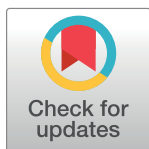


RESEARCH ARTICLE

# A rare *IL33* loss-of-function mutation reduces blood eosinophil counts and protects from asthma

Dirk Smith<sup>1☯<sup>aa</sup></sup>, Hannes Helgason<sup>2,3☯</sup>, Patrick Sulem<sup>2</sup>, Unnur Steina Bjornsdottir<sup>4</sup>, Ai Ching Lim<sup>1</sup>, Gardar Sveinbjornsson<sup>2</sup>, Haruki Hasegawa<sup>1</sup>, Michael Brown<sup>1<sup>ab</sup></sup>, Randal R. Ketchem<sup>1<sup>ac</sup></sup>, Monica Gavala<sup>1</sup>, Logan Garrett<sup>1<sup>ad</sup></sup>, Adalbjorg Jonasdottir<sup>2</sup>, Aslaug Jonasdottir<sup>2</sup>, Asgeir Sigurdsson<sup>2</sup>, Olafur T. Magnusson<sup>2</sup>, Gudmundur I. Eyjolfsson<sup>5</sup>, Isleifur Olafsson<sup>6</sup>, Pall Torfi Onundarson<sup>7,8</sup>, Olof Sigurdardottir<sup>8,9</sup>, David Gislason<sup>10</sup>, Thorarinn Gislason<sup>8,10</sup>, Bjorn Runar Ludviksson<sup>8,11</sup>, Dora Ludviksdottir<sup>8,10</sup>, H. Marike Boezen<sup>12,13</sup>, Andrea Heinzmann<sup>14</sup>, Marcus Krueger<sup>14</sup>, Celeste Porsbjerg<sup>15</sup>, Tarunveer S. Ahluwalia<sup>16</sup>, Johannes Waage<sup>16</sup>, Vibeke Backer<sup>15</sup>, Klaus A. Deichmann<sup>14</sup>, Gerard H. Koppelman<sup>12,17</sup>, Klaus Bønnelykke<sup>16</sup>, Hans Bisgaard<sup>16</sup>, Gisli Masson<sup>2</sup>, Unnur Thorsteinsdottir<sup>2,8</sup>, Daniel F. Gudbjartsson<sup>2,3</sup>, James A. Johnston<sup>1</sup>, Ingileif Jonsdottir<sup>2,8,11\*</sup>, Kari Stefansson<sup>2,8\*</sup>



**OPEN ACCESS**

**Citation:** Smith D, Helgason H, Sulem P, Bjornsdottir US, Lim AC, Sveinbjornsson G, et al. (2017) A rare *IL33* loss-of-function mutation reduces blood eosinophil counts and protects from asthma. *PLoS Genet* 13(3): e1006659. <https://doi.org/10.1371/journal.pgen.1006659>

**Editor:** Tuuli Lappalainen, New York Genome Center & Columbia University, UNITED STATES

**Received:** October 27, 2016

**Accepted:** February 26, 2017

**Published:** March 8, 2017

**Copyright:** © 2017 Smith et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The Dutch asthma study was supported by the Netherlands Asthma Foundation grant AF (AF 95.09, AF 98.48, AF 3.2.02.51 and AF 3.2.07.015) and a grant from the University Medical Center Groningen. The Vlagtwedde-Vlaardingen cohort study was supported by the Ministry of Health and Environmental Hygiene of the Netherlands and the Netherlands Asthma Fund

**1** Amgen Inc., Discovery Research, South San Francisco, California, United States of America, **2** deCODE genetics / Amgen Inc., Reykjavik, Iceland, **3** School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland, **4** Department of Medicine, Landspítali, The National University Hospital of Iceland, Reykjavik, Iceland, **5** The Laboratory in Mjodd, RAM, Reykjavik, Iceland, **6** Department of Clinical Biochemistry, Landspítali, The National University Hospital of Iceland, Reykjavik, Iceland, **7** Laboratory Hematology, Landspítali, The National University Hospital of Iceland, Reykjavik, Iceland, **8** Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland, **9** Department of Clinical Biochemistry, Akureyri Hospital, Akureyri, Iceland, **10** Department of Respiratory Medicine and Sleep, Landspítali, The National University Hospital of Iceland, Reykjavik, Iceland, **11** Department of Immunology, Landspítali, The National University Hospital of Iceland, Reykjavik, Iceland, **12** GRIAC research institute, Groningen, The Netherlands, **13** University Medical Center Groningen, University of Groningen, Department of Epidemiology, Groningen, The Netherlands, **14** Center for Pediatrics, Department of General Pediatrics, Adolescent Medicine and Neonatology, Medical Center – University of Freiburg, Faculty of Medicine, Freiburg, Germany, **15** Department of Respiratory Medicine, Bispebjerg University Hospital, Copenhagen University, Copenhagen, Denmark, **16** COPSAC (Copenhagen Prospective Studies on Asthma in Childhood), Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark, **17** University Medical Center Groningen, University of Groningen, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children’s Hospital, Groningen, The Netherlands

☯ These authors contributed equally to this work.

<sup>aa</sup> Current Address: Insight BioConsulting, LLC, Bainbridge Island, Washington, United States of America,

<sup>ab</sup> Current Address: Adimab, Lebanon, New Hampshire, United States of America

<sup>ac</sup> Current Address: Just Biotherapeutics, Seattle, Washington, United States of America

<sup>ad</sup> Current Address: Absci, Portland, Oregon, United States of America

\* [kstefans@decode.is](mailto:kstefans@decode.is) (KS); [ingileif.jonsdottir@decode.is](mailto:ingileif.jonsdottir@decode.is) (IJ)

## Abstract

IL-33 is a tissue-derived cytokine that induces and amplifies eosinophilic inflammation and has emerged as a promising new drug target for asthma and allergic disease. Common variants at *IL33* and *IL1RL1*, encoding the IL-33 receptor ST2, associate with eosinophil counts and asthma. Through whole-genome sequencing and imputation into the Icelandic population, we found a rare variant in *IL33* (NM\_001199640:exon7:c.487-1G>C (rs146597587-C), allele frequency = 0.65%) that disrupts a canonical splice acceptor site before the last coding exon. It is also found at low frequency in European populations. rs146597587-C

(grant 187) and the Netherlands Asthma Fund grant no. 3.2.02.51, the Stichting Astma Bestrijding, BBMRI-NL (Complementation project), and the European Respiratory Society COPD research award 2011 to H.M. Boezen. COPSAC is funded by private and public research funds all listed on [www.copsac.com](http://www.copsac.com). The Lundbeck Foundation; Danish State Budget; Danish Council for Strategic Research; Danish Council for Independent Research and The Capital Region Research Foundation have provided core support for COPSAC. The Denmark-1 study has been supported with an unrestricted grant from AstraZeneca and ALK-Abello, Denmark. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** I have read the journal's policy and the authors of this manuscript have the following competing interests: Authors affiliated with deCODE genetics/Amgen Inc or Amgen Inc. declare competing financial interests as employees. DS owns some Amgen stock. RRK has outstanding patent applications with Amgen but has signed away all rights to Amgen. HHaruki owns stocks/shares and has paid employment/consultancy. GHK and HMB have received for the Dutch asthma study support from Netherlands Asthma Foundation grant AF (AF 95.09, AF 98.48, AF 3.2.02.51 and AF 3.2.07.015) and a grant from the University Medical Center Groningen. HMB has received for the The Vlagtweede-Vlaardingeng cohort study support from the Ministry of Health and Environmental Hygiene of the Netherlands and the Netherlands Asthma Fund (grant 187) and the Netherlands Asthma Fund grant no. 3.2.02.51, the Stichting Astma Bestrijding, BBMRI-NL (Complementation project), and the European Respiratory Society COPD research award 2011. The COPSAC study (HB, KB, TSA, JW) receives core funding from The Lundbeck Foundation; Danish State Budget; Danish Council for Strategic Research; Danish Council for Independent Research and The Capital Region Research Foundation. In addition COPSAC is funded by private and public research funds all listed on [www.copsac.com](http://www.copsac.com). Denmark-1 study (VB, CP) has been supported with an unrestricted grant from AstraZeneca and ALK-Abello, Denmark.

associates with lower eosinophil counts ( $\beta = -0.21$  SD,  $P = 2.5 \times 10^{-16}$ ,  $N = 103,104$ ), and reduced risk of asthma in Europeans (OR = 0.47; 95%CI: 0.32, 0.70,  $P = 1.8 \times 10^{-4}$ ,  $N$  cases = 6,465,  $N$  controls = 302,977). Heterozygotes have about 40% lower total *IL33* mRNA expression than non-carriers and allele-specific analysis based on RNA sequencing and phased genotypes shows that only 20% of the total expression is from the mutated chromosome. In half of those transcripts the mutation causes retention of the last intron, predicted to result in a premature stop codon that leads to truncation of 66 amino acids. The truncated IL-33 has normal intracellular localization but neither binds IL-33R/ST2 nor activates ST2-expressing cells. Together these data demonstrate that rs146597587-C is a loss of function mutation and support the hypothesis that IL-33 haploinsufficiency protects against asthma.

## Author summary

Only a few genes have been found to play a role in asthma. These include the genes *IL33* and *IL1RL1*, and sequence variants in the human genome close to these genes were initially found to affect the number of eosinophils, cells that play a role in inflammation of the airways in asthma. Based on this knowledge, we decided to use high resolution sequencing technology to search for variants in these genes that cause changes in structure and function of the proteins they encode. We found a rare (0.65%) sequence variant in the *IL33* gene, that causes less production of the IL33 protein and some of the protein formed lacks the capacity to bind to its receptor on cells and promote inflammation. This rare mutation causes reduced number of eosinophils in blood and protects against asthma. These results suggest that drugs that could interfere with the inflammatory activity of the IL33 protein may be beneficial for treatment of asthma.

## Introduction

Asthma is characterized by airflow obstruction, airway hyper-responsiveness and airway inflammation, that promotes mucus obstruction. The inflammatory cytokine IL-33 is widely expressed[1, 2] and abundantly in the bronchial epithelium. IL-33 resides in the nucleus due to chromatin-binding motifs[3] but is released after exposure to e.g. viruses or allergens. IL-33 binds to its receptor ST2[4] (also called IL1RL1, IL-33R) and activates eosinophils and other immune cells promoting inflammation[5], particularly in the lung[6, 7]. IL-33 also binds to a soluble isoform IL1RL1-a/sST2, which is thought to act as a decoy receptor and ameliorates airway inflammation[8].

Through GWAS, we previously discovered common sequence variants in *IL1RL1* and *IL33* that associate strongly with blood eosinophil counts and risk of asthma, consistent with the link between eosinophilic inflammation and asthma [9]. The association between *IL1RL1* and *IL33* variants and asthma risk is well established [10] and has been replicated in ethnically diverse populations[11, 12], and in severe forms of adult asthma[11] and in particular with early childhood[13] asthma with exacerbations.

The critical role of eosinophilic airway inflammation in asthma[14] together with robust association of eosinophil counts and asthma with common variants at the *IL33* and *IL1RL1* loci prompted us to search for novel sequence variants at these loci affecting eosinophil counts using high-coverage sequencing [15]. The variants were identified through sequencing of

8,453 Icelanders and imputed into 150,656 chip-typed Icelanders[15]. In light of the role of eosinophils in pathogenesis of asthma and the previously established association of common variants at the *IL33* and *IL1RL1* loci with eosinophil counts and asthma, individual sequence variants at these loci found to associate significantly with eosinophil counts were further assessed for their effects on asthma.

