

Description of Additional Supplementary Files

File name: Supplementary Data 1

Description: Description of the study individuals. Descriptive statistic for the Icelandic and Swedish sample sets with IgG subclass measurements. Showin the the number of individuals used in the analysis, and number of males and females separately, the mean age and standard deviation (SD), and the average IgG subclass values and SD. For Iceland the results are shown separately for children, younger than 18 years, and adults.

File name: Supplementary Data 2

Description: Association of IgG subclass variants with all IgG subclasses. Association results for the 14 variants that associate with the levels of any of the IgG subclasses in adults, for all the IgG subclasses and for adult and children measurements separately. The table shows for each variant, the loci, its position in NCBI Hg38 build, the rs-ID, the nearest gene, other allele (OA), effect allele (EA), effect allele frequency (EAF) as average over frequencies in Iceland and Sweden, and the unadjusted association results (without adjustment of the effects of other variants) as effect β and P values, from linear regression of adjusted and standardized IgG subclass levels on genotype count. For adults the association results shown are from a meta-analysis of associations results in the Icelandic and the Swedish datasets. n=4,334 adults, n=4,571 children.

File name: Supplementary Data 3

Description: Associated variants affecting coding of genes. List of protein altering variants included in the credible sets of each of the IgG subclass signals and that are highly correlated ($r^2 > 0.8$) with the corresponding lead variant. Included in the table is for each protein altering (coding) variant its rs-ID, position in NCBI Hg38 build, other allele (OA), effect allele (EA), effect allele frequency (EAF) as average over frequencies in Iceland and Sweden, the association, effect β and P value, from logistic regression of adjusted and standardized IgG subclass levels on genotype count with two-sided P values, correlation r^2 with the lead IgG subclass associated variant, the gene that is affected, the coding effect and the coding change. n=4,334 adults.

File name: Supplementary Data 4

Description: Panel A) shows the association of the lead IgG subclass variants with variation in mRNA expression levels (eQTL) in whole blood from 17,848 individuals. Included in the table is the IgG1 subclass, chromosome and position in NCBI Hg38 build, rs-ID, other allele (OA), effect allele (EA), effect allele frequency (EAF) as average over frequencies in Iceland and Sweden, the association, effect β and P value, for the lead variant with IgG subclass levels, and the association, effect β and P value, of the variant with the mRNA expression levels of the listed gene. Also included is the lead cis-eQTL variant for the gene, its rs-ID, position, the association, effect β and P value, with the mRNA expression levels, and the correlation r^2 between the lead IgG subclass variant and the best cis-eQTL variant. Panels B), C), D) and E), show the corresponding results for the association of rs4795397, the lead IgG4 variant at 17q21.1 locus, with B) the mRNA expression levels in B cells measured in 758 individuals; C) mRNA expression levels in CD4+ cells measured in 837 individuals; D) mRNA expression levels in CD8+ cells measured in 807 individuals; and E) protein levels measured in plasma

cells from 39,155 individuals. All effects β and P values come from linear regression tests of the adjusted and standardized IgG subclass levels, mRNA expression levels, or protein levels on genotype count.

File name: Supplementary Data 5

Description: Genomic variants associated with each the IgG subclasses in children. Association results for the 5 lead variants that associate with IgG subclasses levels in children. The table shows for each variant, the IgG subclass, the loci, its position in NCBI Hg38 build, the rs-ID, the nearest gene, other allele (OA), effect allele (EA), effect allele frequency (EAF) in Iceland, and the association results, effect β and P values, from logistic regression of adjusted and standardized IgG subclass levels on genotype count with two-sided P values. The table also includes the meta-analysis results of association of the variants with IgG subclass levels in adults, the lead IgG subclass variants in adults at that locus and the correlation r^2 between the lead variants in children and the lead variants in adults. Pbonf is the Bonferroni adjusted P value, i.e. the P value adjusted for the significance threshold for the corresponding variant class the variant belongs to. V_e is the variance explained by the variant of the corresponding IgG subclass measurements in children. $n=4,571$ children.

File name: Supplementary Data 6

Description: Allotypes defined by missense variants in IGHG genes. Allotypes defined missense variants in the IGHG genes at the 14q32.33 locus. For each gene, the table includes the conventional allotype name, the IMGT nomenclature (IMGT no), and the position of the amino acid within each IgG subclass protein (Eu number) given above. The allotype defining amino acid change (in bold) by the variants listed with the respective rsID at the bottom. Single letter codes are used for the AAs, Lysine (K), Aspartic Acid (D), Leucine (L), Valine (V), Alanine (A), Arginine (R), Glutamic Acid (E), Methionine (M), Glycine (G), Isoleucine (I), Serine (S), Proline (P), Tyrosine (Y), Threonine (T), Asparagine (N), Glutamine (Q) and Phenylalanine (F).

