Sulfasalazine-induced agranulocytosis is associated with the human leukocyte antigen locus

Mia Wadelius MD[1], Niclas Eriksson PhD[2], Reinhold Kreutz Professor[3], Emmanuelle Bondon-Guitton PharmD[4], Luisa Ibañez MD[5], Alfonso Carvajal Professor[6], M. Isabel Lucena Professor[7], Esther Sancho Ponce MD[8], Mariam Molokhia PhD[9], Javier Martin MD[10], Tomas Axelsson PhD[11], Hugo Kohnke MSc[1], Qun-Ying Yue MD[12], Patrik K. E. Magnusson PhD [13], Mats Bengtsson MD[14], Pär Hallberg MD[1], on behalf of EuDAC. *(See collaborators on page 15.)*

Affiliations

[1]Department of Medical Sciences, Clinical Pharmacology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [2]Uppsala Clinical Research Center and Department of Medical Sciences, Uppsala University, Uppsala, Sweden; [3]Charité – Universitätsmedizin Berlin, Institut für Klinische Pharmakologie und Toxikologie, Berlin, Germany; [4]Service de Pharmacologie Médicale et Clinique, Centre Hospitalier Universitaire, Faculté de Médecine de l'Université de Toulouse, Toulouse, France; [5]Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Fundació Institut Català de Farmacologia, Barcelona, Spain; [6]Centro de Estudios sobre la Seguridad de los Medicamentos, Universidad de Valladolid, Valladolid, Spain; [7]S Farmacologia Clinica, Instituto de Investigación Biomedica de Málaga (IBIMA), H Universitario Virgen de la Victoria, Universidad de Málaga, CIBERehd, Madrid, Spain; [8]Servei d'Hematologia i Banc de Sang, Hospital General de Catalunya, Sant Cugat del Vallès, Spain; [9]NIHR Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London Department of Primary Care and Public Health Sciences, London, UK; [10]Instituto de Parasitologia y Biomedicina Lopez-Neyra, CSIC, Granada, Spain; [11]Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [12]Medical Products Agency, Uppsala, Sweden; [13]Swedish Twin Registry, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [14]Department of Immunology, Genetics and Pathology, Clinical Immunology, Uppsala University, Uppsala, Sweden.

Characteristics of the discovery cohort are presented in Supplementary table S1. The top 60 single nucleotide polymorphisms (SNPs) in the discovery cohort are in Supplementary table S2. Supplementary figure S1 is a locus plot of the top hit on chromosome 6. Supplementary table S3 shows associations with human leukocyte antigen (HLA) imputed data in the discovery cohort. A multiple model of associated HLA alleles and the top GWAS SNP in the discovery cohort is shown in Supplementary figure S2. The power calculation is shown in Supplementary figure S3.

Quality control (QC) and data management was performed using PLINK v1.9. Markers that failed the following quality control criteria were discarded: call rate <95%, MAF <1%, a p-value for Hardy–Weinberg equilibrium $>1*10^{-7}$. All included cases and controls had <5% missing markers. All GWAS datasets were converted to top strand, Genome Reference Consortium human assembly 37 (GRCh37), and checked against 1000 genomes reference data to ensure that all markers were on the same strand. Markers were matched on chromosome and position, alleles were flipped where there were strand inconsistencies, and markers with residual discrepancies were removed. All single nucleotide polymorphisms (SNPs) were renamed with the rs-numbers according to the 1000 genomes reference data. The resulting merged data included 596,010 SNPs.

The data was pruned based on pairwise linkage disequilibrium (LD) and principal components (PCs) were calculated. The pruned data contained 110,336 SNPs. In addition, the data was merged with HapMap (release 23, 270 individuals) and PCs were recalculated using the above approach. Resulting graphs of stratification are shown in Supplementary figures S4 and S5. No pronounced genetic outliers were seen.

Imputation of genotypes in the merged and quality controlled genome-wide data was performed using PhaseIT and Impute v2. The 1000 genomes project reference set (phase 3, version October 2014) was downloaded from the impute2 website and used for imputations. Post imputation, hard genotype calls were made if the probability of a certain genotype was >0.8. Also, SNPs with MAF <1% were removed. The total number of SNPs after this process was 9,380,034**.** The Q-Q plot is shown in Supplementary figure S6.

Supplementary table S1. Characteristics of the cases and population controls in the discovery cohort for sulfasalazine-induced agranulocytosis. No genetic outlier was seen in analysis of principal components 1-2 (Supplementary figures S4 and S5). The controls were presumed to originate from the countries where they were recruited.

