

A loss-of-function variant in *ALOX15* protects against nasal polyps and chronic rhinosinusitis

Ragnar P. Kristjansson¹, Stefania Benonisdottir¹, Olafur B. Davidsson¹, Asmundur Oddsson¹, Vinicius Tragante^{1,2}, Jon K. Sigurdsson¹, Lilja Stefansdottir¹, Stefan Jonsson¹, Brynjar O. Jensson¹, Joseph G. Arthur¹, Gudny A. Arnadottir¹, Gerald Sulem¹, Bjarni V. Halldorsson^{1,3}, Bjarni Gunnarsson¹, Gisli H. Halldorsson¹, Olafur A. Stefansson¹, Gudjon R. Oskarsson¹, Aimee M. Deaton¹, Isleifur Olafsson⁴, Gudmundur I. Eyjolfsson⁵, Olof Sigurdardottir⁶, Pall T. Onundarson^{7,8}, David Gislason^{7,9}, Thorarinn Gislason^{7,10}, Bjorn R. Ludviksson^{7,11}, Dora Ludviksdottir^{7,12}, Thorunn A. Olafsdottir^{1,7}, Thorunn Rafnar¹, Gisli Masson¹, Florian Zink¹, Gyda Bjornsdottir¹, Olafur Th. Magnusson¹, Unnur S. Bjornsdottir^{9,13}, Gudmar Thorleifsson¹, Gudmundur L. Norddahl¹, Daniel F. Gudbjartsson^{1,14}, Unnur Thorsteinsdottir^{1,7}, Ingileif Jonsdottir^{1,7}, Patrick Sulem^{1*} and Kari Stefansson^{1,7*}

Nasal polyps (NP) are lesions on the nasal and paranasal sinus mucosa and are a risk factor for chronic rhinosinusitis (CRS). We performed genome-wide association studies on NP and CRS in Iceland and the UK (using UK Biobank data) with 4,366 NP cases, 5,608 CRS cases, and >700,000 controls. We found 10 markers associated with NP and 2 with CRS. We also tested 210 markers reported to associate with eosinophil count, yielding 17 additional NP associations. Of the 27 NP signals, 7 associate with CRS and 13 with asthma. Most notably, a missense variant in *ALOX15* that causes a p.Thr560Met alteration in arachidonate 15-lipoxygenase (15-LO) confers large genome-wide significant protection against NP ($P=8.0 \times 10^{-27}$, odds ratio = 0.32; 95% confidence interval = 0.26, 0.39) and CRS ($P=1.1 \times 10^{-8}$, odds ratio = 0.64; 95% confidence interval = 0.55, 0.75). p.Thr560Met, carried by around 1 in 20 Europeans, was previously shown to cause near total loss of 15-LO enzymatic activity. Our findings identify 15-LO as a potential target for therapeutic intervention in NP and CRS.

Nasal polyps (NP) are non-malignant, typically bilateral growths on the mucosa of the nasal passages and paranasal sinuses¹. NP affect up to 4% of the population¹. They may be asymptomatic, but they often cause nasal obstruction, rhinorrhea, and postnasal drip². NP are a significant risk factor for chronic rhinosinusitis (CRS) and can affect quality of life considerably by causing facial pain, headaches, postnasal drainage, and anosmia³. The presence of polyps has been linked to a greater than average severity of sinusitis as well as delayed recovery^{4,5}. The cause of NP remains unclear, although chronic inflammation is thought to play a role⁶. It has been postulated that NP and CRS may be two expressions of the same disease⁷.

CRS with nasal polyps (CRSwNP) and asthma often coexist and pose a therapeutic challenge, as asthma accompanied by nasal polyposis is often severe and resistant to treatment⁸. NP are also often seen in conjunction with allergies and other inflammatory respiratory conditions, notably CRS. They are commonly infiltrated with

eosinophils and type 2 helper T cells that secrete cytokines such as interleukin-5 (IL-5), IL-13, and IL-17¹. Current therapeutic strategies for nasal polyposis focus on treatment as well as prevention, notably with local or systemic corticosteroids, aspirin desensitization when indicated, and surgery. More recently, treatment includes novel anti-immunoglobulin E (omalizumab) and anti-IL-5 (mepolizumab or reslizumab) biologics^{9–11}. Recurrence of NP after treatment is common, and corticosteroids are the primary preventive measure¹. Long-term use of systemic corticosteroids has serious side effects and is generally avoided¹.

To our knowledge, no sequence variants have been reported to significantly associate with risk of NP or CRS. To search for sequence variants affecting risk of these diseases, we performed genome-wide association studies (GWAS) of NP and CRS using combined datasets from Iceland and the UK. We further assessed the variants detected for association with asthma in the same two populations.

¹deCODE genetics/Amgen Inc., Reykjavik, Iceland. ²Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands. ³School of Science and Engineering, Reykjavik University, Reykjavik, Iceland. ⁴Department of Clinical Biochemistry, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ⁵The Laboratory in Mjodd, RAM, Reykjavik, Iceland. ⁶Department of Clinical Biochemistry, Akureyri Hospital, Akureyri, Iceland. ⁷Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. ⁸Department of Laboratory Hematology, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ⁹Department of Medicine, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹⁰Department of Sleep, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹¹Department of Immunology, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹²Department of Respiratory Medicine, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹³The Medical Center Mjodd, Reykjavik, Iceland. ¹⁴School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. *e-mail: patrick.sulem@decode.is; kari.stefansson@decode.is

Results

We performed a meta-analysis by combining GWAS results on NP in individuals of European descent, from deCODE genetics in Iceland ($n=1,175$ cases and 309,305 controls) and the UK Biobank ($n=3,191$ cases and 405,376 controls). We also performed a GWAS on CRS by combining data from the same populations (Iceland, $n=3,188$ cases and 353,939 controls; UK, $n=2,420$ cases and 406,147 controls). The GWAS in Iceland were performed with 32.5 million markers identified through whole-genome sequencing of 15,220 Icelanders and subsequently imputed into chip-typed individuals, as well as first- and second-degree relatives of those individuals¹². The UK Biobank GWAS were performed with 37.1 million markers from the haplotype reference panel that were well imputed into chip-typed individuals of European ancestry in the UK Biobank. Associations were considered significant if the P value in the combined dataset was below a weighted genome-wide significance threshold based on variant annotation (see Methods and Supplementary Note)¹³. The adjusted significance thresholds are 2.6×10^{-7} for variants with high impact, 5.1×10^{-8} for variants with moderate impact, 4.6×10^{-9} for variants with low impact, 2.3×10^{-9} for other variants in deoxyribonuclease I hypersensitivity sites, and 7.9×10^{-10} for all other variants.

In the NP meta-analysis we found variants at ten loci that associate with NP at genome-wide significance (*ALOX15*, *HLA*, *IL18R1-IL1RL1*, *IL33*, *FOXP1*, *CYP2S1*, *SLC22A4*, *MYRF*, *TSLP*, and *10p14*), and in the CRS meta-analysis we found two signals at two loci corresponding to signals also detected in the NP GWAS (*ALOX15*, *HLA*; Table 1, Figs. 1 and 2, Supplementary Fig. 1, and Supplementary Data 1–3). All of the NP signals except the one in *ALOX15* are represented by common variants (minor allele frequency (MAF) $\geq 5\%$), and four of the signals are represented by or highly correlated with ($r^2 > 0.8$) coding variants. We also looked up the associations of these ten variants with NP without confirmed CRS (NPsCRS; $n=460$ cases and 306,306 controls) and CRS without confirmed NP (CRSsNP; $n=2,587$ cases and 354,540 controls) (Supplementary Table 1). Using these phenotypes, the associations observed for the ten variants never significantly differed from those in the original NP or CRS GWAS (all P values of heterogeneity (P_{het}) > 0.10), and in most instances the risk conferred by the variants therefore extends to both NP and CRS. Furthermore, excluding NP cases from CRS controls and vice versa did not change any associations (Supplementary Table 2).

NP is more common among asthmatics, although this finding may be due to ascertainment bias as this group is more likely to receive airway examination (including of the ear, nose, and throat) and subsequent diagnosis^{14,15}. To assess whether the association of these ten variants with NP is a consequence of their asthma association, we stratified the NP association analysis based on asthma status (Supplementary Tables 3 and 4). All ten variants remained nominally significant among the non-asthmatic NP cases ($P < 6.6 \times 10^{-5}$), suggesting that the association of these variants with NP is not substantially driven by increased NP ascertainment because of their asthma association. Two variants had greater effects among asthmatic NP cases ($P_{\text{het}} < 0.05$). The one variant that does not associate significantly with asthma (regardless of NP status) in our data (at *CYP2S1*) was completely driven by non-asthmatic NP cases ($P=0.69$ among asthmatic NP cases, $P_{\text{het}}=1.8 \times 10^{-8}$; Supplementary Table 3). Conversely, some of the variants showing the strongest association with asthma do not associate strongly with NP. We performed the same analysis of 34 asthma-associated variants reported in two recent studies^{16,17}, 13 of which associate significantly with NP in the combined dataset based on the number of tests performed ($P < 1.4 \times 10^{-3}$). Of these, four show significant heterogeneity between the asthmatic and non-asthmatic groups ($P_{\text{het}} < 0.05$), but never does the NP association appear to be fully driven by the asthma-positive NP cases (Supplementary Tables 5 and 6).

To further investigate the effects of any possible ascertainment bias due to asthma diagnosis, we performed an asthma analysis stratified on NP status (Supplementary Tables 7 and 8). All ten variants had greater effect estimates among asthmatics with diagnosed NP tested against population controls than asthmatics without diagnosed NP tested against population controls; the differences were significant for six of the variants ($P < 0.05/10 = 5.0 \times 10^{-3}$). We also tested the variants for association in asthmatics with NP versus asthmatics without NP (Supplementary Table 9). The same six variants associated with NP, although all ten variants showed effect directions consistent with those in the original GWAS of NP. Of the four variants that were not significant in either analysis when adjusting for multiple testing, two showed nominally significant associations ($P < 0.05$) with a stronger effect on asthma with NP than on asthma without NP (the variants at *IL33* and *MYRF*). For both of these variants, their greatest effect estimates were on NP without asthma versus controls (Supplementary Table 3), and both had greater estimated effects on asthma with NP than on asthma without NP (Supplementary Table 7). The variant at *CYP2S1* did not significantly associate with asthma in our combined dataset (Supplementary Table 7), and in the stratified analyses was significant only among NP cases without diagnosed asthma (Supplementary Tables 3 and 7). The NP associations of these three variants therefore cannot be fully driven by ascertainment bias because of asthma. The variant at *IL18R1* associated at a genome-wide significant level with asthma in our combined dataset and associated more significantly with asthma without NP than asthma with NP (Supplementary Table 7). For this variant, the interplay between asthma and NP warrants further study. Overall, our data demonstrate that ascertainment bias does not substantially affect our results.

Since NP are commonly infiltrated by eosinophils, we tested the association of the NP variants with circulating eosinophil count in Icelanders (Table 1). Of the ten variants, eight associate with eosinophil count, with the NP protective allele always associating with fewer eosinophils. Furthermore, using combined datasets of CRS and asthma from Iceland and the UK, we found that six of the ten NP signals associate with CRS, and seven of the ten associate with asthma with the same effect directions ($P < 0.05/(10 \times 2) = 2.5 \times 10^{-3}$; Table 1 and Supplementary Tables 10–12).

***ALOX15* variant association.** The variant with the largest effects on the risk of both NP and CRS is a low-frequency missense variant rs34210653[A] (NP_001131.3:p.Thr560Met, MAF=2.99%) in *ALOX15* that confers a 68% reduction in NP risk ($P=8.0 \times 10^{-27}$, odds ratio (OR)=0.32, 95% confidence interval (CI)=0.26, 0.39) and 36% reduction in CRS risk ($P=1.1 \times 10^{-8}$, OR=0.64, 95% CI=0.55, 0.75) (Table 1, Supplementary Tables 10 and 11, and Supplementary Fig. 1). The estimated ORs for both conditions are almost identical in the two populations (NP, OR_{Iceland}=0.33 and OR_{UK}=0.32; CRS, OR_{Iceland}=0.66 and OR_{UK}=0.62). *ALOX15* encodes arachidonate 15-lipoxygenase (15-LO), an enzyme that is responsible for metabolism of polyunsaturated fatty acids, notably arachidonic acid¹⁸, and is highly expressed in human airway epithelial cells, mast cells, and eosinophils¹⁹. In Iceland, p.Thr560Met associates with a decrease in the number of circulating eosinophils (effect=−0.073 s.d., $P=1.1 \times 10^{-16}$; Table 1 and Supplementary Fig. 2), consistent with results from a recent study that includes data from the UK Biobank²⁰. The missense variant has only one strong correlate ($r^2 > 0.8$), an intergenic variant rs71368508[A] (NC_000017.10:g.4521473C>A, $r^2=1.00$) that was recently reported to associate with a very small decrease in risk of a combined phenotype of allergic diseases (asthma, hay fever, and eczema; OR=0.89)²¹. We, however, do not detect effects of p.Thr560Met on asthma risk, other white blood cell counts, or allergic rhinitis based on the number of tests performed (all $P \geq 0.01$; Table 1 and Supplementary Tables 12 and 13). p.Thr560Met does not

Table 1 Association of the ten variants identified in the combined GWAS with NP, CRS, and asthma in the combined dataset															
Marker (position)	Allele (min/maj)	MAF (%)	Gene/ Locus	Coding effect	NP (Iceland + UK)			CRS (Iceland + UK)			Asthma (Iceland + UK)			Eosinophil count (Iceland; 251,307 individuals)	
					OR (95% CI)	P	P _{het}	OR (95% CI)	P	P _{het}	OR (95% CI)	P	P _{het}	Effect [SD] (95% CI)	P
rs1391371 (chr6:32636021)	T/A	19.5	HLA-DQA1	Upstream	1.47 (1.39, 1.54)	7.3 × 10 ^{−49}	0.43	1.22 (1.16, 1.28)	3.8 × 10 ^{−15}	0.33	1.17 (1.15, 1.20)	9.9 × 10 ^{−60}	0.27	0.03 (0.02, 0.04)	2.7 × 10 ^{−16}
rs1888909 (chr9:6197392)	T/C	25.3	[IL33]	Intergenic	1.34 (1.27, 1.40)	7.4 × 10 ^{−32}	0.59	1.12 (1.07, 1.18)	5.9 × 10 ^{−7}	0.98	1.12 (1.10, 1.14)	4.2 × 10 ^{−34}	0.69	0.05 (0.04, 0.06)	7.1 × 10 ^{−48}
rs1837253 (chr5:111066174)	T/C	27.7	[TSLP]	Intergenic	0.74 (0.70, 0.78)	3.7 × 10 ^{−29}	0.02	0.93 (0.89, 0.98)	3.9 × 10 ^{−3}	0.23	0.91 (0.89, 0.93)	2.3 × 10 ^{−23}	2.1 × 10 ^{−3}	−0.02 (−0.03, −0.01)	4.4 × 10 ^{−9}
rs34210653 (chr17:4632019)	A/G	2.51	ALOX15	Missense (Thr560Met)	0.32 (0.26, 0.39)	8.0 × 10 ^{−27}	0.89	0.64 (0.55, 0.75)	1.1 × 10 ^{−8}	0.70	0.94 (0.89, 0.99)	0.03	0.52	−0.07 (−0.09, −0.06)	1.1 × 10 ^{−16}
rs1444782 (chr10:9016708)	A/G	40.3	10p14	Intergenic	0.84 (0.80, 0.88)	1.1 × 10 ^{−14}	0.54	0.90 (0.86, 0.94)	3.5 × 10 ^{−7}	0.12	0.90 (0.89, 0.91)	1.8 × 10 ^{−36}	0.25	−0.02 (−0.03, −0.02)	1.1 × 10 ^{−13}
rs17718444 (chr3:71450250)	T/C	30.0	FOXP1	Intron	0.85 (0.81, 0.89)	1.3 × 10 ^{−11}	0.19	0.96 (0.92, 1.00)	0.06	0.30	0.98 (0.97, 1.00)	0.04	0.39	−0.01 (−0.02, 0.00)	0.05
rs338598 (chr19:41195623)	A/C	50.3	CYP251	Intron	0.86 (0.82, 0.90)	5.6 × 10 ^{−11}	0.81	0.92 (0.88, 0.96)	3.5 × 10 ^{−5}	0.94	0.99 (0.98, 1.01)	0.50	0.88	−0.01 (−0.02, 0.00)	0.12
rs6543124 (chr2:102370999)	A/T	38.5	IL18R1	Intron	0.87 (0.83, 0.91)	7.4 × 10 ^{−10}	0.22	0.94 (0.90, 0.98)	4.2 × 10 ^{−3}	0.75	0.93 (0.91, 0.94)	1.3 × 10 ^{−19}	0.02	−0.04 (−0.05, −0.03)	3.8 × 10 ^{−45}
rs1050152 (chr5:132340627)	T/C	40.6	SLC22A4	Missense (Leu503Phe)	0.88 (0.84, 0.92)	1.1 × 10 ^{−8}	0.77	0.92 (0.89, 0.96)	2.3 × 10 ^{−4}	0.67	0.95 (0.94, 0.97)	7.3 × 10 ^{−9}	0.24	−0.04 (−0.05, −0.03)	4.4 × 10 ^{−32}
rs174535 (chr11:61783884)	C/T	36.9	MYRF [FADS1-FADS2] ^a	Synonymous	0.87 (0.83, 0.91)	3.5 × 10 ^{−9}	0.23	0.97 (0.93, 1.01)	0.15	0.58	0.95 (0.94, 0.97)	7.0 × 10 ^{−8}	0.44	−0.01 (−0.01, −0.01)	6.1 × 10 ^{−6}
Eosinophil count association data are reported for the Icelandic dataset only. Effects are presented for the minor allele. Results by country are presented in Supplementary Tables 10–12. Average MAFs for Iceland and the UK are presented. Genes linked to intergenic variants are listed in square brackets. ^a rs174535 has been linked to these genes through eQTL associations.															

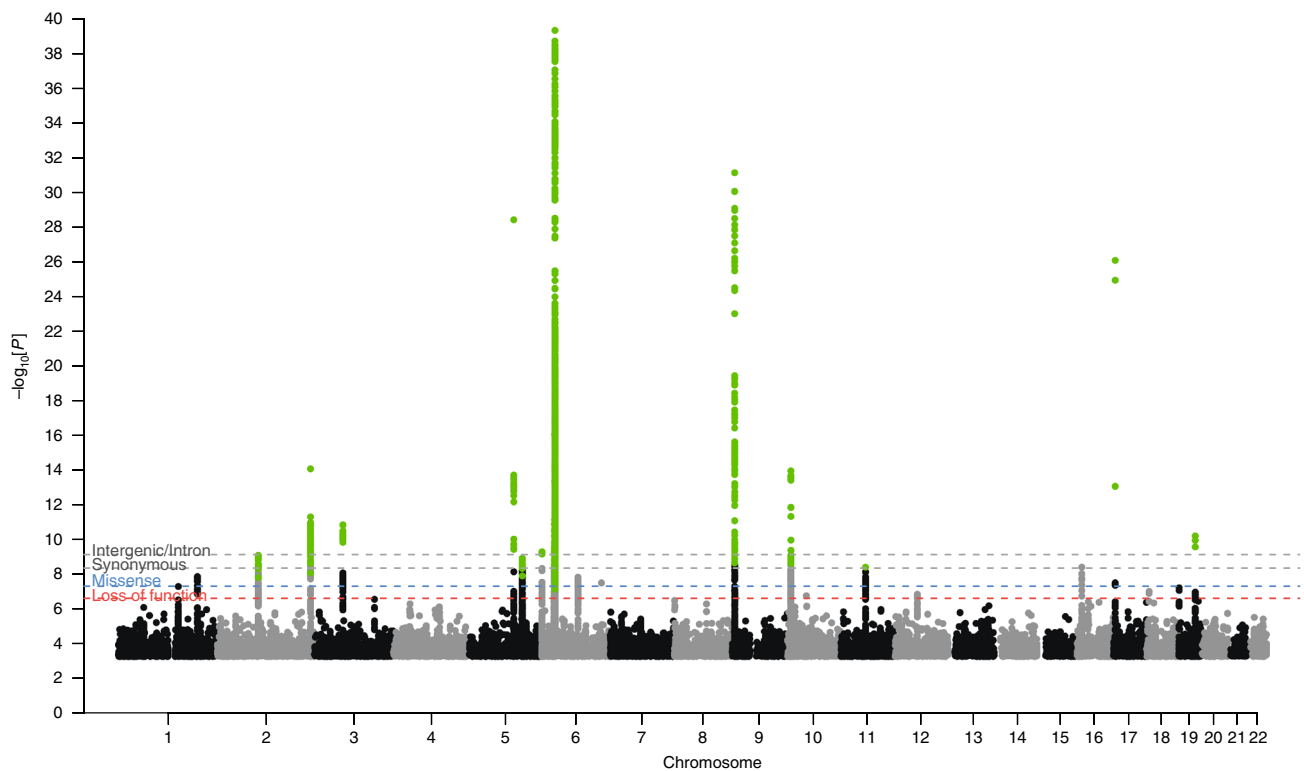


Fig. 1 | Manhattan plot for the combined nasal polyp GWAS (Iceland, $n = 1,175$ cases and 309,305 controls; UK, $n = 3,191$ cases and 405,376 controls). Ten regions harbor genome-wide significant signals. Variants with P values below their respective variant-class thresholds, indicated by horizontal dashed lines, are marked in green. Variants are plotted by chromosomal position (x axis) and $-\log_{10}[P]$ values (y axis; two-sided logistic regression). For clarity, only variants with $P < 0.005$ are plotted.

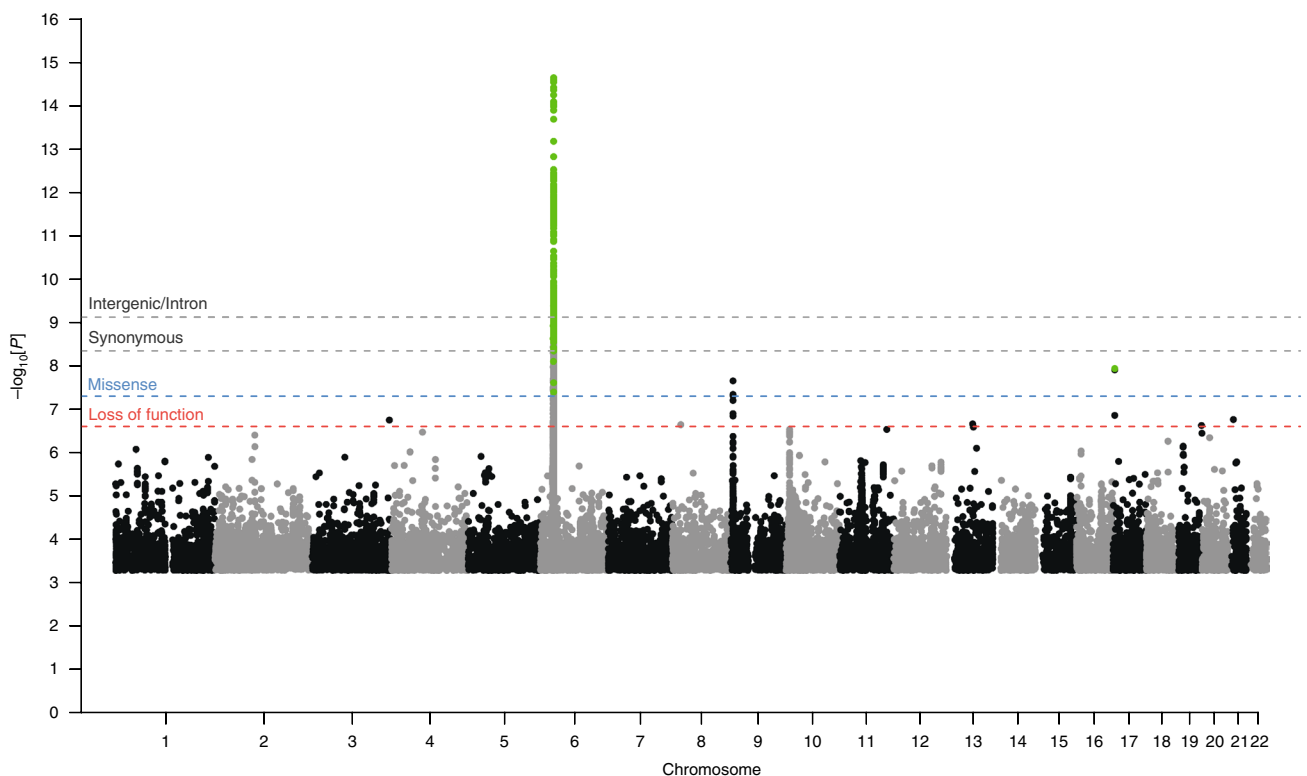


Fig. 2 | Manhattan plot for the combined chronic rhinosinusitis GWAS (Iceland, $n = 3,188$ cases and 353,939 controls; UK, $n = 2,420$ cases and 406,147 controls). Ten regions harbor genome-wide significant signals. All variants with P values below their respective variant-class thresholds, indicated by horizontal dashed lines, are marked in green. Variants are plotted by chromosomal position (x axis) and $-\log_{10}[P]$ values (y axis; two-sided logistic regression). For clarity, only variants with $P < 0.005$ are plotted.