File name: Supplementary Data 7

Description: Association of missense variants defining the IGHG allotypes with the respective IgG subclass levels in adults. The table includes for each of the missense variants in the IGHG genes at 14q32.33, the position in NCBI Hg38 build, the rs-ID, other allele (OA), effect allele (EA), effect allele frequency (EAF) as average over frequencies in Iceland and Sweden, a meta-analysis of the association, effect β and P value, from logistic regression of adjusted and standardized IgG subclass levels on genotype count with two-sided P values in the Icelandic and Swedish datasets, the corresponding amino acid (Pgnomad) and nucleotide substitution (coding change). NA indicates the variant is either absent in the Icelandic and Swedish datasets or is too rare to be tested for association with IgG subclass levels. $n=4,334$ adults.

File name: Supplementary Data 8

Description: Conditional analysis of allotype association with IgG subclass levels. A) Comparison of the lead allotype association for each of the IGHG genes with the corresponding association of the

lead variants at the locus. Shown is the tested allotype, indicated by conventional name and IMGT no, the IgG subclass, the frequency of the allotype as average over the frequency in Iceland and Sweden, the association, effect β , and P value, from linear regression of the IgG subclass levels on allotype count, and the association, effect β_{adj} and P_{adj} , tested conditional on the association of the corresponding lead genetic variant. Also included is the lead genetic variant, its EAF, the correlation r^2 between the variant and the allotype, effect β and P value, from linear regression of the IgG subclass levels on genotype count of the variant, and the association, effect β_{adj} and P_{adj} of the variant tested conditional on the association with the allotype. B) For allotypes that associate with their corresponding IgG subclass levels, this panel shows the association, effect β and P value, and the association, effect β_{adj} and P_{adj} , tested conditional on other allotypes, listed in the covariate column, that associate with same IgG subclass levels. C) For allotypes in one IGHG gene that associate with IgG subclass levels for a different gene, this panel shows the association, effect β and P value, and the association, effect β_{adj} and P_{adj} , tested conditional on the allotype, listed in the covariate column, that associate strongest with the IgG subclass level tested. $n=4,334$ adults.

File name: Supplementary Data 9

Description: Association of allotypes in IGHG genes with IgG subclass levels. Associations, β and P values, from linear regression of each of the IgG subclass levels on allotype count with two-sided P values for each of the tested allotypes. The table includes the conventional allotype name, the IMGT nomenclature (IMGT no) and the allotype defining amino acids given as single letters codes as explained in supplementary table 6, the alleles of corresponding missense variants that define the allotype, the frequency of the allotype in the Icelandic and Swedish datasets separately, and the association results for each of the IgG subclass for adults ($n=4,334$), results for Iceland and Sweden separately and combined, and for children ($n=4,571$).

File name: Supplementary Data 10

Description: Associations of IgG subclass associated variants in the GWAS Catalogue. Look up in the GWAS catalog for association of variants in the 95% credible sets for each of the IgG subclass signals. The table includes the GWAS catalog trait, the position in Hg Build 38 and the rs-ID of the listed variant, the effect allele (EA), other allele (OA), effect (OR or β) and P value for the reported association, the correlation with the lead IgG subclass variant, the association, β and P values, of the variant with IgG subclass levels, the concordance of the effect directions between the reported association and the IgG subclass association, and a reference as PubMed Unique Identifier (PMID) for the published results. Association is calculated using logistic regression of adjusted and standardized IgG subclass levels on genotype count with two-sided P values.

File name: Supplementary Data 11

Description: Association of known variants at the FCGR loci with IgG subclass values in children and adults. Association of known coding variants at the FCGR locus at 1q23.3 with IgG subclass levels in children and adults. For each variant the table shows the rsID, the gene, the position in NCBI Hg38 build, other allele (OA), effect allele (EA), effect allele frequency (EAF) as average over frequencies in Iceland and Sweden, variant type, and association, effect β and P values, from linear regression of adjusted and standardized IgG subclass levels on genotype count. For adults the association results

shown are from an meta-analysis of associations results in the Icelandic and the Swedish datasets. n=4,334 adults, n=4,571 children.

File name: Supplementary Data 12

Description: Association of classic HLA alleles with IgG subclass levels. Association of classical HLA alleles that are strongly correlated with IgG subclass associated sequence variants, rs2647032 for IgG1 and rs3763321 for IgG4. The table shows the HLA allele, the correlation r^2 with the lead sequence variant, rs2647032 for IgG1 and rs3763321 for IgG4, in the Icelandic and Swedish datasets, separately, the frequency and association, effect β and P values from linear regression of adjusted and standardized IgG subclass levels on genotype count, for Iceland and Sweden, separately, and effect β and P values from a fixed effect inverse variance weighted meta-analysis of the Icelandic and Swedish results. Also included is a P values, Phet, for a test of heterogeneity in the effect estimates in Iceland and Sweden. n=4,334 adults, n=4,571 children.