	Cases (n=36)	Controls (n=5,170)
Gender, male [%]	17 [47.2]	-
Age, years	at agranulocytosis [%]	-
· *<25*	1 [2.8]	
· *25-29*	2 [5.6]	
· *30-34*	4 [11.1]	
· *35-39*	2 [5.6]	
· *40-44*	1 [2.8]	
· *45-49*	1 [2.8]	
· *50-54*	6 [16.7]	
· *55-59*	7 [19.4]	
· *60-64*	5 [13.9]	
· *65-69*	3 [8.3]	
· *70-74*	0	
· *>74*	4 [11.1]	
Ethnicity	[% of cases]	[% of controls]
· *Swedish*	27 [75.0]	4,891 [94.6]
· *German*	3 [8.3]	96 [1.9]
· *French*	3 [8.3]	0
· *Spanish*	0	183 [3.5]
· *Other European*[#]	3 [8.3]	0
Diagnosis[¤]	[% of cases]	-
· Rheumatoid arthritis	17 [47.2]	
· Crohn's disease	3 [8.3]	
· Ulcerative colitis	3 [8.3]	
· Psoriasis arthritis	6 [16.7]	
· Ankylosing spondylitis	3 [8.3]	
· Arthritis unspecified	5 [13.9]	
· Sjögren's syndrome	1[2.8]	
Total number of diagnosis	38	
Mean time to onset in days [range]	50 [19-120]†	-
Mean lowest neutrophil count $x10^9$cells/l [range]	0.09 [0-0.5]	-
Mean daily dose in mg [range]	2,026 [1,600-3,000]	-
Cases with co-suspected drugs [%]	5 [13.9]	-

[¤]One case had both rheumatoid arthritis and psoriasis arthritis, and one case had both Crohn's disease and unspecified arthritis.

[#]One of Swedish–Italian, one of Swedish–Finnish, and one of Finnish origin.

† One outlier with time to onset 1,800 days was excluded.

Supplementary table S2. Top 60 markers after analysis of imputed data adjusted for sex and genetic principal components 1-4.

CHR	SNP	BP	GTPS	N	MAF case	MAF control	OR [95% CI]	P	*Gene*
6	rs9266634	31346978	A/T	5183	0.81	0.44	5.37 [2.97, 9.69]	2.55E-08	
20	rs202233001	21735353	C/CACA	5081	0.14	0.03	6.76 [3.41, 13.38]	4.19E-08	
17	rs9675270	1868035	A/T	5110	0.23	0.06	4.64 [2.64, 8.18]	1.04E-07	*RTN4RL1*
6	rs1140412	31324200	G/C	4950	0.67	0.35	3.96 [2.38, 6.59]	1.29E-07	*HLA-B*
17	rs9303228	1875308	C/T	5124	0.23	0.06	4.43 [2.53, 7.75]	1.81E-07	*RTN4RL1*
1	rs74512480	113416257	C/A	5101	0.13	0.02	7.11 [3.40, 14.88]	1.97E-07	
6	rs9372971	130964437	A/G	4900	0.12	0.02	7.02 [3.37, 14.64]	1.98E-07	
17	rs9907154	1875031	T/C	5142	0.23	0.06	4.37 [2.50, 7.64]	2.28E-07	*RTN4RL1*
6	rs1634768	31304560	T/C	4980	0.79	0.46	4.52 [2.55, 8.01]	2.33E-07	
20	rs76801208	57473728	T/C	5153	0.13	0.02	7.31 [3.44, 15.53]	2.35E-07	*GNAS*
9	rs117767021	91300786	A/T	5137	0.09	0.01	10.78 [4.35, 26.68]	2.75E-07	
7	rs74997117	154851728	T/C	5181	0.17	0.04	5.32 [2.81, 10.06]	2.78E-07	
12	12:55328807:G:A	55328807	A/G	5051	0.09	0.01	10.92 [4.37, 27.29]	3.14E-07	
13	rs200629715	21135339	C/CAAACAAAA	4849	0.19	0.04	5.34 [2.81, 10.15]	3.19E-07	
6	rs2523579	31328517	T/C	5193	0.67	0.36	3.67 [2.23, 6.04]	3.28E-07	
8	rs572589	102485762	G/T	5206	0.26	0.09	4.13 [2.40, 7.13]	3.28E-07	
1	rs113887891	117006870	G/T	5094	0.13	0.02	6.95 [3.30, 14.64]	3.30E-07	
6	rs73728881	31319079	C/T	5116	0.72	0.41	3.92 [2.32, 6.64]	3.41E-07	
6	rs9266052	31319080	C/T	5116	0.72	0.41	3.92 [2.32, 6.64]	3.41E-07	
3	rs114431066	197108332	T/C	5184	0.10	0.01	8.64 [3.77, 19.79]	3.45E-07	
13	rs73165468	21254366	T/C	4753	0.22	0.06	4.88 [2.65, 8.97]	3.53E-07	*IFT88*
16	rs150938367	85639393	T/G	5083	0.09	0.01	9.72 [4.04, 23.39]	3.92E-07	
6	rs9266042	31318621	T/A	5199	0.72	0.41	3.87 [2.29, 6.53]	4.30E-07	
17	rs118137040	56399092	T/C	5120	0.13	0.02	6.57 [3.17, 13.64]	4.38E-07	*BZRAP1*
6	rs9266639	31347069	T/C	5184	0.58	0.29	3.41 [2.12, 5.50]	4.45E-07	
6	rs4992474	31318933	T/G	5184	0.72	0.41	3.86 [2.28, 6.51]	4.48E-07	