## Results

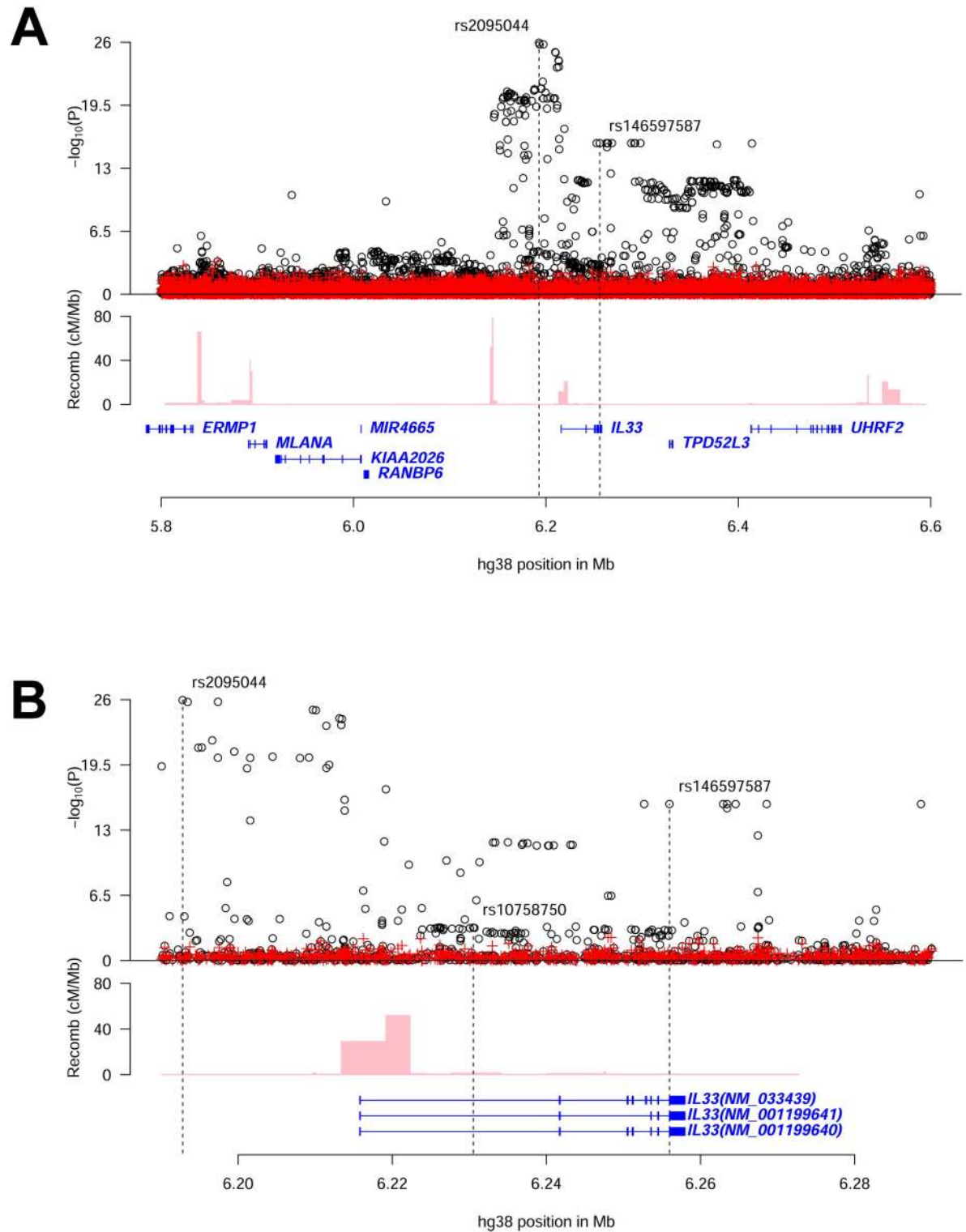
We focus on the association of sequence variants in two 800kb regions centered on *IL33* (chr9:5.8–6.6Mb(hg38)) and *IL1RL1* (chr2:101.9–102.7Mb(hg38)) with eosinophil counts in 103,104 Icelanders[16]. Sequence variants were weighted according to their prior probability of affecting gene function by applying different thresholds for genome-wide (gw) significance that depend on the variant class[17].

Stepwise conditional analysis at the *IL33* locus revealed two significant uncorrelated (pairwise  $r^2 < 0.02$ ) variants (Fig 1 and S1 Fig, Table 1 and S1 Table). One association is novel: rs146597587-C a rare variant (allelic frequency (AF) = 0.65%) that is predicted to disrupt a canonical splice acceptor site at the beginning of the last exon of *IL33*[18, 19] (NM\_001199640: exon7:c.487-1G>C,  $\beta_{\text{adj}} = -0.21$  SD;  $P_{\text{adj}} = 2.5 \times 10^{-16}$ ) (Table 2). rs146597587-C associates with lower eosinophil counts and is not correlated with previously reported variants at the *IL33* locus (S2 Table). The second variant, rs2095044-T (AF = 24.5%), upstream of *IL33*, is associated with increased eosinophil count ( $\beta_{\text{adj}} = 0.051$  SD;  $P_{\text{adj}} = 3.6 \times 10^{-27}$ ) and is highly correlated with rs2381416 ( $r^2 = 0.94$ ), originally described to associate with eosinophil counts and asthma [9] (Table 1, S3 and S4 Tables). A potential secondary signal was noted, rs10758750-G (AF = 27.7%), an intronic variant in *IL33* ( $\beta_{\text{adj}} = -0.023$  SD;  $P_{\text{adj}} = 5.7 \times 10^{-7}$ ) (Table 1) (passing a significance threshold of  $2.8 \times 10^{-6}$ , corresponding to a Bonferroni correction for the number of variants tested (17,935) at the two loci[20]), uncorrelated (pairwise  $r^2 < 0.02$ ) to the novel variant reported here (Fig 1 and S1 Fig, Table 1 and S1 Table) and to the previously reported variants at the *IL33* locus (S2 Table).

Among the two significant variants, the splice acceptor mutation rs146597587-C has the largest effect. The splice acceptor mutation is present in both Europeans (AF = 0.35%) and South-Asians (AF = 0.10%) in the Exome Aggregation Consortium (ExAC) database where it is >10 times more frequent than any other predicted loss-of-function variant in *IL33* (S5 Table, URLs). In total we tested six predicted missense, splice region or loss-of-function variants in *IL33* (S6 Table) and only the association with the splice acceptor mutation rs146597587-C is significant ( $P > 0.2$  for the others). None of the twelve variants highly correlated ( $r^2 > 0.8$ ) with the splice acceptor mutation were coding, making it most likely to be responsible for the effect (S7 Table). Neither of the two other associating variants, rs2095044 nor rs10758750, were correlated with a coding variant (S4 and S8 Tables).

The *IL33* splice acceptor mutation, rs146597587-C was genotyped in 1,370 Dutch samples and its effect on eosinophil counts was replicated ( $\beta = -0.48$  SD,  $P = 0.036$ , AF = 0.69%) (Table 2).

Once the eosinophil count association of the splice acceptor variant was established, we assessed its effect on asthma, based on the prior association of *IL33* and with asthma risk and the role of eosinophils in the pathogenesis of asthma. We assessed the effect of rs146597587-C on asthma in Iceland[9], The Netherlands[21–23], Germany [24] and Denmark[25, 26], including Danish children with severe asthma with at least 2 exacerbations leading to hospitalization between 2 and 6 years of age (COPSAC<sub>exacerbation</sub>)[13]. The *IL33* splice acceptor mutation protects against asthma (OR = 0.47; 95% CI: 0.32–0.70,  $P = 1.8 \times 10^{-4}$ ) (Table 2). We did not observe heterogeneity of the effect in the different sample sets ( $P_{\text{het}} = 0.24$ ,  $I^2 = 26.8$ ). Thus,



**Fig 1. Overview of eosinophil counts associations in the region around *IL33*.** (A) shows an 800kb overview centered on *IL33* on chromosome 9 and (B) shows a 100kb overview around the *IL33* gene. Black circles show  $-\log_{10} P$  as a function of hg38 coordinates for associations with eosinophil counts and red crosses correspond to eosinophil counts associations after adjusting for the three variants rs2095044, rs146597587 (splice acceptor) and rs10758750 that are indicated by vertical broken lines in (B). Genes are shown in blue and recombination rates are reported in cM/Mb. (See S1 Fig for intermediate results from stepwise regression.)

<https://doi.org/10.1371/journal.pgen.1006659.g001>

**Table 1. Association of common sequence variants in *IL33* and *IL1RL1* with eosinophil counts and asthma in Iceland.**

Marker	Pos. (hg38)	A1	Freq. A1 <sup>b</sup> [%]	gene context	Eosinophil counts (N = 103,104)				Asthma <sup>a</sup> (N <sub>affected</sub> = 3,512, N <sub>controls</sub> = 298,026)			
					unadjusted		adjusted		unadjusted		adjusted	
					B <sup>c</sup> (95%CI)	P	β <sup>c</sup> (95%CI)	P	OR <sup>d</sup> (95% CI)	P	OR <sup>d</sup> (95% CI)	P
<b><i>IL33</i> locus:</b>												
rs2095044	chr9: 6,192,796	T	24.5	intergenic	0.051 (0.04,0.06)	1.1×10 <sup>-26</sup>	0.051 (0.04,0.06)	3.6×10 <sup>-27</sup>	1.12 (1.04,1.20)	0.0018	1.10 (1.03,1.18)	0.0048
rs10758750 <sup>e</sup>	chr9: 6,230,513	G	27.7	Intronic ( <i>IL33</i> )	-0.016 (-0.03,-0.01)	5.1×10 <sup>-4</sup>	-0.023 (-0.03,-0.01)	5.7×10 <sup>-7</sup>	1.06 (0.98,1.14)	0.12	1.04 (0.97,1.12)	0.27
<b><i>IL1RL1</i> locus:</b>												
rs13020553	chr2: 102,315,366	G	41.9	intronic ( <i>IL1RL1</i> )	0.048 (0.04,0.06)	3.5×10 <sup>-31</sup>	0.043 (0.03,0.05)	1.6×10 <sup>-24</sup>	1.06 (0.99,1.13)	0.078	1.04 (0.98,1.10)	0.19
rs6719123	chr2: 102,259,080	G	14.2	intergenic	-0.048 (-0.06,-0.04)	1.3×10 <sup>-16</sup>	-0.037 (-0.05,-0.03)	7.0×10 <sup>-10</sup>	0.92 (0.85,1.00)	0.049	0.92 (0.85,1.01)	0.083

Association results for each variant is presented with and without adjusting for the other variants at its locus in this table. All the variants were well imputed (imputation information = 1.00).

<sup>a</sup>) The asthma sample set includes young adults, 45 years old or younger[9] (Materials and methods).

<sup>b</sup>) Freq. A1: Allelic frequency for allele A1.

<sup>c</sup>) β: Effect in SD with respect to allele A1.

<sup>d</sup>) OR: Odds ratio with respect to allele A1.

<sup>e</sup>) rs10758750 shows suggestive association with eosinophil counts, passing a significance threshold of 2.8×10<sup>-6</sup>, corresponding to a Bonferroni correction for the number of variants tested (17,935) at the two loci[20].

<https://doi.org/10.1371/journal.pgen.1006659.t001>

this rare splice acceptor mutation reduces eosinophil counts and protects against asthma, whereas the minor allele of the common rs2095044-T associates with higher eosinophil counts and increased risk of asthma (Tables 1 and 2).

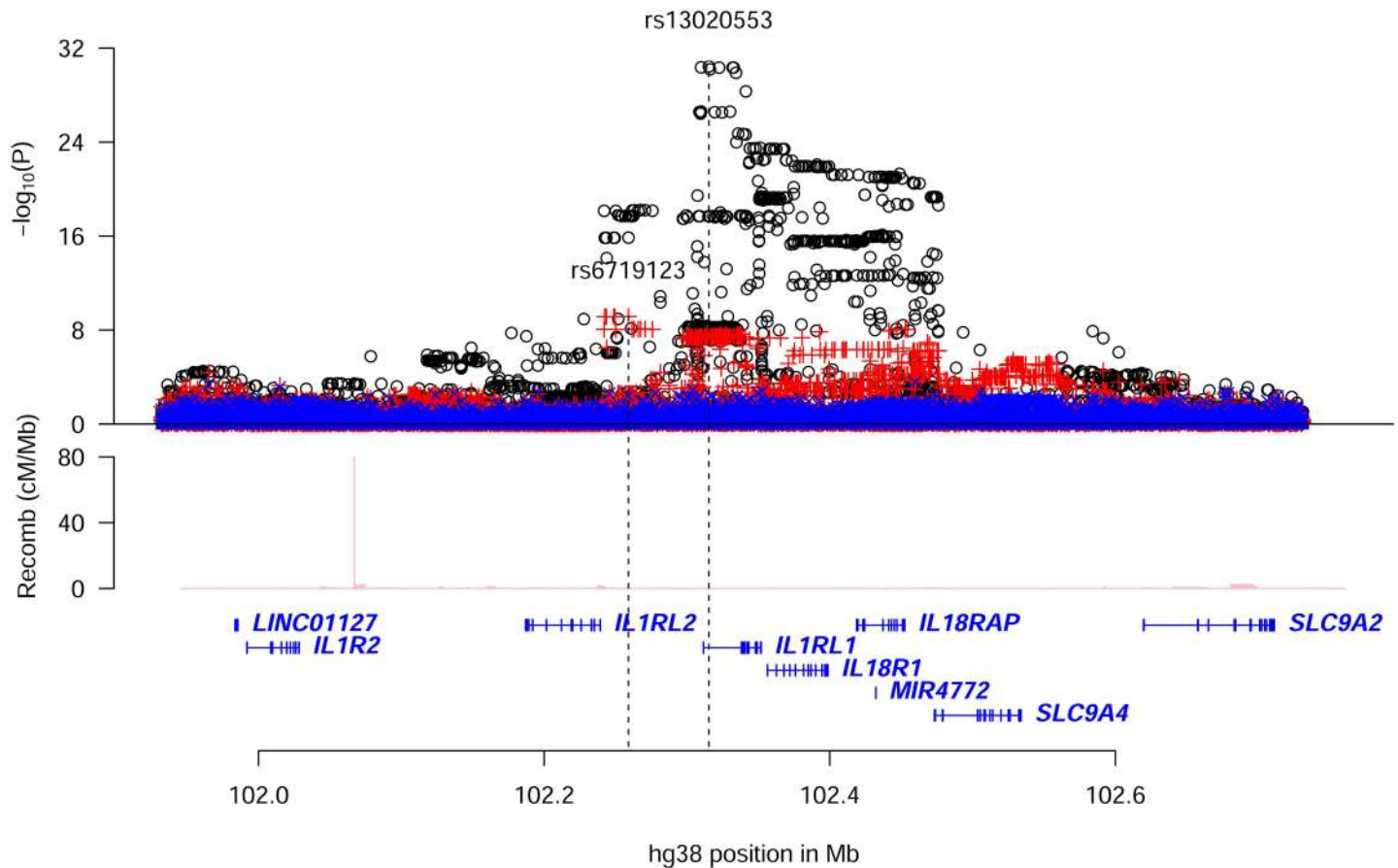
Stepwise conditional analysis of the association with eosinophil counts (N = 103,104) at the *IL1RL1* locus revealed two significant variants, an intronic variant, rs13020553-G

