Supplementary Material

Genetic variants associated with antithyroid drug-induced agranulocytosis: a genome-wide association study in a European population

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

[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]Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Fundació Institut Català de Farmacologia, Barcelona, Spain; [4]Service de Pharmacologie Médicale et Clinique, Centre Hospitalier Universitaire, Faculté de Médecine de l'Université de Toulouse, Toulouse, France; [5]Charité – Universitätsmedizin Berlin, Institut für Klinische Pharmakologie und Toxikologie, Berlin, Germany; [6]Centro de Estudios sobre la Seguridad de los Medicamentos, Universidad de Valladolid, Valladolid, Spain; [7]S Farmacologia Clinica, IBIMA, H Universitario Virgen de la Victoria, Universidad de Málaga, Málaga, and 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.

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, Hugo Kohnke, 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); Bruno Stricker (Erasmus Medical Center, The Netherlands); Julia Ruiz-Nuñes, Camilla Stephens (S Farmacologia Clinica, IBIMA, H Universitario Virgen de la Victoria, Universidad de Málaga, CIBERehd, Málaga, Spain), Inger Öhman (Medical Products Agency, Uppsala, Sweden).

Supplementary methods

Genome-wide array data and analyses

GWAS 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, build 37, 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 596010 SNPs.

The data was pruned based on pairwise linkage disequilibrium (LD) and principal components (PCs) were calculated on these data. The pruned data contained 110336 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 S2 and S3. In figure S3 6 pronounced genetic outliers can be seen, 3 cases from Spain and 3 cases from France. These cases were not excluded from the main GWAS and subsequent analyses of the manuscript, however, sensitivity analyses were performed by reanalyzing each top hit with the 6 cases excluded.

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 the 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 9380034.

Supplementary tables

Supplementary Table S1. Top GWAS results based on 9380034 single nucleotide polymorphisms (SNPs) after imputation for all cases. A) Adjusted by sex and genetic principal components 1-4; B) adjusted by sex and genetic principal components 1-4 and rs114291795. After adjusting for both rs114291795 and rs1811197 no genome-wide significant signals were left.

	CHR	SNP	BP	Alleles (minor/major)	N	MAF case	MAF control	OR [95% CI]	P	Nearby gene
A	6	rs114291795	31377640	G/C	5376	0·14	0·06	3·57 [2·61, 4·90]	2·32E-15	*MICA-HCP5*
	6	rs148015908	31376564	A/T	5394	0·14	0·06	3·50 [2·57, 4·77]	2·52E-15	*MICA-HCP5*
	6	rs115613847	31364812	T/C	5400	0·14	0·06	3·47 [2·55, 4·73]	3·86E-15	
	6	rs145575084	31365089	C/G	5400	0·14	0·06	3·47 [2·55, 4·73]	3·86E-15	
	6	rs4349859	31365787	A/G	5400	0·14	0·06	3·47 [2·55, 4·73]	3·86E-15	
	6	rs116285304	31368391	T/A	5399	0·14	0·06	3·47 [2·55, 4·73]	3·88E-15	
	6	rs116135464	31370984	T/C	5400	0·14	0·06	3·47 [2·54, 4·73]	3·98E-15	*HCP5*
	6	rs189600525	31372642	T/C	5396	0·14	0·06	3·47 [2·54, 4·73]	4·10E-15	*MICA-HCP5*
	6	rs116666910	31327034	A/C	5400	0·14	0·06	3·46 [2·53, 4·72]	5·80E-15	
	6	rs1071816	31324536	C/T	5221	0·27	0·11	2·65 [2·07, 3·38]	5·96E-15	*HLA-B*
B	6	rs1811197	31335997	A/G	5366	0·29	0·16	2·40 [1·89, 3·06]	9·42E-13	
	6	rs2596487	31335997	T/C	5376	0·29	0·16	2·33 [1·83, 2·96]	4·87E-12	
	6	rs3094600	31335997	A/G	5376	0·38	0·24	2·12 [1·70, 2·64]	3·00E-11	
	6	rs2249935	31335997	A/G	5355	0·30	0·17	2·20 [1·73, 2·79]	1·36E-10	
	6	rs2518028	31335997	T/C	5376	0·41	0·32	2·00 [1·61, 2·48]	4·14E-10	*HCP5*
	6	rs532086	31335997	C/T	5352	0·25	0·15	2·13 [1·67, 2·71]	8·52E-10	*C2*
	6	rs2844608	31335997	T/C	5375	0·24	0·43	0·48 [0·38, 0·61]	1·65E-09	
	6	rs67200295	31335997	A/AA*	5360	0·09	0·02	3·58 [2·34, 5·47]	4·25E-09	
	6	rs76217799	31335997	C/T	5373	0·16	0·11	2·35 [1·77, 3·13]	4·35E-09	
	6	rs10649173	31335997	A/AA*	5370	0·17	0·12	2·30 [1·74, 3·05]	4·39E-09	

***** Insertion/deletion variants truncated to max two alleles.

CHR=chromosome; SN*P*=single nucleotide polymorphism; B*P*=base pair; MAF=minor allele frequency; OR=odds ratio; CI=confidence interval.

Supplementary Table S2. Top GWAS results for antithyroid induced cases vs all controls based on 9380034 SNPs after imputation. Adjusted by A) sex and genetic principal components 1-4; B) sex, genetic principal components 1-4 and rs652888; C) sex, genetic principal components 1-4, rs652888 and rs199564443. D) sex, genetic principal components 1-4, rs652888, rs199564443 and rs1071816. After adjusting for variants in A, B, C and D no genome-wide significant signals were left.