6	rs9266055	31319151	G/A	5176	0.72	0.41	3.85 [2.28, 6.50]	4.61E-07	
6	6:31319143:AG:A	31319143	AG/A	5175	0.72	0.41	3.85 [2.28, 6.50]	4.63E-07	
14	rs143510029	38334751	T/C	5122	0.10	0.01	9.11 [3.86, 21.50]	4.72E-07	*TTC6*
6	rs9266053	31319102	C/G	5161	0.72	0.41	3.84 [2.27, 6.48]	4.96E-07	
6	rs9266054	31319109	A/G	5161	0.72	0.41	3.84 [2.27, 6.48]	4.96E-07	
8	rs10106789	102481009	G/T	5182	0.31	0.11	3.75 [2.24, 6.28]	5.11E-07	
3	rs145653734	197107062	G/A	5185	0.10	0.01	8.36 [3.65, 19.16]	5.20E-07	
3	rs114059153	197119291	T/C	5182	0.10	0.01	8.35 [3.64, 19.13]	5.29E-07	
3	rs115085889	197115255	C/G	5181	0.10	0.01	8.35 [3.64, 19.13]	5.30E-07	
3	rs139120906	197129053	C/T	5199	0.10	0.01	8.29 [3.63, 18.95]	5.35E-07	
8	rs2447961	102481308	T/C	5189	0.31	0.11	3.74 [2.23, 6.27]	5.41E-07	
17	rs4622541	1887642	T/C	5017	0.16	0.03	5.42 [2.79, 10.50]	5.68E-07	*RTN4RL1*
11	rs149020156	91673521	A/G	5175	0.10	0.01	7.41 [3.38, 16.25]	5.87E-07	
8	rs11352430	144653527	G/GC	4965	0.39	0.15	3.63 [2.19, 6.02]	5.95E-07	*MROH6*
3	rs114856590	197116828	C/T	5181	0.10	0.01	8.26 [3.60, 18.93]	6.05E-07	
17	rs4790857	1885634	A/C	5027	0.16	0.03	5.40 [2.78, 10.50]	6.41E-07	*RTN4RL1*
9	rs144251347	111881397	A/G	5193	0.17	0.04	5.10 [2.69, 9.70]	6.54E-07	*TMEM245*
9	rs79792689	111891023	T/C	5193	0.17	0.04	5.10 [2.69, 9.70]	6.54E-07	
6	rs2523609	31237255	G/A	5202	0.68	0.38	3.48 [2.13, 5.69]	6.82E-07	*HLA-C*
9	rs74697586	111870894	C/A	5190	0.17	0.04	5.09 [2.68, 9.68]	6.92E-07	*TMEM245*
6	rs622871	31878495	A/G	5154	0.60	0.31	3.29 [2.06, 5.27]	6.99E-07	*C2*
7	rs142915233	24843000	A/G	5140	0.08	0.01	9.86 [3.99, 24.34]	6.99E-07	*OSBPL3*
8	rs74461346	70361625	T/C	4940	0.17	0.04	5.22 [2.72, 10.02]	7.02E-07	
8	rs615581	102482839	C/T	5206	0.31	0.11	3.70 [2.21, 6.20]	7.02E-07	
8	rs616153	102482650	G/A	5206	0.31	0.11	3.70 [2.21, 6.20]	7.02E-07	
12	12:1662065:C:T	1662065	T/C	5197	0.22	0.07	4.57 [2.51, 8.33]	7.09E-07	*WNT5B*
6	rs630379	31922254	A/C	5206	0.60	0.31	3.28 [2.05, 5.24]	7.17E-07	*NELFE*
5	rs111876221	79535200	A/G	5175	0.10	0.02	8.71 [3.70, 20.52]	7.35E-07	*SERINC5*
2	rs143354638	107124840	A/G	5063	0.09	0.01	10.00 [4.02, 24.88]	7.43E-07	

2	rs149050899	107124839	T/G	5063	0.09	0.01	10.00 [4.02, 24.88]	7.43E-07	
3	rs141498805	197140043	C/T	5200	0.10	0.01	8.07 [3.53, 18.46]	7.51E-07	
7	rs181271864	24871395	C/A	5133	0.08	0.01	9.78 [3.96, 24.14]	7.52E-07	*OSBPL3*
8	rs16919354	54466588	G/T	5054	0.13	0.02	6.58 [3.11, 13.89]	7.92E-07	