**Table 2. Associations of the *IL33* splice acceptor variant rs146597587[C] with eosinophil counts and asthma in Iceland and abroad.**

Eosinophil counts	AF	β (SD)	(95%CI)	P	N individuals		P <sub>het</sub> , I <sup>2</sup>
Iceland	0.65%	-0.21	(-0.27, -0.16)	2.5×10 <sup>-16</sup>	103,104		
The Netherlands	0.69%	-0.48	(-0.93, -0.03)	0.036	1,370		
<b>Combined</b>		-0.22	(-0.27, -0.17)	5.3×10 <sup>-17</sup>	104,474		0.25, 25.0
Asthma	AF	OR	(95%CI)	P	N cases	N controls	
Iceland:	0.65%	0.36	(0.21, 0.61)	1.2×10 <sup>-4</sup>	3,512	298,026	
The Netherlands	0.53%	1.08	(0.36, 3.21)	0.89	351	2,830	
Germany	0.40%	0.89	(0.14, 5.48)	0.90	284	252	
Denmark-1	0.50%	0.72	(0.29, 1.79)	0.48	1,121	1,004	
Denmark-2 (COPSAC)	0.45%	0.24	(0.06, 0.94)	0.04	1,197	865	
<b>Combined</b>		0.47	(0.32, 0.70)	1.8×10 <sup>-4</sup>	6,465	302,977	0.24, 26.8

Allele frequency (AF) of rs146597587[C], the effect (β (SD)) on eosinophil counts and odds ratio (OR) for asthma and the corresponding P-values are provided, in addition to the number of individuals, or cases and controls tested. All the asthma sample sets include children and/or young adults: Iceland 45 years age or younger[9], The Netherlands younger than 45 years of age[23, 42], Germany 5–18 years of age[24], Denmark-1 14 to 44 years of age[25, 26] and Denmark-2 (COPSAC) children with severe asthma with at least 2 exacerbations leading to hospitalization between 2 and 6 years of age[13] (Materials and methods).

<https://doi.org/10.1371/journal.pgen.1006659.t002>



**Fig 2. Conditional analysis for eosinophil counts associations in the region around *IL1RL1*.** Plot shows an 800kb overview around the *IL1RL1* gene on chromosome 2. Black circles (o) show  $-\log_{10} P$  as a function of hg38 coordinates for unadjusted associations with eosinophil counts; red crosses (+) correspond to eosinophil counts associations after adjusting for the variant rs13020553; blue 'x' symbols correspond to eosinophil counts associations after adjusting for both rs13020553 and rs6719123. The position of the two variants rs13020553 and rs6719123 are indicated by vertical broken lines. Genes are shown in blue and recombination rates are reported in cM/Mb.

<https://doi.org/10.1371/journal.pgen.1006659.g002>

(MAF = 41.9%), increasing eosinophil counts ( $\beta = 0.043$ ,  $P_{\text{adj}} = 1.6 \times 10^{-24}$ ) and an intergenic variant, rs6719123-G (MAF = 14.2%), decreasing eosinophil counts ( $\beta = -0.037$ ,  $P_{\text{adj}} = 7.0 \times 10^{-10}$ ) (Table 1 and S9 Table); no other variant at the locus remained significant after adjusting for these two (Fig 2 and S10 Table). rs13020553 is highly correlated with rs1420101 ( $r^2 = 0.96$ ,  $D' = 1.00$ ), originally reported to associate with eosinophil counts and asthma [9] (S3 and S10 Tables). Despite the high correlation, rs1420101-T (that is on the background of rs13020553-G) does not fully explain the eosinophil association of rs13020553, whereas rs13020553 fully explains the eosinophil association for rs1420101 (S10 Table). Together rs13020553 and rs6719123 explain all the reported effects on eosinophil counts and asthma at the locus (S3 Table); however, the reported variants do not fully explain the signal captured by rs13020553 and rs6719123 (S11 Table).

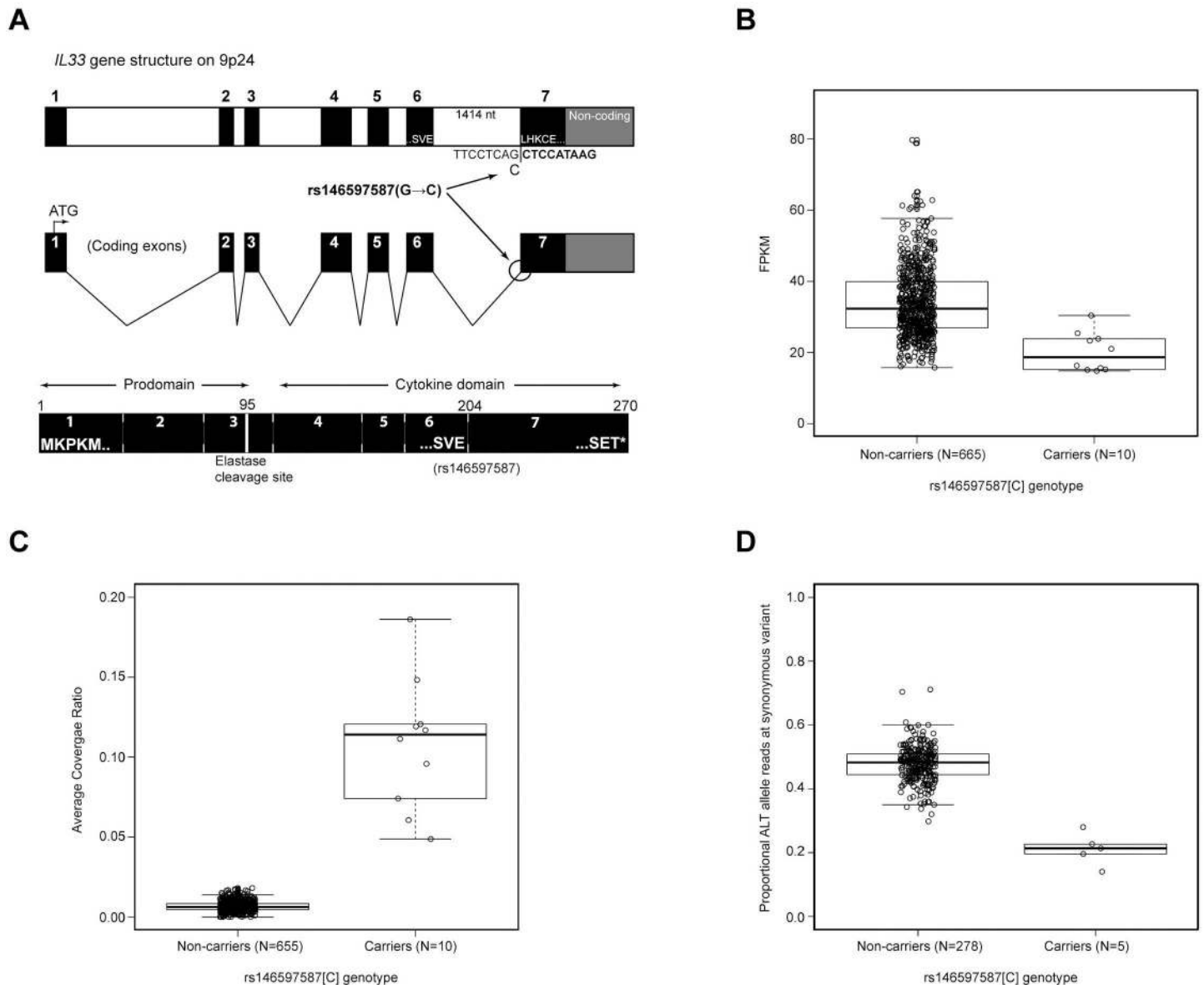
We found and tested 75 coding or splice region variants at the *IL1RL1* locus. Among those not in *IL1RL1*, only one was significant ( $P > 0.001$  for all other coding variants not in *IL1RL1*), rs1420098-C, a splice region variant in *IL18R1* (MAF = 42.5%,  $P = 3.8 \times 10^{-24}$ ,  $\beta = 0.042$ ). The eosinophil counts association of rs1420098 was fully accounted for by rs13020553 ( $r^2 = 0.81$ ,  $P_{\text{adj}} = 0.53$ ), whereas the association of rs13020553 could not be accounted for by rs1420098

( $P_{\text{adj}} = 1.2 \times 10^{-8}$ ,  $\beta_{\text{adj}} = 0.054$ ). We observed two significant coding signals in *IL1RL1* among the sixteen detected and tested (14 missense, one splice region variant and one frameshift variant) (S12 Table). Both coding signals were previously reported to associate with concentration of ST2 in soluble form (sST2) [27]. The first, represented by rs10192157-T (MAF = 39.0%,  $P = 4.2 \times 10^{-20}$ ; HGVS: NP\_057316.3:p.Thr549Ile), corresponds to five perfectly correlated missense variants ( $r^2 = 1.00$ ,  $D' = 1.00$  for all pairs) moderately correlated with rs13020553 ( $r^2 = 0.35$ ); the second rs1041973-A (MAF = 17.7%,  $P = 3.5 \times 10^{-9}$ , HGVS: NP\_057316.3:p.Ala78Glu) is correlated with rs6719123 ( $r^2 = 0.68$ ) (S13 Table). These two eosinophil counts associations are fully accounted for by rs13020553 and rs6719123 (S14 Table). We conclude that the eosinophil counts signals corresponding to rs13020553 and rs6719123 (correlated variants: S15 and S16 Tables) could not be explained by the observed coding variants or reported variants at the locus. Our results are in agreement with the notion that variants at this locus influence eosinophil counts by affecting *IL1RL1*; and IL-33 is known to mediate its biological effects through ST2/*IL1RL1* [4].

The splice acceptor mutation in *IL33* changes AG to AC at the 3' splice junction between the last two exons and is thus predicted to disrupt splicing between them [18, 19]. *IL33* is primarily expressed by stromal cells and is expressed at a relatively low level in hematopoietic tissues. Of the two tissues with mRNA sequencing data available to us *IL33* is essentially absent from whole blood but is expressed in subcutaneous adipose tissue [1] (median RPKM of 12.9 in 350 samples), according to the GTExV6 database [2]. We therefore analyzed the effect of the splice acceptor mutation on the *IL33* transcript quantity and processing using the adipose tissue mRNA sequencing (N = 675) dataset (Materials and methods, and Fig 3). Heterozygote carriers (N = 10) of the splice acceptor mutation showed about 40% lower total *IL33* expression than non-carriers ( $P = 6.8 \times 10^{-6}$ ) (Fig 3). Comparable results were obtained by accessing total IL-33 expression with microarray (S2 Fig). Allele-specific analysis of the RNA sequencing data shows that only 20% of total *IL33* transcripts are from the mutated chromosomes ( $P = 1.3 \times 10^{-4}$ ) indicating that a large fraction of mRNA originating from the mutated chromosomes is likely eliminated through nonsense-mediated decay (NMD) [28]. Analysis of read coverage in the RNA sequencing data, shows retention of the last intron of *IL33* in about half of the RNA generated from the mutated chromosome (~11% of total RNA) in heterozygotes compared to 0.6% of that in non-carriers ( $P = 5.6 \times 10^{-8}$ ). Taken together these data demonstrate that the splice acceptor mutation leads both to elimination of the mutated transcripts and a retention of the last intron, introducing a premature stop codon in about half of the remaining mutated transcripts.

The IL-33 protein generated from the intronic retention transcripts is predicted to lack the last 66 amino acids (out of 270). IL-33 folds as a 12-stranded "β-barrel" structure, prototypical for the IL-1 family. Elimination of the last 66 amino acids removes 5 of the 12 core β-strands and is predicted to disrupt the tertiary structure of the protein. To determine whether IL-33 lacking these residues is functional we expressed recombinant IL-33 lacking the last exon (termination at residue 204, Fig 4). IL-33<sub>1–204</sub> was detectable in transfected mammalian cellular lysates as a 27kD protein, compared to the 35 kD wild type IL-33<sub>1–270</sub> (S3 Fig). IL-33<sub>1–204</sub> was in the nucleus in a manner indistinguishable from IL-33<sub>1–270</sub> (Fig 4), indicating that the truncation does not disrupt nuclear trafficking.