	CHR	SNP	BP	Alleles (minor/major)	N	MAF case	MAF control	OR [95% CI]	P	Nearby gene
A	6	rs652888	31851234	G/A	5203	0·54	0·20	4·73 [3·00, 7·44]	1·92E-11	*EHMT2*
	6	rs199564443	1388978	C/CTTTT*	5155	0·12	0·01	16·06 [7·14, 36·14]	1·95E-11	*FOXF2*
	6	rs532086	31881309	C/T	5186	0·45	0·15	4·69 [2·99, 7·37]	1·99E-11	*C2*
	6	rs115308096	1391939	G/A	5151	0·12	0·01	14·28 [6·39, 31·93]	9·46E-11	*FOXF2*
	6	rs1071816	31324536	C/T	5032	0·40	0·11	4·65 [2·88, 7·50]	2·88E-10	*HLA-B*
	6	rs9394078	31868938	T/C	5206	0·21	0·04	6·81 [3·74, 12·42]	3·76E-10	*C2-ZBTB12*
	6	rs115549271	32452443	A/G	5030	0·49	0·19	4·40 [2·74, 7·05]	7·88E-10	
	6	rs183988210	31819813	G/C	5189	0·15	0·03	8·28 [4·16, 16·46]	1·68E-09	
	6	rs62405568	32453409	C/T	4975	0·47	0·19	4·27 [2·65, 6·89]	2·75E-09	
	6	rs150481781	31315099	AG*/A	5182	0.73	0·34	4·68 [2·81, 7·79]	2·76E-09	
B	6	rs199564443	1388978	C/CTTTT*	5149	0·12	0·01	17·42 [7·38, 41·12]	7·04E-11	*FOXF2*
	6	rs115308096	1391939	G/A	5145	0·12	0·01	15·35 [6·57, 35·87]	2·89E-10	*FOXF2*
	6	rs1071816	31324536	C/T	5026	0·40	0·11	4·54 [2·73, 7·55]	5·13E-09	*HLA-B*
	6	rs56705452	31391591	A/G	5203	0·45	0·20	4·31 [2·59, 7·16]	1·77E-08	*HCP5*
	16	rs78694352	55370494	C/T	5168	0·13	0·02	9·36 [4·26, 20·58]	2·64E-08	
	8	kgp30515713	30515713	C/A	4669	0·11	0·02	10·87 [4·65, 25·42]	3·70E-08	*GTF2E2*
	8	kgp30515715	30515715	C/T	4669	0·11	0·02	10·87 [4·65, 25·42]	3·70E-08	*GTF2E2*
	8	kgp30515718	30515718	C/G	4669	0·11	0·02	10·87 [4·65, 25·42]	3·70E-08	*GTF2E2*
	20	rs116943356	7664053	A/G	5165	0·09	0·01	13·00 [5·21, 32·46]	3·88E-08	
	16	rs78820965	55341430	C/T	5197	0·13	0·03	8·96 [·08, 19·69]	4·87E-08	
C	6	rs1071816	31324536	C/T	4975	0·40	0·11	5·27 [3·06, 9·10]	2·35E-09	*HLA-B*
	6	kgp31249991	31249991	*GG/A	5148	0·23	0·06	6·21 [3·38, 11·42]	4·16E-09	
	6	rs200628408	31249990	*TC/G	5148	0·23	0·06	6·21 [3·38, 11·42]	4·16E-09	
	6	rs185924983	31159511	T/C	5138	0·23	0·06	5·98 [3·28, 10·92]	5·65E-09	
	6	rs142973694	31362941	A/G	5129	0·22	0·06	6·96 [3·61, 13·44]	7·13E-09	
	8	kgp30515713	30515713	C/A	4626	0·11	0·02	13·10 [5·43, 31·62]	1·06E-08	*GTF2E2*
	8	kgp30515715	30515715	C/T	4626	0·11	0·02	13·10 [5·43, 31·62]	1·06E-08	*GTF2E2*
	8	kgp30515718	30515718	C/G	4626	0·11	0·02	13·10 [5·43, 31·62]	1·06E-08	*GTF2E2*
	6	rs115613847	31364812	T/C	5146	0·23	0·06	6·36 [3·35, 12·07]	1·50E-08	
	6	rs116285304	31368391	T/A	5145	0·23	0·06	6·36 [3·35, 12·07]	1·50E-08	
D	8	rs111618861	57044382	C/CA	4715	0·11	0·01	13·75 [5·46, 34·64]	2·71E-08	
	8	rs117204475	17346323	C/T	4800	0·10	0·02	11·32 [4·56, 28·07]	1·645E-07	
	16	rs12919345	8932933	C/T	4886	0·09	0·01	12·97 [4·88, 34·53]	2·872E-07	*PMM2*
	22	rs148706762	31308214	A/G	4952	0·06	0·01	16·66 [5·68, 48·91]	3·053E-07	
	1	rs2358816	114419368	T/A	4915	0·46	0·21	4·30 [2·45, 7·54]	3·744E-07	*AP4B1-AS1*
	15	rs74375844	80792298	A/G	4961	0·08	0·01	14·98 [5·19, 43·20]	5·509E-07	*ARNT2*
	8	rs150304897	118167795	A/G	4945	0·09	0·02	11·17 [4·34, 28·79]	5·814E-07	*SLC30A8*
	19	rs1019139	28433354	T/G	4975	0·08	0·01	13·01 [4·75, 35·59]	5·889E-07	

19	rs149637583	28433967	T/C	4975	0·08	0·01	13·01 [4·75, 35·59]	5·889E-07
19	rs57303086	28435597	T/C	4975	0·08	0·01	13·01 [4·75, 35·59]	5·889E-07

***** Insertion/deletion variants truncated to max two alleles.

CHR=chromosome; SN*P*=single nucleotide polymorphism; B*P*=base pair; MAF=minor allele frequency; OR=odds ratio; CI=confidence interval.

Supplementary Table S3. Analysis of human leukocyte antigen (HLA) imputed data only showing four digit HLA alleles. The effect is modeled per increase of one allele. A) Top markers for all drugs; B) Top markers for all drugs adjusted for *HLA-B*27:05*; C) Top markers for all drugs adjusted for *HLA-B*27:05* and *HLA-B*08:01***;** D) Top markers for all drugs adjusted for *HLA-B*27:05*, *HLA-B*08:01* and *HLA-B*39:01***.**

