Supplementary Note

***HFE* and celiac disease**

Butterworth et al. (commentary by Bowlus and Lie *Gastroenterology* 2003; Ravine and Darke *Gastroenterology* 2003) suggested that *HFE* gene mutations are common in individuals with celiac disease and are in linkage disequilibrium with different HLA alleles as compared to individuals without celiac disease; it was suggested that *HFE* gene mutations may provide a survival advantage in celiac disease by reducing iron deficiency.

If mutations within *HFE* affect the probability to have celiac disease (independent of their effect on systemic iron status), the exclusion restriction assumption of mendelian randomization is violated. Notably, if variants within *HFE* are in strong LD with various HLA alleles associated with celiac disease, as was suggested (see 2 related commentaries for weaknesses in Butterworth et al. and potential alternative explanations for the observations made); this could result in a spurious association of *HFE* mutations and celiac disease. Specifically, Butterworth et al. studied the *HFE* mutations, C282Y (also known as rs1800562) and H63D (also known as rs1799945), both of which are instruments in our study.

To test whether the genetic instruments within *HFE* drove our association, we ran analyses excluding both instruments within *HFE;* results appeared consistent (Figure 1b), suggesting the SNPs within *HFE* did not drive our association.

Using each of the 4 biomarkers to quantify systemic iron status

We primarily used serum iron levels to quantify the genetic associations of the instruments with systemic iron status; analyses using the other 3 biomarkers are shown in Supplementary Table 5. All I_{GX}^2 statistics are sufficiently large.

When using ferritin levels to quantify the associations, we found significant heterogeneity (Q=8.6, p = 0.036), thereby invalidating the IVW estimate; using the WM estimator we found no significant relationship between iron status and risk of celiac disease (OR: 0.74, 95% CI 0.38-1.5, P = 0.4; Supplementary Table 5). We note that the power to detect an OR of 0.8 using ferritin as the quantifier of associations was low (11% power; Supplementary Table 6).

When using transferrin saturation levels to quantify the associations, which had a higher power to detect an OR of 0.8 (49% power; Supplementary Table 6), we found a significant relationship between iron status and risk of celiac disease (OR: 0.74, 95% CI 0.56-0.99, P = 0.04; Supplementary Table 5).

When using total iron-binding capacity to quantify the associations, we found significant heterogeneity (Q=12, p = 0.009) as well as a significant MR-Egger intercept (P=0.041), thereby invalidating the IVW estimate. Using the WM estimator, we found no significant relationship between iron status and risk of celiac disease (OR: 1.2, 95% CI 0.89-1.5, P = 0.28; Supplementary Table 5). The power to detect an OR of 1.2 using variants that explain 3.01% in variance of the exposure (similar to total iron-binding capacity) is 32% (Supplementary Table 6).

References:

Butterworth J, Cooper B, Rosenberg W, et al. The role of hemochromatosis susceptibility gene mutations in protecting against iron deficiency in celiac disease. *Gastroenterology*. 2002;123:444-449.

Bowlus, C.L. and Lie, B.A.. Discussion on the Role of Hemochromatosis Susceptibility Gene Mutation in Protecting Against Iron Deficiency in Celiac Disease. *Gastroenterology* 2003; 124: 1562 – 1563.

Ravine, D. and Darke, C.. Discussion on the Role of Hemochromatosis Susceptibility Gene Mutation in Protecting Against Iron Deficiency in Celiac Disease. *Gastroenterology* 2003; 124: 1563.

Supplementary Table 1: **Results from Bell et al. 2021**. For all selected genetic instruments, the position, minor and major allele, minor allele frequency, and effect on 4 relevant iron biomarkers is shown. The effect is shown for the minor allele.