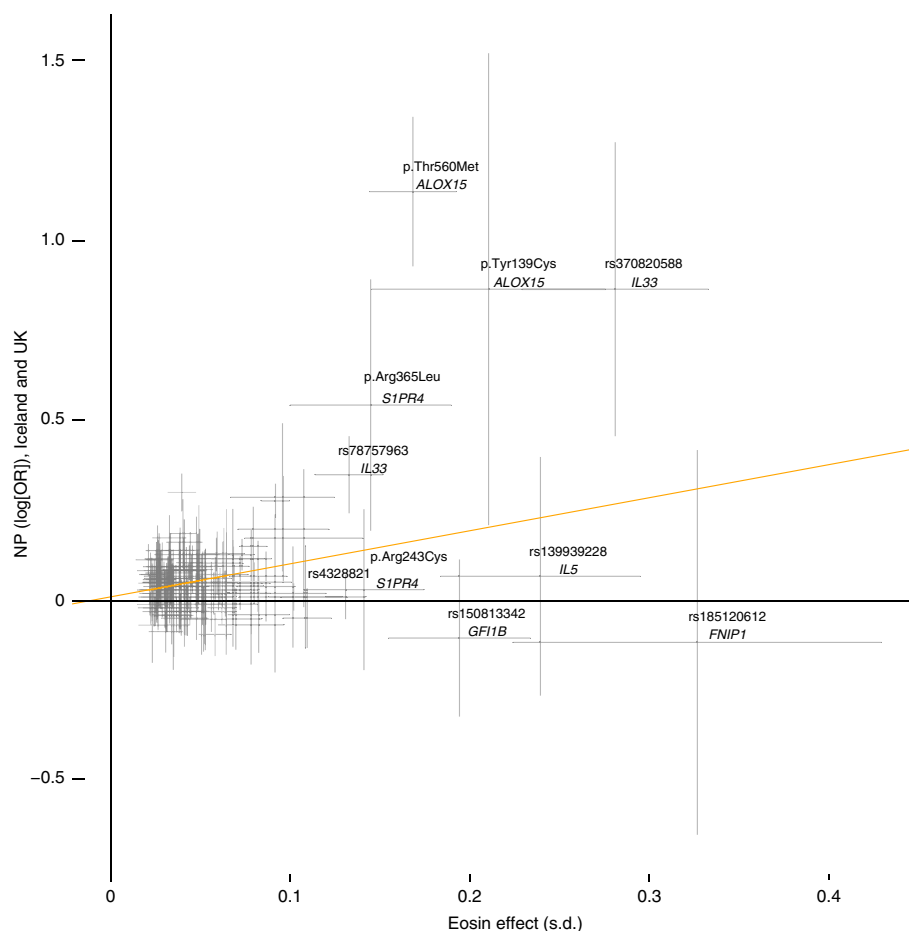


Fig. 3 | Reported eosinophil count variants and risk of nasal polyps. Scatter plot showing 180 previously reported eosinophil count SNPs²⁰. The x axis shows their effect (s.d.) on eosinophil count as reported by Astle et al.²⁰, and the y axis shows the logarithm of estimated OR for NP in a combined dataset of Iceland and the UK. All effects are shown for the eosinophil-increasing allele. Variants with the top ten largest eosinophil effects are labeled. Error bars represent 95% CIs. The solid yellow line ($y = 0.01 + 0.92x$) represents results from a simple linear regression using MAF ($1 - \text{MAF}$) as weights with $R = 0.25$ ($P = 8.2 \times 10^{-4}$; two-sided t test). These data are consistent with the fact that a set of sequence variants (for example, at *IL33* and *ALOX15*) have an effect on NP risk relative to their eosinophil effect, whereas another group of variants clearly do not associate with NP (for example, at *FNIP1*, *IL5*, and *GF11B*).

associate with any expression quantitative trait loci (eQTLs) in Icelandic blood or adipocyte expression datasets, or in the Genotype-Tissue Expression (GTEx) Portal²², and no variants were found to affect expression of *ALOX15*. Functional studies have shown that p.Thr560Met leads to near complete loss of 15-LO enzymatic activity through reduced substrate binding (residual activity is 3.8% of wild type)^{23,24}. In summary, p.Thr506Met decreases protein function, decreases risk of both NP and CRS, and decreases the number of circulating eosinophils.

For NP and CRS, the genotypic effect of p.Thr560Met is consistent with an additive model (Supplementary Fig. 2). We observe 146 homozygous carriers of p.Thr560Met in Iceland and 177 in the UK. None of these individuals have to our knowledge been diagnosed with NP, and only 7 have been diagnosed with CRS. p.Thr560Met has previously been reported in the context of coronary artery disease in a small candidate gene study with nominal association ($n = 1,809$ coronary artery disease cases and 1,734 controls, $\text{OR} = 1.62$, $P = 0.02$)^{23,24}. We were not able to replicate this association in a combined coronary artery disease dataset from Iceland and the UK ($P = 0.54$, $\text{OR} = 1.01$, 95% $\text{CI} = 0.97, 1.06$, $n_{\text{Iceland}} = 38,918$ cases and 317,957 controls, $n_{\text{UK}} = 28,110$ cases and 380,457 controls; Supplementary Table 14). A role for 15-LO as a tumor suppressor, particularly in colorectal cancer, has also been suggested²⁵.

p.Thr560Met does not associate with risk of any of the most common cancers in Iceland or in the UK (all $P \geq 0.05$; Supplementary Table 14). In addition, we did not detect any associations with other available phenotypes tested in house ($n = 4,094$) under the additive and recessive models.

Other variant associations. The nine other signals that reach genome-wide significance in the meta-analysis of NP are common; of these, seven are represented by several markers in high linkage disequilibrium (LD), whereas two of the signals are not highly correlated with other variants (variants at the *CYP2S1* and *TSLP* loci; Table 1). Two of the signals are represented by or highly correlated with ($r^2 > 0.8$) a coding variant. Seven of the signals associate with asthma (*HLA*, *IL18R1-IL1RL1*, *IL33*, *TSLP*, *SLC22A4*, *MYRF-FADS1-FADS2*, and 10p14), of which four are within or close to genes with known inflammatory function (*HLA*, *IL18R1-IL1RL1*, *IL33*, and *TSLP*) (Table 1). Conditional analysis identified no additional independent signals at the ten NP loci. Analysis of regulatory effects proved largely inconclusive (Supplementary Table 15), but most of these nine variants were linked to a single gene through a combination of biological and eQTL evidence (for more details on these common signals, see the Supplementary Note).

Table 2 | Sequence variants reported to associate with eosinophil count that associate with NP

Marker	Allele (min/ maj)	MAF (%)	Gene / [locus]	Coding effect	Eosinophil effect (s.d.)	LD class ($r^2 > 0.8$ with variant)			CRS			Asthma		
						N variants	N coding variants	NP	OR (95% CI)	P	P_{het}	OR (95% CI)	P	P_{het}
rs78757963 (chr9:6282511)	A/G	3.54	[IL33]	Intergenic	0.050	3	0	1.42 (1.27, 1.58)	1.22 (1.11, 1.36)	1.6×10^{-10}	0.15	1.17 (1.12, 1.21)	3.8×10^{-13}	0.92
rs8046011 (chr16:11226805)	G/A	28.6	16p13	Intergenic	0.021	2	0	1.15 (1.10, 1.21)	0.95 (0.91, 0.99)	3.6×10^{-9}	0.17	0.97 (0.96, 0.99)	3.5×10^{-3}	0.76
rs28383314 (chr6:32619436)	T/C	37.0	6p21	Intergenic	-0.035	28	0	0.88 (0.84, 0.92)	1.07 (1.03, 1.12)	5.1×10^{-8}	0.64	1.01 (0.99, 1.03)	0.35	0.02
rs62408225 (chr6:90246690)	G/A	34.4	BACH2	Intron	-0.020	65	0	0.88 (0.84, 0.92)	0.96 (0.92, 1.00)	7.8×10^{-8}	1.1×10^{-3}	0.94 (0.92, 0.95)	4.7×10^{-14}	0.17
rs4807542 (chr19:1104079)	A/G	16.5	GPX4	Synonymous	-0.010	2	0	0.84 (0.79, 0.90)	0.92 (0.87, 0.97)	7.8×10^{-8}	0.40	0.97 (0.95, 1.00)	0.02	0.72
rs10905284 (chr10:8073399)	C/A	43.3	GATA3	Intron	0.014	3	0	1.12 (1.07, 1.17)	0.95 (0.91, 0.99)	4.3×10^{-7}	0.66	0.95 (0.94, 0.97)	1.2×10^{-9}	0.98
rs149206763 (chr17:4618946)	A/G	13.9	[ALOX15]	Intergenic	-0.014	2	0	0.85 (0.79, 0.90)	0.92 (0.87, 0.98)	5.4×10^{-7}	0.64	1.01 (0.99, 1.03)	0.38	0.44
rs7302200 (chr12:56055651)	A/G	33.7	12q13	Intergenic	0.019	44	0	1.12 (1.07, 1.17)	1.01 (0.97, 1.06)	9.4×10^{-7}	0.028	1.05 (1.03, 1.07)	2.4×10^{-8}	0.06
rs556670 (chr1:65026231)	A/C	12.5	1p31	Intergenic	0.018	37	0	1.15 (1.08, 1.23)	1.02 (0.96, 1.09)	1.5×10^{-5}	0.24	1.01 (0.98, 1.03)	0.60	0.61
rs10995240 (chr10:62628871)	C/G	35.0	ZNF365	Intron	-0.020	36	0	0.90 (0.86, 0.95)	0.95 (0.91, 0.99)	1.9×10^{-5}	0.19	0.98 (0.96, 0.99)	0.01	0.89
rs11242709 (chr6:209159)	T/C	20.2	6p25	Intergenic	-0.024	1	0	0.88 (0.83, 0.94)	0.96 (0.91, 1.01)	2.4×10^{-5}	0.13	1.01 (0.99, 1.03)	0.41	0.91
rs146597587 (chr9:6255967)	C/G	0.55	[IL33]	Splice acceptor	-0.184	15	0	0.42 (0.28, 0.63)	1.11 (0.85, 1.44)	2.5×10^{-5}	0.12	0.66 (0.58, 0.75)	1.4×10^{-10}	0.06
rs81290309 (chr21:35340290)	T/A	37.6	21q22	Intergenic	-0.014	6	0	0.91 (0.87, 0.95)	1.06 (1.01, 1.10)	4.8×10^{-5}	0.30	1.06 (1.04, 1.07)	1.8×10^{-10}	0.29
rs9532434 (chr13:39781776)	T/C	33.5	COG6	Intron	-0.008	30	0	0.91 (0.87, 0.95)	1.08 (1.03, 1.13)	5.5×10^{-5}	0.44	1.05 (1.03, 1.07)	5.1×10^{-8}	0.77
rs12935413 (chr16:11116590)	A/G	33.4	CLEC16A	Intron	-0.028	26	0	0.91 (0.87, 0.95)	0.95 (0.91, 0.99)	8.0×10^{-5}	0.039	0.94 (0.93, 0.96)	3.3×10^{-12}	0.29
rs175705 (chr14:75509305)	C/G	27.1	14q24	Intergenic	-0.023	9	0	0.91 (0.86, 0.95)	1.02 (0.97, 1.07)	1.2×10^{-4}	0.06	1.00 (0.99, 1.02)	0.64	0.61
rs1961220 (chr5:35844023)	A/G	36.5	5p13	Intergenic	-0.018	13	0	0.91 (0.87, 0.96)	0.96 (0.92, 1.01)	1.4×10^{-4}	0.75	0.96 (0.94, 0.97)	1.9×10^{-7}	0.49

Also presented are associations with CRS and asthma in the combined dataset of Iceland and the UK. Effects are presented for the minor allele. Average minor allele frequencies for Iceland and the UK are presented. Results for NP by country are presented in Supplementary Table T7. Genes linked to intergenic variants are listed in square brackets.

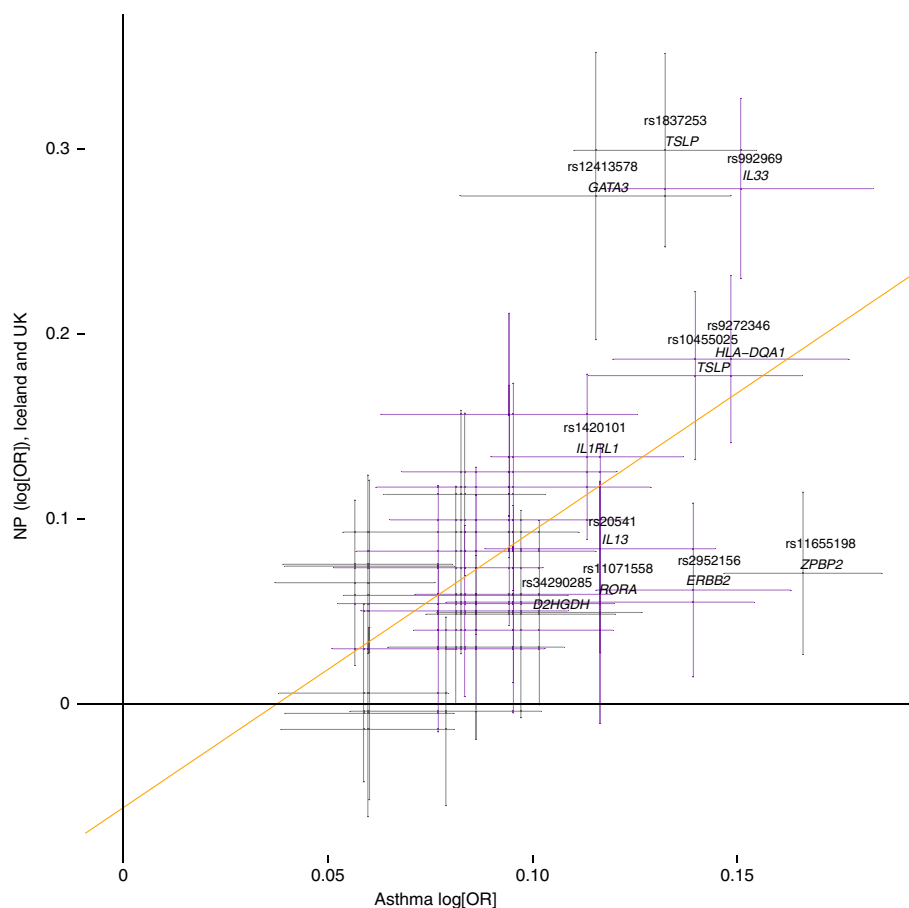


Fig. 4 | Scatter plot showing 34 previously reported asthma SNPs. Analyzed SNPs include 18 reported by Demenais et al.¹⁷ (shown in purple) and 16 reported by Pickrell et al.¹⁶ (shown in gray). The x axis shows the logarithm of the reported asthma ORs, and the y axis shows the logarithm of estimated ORs for NP in a combined data set of Iceland and UK (Iceland, $n=1,175$; UK, $n=3,191$). All effects are shown for the asthma risk-increasing allele. Variants associated with the highest asthma risk are labeled. Error bars represent 95% CIs. The orange yellow line ($y = -0.06 + 1.50x$) represents results from a simple linear regression using MAF ($1 - \text{MAF}$) as weights with $R = 0.61$ ($P = 1.4 \times 10^{-4}$; two-sided t test).

Most notably, in the HLA region, a signal of 72 highly correlated variants (r^2 with leading variant > 0.90) associates with increased risk of NP (Table 1 and Supplementary Table 10). The signal is represented by a common intergenic variant rs1391371[T] (NC_000006.11:g.32603798A>T; MAF in Iceland = 19.2%), located 1.37 kilobases (kb) upstream of *HLA-DQA1* (OR = 1.47, $P = 7.3 \times 10^{-49}$). The HLA type correlating most strongly with rs1391371 is *HLA-DQA1*03:01:01* (MAF in Iceland = 11.2%, $r^2 = 0.54$). The variant is highly correlated with four coding variants (outlined in Supplementary Table 16).

Reported eosinophil, asthma, and allergy variants. We assessed the NP associations of 209 sequence variants reported to associate with eosinophil count in a recent study on blood cell traits²⁰, along with the rare splice acceptor variant rs146597587[C] (NC_000009.12:g.6255967G>C) in *IL33* recently reported by us and others to associate with both eosinophil count and asthma^{26,27} (Supplementary Data 4). Of these markers, some associate with NP risk in proportion to the eosinophil count, whereas others do not (Fig. 3). In particular, the variants with the greatest effect on eosinophil count (variants in *FNIP1*, *IL5*, and *GF1B*) do not associate significantly with NP, and the variants with the largest NP effect (variants in *ALOX15* and *IL33*) do not all have large effects on eosinophil count. Overall, the effects of the reported eosinophil variants on NP risk do not correlate well with their effects on eosinophil count ($R = 0.25$, $P = 8.2 \times 10^{-4}$; Fig. 3), and our results are

inconsistent with a direct causal role of elevated circulating eosinophil count in NP.

Of the 210 eosinophil variants, 24 associate significantly with NP in the combined Iceland and UK dataset based on an adjusted significance threshold ($P < 0.05/210 = 2.4 \times 10^{-4}$). Of those, 7 are highly correlated ($r^2 > 0.7$) with one of the 10 signals detected in the current NP study. The remaining variants represent 17 novel association signals that would not have reached significance in the absence of a prior (Table 2, Supplementary Table 17, and Supplementary Data 5). All of the variants have effects in the same direction for both NP and eosinophil count.

Of the 27 signals associating with NP discovered at genome-wide significance or given the eosinophil count prior, we note that two loci (*ALOX15* and *IL33*) harbor several signals. A common intergenic variant rs149206763[A] (NC_000017.10:g.4522241G>A; MAF in Iceland = 14.0%) 10 kb upstream of *ALOX15* associates with decreased risk of NP ($P = 5.4 \times 10^{-7}$, OR = 0.85, 95% CI = 0.79, 0.90). The variant is not correlated with p.Thr560Met ($r^2 < 0.01$, D-prime (D') = 0.99). We also detect two additional independent variants at the *IL33* locus (Supplementary Tables 18 and 19): the aforementioned rare splice acceptor variant rs146597587[C] (MAF in Iceland = 0.65%) associating with decreased risk of NP ($P = 2.5 \times 10^{-5}$, OR = 0.42, 95% CI = 0.25, 0.63) and the low-frequency intergenic variant rs78757963[A] (NC_000009.11:g.6282511G>A; MAF in Iceland = 3.69%) associating with increased risk of NP ($P = 1.6 \times 10^{-10}$, OR = 1.42, 95% CI = 1.27, 1.58).

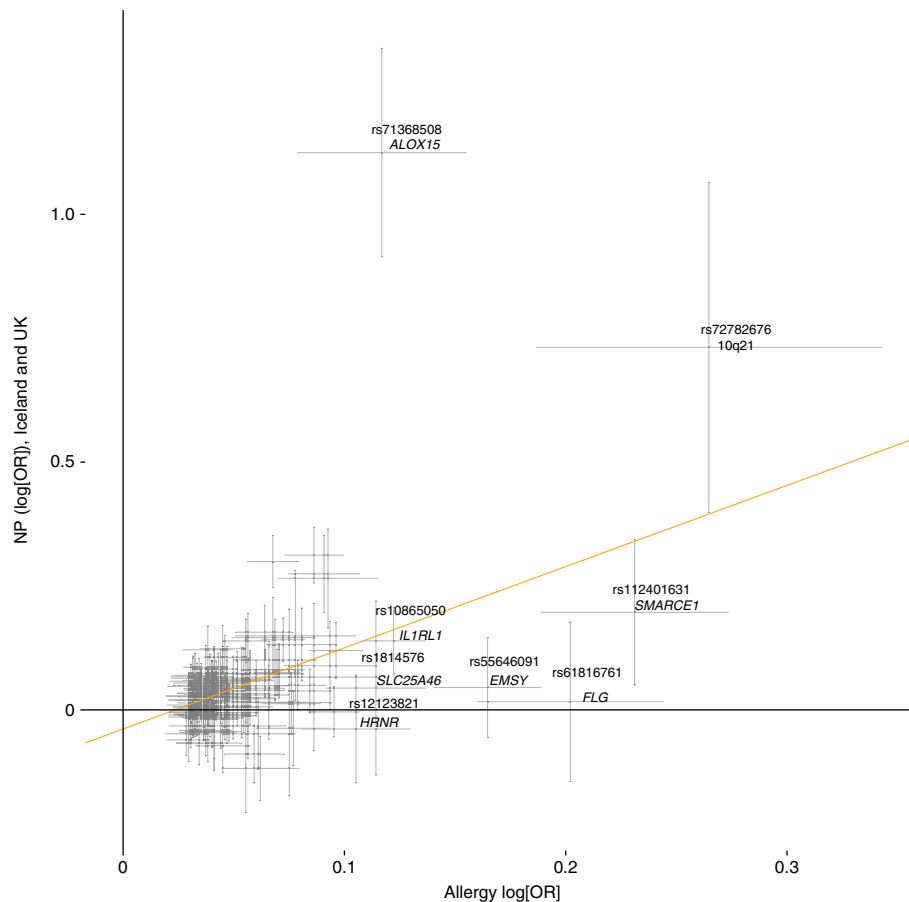


Fig. 5 | Reported allergy variants and risk of nasal polyps. Scatter plot showing 128 previously reported allergy SNPs²¹. The x axis shows the logarithm of their allergy ORs reported by Ferreira et al.²¹, and the y axis shows the logarithm of estimated ORs for NP in a combined data set of Iceland and UK (Iceland, $n=1,175$; UK, $n=3,191$). All effects are shown for the allergy risk-increasing allele. Variants associated with the largest allergy risk are labeled. Error bars represent 95% CIs. The solid yellow line ($y = -0.039 + 1.64x$) represents results from a simple linear regression using MAF ($1 - \text{MAF}$) as weights with $R=0.43$ ($P=3.5 \times 10^{-7}$; two-sided t test).

To assess the effects of reported asthma and allergic disease variants on the risk of NP, we tested 34 sequence variants reported to associate with asthma in two recent studies^{16,17}, and 136 variants reported to associate with allergic diseases (a combined phenotype of asthma, hay fever, and eczema) in a recent meta-analysis of Europeans²¹. We observed a positive correlation between the effects of asthma variants on asthma and NP risk ($R=0.61$, $P=1.4 \times 10^{-4}$; Fig. 4). For the allergic disease variants, the correlation is weaker ($R=0.43$, $P=3.5 \times 10^{-7}$; Fig. 5), with clear outliers. Notably, a stop-gained variant rs61816761[A] (NC_000001.10:g.152285861G>A) in *FLG* affects only allergic diseases, not NP, and mutations in this gene have primarily been reported to associate with atopic dermatitis^{28,29}. Conversely, the signal at *ALOX15* has a drastically larger effect (decrease in risk by 68%) on NP risk than on allergic diseases (decrease in risk by 11%) relative to other markers (Fig. 5). This is consistent with 15-LO having restricted nasal mucosa effects rather than broad allergic effects.

Discussion

Through a meta-analysis of data from Iceland and the UK, we found genome-wide significant associations between sequence variants at ten loci and risk of NP, two of which are also genome-wide significant for CRS. When assessing variants previously reported to associate with circulating eosinophil count, we found 17 additional signals for NP, although we did not demonstrate a consistent relationship between the effect on eosinophil count and risk of NP.

We therefore uncovered a total of 27 sequence variants associating with NP and/or CRS in Iceland and the UK. Some but not all ($n=13$) of these variants also associate with asthma, confirming the correlation between these phenotypes. Conversely, the signal at *ALOX15* has greater effect on NP than allergic disease, and the difference between the effects on the two conditions is greater than in the case of the other variants.

Of all the signals, most notable is a low-frequency missense variant p.Thr560Met in *ALOX15* carried by 1 in 20 Europeans that associates with a 68% reduction of risk of NP ($\text{OR}=0.32$) and 36% reduction of risk of CRS ($\text{OR}=0.64$). The protection conferred by the variant is unusually strong given its allelic frequency of 2.51% in Europeans (around 1 in 20 individuals are heterozygotes and 1 in 1,700 individuals are homozygotes). p.Thr560Met was among variants recently reported to associate with a combined phenotype of allergic diseases (asthma, eczema, and hay fever)²¹, but with a substantially weaker effect ($\text{OR}=0.89$, $P=2 \times 10^{-9}$) than the protection against NP.

ALOX15 encodes the enzyme 15-LO, shown to play a role in inflammatory processes³⁰ and to be upregulated in NP tissue³¹. Furthermore, mice deficient in the murine ortholog of 15-LO have attenuated airway inflammation and remodeling following repeated allergen challenge³². The Thr560 residue of human 15-LO has been shown to be critical for maintaining structural integrity of the active site through hydrogen bonding²³, and p.Thr560Met greatly reduces catalytic activity of 15-LO, resulting in a near-null variant

of the enzyme²³. p.Thr560Met is therefore a loss-of-function mutation. Interestingly, homozygous carriers of this variant do not have an increased risk of other diseases in a phenome scan, despite this drastic reduction in enzymatic activity.

In eosinophils and mast cells, the conversion of arachidonic acid by 15-LO leads to the formation of eoxins. Similarly, human airway epithelial cells can convert arachidonic acid to both eoxin C4 and 15(S)-hydroxy-eicosatetraenoic acid (15-HETE)³³. Eoxins and 15-HETE are potent pro-inflammatory metabolites¹⁸ suggested to play a role in severe asthma and aspirin-exacerbated respiratory disease³⁴ and have been shown to be produced in human NP¹⁸. Eoxins have also been shown to increase the permeability of human endothelial cell monolayer *in vitro*, suggesting that they may enhance vascular permeability, a hallmark of inflammation¹⁸.

We suggest that the protective effect of the rare p.Thr560Met 15-LO variant against NP and CRS is explained by inactivation of the enzyme's catalytic activity, leading to reduced production of pro-inflammatory mediators such as eoxins and 15-HETE in eosinophils and nasal epithelia.

Leukotriene inhibitors are a class of drugs that act by inhibiting the action of 5-LO, another member of the lipoxygenase family highly similar to 15-LO, and are approved for use in asthma and allergic rhinitis³⁵. To date, no pharmacological agents targeting the 15-LO pathway have been approved³⁶. Previously discovered protective loss-of-function variants have successfully guided drug development, as in the case of PCSK9 in prevention of coronary artery disease³⁷. Our findings point to the potential of 15-LO inhibition as a target to exploit in an attempt to contain NP and CRS. The absence of serious disease in homozygous carriers of the loss-of-function mutation in *ALOX15* indicates that it may be safe.

URLs. BWA 0.7.10 mem, <https://github.com/lh3/bwa>; Genome AnalysisTKLite 2.3.9, <https://github.com/broadgsa/gatk/>; Picard tools 1.117, <https://broadinstitute.github.io/picard/>; SAMtools 1.3, <http://samtools.github.io/>; Bedtools v2.25.0–76-g5e7c696z, <https://github.com/arq5x/bedtools2/>; Variant Effect Predictor, <https://github.com/Ensembl/ensembl-vep>; BOLT-LMM, <https://data.broadinstitute.org/alkesgroup/BOLT-LMM/downloads/>.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41588-018-0314-6>.

Received: 14 February 2018; Accepted: 16 November 2018;

Published online: 14 January 2019

References

- Newton, J. R. & Ah-See, K. W. A review of nasal polyposis. *Ther. Clin. Risk Manag.* **4**, 507–512 (2008).
- Drake-Lee, A. B. Nasal polyps. *Hosp. Med.* **65**, 264–267 (2004).
- Cingi, C., Demirbas, D. & Ural, A. Nasal polyposis: an overview of differential diagnosis and treatment. *Recent Pat. Inflamm. Allergy Drug Discov.* **5**, 241–252 (2011).
- Gliklich, R. E. & Metson, R. The health impact of chronic sinusitis in patients seeking otolaryngologic care. *Otolaryngol. Head Neck Surg.* **113**, 104–109 (1995).
- Deal, R. T. & Kountakis, S. E. Significance of nasal polyps in chronic rhinosinusitis: symptoms and surgical outcomes. *Laryngoscope* **114**, 1932–1935 (2004).
- Stevens, W. W., Schleimer, R. P., Chandra, R. K. & Peters, A. T. Biology of nasal polyposis. *J. Allergy Clin. Immunol.* **133**, 1503.e1–4 (2014).
- Staikūniene, J., Vaitkus, S., Japertiene, L. M. & Ryskiene, S. Association of chronic rhinosinusitis with nasal polyps and asthma: clinical and radiological features, allergy and inflammation markers. *Medicina* **44**, 257–265 (2008).
- Fokkens, W. J. et al. EPOS 2012: European position paper on rhinosinusitis and nasal polyps 2012. A summary for otorhinolaryngologists. *Rhinology* **50**, 1–12 (2012).
- Gevaert, P. et al. Mepolizumab, a humanized anti-IL-5 mAb, as a treatment option for severe nasal polyposis. *J. Allergy Clin. Immunol.* **128**, 989–995 (2011).
- Alobid, I. et al. Nasal polyposis and its impact on quality of life: comparison between the effects of medical and surgical treatments. *Allergy* **60**, 452–458 (2005).
- Gevaert, P. et al. Omalizumab is effective in allergic and nonallergic patients with nasal polyps and asthma. *J. Allergy Clin. Immunol.* **131**, 110–116 (2013).
- Gudbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
- Sveinbjornsson, G. et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).
- Klossek, J. et al. Prevalence of nasal polyposis in France: a cross-sectional, case-control study. *Allergy* **60**, 233–237 (2005).
- Langdon, C. & Mullol, J. Nasal polyps in patients with asthma: prevalence, impact, and management challenges. *J. Asthma Allergy* **9**, 45–53 (2016).
- Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).
- Demenaïs, F. et al. Multi-ancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat. Genet.* **50**, 42–53 (2018).
- Feltenmark, S. et al. Eoxins are proinflammatory arachidonic acid metabolites produced via the 15-lipoxygenase-1 pathway in human eosinophils and mast cells. *Proc. Natl Acad. Sci. USA* **105**, 680–685 (2008).
- Claesson, H.-E. On the biosynthesis and biological role of eoxins and 15-lipoxygenase-1 in airway inflammation and Hodgkin lymphoma. *Prostaglandins Other Lipid Mediat.* **89**, 120–125 (2009).
- Astle, W. J. et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* **167**, 1415–1429 (2016).
- Ferreira, M. A. et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat. Genet.* **49**, 1752–1757 (2017).
- GTEx Portal Version 7, <http://www.gtexportal.org> (NIH Common Fund, accessed 10 September 2017).
- Schurmann, K. et al. Molecular basis for the reduced catalytic activity of the naturally occurring T560M mutant of human 12/15-lipoxygenase that has been implicated in coronary artery disease. *J. Biol. Chem.* **286**, 23920–23927 (2011).
- Assimes, T. L. et al. A near null variant of 12/15-LOX encoded by a novel SNP in *ALOX15* and the risk of coronary artery disease. *Atherosclerosis* **198**, 136–144 (2008).
- Chang, J. et al. 12/15 lipoxygenase regulation of colorectal tumorigenesis is determined by the relative tumor levels of its metabolite 12-HETE and 13-HODE in animal models. *Oncotarget* **6**, 2879–2888 (2015).
- Smith, D. et al. A rare *IL33* loss-of-function mutation reduces blood eosinophil counts and protects from asthma. *PLoS Genet.* **13**, e1006659 (2017).
- Mousas, A. et al. Rare coding variants pinpoint genes that control human hematological traits. *PLoS Genet.* **13**, e1006925 (2017).
- Paternoster, L. et al. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.* **47**, 1449–1456 (2015).
- Baurecht, H. et al. Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. *Am. J. Hum. Genet.* **96**, 104–120 (2015).
- Kelavkar, U. P. & Badr, K. F. Effects of mutant p53 expression on human 15-lipoxygenase-promoter activity and murine 12/15-lipoxygenase gene expression: evidence that 15-lipoxygenase is a mutator gene. *Proc. Natl Acad. Sci. USA* **96**, 4378–4383 (1999).
- Rostkowska-Nadolska, B. et al. A microarray study of gene expression profiles in nasal polyps. *Auris Nasus Larynx* **38**, 58–64 (2011).
- Andersson, C. K. et al. Mice lacking 12/15-lipoxygenase have attenuated airway allergic inflammation and remodeling. *Am. J. Respir. Cell Mol. Biol.* **39**, 648–656 (2008).
- Brunnström, Å. et al. On the biosynthesis of 15-HETE and eoxin C4 by human airway epithelial cells. *Prostaglandins Other Lipid Mediat.* **121**, 83–90 (2015).
- James, A. et al. The influence of aspirin on release of eoxin C4, leukotriene C4 and 15-HETE, in eosinophilic granulocytes isolated from patients with asthma. *Int. Arch. Allergy Immunol.* **162**, 135–142 (2013).
- Steinke, J. W. & Culp, J. A. Leukotriene synthesis inhibitors versus antagonists: the pros and cons. *Curr. Allergy Asthma Rep.* **7**, 126–133 (2007).
- Sadeghian, H. & Jabbari, A. 15-Lipoxygenase inhibitors: a patent review. *Expert Opin. Ther. Pat.* **26**, 65–88 (2016).
- Jaworski, K., Jankowski, P. & Kosior, D. A. PCSK9 inhibitors—from discovery of a single mutation to a groundbreaking therapy of lipid disorders in one decade. *Arch. Med. Sci.* **13**, 914–929 (2017).

Acknowledgements

We thank the individuals who participated in this study and whose contribution made this work possible. We also thank our valued colleagues who contributed to the data collection and phenotypic characterization of clinical samples as well as to the genotyping and analysis of the whole-genome association data. This study was supported in part by the National Institute of Dental and Craniofacial Research of the National Institutes of Health, under award number R01DE022905. This research has been conducted using the UK Biobank Resource under application number 24711.

Author contributions

R.P.K., S.B., O.B.D., G.S., D.F.G., U.T., T.R., P.S., and K.S. designed the study and interpreted the results. R.P.K., S.B., O.B.D., A.O., V.T., I.O., G.I.E., O.S., P.T.O., D.G., T.G., B.R.L., D.L., T.A.O., F.Z., G.B., U.S.B., G.T., I.J., and P.S. carried out the subject ascertainment and recruitment. R.P.K., S.J., B.V.H., B.G., G.H.H., O.A.S., G.M., O.T.M., G.L.N., and P.S. performed the sequencing, genotyping, and expression analyses. R.P.K., S.B., O.B.D., J.K.S., L.S., J.G.A., B.O.J., G.A.A., A.O., G.R.O., A.M.D., D.F.G., and P.S. performed the statistical and bioinformatics analyses. R.P.K., D.F.G., U.T.,

G.L.N., I.J., P.S., and K.S. drafted the manuscript. All authors contributed to the final version of the paper.

Competing interests

Authors affiliated with deCODE genetics/Amgen Inc. declare competing interests as employees. The remaining authors declare no competing interests. A.M.D. is a shareholder of Amgen Inc.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41588-018-0314-6>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.S. or K.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Methods

Study subjects from Iceland. This study is based on whole-genome sequence data from 15,220 Icelanders participating in various disease projects at deCODE genetics, sequenced to an average genome-wide coverage of 34X. In addition, 151,677 Icelanders have been genotyped using Illumina SNP chips, and genotype probabilities for untyped relatives have been calculated based on Icelandic genealogy. The relations of these individuals to those who have been directly genotyped are outlined in Supplementary Table 20 and Supplementary Data 6. The process used to whole-genome sequence the Icelandic population, and the subsequent imputation from which the data for this analysis were generated, has been extensively described in recent publications^{12,38}.

All participating individuals who donated blood, or their guardians, provided written informed consent. All sample identifiers were encrypted in accordance with the regulations of the Icelandic Data Protection Authority. Approval for the study was provided by the National Bioethics Committee (ref. VSNb2015100030/03.03). Personal identities of the participants and biological samples were encrypted by a third-party system approved and monitored by the Icelandic Data Protection Authority³⁹. The National Bioethics Committee approved the study, including the protocol, methodology, and all documents presented to the participants, and all methods were performed in accordance with the relevant guidelines and regulations.

The total number of individuals used in the study was 372,176. Of these, 50.4% were male. The mean year of birth was 1969 (s.d. = 28.4 years). To identify NP and CRS cases, we searched for patients with International Classification of Diseases (ICD10) codes J33, representing NP, and J32, representing CRS, at Landspítali, the National University Hospital of Iceland (LUH), in Reykjavík, and at the Icelandic Medical Center in Mjódd (Laeknasetríð). The records spanned from 1986 to 2016. A total of 1,175 NP cases were included in the association analysis; 86 of these were whole-genome sequenced, 788 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 387 were imputed on the basis of information from genotyped close relatives. Controls comprised individuals recruited through different genetic research projects at deCODE genetics. Individuals in the NP cohort were excluded from the control group. Of the controls, 13,941 were whole-genome sequenced, 137,902 were genotyped by chip, and 171,403 were imputed on the basis of the genotypes of close relatives. The total number of NP controls was 309,305. A total of 3,188 CRS cases were included in the association analysis; 251 of these were whole-genome sequenced, 1,980 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 1,208 were imputed on the basis of information from genotyped close relatives. Controls were selected in the same way as for NP. Of the controls, 13,957 were whole-genome sequenced, 138,044 were genotyped by chip, and 215,895 were imputed on the basis of the genotypes of close relatives. The total number of CRS controls was 353,939.

Genotype imputation. Given the long-range phased haplotypes of the sequenced individuals (H) and the carrier probabilities of the haplotypes in the sequenced set (θ), the probability of a new haplotype n , not in the set of sequenced haplotypes H , is imputed as $\sum_{l \in H} \gamma_{nl} \theta_l$, where γ_{nl} is calculated as above for every $l \in H$. In order to test the quality of our imputations, we compared chip genotypes to those imputed on the basis of the sequence reads. Chip-based genotypes were available from 28,204 of the SNPs detected by sequencing that affect exons and splice regions, which proved to be fully concordant with genotypes from whole-genome sequencing. To assess imputation quality, we compared the imputed genotypes with those from the same set of 28,204 chip SNPs. The concordance between imputed and chip-typed genotypes was high. For example, 76.4% of SNPs with a derived allele frequency (DAF) of 1% were imputed almost perfectly ($r^2 > 0.99$), and 98.4% of SNPs were imputed accurately ($r^2 > 0.9$). Even for the SNPs that have only two copies of the minor allele observed in the set of sequenced individuals (DAF = 0.04%), 47% were imputed almost perfectly ($r^2 > 0.99$), and 64% of SNPs were imputed accurately ($r^2 > 0.9$).

The informativeness of genotype imputation (imputation information) was estimated by the ratio of the variance of imputed expected allele counts and the variance of the actual allele counts:

$$\frac{\text{Var}(E(\theta|\text{chip data}))}{\text{Var}(\theta)}$$

where θ is the allele count. Here, $\text{Var}(E(\theta|\text{chip data}))$ is estimated by the observed variance in the imputed expected counts and $\text{Var}(\theta)$ was estimated by $p(1-p)$, where p is the allele frequency.

In addition to imputing sequence variants from the whole-genome sequencing effort into chip-genotyped individuals, we also performed a second imputation step, denoted as familial imputation, where genotypes were imputed into relatives of those who were chip genotyped. The inputs into the second imputation step are the fully phased (in particular, every allele has been assigned its parent of origin) imputed and chip-type genotypes of the available chip-typed individual. The algorithm used to perform the second imputation step is described in the Supplementary Note.

To account for inflation in test statistics due to cryptic relatedness and stratification, we applied the method of LD score regression⁴⁰. With a set of

1.1 million variants, we regressed the χ^2 statistics from our genome-wide analysis scan against LD score and used the intercept as a correction factor. The LD scores were downloaded from a LD score database (http://atguftp.mgh.harvard.edu/rendan/1k_eur_r2_hm3snps_se_weights.RDS; accessed 23 June 2015). P values from the Icelandic and UK datasets were adjusted separately using LD score regression before being combined. The combined P values were not readjusted. The correction factors used are presented in Supplementary Table 21.

Study subjects from the UK. The UK Biobank project is a large prospective cohort study of ~500,000 individuals from across the UK, aged between 40 and 69 at recruitment⁴¹. Genotyping was performed using a custom-made Affymetrix UK UK BiLEVE Axiom⁴², in the first 50,000 participants and with Affymetrix UK Biobank Axiom array in the remaining participants⁴³; 95% of the signals are on both chips. Imputation was performed by the Wellcome Trust Centre for Human Genetics using a combination of 1000 Genomes phase 3⁴⁴, UK10K⁴⁵, and HaploType Reference Consortium⁴⁶ reference panels, for up to 92,693,895 SNPs⁴⁷, although only 37,123,566 well-imputed variants from the Haplotype Reference Consortium panel were used in the current study.

The populations defined from UK Biobank data consisted of 3,191 NP cases and 405,376 controls, as well as 2,420 CRS cases and 406,147 controls. All individuals were of European ancestry (self-reported white British with similar genetic ancestry based on principal component analysis and with consistent reported and genetically determined gender) recruited between 2006 and 2010. Related individuals were used in the study. The NP and CRS diagnoses were ascertained on the basis of ICD diagnoses obtained from primary or secondary diagnostic codes recorded across all participants' hospital episodes. Diagnoses were from ICD10 codes J32 and J33. Self-reported diagnoses were excluded.

Association testing and meta-analysis. Logistic regression was used to test for association between variants and disease, assuming a multiplicative model, treating disease status as the response and expected genotype counts from imputation as covariates. This was done for both populations using software developed at deCODE genetics¹². In Iceland, sex, age, and county of origin were included in the logistic regression, and for the UK Biobank, sex, age, and the first 40 principal components were included. Variants with imputation information below 0.8 were excluded from the analysis.

Variants in the UK Biobank imputation dataset were mapped to National Center for Biotechnology Information Genome Reference Consortium Build 38 positions and matched to the variants in the Icelandic dataset on the basis of allele variation. Results from the different study groups were combined using a Mantel-Haenszel model⁴⁸ in which the groups were allowed to have different population frequencies for alleles and genotypes but were assumed to have a common OR. Heterogeneity was tested by comparing the null hypothesis of the effect being the same in all populations to the alternative hypothesis of each population having a different effect using a likelihood ratio test. I^2 lies between 0% and 100% and describes the proportion of total variation in study estimates that is due to heterogeneity.

We applied genome-wide significance thresholds corrected for multiple testing using an adjusted Bonferroni procedure weighted for variant classes and predicted functional impact. With 32,463,443 sequence variants being tested in the Icelandic dataset, the weights given by Sveinbjörnsson et al.¹³ were rescaled to control the family-wise error rate. The adjusted significance thresholds are 2.6×10^{-7} for variants with high impact ($n = 8,464$), 5.1×10^{-8} for variants with moderate impact ($n = 149,983$), 4.6×10^{-9} for variants with low impact ($n = 2,283,889$), 2.3×10^{-9} for other variants in deoxyribonuclease I hypersensitivity sites ($n = 3,913,058$), and 7.9×10^{-10} for all other variants ($n = 26,108,038$). For further information, see the Supplementary Note.

RNA-sequencing analysis. Generation of poly(A) + complementary DNA sequencing libraries, RNA sequencing, and data processing were carried out as described previously^{49,50}. Two tissue types were available for this analysis: whole blood and adipose tissue. Whole blood from a total of 2,502 individuals and adipose tissue from 682 individuals were used. To estimate the association between sequence variant and gene expression, a generalized linear regression assuming additive genetic effect was used on rank-transformed gene expression estimates.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. We used publicly available software (see URLs) in conjunction with the above-described algorithms in the sequencing processing pipeline (whole-genome sequencing, association testing, RNA-sequencing mapping and analysis).

Data availability

The Icelandic population whole-genome sequencing data have been deposited in the European Variant Archive under accession code [PRJEB15197](https://doi.org/10.6026/PRJEB15197). The authors declare that the data supporting the findings of this study are available within the article, in its Supplementary Data files, and upon reasonable request.

39. Gulcher, J. R., Kristjánsson, K., Gudbjartsson, H. & Stefánsson, K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.* **8**, 739–742 (2000).
40. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
41. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
42. Wain, L. V. et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir. Med.* **3**, 769–781 (2015).
43. Welsh, S., Peakman, T., Sheard, S. & Almond, R. Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort. *BMC Genomics* **18**, 26 (2017).
44. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
45. UK10K Consortium. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–90 (2015).
46. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
47. Bycroft, C. et al. The UK Biobank Resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
48. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl Cancer Inst.* **22**, 719–748 (1959).
49. Oskarsson, G. R. et al. A truncating mutation in EPOR leads to hyporesponsiveness to erythropoietin with normal haemoglobin. *Commun. Biol.* **1**, 49 (2018).
50. Kehr, B. et al. Diversity in non-repetitive human sequences not found in the reference genome. *Nat. Genet.* **49**, 588–593 (2017).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. [For final submission](#): please carefully check your responses for accuracy; you will not be able to make changes later.

► Experimental design

1. Sample size

Describe how sample size was determined.

NP and CRS cases were identified through ICD10 codes from the 372,176 individuals from Iceland and 408,567 from the UK with available genotypes. This yielded 4,366 NP and 5,608 CRS cases. Details are provided in the Methods. No sample size calculation was performed as all available individuals were used for the study.

2. Data exclusions

Describe any data exclusions.

No data were excluded.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

We performed GWAS studies in two independent populations and combined the results. Results are presented for the populations independently and combined, and success of replication is reported on per-variant basis, both as individual p-values and as heterogeneity of effects between groups.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No randomization was used and is not relevant within the context of a genome-wide association study.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

This is not relevant, we are performing a genome-wide association study.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- | | |
|-------------------------------------|---|
| n/a | Confirmed |
| ☐ | ☒ The <u>exact sample size</u> (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☒ | ☐ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ A statement indicating how many times each experiment was replicated |
| ☐ | ☒ The statistical test(s) used and whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☐ | ☒ Test values indicating whether an effect is present
<i>Provide confidence intervals or give results of significance tests (e.g. <i>P</i> values) as exact values whenever appropriate and with effect sizes noted.</i> |
| ☐ | ☒ A clear description of statistics including <u>central tendency</u> (e.g. median, mean) and <u>variation</u> (e.g. standard deviation, interquartile range) |
| ☐ | ☒ Clearly defined error bars in <u>all</u> relevant figure captions (with explicit mention of central tendency and variation) |

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

Software used to analyze data was extensively described in a recent publication (PMC5607473) that is cited in the Methods of our manuscript. The following text appears in the Code Availability section of our Manuscript:

"We used publicly available software (URLs listed below) in conjunction with the above described algorithms in the sequencing processing pipeline (Whole-genome sequencing, Association testing, RNA-seq mapping and analysis):
BWA 0.7.10 mem, <https://github.com/lh3/bwa>
GenomeAnalysisTKLite 2.3.9, <https://github.com/broadgsa/gatk/>
Picard tools 1.117, <https://broadinstitute.github.io/picard/>
SAMtools 1.3, <http://samtools.github.io/>
Bedtools v2.25.0-76-g5e7c696z, <https://github.com/arq5x/bedtools2/>
Variant Effect Predictor <https://github.com/Ensembl/ensembl-vep>
BOLT-LMM <https://data.broadinstitute.org/alkesgroup/BOLT-LMM/downloads/> "

Variants were imputed based on the IMPUTE HMM model.
We used R extensively to analyze data and create plots.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This is described in detail in the Methods:

Iceland:

"This study is based on whole-genome sequence data from 15,220 Icelanders participating in various disease projects at deCODE genetics, sequenced to an average genome-wide coverage of 34X. In addition, 151,677 Icelanders have been genotyped using Illumina SNP chips, and genotype probabilities for untyped relatives has been calculated based on Icelandic genealogy. The relations of these individuals to those who have been directly genotyped are outlined in Supplementary Table 17 and Supplementary Data 6. The process used to whole-genome sequence the Icelandic population, and the subsequent imputation from which the data for this analysis were generated has been extensively described in recent publications"

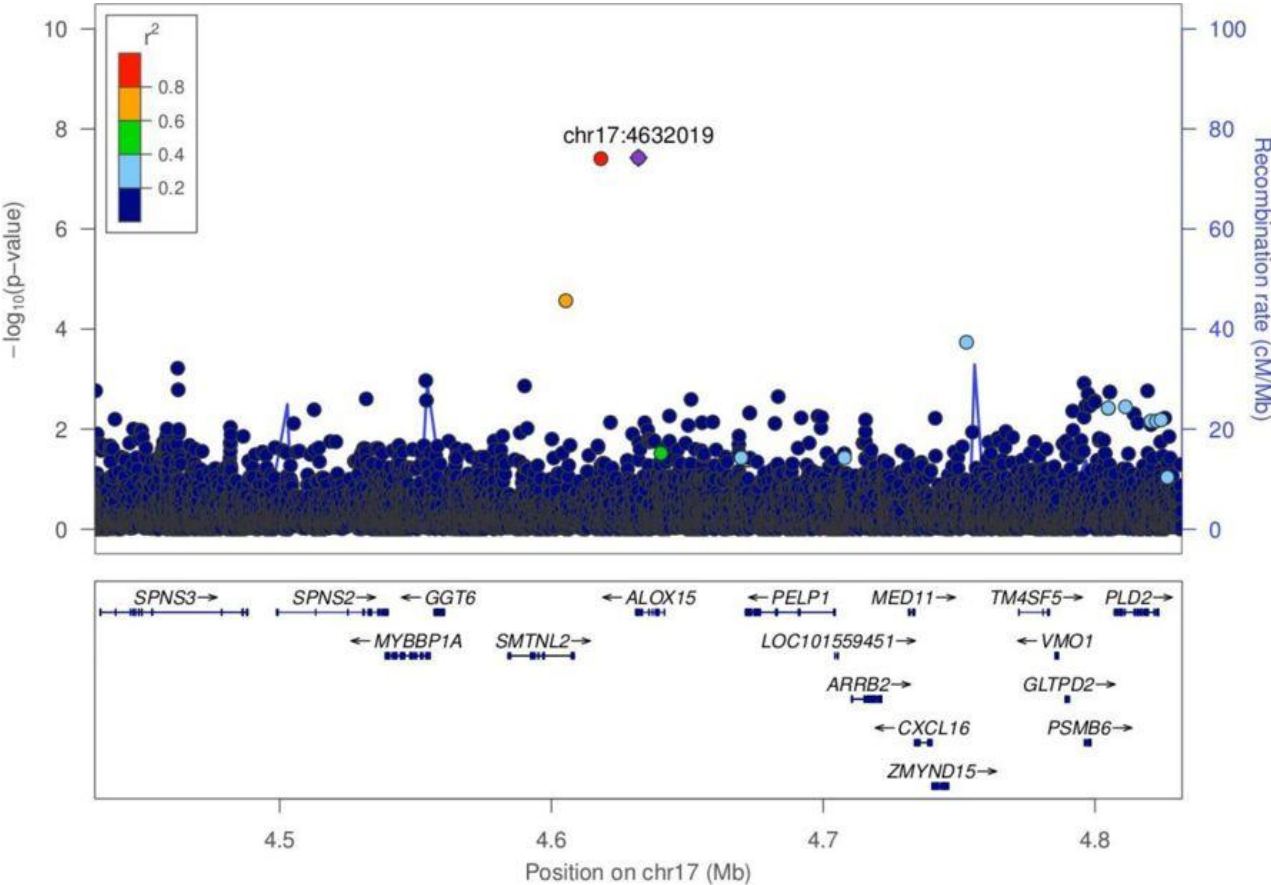
...

"The total number of individuals used in the study was 372,176. Of these, 50.4% were male. The mean year of birth was 1969 (SD = 28.4 years). To identify NP and CRS cases, we searched for patients with International Classification of Diseases (ICD10) codes J33, representing NP, and J32, representing CRS, at Landspítali—The National University Hospital of Iceland in Reykjavik (LUH), and the Icelandic Medical Center in Mjódd (Laeknasetríð). The records spanned from 1986-2016. A total of 1,175 NP cases were included in the association analysis; 86 of these were whole-genome sequenced, 788 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 387 were imputed on the basis of information from genotyped close relatives. Controls comprised individuals recruited through different genetic research projects at deCODE. Individuals in the NP cohort were excluded from the control group. Of the controls, 13,941 were whole-genome sequenced, 137,902 were genotyped by chip, and 171,403 were imputed on the basis of the genotypes of close relatives. The total number of NP controls was 309,305. A total of 3,188 CRS cases were included in the association analysis; 251 of these were whole-genome sequenced, 1,980 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 1208 were imputed on the basis of information from genotyped close relatives. Controls were selected in the same way as for NP. Of the controls, 13,957 were whole-genome sequenced, 138,044 were genotyped by chip, and 215,895 were imputed on the basis of the genotypes of close relatives. The total number of CRS controls was 353,939."

UK:

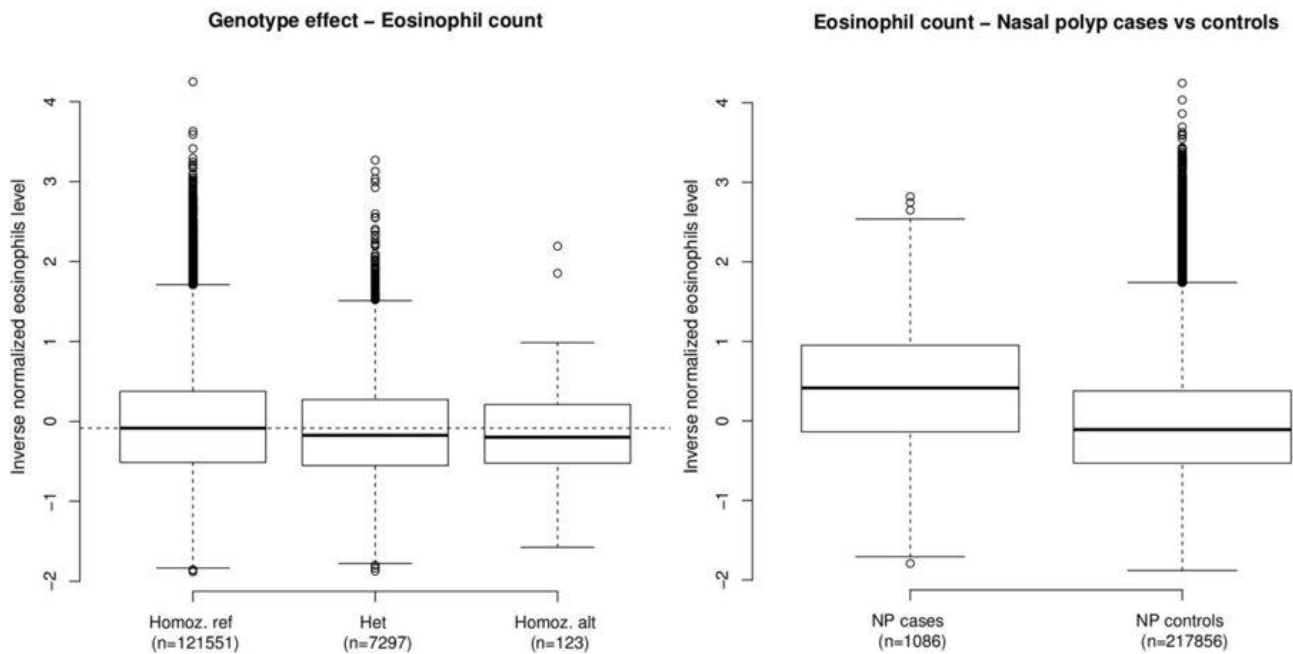
"The UK Biobank project is a large prospective cohort study of ~500,000 individuals from across the United Kingdom, aged between 40-69 at recruitment⁴¹. Genotyping was performed using a custom-made Affymetrix chip, UK BiLEVE Axiom⁴² in the first 50,000 participants, and with Affymetrix UK Biobank Axiom array in the remaining participants⁴³; 95% of the signals are on both chips. Imputation was performed by Wellcome Trust Centre for Human Genetics using a combination of 1000Genomes phase 3⁴⁴, UK10K⁴⁵, and HRC⁴⁶ reference panels, for up to 92,693,895 SNPs⁴⁷, although only 37,123,566 well-imputed variants from the HRC panel were used in the current study. The populations defined from UK Biobank data consisted of 3,191 NP cases and 405,376 controls, as well as 2,420 CRS cases and 406,147 controls. All individuals were of European ancestry (self-reported white British with similar genetic ancestry based on principal component analysis and with consistent reported and genetically determined gender) recruited between 2006 and 2010. Related individuals were used in the study. The NP and CRS diagnoses was ascertained based on ICD diagnoses obtained from primary or secondary diagnostic codes recorded across all participants' hospital episodes. Diagnoses were used from ICD-10 codes J32 and J33. Self-reported diagnoses were excluded.

Supplementary Figure 1: Locus plots for the Icelandic data showing the associations of variants (discovered by whole-genome sequencing of 15,220 Icelanders) at the ALOX15 locus with nasal polyps (n = 1,175).



The leading variant is shown as a purple diamond, and other variants are colored according to correlation (r^2) with the leading marker (legend at top-left). $-\log_{10}P$ values are shown along the left y-axis (two-sided logistic regression), and correspond to the variants depicted in the plot. The right y-axis shows calculated recombination rates at the chromosomal location, plotted as a solid blue line.

Supplementary Figure 2: Box plots depicting the effects of Thr560Met on eosinophil count in Iceland.



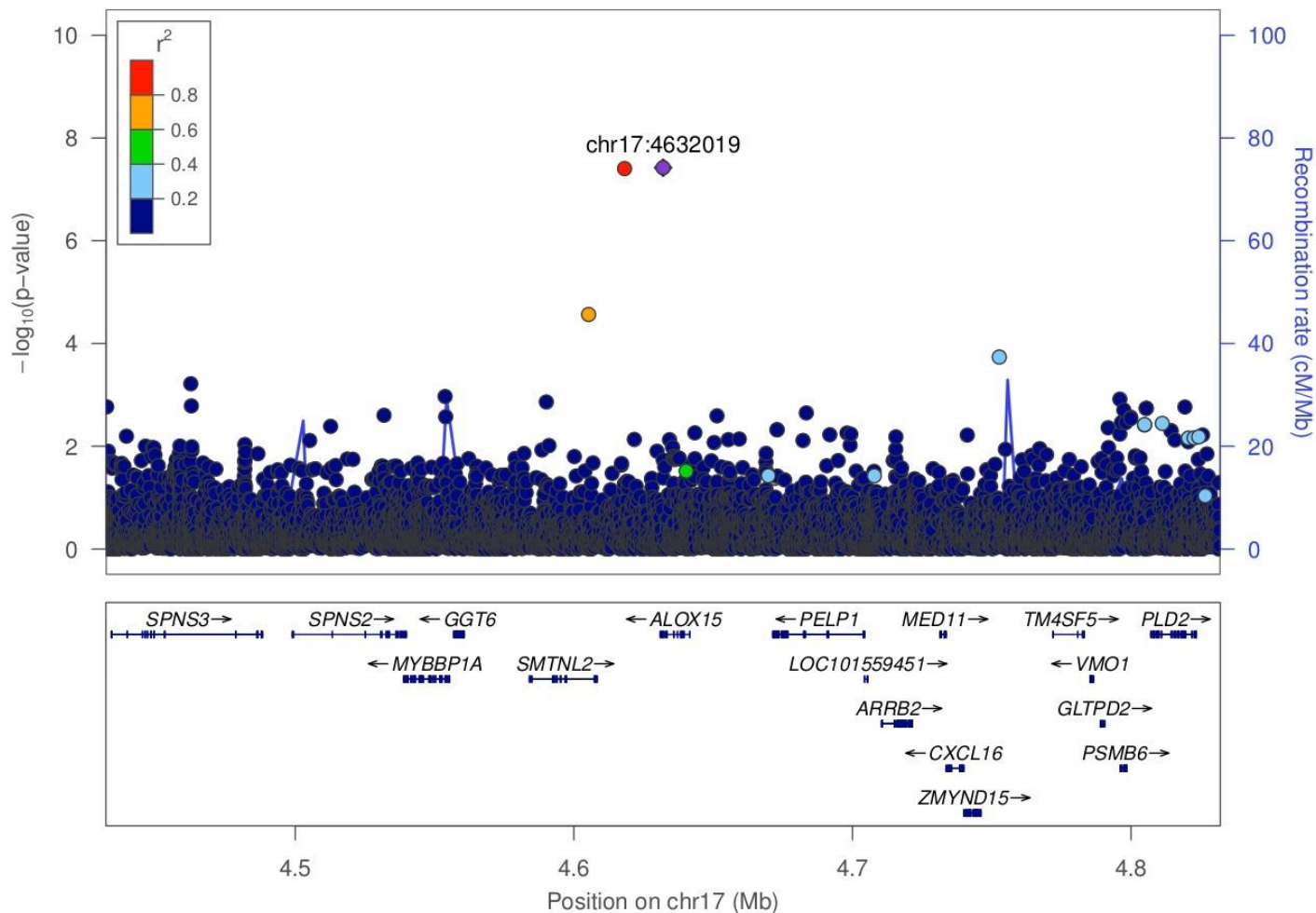
The plot on the left shows the genotypic effect for all chip-typed individuals in the Icelandic dataset with eosinophil count measured ($n = 128,971$). The right plot shows the effect when the eosinophil sample is stratified by nasal polyp status ($n = 1,086$ cases and 217,856 controls). The median is depicted as a horizontal black line, the interquartile range (IQR) is depicted by the upper and lower bounds of the rectangle, and data within 1.5 IQR of the boxes are depicted by the whiskers. Data falling outside this range are depicted as hollow circles.

In the format provided by the authors and unedited.

A loss-of-function variant in *ALOX15* protects against nasal polyps and chronic rhinosinusitis

Ragnar P. Kristjansson¹, Stefania Benonisdottir¹, Olafur B. Davidsson¹, Asmundur Oddsson¹, Vinicius Tragante^{1,2}, Jon K. Sigurdsson¹, Lilja Stefansdottir¹, Stefan Jonsson¹, Brynjar O. Jensson¹, Joseph G. Arthur¹, Gudny A. Arnadottir¹, Gerald Sulem¹, Bjarni V. Halldorsson^{1,3}, Bjarni Gunnarsson¹, Gisli H. Halldorsson¹, Olafur A. Stefansson¹, Gudjon R. Oskarsson¹, Aimee M. Deaton¹, Isleifur Olafsson⁴, Gudmundur I. Eyjolfsson⁵, Olof Sigurdardottir⁶, Pall T. Onundarson^{7,8}, David Gislason^{7,9}, Thorarinn Gislason^{7,10}, Bjorn R. Ludviksson^{7,11}, Dora Ludviksdottir^{7,12}, Thorunn A. Olafsdottir^{1,7}, Thorunn Rafnar¹, Gisli Masson¹, Florian Zink¹, Gyda Bjornsdottir¹, Olafur Th. Magnusson¹, Unnur S. Bjornsdottir^{9,13}, Gudmar Thorleifsson¹, Gudmundur L. Norddahl¹, Daniel F. Gudbjartsson^{1,14}, Unnur Thorsteinsdottir^{1,7}, Ingileif Jonsdottir^{1,7}, Patrick Sulem^{1*} and Kari Stefansson^{1,7*}

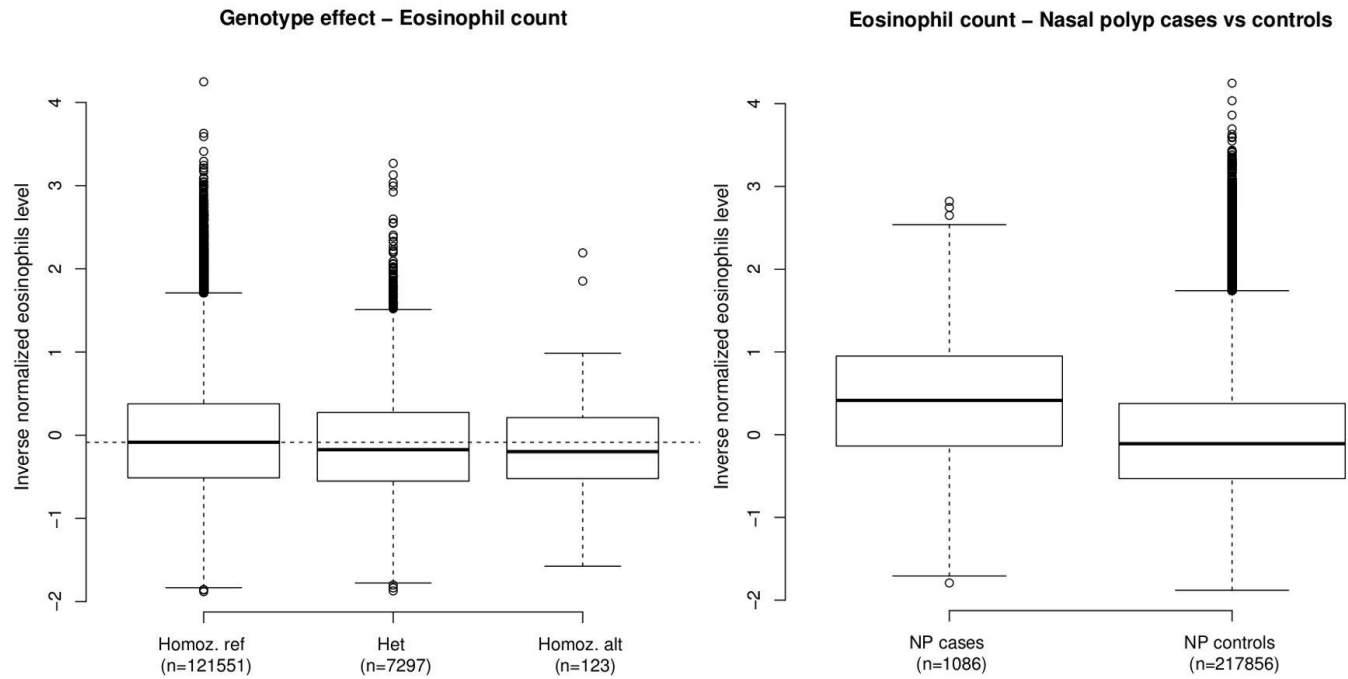
¹deCODE genetics/Amgen Inc., Reykjavik, Iceland. ²Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands. ³School of Science and Engineering, Reykjavik University, Reykjavik, Iceland. ⁴Department of Clinical Biochemistry, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ⁵The Laboratory in Mjodd, RAM, Reykjavik, Iceland. ⁶Department of Clinical Biochemistry, Akureyri Hospital, Akureyri, Iceland. ⁷Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. ⁸Department of Laboratory Hematology, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ⁹Department of Medicine, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹⁰Department of Sleep, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹¹Department of Immunology, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹²Department of Respiratory Medicine, Landspítali, the National University Hospital of Iceland, Reykjavik, Iceland. ¹³The Medical Center Mjodd, Reykjavik, Iceland. ¹⁴School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. *e-mail: patrick.sulem@decode.is; kari.stefansson@decode.is



Supplementary Figure 1

Locus plots for the Icelandic data showing the associations of variants (discovered by whole-genome sequencing of 15,220 Icelanders) at the *ALOX15* locus with nasal polyps ($n = 1,175$).

The leading variant is shown as a purple diamond, and other variants are colored according to correlation (r^2) with the leading marker (legend at top-left). $-\log_{10}P$ values are shown along the left y-axis (two-sided logistic regression), and correspond to the variants depicted in the plot. The right y-axis shows calculated recombination rates at the chromosomal location, plotted as a solid blue line.



Supplementary Figure 2

Box plots depicting the effects of Thr560Met on eosinophil count in Iceland.

The plot on the left shows the genotypic effect for all chip-typed individuals in the Icelandic dataset with eosinophil count measured ($n = 128,971$). The right plot shows the effect when the eosinophil sample is stratified by nasal polyp status ($n = 1,086$ cases and 217,856 controls). The median is depicted as a horizontal black line, the interquartile range (IQR) is depicted by the upper and lower bounds of the rectangle, and data within 1.5 IQR of the boxes are depicted by the whiskers. Data falling outside this range are depicted as hollow circles.

Supplementary Note

Common variant associations with NP

In the HLA region, the signal represented by rs1391371 also associates with eosinophil count and asthma (Table 1, Supplementary Tables 1-3), although the asthma association becomes non-significant when correcting for the top asthma-associating variant in the region (chr6:32523373[T]; $P_{\text{adjusted}} = 0.57$, $OR_{\text{adjusted}} = 1.02$, $r^2 = 0.60$), whereas the top asthma variant remains significant for association with asthma after adjusting for the top NP variant ($P_{\text{unadjusted}} = 1.1 \times 10^{-21}$, $OR_{\text{unadjusted}} = 1.26$; $P_{\text{adjusted}} = 1.5 \times 10^{-3}$, $OR_{\text{adjusted}} = 1.18$).

At 9p24 a common intergenic variant rs1888909[T] (NC_000009.11:g.6197392C>T; MAF in Iceland = 24.5%) associates with increased risk of NP ($P = 7.4 \times 10^{-32}$, $OR = 1.34$, 95% CI = 1.27, 1.40; Table 1). The variant is located 18kb upstream from *IL33*, a gene recently reported by us and others to harbor common and splice acceptor variants associating with asthma and eosinophil count^{1,2}. Another NP associating variant in this gene, the rare splice acceptor variant rs146597587 (Table 2) further associates with eosinophil count and asthma, and was recently confirmed in UK Biobank data³. We do not observe significant eQTL associations for this variant in the Icelandic dataset, and none are reported by Haploreg (Supplementary Data 1-3). At this locus, it has been concluded in the literature that *IL33* is the causal gene¹⁻³. In addition, the rare splice acceptor variant rs146597587[C] has been shown to associate with decreased expression and function of *IL33*¹, and associates in our data with decreased risk of NP (Table 2), providing further evidence for the causal role of *IL33*.

At 2q12, a common intronic variant rs6543124[A] (NC_000002.11:g.102987459T>A; MAF in Iceland = 38.9%) in *IL18R1* associates with decreased risk of NP ($P = 7.4 \times 10^{-10}$, $OR = 0.87$, 95% CI = 0.83, 0.91; Table 1). rs6543124 is strongly correlated with ($r^2 = 0.99$) five missense variants in *IL1RL1* (Supplementary Table

22). Both *IL18R1* and *IL1RL1* encode members of the interleukin 1 receptor family that play a role in processes related to inflammation and immunity⁴. We recently reported the role of *IL1RL1* and its ligand *IL33* in eosinophil count and asthma^{1,2}, although neither of the two reported signals at 2q21 associates strongly with rs6543124 (pairwise $r^2 < 0.34$). Data from Haploreg shows the variants within this LD class to associate significantly with the expression of *IL18R1* and *IL18RAP* in multiple tissue types, but not *IL1RL1* (Supplementary Data 1). rs6543124 is highly correlated with the top cis-eQTL variant reported in GTEx for *IL18R1* in sun-exposed skin (rs6543123; $r^2 = 0.97$), is only weakly correlated with the top overall cis-eQTL variant in blood (rs13405222; NC_000002.11:g.102977196C>T; $r^2 = 0.28$). In the literature, it has not been resolved which gene is causative at this locus, and *IL18R1/IL1RL1* are typically quoted together⁵.

At 5q31, a common missense variant rs1050152[T] in *SLC22A4* (NP_003050.2:p.Leu503Phe, MAF in Iceland = 39.6%) associates with decreased risk of NP ($P = 1.1 \times 10^{-8}$, OR = 0.88, 95% CI = 0.84, 0.92; Table 1). *SLC22A4* encodes the organic cation transporter solute carrier family 22, member 4. Leu503Phe and a strong correlate have been reported to associate with increased risk of Crohn's disease^{6,7}. This specific mutation has been shown to have drastic effects on the function of OCTN1, the protein encoded by *SLC22A4*, including a 50% increase in initial transport capacity at low concentrations of its natural substrate L-ergothioneine⁸. Expression data do not point to a single gene at this locus (Supplementary Data 1-3).

At 5q22, a common intergenic variant rs1837253[T] (NC_000005.9:g.110401872C>T; MAF in Iceland = 28.8%), with no strong correlates, associates with decreased risk of NP ($P = 3.7 \times 10^{-29}$, OR = 0.74, 95% CI = 0.70, 0.78; Table 1). It has no significant eQTL associations in the Icelandic dataset or in any datasets reported in Haploreg (Supplementary Data 1-3), but is located 3.9kb upstream of *TSLP*. *TSLP* encodes thymic stromal lymphopoietin, a protein that has been shown to cause activation of innate lymphoid cells⁹, and has been implicated in allergen-induced asthmatic responses^{10,11}. The variant rs1837253[T]

has been reported to associate with decreased risk of asthma ($P = 7.3 \times 10^{-10}$, OR = 0.84)¹²; an association we replicate ($P = 2.3 \times 10^{-23}$, OR = 0.91; Table 1). Most of the evidence at this locus is supported by *TSLP* biology.

At 19q13, a common intronic variant rs338598[A] in *CYP2S1* (NC_000019.10:g.41195623C>A; MAF in Iceland = 48.0%) associates with decreased risk of NP ($P = 5.6 \times 10^{-11}$, OR = 0.86, 95% CI = 0.82, 0.90; Table 1). The variant is in a very small LD class containing no coding variants (Supplementary Data 1). *CYP2S1* encodes a cytochrome P450 superfamily member, the depletion of which has been shown to affect cell proliferation and migration in bronchial epithelial cells through prostaglandin E2¹³. rs338598[A] has been reported to associate with decreased expression of *CYP2S1* and *AXL*¹⁴ in whole blood. At this locus, we consider the most likely causal gene to be *CYP2S1*.

Finally, we detect three common non-protein altering variants with modest effects on NP (Table 1). Firstly, at 3p13, a common intronic variant rs17718444[T] (NC_000003.11:g.71499401C>T; MAF in Iceland = 28.5%) in *FOXP1* associates with decreased risk of NP ($P = 1.3 \times 10^{-11}$, OR = 0.85, 95% CI = 0.81, 0.89). *FOXP1* encodes a transcription factor that has been shown to play a role in airway epithelial cells, and loss of Foxp1 and Foxp4 has been shown to increase expression of genes known to be upregulated in asthma^{15,16}. Biological evidence points to *FOXP1* as the likely causative gene at this locus, although two of the variant's strong correlates ($r^2 > 0.91$, $D' > 0.96$) have been reported to associate with expression of *ZNF57* in peripheral blood monocytes¹⁷ (Supplementary Data 1).

Secondly, at 11q12, we detect a common synonymous variant rs174535[C] in *MYRF* (NP_001120864.1:p.Ser1051; MAF in Iceland = 39.2%) associates with decreased risk of NP ($P = 3.5 \times 10^{-9}$, OR = 0.87, 95% CI = 0.83, 0.91). The variant and a number of its close correlates have been reported to associate with serum phospholipid levels^{18–20}. The variants are reported to affect expression of the fatty acid desaturase encoding proteins *FADS1* and *FADS2* in a number of studies reported in Haploreg, and

we similarly observe significant association with expression of these genes in the Icelandic dataset (Supplementary Data 2-3).

Finally, at 10p14, we detect a common intergenic variant rs1444782[A] (NC_000010.10:g.9058671G>A; MAF in Iceland = 38.8%) associates with decreased risk of NP ($P = 1.1 \times 10^{-14}$, OR = 0.84, 95% CI = 0.80, 0.88). The variant has no nearby genes, the closest being *GATA3* located 941,507bp away. The variant and a number of its close correlates have been reported to associate with expression of *GATA3* in transformed fibroblasts and whole blood (Supplementary Data 1). *GATA3* is known as the master regulator of Th2 cells^{21,22}, a cell type that has been shown to play a role in NP²³.

Familial Imputation

The algorithm used to perform the second imputation step consists of:

1. For each ungenotyped individual (the proband), find all chip genotyped individuals within two meioses of the individual. The six possible types of two meiotic distance relatives of the proband are (ignoring more complicated relationships due to pedigree loops): Parents, full and half siblings, grandparents, children and grandchildren. If all pedigree paths from the proband to a genotyped relative go through other genotyped relatives, then that relative is excluded. E.g. if a parent of the proband is genotyped, then the proband's grandparents through that parent are excluded. If the number of meiosis in the pedigree around the proband exceeds a threshold (we used 12), then relatives are removed from the pedigree until the number of meiosis falls below 12, in order to reduce computational complexity.
2. At every point in the genome, calculate the probability of each genotyped relative sharing with the proband based on the autosomal SNPs used for phasing. A multipoint algorithm based on the hidden Markov model Lander-Green multipoint linkage algorithm using fast Fourier transforms is used to calculate these sharing probabilities^{24,25}. First single point sharing probabilities are calculated by dividing the genome into 0.5 cM bins and using the haplotypes over these bins as alleles. Haplotypes that are the same, except at most at a single SNP, are treated as identical. When the haplotypes in the pedigree are incompatible over a bin, then a uniform probability distribution was used for that bin. The most common causes for such incompatibilities are recombinations within the pedigree, phasing errors and genotyping errors. Note that since the input genotypes are fully phased, the single point information is substantially more informative than for unphased genotyped, in particular one haplotype of the parent of a genotyped child is

always known. The single point distributions are then convolved using the multipoint algorithm to obtain multipoint sharing probabilities at the center of each bin. Genetic distances were obtained from the most recent version of the deCODE genetic map²⁶.

3. Based on the sharing probabilities at the center of each bin, all the SNPs from the whole genome sequencing are imputed into the proband. To impute the genotype of the paternal allele of a SNP located at x , flanked by bins with centers at x_{left} and x_{right} . Starting with the left bin, going through all possible sharing patterns v , let I_v be the set of haplotypes of genotyped individuals that share identically by descent within the pedigree with the proband's paternal haplotype given the sharing pattern v , and $P(v)$ be the probability of at the left bin – this is the output from step 2 above – and let e_i be the expected allele count of the SNP for haplotype i . Then $e_v = \frac{\sum_{i \in I_v} e_i}{\sum_{i \in I_v} 1}$ is the expected allele count of the paternal haplotype of the proband given v , and an overall estimate of the allele count given the sharing distribution at the left bin is obtained from $e_{left} = \sum P(v)e_v$. If I_v is empty then no relative shares with the proband's paternal haplotype given and thus there is no information about the allele count. We therefore store the probability that some genotyped relative shared the proband's paternal haplotype, $O_{left} = \sum_{v, I_v \neq \emptyset} P(v)$ and an expected allele count, conditional on the proband's paternal haplotype being shared by at least one genotyped relative: $c_{left} = \frac{\sum_{v, I_v \neq \emptyset} P(v)e_v}{\sum_{v, I_v \neq \emptyset} P(v)}$. In the same way calculate O_{right} and c_{right} . Linear interpolation is then used to get an estimates at the SNP from the two flanking bins:

$$O = O_{left} + \frac{x - x_{left}}{x_{right} - x_{left}} (O_{right} - O_{left})$$

$$c = c_{left} + \frac{x - x_{left}}{x_{right} - x_{left}} (c_{right} - c_{left})$$

If θ is an estimate of the population frequency of the SNP then $0c + (1 - 0)\theta$ is an estimate of the allele count for the proband's paternal haplotype. Similarly, an expected allele count can be obtained for the proband's maternal haplotype

Genome-wide significance thresholds

We corrected for multiple testing by applying genome-wide significance thresholds where a weighted Bonferroni procedure was used. Weighted Bonferroni correction is an extension of Bonferroni correction. For T tests, non-negative weights $w = (w_1, \dots, w_T)$ are specified and each hypothesis H_j is rejected if $\frac{P_j}{w_j} \leq \frac{0.05}{T}$. As long as $\frac{1}{T} \sum_j w_j = 1$, this procedure controls the family-wise error rate at 0.05.

As shown by Sveinbjornsson *et al.*, predicted functional impact of 5 different variant classes can be used as weights (variants with high impact, moderate impact, low-impact, other variants in DNase I hypersensitivity sites and all other variants)²⁷. The weights are the estimated enrichment of the sequence annotations among association signals.

Using enrichments as weights controls the FWER: $w_j = e_{c_j}$, where c_j is the category that the j th sequence variant belongs to. Let T_C be the number of sequence variants in category C and we estimate the frequency of category C , q_C , with $\frac{T_C}{T}$. Then

$$\frac{1}{T} \sum_j w_j = \frac{1}{T} \sum_j e_{c_j} = \frac{1}{T} \sum_C T_C e_C = \frac{1}{T} \sum_C T_C \frac{p_C}{q_C} = \frac{1}{T} \sum_C T p_C = \sum_C p_C = 1,$$

showing that using the enrichments as weights in the population they were derived from controls the FWER.

With 32,463,443 sequence variants being tested in the Icelandic dataset the adjusted significance thresholds are 2.6×10^{-7} for variants with high impact (N=8,464), 5.1×10^{-8} for variants with moderate impact (N=149,983), 4.6×10^{-9} for low-impact variants (N=2,283,889), 2.3×10^{-9} for other variants in DNase I hypersensitivity sites (N=3,913,058) and 7.9×10^{-10} for all other variants (N=26,108,038).

Supplementary References

1. Smith, D. *et al.* A rare IL33 loss-of-function mutation reduces blood eosinophil counts and protects from asthma. *PLoS Genet.* **13**, (2017).
2. Gudbjartsson, D. F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–7 (2009).
3. DeBoever, C. *et al.* Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. *Nat. Commun.* **9**, (2018).
4. Boraschi, D. & Tagliabue, A. The interleukin-1 receptor family. *Semin. Immunol.* **25**, 394–407 (2013).
5. Moffatt, M. F. *et al.* A Large-Scale, Consortium-Based Genomewide Association Study of Asthma. *N. Engl. J. Med.* **363**, 1211–1221 (2010).
6. Peltekova, V. *et al.* Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat. Genet.* **36**, 471–475 (2004).
7. Franke, A., McGovern, D. & Barrett, J. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. ...* **42**, 1118–25 (2010).
8. Taubert, D., Grimberg, G., Jung, N., Rubbert, A. & Schömig, E. Functional role of the 503F variant of the organic cation transporter OCTN1 in Crohn's disease. *Gut* **54**, 1505–6 (2005).
9. Kim, B. S. *et al.* TSLP Elicits IL-33-Independent Innate Lymphoid Cell Responses to Promote Skin Inflammation. *Sci. Transl. Med.* **5**, 170ra16-170ra16 (2013).
10. Gauvreau, G. M. *et al.* Effects of an Anti-TSLP Antibody on Allergen-Induced Asthmatic Responses. *N. Engl. J. Med.* **370**, 2102–2110 (2014).

11. West, E. E., Kashyap, M. & Leonard, W. J. TSLP: A Key Regulator of Asthma Pathogenesis. *Drug Discov. Today Dis. Mech.* **9**, e83–e88 (2012).
12. Torgerson, D. G. *et al.* Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat. Genet.* **43**, 887–892 (2011).
13. Madanayake, T. W., Fidler, T. P., Fresquez, T. M., Bajaj, N. & Rowland, A. M. Cytochrome P450 2S1 depletion enhances cell proliferation and migration in bronchial epithelial cells, in part, through modulation of prostaglandin E(2) synthesis. *Drug Metab. Dispos.* **40**, 2119–25 (2012).
14. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
15. Li, S. *et al.* Foxp1/4 control epithelial cell fate during lung development and regeneration through regulation of anterior gradient 2. *Development* **139**, 2500–2509 (2012).
16. Li, S. *et al.* Epithelium-generated neuropeptide Y induces smooth muscle contraction to promote airway hyperresponsivenessNo Title. *J. Clin. Invest.* **126**, 1978–1982 (2016).
17. Zeller, T. *et al.* Genetics and Beyond – The Transcriptome of Human Monocytes and Disease Susceptibility. *PLoS One* **5**, e10693 (2010).
18. Lemaitre, R. *et al.* Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium.No Title. *PLoS Genet.* **7**, e1002193 (2011).
19. Mozaffarian, D. *et al.* Genetic loci associated with circulating phospholipid trans fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. *Am. J. Clin. Nutr.* **101**, 398–406 (2015).

20. Shin, S. *et al.* An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**, 543–550 (2014).
21. Zhu, J. Transcriptional regulation of Th2 cell differentiation. *Immunol. Cell Biol.* **88**, 244–249 (2010).
22. T, H. *et al.* Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat. Genet.* **43**, 893–896 (2011).
23. EP, L. *et al.* IL-25/IL-33-responsive TH2 cells characterize nasal polyps with a default TH17 signature in nasal mucosa. *J. Allergy Clin. Immunol.* **137**, 1514–1524 (2016).
24. Lander, E. & Green, P. Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 2363–7 (1987).
25. Kruglyak, L. & Lander, E. Faster multipoint linkage analysis using Fourier transforms. *J. Comput. Biol.* **5**, 1–7 (1998).
26. Kong, A. *et al.* Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**, 1099–1103 (2010).
27. Sveinbjornsson, G. *et al.* Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).

Supplementary Tables

Supplementary Table 1: Association of the 10 variants identified in the combined GWAS with NP (N=4,366 cases; 714,681 controls), CRS (N=5,608 cases; 760,086 controls), NPsCRS (2,979 cases; 712,354 controls), CRSsNP (4,335 cases; 761,359 controls), and NPwCRS (1,396 cases; 698,839 controls) in the combined dataset from Iceland and the UK. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP (Iceland + UK)		CRS (Iceland + UK)		NP without confirmed CRS (Iceland + UK)			CRS without confirmed NP (Iceland + UK)			NP with confirmed CRS (Iceland + UK)	
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	P _{het} *	OR (95% CI)	P	P _{het} *	OR (95% CI)	P
rs1391371	chr6:32636021	T/A	HLA-DQA1	Upstream	1.47 (1.39, 1.54)	7.3x10 ⁻⁴⁹	1.22 (1.16, 1.28)	3.8x10 ⁻¹⁵	1.48 (1.39, 1.57)	7.0x10 ⁻³⁷	0.87	1.15 (1.08, 1.21)	2.2x10 ⁻⁶	0.13	1.43 (1.30, 1.57)	4.2x10 ⁻¹⁴
rs1888909	chr9:6197392	T/C	[IL33]	Intergenic	1.34 (1.27, 1.40)	7.4x10 ⁻³²	1.12 (1.07, 1.18)	5.9x10 ⁻⁷	1.34 (1.26, 1.42)	2.6x10 ⁻²³	1.00	1.06 (1.01, 1.12)	0.022	0.11	1.33 (1.22, 1.46)	1.1x10 ⁻¹⁰
rs1837253	chr5:111066174	T/C	[TSLP]	Intergenic	0.74 (0.70, 0.78)	3.7x10 ⁻²⁹	0.93 (0.89, 0.98)	3.9x10 ⁻³	0.73 (0.68, 0.77)	9.5x10 ⁻²⁴	0.74	0.99 (0.94, 1.05)	0.84	0.26	0.78 (0.72, 0.86)	2.8x10 ⁻⁷
rs34210653	chr17:4632019	A/G	ALOX15	Missense (Thr560Met)	0.32 (0.26, 0.39)	8.0x10 ⁻²⁷	0.64 (0.55, 0.75)	1.1x10 ⁻⁸	0.31 (0.24, 0.39)	2.8x10 ⁻²⁰	0.85	0.73 (0.62, 0.86)	1.9x10 ⁻⁴	0.25	0.35 (0.24, 0.50)	1.4x10 ⁻⁸
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.84 (0.80, 0.88)	1.1x10 ⁻¹⁴	0.90 (0.86, 0.94)	3.5x10 ⁻⁷	0.84 (0.80, 0.89)	2.3x10 ⁻¹⁰	1.00	0.92 (0.87, 0.96)	3.0x10 ⁻⁴	0.48	0.83 (0.77, 0.90)	1.1x10 ⁻⁵
rs17718444	chr3:71450250	T/C	FOXP1	Intron	0.85 (0.81, 0.89)	1.3x10 ⁻¹¹	0.96 (0.92, 1.00)	0.06	0.84 (0.79, 0.89)	8.8x10 ⁻¹⁰	0.75	0.98 (0.93, 1.03)	0.48	0.57	0.88 (0.81, 0.96)	4.0x10 ⁻³
rs338598	chr19:41195623	A/C	CYP251	Intron	0.86 (0.82, 0.90)	5.6x10 ⁻¹¹	0.92 (0.88, 0.96)	3.5x10 ⁻⁵	0.87 (0.83, 0.92)	3.7x10 ⁻⁷	0.75	0.94 (0.89, 0.98)	5.3x10 ⁻³	0.47	0.84 (0.78, 0.91)	3.2x10 ⁻⁵
rs6543124	chr2:102370999	A/T	IL18R1	Intron	0.87 (0.83, 0.91)	7.4x10 ⁻¹⁰	0.94 (0.90, 0.98)	4.2x10 ⁻³	0.86 (0.82, 0.91)	7.6x10 ⁻⁸	0.75	0.97 (0.92, 1.01)	0.15	0.30	0.88 (0.81, 0.95)	1.6x10 ⁻³
rs1050152	chr5:132340627	T/C	SLC22A4	Missense (Leu503Phe)	0.88 (0.84, 0.92)	1.1x10 ⁻⁸	0.92 (0.89, 0.96)	2.3x10 ⁻⁴	0.88 (0.84, 0.93)	4.4x10 ⁻⁶	1.00	0.95 (0.91, 1.00)	0.048	0.35	0.86 (0.80, 0.94)	3.5x10 ⁻⁴
rs174535	chr11:61783884	C/T	MYRF	Synonymous	0.87 (0.83, 0.91)	3.5x10 ⁻⁹	0.97 (0.93, 1.01)	0.15	0.85 (0.81, 0.90)	2.2x10 ⁻⁸	0.53	0.99 (0.94, 1.04)	0.72	0.56	0.90 (0.83, 0.98)	0.017

* P-value for the X² test of heterogeneity between the phenotype and its corresponding GWAS meta-analysis phenotype (without removal of other cases).

Supplementary Table 2: Association of the 10 variants identified in the combined GWAS with NP and CRS when cases for one are excluded as controls for the other. Associations are presented for the combined dataset from Iceland and the UK, for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP (N=4,366 cases; 714,681 controls)		NP without CRS controls (4,366 cases; 711,105 controls)		CRS (N=5,608 cases; 760,086 controls)		CRS without NP controls (5,608 cases, 757,118 controls)	
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.47 (1.39, 1.54)	7.3x10 ⁻⁴⁹	1.47 (1.39, 1.54)	6.2x10 ⁻⁴⁹	1.22 (1.16, 1.28)	3.8x10 ⁻¹⁵	1.22 (1.16, 1.28)	1.8x10 ⁻¹⁵
rs1888909	chr9:6197392	T/C	[<i>IL33</i>]	Intergenic	1.34 (1.27, 1.40)	7.4x10 ⁻³²	1.34 (1.28, 1.40)	3.0x10 ⁻³²	1.12 (1.07, 1.18)	5.9x10 ⁻⁷	1.13 (1.07, 1.18)	4.6x10 ⁻⁷
rs1837253	chr5:111066174	T/C	[<i>TSLP</i>]	Intergenic	0.74 (0.70, 0.78)	3.7x10 ⁻²⁹	0.74 (0.70, 0.78)	5.9x10 ⁻²⁹	0.93 (0.89, 0.98)	3.9x10 ⁻³	0.93 (0.89, 0.98)	3.2x10 ⁻³
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.32 (0.26, 0.39)	8.0x10 ⁻²⁷	0.32 (0.26, 0.39)	3.6x10 ⁻²⁷	0.64 (0.55, 0.75)	1.1x10 ⁻⁸	0.64 (0.55, 0.75)	8.8x10 ⁻⁹
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.84 (0.80, 0.88)	1.1x10 ⁻¹⁴	0.84 (0.80, 0.88)	1.1x10 ⁻¹⁴	0.90 (0.86, 0.94)	3.5x10 ⁻⁷	0.90 (0.86, 0.93)	2.8x10 ⁻⁷
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.85 (0.81, 0.89)	1.3x10 ⁻¹¹	0.85 (0.81, 0.89)	2.3x10 ⁻¹¹	0.96 (0.92, 1.00)	0.06	0.96 (0.91, 1.00)	5.4x10 ⁻²
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	0.86 (0.82, 0.90)	5.6x10 ⁻¹¹	0.86 (0.82, 0.90)	5.8x10 ⁻¹¹	0.92 (0.88, 0.96)	3.5x10 ⁻⁵	0.92 (0.88, 0.95)	2.9x10 ⁻⁵
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	0.87 (0.83, 0.91)	7.4x10 ⁻¹⁰	0.86 (0.83, 0.91)	5.0x10 ⁻¹⁰	0.94 (0.90, 0.98)	4.2x10 ⁻³	0.94 (0.90, 0.98)	3.5x10 ⁻³
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.88 (0.84, 0.92)	1.1x10 ⁻⁸	0.88 (0.84, 0.92)	6.8x10 ⁻⁹	0.92 (0.89, 0.96)	2.3x10 ⁻⁴	0.92 (0.89, 0.96)	2.2x10 ⁻⁴
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.87 (0.83, 0.91)	3.5x10 ⁻⁹	0.87 (0.83, 0.91)	3.8x10 ⁻⁹	0.97 (0.93, 1.01)	0.15	0.97 (0.93, 1.01)	0.14

Supplementary Table 3: Association of the 10 variants identified in the combined GWAS with NP, NP with confirmed asthma, and NP without confirmed asthma in the combined dataset from Iceland and the UK. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets. NP cases were excluded from the control groups.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP (N=4,366 cases; 714,681 controls)		NP with Asthma (N=1,528 cases; 712,245 controls)		NP without Asthma (2,849 cases; 711,123 controls)		Heterogeneity P_{het}
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.47 (1.39, 1.54)	7.3×10^{-49}	1.52 (1.39, 1.65)	6.0×10^{-22}	1.44 (1.35, 1.53)	1.1×10^{-29}	0.32
rs1888909	chr9:6197392	T/C	[<i>IL33</i>]	Intergenic	1.34 (1.27, 1.40)	7.4×10^{-32}	1.24 (1.14, 1.34)	2.5×10^{-7}	1.39 (1.31, 1.48)	3.4×10^{-28}	2.6×10^{-2}
rs1837253	chr5:111066174	T/C	[<i>TSLP</i>]	Intergenic	0.74 (0.70, 0.78)	3.7×10^{-29}	0.74 (0.68, 0.81)	1.8×10^{-11}	0.74 (0.70, 0.79)	1.6×10^{-19}	1.00
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.32 (0.26, 0.39)	8.0×10^{-27}	0.37 (0.26, 0.51)	3.1×10^{-9}	0.29 (0.22, 0.38)	5.5×10^{-20}	0.15
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.84 (0.80, 0.88)	1.1×10^{-14}	0.75 (0.69, 0.81)	5.6×10^{-14}	0.89 (0.84, 0.94)	3.5×10^{-5}	3.2×10^{-4}
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.85 (0.81, 0.89)	1.3×10^{-11}	0.79 (0.73, 0.86)	1.3×10^{-8}	0.88 (0.83, 0.94)	2.8×10^{-5}	3.6×10^{-2}
rs338598	chr19:41195623	A/C	<i>CYP251</i>	Intron	0.86 (0.82, 0.90)	5.6×10^{-11}	0.99 (0.92, 1.06)	0.69	0.80 (0.76, 0.85)	2.6×10^{-15}	1.8×10^{-8}
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	0.87 (0.83, 0.91)	7.4×10^{-10}	0.91 (0.84, 0.98)	1.3×10^{-2}	0.84 (0.80, 0.89)	2.4×10^{-9}	9.5×10^{-2}
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.88 (0.84, 0.92)	1.1×10^{-8}	0.84 (0.78, 0.91)	9.5×10^{-6}	0.89 (0.85, 0.94)	6.6×10^{-5}	0.24
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.87 (0.83, 0.91)	3.5×10^{-9}	0.88 (0.82, 0.95)	1.6×10^{-3}	0.86 (0.82, 0.91)	3.9×10^{-7}	0.65

Supplementary Table 4: Association of the 10 variants identified in the combined GWAS with NP with confirmed asthma, and NP without confirmed asthma in the individual Icelandic and UK datasets, as well as the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets. NP cases were excluded from the control groups.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP with Asthma (N _{Ice} =450 cases, 306,781 controls ; N _{UK} =1,078 cases, 405,464 controls)						NP without Asthma (N _{Ice} =736 cases, 305,659 controls ; N _{UK} =2,113 cases, 405,464 controls)					
					OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}	OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}
					(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)	
rs1391371	chr6:32636021	T/A	HLA- DQA1	Upstream	1.49 (1.26, 1.76)	3.5x10 ⁻⁶	1.53 (1.38, 1.68)	3.1x10 ⁻¹⁷	1.52 (1.39, 1.65)	6.0x10 ⁻²²	1.53 (1.34, 1.75)	4.2x10 ⁻¹⁰	1.41 (1.31, 1.51)	2.2x10 ⁻²¹	1.44 (1.35, 1.53)	1.1x10 ⁻²⁹
rs1888909	chr9:6197392	T/C	[IL33]	Intergenic	1.12 (0.96, 1.32)	0.16	1.28 (1.17, 1.41)	2.7x10 ⁻⁷	1.24 (1.14, 1.34)	2.5x10 ⁻⁷	1.45 (1.27, 1.64)	1.5x10 ⁻⁸	1.38 (1.29, 1.48)	2.9x10 ⁻²¹	1.39 (1.31, 1.48)	3.4x10 ⁻²⁸
rs1837253	chr5:111066174	T/C	[TSLP]	Intergenic	0.97 (0.82, 1.13)	0.68	0.66 (0.60, 0.73)	8.2x10 ⁻¹⁵	0.74 (0.68, 0.81)	1.8x10 ⁻¹¹	0.74 (0.65, 0.85)	1.8x10 ⁻⁵	0.74 (0.69, 0.80)	1.8x10 ⁻¹⁵	0.74 (0.70, 0.79)	1.6x10 ⁻¹⁹
rs34210653	chr17:4632019	A/G	ALOX15	Missense (Thr560Met)	0.34 (0.18, 0.62)	4.3x10 ⁻⁴	0.38 (0.26, 0.57)	1.8x10 ⁻⁶	0.37 (0.26, 0.51)	3.1x10 ⁻⁹	0.31 (0.18, 0.52)	9.8x10 ⁻⁶	0.28 (0.21, 0.38)	1.1x10 ⁻¹⁵	0.29 (0.22, 0.38)	5.5x10 ⁻²⁰
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.74 (0.64, 0.87)	1.3x10 ⁻⁴	0.75 (0.69, 0.82)	9.8x10 ⁻¹¹	0.75 (0.69, 0.81)	5.6x10 ⁻¹⁴	0.87 (0.77, 0.98)	2.4x10 ⁻²	0.90 (0.84, 0.95)	4.7x10 ⁻⁴	0.89 (0.84, 0.94)	3.5x10 ⁻⁵
rs17718444	chr3:71450250	T/C	FOXP1	Intron	0.86 (0.73, 1.02)	7.6x10 ⁻²	0.77 (0.70, 0.84)	2.9x10 ⁻⁸	0.79 (0.73, 0.86)	1.3x10 ⁻⁸	0.94 (0.83, 1.07)	0.33	0.87 (0.81, 0.93)	2.6x10 ⁻⁵	0.88 (0.83, 0.94)	2.8x10 ⁻⁵
rs338598	chr19:41195623	A/C	CYP2S1	Intron	1.02 (0.88, 1.18)	0.80	0.97 (0.89, 1.06)	0.54	0.99 (0.92, 1.06)	0.69	0.79 (0.70, 0.88)	5.9x10 ⁻⁵	0.81 (0.76, 0.86)	8.8x10 ⁻¹²	0.80 (0.76, 0.85)	2.6x10 ⁻¹⁵
rs6543124	chr2:102370999	A/T	IL18R1	Intron	1.01 (0.87, 1.17)	0.86	0.87 (0.80, 0.95)	2.7x10 ⁻³	0.91 (0.84, 0.98)	1.3x10 ⁻²	0.84 (0.75, 0.95)	5.4x10 ⁻³	0.84 (0.79, 0.90)	1.3x10 ⁻⁷	0.84 (0.80, 0.89)	2.4x10 ⁻⁹
rs1050152	chr5:132340627	T/C	SLC22A4	Missense (Leu503Phe)	0.88 (0.76, 1.02)	8.1x10 ⁻²	0.83 (0.76, 0.91)	3.9x10 ⁻⁵	0.84 (0.78, 0.91)	9.5x10 ⁻⁶	0.89 (0.79, 1.00)	4.9x10 ⁻²	0.90 (0.84, 0.95)	5.1x10 ⁻⁴	0.89 (0.85, 0.94)	6.6x10 ⁻⁵
rs174535	chr11:61783884	C/T	MYRF	Synonymous	0.86 (0.74, 1.00)	5.1x10 ⁻²	0.89 (0.81, 0.98)	1.2x10 ⁻²	0.88 (0.82, 0.95)	1.6x10 ⁻³	0.81 (0.71, 0.91)	4.6x10 ⁻⁴	0.88 (0.83, 0.94)	1.0x10 ⁻⁴	0.86 (0.82, 0.91)	3.9x10 ⁻⁷

Supplementary Table 5: Association of 34 reported asthma-associating variants (PMIDs: 29273806, 27681293) with NP (N=4,366 cases; 714,681 controls), NP with confirmed asthma (N=1,528 cases; 712,245 controls), and NP without confirmed asthma (2,849 cases; 711,123 controls) in the combined dataset from Iceland and the UK. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets. NP cases were excluded from the control groups.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP (N=4,366 cases; 714,681 controls)		NP with Asthma (N=1,528 cases; 712,245 controls)		NP without Asthma (2,849 cases; 711,123 controls)		Heterogeneity P_{het}
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	
rs7705042	chr5:142112854	C/A	<i>NDIFP1</i>	Intronic	0.92 (0.88, 0.96)	3.7×10^{-4}	0.90 (0.83, 0.97)	4.9×10^{-3}	0.94 (0.89, 0.99)	2.1×10^{-2}	0.35
rs1233578	chr6:28744470	G/A	-	Intergenic	1.06 (1.00, 1.12)	0.07	0.98 (0.89, 1.08)	0.66	1.10 (1.02, 1.18)	1.1×10^{-2}	5.1×10^{-2}
rs167769	chr12:57109992	C/T	<i>STAT6</i>	Intronic	0.93 (0.89, 0.97)	1.2×10^{-3}	0.93 (0.86, 1.00)	5.3×10^{-2}	0.93 (0.88, 0.98)	7.4×10^{-3}	1.00
rs17637472	chr17:49384071	A/G	-	Intergenic	1.03 (0.99, 1.08)	0.19	1.02 (0.94, 1.10)	0.64	1.04 (0.98, 1.10)	0.21	0.71
rs3131064	chr6:30796116	C/T	-	Intergenic	1.12 (1.06, 1.19)	4.6×10^{-5}	1.11 (1.01, 1.23)	2.6×10^{-2}	1.13 (1.05, 1.21)	5.9×10^{-4}	0.76
rs1420101	chr2:102341256	T/C	<i>IL1RL1</i>	Intronic	1.15 (1.10, 1.20)	3.5×10^{-9}	1.08 (1.01, 1.17)	3.3×10^{-2}	1.18 (1.11, 1.24)	7.4×10^{-9}	5.5×10^{-2}
rs10455025	chr5:111069301	C/A	<i>TSLP</i>	Upstream	1.19 (1.14, 1.25)	1.9×10^{-14}	1.27 (1.18, 1.37)	3.8×10^{-10}	1.15 (1.09, 1.22)	8.0×10^{-7}	3.7×10^{-2}
rs9272346	chr6:32636595	A/G	<i>HLA-DQA1</i>	Upstream	1.20 (1.15, 1.26)	7.2×10^{-16}	1.24 (1.15, 1.33)	3.8×10^{-8}	1.19 (1.12, 1.26)	1.2×10^{-9}	0.40
rs7927894	chr11:76590272	T/C	-	Intergenic	1.04 (1.00, 1.09)	8.1×10^{-2}	1.07 (0.99, 1.15)	8.9×10^{-2}	1.02 (0.97, 1.08)	0.39	0.30
rs2033784	chr15:67157322	G/A	<i>SMAD3</i>	Intronic	1.06 (1.01, 1.11)	1.6×10^{-2}	1.05 (0.97, 1.14)	0.20	1.06 (1.00, 1.13)	4.3×10^{-2}	0.84
rs4233366	chr1:161189357	T/C	<i>ADAMTS4</i>	Downstream	1.03 (0.98, 1.09)	0.21	1.02 (0.94, 1.11)	0.68	1.04 (0.98, 1.10)	0.23	0.74

rs2244012	chr5:132565533	G/A	<i>RAD50</i>	Intronic	1.05 (0.99, 1.11)	0.09	1.10 (1.01, 1.21)	3.8×10^{-2}	1.02 (0.95, 1.10)	0.54	0.18
rs6959584	chr7:106036059	T/C	<i>CDHR3</i>	3' UTR	1.10 (1.03, 1.17)	5.7×10^{-3}	1.22 (1.10, 1.36)	2.7×10^{-4}	1.03 (0.95, 1.11)	0.49	1.5×10^{-2}
rs662064	chr1:10497194	T/C	<i>PEX14</i>	Intronic	1.01 (0.97, 1.06)	0.56	0.98 (0.90, 1.06)	0.59	1.03 (0.98, 1.10)	0.26	0.28
rs6691738	chr1:173182897	G/T	<i>TNFSF4</i>	Downstream	1.01 (0.96, 1.06)	0.82	1.06 (0.97, 1.14)	0.19	0.98 (0.93, 1.04)	0.53	0.15
rs3784099	chr14:68283210	A/G	<i>RAD51B</i>	Intronic	1.09 (1.03, 1.14)	1.5×10^{-3}	1.07 (0.99, 1.16)	8.8×10^{-2}	1.07 (1.01, 1.14)	1.5×10^{-2}	1.00
rs1837253	chr5:111066174	T/C	-	Intergenic	0.74 (0.70, 0.78)	3.8×10^{-29}	0.74 (0.68, 0.81)	2.3×10^{-11}	0.75 (0.70, 0.79)	3.3×10^{-19}	0.81
rs2428494	chr6:31354420	A/T	<i>HLA-B</i>	Intronic	1.12 (1.07, 1.17)	4.8×10^{-7}	1.16 (1.08, 1.25)	7.4×10^{-5}	1.10 (1.04, 1.16)	7.4×10^{-4}	0.26
rs2325291	chr6:90276967	A/G	<i>BACH2</i>	Intronic	0.88 (0.84, 0.93)	1.5×10^{-7}	0.87 (0.81, 0.94)	4.5×10^{-4}	0.89 (0.84, 0.94)	6.0×10^{-5}	0.64
rs2589561	chr10:9004682	A/G	-	Intergenic	1.17 (1.11, 1.23)	2.1×10^{-8}	1.28 (1.17, 1.40)	5.7×10^{-8}	1.11 (1.04, 1.19)	2.3×10^{-3}	1.2×10^{-2}
rs12543811	chr8:80366650	G/A	-	Intergenic	1.05 (1.00, 1.10)	3.5×10^{-2}	1.10 (1.02, 1.19)	1.2×10^{-2}	1.02 (0.96, 1.08)	0.46	0.10
rs17806299	chr16:11106123	A/G	<i>CLEC16A</i>	Intronic	0.91 (0.86, 0.96)	6.4×10^{-4}	0.92 (0.84, 1.01)	8.6×10^{-2}	0.90 (0.84, 0.96)	2.3×10^{-3}	0.71
rs20541	chr5:132660272	A/G	<i>IL13</i>	Missense	1.09 (1.03, 1.15)	3.6×10^{-3}	1.10 (1.00, 1.21)	3.9×10^{-2}	1.08 (1.01, 1.16)	3.1×10^{-2}	0.75
rs992969	chr9:6209697	A/G	-	Intergenic	1.32 (1.26, 1.39)	3.0×10^{-29}	1.22 (1.13, 1.33)	1.7×10^{-6}	1.38 (1.30, 1.46)	5.2×10^{-26}	1.7×10^{-2}
rs11071558	chr15:60777222	G/A	<i>RORA</i>	Intronic	0.95 (0.89, 1.01)	0.10	0.86 (0.77, 0.97)	0.01	1.00 (0.93, 1.07)	0.91	0.90
rs2952156	chr17:39720582	A/G	<i>ERBB2</i>	Downstream	1.06 (1.01, 1.11)	1.1×10^{-2}	1.13 (1.05, 1.22)	1.9×10^{-3}	1.03 (0.97, 1.09)	0.30	5.7×10^{-2}
rs1723018	chr1:167464183	G/A	<i>CD247</i>	Intronic	0.94 (0.90, 0.98)	4.1×10^{-3}	0.93 (0.86, 1.00)	4.8×10^{-2}	0.94 (0.89, 0.99)	3.1×10^{-2}	0.82
rs12413578	chr10:9007290	T/C	-	Intergenic	0.76 (0.70, 0.82)	4.2×10^{-12}	0.66 (0.58, 0.76)	2.6×10^{-9}	0.81 (0.74, 0.89)	1.5×10^{-5}	1.6×10^{-2}
rs11655198	chr17:39869916	C/T	<i>ZPBP2</i>	Intronic	1.07 (1.03, 1.12)	1.7×10^{-3}	1.11 (1.03, 1.19)	7.2×10^{-3}	1.06 (1.00, 1.12)	3.8×10^{-2}	0.34

rs73196739	chr3:188684683	T/C	<i>LPP</i>	Intronic	0.94 (0.89, 1.00)	5.1x10 ⁻²	0.97 (0.88, 1.06)	0.48	0.93 (0.87, 1.00)	6.1x10 ⁻²	0.47
rs6683383	chr1:203131376	A/T	<i>ADORA1</i>	Intronic	1.01 (0.94, 1.08)	0.83	1.04 (0.96, 1.12)	0.35	0.99 (0.93, 1.05)	0.63	0.29
rs34290285	chr2:241759225	A/G	<i>D2HGDH</i>	Intronic	0.95 (0.91, 1.00)	5.7x10 ⁻²	0.89 (0.81, 0.97)	6.0x10 ⁻³	0.99 (0.93, 1.05)	0.70	3.2x10 ⁻²
rs13412757	chr2:8317950	A/G	-	Intergenic	0.93 (0.89, 0.97)	1.7x10 ⁻³	0.92 (0.85, 1.00)	3.8x10 ⁻²	0.93 (0.88, 0.99)	1.4x10 ⁻²	0.83
rs5743618	chr4:38797027	A/C	<i>TLR1</i>	Missense	1.00 (0.96, 1.06)	0.88	1.01 (0.93, 1.10)	0.87	1.00 (0.94, 1.07)	0.98	1.2x10 ⁻²

Supplementary Table 6: Association of 34 reported asthma-associating variants (PMIDs: 29273806, 27681293) with NP with confirmed asthma, and NP without confirmed asthma in the individual Icelandic and UK datasets, as well as the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets. NP cases were excluded from the control groups.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP with Asthma (N _{Ice} =450 cases, 306,781 controls ; N _{UK} =1,078 cases, 405,464 controls)						NP without Asthma (N _{Ice} =736 cases, 305,659 controls ; N _{UK} =2,113 cases, 405,464 controls)					
					OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}	OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}
					(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)	
rs7705042	chr5:142112854	C/A	<i>NDFIP1</i>	Intronic	0.90 (0.77, 1.04)	0.15	0.90 (0.82, 0.98)	1.5x10 ⁻²	0.90 (0.83, 0.97)	4.9x10 ⁻³	0.99 (0.88, 1.11)	0.81	0.92 (0.87, 0.98)	1.2x10 ⁻²	0.94 (0.89, 0.99)	2.1x10 ⁻²
rs1233578	chr6:28744470	G/A	-	Intergenic	1.19 (0.96, 1.47)	0.12	0.93 (0.83, 1.04)	0.19	0.98 (0.89, 1.08)	0.66	1.12 (0.94, 1.34)	0.21	1.09 (1.01, 1.18)	2.6x10 ⁻²	1.10 (1.02, 1.18)	1.1x10 ⁻²
rs167769	chr12:57109992	C/T	<i>STAT6</i>	Intronic	0.87 (0.76, 1.01)	0.06	0.95 (0.87, 1.04)	0.26	0.93 (0.86, 1.00)	5.3x10 ⁻²	0.95 (0.85, 1.07)	0.38	0.92 (0.87, 0.98)	1.0x10 ⁻²	0.93 (0.88, 0.98)	7.4x10 ⁻³
rs17637472	chr17:49384071	A/G	-	Intergenic	0.90 (0.77, 1.05)	0.18	1.06 (0.97, 1.16)	0.18	1.02 (0.94, 1.10)	0.64	1.06 (0.94, 1.20)	0.33	1.03 (0.97, 1.10)	0.36	1.04 (0.98, 1.10)	0.21
rs3131064	chr6:30796116	C/T	-	Intergenic	1.49 (1.21, 1.84)	1.9x10 ⁻⁴	1.03 (0.93, 1.15)	0.55	1.11 (1.01, 1.23)	2.6x10 ⁻²	1.28 (1.08, 1.53)	4.5x10 ⁻³	1.10 (1.02, 1.19)	1.2x10 ⁻²	1.13 (1.05, 1.21)	5.9x10 ⁻⁴
rs1420101	chr2:102341256	T/C	<i>IL1RL1</i>	Intronic	0.94 (0.81, 1.08)	0.38	1.14 (1.05, 1.25)	2.6x10 ⁻³	1.08 (1.01, 1.17)	3.3x10 ⁻²	1.16 (1.03, 1.31)	1.2x10 ⁻²	1.18 (1.11, 1.25)	1.9x10 ⁻⁷	1.18 (1.11, 1.24)	7.4x10 ⁻⁹
rs10455025	chr5:111069301	C/A	<i>TSLP</i>	Upstream	1.14 (0.97, 1.32)	0.10	1.32 (1.21, 1.44)	3.4x10 ⁻¹⁰	1.27 (1.18, 1.37)	3.8x10 ⁻¹⁰	1.17 (1.04, 1.33)	1.0x10 ⁻²	1.15 (1.08, 1.22)	2.4x10 ⁻⁵	1.15 (1.09, 1.22)	8.0x10 ⁻⁷
rs9272346	chr6:32636595	A/G	<i>HLA-DQA1</i>	Upstream	1.25 (1.08, 1.44)	2.3x10 ⁻³	1.23 (1.13, 1.35)	4.7x10 ⁻⁶	1.24 (1.15, 1.33)	3.8x10 ⁻⁸	1.34 (1.19, 1.50)	8.6x10 ⁻⁷	1.15 (1.08, 1.22)	2.3x10 ⁻⁵	1.19 (1.12, 1.26)	1.2x10 ⁻⁹
rs7927894	chr11:76590272	T/C	-	Intergenic	1.04 (0.90, 1.20)	0.63	1.08 (0.99, 1.18)	0.09	1.07 (0.99, 1.15)	8.9x10 ⁻²	0.95 (0.84, 1.07)	0.39	1.05 (0.98, 1.11)	0.16	1.02 (0.97, 1.08)	0.39
rs2033784	chr15:67157322	G/A	<i>SMAD3</i>	Intronic	1.12 (0.96, 1.32)	0.15	1.03 (0.94, 1.13)	0.53	1.05 (0.97, 1.14)	0.20	1.12 (0.98, 1.27)	8.6x10 ⁻²	1.05 (0.98, 1.12)	0.17	1.06 (1.00, 1.13)	4.3x10 ⁻²
rs4233366	chr1:161189357	T/C	<i>ADAMTS4</i>	Downstream	1.13 (0.96, 1.33)	0.14	0.98 (0.89, 1.08)	0.69	1.02 (0.94, 1.11)	0.68	1.24 (1.09, 1.42)	1.3x10 ⁻³	0.99 (0.93, 1.06)	0.77	1.04 (0.98, 1.10)	0.23
rs2244012	chr5:132565533	G/A	<i>RAD50</i>	Intronic	0.95 (0.79, 1.16)	0.63	1.15 (1.04, 1.28)	8.8x10 ⁻³	1.10 (1.01, 1.21)	3.8x10 ⁻²	0.95 (0.81, 1.12)	0.55	1.04 (0.96, 1.12)	0.33	1.02 (0.95, 1.10)	0.54
rs6959584	chr7:106036059	T/C	<i>CDHR3</i>	3' UTR	1.27 (1.01, 1.59)	0.04	1.21 (1.07, 1.36)	2.6x10 ⁻³	1.22 (1.10, 1.36)	2.7x10 ⁻⁴	1.13 (0.94, 1.37)	0.19	1.01 (0.92, 1.10)	0.87	1.03 (0.95, 1.11)	0.49

rs662064	chr1:10497194	T/C	PEX14	Intronic	1.04 (0.89, 1.21)	0.65	0.96 (0.88, 1.05)	0.37	0.98 (0.90, 1.06)	0.59	0.99 (0.87, 1.12)	0.84	1.05 (0.98, 1.12)	0.17	1.03 (0.98, 1.10)	0.26
rs6691738	chr1:173182897	G/T	TNFSF4	Downstream	1.14 (0.98, 1.33)	0.10	1.03 (0.93, 1.13)	0.59	1.06 (0.97, 1.14)	0.19	0.91 (0.80, 1.04)	0.15	1.00 (0.94, 1.07)	0.98	0.98 (0.93, 1.04)	0.53
rs3784099	chr14:68283210	A/G	RAD51B	Intronic	1.03 (0.88, 1.20)	0.72	1.09 (0.99, 1.20)	7.5x10 ⁻²	1.07 (0.99, 1.16)	8.8x10 ⁻²	1.00 (0.89, 1.12)	0.99	1.10 (1.03, 1.18)	4.6x10 ⁻³	1.07 (1.01, 1.14)	1.5x10 ⁻²
rs1837253	chr5:111066174	T/C	-	Intergenic	0.97 (0.83, 1.14)	0.69	0.66 (0.60, 0.73)	1.1x10 ⁻¹⁴	0.74 (0.68, 0.81)	2.3x10 ⁻¹¹	0.75 (0.65, 0.86)	2.8x10 ⁻⁵	0.74 (0.69, 0.80)	2.4x10 ⁻¹⁵	0.75 (0.70, 0.79)	3.3x10 ⁻¹⁹
rs2428494	chr6:31354420	A/T	HLA-B	Intronic	1.09 (0.94, 1.26)	0.26	1.19 (1.09, 1.29)	8.3x10 ⁻⁵	1.16 (1.08, 1.25)	7.4x10 ⁻⁵	1.18 (1.05, 1.33)	5.0x10 ⁻³	1.08 (1.01, 1.14)	2.0x10 ⁻²	1.10 (1.04, 1.16)	7.4x10 ⁻⁴
rs2325291	chr6:90276967	A/G	BACH2	Intronic	0.99 (0.85, 1.15)	0.88	0.83 (0.76, 0.91)	5.6x10 ⁻⁵	0.87 (0.81, 0.94)	4.5x10 ⁻⁴	1.03 (0.91, 1.16)	0.62	0.85 (0.80, 0.91)	1.6x10 ⁻⁶	0.89 (0.84, 0.94)	6.0x10 ⁻⁵
rs2589561	chr10:9004682	A/G	-	Intergenic	1.29 (1.08, 1.54)	5.0x10 ⁻³	1.28 (1.15, 1.42)	3.4x10 ⁻⁶	1.28 (1.17, 1.40)	5.7x10 ⁻⁸	1.15 (0.99, 1.33)	7.0x10 ⁻²	1.10 (1.02, 1.19)	1.3x10 ⁻²	1.11 (1.04, 1.19)	2.3x10 ⁻³
rs12543811	chr8:80366650	G/A	-	Intergenic	0.96 (0.82, 1.11)	0.57	1.16 (1.06, 1.26)	1.3x10 ⁻³	1.10 (1.02, 1.19)	1.2x10 ⁻²	0.99 (0.87, 1.14)	0.94	1.03 (0.96, 1.10)	0.39	1.02 (0.96, 1.08)	0.46
rs17806299	chr16:11106123	A/G	CLEC16A	Intronic	0.96 (0.79, 1.17)	0.69	0.91 (0.81, 1.01)	8.1x10 ⁻²	0.92 (0.84, 1.01)	8.6x10 ⁻²	1.04 (0.89, 1.21)	0.62	0.86 (0.80, 0.93)	2.3x10 ⁻⁴	0.90 (0.84, 0.96)	2.3x10 ⁻³
rs20541	chr5:132660272	A/G	IL13	Missense	0.99 (0.82, 1.20)	0.90	1.14 (1.03, 1.27)	1.5x10 ⁻²	1.10 (1.00, 1.21)	3.9x10 ⁻²	1.09 (0.94, 1.27)	0.24	1.07 (0.99, 1.16)	6.8x10 ⁻²	1.08 (1.01, 1.16)	3.1x10 ⁻²
rs992969	chr9:6209697	A/G	-	Intergenic	1.11 (0.94, 1.31)	0.20	1.26 (1.15, 1.39)	1.5x10 ⁻⁶	1.22 (1.13, 1.33)	1.7x10 ⁻⁶	1.40 (1.23, 1.59)	2.1x10 ⁻⁷	1.37 (1.28, 1.47)	3.9x10 ⁻²⁰	1.38 (1.30, 1.46)	5.2x10 ⁻²⁶
rs11071558	chr15:60777222	G/A	RORA	Intronic	0.98 (0.79, 1.21)	0.83	0.82 (0.72, 0.94)	3.7x10 ⁻³	0.86 (0.77, 0.97)	1.0x10 ⁻²	0.97 (0.82, 1.14)	0.71	1.00 (0.92, 1.09)	0.96	1.00 (0.93, 1.07)	0.91
rs2952156	chr17:39720582	A/G	ERBB2	Downstream	1.20 (1.03, 1.39)	2.0x10 ⁻²	1.11 (1.01, 1.21)	2.6x10 ⁻²	1.13 (1.05, 1.22)	1.9x10 ⁻³	1.07 (0.95, 1.21)	0.28	1.02 (0.96, 1.09)	0.55	1.03 (0.97, 1.09)	0.30
rs1723018	chr1:167464183	G/A	CD247	Intronic	0.95 (0.82, 1.10)	0.50	0.92 (0.84, 1.00)	5.8x10 ⁻²	0.93 (0.86, 1.00)	4.8x10 ⁻²	0.94 (0.83, 1.05)	0.27	0.94 (0.89, 1.00)	6.3x10 ⁻²	0.94 (0.89, 0.99)	3.1x10 ⁻²
rs12413578	chr10:9007290	T/C	-	Intergenic	0.59 (0.45, 0.78)	1.9x10 ⁻⁴	0.69 (0.59, 0.80)	2.2x10 ⁻⁶	0.66 (0.58, 0.76)	2.6x10 ⁻⁹	0.77 (0.63, 0.95)	1.6x10 ⁻²	0.82 (0.74, 0.91)	2.8x10 ⁻⁴	0.81 (0.74, 0.89)	1.5x10 ⁻⁵
rs11655198	chr17:39869916	C/T	ZBP2	Intronic	1.24 (1.08, 1.43)	2.7x10 ⁻³	1.06 (0.97, 1.15)	0.18	1.11 (1.03, 1.19)	7.2x10 ⁻³	1.10 (0.98, 1.24)	9.7x10 ⁻²	1.05 (0.98, 1.11)	0.14	1.06 (1.00, 1.12)	3.8x10 ⁻²
rs73196739	chr3:188684683	T/C	LPP	Intronic	1.04 (0.87, 1.26)	0.64	0.94 (0.83, 1.05)	0.26	0.97 (0.88, 1.06)	0.48	0.92 (0.79, 1.07)	0.28	0.94 (0.86, 1.02)	0.12	0.93 (0.87, 1.00)	6.1x10 ⁻²
rs6683383	chr1:203131376	A/T	ADORA1	Intronic	1.07 (0.92, 1.25)	0.38	1.03 (0.94, 1.12)	0.56	1.04 (0.96, 1.12)	0.35	0.95 (0.84, 1.08)	0.42	1.00 (0.93, 1.07)	0.91	0.99 (0.93, 1.05)	0.63

rs34290285	chr2:241759225	A/G	<i>D2HGDH</i>	Intronic	1.17 (1.00, 1.37)	5.0x10 ⁻²	0.79 (0.71, 0.87)	4.8x10 ⁻⁶	0.89 (0.81, 0.97)	6.0x10 ⁻³	0.90 (0.79, 1.03)	0.13	1.01 (0.94, 1.09)	0.70	0.99 (0.93, 1.05)	0.70
rs13412757	chr2:8317950	A/G	-	Intergenic	0.87 (0.75, 1.01)	7.0x10 ⁻²	0.94 (0.86, 1.03)	0.18	0.92 (0.85, 1.00)	3.8x10 ⁻²	0.95 (0.85, 1.08)	0.46	0.92 (0.87, 0.99)	1.7x10 ⁻²	0.93 (0.88, 0.99)	1.4x10 ⁻²
rs5743618	chr4:38797027	A/C	<i>TLR1</i>	Missense	0.98 (0.84, 1.14)	0.77	1.02 (0.92, 1.13)	0.70	1.01 (0.93, 1.10)	0.87	0.98 (0.86, 1.11)	0.74	1.01 (0.94, 1.08)	0.83	1.00 (0.94, 1.07)	0.98

Supplementary Table 7: Association of the 10 variants identified in the combined GWAS with NP, asthma, asthma with confirmed NP, and asthma without confirmed NP in the combined dataset from Iceland and the UK. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	NP vs. controls (N=4,366 cases; 714,681 controls)		Asthma vs. controls (39,342 cases; 731,538 controls)		Asthma with NP vs. controls (N=1,528 cases, 688,086 controls)		Asthma without NP vs. controls (37,881 cases, 730,507 controls)		Heterogeneity P_{het}
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.47 (1.39, 1.54)	7.3×10^{-49}	1.17 (1.15, 1.20)	9.9×10^{-60}	1.53 (1.40, 1.66)	1.3×10^{-22}	1.16 (1.14, 1.19)	3.2×10^{-50}	5.3×10^{-10}
rs1888909	chr9:6197392	T/C	<i>[IL33]</i>	Intergenic	1.34 (1.27, 1.40)	7.4×10^{-32}	1.12 (1.10, 1.14)	4.2×10^{-34}	1.25 (1.15, 1.35)	1.3×10^{-7}	1.12 (1.10, 1.14)	3.9×10^{-31}	1.1×10^{-2}
rs1837253	chr5:111066174	T/C	<i>[TSLP]</i>	Intergenic	0.74 (0.70, 0.78)	3.7×10^{-29}	0.91 (0.89, 0.93)	2.3×10^{-23}	0.74 (0.68, 0.81)	8.8×10^{-12}	0.92 (0.90, 0.94)	2.7×10^{-19}	1.4×10^{-6}
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.32 (0.26, 0.39)	8.0×10^{-27}	0.94 (0.89, 0.99)	0.03	0.37 (0.26, 0.51)	3.1×10^{-9}	0.96 (0.91, 1.02)	0.17	2.2×10^{-8}
rs1444782	chr10:9016708	A/G	<i>10p14</i>	Intergenic	0.84 (0.80, 0.88)	1.1×10^{-14}	0.90 (0.89, 0.91)	1.8×10^{-36}	0.74 (0.69, 0.80)	2.3×10^{-14}	0.91 (0.89, 0.92)	4.0×10^{-30}	2.9×10^{-7}
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.85 (0.81, 0.89)	1.3×10^{-11}	0.98 (0.97, 1.00)	0.04	0.79 (0.73, 0.86)	1.1×10^{-8}	0.99 (0.97, 1.01)	0.30	1.0×10^{-7}
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	0.86 (0.82, 0.90)	5.6×10^{-11}	0.99 (0.98, 1.01)	0.50	0.98 (0.92, 1.06)	0.68	1.00 (0.99, 1.02)	0.86	0.68
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	0.87 (0.83, 0.91)	7.4×10^{-10}	0.93 (0.91, 0.94)	1.3×10^{-19}	0.90 (0.84, 0.97)	8.5×10^{-3}	0.93 (0.91, 0.94)	8.7×10^{-19}	0.42
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.88 (0.84, 0.92)	1.1×10^{-8}	0.95 (0.94, 0.97)	7.3×10^{-9}	0.84 (0.78, 0.91)	8.1×10^{-6}	0.96 (0.94, 0.97)	2.2×10^{-7}	8.1×10^{-4}
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.87 (0.83, 0.91)	3.5×10^{-9}	0.95 (0.94, 0.97)	7.0×10^{-8}	0.88 (0.81, 0.95)	1.2×10^{-3}	0.96 (0.94, 0.97)	8.6×10^{-7}	3.1×10^{-2}

Supplementary Table 8: Association of the 10 variants identified in the combined GWAS with asthma with confirmed NP, and asthma without confirmed NP in the individual Icelandic and UK datasets, as well as the combined dataset, both tested against population controls. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	Asthma with NP vs. controls (N _{Ice} =450 cases, 307,696 controls ; N _{UK} =1,078 cases, 380,390 controls)						Asthma without NP vs. controls (N _{Ice} =12,807 cases, 350,117 controls ; N _{UK} =25,074 cases, 380,390 controls)					
					OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}	OR _{Ice}	P _{Ice}	OR _{UK}	P _{UK}	OR _{meta}	P _{meta}
					(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)		(95% CI)	
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.49 (1.26, 1.76)	3.2x10 ⁻⁶	1.54 (1.40, 1.70)	7.2x10 ⁻¹⁸	1.53 (1.40, 1.66)	1.3x10 ⁻²²	1.18 (1.14, 1.23)	8.4x10 ⁻¹⁶	1.16 (1.13, 1.18)	2.7x10 ⁻³⁶	1.16 (1.14, 1.19)	3.2x10 ⁻⁵⁰
rs1888909	chr9:6197392	T/C	<i>[IL33]</i>	Intergenic	1.13 (0.96, 1.32)	0.15	1.29 (1.17, 1.42)	1.3x10 ⁻⁷	1.25 (1.15, 1.35)	1.3x10 ⁻⁷	1.13 (1.08, 1.17)	9.6x10 ⁻¹⁰	1.11 (1.09, 1.14)	5.1x10 ⁻²³	1.12 (1.10, 1.14)	3.9x10 ⁻³¹
rs1837253	chr5:111066174	T/C	<i>[TSLP]</i>	Intergenic	0.96 (0.82, 1.13)	0.66	0.66 (0.59, 0.73)	3.3x10 ⁻¹⁵	0.74 (0.68, 0.81)	8.8x10 ⁻¹²	0.96 (0.92, 0.99)	2.0x10 ⁻²	0.91 (0.89, 0.92)	1.6x10 ⁻¹⁹	0.92 (0.90, 0.94)	2.7x10 ⁻¹⁹
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.34 (0.18, 0.62)	4.3x10 ⁻⁴	0.38 (0.26, 0.57)	1.8x10 ⁻⁶	0.37 (0.26, 0.51)	3.1x10 ⁻⁹	0.94 (0.85, 1.04)	0.22	0.97 (0.91, 1.04)	0.41	0.96 (0.91, 1.02)	0.17
rs1444782	chr10:9016708	A/G	<i>10p14</i>	Intergenic	0.75 (0.64, 0.87)	1.4x10 ⁻⁴	0.74 (0.68, 0.81)	3.7x10 ⁻¹¹	0.74 (0.69, 0.80)	2.3x10 ⁻¹⁴	0.93 (0.90, 0.96)	2.4x10 ⁻⁵	0.90 (0.89, 0.91)	1.2x10 ⁻²⁶	0.91 (0.89, 0.92)	4.0x10 ⁻³⁰
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.86 (0.73, 1.01)	6.9x10 ⁻²	0.76 (0.69, 0.84)	2.8x10 ⁻⁸	0.79 (0.73, 0.86)	1.1x10 ⁻⁸	0.98 (0.94, 1.01)	0.20	0.99 (0.97, 1.02)	0.64	0.99 (0.97, 1.01)	0.30
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	1.01 (0.88, 1.17)	0.84	0.97 (0.89, 1.06)	0.55	0.98 (0.92, 1.06)	0.68	0.99 (0.96, 1.02)	0.52	1.00 (0.99, 1.02)	0.59	1.00 (0.99, 1.02)	0.86
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	1.01 (0.87, 1.18)	0.90	0.87 (0.79, 0.95)	1.8x10 ⁻³	0.90 (0.84, 0.97)	8.5x10 ⁻³	0.96 (0.92, 0.99)	1.2x10 ⁻²	0.92 (0.90, 0.93)	2.1x10 ⁻¹⁸	0.93 (0.91, 0.94)	8.7x10 ⁻¹⁹
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.88 (0.76, 1.02)	8.5x10 ⁻²	0.83 (0.76, 0.91)	3.1x10 ⁻⁵	0.84 (0.78, 0.91)	8.1x10 ⁻⁶	0.94 (0.91, 0.97)	2.1x10 ⁻⁴	0.96 (0.94, 0.98)	1.1x10 ⁻⁴	0.96 (0.94, 0.97)	2.2x10 ⁻⁷
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.86 (0.74, 1.00)	4.6x10 ⁻²	0.89 (0.81, 0.97)	1.0x10 ⁻²	0.88 (0.81, 0.95)	1.2x10 ⁻³	0.97 (0.94, 1.01)	0.11	0.95 (0.94, 0.97)	1.9x10 ⁻⁶	0.96 (0.94, 0.97)	8.6x10 ⁻⁷

Supplementary Table 9: Association of the 10 variants identified in the combined GWAS with asthma with confirmed NP tested against asthma without confirmed NP in the individual Icelandic and UK datasets, as well as the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	Asthma with NP (N=1,078) vs. Asthma without NP (N=25,074)					
					OR _{Ice} (95% CI)	P _{Ice}	OR _{UK} (95% CI)	P _{UK}	OR _{meta} (95% CI)	P _{meta}
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.25 (1.06, 1.49)	9.2x10 ⁻³	1.34 (1.21, 1.48)	1.2x10 ⁻⁸	1.32 (1.21, 1.44)	4.5x10 ⁻¹⁰
rs1888909	chr9:6197392	T/C	[<i>IL33</i>]	Intergenic	1.03 (0.87, 1.21)	0.76	1.16 (1.05, 1.27)	2.8x10 ⁻³	1.12 (1.03, 1.22)	6.0x10 ⁻³
rs1837253	chr5:111066174	T/C	[<i>TSLP</i>]	Intergenic	1.00 (0.86, 1.16)	0.98	0.73 (0.65, 0.81)	7.1x10 ⁻⁹	0.81 (0.74, 0.88)	2.3x10 ⁻⁶
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.37 (0.20, 0.68)	1.6x10 ⁻³	0.39 (0.26, 0.58)	4.5x10 ⁻⁶	0.38 (0.27, 0.54)	2.6x10 ⁻⁸
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.82 (0.70, 0.95)	9.1x10 ⁻³	0.83 (0.76, 0.91)	4.5x10 ⁻⁵	0.83 (0.76, 0.89)	1.3x10 ⁻⁶
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.88 (0.75, 0.95)	0.14	0.77 (0.70, 0.85)	1.1x10 ⁻⁷	0.80 (0.73, 0.87)	1.1x10 ⁻⁷
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	1.03 (0.89, 1.19)	0.68	0.96 (0.88, 1.05)	0.42	0.98 (0.91, 1.06)	0.64
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	1.05 (0.91, 1.22)	0.50	0.95 (0.87, 1.04)	0.30	0.98 (0.91, 1.06)	0.60
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.94 (0.80, 1.09)	0.38	0.86 (0.79, 0.94)	1.3x10 ⁻³	0.88 (0.82, 0.95)	1.3x10 ⁻³
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.88 (0.75, 1.02)	0.09	0.93 (0.85, 1.02)	0.14	0.92 (0.85, 0.99)	3.4x10 ⁻²

Supplementary Table 10: Association of the 10 variants identified in the combined GWAS with NP in Iceland, the UK, and the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Allele (min/maj)	Gene / [Locus]	Coding effect	LD class ($r^2 > 0.8$ with variant)		Iceland (1,175 cases; 309,305 controls)			UK Biobank (3,191 cases, 405,376 controls)			Combined (N=4,366 cases; 714,681 controls)		
				N variants	N coding variants	MAF (%)	OR (95% CI)	P	MAF (%)	OR (95% CI)	P	OR (95% CI)	P	P _{het}
rs1391371 (chr6:32636021)	T/A	HLA-DQA1	Upstream	196	5	19.2	1.52 (1.37, 1.69)	1.1x10 ⁻¹⁴	19.8	1.45 (1.37, 1.54)	6.0x10 ⁻³⁶	1.47 (1.39, 1.54)	7.3x10 ⁻⁴⁹	0.43
rs1888909 (chr9:6197392)	T/C	[IL33]	Intergenic	9	0	24.5	1.30 (1.18, 1.44)	3.4x10 ⁻⁷	26.0	1.35 (1.27, 1.42)	3.2x10 ⁻²⁶	1.34 (1.27, 1.40)	7.4x10 ⁻³²	0.59
rs1837253 (chr5:111066174)	T/C	[TSLP]	Intergenic	1	0	28.8	0.83 (0.74, 0.92)	3.4x10 ⁻⁴	26.6	0.71 (0.67, 0.76)	1.5x10 ⁻²⁷	0.74 (0.70, 0.78)	3.7x10 ⁻²⁹	0.02
rs34210653 (chr17:4632019)	A/G	ALOX15	Missense (Thr560Met)	2	1	2.99	0.33 (0.22, 0.49)	3.8x10 ⁻⁸	2.03	0.32 (0.25, 0.40)	3.4x10 ⁻²⁰	0.32 (0.26, 0.39)	8.0x10 ⁻²⁷	0.89
rs1444782 (chr10:9016708)	A/G	10p14	Intergenic	12	0	38.8	0.81 (0.74, 0.90)	2.9x10 ⁻⁵	41.8	0.84 (0.80, 0.89)	6.4x10 ⁻¹¹	0.84 (0.80, 0.88)	1.1x10 ⁻¹⁴	0.54
rs17718444 (chr3:71450250)	T/C	FOXP1	Intron	42	0	28.5	0.90 (0.81, 1.00)	4.2x10 ⁻²	31.4	0.83 (0.79, 0.88)	4.8x10 ⁻¹¹	0.85 (0.81, 0.89)	1.3x10 ⁻¹¹	0.19
rs338598 (chr19:41195623)	A/C	CYP2S1	Intron	1	0	48.0	0.87 (0.79, 0.95)	3.1x10 ⁻³	47.4	0.86 (0.82, 0.90)	4.8x10 ⁻⁹	0.86 (0.82, 0.90)	5.6x10 ⁻¹¹	0.81
rs6543124 (chr2:102370999)	A/T	IL18R1	Intron	84	5	38.9	0.91 (0.83, 1.00)	5.5x10 ⁻²	38.1	0.85 (0.81, 0.90)	2.2x10 ⁻⁹	0.87 (0.83, 0.91)	7.4x10 ⁻¹⁰	0.22
rs1050152 (chr5:132340627)	T/C	SLC22A4	Missense (Leu503Phe)	43	1	39.6	0.89 (0.81, 0.98)	1.5x10 ⁻²	41.5	0.88 (0.83, 0.92)	2.4x10 ⁻⁷	0.88 (0.84, 0.92)	1.1x10 ⁻⁸	0.77
rs174535 (chr11:61783884)	C/T	MYRF	Synonymous	65	0	39.2	0.82 (0.75, 0.91)	8.7x10 ⁻⁵	34.5	0.88 (0.84, 0.93)	4.8x10 ⁻⁶	0.87 (0.83, 0.91)	3.5x10 ⁻⁹	0.23

Supplementary Table 11: Association of the 10 variants identified in the combined GWAS with CRS in Iceland, the UK, and the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	Iceland (3,188 cases; 353,939 controls)		UK (2,420 cases; 406,147 controls)		Combined (5,608 cases; 760,086 controls)		
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	P _{het}
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.25 (1.16, 1.34)	7.4x10 ⁻¹⁰	1.19 (1.11, 1.27)	6.2x10 ⁻⁷	1.22 (1.16, 1.28)	3.8x10 ⁻¹⁵	0.33
rs1888909	chr9:6197392	T/C	[<i>IL33</i>]	Intergenic	1.12 (1.05, 1.20)	5.6x10 ⁻⁴	1.12 (1.06, 1.20)	3.0x10 ⁻⁴	1.12 (1.07, 1.18)	5.9x10 ⁻⁷	0.98
rs1837253	chr5:111066174	T/C	[<i>TSLP</i>]	Intergenic	0.96 (0.90, 1.03)	0.23	0.91 (0.85, 0.97)	3.9x10 ⁻³	0.93 (0.89, 0.98)	3.9x10 ⁻³	0.23
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.66 (0.54, 0.81)	5.7x10 ⁻⁵	0.62 (0.50, 0.78)	4.7x10 ⁻⁵	0.64 (0.55, 0.75)	1.1x10 ⁻⁸	0.70
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.87 (0.81, 0.92)	3.5x10 ⁻⁶	0.93 (0.87, 0.98)	8.6x10 ⁻³	0.90 (0.86, 0.94)	3.5x10 ⁻⁷	0.12
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.98 (0.92, 1.05)	0.59	0.94 (0.88, 1.00)	0.04	0.96 (0.92, 1.00)	0.06	0.30
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	0.92 (0.87, 0.97)	3.8x10 ⁻³	0.92 (0.87, 0.97)	3.1x10 ⁻³	0.92 (0.88, 0.96)	3.5x10 ⁻⁵	0.94
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	0.95 (0.89, 1.01)	0.08	0.93 (0.88, 0.99)	0.02	0.94 (0.90, 0.98)	4.2x10 ⁻³	0.75
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.92 (0.86, 0.97)	4.3x10 ⁻³	0.93 (0.88, 0.99)	0.02	0.92 (0.89, 0.96)	2.3x10 ⁻⁴	0.67
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.96 (0.90, 1.02)	0.16	0.98 (0.93, 1.04)	0.52	0.97 (0.93, 1.01)	0.15	0.58

Supplementary Table 12: Association of the 10 variants identified in the combined GWAS with asthma in Iceland, the UK, and the combined dataset. Effects are presented for the Icelandic minor allele. Genes linked to intergenic variants are listed in square brackets.

Marker	Position	Allele (min/maj)	Gene	Coding effect	Iceland (13,190 cases; 349,035 controls)		UK (26,152 cases; 382,503 controls)		Combined (39,342 cases; 731,538 controls)		
					OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	P _{het}
rs1391371	chr6:32636021	T/A	<i>HLA-DQA1</i>	Upstream	1.20 (1.15, 1.25)	2.2x10 ⁻¹⁸	1.17 (1.14, 1.19)	2.6x10 ⁻⁴³	1.17 (1.15, 1.20)	9.9x10 ⁻⁶⁰	0.27
rs1888909	chr9:6197392	T/C	[<i>IL33</i>]	Intergenic	1.13 (1.08, 1.17)	7.0x10 ⁻¹⁰	1.12 (1.09, 1.14)	8.0x10 ⁻²⁶	1.12 (1.10, 1.14)	4.2x10 ⁻³⁴	0.69
rs1837253	chr5:111066174	T/C	[<i>TSLP</i>]	Intergenic	0.96 (0.92, 0.99)	0.02	0.90 (0.88, 0.92)	2.6x10 ⁻²⁴	0.91 (0.89, 0.93)	2.3x10 ⁻²³	2.1x10 ⁻³
rs34210653	chr17:4632019	A/G	<i>ALOX15</i>	Missense (Thr560Met)	0.92 (0.83, 1.01)	0.09	0.95 (0.89, 1.01)	0.13	0.94 (0.89, 0.99)	0.03	0.52
rs1444782	chr10:9016708	A/G	10p14	Intergenic	0.92 (0.88, 0.95)	6.2x10 ⁻⁷	0.90 (0.88, 0.91)	2.5x10 ⁻³¹	0.90 (0.89, 0.91)	1.8x10 ⁻³⁶	0.25
rs17718444	chr3:71450250	T/C	<i>FOXP1</i>	Intron	0.97 (0.93, 1.00)	0.08	0.99 (0.97, 1.01)	0.15	0.98 (0.97, 1.00)	0.04	0.39
rs338598	chr19:41195623	A/C	<i>CYP2S1</i>	Intron	0.99 (0.96, 1.03)	0.65	0.99 (0.98, 1.01)	0.60	0.99 (0.98, 1.01)	0.50	0.88
rs6543124	chr2:102370999	A/T	<i>IL18R1</i>	Intron	0.96 (0.93, 0.99)	0.02	0.92 (0.90, 0.93)	1.2x10 ⁻¹⁹	0.93 (0.91, 0.94)	1.3x10 ⁻¹⁹	0.02
rs1050152	chr5:132340627	T/C	<i>SLC22A4</i>	Missense (Leu503Phe)	0.94 (0.90, 0.97)	1.4x10 ⁻⁴	0.96 (0.94, 0.98)	6.4x10 ⁻⁶	0.95 (0.94, 0.97)	7.3x10 ⁻⁹	0.24
rs174535	chr11:61783884	C/T	<i>MYRF</i>	Synonymous	0.97 (0.93, 1.00)	0.05	0.95 (0.93, 0.97)	3.7x10 ⁻⁷	0.95 (0.94, 0.97)	7.0x10 ⁻⁸	0.44

Supplementary Table 13: Association of Thr560Met in *ALOX15* with white blood cell counts and other phenotypes related to NP in the Icelandic dataset.

Phenotype	Measurements/ Cases	OR/Effect	<i>P</i>
Eosinophil count	251,307	-0.073 SD	1.1×10^{-16}
Nasal polyps	1,175	0.327	3.8×10^{-8}
Chronic sinusitis	3,188	0.662	5.7×10^{-5}
Basophil count	251,273	-0.015 SD	0.01
Allergic rhinitis	6,636	0.851	0.02
Asthma	13,190	0.915	0.09
Monocyte count	251,307	-0.014 SD	0.13
Lymphocyte count	251,307	-0.010 SD	0.22
Neutrophil count	251,307	0.007 SD	0.42

Supplementary Table 14: Association of Thr560Met with the most common cancers and with coronary artery disease in the Icelandic, UK, and combined datasets. None of the associations are significant.

Phenotype	Iceland			UK			Combined	
	OR (95% CI)	P	N cases	OR (95% CI)	P	N cases	OR (95% CI)	P
Breast cancer	0.93 (0.82, 1.06)	0.27	3,810	0.99 (0.87, 1.13)	0.88	9,997	0.96 (0.88, 1.05)	0.37
Colorectal cancer	0.91 (0.78, 1.06)	0.22	4,146	0.98 (0.91, 1.06)	0.60	12,753	0.97 (0.90, 1.03)	0.31
Kidney cancer	1.17 (0.94, 1.46)	0.16	1,673	1.04 (0.76, 1.43)	0.81	999	1.13 (0.94, 1.35)	0.20
Lung cancer	1.05 (0.92, 1.20)	0.48	4,958	1.26 (1.04, 1.53)	0.02	2,083	1.11 (1.00, 1.25)	0.06
Pancreatic cancer	1.03 (0.81, 1.31)	0.81	1,307	1.03 (0.72, 1.47)	0.87	586	1.03 (0.84, 1.26)	0.77
Prostate cancer	1.08 (0.95, 1.23)	0.25	5,689	0.96 (0.85, 1.08)	0.50	5,817	1.01 (0.93, 1.11)	0.79
Thyroid cancer	0.80 (0.60, 1.07)	0.13	1,209	1.48 (0.94, 2.33)	0.09	347	0.96 (0.75, 1.22)	0.71
Melanoma	0.89 (0.71, 1.11)	0.31	1,786	1.67 (0.78, 3.60)	0.19	115	0.94 (0.75, 1.16)	0.55
Coronary artery disease	1.05 (0.98, 1.13)	0.17	38,918	0.99 (0.94, 1.04)	0.69	28,110	1.01 (0.97, 1.05)	0.64

Supplementary Table 15: Summary of annotations for variants and loci. Functional annotations in terms of coding and regulatory effects are listed out for the lead variant and variants in LD ($r^2>0.8$) for each locus.

Chr	rsName	Freq	Gene alterations	Promoter alterations	Enhancer alterations		LD class (span)
			Gene with coding or LOF variants in LD class	Variants in LD class intersecting with promoters for genes*	Variants in LD class intersecting with enhancer (eRNA)*	Enhancer-Gene target predictions (JEME)**	
2	rs6543124	0.389	IL1RL1 (missense variants in LD; rs10192036, rs10204137, rs10192157, rs10206753)	rs4851566 (IL18R1, alternative promoter)			81 (20.2kb)
3	rs17718444	0.285			chr3:71460229(indel), rs62247168, rs62247169, rs62247170, rs62247171, rs60281674, rs60405683, rs6549392, chr3:71489869(snv), rs6549393	EIF4E3, FOXP1, GPR27	38 (78.0kb)
5	rs1837253	0.288					1
5	rs1050152	0.396	SLC22A4 (lead variant is missense; rs1050152)	rs2631367 (SLC22A5 alternative promoter)	rs2706336	IL5, PDLIM4, IL13, IRF1, SLC22A4, SLC22A5, P4HA2, RAD50	43 (236.0kb)
6	rs1391371	0.192	HLA-DQA1 (missense variant in LD; rs9272785)	rs9272425 (HLA-DQA1, primary promoter); chr6:32643071:I:2, chr6:32643074:S:2 (HLA-DQA1, alternative promoter)	rs17425622, rs908566839, rs532965, rs3104413, rs9271494, rs9271509, rs9271579, rs9271580, chr6:32623009(indel), rs9271594, rs9271608, rs17426593, rs41269951, rs41269955, rs201195623, rs41269959, rs34141382, chr6:32640730(snv)	TAP2, TAP1, SLC44A4, BAK1, HLA-DQB2, HLA-DQB1, C2, AIF1, HLA-DMB, HLA-DOB, HLA-DPB1, HLA-DQA1, HLA-DRB1, HLA-DRB5, PSMB9, NELFE, TAPBP, B3GALT4, PSMB8, STK19, BRD2, HLA-DMA, C6orf10, RNF5, SKIV2L, HSD17B8, HLA-DRA, BTNL2, HLA-DQA2, LY6G5B, FKBPL, AGPAT1, HLA-DPA1, GPANK1, SLC39A7, RPS18, NOTCH4	188 (192.2kb)
9	rs1888909	0.245					9 (20.7kb)
10	rs1444782	0.388			rs1031163, rs1444782		12 (19.4kb)
11	rs174535	0.391		rs174538 (C11orf10, primary promoter)	rs199985711, rs149597144	MYRF, RAB3IL1, FADS2, FTH1, FEN1, DAGLA, FADS1, TMEM258, INCENP, FADS3	61 (81.1kb)
17	rs34210653	0.030	ALOX15 (lead variant is missense; rs34210653)				2 (13.8kb)
19	rs338598	0.480			rs338598	CYP2A13, CYP2F1, TMEM91, AXL, HNRNPUL1, CYP2S1	1

* Promoter and enhancer data were derived from the Fantom5 project and intersected with variant positions. The promoter data cover approximately 0.5% of the genome, while the enhancer data (extended by +/-250bp) cover approximately 3.4% of the genome.

** Joint effect of multiple enhancers (JEME) provides predictions, based on the data derived from the Fantom5 project, for enhancers in terms of their effects on gene expression.

Supplementary Table 16: Associations of variants in the *HLA* region with NP in Iceland, the UK, and the combined dataset. The signal is represented by the intergenic variant rs1391371[T] 1.37kb upstream of in *HLA-DQA1*, which is highly correlated to four missense variants in *HLA-DQA1*. Correlation is calculated in the Icelandic dataset. Two of the four correlated missense variants are not tested in the UK Biobank dataset.

Marker	Position	Allele (min/maj)	Iceland (N=1,175 cases; 309,305 controls)			UK (N=3,191 cases; 405,376 controls)			Combined (N=4,366 cases; 714,681 controls)		Relation to rs1391371		
			MAF (%)	OR (95% CI)	P	MAF (%)	OR (95% CI)	P	OR (95% CI)	P	Distance (BP)	r ²	D'
rs1391371	chr6:32636021	T/A	19.2	1.52 (1.37, 1.69)	1.1x10 ⁻¹⁴	19.8	1.45 (1.37, 1.54)	6.0x10 ⁻³⁶	1.47 (1.39, 1.54)	7.3x10 ⁻⁴⁹	-	-	-
rs9272785	chr6:32642624	A/G	19.2	1.52 (1.37, 1.69)	1.0x10 ⁻¹⁴	18.6	1.43 (1.35, 1.52)	7.5x10 ⁻³¹	1.45 (1.38, 1.53)	1.0x10 ⁻⁴³	6603	1.00	1.00
rs1048023	chr6:32641373	G/C	19.2	1.52 (1.37, 1.69)	9.0x10 ⁻¹⁵	-	-	-	1.52 (1.37, 1.69)	9.0x10 ⁻¹⁵	5352	1.00	1.00
rs3207983	chr6:32641436	A/G	17.2	1.54 (1.38, 1.72)	3.3x10 ⁻¹⁴	23.4	0.95 (0.90, 1.00)	0.07	1.05 (1.00, 1.10)	0.05	5415	0.86	0.99
rs12722061	chr6:32641444	C/G	16.4	1.55 (1.38, 1.74)	8.6x10 ⁻¹⁴	34.9	1.21 (1.15, 1.27)	6.5x10 ⁻¹³	1.26 (1.20, 1.32)	7.5x10 ⁻²²	5423	0.81	0.99
rs1064944	chr6:62641522	G/A	19.7	1.52 (1.37, 1.69)	7.8x10 ⁻¹⁵	-	-	-	1.52 (1.37, 1.69)	7.8x10 ⁻¹⁵	5501	0.97	1.00

MAF = minor allele frequency; OR = odds ratio; CI = confidence interval

Supplementary Table 17: Sequence variants reported to associate with eosinophil count that associate significantly with NP in the combined dataset ($P < 0.05/210 = 2.4 \times 10^{-4}$). Effects are presented for the minor allele in Iceland. Genes linked to intergenic variants are listed in square brackets.

Marker	Chr:Position	Allele (min/maj)	Gene / [Locus]	Coding effect	Eosinophil effect (SD)	LD class ($r^2 > 0.8$ with variant)		Iceland (1,175 cases; 309,305 controls)			UK Biobank (3,191 cases, 405,376 controls)			Combined (N=4,366 cases; 714,681 controls)		
						N variants	N coding variants	MAF (%)	OR (95% CI)	P	MAF (%)	OR (95% CI)	P	OR (95% CI)	P	P _{het}
rs78757963	chr9:6282511	A/G	[IL33]	intergenic	0.050	3	0	3.69	1.22 (0.97, 1.54)	0.09	3.38	1.48 (1.31, 1.66)	2.5×10^{-10}	1.42 (1.27, 1.58)	1.6×10^{-10}	0.15
rs8046011	chr16:11226805	G/A	16p13	Intergenic	0.021	2	0	27.9	1.08 (0.98, 1.20)	0.11	29.3	1.18 (1.11, 1.24)	4.9×10^{-9}	1.15 (1.10, 1.21)	3.6×10^{-9}	0.17
rs28383314	chr6:32619436	T/C	6p21	Intergenic	-0.035	28	0	36.1	0.90 (0.82, 0.99)	0.03	37.8	0.88 (0.83, 0.92)	5.4×10^{-7}	0.88 (0.84, 0.92)	5.1×10^{-8}	0.64
rs62408225	chr6:90246690	G/A	BACH2	Intron	-0.020	65	0	34.4	1.01 (0.92, 1.12)	0.78	34.3	0.84 (0.80, 0.89)	3.4×10^{-10}	0.88 (0.84, 0.92)	7.8×10^{-8}	1.1×10^{-3}
rs4807542	chr19:1104079	A/G	GPX4	Synonymous	-0.010	2	0	16.5	0.89 (0.78, 1.01)	0.06	16.5	0.83 (0.78, 0.89)	3.2×10^{-7}	0.84 (0.79, 0.90)	7.8×10^{-8}	0.40
rs10905284	chr10:8073399	C/A	GATA3	Intron	0.014	3	0	43.4	1.14 (1.04, 1.25)	4.8×10^{-3}	43.2	1.11 (1.06, 1.17)	2.4×10^{-5}	1.12 (1.07, 1.17)	4.3×10^{-7}	0.66
rs149206763	chr17:4618946	A/G	[ALOX15]	Intergenic	-0.014	2	0	13.8	0.87 (0.76, 1.00)	0.05	14.0	0.84 (0.78, 0.90)	3.6×10^{-6}	0.85 (0.79, 0.90)	5.4×10^{-7}	0.64
rs7302200	chr12:56055651	A/G	12q13	Intergenic	0.019	44	0	34.4	1.02 (0.93, 1.12)	0.65	33.0	1.15 (1.09, 1.21)	8.7×10^{-8}	1.12 (1.07, 1.17)	9.4×10^{-7}	0.028
rs556670	chr1:65026231	A/C	1p31	Intergenic	0.018	37	0	11.4	1.24 (1.08, 1.42)	1.9×10^{-3}	13.6	1.13 (1.05, 1.22)	1.2×10^{-3}	1.15 (1.08, 1.23)	1.5×10^{-5}	0.24
rs10995240	chr10:62628871	C/G	ZNF365	Intron	-0.020	36	0	33.5	0.96 (0.87, 1.06)	0.39	36.4	0.89 (0.84, 0.94)	1.1×10^{-5}	0.90 (0.86, 0.95)	1.9×10^{-5}	0.19
rs11242709	chr6:209159	T/C	6p25	Intergenic	-0.024	1	0	20.2	0.96 (0.85, 1.08)	0.46	20.2	0.86 (0.81, 0.92)	9.3×10^{-6}	0.88 (0.83, 0.94)	2.4×10^{-5}	0.13
rs146597587	chr9:6255967	C/G	[IL33]	Splice acceptor	-0.184	15	0	0.65	0.63 (0.32, 1.24)	0.18	0.44	0.33 (0.19, 0.54)	1.8×10^{-5}	0.42 (0.28, 0.63)	2.5×10^{-5}	0.12
rs8129030	chr21:35340290	T/A	21q22	Intergenic	-0.014	6	0	39.2	0.95 (0.86, 1.04)	0.29	36.0	0.90 (0.85, 0.95)	4.9×10^{-5}	0.91 (0.87, 0.95)	4.8×10^{-5}	0.30
rs9532434	chr13:39781776	T/C	COG6	Intron	-0.008	30	0	31.9	0.88 (0.79, 0.97)	0.01	35.1	0.92 (0.87, 0.97)	1.4×10^{-3}	0.91 (0.87, 0.95)	5.5×10^{-5}	0.44
rs12935413	chr16:11116590	A/G	CLEC16A	Intron	-0.028	26	0	31.5	0.99 (0.90, 1.09)	0.88	35.3	0.89 (0.84, 0.93)	8.6×10^{-6}	0.91 (0.87, 0.95)	8.0×10^{-5}	0.039
rs175705	chr14:75509305	C/G	14q24	Intergenic	-0.023	9	0	25.8	0.99 (0.89, 1.09)	0.82	28.3	0.88 (0.84, 0.94)	1.9×10^{-5}	0.91 (0.86, 0.95)	1.2×10^{-4}	0.06
rs1961220	chr5:35844023	A/G	5p13	Intergenic	-0.018	13	0	39.7	0.90 (0.82, 0.99)	0.03	33.3	0.92 (0.87, 0.97)	1.5×10^{-3}	0.91 (0.87, 0.96)	1.4×10^{-4}	0.75

Supplementary Table 18: Conditional association results for association the three variants at the *IL33* locus with NP in Iceland, the UK, and the combined datasets. Results are presented for each variant when conditioned on the other two.

Marker	Position	Allele (min/maj)	Iceland			UK			Combined		
			MAF (%)	OR (95% CI)	<i>P</i>	MAF (%)	OR (95% CI)	<i>P</i>	OR (95% CI)	<i>P</i>	<i>P_{het}</i>
rs1888909	chr9:6197392	T/C	24.5	1.31 (1.18, 1.46)	1.1x10 ⁻⁶	26.0	1.33 (1.25, 1.41)	1.9x10 ⁻²⁰	1.33 (1.26, 1.40)	1.2x10 ⁻²⁵	0.81
rs78757963	chr9:6282511	A/G	3.69	0.96 (0.78, 1.18)	0.70	3.38	1.23 (1.09, 1.38)	5.4x10 ⁻⁴	1.16 (1.05, 1.28)	4.8x10 ⁻³	0.04
rs146597587	chr9:6255967	C/G	0.65	0.69 (0.38, 1.26)	0.29	0.44	0.52 (0.36, 0.75)	5.4x10 ⁻⁴	0.56 (0.41, 0.77)	3.5x10 ⁻⁴	0.43

MAF = minor allele frequency; OR = odds ratio; CI = confidence interval

Supplementary Table 19: Linkage information for the three variants at the *IL33* locus. Values are calculated for the Icelandic dataset. The value of r^2 is presented first, and the D' value is subsequently shown in parentheses.

	rs1888909	rs78757963	rs146597587
rs1888909	-	0.12 (-0.98)	2.2×10^{-3} (0.98)
rs78757963	0.12 (-0.98)	-	2.8×10^{-4} (-1.00)
rs146597587	2.2×10^{-3} (0.98)	2.8×10^{-4} (-1.00)	-

Supplementary Table 20: Relatedness to directly genotyped individuals of those individuals whose genotypes were calculated based on genotypes of close relatives. In this case, 1st degree relatives are defined as parents, children, or siblings, whereas 2nd degree relatives are defined as grandparents and grandchildren. Individuals are listed according to how many 1st and 2nd degree relatives of theirs are directly genotyped (WGS or chip-typed). Further detail can be found in Supplementary Data 2.

		NP Cases	NP Controls	CRS Cases	CRS Controls
Genotyped first-degree relatives (N)	0	13	19,506	92	34,208
	1	179	87,226	678	113,649
	2	62	23,581	148	25,126
	3	64	17,336	137	18,000
	4	27	10,192	61	10,687
	5+	42	13,562	92	14,225
Genotyped second-degree relatives (N)	0	303	110,329	835	127,602
	1	52	30,651	188	41,243
	2	29	21,687	140	33,722
	3	3	6,621	34	10,680
	4	0	1,387	5	1,670
	5+	0	728	6	978

Supplementary Table 21: The LD score regression intercepts used as genomic inflation factors in the Icelandic and UK GWAS analyses of NP and CRS. The studies were adjusted individually before being combined in the meta-analyses, which were not subsequently readjusted.

Phenotype	Dataset	Correction Factor
Nasal polyps	Iceland	1.063325270
	UK	1.028747499
Chronic rhinosinusitis	Iceland	1.108944145
	UK	1.007399653
Asthma	Iceland	1.325291972
	UK	1.060497953
Eosinophil count	Iceland	0.610445567

Supplementary Table 22: Associations of variants at the *IL18R1/IL1RL1* locus with NP in Iceland, the UK, and the combined dataset. The signal is represented by the intron variant rs6543124 in *IL18R1*, which is highly correlated to five missense variants in *IL1RL1*.

Marker	Position	Allele (min/maj)	Iceland (1,175 cases; 309,305 controls)			UK (3,191 cases; 405,376 controls)			Combined (N=4,366 cases; 714,681 controls)		Relation to rs6543124		
			MAF (%)	OR (95% CI)	P	MAF (%)	OR (95% CI)	P	OR (95% CI)	P	Distance (BP)	r ²	D'
rs6543124	chr2:102370999	A/T	38.9	0.91 (0.83, 1.00)	0.055	8.1	0.85 (0.81, 0.90)	2.2x10 ⁻⁹	0.87 (0.83, 0.91)	7.4x10 ⁻¹⁰	-	-	-
rs4988956	chr2:102351547	A/G	39.0	0.92 (0.81, 0.90)	0.073	38.3	0.86 (0.81, 0.90)	6.6x10 ⁻⁹	0.87 (0.83, 0.91)	2.8x10 ⁻⁹	19,452	0.97	0.99
rs10192036	chr2:102351751	A/C	39.0	0.92 (0.81, 0.90)	0.074	38.3	0.86 (0.81, 0.90)	7.3x10 ⁻⁹	0.87 (0.83, 0.91)	3.2x10 ⁻⁹	19,248	0.97	0.99
rs10204137	chr2:102351752	G/A	39.0	0.92 (0.81, 0.90)	0.073	38.3	0.86 (0.81, 0.90)	5.9x10 ⁻⁹	0.87 (0.83, 0.91)	2.6x10 ⁻⁹	19,247	0.97	0.99
rs10192157	chr2:102351896	T/C	39.0	0.92 (0.81, 0.90)	0.074	38.3	0.86 (0.81, 0.90)	7.0x10 ⁻⁹	0.87 (0.83, 0.91)	3.0x10 ⁻⁹	19,103	0.97	0.99
rs10206753	chr2:102351902	C/T	39.0	0.92 (0.81, 0.90)	0.073	38.3	0.86 (0.81, 0.90)	6.5x10 ⁻⁹	0.87 (0.83, 0.91)	2.8x10 ⁻⁹	19,097	0.97	0.99

MAF = minor allele frequency; OR = odds ratio; CI = confidence interval

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. [For final submission](#): please carefully check your responses for accuracy; you will not be able to make changes later.

► Experimental design

1. Sample size

Describe how sample size was determined.

NP and CRS cases were identified through ICD10 codes from the 372,176 individuals from Iceland and 408,567 from the UK with available genotypes. This yielded 4,366 NP and 5,608 CRS cases. Details are provided in the Methods. No sample size calculation was performed as all available individuals were used for the study.

2. Data exclusions

Describe any data exclusions.

No data were excluded.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

We performed GWAS studies in two independent populations and combined the results. Results are presented for the populations independently and combined, and success of replication is reported on per-variant basis, both as individual p-values and as heterogeneity of effects between groups.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No randomization was used and is not relevant within the context of a genome-wide association study.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

This is not relevant, we are performing a genome-wide association study.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| ☐ | ☒ The <u>exact sample size</u> (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☒ | ☐ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ A statement indicating how many times each experiment was replicated |
| ☐ | ☒ The statistical test(s) used and whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☐ | ☒ Test values indicating whether an effect is present
<i>Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.</i> |
| ☐ | ☒ A clear description of statistics including <u>central tendency</u> (e.g. median, mean) and <u>variation</u> (e.g. standard deviation, interquartile range) |
| ☐ | ☒ Clearly defined error bars in <u>all</u> relevant figure captions (with explicit mention of central tendency and variation) |

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

Software used to analyze data was extensively described in a recent publication (PMC5607473) that is cited in the Methods of our manuscript. The following text appears in the Code Availability section of our Manuscript:

"We used publicly available software (URLs listed below) in conjunction with the above described algorithms in the sequencing processing pipeline (Whole-genome sequencing, Association testing, RNA-seq mapping and analysis):
BWA 0.7.10 mem, <https://github.com/lh3/bwa>
GenomeAnalysisTKLite 2.3.9, <https://github.com/broadgsa/gatk/>
Picard tools 1.117, <https://broadinstitute.github.io/picard/>
SAMtools 1.3, <http://samtools.github.io/>
Bedtools v2.25.0-76-g5e7c696z, <https://github.com/arq5x/bedtools2/>
Variant Effect Predictor <https://github.com/Ensembl/ensembl-vep>
BOLT-LMM <https://data.broadinstitute.org/alkesgroup/BOLT-LMM/downloads/> "

Variants were imputed based on the IMPUTE HMM model.
We used R extensively to analyze data and create plots.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This is described in detail in the Methods:

Iceland:

"This study is based on whole-genome sequence data from 15,220 Icelanders participating in various disease projects at deCODE genetics, sequenced to an average genome-wide coverage of 34X. In addition, 151,677 Icelanders have been genotyped using Illumina SNP chips, and genotype probabilities for untyped relatives has been calculated based on Icelandic genealogy. The relations of these individuals to those who have been directly genotyped are outlined in Supplementary Table 17 and Supplementary Data 6. The process used to whole-genome sequence the Icelandic population, and the subsequent imputation from which the data for this analysis were generated has been extensively described in recent publications"

...

"The total number of individuals used in the study was 372,176. Of these, 50.4% were male. The mean year of birth was 1969 (SD = 28.4 years). To identify NP and CRS cases, we searched for patients with International Classification of Diseases (ICD10) codes J33, representing NP, and J32, representing CRS, at Landspítali—The National University Hospital of Iceland in Reykjavik (LUH), and the Icelandic Medical Center in Mjódd (Laeknasetríð). The records spanned from 1986-2016. A total of 1,175 NP cases were included in the association analysis; 86 of these were whole-genome sequenced, 788 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 387 were imputed on the basis of information from genotyped close relatives. Controls comprised individuals recruited through different genetic research projects at deCODE. Individuals in the NP cohort were excluded from the control group. Of the controls, 13,941 were whole-genome sequenced, 137,902 were genotyped by chip, and 171,403 were imputed on the basis of the genotypes of close relatives. The total number of NP controls was 309,305. A total of 3,188 CRS cases were included in the association analysis; 251 of these were whole-genome sequenced, 1,980 were genotyped using various Illumina chips and imputed using long-range phased haplotypes, and genotype probabilities for 1208 were imputed on the basis of information from genotyped close relatives. Controls were selected in the same way as for NP. Of the controls, 13,957 were whole-genome sequenced, 138,044 were genotyped by chip, and 215,895 were imputed on the basis of the genotypes of close relatives. The total number of CRS controls was 353,939."

UK:

"The UK Biobank project is a large prospective cohort study of ~500,000 individuals from across the United Kingdom, aged between 40-69 at recruitment⁴¹. Genotyping was performed using a custom-made Affymetrix chip, UK BiLEVE Axiom⁴² in the first 50,000 participants, and with Affymetrix UK Biobank Axiom array in the remaining participants⁴³; 95% of the signals are on both chips. Imputation was performed by Wellcome Trust Centre for Human Genetics using a combination of 1000Genomes phase 3⁴⁴, UK10K⁴⁵, and HRC⁴⁶ reference panels, for up to 92,693,895 SNPs⁴⁷, although only 37,123,566 well-imputed variants from the HRC panel were used in the current study. The populations defined from UK Biobank data consisted of 3,191 NP cases and 405,376 controls, as well as 2,420 CRS cases and 406,147 controls. All individuals were of European ancestry (self-reported white British with similar genetic ancestry based on principal component analysis and with consistent reported and genetically determined gender) recruited between 2006 and 2010. Related individuals were used in the study. The NP and CRS diagnoses was ascertained based on ICD diagnoses obtained from primary or secondary diagnostic codes recorded across all participants' hospital episodes. Diagnoses were used from ICD-10 codes J32 and J33. Self-reported diagnoses were excluded.