	CHR	SNP	BP	N	MAF case	MAF control	OR [95% CI]	P
A	6	*HLA-B*27:05*	31431272	5404	0·10	0·06	3·24 [2·31, 4·55]	1·20E-11
	6	*HLA-C*02:02*	31346171	5404	0·14	0·06	2·55 [1·87, 3·49]	4·21E-09
	6	*HLA-B*08:01*	31431272	5404	0·16	0·11	1·90 [1·44, 2·50]	5·68E-06
	6	*HLA-A*31:01*	30019970	5404	0·06	0·03	2·67 [1·71, 4·16]	1·59E-05
	6	*HLA-B*39:01*	31431272	5404	0·03	0·01	4·23 [2·12, 8·44]	4·16E-05
	6	*HLA-C*07:01*	31346171	5404	0·20	0·14	1·67 [1·29, 2·15]	8·99E-05
	6	*HLA-C*04:01*	31346171	5404	0·07	0·09	0·49 [0·32, 0·74]	7·10E-04
	6	*HLA-B*15:01*	31431272	5404	0·03	0·11	0·42 [0·24, 0·73]	1·85E-03
	6	*HLA-A*01:01*	30019970	5404	0·19	0·14	1·49 [1·15, 1·93]	2·96E-03
	6	*HLA-B*39:06*	31431272	5404	0·02	0·01	3·88 [1·56, 9·63]	3·47E-03
B	6	*HLA-B*08:01*	31431272	5404	0·16	0·11	2·14 [1·61, 2·84]	1·48E-07
	6	*HLA-C*07:01*	31346171	5404	0·20	0·14	1·88 [1·45, 2·45]	2·49E-06
	6	*HLA-B*39:01*	31431272	5404	0·03	0·01	4·49 [2·24, 9·01]	2·39E-05
	6	*HLA-A*31:01*	30019970	5404	0·06	0·03	2·48 [1·59, 3·86]	5·99E-05
	6	*HLA-B*39:06*	31431272	5404	0·02	0·01	4·60 [1·83, 11·57]	1·20E-03
	6	*HLA-A*01:01*	30019970	5404	0·19	0·14	1·54 [1·19, 2·01]	1·29E-03
	6	*HLA-C*04:01*	31346171	5404	0·07	0·09	0·51 [0·34, 0·78]	1·89E-03
	6	*HLA-C*15:02*	31346171	5404	0·03	0·02	2·18 [1·26, 3·77]	5·51E-03
	6	*HLA-B*15:01*	31431272	5404	0·03	0·11	0·46 [0·27, 0·80]	5·60E-03
	6	*HLA-C*02:02*	31346171	5404	0·14	0·06	1·67 [1·15, 2·43]	7·41E-03
C	6	*HLA-B*39:01*	31431272	5404	0·03	0·01	4·97 [2·47, 10·00]	6·97E-06
	6	*HLA-A*31:01*	30019970	5404	0·06	0·03	2·71 [1·74, 4·24]	1·16E-05
	6	*HLA-B*39:06*	31431272	5404	0·02	0·01	5·31 [2·10, 13·43]	4·15E-04
	6	*HLA-C*15:02*	31346171	5404	0·03	0·02	2·54 [1·46, 4·43]	9·56E-04
	6	*HLA-C*02:02*	31346171	5404	0·14	0·06	1·84 [1·26, 2·70]	1·74E-03
	6	*HLA-B*40:02*	31431272	5404	0·05	0·02	2·16 [1·27, 3·69]	4·56E-03
	6	*HLA-C*04:01*	31346171	5404	0·07	0·09	0·57 [0·37, 0·87]	9·58E-03
	6	*HLA-C*12:03*	31346171	5404	0·07	0·03	1·73 [1·12, 2·65]	1·31E-02
	6	*HLA-C*04:03*	31346171	5404	0·00	0·00	22·74 [1·90, 271·70]	1·36E-02
	6	*HLA-C*02:06*	31346171	5404	0·00	0·00	22·69 [1·89, 272·70]	1·38E-02
D	6	*HLA-A*31:01*	30019970	5404	0·06	0·03	2·61 [1·67, 4·09]	2·54E-05
	6	*HLA-B*39:06*	31431272	5404	0·02	0·01	5·66 [2·23, 14·34]	2·60E-04
D	6	*HLA-C*15:02*	31346171	5404	0·03	0·02	2·66 [1·53, 4·63]	5·54E-04
	6	*HLA-B*40:02*	31431272	5404	0·05	0·02	2·33 [1·37, 3·98]	1·88E-03
	6	*HLA-C*02:02*	31346171	5404	0·14	0·06	1·84 [1·25, 2·70]	1·92E-03
	6	*HLA-C*04:03*	31346171	5404	0·00	0·00	23·99 [2·00, 288·00]	1·22E-02
	6	*HLA-C*02:06*	31346171	5404	0·00	0·00	23·95 [1·99, 288·90]	1·24E-02
	6	*HLA-C*03:05*	31346171	5404	0·00	0·00	23·95 [1·99, 288·90]	1·24E-02

CHR	SNP	BP	N	MAF case	MAF control	OR [95% CI]	P
6	*HLA-C*03:06*	31346171	5404	0·00	0·00	23·95 [1·99, 288·90]	1·24E-02
6	*HLA-C*04:04*	31346171	5404	0·00	0·00	23·95 [1·99, 288·90]	1·24E-02

CHR=chromosome; SNP=single nucleotide polymorphism; BP=base pair; MAF=minor allele frequency; OR=odds ratio; CI=confidence interval.

Supplementary Table S4. LD (r^2) between the top SNPs and HLA alleles

	rs199564443	rs1071816	rs1811197	rs114291795	rs652888	*HLA-B*27:05*	*HLA-B*08:01*
rs199564443	NA	0	0	0	0	0	0
rs1071816	0	NA	0·044	0·499	0·004	0·501	0·017
rs1811197	0	0·044	NA	0·013	0·325	0·013	0·679
rs114291795	0	0·499	0·013	NA	0·006	0·852	0·009
rs652888	0	0·004	0·325	0·006	NA	0·004	0·4
*HLA-B*27:05*	0	0·501	0·013	0·852	0·004	NA	0·009
*HLA-B*08:01*	0	0·017	0·679	0·009	0·4	0·009	NA

Supplementary Table S5. Univariate odds ratios for the top HLA alleles for antithyroid agents.

