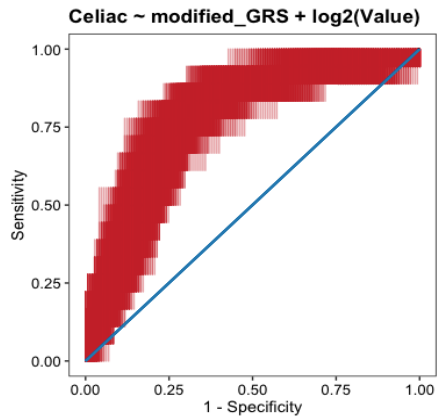
- 84.2% of Celiac cases (16/19) carried one or more DQB1\*02 allele, with the DQ genotypes:
  - DQ2.5/X (n=5)
  - DQ2.2/X (n=4)
  - DQ2.2/DQ8 (n=3)
  - DQ2.5/DQ2.2 (n=1)
  - DQ2.5/DQ8 (n=1)
  - DQ7.5/DQ2.2 (n=1)
  - DQ7.5/DQ2.5 (n=1)
- Only three Celiac cases were not carriers of DQB1\*02, instead presenting with the genotypes:
  - DQ7.5/DQ8 (n=1)
  - DQ7.5/X (n=1)
  - X/X (n=1)

Alleles encoding traditionally CD-associated heterodimers

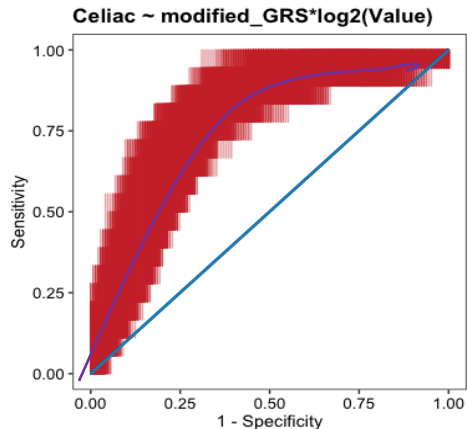
Heterodimer	DQA1*	DQB1*
DQ2.5	05	0201
DQ2.2	0201	0202
DQ8	03	0302
DQ7.5	05	03

# Models involving both GRS and single protein abundance improve prediction of celiac disease in Down syndrome (Please email [Jessica.rose.shaw@gmail.com](mailto:Jessica.rose.shaw@gmail.com) for details)

Prediction of celiac status using GRS\* and abundance of any one protein as precision covariate



Prediction of celiac status with interaction of GRS\* and abundance of any one protein



AUC values

UID <chr>	Celiac ~ modified_GRS <dbl>	Celiac ~ modified_GRS + log2(Value) <dbl>
WAP four-disulfide core domain protein 6 (WFDC6.13412.5)	75.2	84.5
4-hydroxy-2-oxoglutarate aldolase, mitochondrial (HOGA1.10024.44)	75.2	83.9
Tumor necrosis factor receptor superfamily member 17 (TNFRSF17.2665.26)	75.2	83.7
Membrane-associated progesterone receptor component 2 (PGRMC2.8681.93)	75.2	83.5
Unique cartilage matrix-associated protein (UCMA.10977.55)	75.2	83.1
Zinc finger protein 134 (ZNF134.12787.47)	75.2	82.9
Marginal zone B- and B1-cell-specific protein (PACAP.16322.10)	75.2	82.8
Furin (FURIN.6276.16)	75.2	82.7
GRB2-related adapter protein (GRAP.12820.1)	75.2	82.6
Brain-specific angiogenesis inhibitor 1-associated protein 2 (BAIAP2.19331.18)	75.2	82.3

AUC values

UID <chr>	Celiac ~ modified_GRS <dbl>	Celiac ~ modified_GRS*log2(Value) <dbl>
SUN domain-containing protein 3 (SUN3.8852.10)	75.2	86.7
F-box/LRR-repeat protein 5 (FBXL5.12846.3)	75.2	85.7
Regulator of G-protein signaling 4 (RGS4.17855.28)	75.2	85.4
Macrophage-capping protein (CAPG.4968.50)	75.2	85.2
Tumor necrosis factor receptor superfamily member 17 (TNFRSF17.2665.26)	75.2	85.2
Polyadenylate-binding protein 3 (PABPC3.10447.18)	75.2	84.6
Polyadenylate-binding protein 4 (PABPC4.12603.87)	75.2	84.6
WAP four-disulfide core domain protein 6 (WFDC6.13412.5)	75.2	84.6
Mitogen-activated protein kinase 14 (MAPK14.5007.1)	75.2	84.3
ADP-dependent glucokinase (ADPGK.6221.1)	75.2	84.2

# Questions?