Computational modeling [29] indicates that truncated IL-33 lacks core structural folds and key surface residues predicted to contribute to receptor binding (Fig 4A). To determine whether the truncated form binds ST2 we generated IL-33<sub>95–270</sub> (wild type) and IL-33<sub>95–204</sub> (truncated mutant) recombinant proteins. The N-terminus starting residue at amino acid position 95 was chosen based on natural processing of IL-33 before receptor binding [30] (S4 Fig).

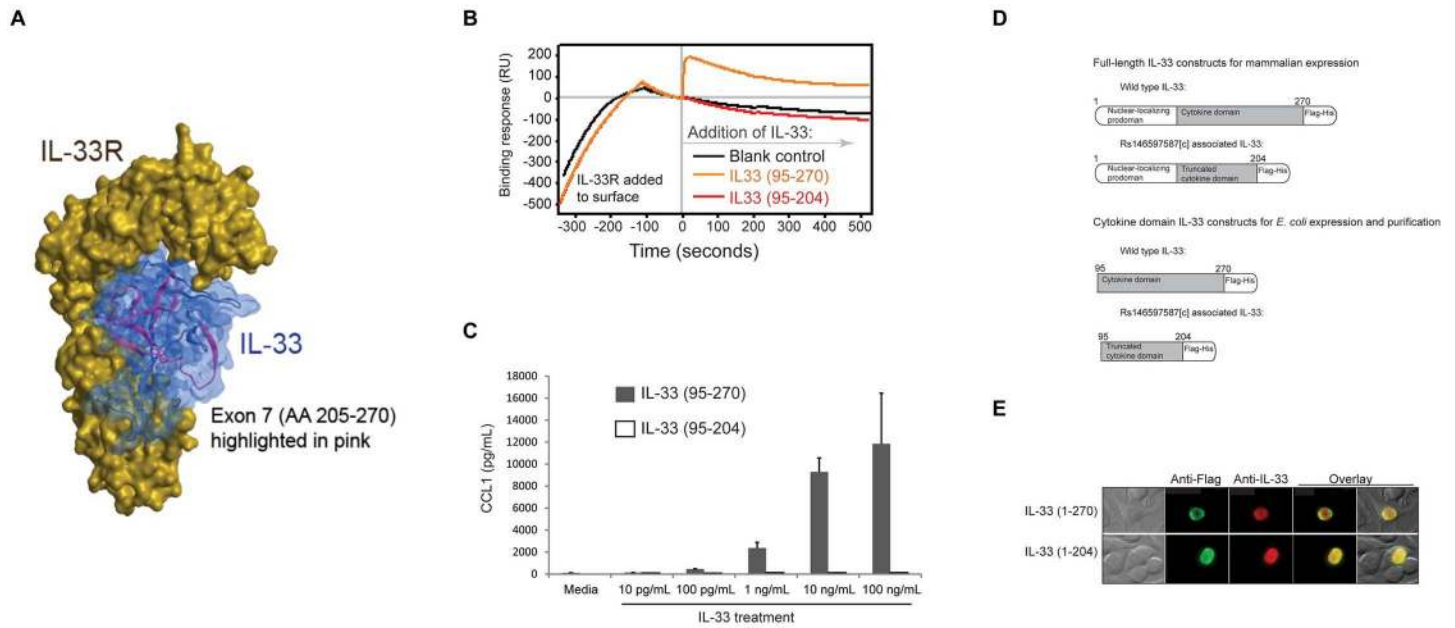


**Fig 3. RNA analysis of *IL33* mRNA expression and carrier status of the splice acceptor variant rs146597587.** (A) Gene structure of *IL33* and location of rs146597587 SNP that changes a canonical splice acceptor site from AG to AC. The mutation lies at the junction of coding exons 6 and 7. Exon 7 begins LHKCE and encodes a significant portion of the cytokine domain. Also, indicated is the location of an elastase cleavage site at amino acid 95 that releases the cytokine domain from the N terminus prodomain and leads to a more active cytokine. (B) Expression in FPKM of *IL33* in adipose tissue for rs146597587 non-carriers (GG, N = 744) and carriers (GC, N = 6) based on RNA-Seq data;  $P(\text{Wilcox}) = 0.0015$ , median non-carriers: 34.1, carriers: 21.1 (i.e. heterozygotes have 38% lower expression than non-carriers). (C) Ratio of RNA-Seq read coverage in last intron of *IL33* versus its last exon (splice acceptor resides at this intron-exon boundary) for rs146597587 non-carriers (GG, N = 738) and carriers (GC, N = 9);  $P(\text{Wilcox}) = 2.5 \times 10^{-7}$ , mean non-carriers: 0.67%, carriers: 9.0%. (D) Proportion of RNA-Seq read with ALT allele for the synonymous variant rs10975519 (MAF = 29%) for rs146597587 non-carriers (GG, N = 309) and carriers (GC, N = 5); a chromosome carrying rs146597587-C carries the ALT allele of rs10975519;  $P(\text{Wilcox}) = 1.4 \times 10^{-4}$ , mean non-carriers: 0.48, non-carriers: 0.20.

<https://doi.org/10.1371/journal.pgen.1006659.g003>

Using surface plasmon resonance we found that IL-33<sub>95-270</sub> bound rapidly to IL-33R/ST2 (Fig 4), consistent with high affinity binding to IL-33R that we previously reported[31], whereas IL-33<sub>95-204</sub> showed no interaction signal, indicating a complete lack of receptor binding. Furthermore, IL-33<sub>95-270</sub> induced a concentration- and IL-33R-dependent CCL1 release





**Fig 4. (A)** Computational model of the IL-33 cytokine domain bound to its receptor. Shown in pink are amino acids 205–270, which are derived from exon 7 and absent in the rs146597587-C-associated truncated form of IL-33. Besides contributing to overall structure, this part of IL-33 also contributes to receptor interaction. **(B)** Binding profile from Biacore surface plasmon resonance measurement of IL-33 binding to IL-33R. Recombinant IL-33R encompassing the extracellular domain was immobilized on the chip surface and then recombinant IL-33 was allowed to flow over the chip (beginning at time 0). Productive receptor binding is indicated by a prolonged increase in resonance following cytokine addition. Wild type IL-33 (95–270) bound receptor, but the mutant IL-33 (95–204) did not. **(C)** Biological response to IL-33. LAD2 human mast cells were treated with the indicated concentrations of either wild type IL-33 (95–207) or mutant IL-33 (95–204) and concentration of CCL1 released into the culture medium was measured by ELISA. **(D)** Schemata of constructs made to evaluate expression of full-length IL-33 proteins in mammalian cells or, for expression and purification of cytokine domain-only proteins in *E. coli*. **(E)** Immunofluorescent staining indicating equivalent nuclear localization of full-length wild type IL-33 (1–270) and the rs146597587[C]-associated form of IL-33 (1–204) in transfected HEK293 human epithelial cells.

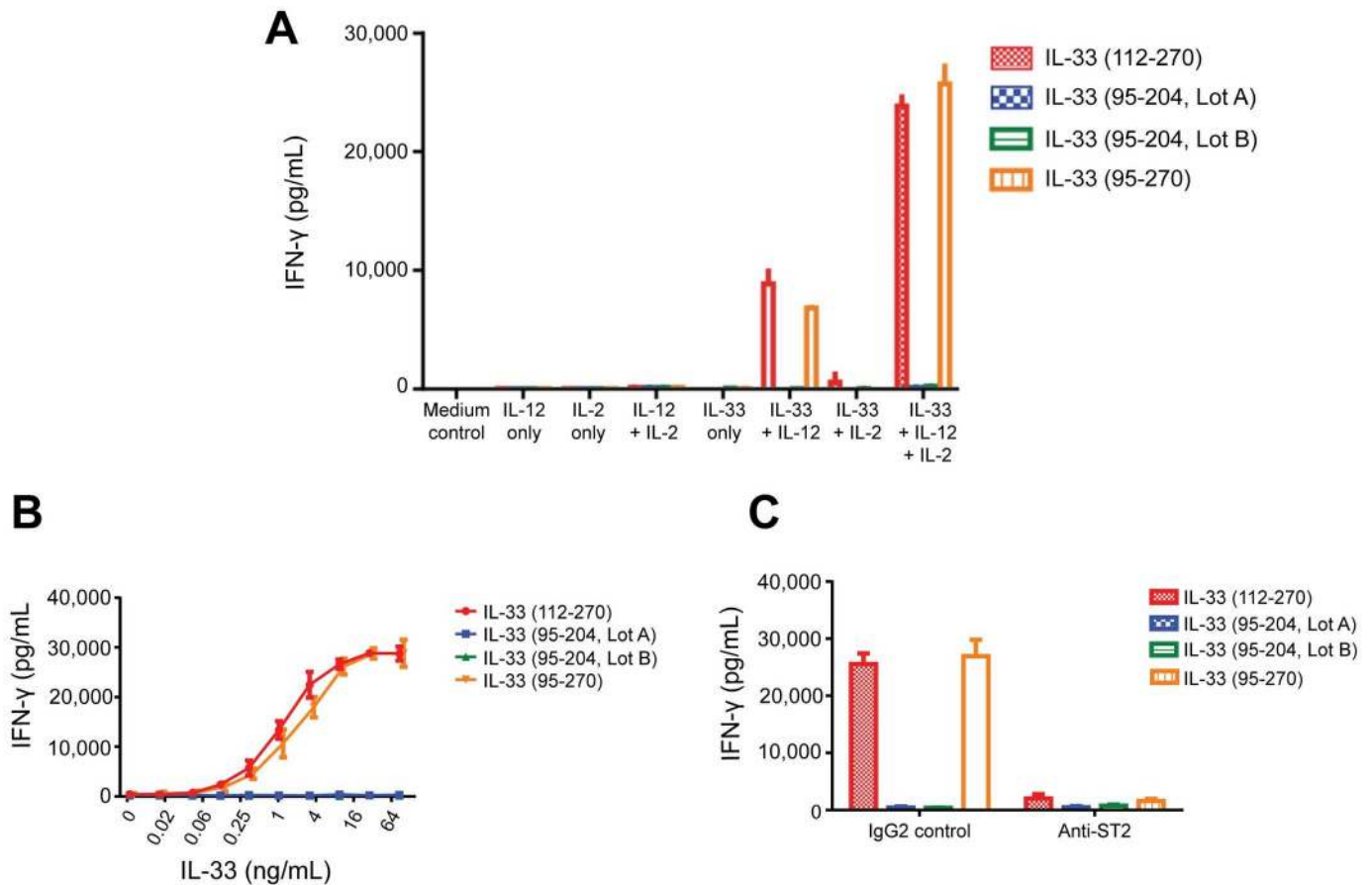
<https://doi.org/10.1371/journal.pgen.1006659.g004>

in human mast cell line expressing IL-33R (Fig 4C and S5 Fig), whereas IL-33<sub>95–204</sub> did not at any concentration tested, consistent with a complete loss of cytokine activity. The mutant IL-33 was also inactive compared to wild type inducing IFN- $\gamma$  release from human CD4<sup>+</sup> T cells (Fig 5).

In *il33* knock-out mice, 30% reduced fertility has been reported [32]. Nine imputed rs146597587-C homozygotes were found in the Icelandic data and they neither show reduced life expectancy nor reduced fertility (S16 Table), indicating that a predicted very low level of IL-33 is compatible with long life and healthy reproduction. Moreover, rs146597587-C homozygosity (recessive model) does not confer risk of any other disease that was tested in the Icelandic data [15, 33], indicating that IL-33 is largely dispensable. Two homozygotes of the splice acceptor mutation are reported in ExAC (S5 Table, URLs). An association of *il33* polymorphism with eosinophil numbers in rats has been reported [34] and we found that deletion of *il33* leads to significant reduction in blood eosinophils in unchallenged mice ( $P = 0.014$  males,  $P = 0.00010$  females) (Fig 6), mimicking our human observations.

## Discussion

Resequencing of 100 genes implicated in asthma only revealed a few coding variants associating with asthma [35]. Since eosinophils are known to play a key role in inflammation of the airway in asthma [12] we used high-coverage sequencing [13] to search for novel sequence variants affecting eosinophil counts at the well established asthma loci, the *IL33* and *IL1RL1* loci,

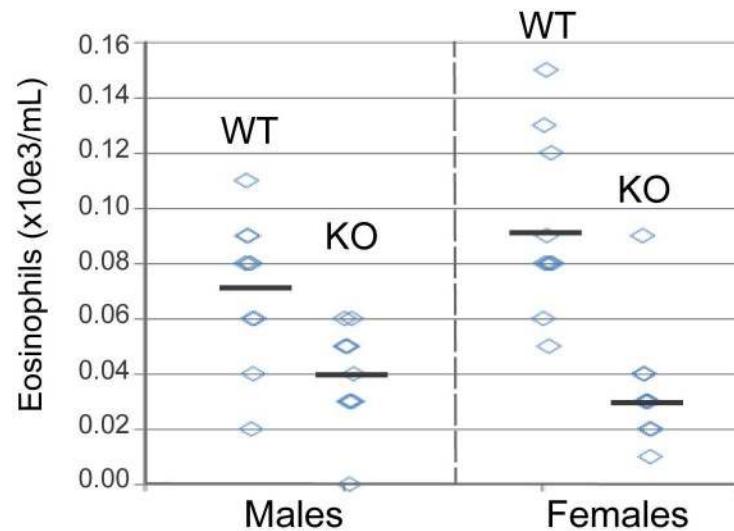


**Fig 5. Activity of IL-33 variants in human CD4 T cell bioassay.** Purified human blood CD4 T cells were incubated in the presence or absence of IL-2 and IL-12 for 72 hours and IFN- $\gamma$  was measured in the cell-free supernatants (as described in online methods). Both IL-33 (95–270) (Amgen) and IL-33(112–270) (R&D Systems) were used as positive controls and two independent lots of IL-33(95–204) were tested. **(A)** IL-33, IL-2 and IL-12 dependence of assay with IL-33 at 10ng/mL; **(B)** Dose titration of IL-33 variants (in the presence of 10 ng/mL IL-2 and 10 ng/mL IL-12). **(C)** ST2 dependence of response. IL-33 variants (10 ng/mL plus IL-2 and IL-12) were tested in the presence of 25 ug/mL human IgG2 control antibody or IgG2 anti-human ST2 blocking antibody.

<https://doi.org/10.1371/journal.pgen.1006659.g005>

and tested their effects on asthma. Our results indicate that variants at the *IL1RL1* locus affect eosinophil counts and the risk of asthma most likely by affecting *IL1RL1* itself. IL-33 is known to mediate its biological effects through ST2/*IL1RL1* [4] and sensitized mice with *il1rl1* (*ST2*<sup>-/-</sup>) knocked out show less eosinophil numbers in bronchoalveolar lavage fluid upon allergen exposure than wild-type mice, reduced levels of Th2 cytokines and chemoattractants in the lungs, and reduced goblet cell hyperplasia around the peripheral airways in murine models of allergy and asthma [36, 37].

We report a rare variant, rs146597587, in *IL33* representing a loss-of-function mutation in this known asthma gene. This splice acceptor mutation associates with lower eosinophil counts and protection against asthma, with the largest protective effect from severe asthma with frequent exacerbations in young Danish children (OR = 0.24, Table 2). This is in agreement with increasing risk conferred by the common *IL33* variant rs2381416 with increasing severity of asthma in these children, with OR from 1.27 to 1.69 for severity groups 1 to 4 (S17 Table). The splice acceptor mutation caused reduced expression of *IL33* transcripts, likely due to NMD, and production of truncated IL-33 that lacks the cytokine function due to lack of binding to IL-33R/ST2 resulting in abrogation of IL-33R-dependent release of CCL1 from mast cells and



**Fig 6. Blood eosinophil counts from wild type and *il33*-deficient mice.** Blood were collected from male and female 10–12 week old mice and eosinophils were enumerated as described in the online methods. The horizontal bars indicate means within groups. Significance calculated using the unpaired T test,  $p = 0.014$  males,  $p = 0.0001$  females.

<https://doi.org/10.1371/journal.pgen.1006659.g006>

IFN- $\gamma$  release from human CD4<sup>+</sup> T cells. We therefore infer that asthma risk is mediated through IL-33. Accordingly, common variants in *IL1RL1* that associate with increased asthma risk associate with reduced expression of soluble ST2 [38, 39], with the predicted effect being increased IL-33 activity, due to reduced level of this decoy receptor. Thus, human genetics support the rationale for therapeutically inhibiting the IL-33-ST2 pathway in an attempt at containing asthma.

## Materials and methods

### Ethics statement

The Icelandic study was approved by the National Bioethics Committee (VSN\_14–099) and the Data Protection Authority (no. PV\_2014060841/PS) in Iceland. All participating subjects who donated blood provided informed consent. Personal identities of the participants and biological samples were encrypted by a third-party system approved and monitored by the Icelandic Data Protection Authority.

The Denmark-1 study was approved by the local ethics committee of Copenhagen, Denmark (Approval no. KF01-400/98 and KF01-074/01). All participants signed informed consent.

The COPSAC study research protocol was approved by The Danish National Ethical Committee on Health Research (KF 01-289/96, H-B-2008-093, H-16039498 and H-B-2998-103) and is in accordance with the ethical scientific principles of the Helsinki Declaration II. All parents in the cohorts signed informed consent.

The German study was approved by the Ethical Commission of the University of Freiburg (Approval no. 96/05). All participants signed informed consent.

The Dutch European ancestry asthma study was approved by the Medical Ethics Committee of the University Hospital Groningen (Approvals no. MEC 90/09/178, MEC 96/04/077, MEC 97/10/184). All participants signed informed consent. The Dutch Vlagtwedde/Vlaardingen study protocol was approved by the local university medical hospital ethics committee, University of Groningen, University Medical Center Groningen, The Netherlands and all

participants gave their written informed consent. In 1984, the Committee on Human Subjects in Research of the University of Groningen reviewed the study and affirmed the safety of the protocol and study design.

The mouse experiments were performed at Amgen Inc. All animal use procedures were in accordance with Amgen animal use and care guidelines and approved by IACUC.

## Study sample sets for eosinophil count and asthma

**Iceland.** We obtained eosinophil counts from three of the largest laboratories in Iceland (measurements performed between the years 1993 and 2015). The median eosinophil counts were 200/ $\mu$ l (lower and upper quartiles: 100/ $\mu$ l and 300/ $\mu$ l). The circulating eosinophil counts were standardized to a standard normal distribution using quantile-quantile standardization and then adjusted for sex, year of birth and age at measurement, as previously described [9, 40]. A total of 103,104 individuals with eosinophil counts were included in the study, where 82,642 were chip-typed and directly imputed; the remaining 20,462 were first and second degree relatives of chip-typed individuals and had their genotypes inferred based on genealogy [16].

Icelandic asthma patients over 18 years of age who attended an asthma clinic or emergency room at the National University Hospital of Iceland or the Icelandic Medical Center (Laekna-setrid) during the years 1977 to 2014 were recruited [9]. Asthma diagnosis was based on a combination of physician's diagnosis, a positive reply to the question: „Has a doctor confirmed your asthma diagnosis?“, questionnaires pertaining to asthma symptoms and ICD diagnosis when receiving emergency care [9, 40]. Atopy status (defined by at least one positive response to allergens) determined by skin prick testing was available for part of the cohort. A total of 3,512 (1,842 chip-typed) asthma cases 45 years old or younger were used in the study (18.5% with known age of onset 18 years of age or younger) [9, 40] and 298,026 (134,762 chip-typed) controls. Icelandic controls were participants from various deCODE genetics programs without known asthma.

The study was approved by the National Bioethics Committee (VSN\_14–099) and the Data Protection Authority (no. PV\_2014060841/PS) in Iceland. All participating subjects who donated blood provided informed consent. Personal identities of the participants and biological samples were encrypted by a third-party system approved and monitored by the Icelandic Data Protection Authority.

**Denmark 1.** Asthma subjects and controls from three studies at the Copenhagen Research Unit, aged 14 to 44 years and resided in Copenhagen, Denmark, were used in the analysis [25, 26]. The asthma patients had doctor diagnosed asthma by a respiratory specialist and had undergone a physical examination and spirometry. Danish control subjects were recruited through the Danish Blood Donor Corps in the Copenhagen area.

**Denmark 2 (COPSAC).** Cases were from the COPSAC<sub>exacerbation</sub> cohort. Children with repeated acute hospitalizations (cases) were identified in the Danish National Patient Register covering all diagnoses of discharges from Danish hospitals [41]. Information on birth-related events was obtained from the national birth register. Inclusion criteria were at least two acute hospitalizations for asthma (ICD-8-code 493, ICD-10 codes J45-46) from 2 to 6 years of age, as previously described [13]. Duration of hospitalization had to be more than 1 day, and two hospitalizations had to be separated by at least 6 months. Exclusion criteria were side diagnosis during hospitalization, registered chronic diagnosis considered to affect risk of hospitalization for asthma, low birth weight (<2.5 kg) or gestational age of under 36 weeks at birth. Control subjects were children without asthma exacerbations in the COPSAC<sub>2000</sub> and COPSAC<sub>2010</sub> birth cohorts. Study participants gave informed consent for use of their biological samples for

genetic studies. The research protocol was approved by The Danish National Ethical Committee on Health Research and is in accordance with the ethical scientific principles of the Helsinki Declaration II.

**Germany.** The German case-control population consisted of children (aged 5–18 years) with suspected asthma recruited from the southwestern part of Germany between July 2000 and January 2005[24]. The probands were characterized at the Centre of Pediatrics and Adolescent Medicine, Freiburg, Germany. The diagnosis of asthma was based on a history of respiratory symptoms, the use of anti-asthmatic medication and the presence of bronchial hyperresponsiveness (defined as a fall of at least 15% in baseline FEV1 after either inhalation of  $\leq 8$  mg/ml histamine or  $\leq 6$  minutes of exercise provocation). Participants were asked in advance to discontinue any asthma or allergy medication before the clinical testing. Atopy status was determined by skin prick tests to 17 common allergens. Measurement of total serum IgE was carried out by using an enzyme allergosorbent test (Phadezym; Pharmacia, Uppsala, Sweden). The control sample consisted of subjects (aged 19–40 years) randomly chosen from the same area in the southwestern part of Germany and children recruited from clinics at the Centre of Pediatrics and Adolescent Medicine, Freiburg, Germany. No medical history was taken, and no medical testing was performed on the adult controls whereas the pediatric controls had no previous history of asthma, recurrent wheezing, atopic dermatitis or atopy. Approval was granted by the Ethical Commission of the University of Freiburg. A statement of informed consent was signed by all participants or signed by their parents in the case of children.

**The Netherlands.** The Dutch cases have been previously described[21–23]. Briefly asthma patients of European ancestry were initially studied between 1962 and 1975. At the time of recruitment and initial testing, all patients had asthma symptoms, were hyperresponsive to histamine (PC20 histamine  $< 32$  mg/ml), and were younger than 45 years of age. Between 1990 and 1999, patients were invited for follow-up tests: including pulmonary function, bronchial responsiveness to histamine, and total and specific serum IgE and skin tests to 16 common aero allergens [22, 23]. Participants were asked in advance to discontinue asthma or allergy medication, therapy with oral corticosteroids was however continued.

A second independent Dutch asthmatic population was ascertained between 1998 and 2001, consisting of probands with asthma, ascertained through local hospitals and media appeals. Controls were derived from Vlagtwedde/Vlaardingen cohort study, a population-based cohort of adult subjects of European ancestry in the Netherlands, recruited from 1965 with a follow-up of over 25 years. Surveys were performed every 3 years, in which information was collected on respiratory symptoms, smoking status, FEV1, and allergy skin tests [42].

## Whole-genome sequencing and imputation

Genotyping of all Icelandic samples was carried out at deCODE genetics in Reykjavik, Iceland, using methods recently described[15]. In brief, whole-genome sequencing was performed for 8,453 Icelanders who were recruited as part of various genetic programs at deCODE genetics, to an average depth of at least 10 $\times$  (median 32 $\times$ ) using Illumina technology. The sequencing was performed using the following three different library preparation methods and sequencing instruments from Illumina: (i) standard TruSeq DNA library preparation method; Illumina GAIIx and/or HiSeq 2000 sequencers; (ii) TruSeq DNA PCR-free library preparation method; Illumina HiSeq 2500 sequencers; and (iii) TruSeq Nano DNA library preparation method; Illumina HiSeq X sequencers (S1 Method).

SNPs and indels identified in the whole-genome sequencing data were identified using the Genome Analysis Toolkit HaplotypeCaller (GATK version 3.3.0)[43] and imputed into

150,656 Icelanders who had been genotyped with various Illumina SNP chips and their genotypes phased using long-range phasing chip-genotyped individuals using long-range phasing [44]. In addition, using the Icelandic genealogical database, genotype probabilities were calculated for first- and second-degree relatives of chip-genotyped individuals [15]. The effects of sequence variants on protein-coding RefSeq genes [18] were annotated using the Variant Effect Predictor (VEP) version 80 [19].

## Genotyping of single variants

Single SNP genotyping in the replication sample sets was carried out by deCODE Genetics in Reykjavik, Iceland, applying the Centaurus (Nanogen) platform [45] except that genotyping single SNPs in the Danish cohort of children with severe asthma was carried out at the AROS Applied Biotechnology AS center (<http://arosab.com/services/microarrays/genotyping/>), in Aarhus, Denmark, applying the Illumina Infinium HumanOmniExpressExome Bead chip platform and genome studio software [46].

## Association testing in Iceland

Generalized linear regression models were used to test for associations between sequence variants and quantitative traits, assuming an additive genetic model. Let  $y$  be the vector of quantitative measurements, and let  $g$  be the vector of expected allele counts for the sequence variant being tested. We assume the quantitative measurements follow a normal distribution with a mean that depends linearly on the expected allele at the variant and a variance covariance matrix proportional to the kinship matrix:

$$y \sim \mathcal{N}(\alpha + \beta g, 2\sigma^2\Phi),$$

where

$$\Phi_{ij} = \begin{cases} \frac{1}{2}, & i = j \\ 2k_{ij}, & i \neq j \end{cases}.$$

is the kinship matrix as estimated from the Icelandic genealogical database. Logistic regression was used to test for association between sequence variants and binary traits. Other available individual characteristics that correlate with disease status were also included in the model as nuisance variables. These characteristics were: sex, county of birth, current age or age at death (first- and second-order terms included), blood sample availability for the individual and an indicator function for the overlap of the lifetime of the individual with the timespan of phenotype collection. Testing was performed using the likelihood ratio statistic. Conditional analysis was performed by including the sequence variant being conditioned on as a covariate in the model under the null and the alternative in the generalized linear regression. Sequence variants were weighted according to their prior probability of affecting gene function by applying different thresholds for genome-wide (gw) significance that depend on the variant class [17].

Association summary statistics are provided in supplementary tables for the *IL33* locus (S18 Table) and the *IL1RL1* locus (S19 Table).

**LD score regression.** To account for inflation in test statistics due to cryptic relatedness and stratification, we applied the method of LD score regression [47]. With a set of 1.1M variant we regressed the  $\chi^2$  statistics from GWAS scans against LD score and used the intercept as a correction factor. The LD scores were downloaded from an LD score database (see URLs). The correction factors were 1.2844 for the eosinophil counts, and 1.1120 for asthma.

## Gene expression microarrays

Samples of RNA from human peripheral blood and adipose tissue were hybridized to Agilent Technologies Human 25K microarrays as described previously[48] and the effect of SNPs on the *IL33* expression in adipocytes evaluated. We quantified expression changes between two samples as the mean logarithm ( $\log_{10}$ ) expression ratio (MLR) compared to a reference pool RNA sample. In comparing expression levels between groups of individuals with different genotypes, we denoted the expression level for each genotype as 10 (average MLR), where the MLR is averaged over individuals with the particular genotype. We determined s.e.m. and significance by regressing the MLR values against the number of risk alleles carried. We took into account the effects of age, gender and differential cell type count in blood as explanatory variables in the regression.

## RNA sequencing

The effect of on the *IL33* expression in adipocytes was evaluated by RNA sequencing as previously described[20].

**Preparation of Poly-A cDNA sequencing libraries.** The quality and quantity of isolated total RNA samples was assessed using the Total RNA 6000 Nano chip for the Agilent 2100 Bioanalyzer. cDNA libraries derived from Poly-A mRNA were generated using Illumina's TruSeq RNA Sample Prep Kit. Briefly, Poly-A mRNA was isolated from total RNA samples (1–4  $\mu\text{g}$  input) using hybridization to Poly-T beads. The Poly-A mRNA was fragmented at 94°C, and first-strand cDNA was prepared using random hexamers and the SuperScript II reverse transcriptase (Invitrogen). Following second-strand cDNA synthesis, end repair, addition of a single A base, adaptor ligation, AMPure bead purification, and PCR amplification, the resulting cDNA was measured on a Bioanalyzer using the DNA 1000 Lab Chip.

**Sequencing.** Samples were clustered on to flowcells using Illumina's cBot and the TruSeq PE cluster kits v2, respectively. Paired-end sequencing was performed with either GAIIX instruments using the TruSeq SBS kits v5 from Illumina or HiSeq 2000 instruments using TruSeq v3 flowcells/SBS kits; read lengths were 2x76, 2x101 or 2x125 cycles.

**Read alignment.** RNA sequencing reads were aligned to Homo sapiens Build 38 with TopHat[49] version 2.0.12 with a supplied set of known transcripts in GTF format (RefSeq hg38; Homo sapiens, NCBI, build 38). TopHat was configured such that it attempts first to align reads to the provided transcriptome then, for reads that do not map fully to the transcriptome, it attempts to map them onto the genome. Read mapping statistics used for read count normalization were calculated using the CollectRnaSeqMetrics tool in Picard version 1.79 (<http://broadinstitute.github.io/picard/command-line-overview.html#CollectRnaSeqMetrics>). After the read alignment step by TopHat, gene expression estimation in FPKM values (Fragments Per Kilobase of transcript per Million mapped reads) was done using Cufflinks[49, 50] version 2.2.1 for the same set of transcripts used in the read alignment by TopHat.

## Mammalian expression

*IL-33* variant-expressing vectors as shown in Fig 4D were transfected into HEK293-EBNA1 cells (obtained from National Research Council of Canada) using the protocols described elsewhere[51]. At 24 hours post transfection, cells were either seeded onto polylysine-coated glass coverslips and cultured for 24 hours for imaging, or fed with Difco yeastolate cell culture supplement (BD Biosciences). Cell culture media were collected and whole cell lysates were prepared 2, 4, and 7 days post transfection.

## Cell imaging

At 48 hrs post transfection, cells were fixed with 4% paraformaldehyde in 0.1 M sodium phosphate, pH 7.2, for 30 minutes at room temperature. After a wash in PBS containing 0.1 M glycine, the fixed cells were incubated with permeabilization buffer (PBS containing 0.4% saponin, 1% BSA, 5% fish gelatin) for 15 minutes, followed by incubation with primary antibody in permeabilization buffer for 60 min. Primary antibodies included anti-IL33 mature domain (R&D Systems), anti-FLAG (clone M2, Amgen), anti-GAPDH (clone 6C5, Amgen), anti-gianin (Covance) and anti-GM130 (BD Transduction Lab). After 3 washes in permeabilization buffer, the cells were incubated with Alexa Fluor 488- or 594-conjugated secondary antibody in permeabilization buffer for 60 minutes. Slides were analyzed using a Nikon Eclipse 80i microscope using a 100× or 60× CFI Plan Apo oil objective lens. Images were acquired using a Cool SNAP HQ2 digital camera (Photometrics) and Nikon Elements imaging software.

## Western blotting

Harvested cell culture media and processed cell pellets were heated for 5 min at 90°C in lithium dodecyl sulfate sample buffer (Life Technologies) containing 5% (v/v) beta-mercaptoethanol. NuPAGE 4–12% Bis–Tris gradient gel and the accompanying running buffer system (both from Life Technologies) were used to perform SDS-PAGE. Resolved proteins were electrotransferred to a nitrocellulose membrane, blocked with fluorescent Western blocking buffer (Rockland), and probed with primary antibodies (described above). After 3–4 washes in PBS containing 0.05% (v/v) Tween-20, the nitrocellulose membranes were incubated with Alexa-Fluor 680- conjugated secondary antibodies (Life Technologies), followed by 2–3 additional washing steps in PBS containing 0.05% (v/v) Tween-20. The fluorescent Western images were acquired by using an Odyssey infrared imaging system from LI-COR Biosciences.

## *E. coli* expression and purification

IL-33 was cloned into the pET24a expression plasmid in frame with an amino terminus poly-His tag and thrombin cleavage site such that thrombin cleavage would release IL-33 proteins corresponding to residues 95–270 (wild type) or 95–204 (mutant).

For IL-33<sub>95–270</sub>, following *E. coli* culture, cell pellets were lysed and the supernatants were collected and purified on a nickel HisTrap column (GE Healthcare Life Sciences). The material was eluted from the column using an imidazole gradient. Appropriate fractions were pooled and further purified by preparative size exclusion chromatography (SEC) using a Superdex 75 column (GE Healthcare Life Sciences) with PBS, pH 7.2, 1 mM DTT as the running buffer. Recovered material was cleaved using the Thrombin CleanCleave Kit (Sigma) with a 1 hour incubation at 4°C. Preparative SEC was used to remove the free His tag and any aggregated material, followed by an additional SEC step to further reduce endotoxin contamination. IL-33<sub>95–270</sub> was formulated in PBS, pH 7.2, 1 mM DTT, 1 mM EDTA and filtered through a 0.2 µm filter. Endotoxin was undetectable.

IL-33<sub>95–204</sub> was not recoverable from the cell pellet supernatant due to low expression in the soluble fraction. Instead, inclusion bodies were solubilized 1/10 (w/v) with 8 M guanidine, 50 mM Tris, 8mM DTT, pH 9.0 for 1 hour. This was followed by dilution 1/25 (v/v) with PBS buffer, pH 7.6 and the material was concentrated and loaded onto a Superdex 75 column using PBS, pH 7.2 as the running buffer. The huIL-33 peak was then buffer-exchanged into 1X thrombin cleavage buffer (50 mM Tris-HCl, pH 8.0, 100 mM CaCl<sub>2</sub>) and concentrated to 1 mg/mL. 0.1 mL of thrombin agarose resin (Sigma) was added to 1 mg of huIL-33 and rotated at 4°C for 1 hour. Cleaved huIL-33 was then passed over a HisTrap column to capture free



cleaved His tag. The flow-through material was further purified using preparative SEC, as for IL-33<sub>95–270</sub>, prior to formulation in PBS, pH 7.2, 1 mM EDTA. Fractions were analyzed by SDS-PAGE and Coomassie staining (total protein) or Western blot (with anti-IL-33) before and after thrombin cleavage to confirm cleavage and IL-33 identity (S4 Fig).

## Computational modeling

The structure model of the ST2/IL-33 complex was provided by Lingel et al.[29]. The receptor and ligand are shown with a molecular surface, computed with default parameters in the Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, QC, Canada, H3A 2R7, 2011. (n.d.)), with ST2 shown in gold and IL-33 shown in blue, with amino acids 205–270, which are derived from exon 7, shown in magenta (Fig 4A).

## Surface plasmon resonance

All experiments were performed using a Biacore T200 optical biosensor (Biacore AB). HBS-P buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Surfactant P20) was purchased from Teknova. The CM5 sensor chip, sodium dodecyl sulphate (SDS) (0.5% w/v), NaOH (50 mM), coupling reagents (N-ethyl-N<sup>3</sup>-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC)/N-hydroxy-succinimide(NHS)), ethanolamine (1.0 M Ethanolamine-HCl, pH 8.5), Acetate 5.0 (10 mM sodium acetate) and Glycine 1.5 (10 mM glycine, pH 1.5) were purchased from GE Healthcare Life Sciences. The HCl solution (100 mM) was purchased from Bio Rad Labs. The penta-His antibody (BSA-free) was purchased from 5Prime. All other reagents were prepared by Amgen.

Biosensor analysis was conducted at 25°C in HBS-P buffer. The CM5 sensor chip was conditioned with twice with serial injections (30 seconds) of SDS (0.1%), NaOH (10 mM), HCl (10 mM) and again with SDS (0.1%) at a flow rate of 30  $\mu$ L/min (first two injections) or 60  $\mu$ L/min (last three injections). The penta-His antibody was diluted (200  $\mu$ g/mL) in water and then buffer exchanged via desalting spin column (Zeba, 0.5 mL, 40 K) into Acetate 5.0 buffer. This buffer-exchanged antibody was immobilized to flow cells 1 (3148 RU) and 2 (7554 RU) of the sensor chip via standard amine coupling (EDC/NHS) and ethanolamine blocking[52]. Human ST2-Flag-His (8  $\mu$ g/mL in HBS-P, Amgen) was injected (300s at 10  $\mu$ L/min) over flow cell 2. This captured between 600 and 900 RU of antibody. After the capture step, the human *E. coli*-derived IL-33 variants (200 nM) were injected (180s, 50  $\mu$ L/min) over flow cells 1–2 to observe the association (180 s) of human IL-33 to human ST2. Each flow cell was then flushed with running buffer to observe the dissociation of human IL-33 (300 s) from the chip surface. Each sample of IL-33 was tested individually as a single replicate. A blank buffer injection was tested before the first sample injection and the surface was regenerated (15  $\mu$ L at 50  $\mu$ L/min) with 10 mM glycine pH 1.5 and a reloading of ST2 was captured to the chip surface before each sample injection.

The data was prepared and analyzed with Scrubber 2.0 software (BioLogic Software Pty Ltd, Campbell, Australia) as follows. The raw data from each experiment was x and y-axis normalized just prior to the injection of IL-33 and then cropped to include the ST2 capture step and the association/dissociation of IL-33.

## Mast cell bioassay

LAD2 human mast cells were seeded into 48-well tissue culture plates at 250,000 cells/375  $\mu$ L media (Stem-Pro-34 serum-free media (Life Technologies, Grand Island, NY, USA) supplemented with 2 mM L-glutamine, 100 IU/ml penicillin, 50  $\mu$ g/ml streptomycin (complete SFM) and 100ng/ml SCF (Peprotech). The following day cells were stimulated with medium only or

increasing concentrations (10 pg/mL– 100 ng/mL) of wild-type (95–270) or mutant (95–204) IL-33 for 24 hours at 37°C. Culture supernatants were then collected followed by removal of cell debris by centrifugation and stored at -80°C until ready for use. The concentration CCL1 in the collected supernatants was determined using an R&D Human CCL1 DuoSet ELISA kit under manufacturer's specifications.

To demonstrate the ST2-dependence of the assay, stimulations were performed as described with the addition of a final concentration of 20 ug/mL human IgG2 isotype control (Amgen) or human anti-huST2 blocking antibody (Amgen) 30 minutes prior to addition of 100 ng/mL IL-33.

## Human CD4 T cell bioassay

Highly purified human CD4+ T cells were seeded at 200,000 cells/well in 96-well round bottom plates. E. coli-produced human IL-33, either beginning at amino acid 95, as described above, or using IL-33 (112–270) (R&D Systems) was either titrated as a dose-response or added to a final concentration of 10ng/mL along with human IL-12 and human IL-2, each at a final concentration of 10ng/mL. For ST2 blocking, anti-huST2 IgG2 or IgG2 control was added to a final concentration of 25 ug/mL 30 minutes prior to addition of cytokines. Cultures were incubated for 72hrs at 37°C in a 5% CO2 incubator. Cell-free supernatants were collected and IFN- $\gamma$  was quantitated by ELISA (R&D Systems).

## Mouse eosinophil enumeration

Blood was collected from 10–12 week old wild-type and *il33*-deficient Bl/6 under isoflurane anesthesia and analyzed on a Siemens Advia 120 hematology analyzer with mouse-specific software. Blood smears were air-dried and stained with Wright-Giemsa. Automated differentials generated on the Advia 120 were confirmed by slide evaluation for any sample with eosinophil counts greater than 4% or with evidence of platelet clumps on the Advia scatterplots or on the blood smear.

## URLs

Predicted loss-of-function variants in *IL33* and their genotype counts in the Exome Aggregation Consortium (ExAC) database were retrieved from <http://exac.broadinstitute.org/gene/ENSG00000137033>, accessed November 24, 2015.

LD Score database (accessed 23 June 2015), [ftp://atguftp.mgh.harvard.edu/brendan/1k\\_eur\\_r2\\_hm3snps\\_se\\_weights.RDS](ftp://atguftp.mgh.harvard.edu/brendan/1k_eur_r2_hm3snps_se_weights.RDS)

## Supporting information

**S1 Methods. Whole-genome sequencing and imputation.** Sample preparation and DNA whole-genome sequencing methods:

- Sample preparation and sequencing using the standard TruSeq DNA library preparation method.
- Sample preparation and sequencing using the TruSeq DNA PCR-free method.
- Sample preparation and sequencing using the TruSeq Nano DNA method. (DOCX)

**S1 Fig. Conditional analysis for eosinophil counts associations in the region around *IL33*.**  
(DOCX)

**S2 Fig. *IL33* mRNA expression in adipose tissue based on microarray data and carrier status of the splice acceptor variant rs146597587.**  
(DOCX)

**S3 Fig. Western blot of mammalian cell-expressed human (FLAG-tagged) *IL-33* protein variants.**  
(DOCX)

**S4 Fig. SDS-PAGE analysis of *E.coli*-produced *IL-33* 95–204 variant.**  
(DOCX)

**S5 Fig. ST2-dependence of human mast cell bioassay.**  
(DOCX)

**S1 Table.  $D'$  and  $r^2$  between the three variants at *IL33* associating with eosinophil counts.**  
(DOCX)

**S2 Table. Correlations ( $r^2$ ) between reported variants at *IL33* and the three significant eosinophil counts variants.**  
(DOCX)

**S3 Table. Association results in Iceland for variants at the *IL33* and *IL1RL1* loci reported to associate with asthma, allergy or eosinophil counts in Iceland.**  
(DOCX)

**S4 Table. Variants that have  $r^2 > 0.8$  with the intergenic variant rs2095044 in an 800kb window centered on *IL33* (chr9:5.8–6.6Mb (hg38)).**  
(DOCX)

**S5 Table. Loss-of-function variants in *IL33* reported in the Exome Aggregation Consortium (ExAC) dataset.**  
(DOCX)

**S6 Table. Predicted loss-of-function, missense and splice region variants in *IL33* that are detected, imputed and tested in Iceland.**  
(DOCX)

**S7 Table. Variants that have  $r^2 > 0.8$  with the splice acceptor variant rs146597587 in an 800kb window centered on *IL33* (chr9:5.8–6.6Mb (hg38)).**  
(DOCX)

**S8 Table. Variants that have  $r^2 > 0.8$  with the intronic variant rs10758750 in a 800kb window centered on *IL33* (chr9:5.8–6.6Mb (hg38)).**  
(DOCX)

**S9 Table. Correlations ( $r^2$ ) of reported variants at *IL1RL1* with the two variants from stepwise regression and two top coding signals.**  
(DOCX)

**S10 Table. Conditional analysis based on eosinophil counts for top variants from stepwise regression and reported variants at the *IL1RL1* locus.**  
(DOCX)

**S11 Table. Variants that have  $r^2 > 0.8$  with the intronic variant rs13020553 in an 800kb window centered on *IL1RL1* (101.9–102.7Mb).**

(DOCX)

**S12 Table. Variants that have  $r^2 > 0.8$  with the intergenic variant rs6719123 in an 800kb window centered on *IL1RL1* (101.9–102.7Mb).**

(DOCX)

**S13 Table. Correlations ( $r^2$ ) for the two variants from stepwise regression and two top coding signals at *IL1RL1*.**

(DOCX)

**S14 Table. Conditional analysis based on eosinophil counts for top variants from stepwise regression and top coding variants at the *IL1RL1* locus.**

(DOCX)

**S15 Table. Predicted loss-of-function, missense and splice region variants in *IL1RL1* that are detected, imputed and tested in Iceland.**

(DOCX)

**S16 Table. Fertility and longevity data for the nine imputed homozygotes for the splice acceptor variant rs146597587 in *IL33* found in Iceland.**

(DOCX)

**S17 Table. Association of the common *IL-33* variant rs2381416 with pediatric asthma of increasing severity.**

(DOCX)

**S18 Table. Associations of sequence variants in a 800kb region centered on *IL33* (chr9:5.8–6.6Mb (hg38) with eosinophil counts in 103,104 Icelanders. Shown are results with  $P < 0.05$ .**

(XLSX)

**S19 Table. Associations of sequence variants in a 800kb region centered on *IL1RL1* (chr2:101.9–102.7Mb (hg38) with eosinophil counts in 103,104 Icelanders. Shown are results with  $P < 0.05$ .**

(XLSX)

## Acknowledgments

We thank all study participants who provided data for this study and our valued colleagues who contributed to data collection and phenotypic characterization of clinical samples, genotyping and analysis of genome sequences data. We thank Tod Martin and Nancy Everds for their contribution to this work.

We would like to thank all study participants, field workers and lab technicians who contributed to the Dutch study.

We gratefully express our gratitude to the children and families of the COPSAC cohorts for all their support and commitment. We also acknowledge and appreciate the unique efforts and teamwork of the COPSAC research team.

## Author Contributions

**Conceptualization:** DS HHe PS USB UT DFG JAJ IJ KS.

**Data curation:** DS HHe PS UT DFG IJ.

**Formal analysis:** HHe PS GS GM UT DFG IJ.

**Funding acquisition:** KS JAJ GHK HB KAD VB HMB.

**Investigation:** USB GIE IO PTO OS DG TG BRL DL HMB GHK AH MK KAD CP VB TSA JW KB HB DS ACL HHa MB RRK MG LG JAJ Adj AsJ AS OTM.

**Project administration:** DS JAJ IJ KS.

**Resources:** HHe PS GS GM UT DFG IJ KS DS ACL HHa MB RRK MG LG JAJ Adj AsJ AS OTM USB GIE IO PTO OS DG TG BRL DL HMB GHK AH MK KAD CP VB TSA JW KB HB.

**Supervision:** DS UT DFG JAJ IJ KS.

**Validation:** DS HHe PS UT DFG JAJ IJ KS.

**Visualization:** DS HHe DFG IJ.

**Writing – original draft:** DS HHe PS DFG IJ KS.

**Writing – review & editing:** GS Adj AsJ AS OTM GM UT ACL HHa MB RRK MG LG JAJ USB GIE IO PTO OS DG TG BRL DL HMB GHK AH MK KAD CP VB TSA JW KB HB.

## References

1. Wood IS, Wang B, Trayhurn P. IL-33, a recently identified interleukin-1 gene family member, is expressed in human adipocytes. *Biochemical and biophysical research communications*. 2009; 384(1):105–9. Epub 2009/04/28. <https://doi.org/10.1016/j.bbrc.2009.04.081> PMID: [19393621](https://pubmed.ncbi.nlm.nih.gov/19393621/)
2. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015; 348(6235):648–60. Epub 2015/05/09. <https://doi.org/10.1126/science.1262110> PMID: [25954001](https://pubmed.ncbi.nlm.nih.gov/25954001/)
3. Carriere V, Roussel L, Ortega N, Lacorre DA, Americh L, Aguilar L, et al. IL-33, the IL-1-like cytokine ligand for ST2 receptor, is a chromatin-associated nuclear factor in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(1):282–7. Epub 2006/12/23. <https://doi.org/10.1073/pnas.0606854104> PMID: [17185418](https://pubmed.ncbi.nlm.nih.gov/17185418/)
4. Schmitz J, Owyang A, Oldham E, Song Y, Murphy E, McClanahan TK, et al. IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. *Immunity*. 2005; 23(5):479–90. Epub 2005/11/16. <https://doi.org/10.1016/j.immuni.2005.09.015> PMID: [16286016](https://pubmed.ncbi.nlm.nih.gov/16286016/)
5. Makrinioti H, Toussaint M, Jackson DJ, Walton RP, Johnston SL. Role of interleukin 33 in respiratory allergy and asthma. *The Lancet Respiratory medicine*. 2014; 2(3):226–37. Epub 2014/03/14. [https://doi.org/10.1016/S2213-2600\(13\)70261-3](https://doi.org/10.1016/S2213-2600(13)70261-3) PMID: [24621684](https://pubmed.ncbi.nlm.nih.gov/24621684/)
6. Cherry WB, Yoon J, Bartemes KR, Iijima K, Kita H. A novel IL-1 family cytokine, IL-33, potently activates human eosinophils. *The Journal of allergy and clinical immunology*. 2008; 121(6):1484–90. Epub 2008/06/10. <https://doi.org/10.1016/j.jaci.2008.04.005> PMID: [18539196](https://pubmed.ncbi.nlm.nih.gov/18539196/)
7. Stolarski B, Kurowska-Stolarska M, Kewin P, Xu D, Liew FY. IL-33 exacerbates eosinophil-mediated airway inflammation. *J Immunol*. 2010; 185(6):3472–80. Epub 2010/08/10. <https://doi.org/10.4049/jimmunol.1000730> PMID: [20693421](https://pubmed.ncbi.nlm.nih.gov/20693421/)
8. Lee HY, Rhee CK, Kang JY, Byun JH, Choi JY, Kim SJ, et al. Blockade of IL-33/ST2 ameliorates airway inflammation in a murine model of allergic asthma. *Experimental lung research*. 2014; 40(2):66–76. Epub 2014/01/23. <https://doi.org/10.3109/01902148.2013.870261> PMID: [24446582](https://pubmed.ncbi.nlm.nih.gov/24446582/)
9. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadóttir A, Sulem P, Jonsdóttir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nature genetics*. 2009; 41(3):342–7. Epub 2009/02/10. <https://doi.org/10.1038/ng.323> PMID: [19198610](https://pubmed.ncbi.nlm.nih.gov/19198610/)
10. Grotenboer NS, Ketelaar ME, Koppelman GH, Nawijn MC. Decoding asthma: translating genetic variation in IL33 and IL1RL1 into disease pathophysiology. *The Journal of allergy and clinical immunology*. 2013; 131(3):856–65. Epub 2013/02/06. <https://doi.org/10.1016/j.jaci.2012.11.028> PMID: [23380221](https://pubmed.ncbi.nlm.nih.gov/23380221/)

11. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nature genetics*. 2011; 43(9):887–92. Epub 2011/08/02. <https://doi.org/10.1038/ng.888> PMID: [21804549](https://pubmed.ncbi.nlm.nih.gov/21804549/)
12. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. *The New England journal of medicine*. 2010; 363(13):1211–21. Epub 2010/09/24. <https://doi.org/10.1056/NEJMoa0906312> PMID: [20860503](https://pubmed.ncbi.nlm.nih.gov/20860503/)
13. Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nature genetics*. 2014; 46(1):51–5. Epub 2013/11/19. <https://doi.org/10.1038/ng.2830> PMID: [24241537](https://pubmed.ncbi.nlm.nih.gov/24241537/)
14. Brusselle GG, Maes T, Bracke KR. Eosinophils in the spotlight: Eosinophilic airway inflammation in non-allergic asthma. *Nature medicine*. 2013; 19(8):977–9. Epub 2013/08/08. <https://doi.org/10.1038/nm.3300> PMID: [23921745](https://pubmed.ncbi.nlm.nih.gov/23921745/)
15. Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, Gylfason A, et al. Large-scale whole-genome sequencing of the Icelandic population. *Nature genetics*. 2015; 47(5):435–44. Epub 2015/03/26. <https://doi.org/10.1038/ng.3247> PMID: [25807286](https://pubmed.ncbi.nlm.nih.gov/25807286/)
16. Styrkarsdottir U, Thorleifsson G, Sulem P, Gudbjartsson DF, Sigurdsson A, Jonasdottir A, et al. Non-sense mutation in the LGR4 gene is associated with several human diseases and other traits. *Nature*. 2013; 497(7450):517–20. Epub 2013/05/07. <https://doi.org/10.1038/nature12124> PMID: [23644456](https://pubmed.ncbi.nlm.nih.gov/23644456/)
17. Sveinbjornsson G, Albrechtsen A, Zink F, Gudjonsson SA, Oddson A, Masson G, et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nature genetics*. 2016; 48(3):314–7. Epub 2016/02/09. <https://doi.org/10.1038/ng.3507> PMID: [26854916](https://pubmed.ncbi.nlm.nih.gov/26854916/)
18. Eilbeck K, Lewis SE, Mungall CJ, Yandell M, Stein L, Durbin R, et al. The Sequence Ontology: a tool for the unification of genome annotations. *Genome biology*. 2005; 6(5):R44. Epub 2005/05/17. <https://doi.org/10.1186/gb-2005-6-5-r44> PMID: [15892872](https://pubmed.ncbi.nlm.nih.gov/15892872/)
19. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010; 26(16):2069–70. Epub 2010/06/22. <https://doi.org/10.1093/bioinformatics/btq330> PMID: [20562413](https://pubmed.ncbi.nlm.nih.gov/20562413/)
20. Gretarsdottir S, Helgason H, Helgadóttir A, Sigurdsson A, Thorleifsson G, Magnúsdóttir A, et al. A Splice Region Variant in LDLR Lowers Non-high Density Lipoprotein Cholesterol and Protects against Coronary Artery Disease. *PLoS genetics*. 2015; 11(9):e1005379. Epub 2015/09/04. <https://doi.org/10.1371/journal.pgen.1005379> PMID: [26327206](https://pubmed.ncbi.nlm.nih.gov/26327206/)
21. Jongepier H, Koppelman GH, Nolte IM, Bruinenberg M, Bleecker ER, Meyers DA, et al. Polymorphisms in SPINK5 are not associated with asthma in a Dutch population. *The Journal of allergy and clinical immunology*. 2005; 115(3):486–92. Epub 2005/03/09. <https://doi.org/10.1016/j.jaci.2004.12.013> PMID: [15753894](https://pubmed.ncbi.nlm.nih.gov/15753894/)
22. Koppelman GH, Stine OC, Xu J, Howard TD, Zheng SL, Kauffman HF, et al. Genome-wide search for atopy susceptibility genes in Dutch families with asthma. *The Journal of allergy and clinical immunology*. 2002; 109(3):498–506. Epub 2002/03/19. PMID: [11897998](https://pubmed.ncbi.nlm.nih.gov/11897998/)
23. Postma DS, Meyers DA, Jongepier H, Howard TD, Koppelman GH, Bleecker ER. Genomewide screen for pulmonary function in 200 families ascertained for asthma. *American journal of respiratory and critical care medicine*. 2005; 172(4):446–52. Epub 2005/05/20. <https://doi.org/10.1164/rccm.200407-864OC> PMID: [15901612](https://pubmed.ncbi.nlm.nih.gov/15901612/)
24. Heinzmann A, Jerkic SP, Ganter K, Kurz T, Blattmann S, Schuchmann L, et al. Association study of the IL13 variant Arg110Gln in atopic diseases and juvenile idiopathic arthritis. *The Journal of allergy and clinical immunology*. 2003; 112(4):735–9. Epub 2003/10/18. [https://doi.org/10.1016/S0091-1456\(4\)352](https://doi.org/10.1016/S0091-1456(4)352)
25. Hansen JW, Thomsen SF, Nolte H, Backer V. Rhinitis: a complication to asthma. *Allergy*. 2010; 65(7):883–8. Epub 2010/02/04. <https://doi.org/10.1111/j.1398-9995.2009.02290.x> PMID: [20121767](https://pubmed.ncbi.nlm.nih.gov/20121767/)
26. Hansen JW, Thomsen SF, Porsbjerg C, Rasmussen LM, Harmsen L, Johansen JS, et al. YKL-40 and genetic status of CHI3L1 in a large group of asthmatics. *European clinical respiratory journal*. 2015; 2:25117. Epub 2015/12/18. <https://doi.org/10.3402/ecrj.v2.25117> PMID: [26672955](https://pubmed.ncbi.nlm.nih.gov/26672955/)
27. Ho JE, Chen WY, Chen MH, Larson MG, McCabe EL, Cheng S, et al. Common genetic variation at the IL1RL1 locus regulates IL-33/ST2 signaling. *The Journal of clinical investigation*. 2013; 123(10):4208–18. Epub 2013/09/04. <https://doi.org/10.1172/JCI67119> PMID: [23999434](https://pubmed.ncbi.nlm.nih.gov/23999434/)
28. Miller JN, Pearce DA. Nonsense-mediated decay in genetic disease: friend or foe? *Mutation research Reviews in mutation research*. 2014; 762:52–64. Epub 2014/12/09. <https://doi.org/10.1016/j.mrrev.2014.05.001> PMID: [25485595](https://pubmed.ncbi.nlm.nih.gov/25485595/)

29. Lingel A, Weiss TM, Niebuhr M, Pan B, Appleton BA, Wiesmann C, et al. Structure of IL-33 and its interaction with the ST2 and IL-1RAcP receptors—insight into heterotrimeric IL-1 signaling complexes. *Structure*. 2009; 17(10):1398–410. Epub 2009/10/20. <https://doi.org/10.1016/j.str.2009.08.009> PMID: [19836339](https://pubmed.ncbi.nlm.nih.gov/19836339/)
30. Lefrancais E, Roga S, Gautier V, Gonzalez-de-Peredo A, Monsarrat B, Girard JP, et al. IL-33 is processed into mature bioactive forms by neutrophil elastase and cathepsin G. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(5):1673–8. Epub 2012/02/07. <https://doi.org/10.1073/pnas.1115884109> PMID: [22307629](https://pubmed.ncbi.nlm.nih.gov/22307629/)
31. Palmer G, Lipsky BP, Smithgall MD, Meininger D, Siu S, Talabot-Ayer D, et al. The IL-1 receptor accessory protein (AcP) is required for IL-33 signaling and soluble AcP enhances the ability of soluble ST2 to inhibit IL-33. *Cytokine*. 2008; 42(3):358–64. Epub 2008/05/03. <https://doi.org/10.1016/j.cyto.2008.03.008> PMID: [18450470](https://pubmed.ncbi.nlm.nih.gov/18450470/)
32. Wu J, Carlock C, Zhou C, Nakae S, Hicks J, Adams HP, et al. IL-33 is required for disposal of unnecessary cells during ovarian atresia through regulation of autophagy and macrophage migration. *J Immunol*. 2015; 194(5):2140–7. Epub 2015/01/27. <https://doi.org/10.4049/jimmunol.1402503> PMID: [25617473](https://pubmed.ncbi.nlm.nih.gov/25617473/)
33. Sulem P, Helgason H, Oddson A, Stefansson H, Gudjonsson SA, Zink F, et al. Identification of a large set of rare complete human knockouts. *Nature genetics*. 2015; 47(5):448–52. Epub 2015/03/26. <https://doi.org/10.1038/ng.3243> PMID: [25807282](https://pubmed.ncbi.nlm.nih.gov/25807282/)
34. Luo H, Higuchi K, Matsumoto K, Mori M. An interleukin-33 gene polymorphism is a modifier for eosinophilia in rats. *Genes and immunity*. 2013; 14(3):192–7. Epub 2013/03/01. <https://doi.org/10.1038/gene.2013.7> PMID: [23446743](https://pubmed.ncbi.nlm.nih.gov/23446743/)
35. Torgerson DG, Capurso D, Mathias RA, Graves PE, Hernandez RD, Beaty TH, et al. Resequencing candidate genes implicates rare variants in asthma susceptibility. *American journal of human genetics*. 2012; 90(2):273–81. Epub 2012/02/14. <https://doi.org/10.1016/j.ajhg.2012.01.008> PMID: [22325360](https://pubmed.ncbi.nlm.nih.gov/22325360/)
36. Morita H, Arae K, Ohno T, Kajiwara N, Oboki K, Matsuda A, et al. ST2 requires Th2-, but not Th17-, type airway inflammation in epicutaneously antigen-sensitized mice. *Allergology international: official journal of the Japanese Society of Allergology*. 2012; 61(2):265–73. Epub 2012/03/01.
37. Zoltowska AM, Lei Y, Fuchs B, Rask C, Adner M, Nilsson GP. The interleukin-33 receptor ST2 is important for the development of peripheral airway hyperresponsiveness and inflammation in a house dust mite mouse model of asthma. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*. 2016; 46(3):479–90. Epub 2015/11/27.
38. Akhbar L, Berube JC, Bosse Y, Laviolette M, Hao K, Nickle DC, et al. Lung expression quantitative trait loci data set identifies important functional polymorphisms in the asthma-associated IL1RL1 region. *The Journal of allergy and clinical immunology*. 2014; 134(3):729–31. Epub 2014/04/22. <https://doi.org/10.1016/j.jaci.2014.02.039> PMID: [24746754](https://pubmed.ncbi.nlm.nih.gov/24746754/)
39. Savenije OE, Kerkhof M, Reijmerink NE, Brunekreef B, de Jongste JC, Smit HA, et al. Interleukin-1 receptor-like 1 polymorphisms are associated with serum IL1RL1-a, eosinophils, and asthma in childhood. *The Journal of allergy and clinical immunology*. 2011; 127(3):750–6 e1–5. Epub 2011/02/02. <https://doi.org/10.1016/j.jaci.2010.12.014> PMID: [21281963](https://pubmed.ncbi.nlm.nih.gov/21281963/)
40. Halapi E, Gudbjartsson DF, Jonsdottir GM, Bjornsdottir US, Thorleifsson G, Helgadottir H, et al. A sequence variant on 17q21 is associated with age at onset and severity of asthma. *European journal of human genetics: EJHG*. 2010; 18(8):902–8. Epub 2010/04/08. <https://doi.org/10.1038/ejhg.2010.38> PMID: [20372189](https://pubmed.ncbi.nlm.nih.gov/20372189/)
41. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian journal of public health*. 2011; 39(7 Suppl):30–3. Epub 2011/08/04. <https://doi.org/10.1177/1403494811401482> PMID: [21775347](https://pubmed.ncbi.nlm.nih.gov/21775347/)
42. van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Schouten JP, Boezen HM. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *American journal of respiratory and critical care medicine*. 2005; 172(3):329–33. Epub 2005/05/10. <https://doi.org/10.1164/rccm.200411-1486OC> PMID: [15879414](https://pubmed.ncbi.nlm.nih.gov/15879414/)
43. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*. 2010; 20(9):1297–303. Epub 2010/07/21. <https://doi.org/10.1101/gr.107524.110> PMID: [20644199](https://pubmed.ncbi.nlm.nih.gov/20644199/)
44. Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, et al. Parental origin of sequence variants associated with complex diseases. *Nature*. 2009; 462(7275):868–74. Epub 2009/12/18. <https://doi.org/10.1038/nature08625> PMID: [20016592](https://pubmed.ncbi.nlm.nih.gov/20016592/)
45. Kutuyav IV, Milesi D, Belousov Y, Podyminogin M, Vorobiev A, Gorn V, et al. A novel endonuclease IV post-PCR genotyping system. *Nucleic acids research*. 2006; 34(19):e128. Epub 2006/10/03. <https://doi.org/10.1093/nar/gkl679> PMID: [17012270](https://pubmed.ncbi.nlm.nih.gov/17012270/)

46. Loisel DA, Du G, Ahluwalia TS, Tisler CJ, Evans MD, Myers RA, et al. Genetic associations with viral respiratory illnesses and asthma control in children. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*. 2016; 46(1):112–24. Epub 2015/09/25.
47. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*. 2015; 47(3):291–5. Epub 2015/02/03. <https://doi.org/10.1038/ng.3211> PMID: [25642630](https://pubmed.ncbi.nlm.nih.gov/25642630/)
48. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. *Nature*. 2008; 452(7186):423–8. Epub 2008/03/18. <https://doi.org/10.1038/nature06758> PMID: [18344981](https://pubmed.ncbi.nlm.nih.gov/18344981/)
49. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*. 2012; 7(3):562–78. Epub 2012/03/03. <https://doi.org/10.1038/nprot.2012.016> PMID: [22383036](https://pubmed.ncbi.nlm.nih.gov/22383036/)
50. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*. 2010; 28(5):511–5. Epub 2010/05/04. <https://doi.org/10.1038/nbt.1621> PMID: [20436464](https://pubmed.ncbi.nlm.nih.gov/20436464/)
51. Stoops J, Byrd S, Hasegawa H. Russell body inducing threshold depends on the variable domain sequences of individual human IgG clones and the cellular protein homeostasis. *Biochimica et biophysica acta*. 2012; 1823(10):1643–57. Epub 2012/06/26. <https://doi.org/10.1016/j.bbamcr.2012.06.015> PMID: [22728328](https://pubmed.ncbi.nlm.nih.gov/22728328/)
52. Johnsson B, Lofas S, Lindquist G. Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Analytical biochemistry*. 1991; 198(2):268–77. Epub 1991/11/01. PMID: [1724720](https://pubmed.ncbi.nlm.nih.gov/1724720/)