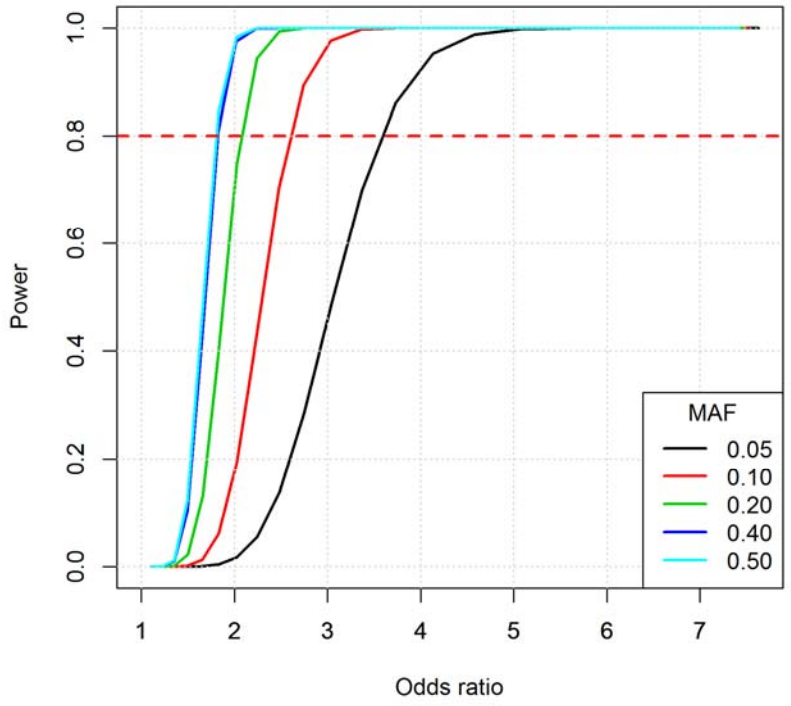
The effect is modelled per increase of one present HLA allele. A) Adjusted by sex and genetic principal components 1-4; B) adjusted by sex, genetic principal components 1-4 and *HLA-B*27:05*;

	CHR	SNP	BP	N	MAF case	MAF control	OR [95% CI]	P
A	6	*HLA-B*27:05*	31431272	5209	0·19	0·06	5·44 [2·94, 10·06]	6·88E-08
	6	*HLA-C*02:02*	31346171	5209	0·23	0·06	4·34 [2·51, 7·51]	1·56E-07
	6	*HLA-B*08:01*	31431272	5209	0·26	0·11	2·93 [1·73, 4·96]	6·18E-05
	6	*HLA-A*01:01*	30019970	5209	0·28	0·14	2·33 [1·40, 3·88]	1·18E-03
	6	*HLA-B*39:01*	31431272	5209	0·04	0·01	6·77 [1·96, 23·40]	2·54E-03
	6	*HLA-C*07:01*	31346171	5209	0·26	0·14	2·15 [1·28, 3·62]	3·72E-03
	6	*HLA-B*39:06*	31431272	5209	0·03	0·01	5·88 [1·34, 25·87]	1·91E-02
	6	*HLA-B*07:02*	31431272	5209	0·04	0·15	0·27 [0·08, 0·84]	2·45E-02
	6	*HLA-A*31:01*	30019970	5209	0·06	0·03	2·56 [1·00, 6·58]	5·07E-02
	6	*HLA-B*15:26*	31431272	5209	0·01	0·00	17·26 [0·92, 322·70]	5·66E-02
B	6	*HLA-B*08:01*	31431272	5209	0·26	0·11	3·88 [2·21, 6·81]	2·20E-06
	6	*HLA-C*07:01*	31346171	5209	0·26	0·14	2·82 [1·63, 4·89]	2·25E-04
	6	*HLA-A*01:01*	30019970	5209	0·28	0·14	2·46 [1·47, 4·12]	6·30E-04
	6	*HLA-B*39:01*	31431272	5209	0·04	0·01	7·66 [2·17, 27·06]	1·58E-03
	6	*HLA-B*39:06*	31431272	5209	0·03	0·01	7·73 [1·70, 35·09]	8·05E-03
	6	*HLA-C*02:02*	31346171	5209	0·23	0·06	2·45 [1·18, 5·10]	1·67E-02
	6	*HLA-B*40:02*	31431272	5209	0·06	0·02	2·67 [1·05, 6·78]	3·86E-02
	6	*HLA-A*02:01*	30019970	5209	0·23	0·35	0·58 [0·34, 1·01]	5·33E-02
	6	*HLA-B*15:26*	31431272	5209	0·01	0·00	18·43 [0·95, 356·40]	5·38E-02
	6	*HLA-B*07:02*	31431272	5209	0·04	0·15	0·33 [0·10, 1·06]	6·16E-02

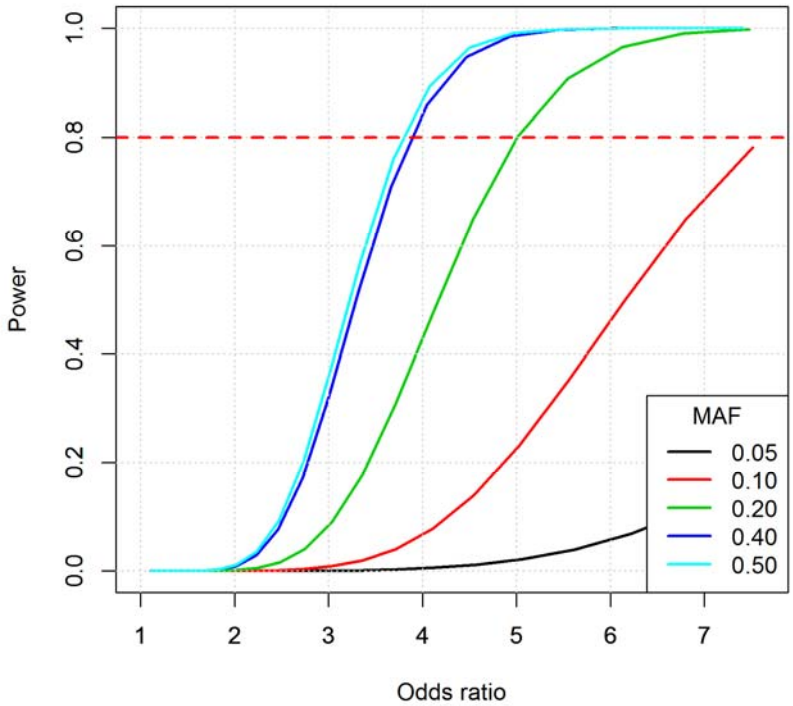
CHR=chromosome; SN*P*=single nucleotide polymorphism; B*P*=base pair; MAF=minor allele frequency; OR=odds ratio; CI=confidence interval.