Rsid	Position (hg38)	Minor: Major Allele	MAF (%)	Phenotype	Effect in SD (95% CI)	P-value
rs1799945	chr6:26090951	G:C	13.7	TSAT	0.21 (0.2; 0.23)	6.1e-229
				TIBC	-0.12 (-0.13; -0.1)	4.29e-66
				Ferritin	0.059 (0.049; 0.069)	1.51e-31
				Iron	0.17 (0.16; 0.18)	1.26e-187
rs1800562	chr6:26092913	A:G	6.77	TIBC	-0.45 (-0.47; -0.43)	1e-300
				TSAT	0.45 (0.42; 0.47)	1e-300
				Ferritin	0.13 (0.12; 0.15)	1.85e-84
				Iron	0.27 (0.26; 0.29)	3.66e-276
rs57659670	chr15:45106240	C:T	7.53	Ferritin	-0.14 (-0.16; -0.13)	1.05e-113
				Iron	-0.042 (-0.056; -0.028)	1.08e-08
				TIBC	0.077 (0.06; 0.094)	3.67e-19
				TSAT	-0.058 (-0.074; -0.041)	5.73e-12
rs855791	chr22:37066896	A:G	43.1	TSAT	-0.17 (-0.18; -0.16)	1e-300
				Iron	-0.17 (-0.18; -0.16)	1e-300
				TIBC	0.026 (0.017; 0.035)	2.88e-08
				Ferritin	-0.044 (-0.051; -0.038)	6.14e-37

MAF: minor allele frequency; SD: standard deviation; CI: confidence interval; TIBC: total iron-binding capacity; TSAT transferrin saturation.

Supplementary Table 2: **Strength of instruments.** For 4 iron biomarkers, the amount of variance explained by each instrument as well as the F-statistic is computed. As effect sizes are in units of standard deviation, the percentage of variation in iron biomarker explained by the SNP (R^2); we computed this as $2 \times \beta^2 \times AF \times (1 - AF)$. Total R^2 values are 0.65, 3.04, 3.01, and 5.07 for ferritin, iron, TIBC, and TSAT, respectively. F-statistics (F) were computed as a function of the variance explained and sample size (Palmer et al. *Stat Methods Med Res* 2012).

	Effect on exposure (units of SD)				
SNP	Estimate	SE	P-value	R^2	F
Ferritin					
rs1799945	0.059	0.005	1.51e-31	0.08	202.77
rs1800562	0.13	0.0067	1.85e-84	0.21	526.22
rs57659670	-0.14	0.0062	1.05e-113	0.27	673.67
rs855791	-0.044	0.0035	6.14e-37	0.09	233.95
Iron					
rs1799945	0.17	0.0058	1.26e-187	0.68	1125.08
rs1800562	0.27	0.0076	3.66e-276	0.92	1518.66
rs57659670	-0.042	0.0073	1.08e-08	0.02	40.18
rs855791	-0.17	0.0046	1e-300	1.42	2351.05
TIBC (total iron-binding capacity)					

rs1799945	-0.12	0.007	4.29e-66	0.34	462.72
rs1800562	-0.45	0.012	1e-300	2.56	3552.69
rs57659670	0.077	0.0086	3.67e-19	0.08	111.91
rs855791	0.026	0.0047	2.88e-08	0.03	44.92
TSAT (transferrin saturation)					
rs1799945	0.21	0.0065	6.1e-229	1.04	1385.41
rs1800562	0.45	0.012	1e-300	2.56	3448.83
rs57659670	-0.058	0.0084	5.73e-12	0.05	61.62
rs855791	-0.17	0.0046	1e-300	1.42	1890.36

Supplementary Table 3: **Reported exposure and outcome relationships.** For all 4 genetic instruments, the reported effect allele, allele frequency of effect allele (EAF), and association statistics are reported for serum iron levels and celiac disease. Bolded SNP had to be harmonized across datasets.

	Exposure (serum iron levels)					
SNP	Effect allele	Other allele	EAF	Beta	SE	P-val
rs1799945	G	C	0.1370	0.170	0.00582	1.26e-187
rs1800562	A	G	0.0677	0.270	0.00760	3.66e-276
rs57659670	C	T	0.0753	-0.042	0.00735	1.08e-08
rs855791	**A**	**G**	**0.4310**	**-0.170**	**0.00459**	**1.00e-300**
Outcome (celiac disease)						
rs1799945	G	C	0.1510	-0.1550	0.0464	0.000842
rs1800562	A	G	0.0778	-0.0393	0.0618	0.525000
rs57659670	C	T	0.0750	0.0360	0.0628	0.566000
rs855791	**G**	**A**	**0.5601**	**-0.0627**	**0.0335**	**0.061700**

Beta: effect size (units of standard deviation for exposure; log of the odds ratio for outcome); SE standard error; P-val p-value for the association.

Supplementary Table 4: Phenotypes associated with included SNPs as determined through Phenoscanner.

SNP	Associated phenotypes
rs1800562	mean corpuscular hemoglobin; disorders of mineral metabolism; iron status biomarkers; transferrin saturation; red cell distribution width; mean corpuscular hemoglobin concentration; transferrin levels; hemoglobin concentration; iron levels; reticulocyte fraction of red cells; reticulocyte count; hematocrit; mean corpuscular hemoglobin; transferrin; self-reported hereditary or genetic haematological disorder; transferrin saturation with iron; mean corpuscular volume; high light scatter percentage of red cells; ferritin levels; high light scatter reticulocyte count; alcohol consumption; transferrin glycosylation; carbohydrate deficient transferrin supplementary concept; HbA1c; Glycated hemoglobin levels; Hematological parameters; Erythrocyte

	indices; Hemoglobin a glycosylated; Hemoglobin; Iron; Total iron binding capacity; Ferritin; Diastolic blood pressure; Comparative height size at age 10; Transferrin saturation; Hepcidin levels; Red blood cell count; Polycythaemia vera; LDL cholesterol; Low density lipoprotein; Total cholesterol; Cholesterol total; Cardiovascular disease risk factors; Height; Pulse rate; Impedence of arm left; Ferritin log10; Hematocrit; Iron deficiency; Unsaturated iron binding capacity; Vascular or heart problems diagnosed by doctor: high blood pressure; Soluble transferrin receptor; Forced vital capacity; Self-reported hypertension; Platelet count; Cause of death: liver cell carcinoma; Monocyte count; Forced expiratory volume in 1 second; Impedence of arm right; Vascular or heart problems diagnosed by doctor: none of the above; Forced vital capacity, best measure; Hepcidin supplementary concept; Impedence of whole body; Red blood cell traits; Plateletcrit; Other complications of surgical and medical care; Hamatology traits; Forced expiratory volume in 1-second, best measure; Long-standing illness, disability or infirmity; Medication for cholesterol, blood pressure or diabetes: blood pressure medication; Platelet distribution width
rs1799945	mean corpuscular hemoglobin; mean corpuscular volume; red cell distribution width; mean corpuscular hemoglobin concentration; hemoglobin concentration; hematocrit; reticulocyte fraction of red cells; iron status biomarkers transferrin saturation; reticulocyte count; iron status biomarkers iron levels; vascular or heart problems diagnosed by doctor: high blood pressure; self-reported hypertension; HbA1c; vascular or heart problems diagnosed by doctor: none of the above; diastolic blood pressure; blood pressure; treatment with blood pressure medication; platelet count; mean arterial pressure; systolic blood pressure; red blood cell count; medication for cholesterol, blood pressure or diabetes: blood pressure medication; hypertension; iron status biomarkers ferritin levels; high light scatter percentage of red cells; hemoglobin; iron status biomarkers; iron; serum iron; plateletcrit; immature fraction of reticulocytes; treatment with Bendroflumethiazide; no treatment with medication for cholesterol, blood pressure, diabetes, or take exogenous hormones; high light scatter reticulocyte count; illness of siblings: high blood pressure; illness of mother: high blood pressure; treatment with lisinopril; platelet distribution width; serum urate
rs855791	Mean corpuscular hemoglobin; mean corpuscular volume, red cell distribution width; mean corpuscular hemoglobin concentration; hemoglobin concentration; iron status biomarkers iron levels; iron status biomarkers transferrin saturation; hematocrit; red blood cell traits; clinical laboratory measurements; reticulocyte fraction of red cells; hemoglobin; reticulocyte count; HbA1c; erythrocyte indices; transferrin saturation with iron ; iron; plateletcrit; platelet count; immature fraction of reticulocytes; soluble transferrin receptor; iron status biomarkers; iron regulatory proteins; iron status biomarkers ferritin levels; glycated hemoglobin HbA1c; hemoglobin a glycosylated; red blood cell count; hepcidin transferrin saturation ratio; transferrin saturation; hepcidin levels; transferrin; hematology traits; platelet

	distribution width; high light scatter percentage of red cells; mineral and other dietary supplements: iron; ferritin
rs57659670	None

Supplementary Table 5: **Association of genetic instruments with sex and age.**

	Exposure		Outcome			
SNP	Eff: Other	AF	Variant	AF	Sex p-value	Age p-value
rs1799945	G:C	0.1370	6:26091179:C:G	0.1512	0.695	0.732
rs1800562	A:G	0.0677	6:26093141:G:A	0.0788	0.852	0.248
rs57659670	C:T	0.0753	15:45398438:T:C	0.0759	0.556	0.570
rs855791	A:G	0.4310	22:37462936:A:G	0.5615	0.107	0.727

Variant identifier in the form "chr:pos:ref:alt", where "ref" is aligned to the forward strand of GRCh37 and "alt" is the effect allele.

Supplementary Table 6: **MR results when using each of the 4 biomarkers in turn to quantify the genetic associations of the instruments with systemic iron status.** IVW, weighted median, and MR-Egger results for all 4 biomarkers is shown. Highlighted row shows main results (Table 2).

Biomarker	MR Method	OR	95% CI	P	Cochran Q Statistic	Cochran Q Statistic P value	MR-Egger	
							Intercept P value	I_{gx}^2,%
Ferritin	IVW	0.48	(0.19:1.2)	0.12	8.6	0.036		
	Weighted median	0.74	(0.38:1.5)	0.4				
	MR-Egger	1.6	(0.38:7)	0.51			0.061	96
Iron	IVW	0.65	(0.47:0.91)	0.01	4.8	0.18		
	Weighted median	0.71	(0.52:0.98)	0.035				
	MR-Egger	0.95	(0.34:2.7)	0.92			0.46	100
TIBC	IVW	1.3	(0.79:2.1)	0.32	12	0.0091		
	Weighted median	1.2	(0.89:1.5)	0.28				
	MR-Egger	0.96	(0.62:1.5)	0.85			0.041	100
TSAT	IVW	0.74	(0.56:0.99)	0.044	6.5	0.088		
	Weighted median	0.77	(0.61:0.99)	0.038				
	MR-Egger	1	(0.57:1.8)	0.97			0.23	100

CI confidence interval; TIBC total iron-binding capacity; TSAT transferrin saturation.

Supplementary Table 7: **Power analyses**: We conduct various power analyses (Brion et al. *IJE* 2013; https://shiny.cnsgenomics.com/mRnd/) for a variety of R^2 values. We assumed a sample size of 336638, with a case prevalence of 1855/336638, and type-I error rate of 5%. We tested R^2 values of 0.65, 3.04, 3.01, and 5.07 to mirror ferritin, iron, TIBC, and TSAT, respectively.

R^2	OR	Power
0.65 (ferritin)	0.8	0.11
3.04 (iron)	0.8	0.32
5.07 (TSAT)	0.8	0.49
3.01 (TIBC)	1.2	0.32