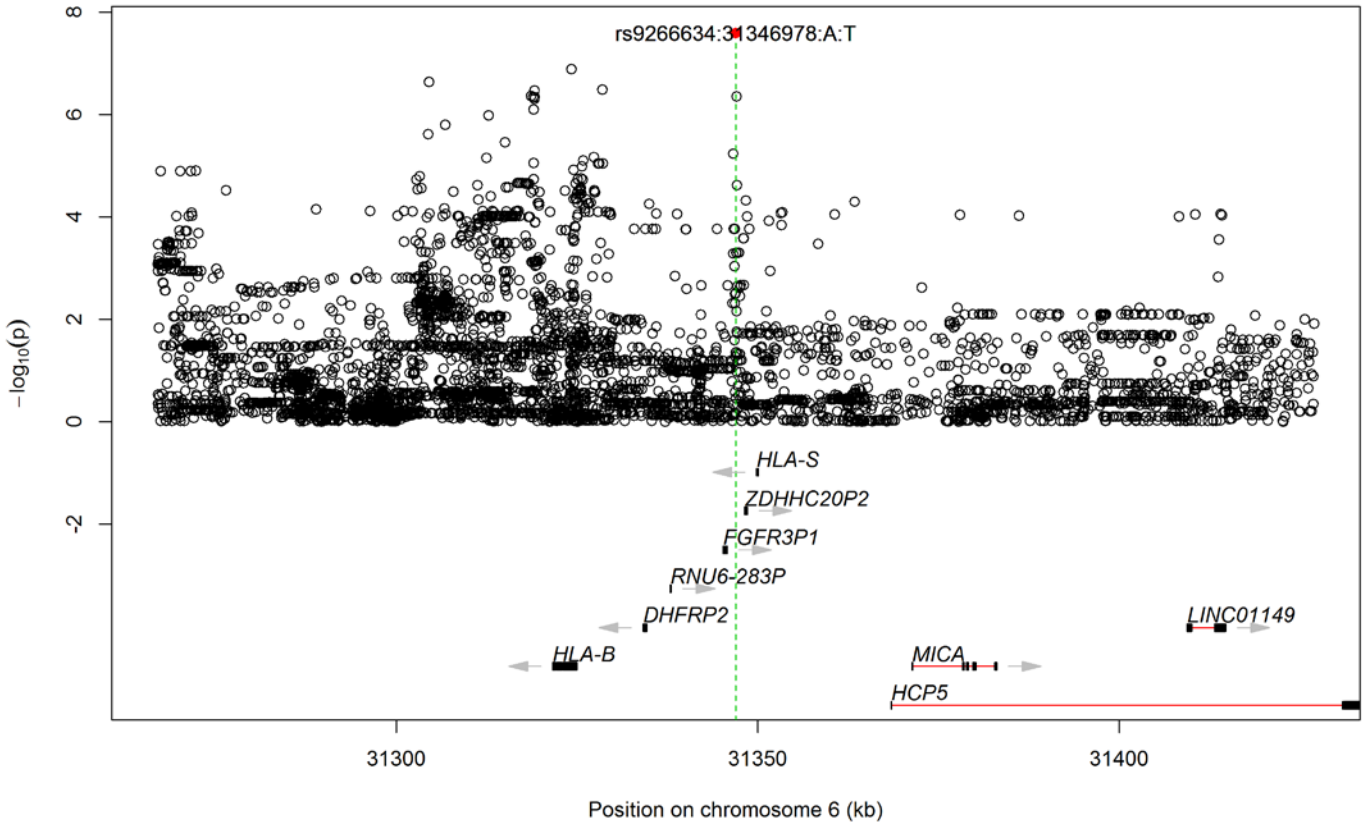
CHR = chromosome, SNP = single nucleotide polymorphism, BP = base pair, GTPS = nucleotide exchange, N = number, MAF = minor allele frequency, OR [95% CI] = odds ratio with 95% confidence interval, P = p-value.

Supplementary table S3. Top 30 imputed human leukocyte antigen (HLA) alleles at second field resolution adjusted for sex and genetic principal components 1-4. The effect is modeled per increase of one allele.