Supplementary figures

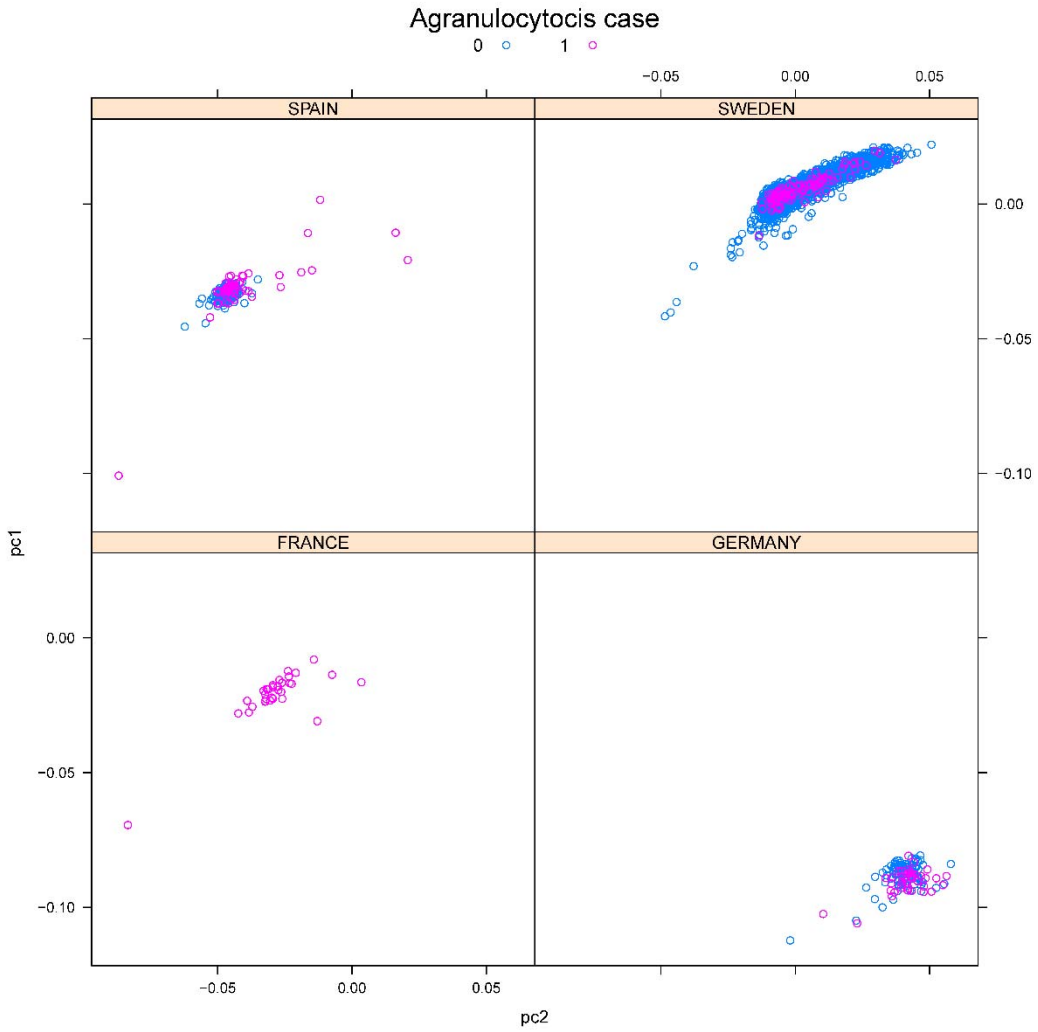
A



B

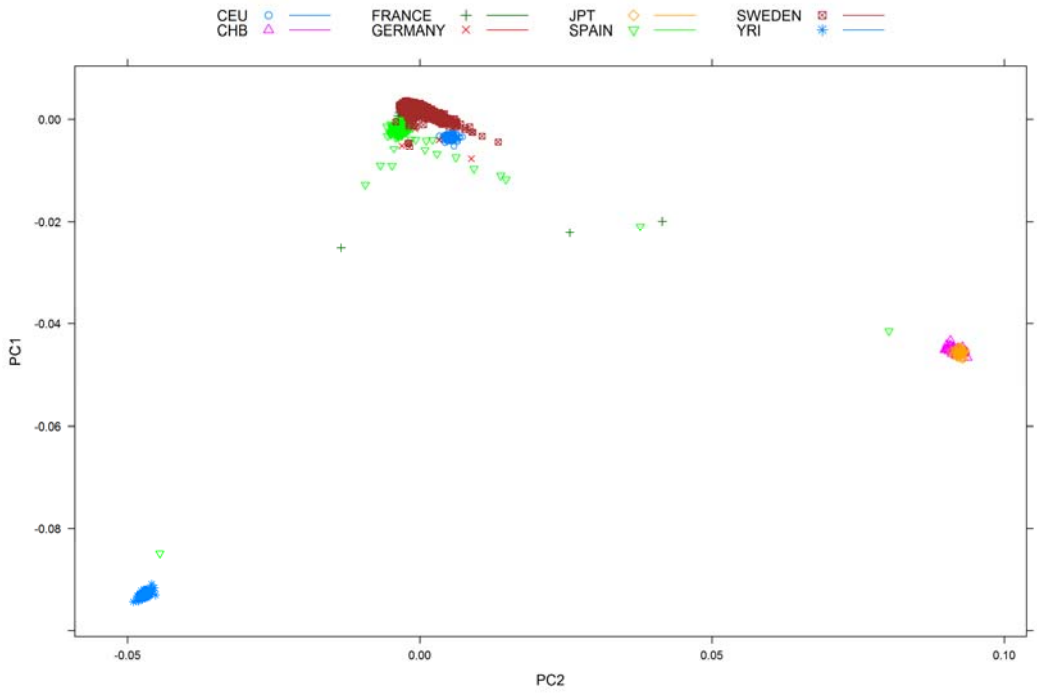


Supplementary Figure S1. Power curves for GWAS. These curves shows the power to find a genome-wide significant association ($p=5*10^{-8}$) given a minor allele frequency between 0·05 and 0·50 with a study size of A) 234 cases and 5100 controls, B) 39 cases and 5100 controls. The statistical test used is the Chi^2 approximation.