File name: Supplementary Data 13

Description: Association of IGHG1 deletion with gene expression. Association of IGHG1 deletion and mRNA expression of CD19, MS4A1 and IGHG1-4. Effect on expression is given as β that represents log2 difference in mRNA expression between carriers and non-carriers of the IGHG1 deletion. Effect_p is the change in mRNA expression calculated as a percentage change on a linear scale. Association is calculated using logistic regression of expression on genotype count with two-sided P values. n=17,381 individuals.

File name: Supplementary Data 14

Description: Disease diagnosis of the 4 homozygot carriers of the IGHG1 deletion. IGHG1 deletion homozygotes often have asthma, acute sinusitis, cardiopathies myalgia and pain.

File name: Supplementary Data 15

Description: Association of variants at the FCGR loci with IgG1 in children and adults. Association of three correlated lead variants at the FCGR locus with IgG1 in children (n=4,571) and adults (4,334). The variants are a missense variant rs9427397 in FCGR2A that has the most significant correlation with IgG1 in children, and rs143596860 and rs7554873 which capture the correlation with IgG1 in adults. Shown are the unadjusted effects β and P values, and effects, β_{adj} and P_{adj} for rs9427397 adjusted for rs143596860 and rs7554873, and effects β_{adj} and P_{adj} for rs143596860 and rs7554873 adjusted for rs9427397. OA, EA and EAF are reference allele, effect allele and effect allele frequency, respectively. The effect allele frequency (EAF) is calculated as a simple average over the frequency in the Icelandic and Swedish datasets. Association is calculated using logistic regression of adjusted and standardized IgG1 subclass levels on genotype count with two-sided P values.

File name: Supplementary Data 16

Description: Association of IgG4 variants with IgE levels. Effect of IgG4 associating variants on IgE levels measured in 26,638 Icelanders. Included is the loci (cytoband location), position in NCBI Hg38 Build, variant id, nearest gene, other (OA) and effect (EA) allele, effect allele frequency (EAF), and effect β and P value from linear regression of standardized IgE levels on genotype count of the IgG4 variants, adjusting for age, sex, year of birth and county of origin. The effect estimates are in standard deviations of the distribution of the IgE measurements.

File name: Supplementary Data 17

Description: eQTL association at 17q21.1 in CD138+ plasma cells. Association of variants correlated with IgG4 subclass lead variant rs4795397 at 17q21.1 with expression of CD138+ immunomagnetic bead-enriched plasma cells (PMID: 25480495). The table shows eQTL variants from a plasma cell RNA sequencing dataset, its position in NCBI Hg38 build, other alleles (OA), effect allele (EA), and effect allele frequency (EAF) as a simple average over frequencies in Iceland and Sweden. For the eQTL association, the table shows the gene, the median gene expression TPM, calculated for a subset of 221 samples from the cohort, the effect direction, association P value from logistic regression of expression values on genotype count, and the rank of the variant, which refers for the lowest P value of the eQTL variant among all eQTL variants found for a given gene. For the association of the variants with IgG4 levels the table shows the effect β and P values from a meta-analysis of results from logistic regression of standardized IgG4 levels on genotype counts, for Iceland and Sweden respectively, and the correlation r^2 of the eQTL variant with the lead IgG4 variant in the region, rs4795397. For two of the genes, GSDMB and ORMDL3, no eQTL association was found for the transcripts. Association is calculated using logistic regression of expression on genotype count with two-sided P values.

File name: Supplementary Data 18

Description: Luciferase probe sequences. 121-bp long DNA sequences centered on the variant of interest (highlighted in red) were used in luciferase assays.

File name: Supplementary Data 19

Description: Colocalization analysis at 17q21.1. Pair-wise colocalization analysis for correlated association signals for immunoglobulins (IgG4, IgM), various autoimmune diseases, proteomics pQTLs (FAIM3, IL5RA), asthma and eQTLs in B, CD4 positive T, and CD8 positive T cells at the 17q21.1 locus. The colocalization is performed with the coloc R package and the table shows the posterior probability for a one shared genetic signal (PP4) in the upper right corner and the posterior probability for two distinct genetic signals (PP3) in the lower left corner. For each trait, the table includes the variant (rs-ID), position in Hg38, the two-sided P-value for the lead association signal, and the correlation with the lead IgG4 and Asthma variants. In case of multiple signals for a trait at this locus, conditional analysis was used to remove effects of the secondary signals. The light grey box highlights the association signals that primarily co-localize with the IgG4 association signal, while the dark grey box highlights the association signals that primarily co-localize with the adult asthma signal.