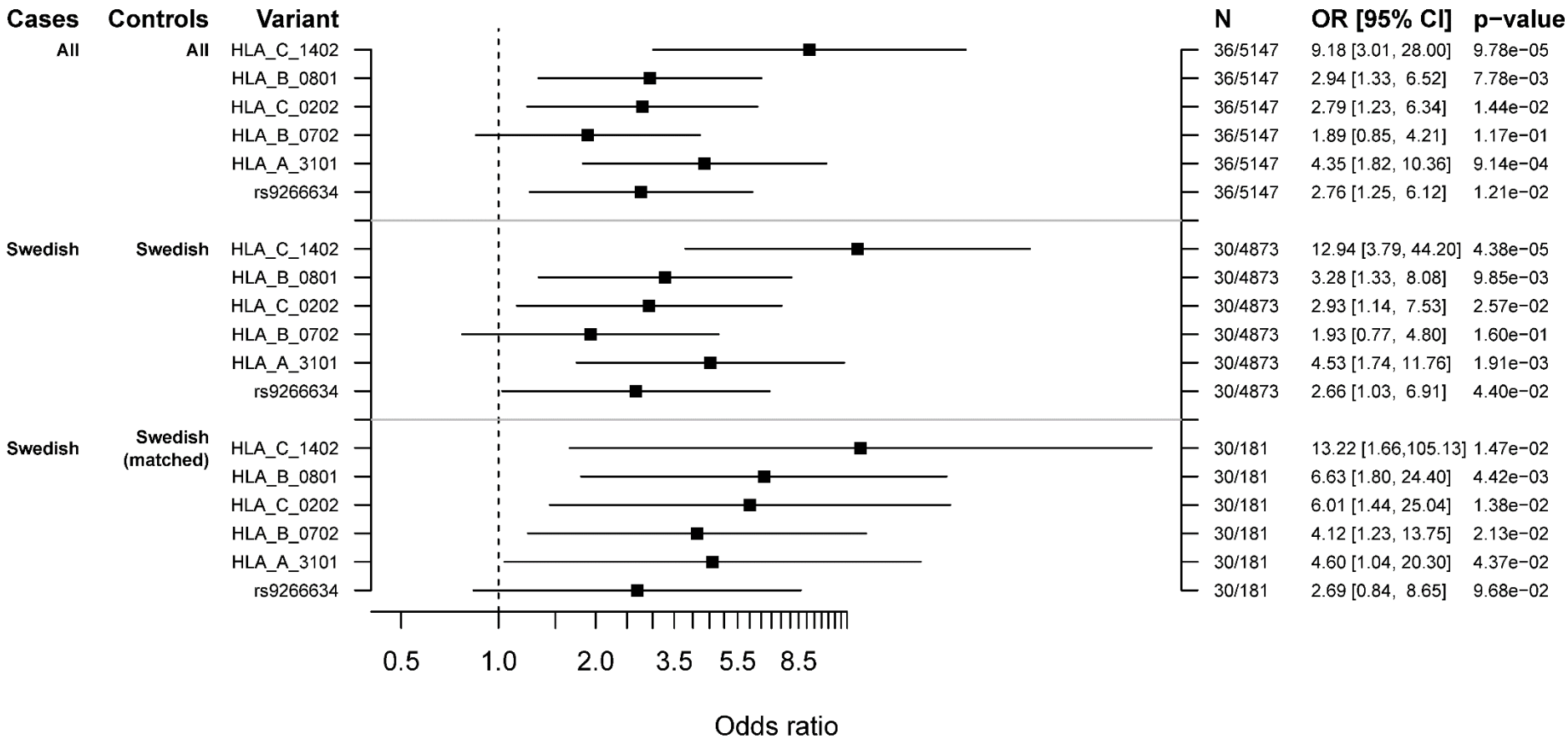
CHR	HLA variant	BP	GTPS	N	MAF case	MAF control	OR [95% CI]	P
6	HLA_C_1402	31346171	P/A	5206	0.07	0.01	7.72 [2.90, 20.58]	4.36E-05
6	HLA_B_0801	31431272	P/A	5206	0.26	0.11	2.92 [1.71, 4.99]	8.37E-05
6	HLA_C_0206	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0305	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0306	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0403	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0404	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0407	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0410	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_0726	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_1204	31346171	P/A	5206	0.01	0.00	92.68 [8.02, 1071.00]	2.86E-04
6	HLA_C_1504	31346171	P/A	5206	0.01	0.00	175.30 [10.56, 2909.00]	3.12E-04
6	HLA_C_1801	31346171	P/A	5206	0.01	0.00	175.30 [10.56, 2909.00]	3.12E-04
6	HLA_C_0701	31346171	P/A	5206	0.29	0.14	2.52 [1.51, 4.20]	3.97E-04
6	HLA_A_3101	30019970	P/A	5206	0.10	0.03	3.76 [1.67, 8.46]	1.40E-03
6	HLA_C_0202	31346171	P/A	5206	0.15	0.06	2.68 [1.41, 5.08]	2.59E-03
6	HLA_B_5101	31431272	P/A	5206	0.14	0.05	2.82 [1.42, 5.60]	3.16E-03
6	HLA_A_7401	30019970	P/A	5206	0.01	0.00	25.10 [2.38, 264.70]	7.34E-03
6	HLA_C_0801	31346171	P/A	5206	0.01	0.00	17.08 [2.10, 139.30]	8.03E-03
6	HLA_C_0702	31346171	P/A	5206	0.25	0.15	1.88 [1.10, 3.22]	2.03E-02
6	HLA_C_1604	31346171	P/A	5206	0.01	0.00	11.68 [1.46, 93.49]	2.06E-02
6	HLA_A_0101	30019970	P/A	5206	0.24	0.14	1.90 [1.10, 3.30]	2.16E-02
6	HLA_C_1602	31346171	P/A	5206	0.01	0.00	11.36 [1.36, 94.51]	2.46E-02
6	HLA_B_1501	31431272	P/A	5206	0.01	0.11	0.12 [0.02, 0.85]	3.44E-02

6	HLA_B_2702	31431272	P/A	5206	0.01	0.00	8.93 [1.15, 69.24]	3.62E-02
6	HLA_B_0702	31431272	P/A	5206	0.24	0.15	1.76 [1.02, 3.01]	4.12E-02
6	HLA_C_0304	31346171	P/A	5206	0.06	0.14	0.36 [0.13, 0.98]	4.59E-02
6	HLA_B_4002	31431272	P/A	5206	0.06	0.02	2.83 [1.02, 7.90]	4.66E-02
6	HLA_C_0302	31346171	P/A	5206	0.01	0.00	7.90 [1.00, 62.16]	4.96E-02
6	HLA_B_4402	31431272	P/A	5206	0.03	0.11	0.25 [0.06, 1.01]	5.10E-02