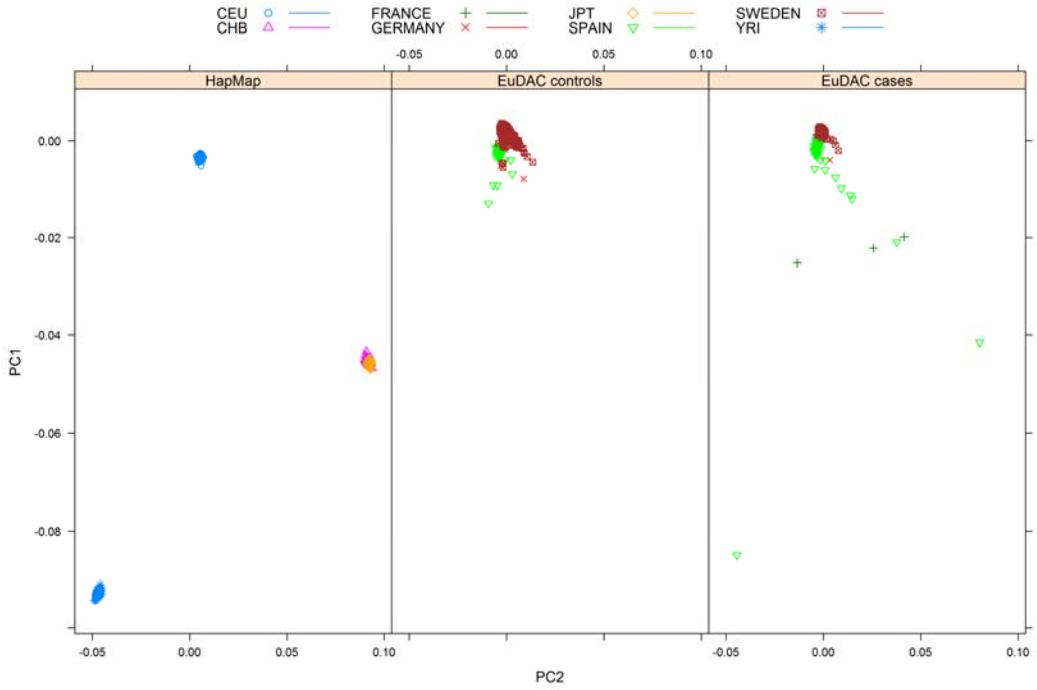


Supplementary Figure S2. Plot of the first two genetic principal components showing the population structure of the EuDAC material. The plot is divided in panels by country and colours are by case (=1) and control (=0).