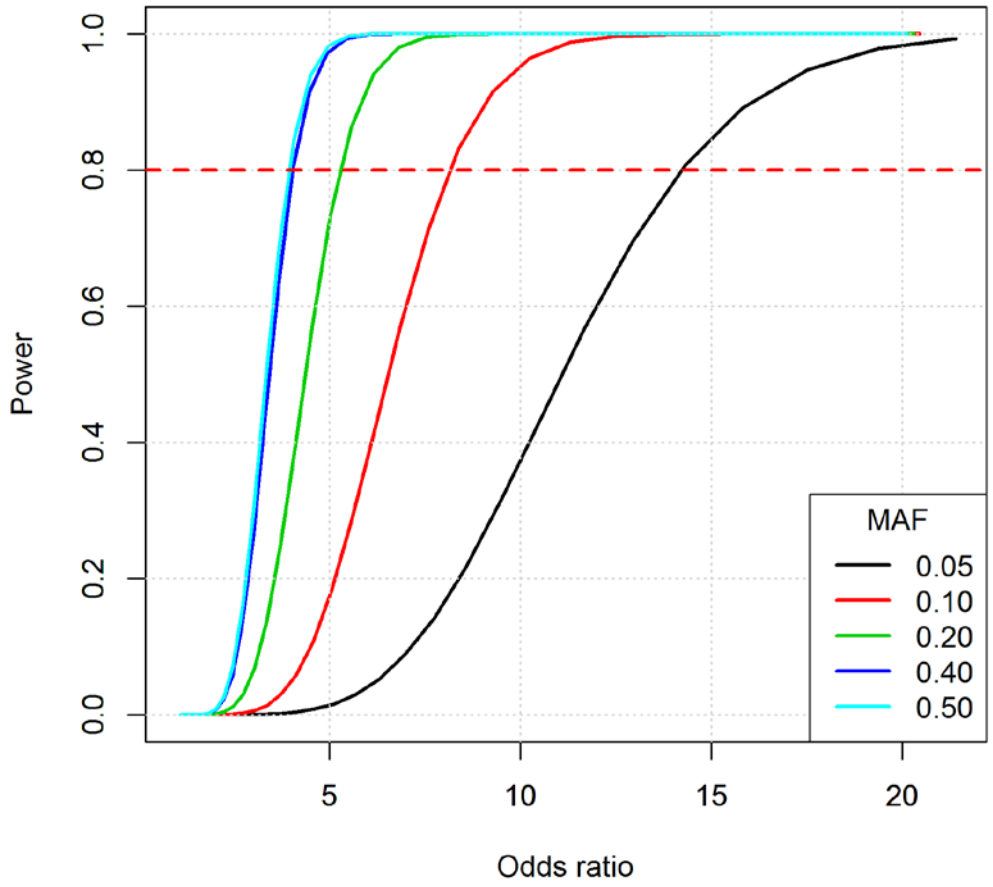
CHR = chromosome, HLA = human leukocyte antigen, BP = base pair, GTPS = nucleotide exchange, N = number, MAF = minor allele frequency, OR [95% CI] = odds ratio with 95% confidence interval, P = p-value.



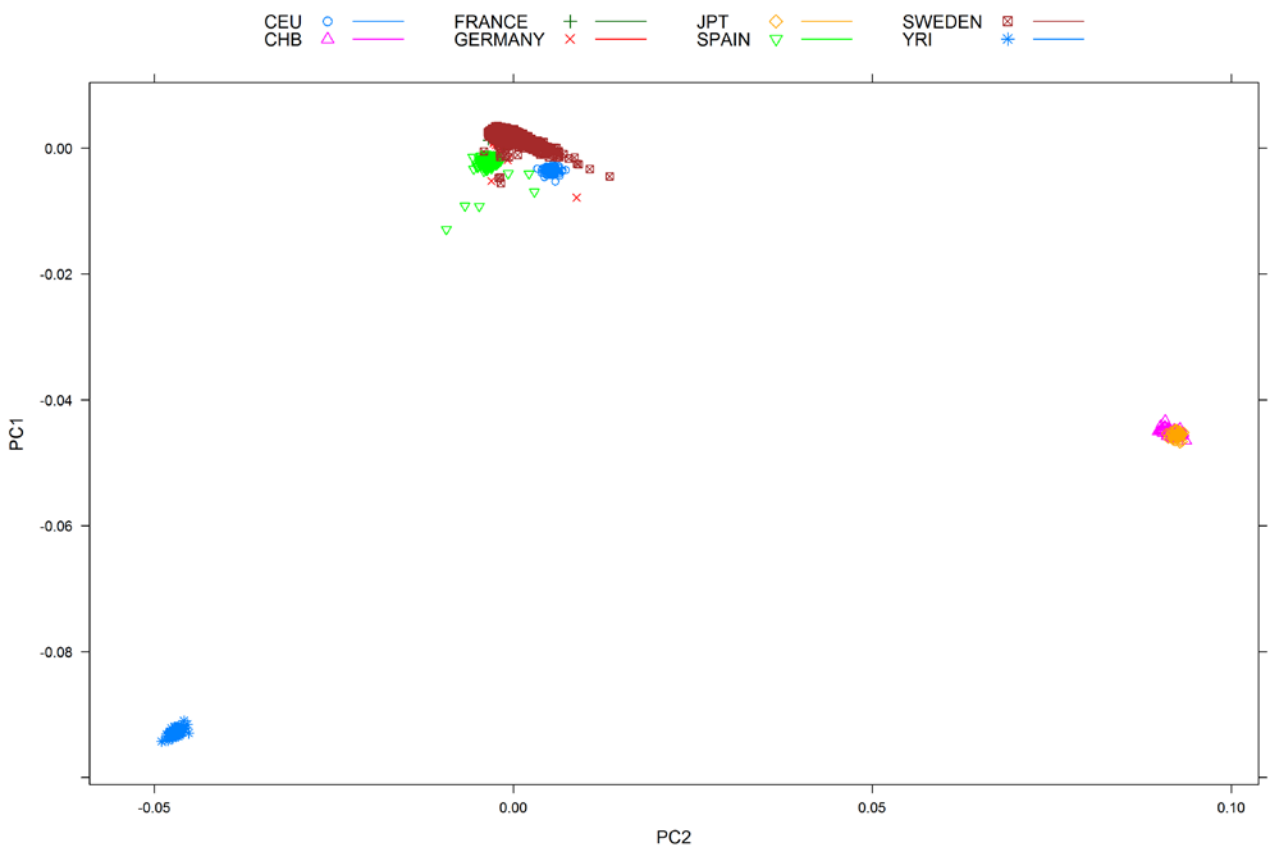
Supplementary figure S1. Locus plot of rs9266634 A>T at 31346978 in the HLA region on chromosome 6. The analysis is corrected for sex, PC and all included variants.



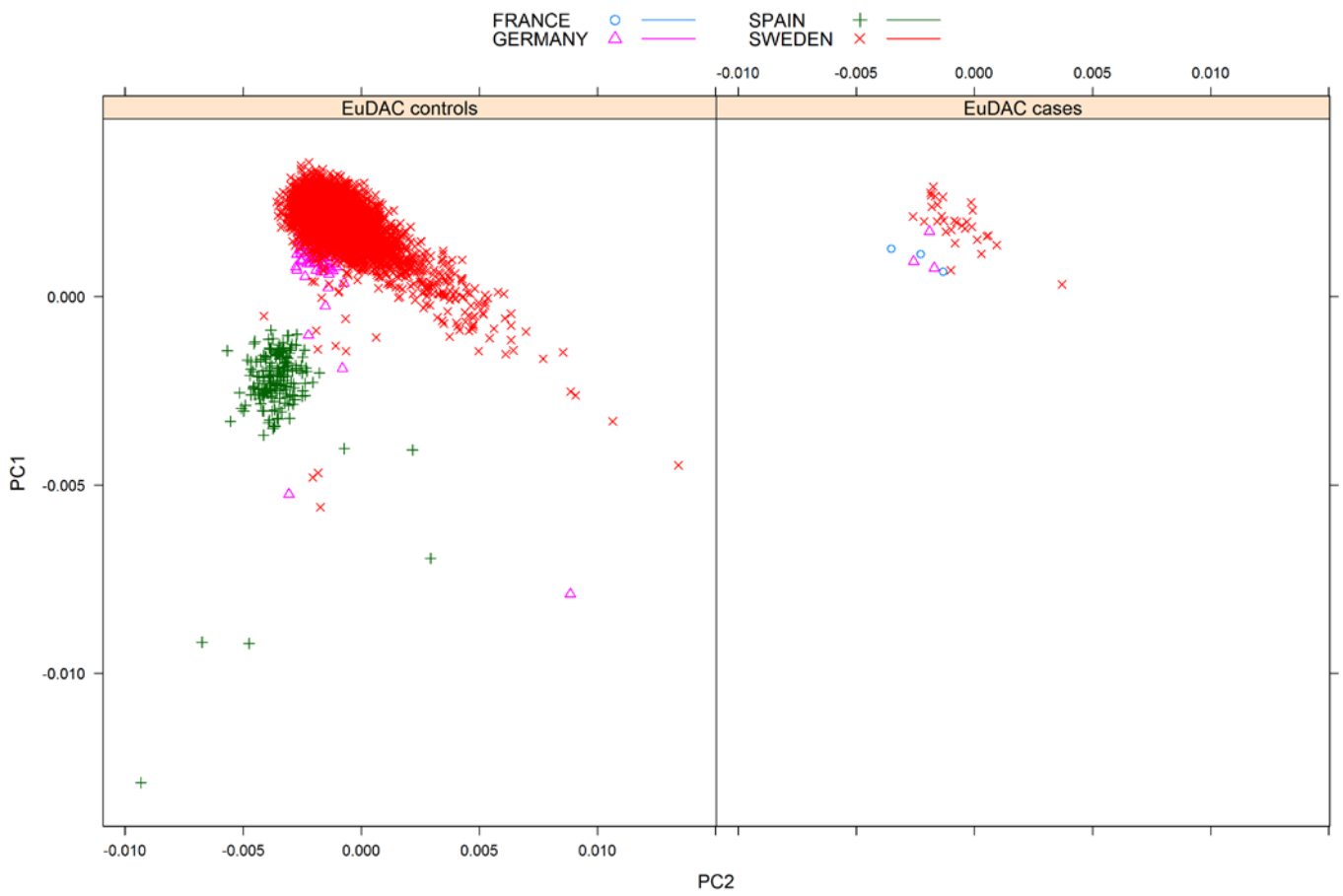
Supplementary figure S2. A multiple model that includes associated HLA alleles and the top GWAS SNP in the discovery cohort. The model is corrected for sex, PC and all included variants.