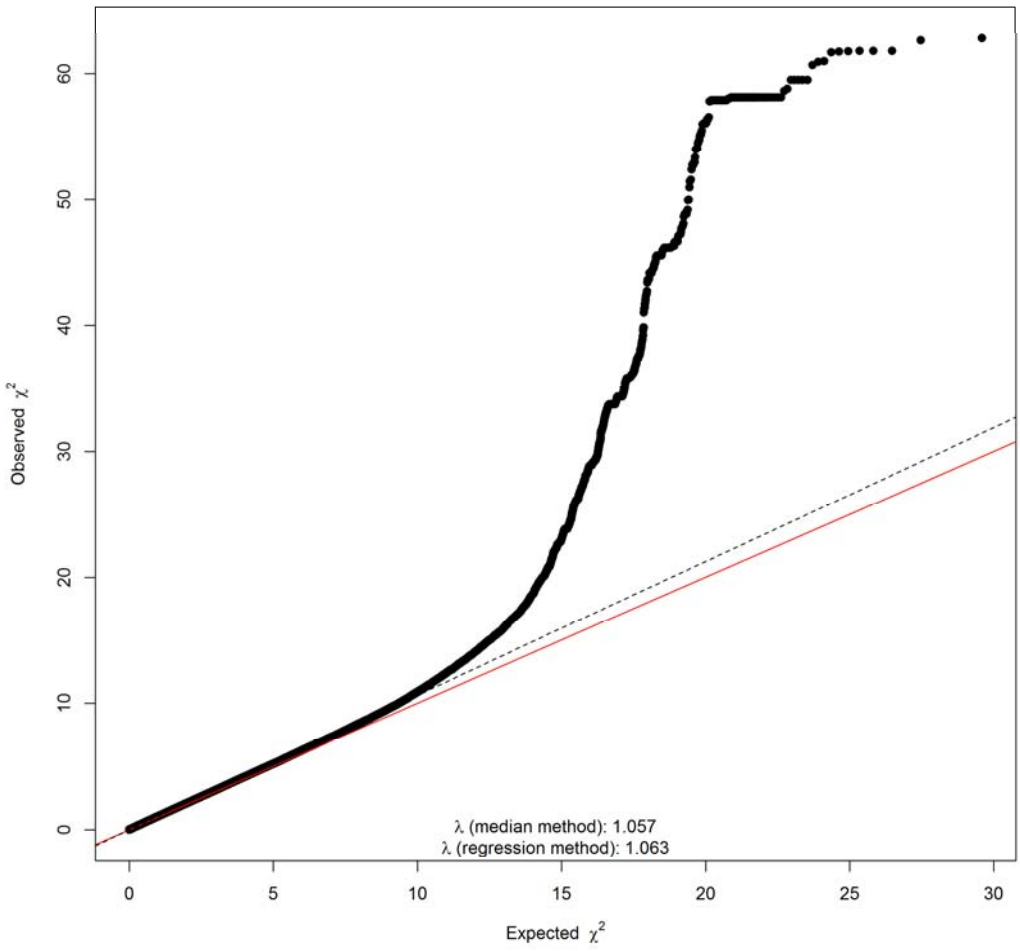
A



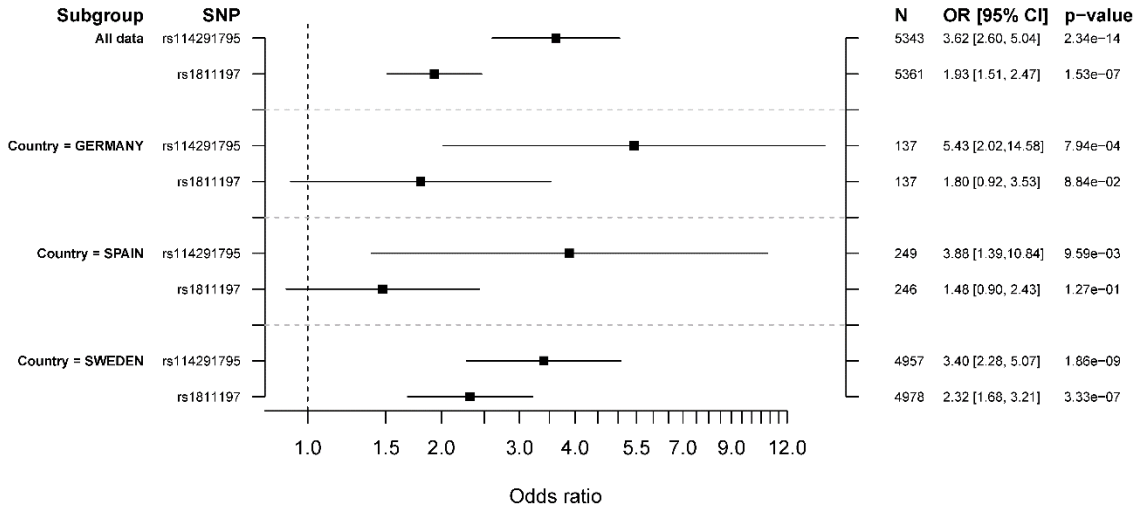
B



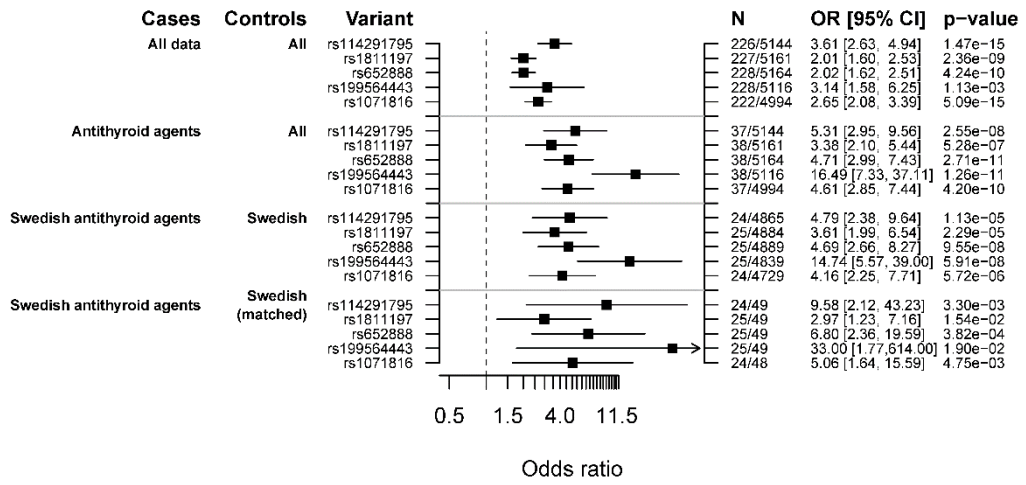
Supplementary Figure S3. Population structure when projected on HapMap coordinates. A) Hapmap and EuDAC data B) Hapmap and EuDAC data divided in panles by Hapmap, EuDAC controls and EuDAC cases. The Hapmap populations are CEU (Utah residents with ancestry from northern and western Europe), JPT (Japanese in Tokyo, Japan), CHB (Han Chinese in Beijing, China) and YRI (Yoruba in Ibadan, Nigeria).



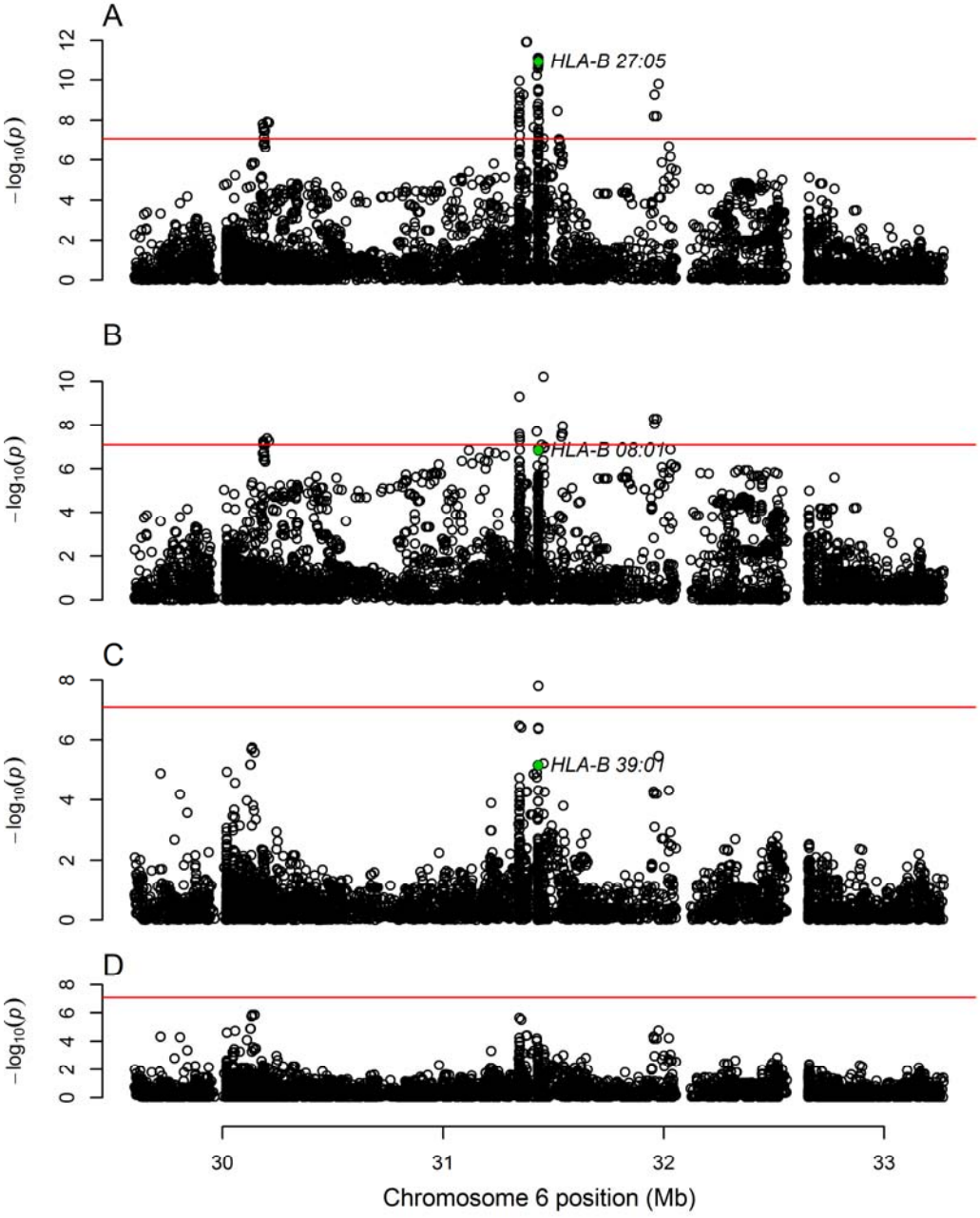
Supplementary Figure S4. Q-Q plot for analysis of all agranulocytosis cases vs all controls after imputation. Adjusted by sex, genetic principal components 1-4.



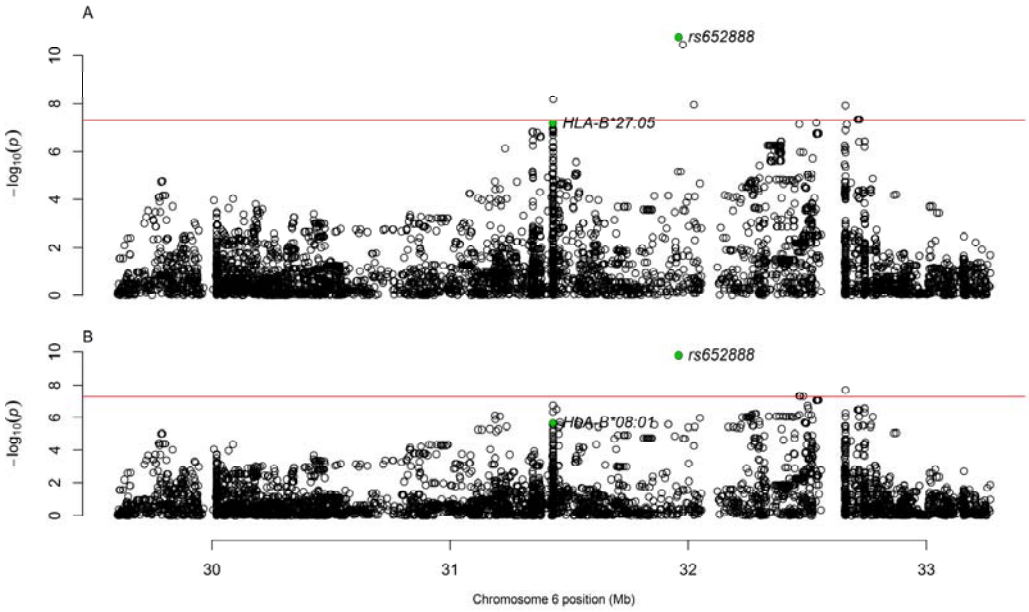
Supplementary Figure S5. Forest plot of estimated odds ratios (OR) in total and by country for the two top single nucleotide polymorphisms (SNPs) rs114291795 and rs1811197. No stratification by drug/drug-group. Cases from France are excluded in this analysis because of lack of controls. SNP effects are modeled as additive.



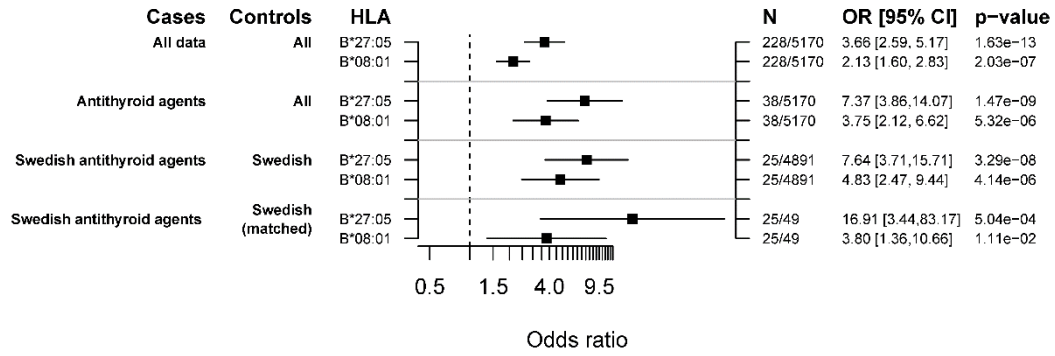
Supplementary Figure S6. Sensitivity analysis excluding the 6 genetic outliers found through PCA. Forest plot of estimated univariate odds ratios (OR) with 95% confidence intervals (CI) showing the top single nucleotide polymorphisms (SNPs) on chromosome 6, rs114291795, rs1811197, rs652888, rs19956443 and rs1071816, for all cases vs all controls, and antithyroid-induced cases vs all controls. Also shown are Swedish antithyroid drug-induced cases vs all Swedish controls as well as Swedish antithyroid drug-induced cases vs Swedish controls matched for hyperthyroidism. Note that none of the matched Swedish controls had the variant rs199564443, OR, CI and P was calculated manually and adjustment for gender and PC 1-4 was therefore not possible. Numbers of cases/controls are given in the N column.



Supplementary Figure S7. Manhattan plots for the analyses of the MCH region when adjusting for each top four digit HLA allele. A) Initial analysis, B) adjusted for *HLA-B*27:05*, C) adjusted for *HLA-B*27:05* and *HLA-B 08:01* and D) adjusted for *HLA-B*27:05, HLA-B*08:01* and *HLA-B*39:01*. Note that although these analyses focuses on the four digit HLA alleles, all variants are still present in the plots.



Supplementary Figure S8. Manhattan plots for the analyses of the major histocompatibility complex (MHC) region including imputed single nucleotide polymorphisms (SNPs) and human leukocyte antigen (HLA) alleles. A) Initial analysis of cases induced by antithyroid agents vs all controls, B) adjusted for the top hit allele *HLA-B*27:05*. The red line displays the threshold for genome-wide significance of $5*10^{-8}$.



Supplementary Figure S9. Sensitivity analysis excluding the 6 genetic outliers found through PCA. Forest plot of estimated odds ratios (OR) with 95% confidence intervals (CI) based on a multiple model for the top HLA-B alleles showing A) all cases vs all controls, B) all cases induced by antithyroid agents vs all controls, C) all Swedish cases induced by antithyroid agents vs all Swedish controls, and D) all Swedish cases induced by antithyroid agents vs matched Swedish controls. The results are from a multiple model including both variants. Matched controls have been treated for hyperthyroidism. Numbers of cases/controls are given in the N column.

Supplementary Figure S10. Plot of sensitivity and specificity in the data when different cutoffs are applied to the predicted probability of the model in figure 7. The red circle and x marks the sensitivity and specificity at the optimal cutoff that maximizes both measures. In our case the maximum is reached when a cutoff of p > 0·005 is used which gives a sensitivity of 84·2% and a specificity of 86·1%.