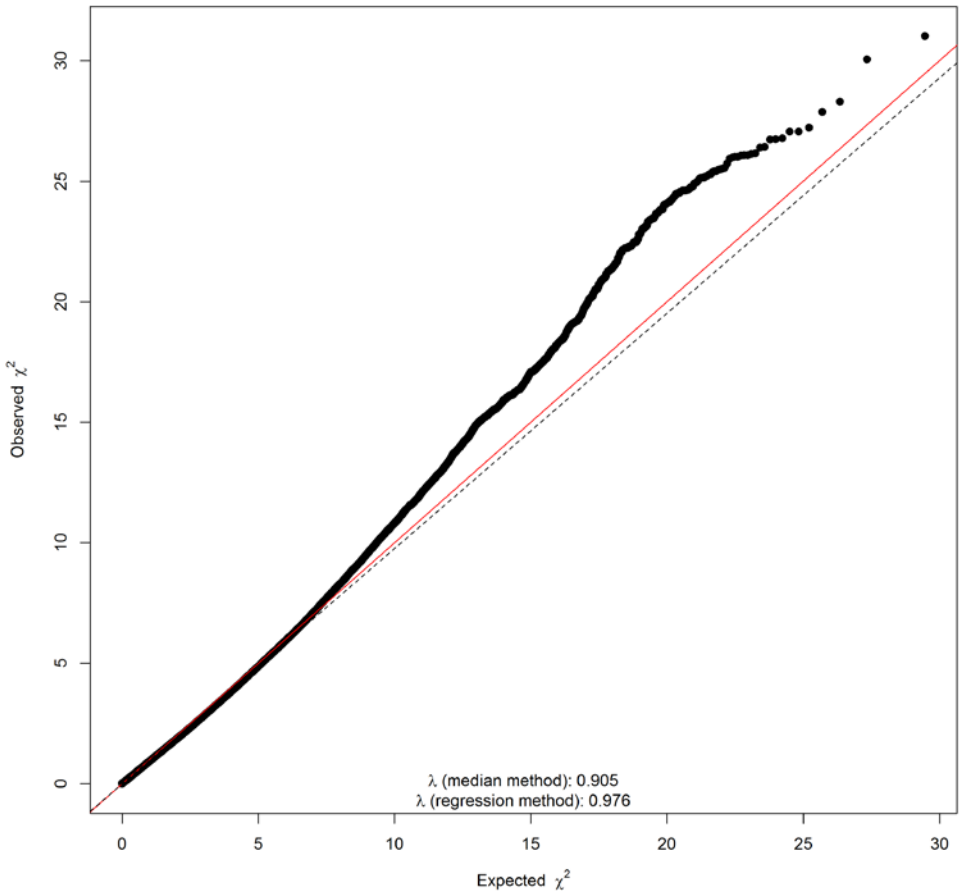
Supplementary figure S3. Power calculation. With 36 cases and 5,000 controls, the power is 80% to detect an odds ratio (OR) of 4 for variants with minor allele frequency (MAF) 40%, and to detect an OR slightly above 5 for variants with MAF 20%.



Supplementary figure S4. Analysis of principal components 1 and 2 (PC 1 and PC2) for the EUDAC sulfasalazine cohort compared with Utah residents with Northern and Western European ancestry from the CEPH collection (CEU), Han Chinese in Beijing, China (CHB), Japanese in Tokyo, Japan (JPT), and Yoruba people in Ibadan, Nigeria (YRI).



Supplementary figure S5. Analysis of principal components 1 and 2 (PC 1 and PC2) for the EUDAC controls and the cases of sulfasalazine-induced agranulocytosis.



Supplementary figure S6. Q-Q plot for analysis of sulfasalazine-induced agranulocytosis cases vs all controls after imputation. Adjusted by sex, genetic principal components 1-4. The red line is the line of identity. The dashed black line is the regression method lambda result.

EuDAC collaborators

Maryse Lapeyre-Mestre, Jean Louis Montastruc (Service de Pharmacologie Médicale et Clinique, Centre Hospitalier Universitaire, Faculté de Médecine de l'Université de Toulouse, UMR Inserm 1027, CIC 1436, Toulouse, France); Edeltraut Garbe (Leibniz-Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany); Lourdes Vendrell (Fundació Institut Català de Farmacologia, Hospital Universitari Vall d'Hebron, Barcelona, Spain); Ramon Puig Treserra (Fundació Institut Català de Farmacologia, Barcelona, Spain); Jose Luis Caro (Banc de Sang i Teixits, Barcelona, Spain); Maria Sainz Gil, Maria-Isabel Jimenez Serrania, Inés Salado (Centro de Estudios sobre la Seguridad de los Medicamentos. Universidad de Valladolid, Spain); Paul McKeigue (University of Edinburgh Medical School, UK); Erik Eliasson (Clinical Pharmacology, Karolinska Institutet, Sweden); Håkan Melhus, Ulrica Ramqvist, Elisabet Stjernberg, Sofie Collin, Eva Prado Lopez, Agnes Wadelius, Martha Wadelius and Agnes Kataja Knight (Department of Medical Sciences, Clinical Pharmacology, Uppsala University, Uppsala, Sweden); Daniel Garwicz (Department of Medical Sciences, Clinical Chemistry, Uppsala University, Uppsala, Sweden); Ulla Lindqvist (Department of Medical Sciences, Rheumatology, Uppsala University, Uppsala, Sweden); Marco Cavalli and Claes Wadelius (Department of Immunology, Genetics and Pathology, Medical Genetics, Uppsala University, Uppsala, Sweden); Bruno Stricker (Erasmus Medical Center, The Netherlands); Julia Ruiz-Nuñez, Camilla Stephens (S Farmacologia Clinica, IBIMA, H Universitario Virgen de la Victoria, Universidad de Málaga, CIBERehd, Madrid, Spain), Inger Öhman (Medical Products Agency, Uppsala, Sweden), Barbro Sandin (Swedish Twin Registry, MEB, Karolinska Institutet, Stockholm, Sweden). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX).