

A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order

Adaikalavan Ramasamy, DPhil,^a Ivan Curjuric, MD,^{b,c} Lachlan J. Coin, DPhil,^d Ashish Kumar, MSc,^{b,e,f} Wendy L. McArdle, PhD,^g Medea Imboden, PhD,^{b,c} Benedicte Leynaert, PhD,^h Manolis Kogevinas, MD,^{i,j,k,l} Peter Schmid-Grendelmeier, MD,^m Juha Pekkanen, MD,^{n,o} Matthias Wjst, MD,^p Andreas J. Bircher, MD,^{c,q} Ulla Sovio, PhD,^{d,r} Thierry Rochat, MD,^s Anna-Liisa Hartikainen, MD,^t David J. Balding, DPhil,^u Marjo-Riitta Jarvelin, MD,^{d,v} Nicole Probst-Hensch, PhD,^{b,c} David P. Strachan, MD,^{w,*} and Deborah L. Jarvis, MD^{a,v,*} London, Oxford, and Bristol, United Kingdom, Basel, Zurich, and Geneva, Switzerland, Paris, France, Barcelona, Spain, Heraklion, Greece, Kuopio, Helsinki, and Oulu, Finland, and Munich-Neuherberg, Germany

Background: Hay fever or seasonal allergic rhinitis (AR) is a chronic disorder associated with IgE sensitization to grass. The underlying genetic variants have not been studied comprehensively. There is overwhelming evidence that those who have older siblings have less AR, although the mechanism for this remains unclear.

Objective: We sought to identify common genetic variant associations with prevalent AR and grass sensitization using existing genome-wide association study (GWAS) data and to determine whether genetic variants modify the protective effect of older siblings.

Method: Approximately 2.2 million genotyped or imputed single nucleotide polymorphisms were investigated in 4 large European adult cohorts for AR (3,933 self-reported cases vs 8,965 control subjects) and grass sensitization (2,315 cases vs 10,032 control subjects).

Results: Three loci reached genome-wide significance for either phenotype. The HLA variant rs7775228, which *cis*-regulates *HLA-DRB4*, was strongly associated with grass sensitization and weakly with AR ($P_{\text{grass}} = 1.6 \times 10^{-9}$; $P_{\text{AR}} = 8.0 \times 10^{-3}$). Variants in a locus near chromosome 11 open reading frame 30 (*C11orf30*) and leucine-rich repeat containing 32 (*LRRC32*), which was previously associated with atopic dermatitis and eczema, were

also strongly associated with both phenotypes (rs2155219; $P_{\text{grass}} = 9.4 \times 10^{-9}$; $P_{\text{AR}} = 3.8 \times 10^{-8}$). The third genome-wide significant variant was rs17513503 ($P_{\text{grass}} = 1.2 \times 10^{-8}$; $\text{PAR} = 7.4 \times 10^{-7}$) which was located near transmembrane protein 232 (*TMEM232*) and solute carrier family 25, member 46 (*SLC25A46*). Twelve further loci with suggestive associations were also identified. Using a candidate gene approach, where we considered variants within 164 genes previously thought to be important, we found variants in 3 further genes that may be of interest: thymic stromal lymphopoietin (*TSLP*), Toll-like receptor 6 (*TLR6*) and nucleotide-binding oligomerization domain containing 1 (*NOD1/CARD4*). We found no evidence for variants that modified the effect of birth order on either phenotype. **Conclusions:** This relatively large meta-analysis of GWASs identified few loci associated with AR and grass sensitization. No birth order interaction was identified in the current analyses. (*J Allergy Clin Immunol* 2011;128:996-1005.)

Key words: Hay fever, IgE sensitization to grass, hygiene hypothesis, older siblings, gene-environment interaction, genome-wide association study, European Community Respiratory Health Survey, British 1958 Birth Cohort, Northern Finland Birth Cohort of 1966, Swiss Study on Air Pollution and Lung Disease in Adults

From ^aRespiratory Epidemiology and Public Health, Imperial College, London;

^bChronic Disease Epidemiology, Swiss Tropical and Public Health Institute, Basel;

^cthe University of Basel; ^dthe Department of Epidemiology and Biostatistics, Imperial

College, London; ^ethe Wellcome Trust Centre for Human Genetics, University of Ox-

ford; ^fthe Oxford Centre for Diabetes, Endocrinology and Metabolism, University of

Oxford; ^gthe Avon Longitudinal Study of Parents and Children (ALSPAC) Laboratory,

Department of Social and Community Medicine, University of Bristol; ^hInstitut Na-

tional de la Santé et de la Recherche Médicale, Unit 700, Epidemiologie, Paris; ⁱthe

Centre for Research in Environmental Epidemiology, Barcelona; ^jthe Municipal Insti-

tute of Medical Research (IMIM-Hospital del Mar), Barcelona; ^kCIBER Epi-

demiología y Salud Pública, Barcelona; ^lthe Department of Social Medicine,

Medical School, University of Crete, Heraklion; ^mthe Allergy Unit, Department of

Dermatology, University Hospital Zurich; ⁿthe Department of Environmental Health,

National Institute for Health and Welfare (THL), Helsinki; ^othe Institute of Public

Health and Clinical Nutrition, University of Eastern Finland, Kuopio; ^pHelmholtz Zen-

trum Munchen German Research Center for Environmental Health, Munich-Neuher-

berg; ^qthe Allergy Unit, Department of Dermatology, University Hospital Basel; ^rthe

Department of Medical Statistics, London School of Hygiene and Tropical Medicine;

^sthe Division of Pulmonary Medicine, University Hospitals of Geneva; ^tthe Depart-

ment of Clinical Sciences, Obstetrics and Gynecology, Institute of Clinical Medicine,

University of Oulu; ^uthe Institute of Genetics, University College London; ^vMRC-HPA

Centre for Environment and Health, Imperial College London; and ^wthe Division of

Community Health Sciences, St George's, University of London.

*These authors contributed equally to this work.

Details of the many charities, governmental bodies, and scientific funding organizations that supported the epidemiologic study, including phenotyping, DNA collection, and genotyping for the British 1958 Birth Cohort (B58C), the European Community Respiratory Health Survey (ECRHS2), the Northern Finland Birth Cohort of 1966 (NFBC1966), and the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA), can be found in this article's Online Repository at www.jacionline.org. A. R. has received research support from the European Commission (through project GABRIEL, contract no. 018996 under the Integrated Program LSH-2004-1.2.5-1) and the Department of Health, United Kingdom. U. S. was supported by Medical Research Council studentship grant G0500539.

Disclosure of potential conflict of interest: M. Wjst receives research support from the Helmholtz Center and EU Project European. T. Rochat receives research support from the Swiss National Foundation for Scientific Research. The rest of the authors have declared that they have no conflict of interest.

Received for publication February 25, 2011; revised August 22, 2011; accepted for publication August 29, 2011.

Corresponding author: Deborah L. Jarvis, MD, Emmanuel Kaye Building, National Heart and Lung Institute, Imperial College London, Manresa Rd, London SW3 6LR, United Kingdom. E-mail: d.jarvis@imperial.ac.uk.

0091-6749/\$36.00

© 2011 American Academy of Allergy, Asthma & Immunology

doi:10.1016/j.jaci.2011.08.030

Abbreviations used

AR:	Allergic rhinitis
B58C:	British 1958 Birth Cohort
ECRHS:	European Community Respiratory Health Survey
GWAS:	Genome-wide association study
NFBC1966:	Northern Finland Birth Cohort of 1966
NOD1:	Nucleotide-binding oligomerization domain containing 1
SAPALDIA:	Swiss Study on Air Pollution and Lung Disease in Adults
SLC25A46:	Solute carrier family 25 member 46
SNP:	Single nucleotide polymorphism
STAT6:	Signal transducer and activator of transcription 6
TLR6:	Toll-like receptor 6
TMEM232:	Transmembrane protein 232
TSLP:	Thymic stromal lymphopoietin

Rhinitis is a common chronic disorder in adults, and seasonal allergic rhinitis (AR) or hay fever, which is characterized by episodes of rhinorrhea, sneezing, and itchy/watery eyes, is strongly associated with IgE sensitization to grass and other pollens. Within Western populations, the prevalence of AR and IgE sensitization to grass has increased substantially during the latter half of the 20th century,^{1,2} with some authors identifying the increase as commencing after the industrial revolution.

AR, in common with other allergic diseases, commonly runs in families, and variants in several genes have been identified as biologically plausible candidates for effects on circulating IgE levels, sensitization to specific allergens, or clinical allergic diseases.³ Genetic linkage studies of total IgE levels within families, as well as candidate gene studies, have implicated a number of genetic variants, particularly in the *IL4*, *IL13*, and signal transducer and activator of transcription 6 (*STAT6*) genes, as potentially important determinants of total plasma IgE concentration.⁴⁻⁶ Two genome-wide association studies (GWASs) investigating total IgE concentration have confirmed the associations with *IL13/RAD50* and *STAT6* loci and identified strong additional associations with functional variants of the α chain of the high-affinity receptor for IgE (*FCER1A*) and a single nucleotide polymorphism (SNP) of unknown function near *HLA-DRB1*.^{7,8} However, these loci did not emerge as significant correlates of circulating allergen-specific IgE levels (to grass, cat, and/or dust mite) in a recent genome-wide study of British cohorts.⁹

Because the substantial increase in the prevalence of AR has only occurred over the last few decades, a period too short for substantial change in the genetic makeup of populations, alterations in environmental and lifestyle factors must also be important in the pathogenesis of disease. One of the most enduring hypotheses to explain the increase in the prevalence of AR has been the hygiene hypothesis, which suggests that exposure early in life to infections and microbes leads to altered immune responses, a decreased risk of IgE sensitization, and a decreased risk of AR throughout life. Evidence for this was provided in a large study¹⁰ in which it was noted that children who had many older brothers and sisters (and therefore a greater risk of being exposed to repeated infections) had a lower prevalence of AR. This protective effect of increasing birth order (or

having many siblings) on AR and IgE sensitization to grass has been consistently replicated in studies of children and adults within developed nations.^{11,12} However, the precise immunologic mechanism and exposure through which this relationship is achieved remain unknown.

The aim of this article is to identify genetic variants that are associated with AR and IgE sensitization to grass and to identify genetic variants that modify the protective effect of increasing birth order to these outcomes using GWAS data from approximately 13,000 subjects taking part in 4 large epidemiologic cohort studies. To complement our findings, we also examined the association of SNPs in previously identified candidate genes.

The findings from this work could help elucidate the immunologic mechanisms involved in the pathogenesis of seasonal AR and enhance our understanding of the hygiene hypothesis.

METHODS

Participants and studies

This analysis uses information collected from population samples of white adults taking part in 4 large epidemiologic projects: the British 1958 Birth Cohort (B58C)^{10,13}; the follow-up of the European Community Respiratory Health Survey (ECRHS2)¹⁴⁻¹⁶; the Northern Finland Birth Cohort of 1966 (NFBC1966)¹⁷; and the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA).^{18,19} Participants provided information on AR, either underwent skin prick testing or had specific IgE levels to grass measured in serum, and provided blood samples suitable for DNA extraction. Informed consent was obtained from participants and described elsewhere. Phenotype definition and study descriptors are provided in Table 1.

Genotyping and imputation

Genome-wide genotyping was conducted on available platforms (5 Illumina [San Diego, Calif] and 1 Affymetrix [Santa Clara, Calif]) during the period 2006-2008. After standard quality control checks on genotype data, we imputed HapMap 2 SNPs using the 60 CEU parents as a reference sample to allow testing at ungenotyped SNPs and combined analysis between the studies. Only SNPs with good imputation quality (MACH Rsq >0.40 or IMPUTE info >0.40) with a minor allele frequency of greater than 5% were considered, and up to 2,217,510 imputed and genotyped autosomal SNPs were analyzed. Information on platforms used, the calling algorithm, imputation, and the software used in each study is provided in Table E1 in this article's Online Repository at www.jacionline.org.

Genome-wide association of AR and IgE sensitization to grass

The association tests for AR and IgE sensitization to grass assumed an additive genetic model and included within-study adjustments for age (except in the birth cohorts) and sex. Additionally, B58C was adjusted for region of birth, NFBC1966 was adjusted for the relevant principal components to allow for population stratification, and ECRHS2 and SAPALDIA were adjusted for recruitment centers and principal components. Participants from the B58C study were genotyped as part of 3 different nonoverlapping genetics consortiums, thus resulting in a total of 6 GWAS datasets. All dataset-specific effect estimates are based on the positive strand of mostly National Center for Biotechnology Information build 36 of the reference sequence (1.6% of the SNPs were found in build 35 only). Dataset-specific estimates were meta-analyzed using a fixed effect inverse-variance technique. Genomic control was applied at the dataset level and after meta-analysis; each adjustment was small, with λ_{GC} value < 1.022. The meta-analysis and figures were produced with R version 2.7.0 software.²⁰ We considered any SNP association to be of genome-wide significance at a *P* value < 5×10^{-8} or as suggestive at $5 \times 10^{-8} < P \text{ value} < 5 \times 10^{-6}$.

Candidate genes

We searched the HuGE Literature Finder²¹ for candidate genes using the following key words: “allergic rhinitis” or “rhinitis” or “hay fever” or “grass pollen”; “specific” and “skin prick test”; “specific” and “IgE”; or “hygiene hypothesis.” Then we identified the SNP variants within 5 kb of the flanking regions of each identified autosomal gene using BioMart.²² Associations of the phenotypes with these SNPs reaching a P value $< 10^{-4}$ were examined in detail.

Effect modification of the association of AR and grass sensitization with increasing birth order

We used the 2-step approach proposed by Murcay et al²³ and modified by Ege et al²⁴ for designs other than 1:1 case-control studies (see this article's Online Repository at www.jacionline.org for details) in a meta-analytic context to assess effect modification by genetic variants on the association of AR and IgE to grass, with increasing birth order analyzed as binary (ie, firstborn or not). Briefly, in the first step we selected SNPs that either showed an association with firstborn status (cases and control subjects analyzed separately and then combined using the inverse variance) or with disease status. The χ^2 statistics of SNPs for modeling disease status (step 1a) and for modeling firstborn status (step 1b) were summed and tested with 2 df at a P value $< 10^{-4}$. The SNPs passing step 1 are carried forward in step 2 in which we fit a logistic regression model that included an interaction term between SNP and firstborn status. The effect estimate of the interaction term for each SNP and disease was obtained from each dataset and combined using a fixed-effects meta-analysis. SNPs were considered significant if the P value of the interaction term in step 2 was again $< 10^{-4}$. Steps 1 and 2 are statistically independent, and therefore the overall P value cutoff after accounting for both steps is 1×10^{-8} .

RESULTS

Combining subjects with genome-wide data from the 4 studies, there were 3,933 cases of AR and 8,965 control subjects. Similarly, there were 2,315 subjects with IgE to grass sensitization and 10,032 control subjects. Table I shows the numbers of subjects and the case definitions used in each of the 4 studies.

Genetic variants associated with grass sensitization and AR

Fig 1 shows the Manhattan plots for AR and grass sensitization, and the results are summarized in Table II (Q-Q plots are shown in Fig E1 in this article's Online Repository at www.jacionline.org). SNPs at 3 loci reach genome-wide significance ($P < 5 \times 10^{-8}$) for either AR or grass sensitization, whereas 12 loci show suggestive association ($5 \times 10^{-8} < P < 5 \times 10^{-6}$), and a further 3 candidate genes have SNPs with $5 \times 10^{-6} < P < 1 \times 10^{-4}$. Fig 2 depicts data from Table II and shows the concordance in genetic associations between AR and grass sensitization. Figs 3 to 5 show the regional association and forest plots for the 3 loci that reach genome-wide significance.

The strongest association was with grass sensitization in the HLA region (rs7775228 with $P_{\text{grass}} = 1.6 \times 10^{-9}$; Fig 3). The association was present in all studies, with no evidence of heterogeneity between studies ($P = .43$). This SNP showed only a moderate association with AR ($P_{\text{AR}} = 8.0 \times 10^{-3}$). The rs9271300 variant, which was recently shown to be associated with total IgE levels,⁷ is 77 kb from rs7775228 but was not in linkage disequilibrium (linkage disequilibrium $R^2 = 0.027$) and did

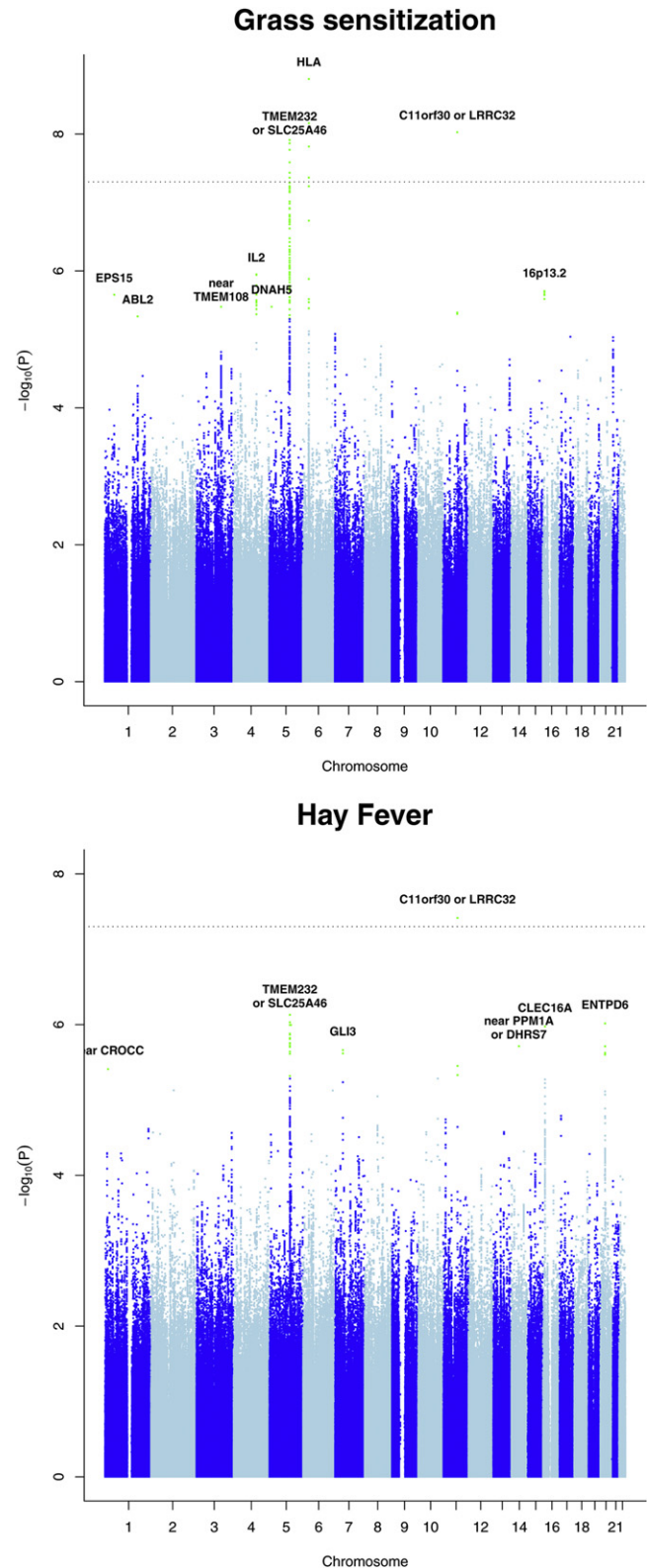


FIG 1. Manhattan plots for main-effect analyses of IgE sensitization to grass and AR.

not show any evidence of association in our dataset ($P_{\text{grass}} = .17$; $P_{\text{AR}} = .029$; see Fig E2, A, in this article's Online Repository at www.jacionline.org).

TABLE I. Phenotype definition for each study, numbers of cases and control subjects available, and size of the protective effect of hay fever

Study	Location	Age at which information was collected	Genotyping platform*	AR definition	No. of AR cases	No. of AR control subjects	Unadjusted OR for AR with presence of ≥ 1 older sibling	IgE to grass definition	No. with IgE grass sensitization	No. without IgE grass sensitization
B58C	Britain (11 regions)	42 for AR 44.5 for IgE	Affymetrix 500k (WTCCC) Illumina 550k (T1DGC) Illumina Quad 610 (GABRIEL)	Have you ever had or been told you had “hay fever”?	1,091	3,699	0.667 (0.581-0.765)	HYTEC automated enzyme immunoassay to mixed grasses with 0.30 kU/L cutoff (specific IgE was tested only if total IgE was >30 kU/L)	927	3,616
ECRHS2	Europe (15 centers)	27-57	Illumina Quad 610	Do you have any nasal allergies, including hay fever?	722	1,412	0.885 (0.732-1.069)	Specific IgE for Timothy grass using Pharmacia CAP system with 0.35 kU/L cutoff	465	1,688
NFBC1966	Northern Finland (5 centers)	31	Illumina CNV 370 duo	Have you ever had the following symptoms or conditions associated with the airways, allergies, or both? AR (associated with animals and pollens, such as hay fever)	1,788	2,912	0.875 (0.771-0.992)	Skin prick test for Timothy grass; response positive if MWD grass–MWD negative control ≥ 3 mm	695	3,727
SAPALDIA	Switzerland (8 centers)	18-60	Illumina Quad 610	Do you have any nasal allergies, including hay fever?	332	942	0.776 (0.600-1.000)	Skin prick test for Timothy grass; response positive if MWD grass–MWD negative control ≥ 3 mm	228	1001
Total					3,933	8,965	0.792 (0.731-0.857)		2,315	10,032

MWD, Mean wheal diameter; OR, odds ratio.

*See this article's Online Repository for more information about genotyping platforms, the calling algorithm, filters applied before imputation, imputation software, and genotype-phenotype software.

TABLE II. List of loci that either achieve genome-wide significance in the meta-analyses of 4 cohorts, have a suggestive association, or are located in candidate genes with $5 \times 10^{-6} < P < 1 \times 10^{-4}$ for either phenotype

Band	Genes	Best SNP	Chromosome (position)	Risk allele/ reference	Risk allele frequency	AR		Grass sensitization	
						OR (95% CI)	P value	OR (95% CI)	P value
Genome-wide significant loci (SNPs with $P < 5 \times 10^{-8}$ for either AR or grass sensitization) from GWASs									
6p21.32	HLA region	rs7775228	6 (32,766,057)	C/T	13.2%	1.11 (1.03-1.20)	8.0×10^{-3}	1.33 (1.21-1.45)	1.6×10^{-9}
11q13.5	<i>C11orf30</i> or <i>LRRC32</i>	rs2155219	11 (75,976,842)	T/G	47.0%	1.17 (1.11-1.24)	3.8×10^{-8}	1.22 (1.14-1.31)	9.4×10^{-9}
5q22.1	<i>TMEM232</i> and <i>SLCA25A46</i>	rs17513503	5 (110,174,345)	G/C	8.6%	1.28 (1.16-1.41)	7.4×10^{-7}	1.39 (1.24-1.56)	1.2×10^{-8}
Suggestive loci (SNPs with $5 \times 10^{-8} < P < 5 \times 10^{-6}$ for either AR or grass sensitization) from GWASs									
20p11.21	<i>ENTPD6</i> (previously known as IL-6 signal transducer)	rs1044573	20 (25,154,654)	A/G	51.3%	1.15 (1.09-1.21)	9.7×10^{-7}	1.04 (0.98-1.12)	2.0×10^{-1}
5q23.1	70 kb to <i>SEMA6A</i>	rs6898653	5 (116,003,555)	G/A	21.8%	1.23 (1.13-1.34)	1.0×10^{-6}	1.09 (0.98-1.21)	1.0×10^{-1}
16p13.13	C-type lectin domain family 16, member A (<i>CLEC16A</i>)	rs887864	16 (11,066,386)	A/G	63.3%	1.16 (1.09-1.23)	1.1×10^{-6}	1.03 (0.96-1.11)	3.4×10^{-1}
4q27	<i>IL2</i>	rs2069772	4 (123,730,738)	C/T	30.3%	1.07 (1.01-1.14)	1.5×10^{-2}	1.19 (1.11-1.28)	1.1×10^{-6}
14q23.1	Near <i>PPM1A</i> and <i>DHRS7</i> (a dehydrogenase/reductase)	rs216518	14 (59,753,183)	A/C	14.5%	1.21 (1.12-1.31)	1.9×10^{-6}	1.09 (0.99-1.20)	8.6×10^{-2}
16p13.2	Intergenic region	rs631208	16 (9,307,225)	A/G	59.5%	1.09 (1.03-1.15)	2.1×10^{-3}	1.18 (1.10-1.26)	2.0×10^{-6}
7p14.1	GLI family zinc finger 3 (<i>GLI3</i>)	rs4724100	7 (42,037,919)	C/T	44.7%	1.14 (1.08-1.21)	2.2×10^{-6}	1.03 (0.96-1.10)	3.9×10^{-1}
1p32.3	Epidermal growth factor receptor pathway substrate 15 (<i>EPS15</i>)	rs6673480	1 (51,571,263)	T/C	7.2%	1.09 (0.98-1.22)	1.1×10^{-1}	1.35 (1.19-1.52)	2.2×10^{-6}
5p15.2	<i>DNAH5</i> (a force generating of respiratory cilia)	rs6554809	5 (13,793,976)	C/T	84.2%	1.07 (0.99-1.17)	9.1×10^{-2}	1.29 (1.16-1.43)	3.3×10^{-6}
3q22.1	30 kb to transmembrane protein 108 (<i>TMEM108</i>)	rs7617456	3 (134,210,601)	G/A	58.5%	1.03 (0.98-1.09)	2.6×10^{-1}	1.18 (1.10-1.26)	3.3×10^{-6}
1p36.13	7 kb to ciliary rootlet coiled-coil, rootletin (<i>CROCC</i>)	rs6586513	1 (16,961,637)	C/A	28.7%	1.24 (1.13-1.36)	3.9×10^{-6}	1.11 (0.99-1.24)	6.2×10^{-2}
1q25.2	v-abl Abelson murine leukemia viral oncogene homolog 2 (<i>ABL2</i>)	rs1325195	1 (175,803,413)	A/G	59.3%	1.01 (0.95-1.06)	8.1×10^{-1}	1.17 (1.10-1.26)	4.6×10^{-6}
Additional candidate loci not included above (SNPs with $5 \times 10^{-6} < P < 1 \times 10^{-4}$ for either AR or grass sensitization)									
5q22.1	<i>TSLP</i>	rs1898671	5 (110,435,901)	T/C	34.7%	1.15 (1.08-1.22)	5.2×10^{-6}	1.10 (1.02-1.18)	9.0×10^{-3}
4p14	<i>TLR6</i>	rs3860069	4 (38,684,687)	A/C	79.8%	1.15 (1.07-1.24)	2.7×10^{-4}	1.21 (1.11-1.33)	4.4×10^{-5}
7p14.3	<i>NOD1</i> , previously known as <i>CARD4</i>	rs7789045	7 (30,267,262)	T/A	54.7%	1.04 (0.98-1.10)	1.5×10^{-1}	1.15 (1.08-1.24)	6.2×10^{-5}

Genes in the HLA region and *IL2* are also candidate genes. Positions of SNPs are reported in National Center for Biotechnology Information build 36 coordinates and aligned to the forward strand. The odds ratios reported are per increase in the risk allele and adjusted for age (in nonbirth cohorts), sex, and locale.

OR, Odds ratio.

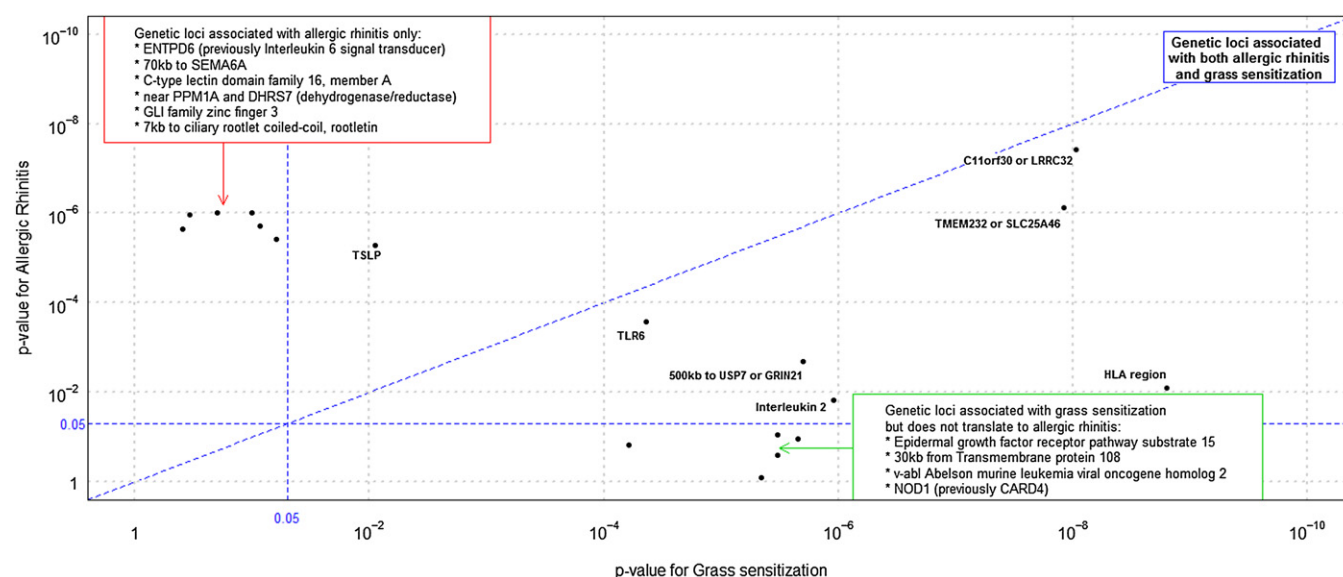


FIG 2. Concordance in statistical significance between AR and grass sensitization for the selected loci. Each loci is represented by the SNP with the lowest P value. The direction of association is consistent between these 2 phenotypes for the loci shown here (see Table II).

SNP rs2155219, located at 11q13.5, 37 kb upstream of chromosome 11 open reading frame 30 (*C11orf30*) and 69 kb downstream of leucine-rich repeat containing 32 (*LRRC32*), was strongly and consistently associated with both grass sensitization and AR ($P_{\text{grass}} = 9.4 \times 10^{-9}$; $P_{\text{AR}} = 3.8 \times 10^{-8}$; Fig 4). A nearby SNP, rs7927894, in high linkage disequilibrium (linkage disequilibrium $R^2 = 0.73$) has previously been identified as being associated with atopic dermatitis²⁵ and eczema²⁶ and, more recently, with hay fever when eczema was also present²⁷ but did not reach genome-wide significance in our results ($P_{\text{grass}} = 4.2 \times 10^{-6}$; $P_{\text{AR}} = 3.5 \times 10^{-6}$; see Fig E2, B).

The third-strongest association is for both phenotypes ($P_{\text{grass}} = 1.2 \times 10^{-8}$; $P_{\text{AR}} = 7.4 \times 10^{-7}$; Fig 5) for rs17513503 situated at the 5q22.1 locus near transmembrane protein 232 (*TMEM232*) and solute carrier family 25, member 46 (*SLC25A46*). The thymic stromal lymphopoietin (*TSLP*) gene, which was previously identified as of possible relevance for allergic disease,²⁸ is located 260 kb away but is located in a different haplotype block and in low linkage disequilibrium. Interestingly, SNPs from *TSLP* also show modest association with AR and weak association with grass sensitization, and for both, the most significantly associated SNP is rs1898671 ($P_{\text{AR}} = 5.2 \times 10^{-6}$; $P_{\text{grass}} = 9 \times 10^{-3}$; see Fig E2, C).

Twelve further loci showed some association ($5 \times 10^{-8} < P < 5 \times 10^{-6}$) for at least one of the 2 phenotypes (Table II). These include *IL2*, which is also a candidate gene (see below), and some biologically plausible genes: *ENTPD6* (previously known as IL-6 signal transducer), epidermal growth factor receptor pathway substrate 15, and *DNAH5* (force-generating protein of respiratory cilia). The regional association and forest plots for these 12 loci are shown in Figs E3 to E14 in this article's Online Repository at www.jacionline.org.

For all reported associations, the direction of effect for hay fever and grass sensitization was consistent, and there was no evidence of heterogeneity of effect between studies ($P > .05$).

Candidate gene analyses

We identified 164 candidate autosomal genes for AR and IgE to grass. Ten of these genes were in the HLA region (*BTNL2*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1*, *HLA-DRB3*, *HLA-DRB4*, *HLA-DRB5*, *TAP1*, and *TAP2*) on chromosome 6. Because this is a gene-dense region, we prefer to visualize the association statistics (Fig 3). As noted earlier, we observe a strong association with grass sensitization and a weak association with AR for SNPs from this region.

The remaining 154 candidate genes (see Table E3 in this article's Online Repository at www.jacionline.org) were mapped to within 5 kb of each gene to 10,839 SNPs in our datasets. SNPs from only 4 candidate genes (*IL2*, nucleotide-binding oligomerization domain containing 1 [*NOD1*]/*CARD4*, Toll-like receptor 6 [*TLR6*], and *TSLP*) had an association P value of less than 10^{-4} for at least 1 of the phenotypes; *IL2* and *TSLP* were already identified earlier within our GWAS as being significant at greater than 1×10^{-6} . The summary statistics for *TSLP*, *TLR*, and *NOD1/CARD4* are shown in Table II and in Figs E15 to E17 in this article's Online Repository at www.jacionline.org.

SNPs that modify the protective effect of increasing birth order

Table I also shows the protective effect of having at least 1 older sibling within each study for an AR (pooled odds ratio of 0.79 [95% CI, 0.73-0.86]). This association persisted when birth order was considered as the number of older siblings and was of similar magnitude in the entire study sample (ie, including subjects who were not genotyped; see Table E3).

Step 1 identifies 647 SNPs with P values $< 10^{-4}$ with firstborn status for either phenotype. In step 2 we tested the SNP-firstborn interaction term for these 647 SNPs, and none of these had a P value $< 10^{-4}$.

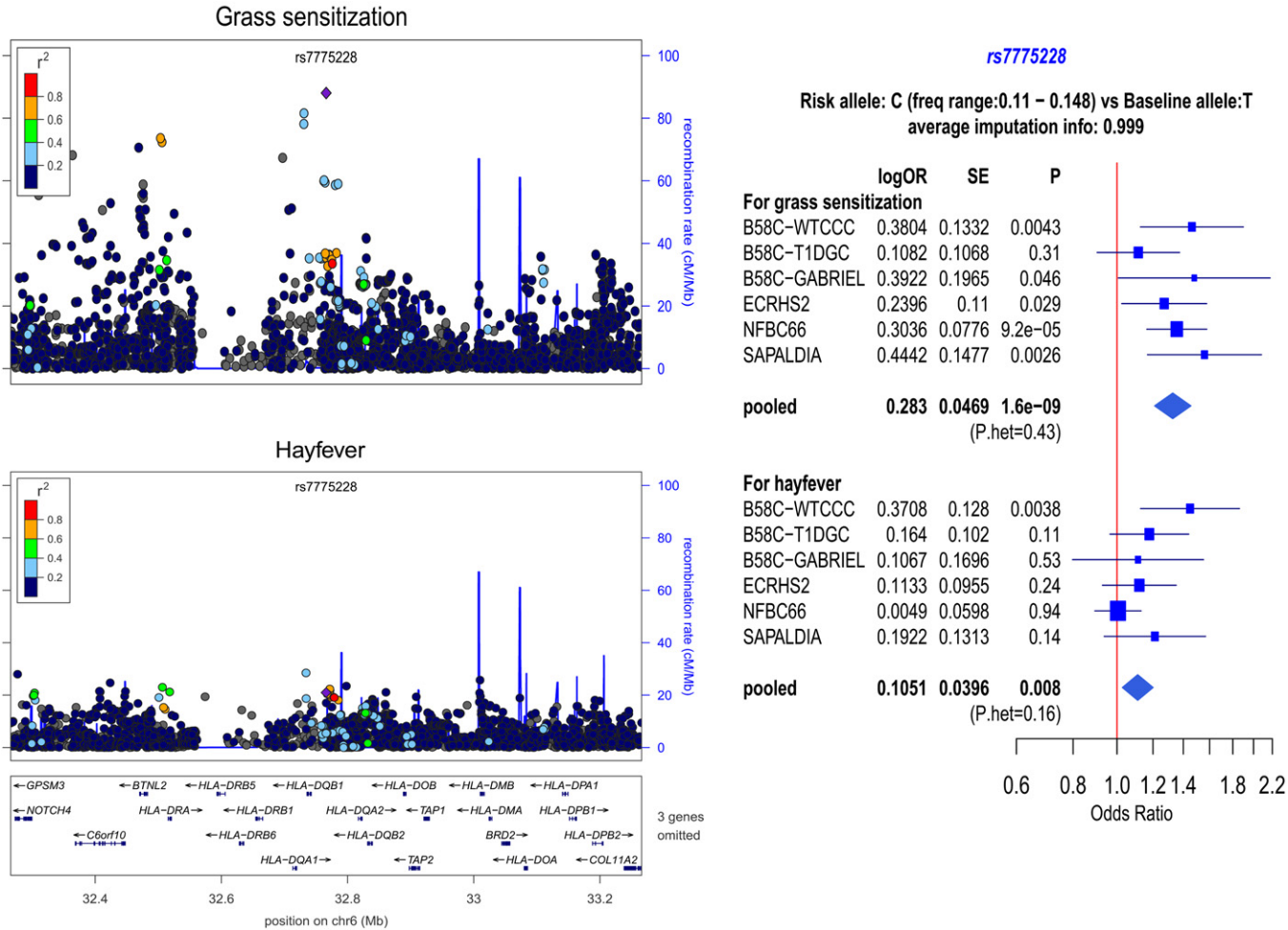


FIG 3. Regional association and forest plots for rs7775228 in the HLA region.

Furthermore, none of the SNPs from the candidate genes showed any evidence (P value of interaction term $< 10^{-4}$) of modifying the protective effect of older siblings.

DISCUSSION

We believe that this is the first genome-wide association meta-analysis of AR and the largest genome-wide interaction study yet conducted for any allergic disease. We investigated the associations of prevalent AR and IgE levels to grass pollens for more than 2.2 million SNPs in almost 13,000 European white adults and also identified genes that might explain the protective effect of increasing birth order on disease. Although we identified several SNPs strongly associated with AR and IgE to sensitization to grass, we found no consistent evidence that any SNPs modify the protective effect of increasing birth order.

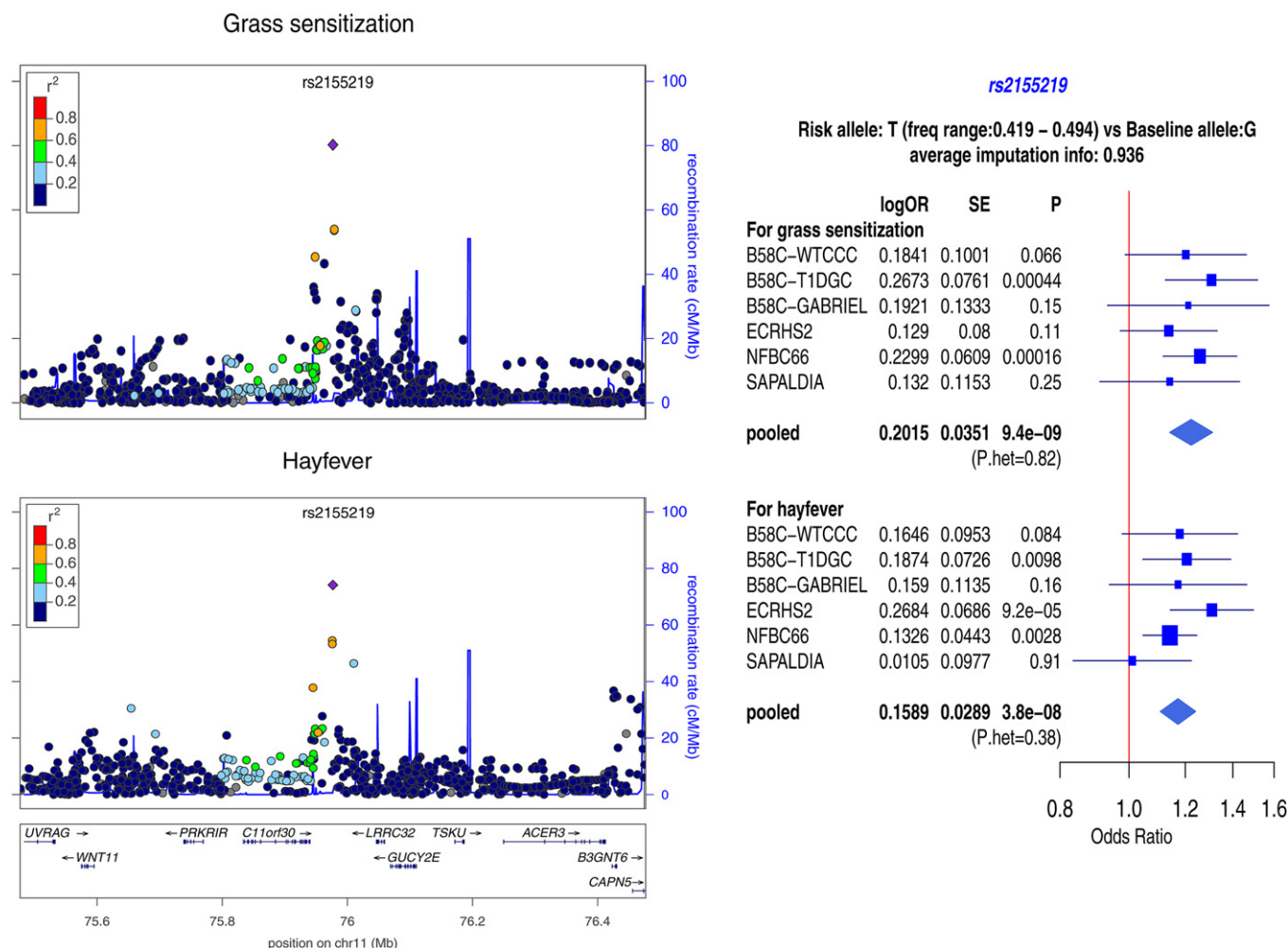
The present study has several strengths. First, it includes GWAS data from almost 13,000 adults of European origin who were recruited into population-based studies (2 birth cohorts and 2 respiratory cohorts) and therefore has good statistical power to detect an association. Second, we investigated the association of SNPs from candidate genes identified in the literature, complementing the genome-wide analysis. Finally, we adopted a

statistically efficient 2-step approach for testing gene-environment interactions in the context of meta-analyzing multiple studies.

It is important to recognize several limitations of the current study. The participants for B58-GABRIEL, ECRHS2, and SAPALDIA were selected for genotyping based on an asthma case-control design. Even though these cohorts are enriched with asthmatic patients and thus are not strictly population representative, we observed highly consistent associations for our top hits in all cohorts (see forest plots) and no significant statistical heterogeneity by asthma status (see Table E4 in this article's Online Repository at www.jacionline.org), suggesting that the association seen is not an artifact of sampling.

There are also several limitations on phenotype definitions. First, the presence of AR is based on self-report and not on a physician's diagnosis and includes the whole spectrum of disease severity and allergy-related comorbidities. However, we note that similar effect sizes were still seen for the majority of the top hits presented here when we restricted our study to subjects without asthma or eczema and also when we used a stricter definition for control subjects and also for cases (see Table E4).

Second, AR could be triggered by exposure to allergens other than grass, but grass sensitization is common among those with



AR and is one of the most commonly tested allergens in epidemiologic studies.

Third, the assessment of IgE sensitization to grass was conducted using a standardized protocol within each of the 4 cohorts, but different methods were used across cohorts. Results from skin prick tests are highly correlated with the presence of specific IgE in serum,²⁹ although it is recognized that they might represent different immunologic processes. Notwithstanding these minor variations in phenotype definition, we observed 3 genome-wide significant associations of SNPS with phenotype, each with small individual risks associated with the risk allele (observed odds ratio, <1.4).

Our strongest association signal was for rs775228 in the HLA region with grass sensitization. It is almost 4 decades since an association between genetic variants in the HLA region and specific allergen sensitization (to ragweed) was reported,³⁰ and various loci in the HLA region have been implicated in candidate gene studies of asthma or allergic phenotypes.³ Expression Quantitative Trait Loci analyses^{31,32} suggest that rs775228 *cis*-regulates *HLA-DRB4* ($P = 1.4 \times 10^{-17}$). Interestingly, the association of this SNP with AR was only moderate. One possible explanation is that the HLA region is important for sensitization

but not necessarily for disease expression. Another explanation could be that the grass pollen is less abundant in Northern Finland whereas birch pollen is more abundant compared with levels seen in central Europe³³ and that both pollen types can trigger AR. There is some support for the second explanation from the forest plot for rs7775228 (Fig 3) in which NFBC1966 shows a marked difference between the association between rs7775228 and grass sensitization compared with its corresponding association with AR.

We also observed strong associations of both AR and IgE sensitization to grass with a common (minor allele frequency, 47%) polymorphism in the 11q13.5 locus. This region has previously been strongly implicated in allergic disorders of the skin. Expression Quantitative Trait Loci analyses suggest that there is a weak association with IL-2 receptor α ($P = 9.7 \times 10^{-5}$).

The third locus to reach genome-wide significance was on 5q22.1 near *TMEM232* or *SLC25A46* (rs17513503). This is of particular interest because of its proximity to a gene related to expression of TSLP. TSLP is a factor thought to be of importance in the severity of allergic disease and showed some associations with asthma in the GABRIEL consortium,⁷ which might also be sex specific.³⁴ However, we found no evidence that the association

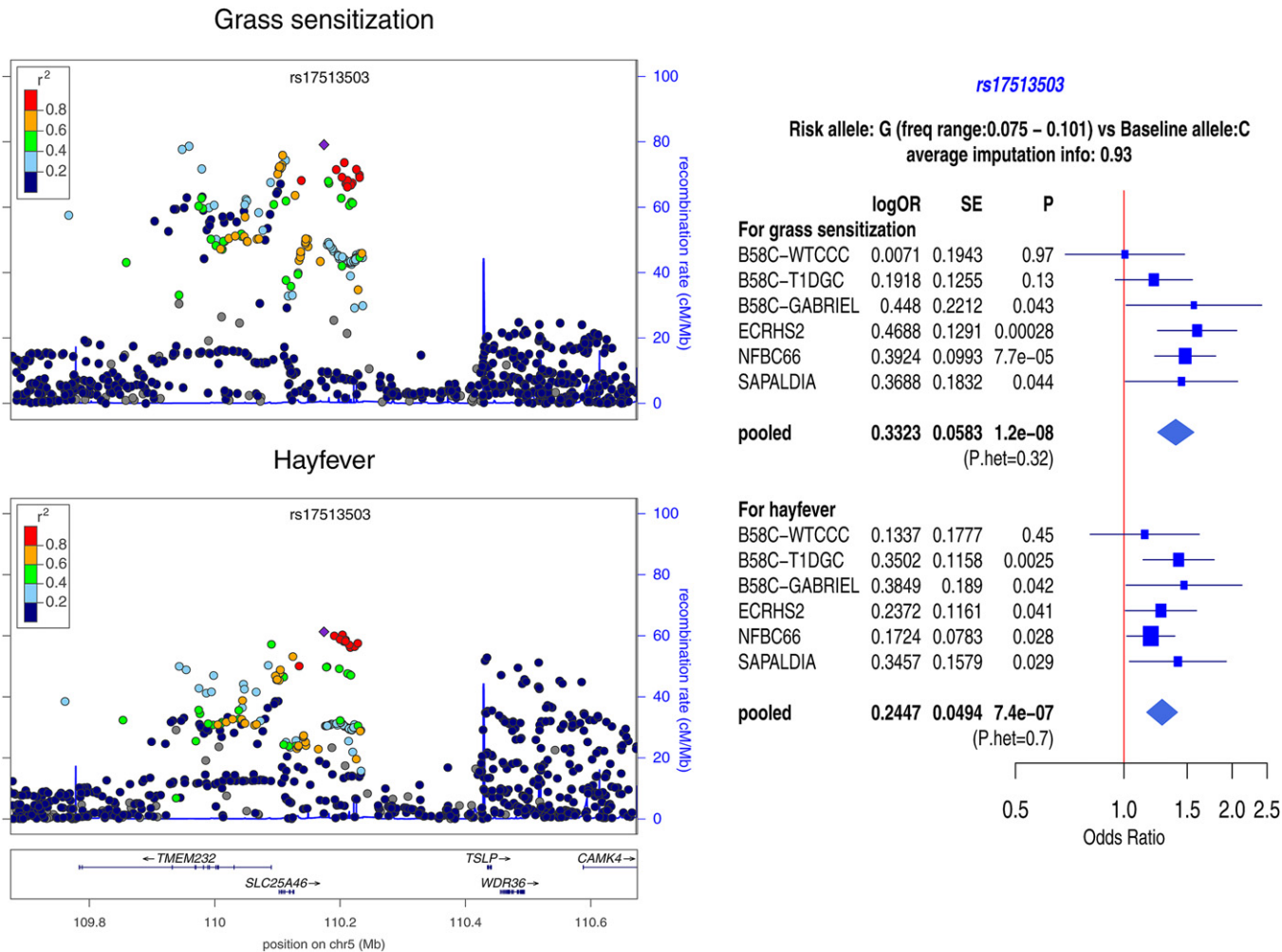


FIG 5. Regional association and forest plots for rs17513503, which is near *SLC25A46* and *TMEM232*.

of this SNP (or other top hits) was changed when stratifying by sex or asthma or eczema status (see Table E4). It might be relevant that rs17513503 very weakly transregulates a transcript of Flotlin 1 (208749_x_at, $P = 7.9 \times 10^{-4}$), an HLA gene of unknown function.

It is striking that in this statistically powerful meta-analysis, variants in only 4 of 154 previously identified candidate genes showed more than a moderate association with either hay fever or grass pollen sensitization. This contrasts with the situation for serum total IgE levels, in which strong signals emerge with biologically plausible candidate genes (*FCERIA*, *IL13*, and *STAT6*), none of which were represented among our prominent associations. Even though the coverage of candidate SNPs from the imputation-based GWAS might be incomplete,³⁵ this is rather surprising given the number of genes tested. It is also intriguing that most of the strong associations for one phenotype were not significant for the other phenotype when the GWASs were based on essentially the same subjects.

Taken together, these observations suggest on the one hand that the genetic determination of allergic disease might perhaps be more complicated than hitherto suspected, with distinct pathways influencing total IgE levels, circulating specific IgE levels, end-organ sensitivity (as manifest by skin prick test responses), and

clinical allergic disease. On the other hand, it is of interest that the 11q13 locus, which has been reported to be associated with eczema, was also associated with both hay fever and grass pollen sensitization in our analyses, illustrating that some variants might have pleiotropic effects on a number of allergic outcomes.

Despite finding a consistent protective effect of increasing birth order on AR in all 4 participating cohorts and adopting a statistically efficient 2-step genome-wide approach to detect possible gene-environment interactions, we were unable to find any SNPs that unequivocally modified the protective effect of older siblings on hay fever and grass pollen sensitization. Furthermore, using a less stringent significance level, we excluded the possibility that SNPs in previously published candidate genes interacted strongly with birth order to determine hay fever risk. A recent study of filaggrin-null mutations and childhood eczema found a significant interaction with older siblings, but this was in the opposite direction as expected.³⁶ A recent case-control study of asthma in central European farming communities failed to replicate previously published gene-environment interactions for asthma and allergic sensitization in relation to living on a farm and, in common with our study, found that a genome-wide search for gene-environment interactions underlying the hygiene hypothesis was unproductive.²⁴

In conclusion, this statistically powerful genome-wide meta-analysis identified only a few loci that are associated with the risk of AR and grass pollen sensitization. This comprehensive exploration of gene-environment interactions with a known correlate of hay fever (birth order) has not provided any convincing evidence for effect modification by common SNPs. Therefore new insights into the biological mechanisms underlying the protective effect of older siblings and the hygiene hypothesis more generally remain elusive.

Please see this article's Online Repository for details of funding and acknowledgement for study recruitment, DNA collection, and genotyping for each study.

Key messages

- We identified few genetic variants that are associated with AR and IgE sensitization to grass, but only 3 of these reach genome-wide significance.
- One of the genome-wide significant loci has also been previously associated with atopic dermatitis and eczema.
- We found no evidence of genetic modification of the protective effect of having older siblings to AR or grass sensitization.

REFERENCES

- Burr ML, Butland BK, King S, Vaughan-Williams E. Changes in asthma prevalence: two surveys 15 years apart. *Arch Dis Child* 1989;64:1452-6.
- Ninan TK, Russell G. Respiratory symptoms and atopy in Aberdeen schoolchildren: evidence from two surveys 25 years apart. *BMJ* 1992;304:873-5.
- Vercelli D. Discovering susceptibility genes for asthma and allergy. *Nat Rev Immunol* 2008;8:169-82.
- Xu J, Postma DS, Howard TD, Koppelman GH, Zheng SL, Stine OC, et al. Major genes regulating total serum immunoglobulin E levels in families with asthma. *Am J Hum Genet* 2000;67:1163-73.
- Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun* 2006;7:95-100.
- Duetsch G, Illig T, Loesgen S, Rohde K, Klopp N, Herbon N, et al. STAT6 as an asthma candidate gene: polymorphism-screening, association and haplotype analysis in a Caucasian sib-pair study. *Hum Mol Genet* 2002;11:613-21.
- Weidinger S, Gieger C, Rodriguez E, Baurecht H, Mempel M, Klopp N, et al. Genome-wide scan on total serum IgE levels identifies FCER1A as novel susceptibility locus. *PLoS Genet* 2008;4:e1000166.
- Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;363:1211-21.
- Wan YI, Strachan DP, Evans DM, Henderson J, McKeever T, Holloway JW, et al. A genome-wide association study to identify genetic determinants of atopy in subjects from the United Kingdom. *J Allergy Clin Immunol* 2011;127:223-31, e1-3.
- Strachan D. Hay fever, hygiene, and household size. *BMJ* 1989;299:1259-60.
- Strachan DP. Family size, infection and atopy: the first decade of the "hygiene hypothesis." *Thorax* 2000;55(suppl 1):S2-10.
- Karmaus W, Botezan C. Does a higher number of siblings protect against the development of allergy and asthma? A review. *J Epidemiol Community Health* 2002;56:209-17.
- Strachan DP, Rudnicka AR, Power C, Shepherd P, Fuller E, Davis A, et al. Life-course influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol* 2007;36:522-31.
- Burney PG, Luczynska C, Chinn S, Jarvis D. The European Community Respiratory Health Survey. *Eur Respir J* 1994;7:954-60.
- Jarvis D, Chinn S, Luczynska C, Burney P. The association of family size with atopy and atopic disease. *Clin Exp Allergy* 1997;27:240-5.
- Genes for asthma? An analysis of the European Community Respiratory Health Survey. *Am J Respir Crit Care Med* 1997;156:1773-80.
- Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. *Acta Paediatr Scand* 1969;193(suppl 193):1+.
- Downs SH, Schindler C, Liu LJ, Keidel D, Bayer-Oglesby L, Brutsche MH, et al. Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med* 2007;357:2338-47.
- Ackermann-Lieblich U, Kuna-Dibbert B, Probst-Hensch NM, Schindler C, Felber Dietrich D, Stutz EZ, et al. Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991-2003: methods and characterization of participants. *Soz Präventivmed* 2005;50:245-63.
- R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: <http://www.R-project.org>. Accessed September 29, 2011.
- HuGE Literature Finder. HuGE navigator. Available at: <http://www.hugenavigator.net/HuGENavigator/startPagePubLit.do>. Accessed September 9, 2010.
- Haider S, Ballester B, Smedley D, Zhang J, Rice P, Kasprzyk A. BioMart Central Portal—unified access to biological data. *Nucleic Acids Res* 2009;37(Web Server issue):W23-7.
- Murcay CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol* 2009;169:219-26.
- Ege MJ, Strachan D, Cookson W, Moffatt M, Gut I, Lathrop M, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. *J Allergy Clin Immunol* 2011;127:138-44.
- Esparza-Gordillo J, Weidinger S, Fölster-Holst R, Bauerfeind A, Ruschendorf F, Patone G, et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat Genet* 2009;41:596-601.
- O'Regan G, Campbell L, Cordell H, Irvine A, McLean W, Brown S. Chromosome 11q13.5 variant associated with childhood eczema: an effect supplementary to filaggrin mutations. *J Allergy Clin Immunol* 2010;125:170-4, e1-2.
- Marenholz I, Bauerfeind A, Esparza-Gordillo J, Kerscher T, Granell R, Nickel R, et al. The eczema risk variant on chromosome 11q13 (rs7927894) in the population-based ALSPAC cohort: a novel susceptibility factor for asthma and hay fever. *Hum Mol Genet* 2011;20:2443-9.
- Fang C, Siew LQ, Corrigan CJ, Ying S. The role of thymic stromal lymphopoietin in allergic inflammation and chronic obstructive pulmonary disease. *Arch Immunol Ther Exp (Warsz)* 2010;58:81-90.
- Weinmayr G, Genuneit J, Nagel G, Björkstén B, van Hage M, Pritanji A, et al. International variations in associations of allergic markers and diseases in children: ISAAC Phase Two. *Allergy* 2010;65:766-75.
- Levine BB, Sember RH, Fotino M. Ragweed hay fever: genetic control and linkage to HL-A haplotypes. *Science* 1972;178:1201-3.
- Dixon A, Liang L, Moffatt M, Chen W, Heath S, Wong K, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202-7.
- Genome-wide evaluation of validity and efficiency of genotype imputation [database on the Internet]. Available at: <http://www.sph.umich.edu/csg/liang/imputation/>. Accessed September 9, 2010.
- Koivikko A, Kupias R, Mäkinen Y, Pohjola A. Pollen seasons: forecasts of the most important allergenic plants in Finland. *Allergy* 1986;41:233-42.
- Hunninghake GM, Soto-Quiros ME, Avila L, Kim HP, Lasky-Su J, Rafaels N, et al. TSLP polymorphisms are associated with asthma in a sex-specific fashion. *Allergy* 2010;65:1566-75.
- Michel S, Liang L, Depner M, Klopp N, Ruether A, Kumar A, et al. Unifying candidate gene and GWAS approaches in asthma. *PLoS One* 2010;5:e13894.
- Cramer C, Link E, Horster M, Koletzko S, Bauer CP, Berdel D, et al. Elder siblings enhance the effect of filaggrin mutations on childhood eczema: results from the 2 birth cohort studies LISAPLUS and GINIPLUS. *J Allergy Clin Immunol* 2010;125:1254-60, e5.

ONLINE REPOSITORY

Original Article

A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization, and their interaction with birth order

Adaikalavan Ramasamy, DPhil,^a Ivan Curjuric, MD,^{b,c} Lachlan J Coin, DPhil,^d Ashish Kumar, MSc,^{b,e,f} Wendy L McArdle, PhD,^g Medea Imboden, PhD,^{b,c} Benedicte Leynaert, PhD,^h Manolis Kogevinas, MD,^{i,j,k,l} Peter Schmid-Grendelmeier, MD,^m Juha Pekkanen, MD,^{n,o} Matthias Wjst, MD,^p Andreas J Bircher, MD,^{c,q} Ulla Sovio, PhD,^{d,r} Thierry Rochat, MD,^s Anna-Liisa Hartikainen, MD,^t David Balding, DPhil,^u Marjo-Riitta Jarvelin, MD,^{d,v} Nicole Probst-Hensch, PhD,^{b,c} David P. Strachan, MD,^{w,*} Deborah L Jarvis, MD,^{a,v,*}

^a Respiratory Epidemiology and Public Health, Imperial College, London SW3 6LR, United Kingdom

^b Chronic Disease Epidemiology, Swiss Tropical and Public Health Institute, Basel, Switzerland

^c University of Basel, Switzerland

^d Department of Epidemiology and Biostatistics, Imperial College, London W2 1PG, United Kingdom

^e Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

^f Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom

^g Avon Longitudinal Study of Parents and Children (ALSPAC) Laboratory, Department of Social and Community Medicine, University of Bristol, Bristol, United Kingdom

^h The Institut National de la Sante et de la Recherche Medicale, Unit 700, Epidemiologie, Paris

ⁱ Centre for Research in Environmental Epidemiology, Barcelona, Spain

^j Municipal Institute of Medical Research (IMIM-Hospital del Mar), Barcelona, Spain

^k CIBER Epidemiología y Salud Pública, Barcelona, Spain

^l Department of Social Medicine, Medical School, University of Crete, Heraklion, Greece

^m Allergy Unit, Department of Dermatology, University Hospital Zurich, Switzerland

ⁿ Department of Environmental Health, National Institute for Health and Welfare (THL), Finland

^o Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Finland

^p Helmholtz Zentrum Munchen German Research Center for Environmental Health, Munich-Neuherberg, Germany

^q Allergy Unit, Department of Dermatology, University Hospital Basel, Switzerland

^r Department of Medical Statistics, London School of Hygiene and Tropical Medicine, United Kingdom

^s Division of Pulmonary Medicine, University Hospitals of Geneva, Switzerland

^t Department of Clinical Sciences, Obstetrics and Gynecology, Institute of Clinical Medicine, University of Oulu, Oulu, Finland.

^u Institute of Genetics, University College London, United Kingdom

^v [MRC-HPA Centre for Environment and Health](#), Imperial College London, United Kingdom

^w Division of Community Health Sciences, St George's, University of London, United Kingdom

*** Joint senior authors**

Corresponding author

Dr. Deborah L Jarvis

Emmanuel Kaye Building, National Heart and Lung Institute, Imperial College London,
Manressa Road, London SW3 6LR.

Telephone: +44 (0) 207 352 8121 ext 3510

Fax: + 44 (0) 207 351 8322

Email: d.jarvis@imperial.ac.uk

METHODS

Studies for inclusion, DNA collection and sources of funding

B58C – a nationwide British birth cohort including participants born in a particular week of 1958 (11). The allergic rhinitis phenotype information relevant for this meta-analysis was collected via an administered interview at ages 33 and 42. Details of the B58C biomedical follow-up at age 44.5 have been previously reported and a full technical report is available online (<http://www.b58cgene.sgu.ac.uk/report.php>). We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02.

NFBC1966 - The Northern Finland Birth Cohort of 1966 (NFBC1966) study programme initiated in the 1960s were established in the provinces of Oulu and Lapland with women expected to give birth in 1966 comprising 12231 children. At age 31, participants were sent a postal questionnaire and invited for clinical assessment where DNA was collected with informed consent. Please see <http://kelo.oulu.fi/NFBC/> for more details about the study.

Financial support was received from the Academy of Finland (project grants 104781, 1114194, 120315 and Center of Excellence in Complex Disease Genetics), Oulu University Hospital, Biocenter Oulu, University of Oulu, Finland, the European Commission (EURO-BLCS, Framework 5 award QLGI-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), ENGAGE project and grant agreement HEALTH-F4-2007-201413, and the Medical Research Council (studentship grant

G0500539). We thank Professor Paula Rantakallio (launch of NFBC1966 and 1986), Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking).

ECRHS – a multicentre, mainly European, population based survey carried out in 1990s with a follow-up starting in 1998. Questionnaire information is based on information collected through interviewer administered questionnaire (forms available at <http://www.ecrhs.org/>). Participants were aged 20 – 48 (mean of 34) at the first survey and 28 – 56 (mean age 43) at the second survey. Samples for genotyping were collected at the follow-up and were mainly European participants and included an enrichment of asthma cases.

The co-ordination of ECRHS II was supported by the European Commission, as part of their Quality of Life programme. The following bodies funded the local studies in ECRHS II: **Albacete:** Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02), Hospital Universitario de Albacete, Consejeria de Sanidad; **Barcelona:** SEPAR, Public Health Service (grant code: R01 HL62633-01), Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02) CIRIT (grant code: 1999SGR 00241) Red Respira ISCII; CIBER Epidemiologia y Salud Pública (CIBERESP), Spain **Basel:** Swiss National Science Foundation, Swiss Federal Office for Education & Science, Swiss National Accident Insurance Fund (SUVA), USC NIEHS Center grant 5P30 ES07048; **Bergen:** Norwegian Research Council, Norwegian Asthma & Allergy Association (NAAF), Glaxo Wellcome AS, Norway Research Fund; **Erfurt:** GSF-National Research Centre for Environment & Health, Deutsche Forschungsgemeinschaft (DFG) (grant code FR 1526/1-1); **Galdakao:** Basque Health Dept; **Grenoble:** Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, CHU de Grenoble, Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, Comite des Maladies Respiratoires de l'Isere; **Hamburg:** GSF-National Research

Centre for Environment & Health, Deutsche Forschungsgemeinschaft (DFG) (grant code MA 711/4-1); **Ipswich and Norwich:** Asthma UK (formerly known as National Asthma Campaign); **Huelva:** Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02); **Oviedo:** Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02) ; **Paris:** Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, UCB-Pharma (France), Aventis (France), Glaxo France, Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, CHU de Grenoble; **Tartu:** Estonian Science Foundation; **Umeå:** Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences & Allergy Research, Swedish Asthma & Allergy Foundation, Swedish Cancer & Allergy Foundation; **Uppsala:** Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences & Allergy Research, Swedish Asthma & Allergy Foundation, Swedish Cancer & Allergy Foundation; *Financial support for ECRHS I for centres in ECRHS II was provided by:* Ministère de la Santé, Glaxo France, Insitut Pneumologique d'Aquitaine, Contrat de Plan Etat-Région Languedoc-Rousillon, CNMATS, CNMRT (90MR/10, 91AF/6), Ministre delegué de la santé, RNSP, France; GSF, and the Bundesminister für Forschung und Technologie, Bonn, Germany; Norwegian Research Council project no. 101422/310; Ministerio Sanidad y Consumo FIS (grants #91/0016060/00E-05E and #93/0393), and grants from Hospital General de Albacete, Hospital General Juan Ramón Jiménez, Consejería de Sanidad Principado de Asturias, Spain; The Swedish Medical Research Council, the Swedish Heart Lung Foundation, the Swedish Association against Asthma and Allergy; Swiss National Science Foundation grant 4026-28099; National Asthma Campaign, British Lung Foundation, Department of Health, South Thames Regional Health Authority, UK.

SAPALDIA – a population-based cohort that recruited subjects aged 18 to 60 from population-based registries in eight Swiss communities. At both, baseline (1991) and follow-up (2002) examination information was collected through interviewer administered questionnaires (<http://www.sapaldia.net/en/>). Participants were mainly Caucasians. DNA samples were collected at follow-up only.

The SAPALDIA Study is supported by grants from the Swiss National Science Foundation (grants no 3347CO-108796/1, 3247BO-104283, 3247BO-104288, 3247BO-104284, 32-65896.01, 32-59302.99, 32-52720.97, 32-4253.94, 4026-28099, PDFMP3-123171), by the Federal Office for Forest, Environment and Landscape, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Landschaft, Geneva, Ticino, and Zurich, Swiss Lung League, the canton's lung leagues of Basel Stadt/ Basel Landschaft, Geneva, Luzern, Ticino and Zurich. Study directorate: Thierry Rochat (p), Jean-Michel Gaspoz (c), Nino Künzli (e/exp), LJ Sally Liu (exp), Nicole M Probst Hensch (e/g), Christian Schindler (s). Scientific team: Ursula Ackermann-Liebrich (e), Jean-Claude Barthélémy (c), Wolfgang Berger (g), Robert Bettschart (p), Andreas Bircher (a), G Bolognini (p), Otto Brändli (p), Martin Brutsche (p), Luc Burdet (p), Martin Frey (p), Margaret W Gerbase (p), Diane Gold (e/c/p), Werner Karrer (p), Roland Keller (p), Bruno Knöpfli (p), Urs Neu (exp), Laurent Nicod (p), Marco Pons (p), Thomas Rothe (p), Erich Russi (p), Peter Schmid-Grendelmeyer (a), Joel Schwartz (e), Diana Stolz (p), Peter Straehl (exp), Jean-Marie Tschopp (p), Arnold von Eckardstein (cc), Jean-Pierre Zellweger (p), Elisabeth Zemp Stutz (e). Scientific team at coordinating centers: Pierre-Olivier Bridevaux (p), Ivan Curjuric (e), Julia Dratva (e), Denise Felber Dietrich (c), Dirk Keidel (s), Medea Imboden (g), Flurina Meier (e), Harish Phuleria (exp), Emmanuel Schaffner (s), Gian-Andri Thun (g), Alex Ineichen (exp), Matthias Ritter (exp). (a) allergology, (c) cardiology, (cc) clinical chemistry, (e) epidemiology, (exp) exposition, (g) genetics and molecular

biology, (m) meteorology, (p) pneumology, (s) statistics. The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites. Local fieldworkers: Aarau: M. Broglie, M. Bünter, D. Gashi, Basel: R. Armbruster, T. Damm, U. Egermann, M. Gut, L. Maier, A. Vögelin, L. Walter, Davos: D. Jud, N. Lutz, Geneva: M. Ares, M. Bennour, B. Galobardes, E. Namer, Lugano: B. Baumberger, S. Boccia Soldati, E. Gehrig-Van Essen, J. Jordan, Wald: R. Gimme, N. Kourkoulos, U. Schafroth. Administration: Nora Bauer, Chantal Gabriel, Rachel Gutknecht.

Funding for genotyping

Genotyping for the **B58C-WTCCC** subset was funded by the Wellcome Trust grant 076113/B/04/Z (32).

The **B58C-T1DGC** genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium (33), a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National

Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895).

DNA extractions, sample quality controls, biobank up-keeping and aliquotting for **NFBC1966** was performed in the National Institute for Health and Welfare, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki.

Genotyping for the **B58C-GABRIEL**, **ECRHS** and **SAPALDIA** was supported by a contract from the European Commission as part of GABRIEL (A multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract number 018996 under the Integrated Program LSH-2004-1.2.5-1 Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control and grants from the French Ministry of Research (see Moffatt, Gut, Demenais et al, NEJM 2010).

Statistical analysis: Summary of method for identifying GxE effects

We present a brief summary of the algorithm for Ege et al (JACI, 2011) in a meta-analytic context. This is a two-step procedure where the first step is to filter SNPs that are to be tested in the second step. Step 1 is divided into two substep Step 1a and Step 1b. The second step involves fitting a model which includes a SNP by environment interaction term and testing if the interaction term is significantly different from zero.

1. FILTERING SNPS

a. Finding SNPs associated with disease

- i. For every study, fit a logistic model to find the SNP association with, say, allergic rhinitis:

$$\text{logit}(\text{allergic rhinitis}) = \beta_{SNP} \times SNP + \beta_{sex} \times sex + \beta_{age} \times age$$

- ii. Apply genomic control at study-level (which inflates the standard errors).

Store the estimates and the (inflated) standard errors as β_{SNP} and

$$SE(\beta_{SNP}).$$

- iii. Meta-analyze β_{SNP} and $SE(\beta_{SNP})$ across the studies using fixed effects inverse-variance technique. Apply genomic control at meta-analyses level.

Denote the resulting pooled estimates and standard error by β_{SNP}^{META} and

$$SE(\beta_{SNP}^{META}).$$

- iv. Calculate $\chi^2_{Step 1a} = \left(\frac{\beta_{SNP}^{META}}{SE(\beta_{SNP}^{META})} \right)^2$ which has 1 degrees of freedom

b. Finding SNPs associated with the exposure

- i. For every study, fit two logistic models to model the binary variable “firstborn” for, say, allergic rhinitis cases and allergic rhinitis controls respectively.

$$\text{logit}(\text{firstborn}) = \beta_{SNP}^{\text{Cases}} \times SNP + \beta_{sex}^{\text{Cases}} \times sex + \beta_{age}^{\text{Cases}} \times age \text{ in allergic rhinitis cases only}$$

$$\text{logit}(\text{firstborn}) = \beta_{SNP}^{\text{Controls}} \times SNP + \beta_{sex}^{\text{Controls}} \times sex + \beta_{age}^{\text{Controls}} \times age \text{ in allergic rhinitis controls only}$$

- ii. Average the SNP coefficients for two models using fixed effect inverse-variance technique.
- iii. Apply genomic control at study-level (which inflates the standard errors).

Store the estimates and (inflated) standard errors as $AvlogOR$ and $SE(AvlogOR)$.

- iv. Meta-analyze $AvlogOR$ and $SE(AvlogOR)$ across the studies using fixed effect inverse-variance technique. Apply genomic control at meta-analyses level. Denote the resulting pooled estimates and standard errors by $AvlogOR^{META}$ and $SE(AvlogOR)^{META}$.

- v. Calculate $X_{Step 1b}^2 = \left(\frac{AvlogOR^{META}}{SE(AvlogOR)^{META}} \right)^2$ which has 1 degree of freedom.

- vi. NOTE: $DlogOR = \beta_{SNP}^{\text{Cases}} - \beta_{SNP}^{\text{Controls}}$ and

$$SE(DlogOR) = \sqrt{SE(\beta_{SNP}^{\text{Cases}})^2 + SE(\beta_{SNP}^{\text{Controls}})^2}$$

are a good and useful approximation of the interaction term tested in STEP2.

c. Combine Step1a and Step1b

- i. Sum up the Chi-squares for Step 1 as $X_{Step\ 1}^2 = X_{Step\ 1a}^2 + X_{Step\ 1b}^2$ which has 2 degrees of freedom.
- ii. Calculate the p-value for $X_{Step\ 1}^2$ using Chi-square distribution with 2 degrees of freedom.
- iii. Select SNPs that have p-value < 0.0001 for the next Step.

2. TESTING FOR INTERACTION TERM IN SNPS SELECTED FROM STEP 1

- a. For every study, fit a logistic model with the interaction term only for SNPs that pass Step 1c.

$$\text{logit}(\text{allergic rhinitis}) = \beta_{SNP} \times SNP + \beta_{firstborn} \times firstborn + \beta_{int} \times firstborn:SNP + \beta_{sex} \times sex + \beta_{age} \times age$$

- b. Calculate the p-value for the β_{int} term from a Gaussian distribution.
- c. Report SNPs that have p-value < 0.0001 as statistically significant.

276 **Table E1** : Information on genotyping, imputation, genotype-phenotype software and genomic control values.

	B58C - WTCCC	B58C - T1DGC	B58C - GABRIEL	ECRHS2	NFBC1966	SAPALDIA
GWAS platform	Affymetrix 500K	Illumina 550K	Illumina Quad 610k	Illumina Quad 610k	Illumina duo 317K CNV	Illumina Quad 610k
Calling algorithm	CHIAMO	ILLUMINUS	Gencall	Gencall	AutoCall	Gencall
Individual call rate filter applied (before imputation)	0.97	0.98	none	none	none	none
SNP call rate filter applied before imputation	none	none	none	none	0.95	none
SNP HWE filter applied (before imputation)	none	none	none	none	0.0001	none
SNP MAF filter applied (before imputation)	none	none	none	none	0.01	none
Other filter (before imputation)	none	none	none	none	None	none
No of SNPs after filtering (before imputation)	490033	520010	582892	582892	328007	582892
Imputation software	IMPUTE 0.2.0	MACH 1.0.13	MACH 1.0	MACH 1.0	IMPUTE v1.0	MACH 1.0
NCBI: HapMap CEU version for imputation	35;21	35; 21	36; 22	36; 22	35; 21	36; 22
Genotype-phenotype association software	QUICKTEST v0.94	ProbABEL v0.0-7	Stata v10	ProbABEL v0.0-7	QUICKTEST v0.94	ProbABEL v0.0-7
Genomic Control Lambda for hay fever	1.003611	1.007229	1.015593	1	1.002972	1.000014
Genomic Control Lambda for grass sensitization	1.001026	1.017668	1.021696	1.00099	1.006281	1

277 **Table E2** : Information on candidate genes (excluding ten genes in the HLA region).

Gene	# SNPs in our analyses that map within 5kb of this gene	Search term for HuGeNet Navigator					Sum
		"allergic rhinitis" or "hayfever" or "hay fever"	"grass pollen"	"specific AND skin prick test"	"specific AND IgE"	"hygiene hypothesis"	
IL13	15	11	0	1	16	0	28
CD14	13	11	0	3	12	1	27
IL4R	69	10	0	1	12	1	24
IL4	23	5	1	1	16	0	23
TNF (previously TNFA) / LTA	42	5	0	1	13	0	19
ADRB2	20	2	0	2	9	0	13
FLG	24	6	0	0	6	1	13
IL18	34	5	0	0	5	0	10
MS4A2	28	2	0	0	8	0	10
TLR4	43	6	0	0	3	1	10
ADAM33	32	4	0	0	4	0	8
IL6	30	5	0	0	3	0	8
TLR2	12	3	0	0	4	1	8
CTLA4	19	3	0	0	4	0	7
IL10	21	3	0	1	3	0	7
GSTP1	16	1	0	0	5	0	6
IFNG	23	3	0	0	3	0	6
NOD1 (previously CARD4)	71	2	0	0	4	0	6
ACE	34	3	0	0	2	0	5
PHF11	52	2	0	0	3	0	5
TLR9	8	1	0	0	3	1	5
FADS1	15	3	0	0	1	0	4
FADS2	46	3	0	0	1	0	4
FCER1A	36	0	0	0	4	0	4
GATA3	40	3	0	0	1	0	4
IL1A	28	1	0	1	2	0	4
IL2	14	2	0	0	2	0	4
IL5	9	3	0	0	1	0	4
NOS2 (previously NOS2A)	50	0	0	0	4	0	4
SPINK5	155	2	0	0	2	0	4
TGFB1	16	2	0	0	2	0	4
CCR2	9	1	0	1	1	0	3
DPP10	1372	1	0	0	2	0	3
FCER1G	21	0	0	0	3	0	3
FCGR1A	2	1	1	0	1	0	3
HAVCR1	34	1	0	0	2	0	3
ICOS	38	2	0	0	1	0	3
MBL2	47	2	0	0	1	0	3
MTHFR	46	1	0	0	2	0	3
NOD2 (previously CARD15)	42	2	0	0	1	0	3
NOS3	16	0	0	1	2	0	3
PTGS2	21	2	0	0	1	0	3
RNASE3	7	1	0	1	1	0	3
AGT	58	1	0	0	1	0	2
C3	73	0	0	0	2	0	2
CCL5	16	1	0	0	1	0	2
CCR3	77	2	0	0	0	0	2
CD55	27	1	0	0	1	0	2
CLC	37	1	1	0	0	0	2
CSF2	18	2	0	0	0	0	2
CYP2D6	8	2	0	0	0	0	2
EPO	5	1	0	0	1	0	2
FOXP1	7	2	0	0	0	0	2

GSTM1	3	0	0	0	2	0	2
GSTT1	11	0	0	0	2	0	2
HNMT	64	2	0	0	0	0	2
IFNGR1	17	1	0	0	1	0	2
IL10RA	38	1	0	0	1	0	2
IL12A	22	0	0	0	2	0	2
IL1B	12	1	0	0	1	0	2
IL1RN	83	1	0	0	1	0	2
IL9	14	0	0	0	2	0	2
IRF1	30	0	0	0	2	0	2
LTC4S	3	1	0	0	1	0	2
NOS1	192	0	0	0	2	0	2
NPSR1	325	0	0	0	2	0	2
PIP	11	1	1	0	0	0	2
PLUNC	21	1	1	0	0	0	2
PTGDR	29	1	0	0	1	0	2
RCBTB1	112	1	0	0	1	0	2
SERPINE1	29	1	0	0	1	0	2
SETDB2	58	1	0	0	1	0	2
SLC11A1	14	0	0	0	2	0	2
STAT1	68	0	0	1	1	0	2
STAT3	34	1	0	0	1	0	2
TBX21	16	0	0	1	1	0	2
TLR3	20	0	0	0	1	1	2
WNT2B	34	1	1	0	0	0	2
ABO	75	0	0	0	1	0	1
ABP1	85	1	0	0	0	0	1
ADAM23	111	0	0	0	1	0	1
ADAMTSL1	737	0	0	0	1	0	1
ADH1B	22	0	0	0	1	0	1
ADH1C	41	0	0	0	1	0	1
ALOX5AP	83	0	0	0	1	0	1
C5orf20	32	0	0	0	1	0	1
CCDC26	6	1	0	0	0	0	1
CCL24	26	1	0	0	0	0	1
CCR1	19	1	0	0	0	0	1
CCR5	7	1	0	0	0	0	1
CD28	42	0	0	0	1	0	1
CD3D	10	1	0	0	0	0	1
CDH26	79	0	0	0	1	0	1
CDKN2A	29	1	0	0	0	0	1
CDKN2B	22	1	0	0	0	0	1
CHIA	69	0	0	0	0	1	1
CHIT1	92	1	0	0	0	0	1
CMA1	27	0	0	0	1	0	1
CSF2RB	67	1	0	0	0	0	1
CTNNA3	2454	0	0	0	1	0	1
CYP1A1	5	0	0	0	1	0	1
DAP3	15	0	0	0	1	0	1
EDN1	31	1	0	0	0	0	1
EGR1	5	1	0	0	0	0	1
EHF	89	0	0	0	1	0	1
ELF5	66	0	0	0	1	0	1
FCGR2A	58	1	0	0	0	0	1
FCGR3A	33	1	0	0	0	0	1
FUT2	19	0	0	0	1	0	1
FUT3	11	0	0	0	1	0	1
GPR44	10	0	0	0	1	0	1
HAVCR2	31	1	0	0	0	0	1
HDC	39	1	0	0	0	0	1

HLA-A	23	0	0	0	1	0	1
HLA-B	34		0	0	1	0	1
HLA-C	15	0	0	0	1	0	1
HMOX1	12	0	0	0	1	0	1
HRH3	7	0	0	0	1	0	1
ICAM1	17	1	0	0	0	0	1
IGHG1	2	0	0	0	1	0	1
IGHG3	1	0	0	0	1	0	1
IL12B	41	1	0	0	0	0	1
IL1R1	201	0	0	0	1	0	1
IL28RA	47	1	0	0	0	0	1
IL2RA	109	0	0	0	1	0	1
IL3	15	1	0	0	0	0	1
IL5RA	68	1	0	0	0	0	1
IL6R	41	0	0	0	1	0	1
IRS2	51	0	0	0	1	0	1
ITGB3	81	0	0	0	1	0	1
ITK	122	1	0	0	0	0	1
KCNJ10	24	0	0	0	1	0	1
LIG4	24	0	0	0	1	0	1
MMP1	39	1	0	0	0	0	1
MPO	11	1	0	0	0	0	1
OPRM1	350	0	0	0	1	0	1
ORMDL3	14	0	0	0	1	0	1
PHLDB1	34	1	0	0	0	0	1
PTPN22	39	0	0	0	1	0	1
RTEL1	41	1	0	0	0	0	1
S100A7	10	1	0	0	0	0	1
SCGB1A1	7	1	0	0	0	0	1
SFTPD	65	1	0	0	0	0	1
STAT4	123	0	0	0	1	0	1
STAT6	17	0	0	0	1	0	1
TBXA2R	12	0	0	0	1	0	1
TERT	17	1	0	0	0	0	1
TGFBR1	56	0	0	0	1	0	1
TGFBR2	132	0	0	0	1	0	1
TLR6	38	0	0	0	0	1	1
TNFSF13B	55	0	0	0	1	0	1
TSLP	33	0	0	0	1	0	1
VEGFA (previously VEGF)	23	0	0	0	1	0	1
XCR1	16	1	0	0	0	0	1

278 **Table E3 :** Table showing the protective effect of number of older siblings for allergic rhinitis. The first rows of each study represent all
 279 individuals with the required information and include individuals without genotyped data. # Odds ratio adjusted for sex, age and centre as
 280 necessary.

Study (subset)	N	Allergic rhinitis status	Number of older siblings					Unadjusted OR per sibling	Adjusted OR per sibling #	Unadjusted OR for non-firstborn
			0	1	2	3	≥ 4			
B58C (all available)	8,922	no	2507	2210	1169	566	572	0.8306 (0.7942, 0.8687)	0.8349 (0.7982, 0.8734)	0.6814 (0.6148,0.7552)
		yes	852	583	249	120	94			
B58C (WTCCC)	1,436	no	385	360	205	92	96	0.8113 (0.7235, 0.9099)	0.8190 (0.7290, 0.9201)	0.7376 (0.5679,0.9581)
		yes	122	101	47	17	11			
B58C (T1DGC)	2,479	no	697	616	329	153	158	0.8160 (0.7487, 0.8893)	0.8201 (0.7521, 0.8942)	0.6221 (0.512,0.7557)
		yes	248	149	73	29	27			
B58C (GABRIEL)	832	no	217	190	86	38	46	0.8125 (0.7101, 0.9297)	0.8181 (0.7128, 0.9391)	0.711 (0.5275,0.9582)
		yes	117	87	25	18	8			
B58C (all gwas)	4747	no	1299	1166	620	283	300	0.8102 (0.7621, 0.8612)	0.8195 (0.7703, 0.8720)	0.6666 (0.5808,0.7649)
		yes	487	337	145	64	46			
ECRHS2 (all available)	10,076	no	2398	2171	1156	553	550	0.9285 (0.8967, 0.9613)	0.9301 (0.8970, 0.9644)	0.8464 (0.7764,0.9226)
		yes	1267	1012	531	212	226			
ECRHS2 (GABRIEL)	2,023	no	467	413	234	106	96	0.9188 (0.8501, 0.9930)	0.9252 (0.8522, 1.0046)	0.8850 (0.7325,1.0691)
		yes	271	240	109	40	47			
NFBC1966 (all available)	8,580	no	1675	1258	858	534	1044	0.9345 (0.9070, 0.9629)	0.9281 (0.8956, 0.9619)	0.8848 (0.8061,0.9712)
		yes	1088	810	508	281	524			
NFBC1966 (with gwas only)	4,689	no	889	665	484	279	584	0.9434 (0.9064, 0.9818)	0.9423 (0.9054, 0.9808)	0.8749 (0.7713,0.9923)
		yes	600	420	294	170	304			
SAPALDIA (all available)	8,168	no	2457	1889	1143	537	649	0.8818 (0.8422, 0.9234)	0.9077 (0.8659, 0.9514)	0.7596 (0.6779,0.8511)
		Yes	648	440	191	98	116			
SAPALDIA (GABRIEL)	1,272	no	333	269	162	87	90	0.8578 (0.7737, 0.9511)	0.8765 (0.7879, 0.9752)	0.7756 (0.6002,1.0022)
		yes	137	114	32	22	26			

Table E4 : Sensitivity analysis for selected hits. Texts in bold black font correspond to association that was highlighted in bold in Table 2; and may not correspond exactly with Table 2 as no genomic correct was applied here and B58C was analyzed a whole. The “Phet” column is the p-value of heterogeneity between various strata (with asthma vs. without asthma, with eczema vs. without eczema, females vs. males). Clean controls are defined as individuals who are negative for both phenotypes. The analysis summarized in the last row of each SNP (i.e. indicated by “(AR & Grass)”) defines cases as individuals who are positive for both phenotypes.

SNP (Gene/Region)	Analysis/Strata	Allergic Rhinitis		Grass Sensitization	
		Odds Ratio (95% CI); p-value	Phet	Odds Ratio (95% CI); p-value	Phet
A) Genome-wide significant loci (SNPs with P < 5 x 10 ⁻⁸ for either allergic rhinitis or grass sensitization) from GWAs					
rs7775228 (HLA region)	Unstratified	1.103 (1.021, 1.191); p=1.3e-02		1.309 (1.197, 1.433); p=4.6e-09	
	... without Asthma	1.084 (0.987, 1.190); p=9.1e-02	0.69	1.314 (1.174, 1.470); p=1.9e-06	0.46
	... with Asthma	1.042 (0.884, 1.229); p=6.2e-01		1.217 (1.029, 1.440); p=2.2e-02	
	... without Eczema	1.159 (1.044, 1.286); p=5.4e-03	0.31	1.341 (1.187, 1.515); p=2.3e-06	0.88
	... with Eczema	1.066 (0.944, 1.204); p=3.0e-01		1.322 (1.151, 1.518); p=7.9e-05	
	... in Females	1.112 (1.000, 1.237); p=5.0e-02	0.81	1.310 (1.146, 1.498); p=7.5e-05	0.90
	... in Males	1.091 (0.974, 1.221); p=1.3e-01		1.326 (1.172, 1.500); p=7.6e-06	
	... vs. clean controls	1.127 (1.040, 1.220); p=3.4e-03		1.308 (1.191, 1.436); p=1.8e-08	
	(AR & Grass) vs. clean controls		1.327 (1.193, 1.477); p=2.1e-07		
rs2155219 (C11orf30 or LRRC32)	Unstratified	1.175 (1.111, 1.243); p=1.8e-08		1.227 (1.147, 1.313); p=3.0e-09	
	... without Asthma	1.154 (1.078, 1.235); p=3.8e-05	0.55	1.209 (1.111, 1.316); p=1.1e-05	0.65
	... with Asthma	1.106 (0.981, 1.246); p=1.0e-01		1.168 (1.032, 1.322); p=1.4e-02	
	... without Eczema	1.174 (1.087, 1.267); p=4.2e-05	0.95	1.204 (1.098, 1.320); p=7.8e-05	0.75
	... with Eczema	1.169 (1.071, 1.276); p=4.8e-04		1.232 (1.111, 1.364); p=6.9e-05	
	... in Females	1.196 (1.107, 1.292); p=5.9e-06	0.58	1.295 (1.171, 1.432); p=5.0e-07	0.19
	... in Males	1.158 (1.067, 1.258); p=4.8e-04		1.181 (1.077, 1.295); p=4.0e-04	
	... vs. clean controls	1.185 (1.118, 1.255); p=1.1e-08		1.258 (1.173, 1.349); p=1.1e-10	
	(AR & Grass) vs. clean controls		1.287 (1.188, 1.395); p=7.2e-10		
rs17513503 (Transmembrane protein 232 & SLC25A46)	Unstratified	1.270 (1.154, 1.397); p=1.0e-06		1.386 (1.240, 1.549); p=8.6e-09	
	... without Asthma	1.249 (1.111, 1.403); p=1.9e-04	0.86	1.298 (1.126, 1.495); p=3.2e-04	0.24
	... with Asthma	1.223 (1.003, 1.492); p=4.7e-02		1.506 (1.230, 1.844); p=7.4e-05	
	... without Eczema	1.305 (1.147, 1.486); p=5.6e-05	0.40	1.387 (1.192, 1.614); p=2.3e-05	0.96
	... with Eczema	1.198 (1.032, 1.392); p=1.8e-02		1.380 (1.165, 1.635); p=2.0e-04	
	... in Females	1.201 (1.049, 1.375); p=7.8e-03	0.19	1.356 (1.146, 1.604); p=3.8e-04	0.54
	... in Males	1.365 (1.191, 1.565); p=8.3e-06		1.455 (1.250, 1.692); p=1.2e-06	

... vs. clean controls	1.268 (1.148, 1.401); p=2.9e-06	1.393 (1.240, 1.564); p=2.1e-08
(AR & Grass) vs. clean controls	1.523 (1.339, 1.733); p=1.5e-10	

B) Suggestive loci (SNPs with $5 \times 10^{-8} < P < 5 \times 10^{-6}$ for either allergic rhinitis or grass sensitization) from GWAs

rs1044573 (ENTPD6 (previously interleukin 6 signal transducer))	Unstratified	1.153 (1.092, 1.218); p=3.4e-07		1.054 (0.987, 1.126); p=1.2e-01	
	... without Asthma	1.131 (1.059, 1.208); p=2.5e-04	0.23	1.046 (0.963, 1.137); p=2.8e-01	0.91
	... with Asthma	1.229 (1.092, 1.383); p=6.1e-04		1.055 (0.935, 1.190); p=3.9e-01	
	... without Eczema	1.209 (1.123, 1.303); p=5.1e-07	0.16	1.062 (0.971, 1.161); p=1.9e-01	0.98
	... with Eczema	1.115 (1.023, 1.215); p=1.3e-02		1.064 (0.962, 1.177); p=2.3e-01	
	... in Females	1.210 (1.122, 1.305); p=7.4e-07	0.07	1.083 (0.982, 1.194); p=1.1e-01	0.49
	... in Males	1.094 (1.010, 1.185); p=2.7e-02		1.033 (0.944, 1.130); p=4.8e-01	
	... vs. clean controls	1.143 (1.080, 1.209); p=3.7e-06		1.081 (1.010, 1.157); p=2.4e-02	
	(AR & Grass) vs. clean controls	1.141 (1.055, 1.233); p=9.2e-04			

rs6898653 (70kb to SEMA6A)	Unstratified	1.240 (1.141, 1.347); p=3.6e-07		1.096 (0.991, 1.212); p=7.4e-02	
	... without Asthma	1.215 (1.099, 1.344); p=1.4e-04	0.47	1.060 (0.933, 1.205); p=3.7e-01	0.53
	... with Asthma	1.310 (1.096, 1.566); p=3.0e-03		1.138 (0.950, 1.365); p=1.6e-01	
	... without Eczema	1.307 (1.169, 1.462); p=2.7e-06	0.13	1.180 (1.029, 1.354); p=1.8e-02	0.08
	... with Eczema	1.147 (1.007, 1.306); p=3.9e-02		0.980 (0.841, 1.143); p=8.0e-01	
	... in Females	1.246 (1.113, 1.396); p=1.4e-04	0.91	1.087 (0.937, 1.261); p=2.7e-01	0.88
	... in Males	1.234 (1.093, 1.395); p=7.2e-04		1.104 (0.961, 1.267); p=1.6e-01	
	... vs. clean controls	1.259 (1.155, 1.372); p=1.5e-07		1.156 (1.041, 1.284); p=6.8e-03	
	(AR & Grass) vs. clean controls	1.224 (1.087, 1.378); p=8.8e-04			

rs887864 (C-type lectin domain family 16, member A)	Unstratified	1.148 (1.084, 1.217); p=2.9e-06		1.031 (0.962, 1.105); p=3.9e-01	
	... without Asthma	1.149 (1.071, 1.233); p=1.1e-04	0.98	1.009 (0.925, 1.101); p=8.4e-01	0.63
	... with Asthma	1.148 (1.017, 1.295); p=2.6e-02		1.048 (0.924, 1.188); p=4.6e-01	
	... without Eczema	1.152 (1.064, 1.246); p=4.5e-04	0.90	1.012 (0.921, 1.112); p=8.0e-01	0.60
	... with Eczema	1.160 (1.060, 1.270); p=1.3e-03		1.051 (0.946, 1.168); p=3.6e-01	
	... in Females	1.175 (1.084, 1.274); p=8.3e-05	0.37	0.986 (0.889, 1.093); p=7.9e-01	0.26
	... in Males	1.115 (1.025, 1.213); p=1.1e-02		1.069 (0.972, 1.174); p=1.7e-01	
	... vs. clean controls	1.147 (1.080, 1.217); p=8.0e-06		1.064 (0.991, 1.143); p=8.7e-02	
	(AR & Grass) vs. clean controls	1.116 (1.027, 1.212); p=9.3e-03			

rs2069772 (Interleukin 2)	Unstratified	1.073 (1.013, 1.137); p=1.7e-02		1.187 (1.107, 1.273); p=1.7e-06	
	... without Asthma	1.043 (0.973, 1.118); p=2.3e-01	0.53	1.192 (1.092, 1.301); p=8.1e-05	0.49
	... with Asthma	1.092 (0.963, 1.238); p=1.7e-01		1.129 (0.992, 1.284); p=6.6e-02	
	... without Eczema	1.044 (0.964, 1.130); p=2.9e-01	0.52	1.194 (1.085, 1.315); p=3.0e-04	0.72

	... with Eczema	1.085 (0.993, 1.187); p=7.3e-02		1.164 (1.048, 1.293); p=4.7e-03	
	... in Females	1.019 (0.940, 1.104); p=6.5e-01	0.06	1.100 (0.990, 1.222); p=7.8e-02	0.04
	... in Males	1.140 (1.047, 1.240); p=2.4e-03		1.274 (1.159, 1.401); p=5.6e-07	
	... vs. clean controls	1.079 (1.016, 1.145); p=1.3e-02		1.193 (1.110, 1.283); p=1.9e-06	
	(AR & Grass) vs. clean controls	1.203 (1.107, 1.307); p=1.4e-05			
rs216518 (near PPM1A and DHRS7 (dehydrogenase/reductase))	Unstratified	1.201 (1.113, 1.296); p=2.2e-06		1.090 (0.994, 1.195); p=6.8e-02	
	... without Asthma	1.222 (1.115, 1.338); p=1.7e-05	0.76	1.174 (1.047, 1.316); p=6.0e-03	0.05
	... with Asthma	1.259 (1.063, 1.491); p=7.5e-03		0.956 (0.805, 1.134); p=6.0e-01	
	... without Eczema	1.248 (1.126, 1.384); p=2.6e-05	0.20	1.115 (0.982, 1.266); p=9.4e-02	0.40
	... with Eczema	1.128 (1.002, 1.269); p=4.6e-02		1.028 (0.895, 1.180); p=7.0e-01	
	... in Females	1.199 (1.080, 1.331); p=6.9e-04	0.96	1.012 (0.880, 1.164); p=8.7e-01	0.15
	... in Males	1.204 (1.077, 1.345); p=1.0e-03		1.162 (1.026, 1.317); p=1.8e-02	
	... vs. clean controls	1.190 (1.100, 1.287); p=1.5e-05		1.133 (1.030, 1.248); p=1.1e-02	
	(AR & Grass) vs. clean controls	1.182 (1.060, 1.318); p=2.5e-03			
rs631208 (500 kb to USP7 and a glutamate receptor)	Unstratified	1.090 (1.032, 1.152); p=2.2e-03		1.180 (1.103, 1.261); p=1.2e-06	
	... without Asthma	1.091 (1.020, 1.167); p=1.1e-02	0.44	1.160 (1.067, 1.262); p=5.3e-04	0.55
	... with Asthma	1.035 (0.921, 1.163); p=5.6e-01		1.213 (1.076, 1.368); p=1.7e-03	
	... without Eczema	1.070 (0.992, 1.153); p=8.1e-02	0.41	1.202 (1.097, 1.317); p=8.2e-05	0.63
	... with Eczema	1.122 (1.030, 1.223); p=8.5e-03		1.162 (1.051, 1.284); p=3.4e-03	
	... in Females	1.089 (1.008, 1.175); p=3.0e-02	0.99	1.243 (1.124, 1.374); p=2.2e-05	0.18
	... in Males	1.089 (1.005, 1.181); p=3.8e-02		1.133 (1.036, 1.240); p=6.4e-03	
	... vs. clean controls	1.107 (1.045, 1.172); p=5.1e-04		1.185 (1.106, 1.270); p=1.4e-06	
	(AR & Grass) vs. clean controls	1.197 (1.106, 1.297); p=8.8e-06			
rs4724100 (GLI family zinc finger 3)	Unstratified	1.134 (1.074, 1.198); p=6.4e-06		1.029 (0.963, 1.098); p=4.0e-01	
	... without Asthma	1.118 (1.046, 1.195); p=9.9e-04	0.11	1.007 (0.927, 1.094); p=8.7e-01	0.25
	... with Asthma	1.247 (1.109, 1.402); p=2.2e-04		1.097 (0.972, 1.236); p=1.3e-01	
	... without Eczema	1.143 (1.061, 1.232); p=4.2e-04	0.78	1.003 (0.917, 1.096); p=9.5e-01	0.38
	... with Eczema	1.125 (1.033, 1.226); p=7.0e-03		1.065 (0.964, 1.177); p=2.1e-01	
	... in Females	1.124 (1.042, 1.212); p=2.4e-03	0.73	0.986 (0.894, 1.087); p=7.7e-01	0.26
	... in Males	1.147 (1.058, 1.242); p=8.0e-04		1.064 (0.973, 1.163); p=1.7e-01	
	... vs. clean controls	1.122 (1.060, 1.187); p=6.8e-05		1.050 (0.981, 1.124); p=1.6e-01	
	(AR & Grass) vs. clean controls	1.091 (1.009, 1.179); p=2.8e-02			
rs6673480 (Epidermal growth factor)	Unstratified	1.088 (0.978, 1.210); p=1.2e-01		1.318 (1.168, 1.486); p=6.9e-06	

receptor pathway substrate 15)	... without Asthma	1.133 (0.998, 1.287); p=5.5e-02	0.91	1.324 (1.140, 1.538); p=2.3e-04	0.26
	... with Asthma	1.151 (0.913, 1.451); p=2.3e-01		1.551 (1.228, 1.958); p=2.3e-04	
	... without Eczema	1.149 (0.997, 1.325); p=5.5e-02	0.24	1.346 (1.146, 1.582); p=3.0e-04	0.75
	... with Eczema	1.007 (0.851, 1.191); p=9.4e-01		1.293 (1.072, 1.560); p=7.3e-03	
	... in Females	1.055 (0.906, 1.229); p=4.9e-01	0.54	1.201 (0.996, 1.448); p=5.5e-02	0.13
	... in Males	1.128 (0.970, 1.311); p=1.2e-01		1.450 (1.235, 1.703); p=5.5e-06	
	... vs. clean controls	1.092 (0.977, 1.220); p=1.2e-01		1.293 (1.140, 1.466); p=6.1e-05	
	(AR & Grass) vs. clean controls			1.322 (1.145, 1.527); p=1.4e-04	

rs6554809 (DNAH5 (force generating of respiratory cilia))	Unstratified	1.053 (0.973, 1.139); p=2.0e-01		1.272 (1.152, 1.404); p=1.9e-06	
	... without Asthma	1.059 (0.962, 1.166); p=2.4e-01	0.74	1.235 (1.090, 1.398); p=9.0e-04	0.25
	... with Asthma	1.093 (0.928, 1.289); p=2.9e-01		1.401 (1.174, 1.672); p=1.9e-04	
	... without Eczema	1.133 (1.016, 1.263); p=2.5e-02	0.04	1.283 (1.122, 1.467); p=2.7e-04	0.94
	... with Eczema	0.958 (0.848, 1.081); p=4.8e-01		1.273 (1.095, 1.479); p=1.7e-03	
	... in Females	0.999 (0.896, 1.115); p=9.9e-01	0.17	1.304 (1.122, 1.516); p=5.3e-04	0.73
	... in Males	1.117 (0.995, 1.253); p=6.0e-02		1.259 (1.103, 1.438); p=6.6e-04	
	... vs. clean controls	1.079 (0.995, 1.171); p=6.7e-02		1.277 (1.152, 1.414); p=2.8e-06	
	(AR & Grass) vs. clean controls			1.256 (1.117, 1.413); p=1.5e-04	

rs7617456 (30kb to transmembrane protein 108)	Unstratified	1.033 (0.977, 1.093); p=2.5e-01		1.168 (1.092, 1.250); p=6.7e-06	
	... without Asthma	1.072 (1.001, 1.147); p=4.6e-02	0.10	1.179 (1.083, 1.284); p=1.5e-04	0.92
	... with Asthma	0.956 (0.850, 1.075); p=4.5e-01		1.188 (1.052, 1.341); p=5.5e-03	
	... without Eczema	0.999 (0.926, 1.078); p=9.8e-01	0.26	1.107 (1.010, 1.213); p=3.0e-02	0.10
	... with Eczema	1.068 (0.979, 1.165); p=1.4e-01		1.243 (1.121, 1.378); p=3.5e-05	
	... in Females	0.991 (0.918, 1.070); p=8.1e-01	0.14	1.168 (1.056, 1.292); p=2.5e-03	0.96
	... in Males	1.079 (0.994, 1.171); p=6.9e-02		1.164 (1.061, 1.276); p=1.3e-03	
	... vs. clean controls	1.046 (0.987, 1.108); p=1.3e-01		1.165 (1.086, 1.250); p=1.9e-05	
	(AR & Grass) vs. clean controls			1.151 (1.062, 1.248); p=6.2e-04	

rs6586513 (7kb to ciliary rootlet coiled-coil, rootletin)	Unstratified	1.239 (1.129, 1.360); p=6.3e-06		1.095 (0.978, 1.226); p=1.1e-01	
	... without Asthma	1.217 (1.088, 1.362); p=5.9e-04	0.51	1.087 (0.944, 1.253); p=2.4e-01	0.97
	... with Asthma	1.316 (1.073, 1.614); p=8.4e-03		1.092 (0.886, 1.345); p=4.1e-01	
	... without Eczema	1.221 (1.075, 1.386); p=2.1e-03	0.63	1.092 (0.935, 1.274); p=2.7e-01	0.99
	... with Eczema	1.279 (1.106, 1.480); p=9.4e-04		1.090 (0.920, 1.293); p=3.2e-01	
	... in Females	1.199 (1.054, 1.364); p=5.9e-03	0.45	1.007 (0.849, 1.194); p=9.3e-01	0.19
	... in Males	1.288 (1.126, 1.474); p=2.3e-04		1.172 (1.007, 1.364); p=4.1e-02	
	... vs. clean controls	1.237 (1.123, 1.361); p=1.5e-05		1.146 (1.020, 1.287); p=2.2e-02	

(AR & Grass) vs. clean controls		1.169 (1.024, 1.335); p=2.1e-02			
rs1325195 (v-abl Abelson murine leukemia viral oncogene homolog 2)	Unstratified	1.007 (0.953, 1.064); p=8.1e-01		1.173 (1.097, 1.255); p=3.2e-06	
	... without Asthma	1.032 (0.965, 1.104); p=3.5e-01	0.56	1.171 (1.075, 1.274); p=2.8e-04	0.50
	... with Asthma	0.992 (0.882, 1.115); p=8.9e-01		1.231 (1.091, 1.390); p=7.7e-04	
	... without Eczema	1.003 (0.930, 1.081); p=9.5e-01	0.75	1.134 (1.034, 1.243); p=7.3e-03	0.23
	... with Eczema	1.021 (0.937, 1.113); p=6.3e-01		1.232 (1.113, 1.365); p=5.9e-05	
	... in Females	0.984 (0.912, 1.061); p=6.7e-01	0.41	1.132 (1.025, 1.252); p=1.5e-02	0.35
	... in Males	1.031 (0.951, 1.118); p=4.6e-01		1.208 (1.103, 1.324); p=5.1e-05	
	... vs. clean controls	1.020 (0.964, 1.080); p=5.0e-01		1.161 (1.083, 1.245); p=2.8e-05	
(AR & Grass) vs. clean controls		1.162 (1.073, 1.259); p=2.4e-04			

C) Additional candidate loci not included above (SNPS with $5 \times 10^{-6} < P < 1 \times 10^{-4}$ for either allergic rhinitis or grass sensitization)

rs1898671 (Thymic stromal lymphopoietin (TSLP))	Unstratified	1.151 (1.085, 1.222); p=3.4e-06		1.111 (1.034, 1.193); p=3.9e-03	
	... without Asthma	1.093 (1.017, 1.174); p=1.6e-02	0.08	1.082 (0.989, 1.184); p=8.8e-02	0.94
	... with Asthma	1.245 (1.096, 1.414); p=7.7e-04		1.088 (0.956, 1.238); p=2.0e-01	
	... without Eczema	1.185 (1.093, 1.286); p=4.2e-05	0.34	1.165 (1.056, 1.285); p=2.2e-03	0.12
	... with Eczema	1.116 (1.018, 1.224); p=1.9e-02		1.038 (0.932, 1.155); p=5.0e-01	
	... in Females	1.100 (1.014, 1.194); p=2.2e-02	0.11	1.105 (0.994, 1.228); p=6.5e-02	0.79
	... in Males	1.212 (1.111, 1.323); p=1.5e-05		1.126 (1.021, 1.241); p=1.7e-02	
	... vs. clean controls	1.160 (1.090, 1.233); p=2.4e-06		1.137 (1.056, 1.225); p=6.5e-04	
(AR & Grass) vs. clean controls		1.181 (1.085, 1.286); p=1.2e-04			

rs3860069 (Toll-like receptor 6 (TLR6))	Unstratified	1.145 (1.061, 1.235); p=4.6e-04		1.190 (1.087, 1.303); p=1.7e-04	
	... without Asthma	1.118 (1.017, 1.228); p=2.0e-02	0.66	1.163 (1.036, 1.306); p=1.1e-02	0.67
	... with Asthma	1.163 (1.002, 1.349); p=4.7e-02		1.214 (1.037, 1.420); p=1.6e-02	
	... without Eczema	1.179 (1.063, 1.308); p=1.9e-03	0.39	1.240 (1.095, 1.403); p=6.7e-04	0.37
	... with Eczema	1.102 (0.981, 1.237); p=1.0e-01		1.139 (0.994, 1.305); p=6.1e-02	
	... in Females	1.136 (1.023, 1.261); p=1.7e-02	0.86	1.205 (1.052, 1.381); p=7.1e-03	0.75
	... in Males	1.152 (1.031, 1.286); p=1.2e-02		1.170 (1.035, 1.323); p=1.2e-02	
	... vs. clean controls	1.143 (1.057, 1.237); p=7.9e-04		1.201 (1.094, 1.318); p=1.2e-04	
(AR & Grass) vs. clean controls		1.286 (1.151, 1.437); p=8.6e-06			

rs7789045 (NOD1 also known as CARD4)	Unstratified	1.041 (0.984, 1.101); p=1.6e-01		1.147 (1.071, 1.227); p=7.7e-05	
	... without Asthma	1.049 (0.980, 1.123); p=1.6e-01	0.79	1.131 (1.038, 1.231); p=4.7e-03	0.63
	... with Asthma	1.029 (0.913, 1.161); p=6.4e-01		1.173 (1.036, 1.328); p=1.2e-02	
	... without Eczema	1.041 (0.964, 1.124); p=3.1e-01	0.94	1.169 (1.065, 1.283); p=1.0e-03	0.63
	... with Eczema	1.036 (0.950, 1.130); p=4.2e-01		1.130 (1.020, 1.251); p=1.9e-02	

... in Females	0.998 (0.924, 1.078); p=9.6e-01	0.12	1.171 (1.058, 1.296); p=2.2e-03	0.61
... in Males	1.092 (1.006, 1.186); p=3.5e-02		1.130 (1.030, 1.239); p=9.7e-03	
... vs. clean controls	1.046 (0.987, 1.109); p=1.3e-01		1.147 (1.069, 1.230); p=1.3e-04	
(AR & Grass) vs. clean controls			1.148 (1.059, 1.245); p=7.8e-04	

Extra loci (shown in Figure 6)

rs9271300 (Top hit for total IgE from Moffatt et al (NEJM, 2010))	Unstratified	1.063 (1.006, 1.124); p=3.0e-02		1.049 (0.981, 1.122); p=1.6e-01	
	... without Asthma	1.029 (0.962, 1.099); p=4.1e-01	0.45	1.042 (0.957, 1.134); p=3.4e-01	0.63
	... with Asthma	1.085 (0.962, 1.223); p=1.9e-01		1.004 (0.887, 1.136); p=9.5e-01	
	... without Eczema	1.058 (0.982, 1.141); p=1.4e-01	0.86	1.073 (0.979, 1.176); p=1.3e-01	0.49
	... with Eczema	1.070 (0.980, 1.168); p=1.3e-01		1.023 (0.923, 1.134); p=6.7e-01	
	... in Females	1.020 (0.946, 1.100); p=6.1e-01	0.12	1.000 (0.906, 1.104); p=1.0e+00	0.17
	... in Males	1.114 (1.026, 1.209); p=1.0e-02		1.100 (1.002, 1.207); p=4.5e-02	
	... vs. clean controls	1.062 (1.003, 1.124); p=4.0e-02		1.066 (0.994, 1.142); p=7.2e-02	
	(AR & Grass) vs. clean controls		1.087 (1.004, 1.177); p=3.9e-02		

rs7927894 (Top hit for atopic dermatitis Esparza-Gordillo et al (NG, 2009))	Unstratified	1.151 (1.085, 1.222); p=3.4e-06		1.187 (1.106, 1.274); p=2.0e-06	
	... without Asthma	1.131 (1.052, 1.215); p=8.6e-04	0.94	1.189 (1.088, 1.300); p=1.3e-04	0.47
	... with Asthma	1.125 (0.992, 1.276); p=6.7e-02		1.122 (0.986, 1.277); p=8.2e-02	
	... without Eczema	1.144 (1.055, 1.240); p=1.1e-03	0.82	1.152 (1.046, 1.268); p=4.0e-03	0.44
	... with Eczema	1.161 (1.058, 1.274); p=1.7e-03		1.220 (1.095, 1.358); p=3.0e-04	
	... in Females	1.206 (1.111, 1.309); p=7.1e-06	0.12	1.223 (1.101, 1.358); p=1.7e-04	0.55
	... in Males	1.098 (1.006, 1.198); p=3.7e-02		1.171 (1.063, 1.290); p=1.4e-03	
	... vs. clean controls	1.168 (1.098, 1.242); p=7.9e-07		1.224 (1.138, 1.317); p=5.9e-08	
	(AR & Grass) vs. clean controls		1.253 (1.152, 1.363); p=1.4e-07		

290 **Table E5** : Sample size corresponding to the analyses presented in Table E4.

Analysis/subgroup	<i>Allergic Rhinitis</i>		<i>Grass Sensitization</i>	
	# controls	# cases	# controls	# cases
Unstratified	8330	3742	9809	2263
... without Asthma	7674	2487	8297	1372
... with Asthma	1203	1477	1644	926
... without Eczema	6170	1925	6469	1198
... with Eczema	2745	2051	3460	1102
... in Females	4532	2155	5278	1005
... in Males	4396	1823	4705	1305
... vs clean control	7690	3742	7690	2263
(AR & Grass) vs. clean controls	# controls = 7690		# cases = 1623	

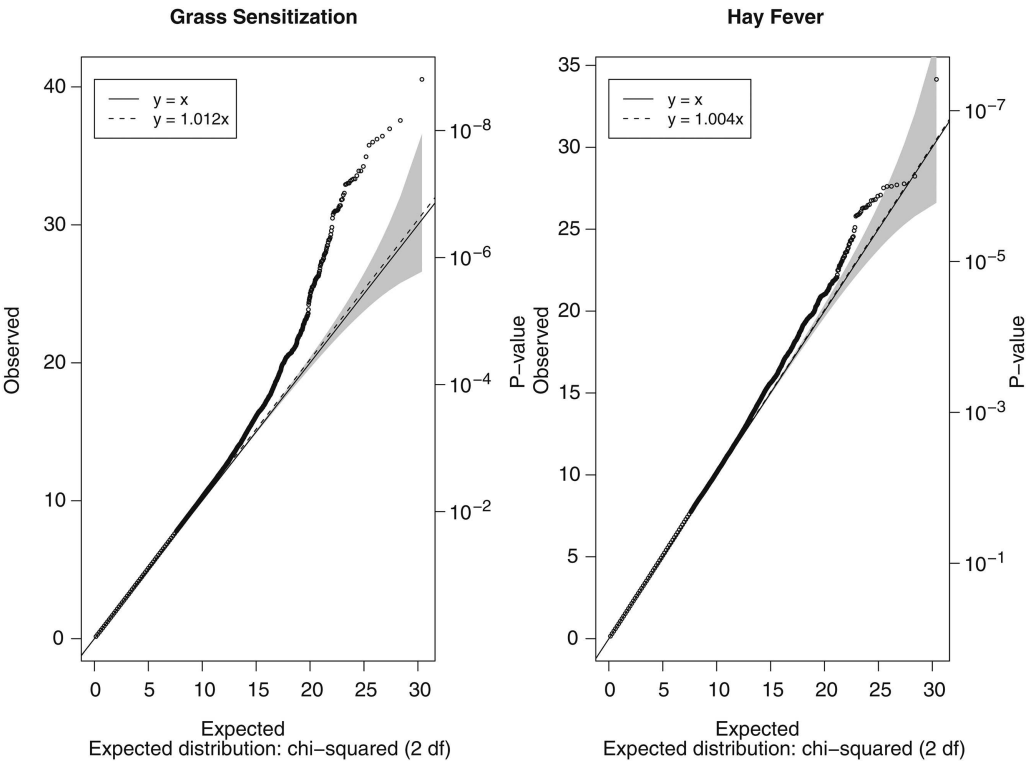


FIG E1. Q-Q plots for grass sensitization and AR.

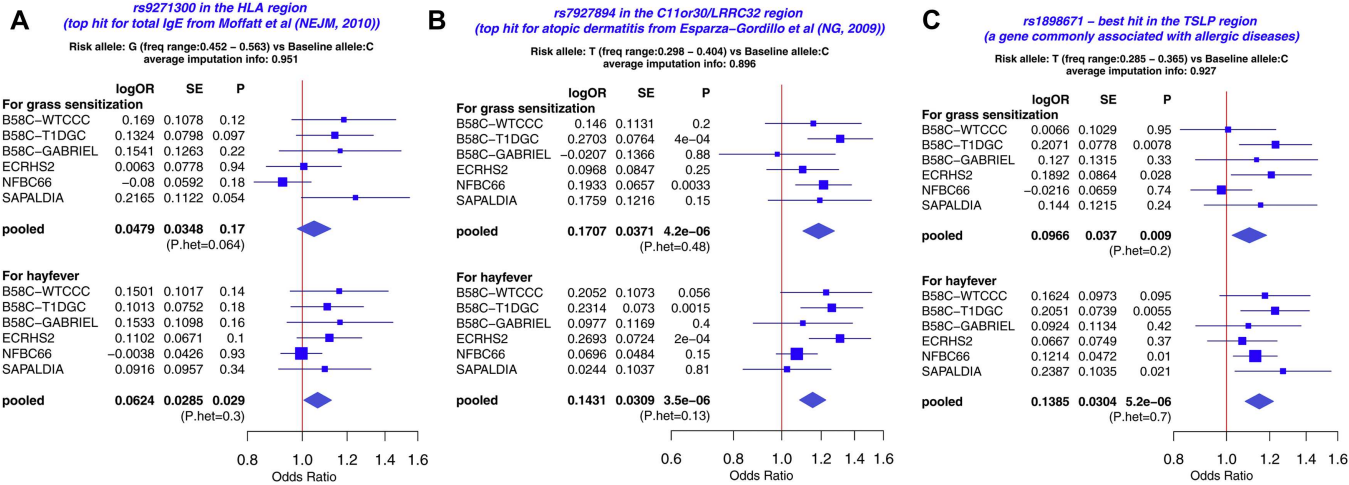


FIG E2. Forest plots (A) rs9271300 (B) rs7927894 and (C) rs1898671 which were previously shown in the literature to be associated with related phenotypes. OR, Odds ratio.

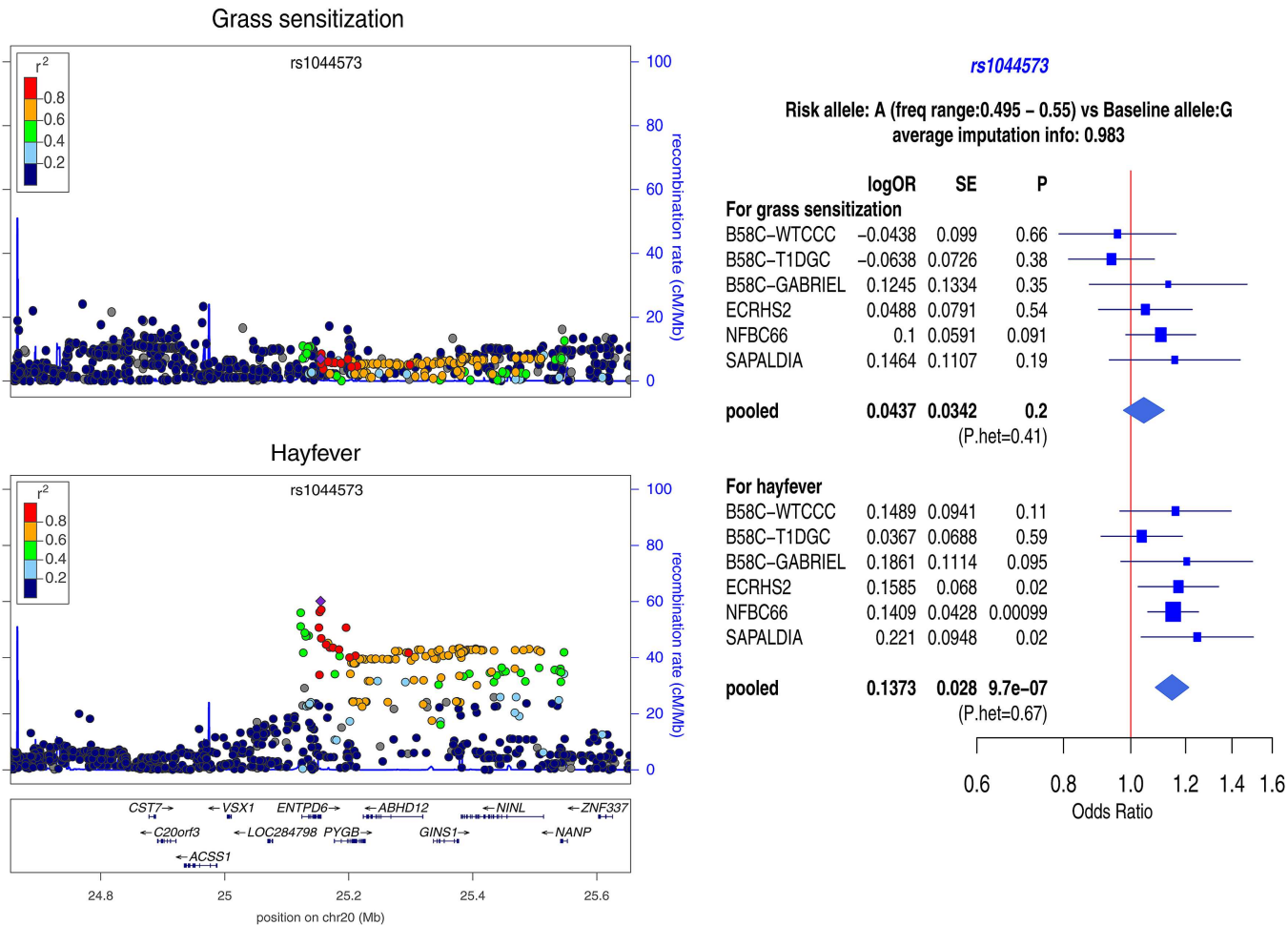


FIG E3. Regional association and forest plots for rs1044573 in *ENTPD6* (previously IL-6 signal transducer). OR, Odds ratio.

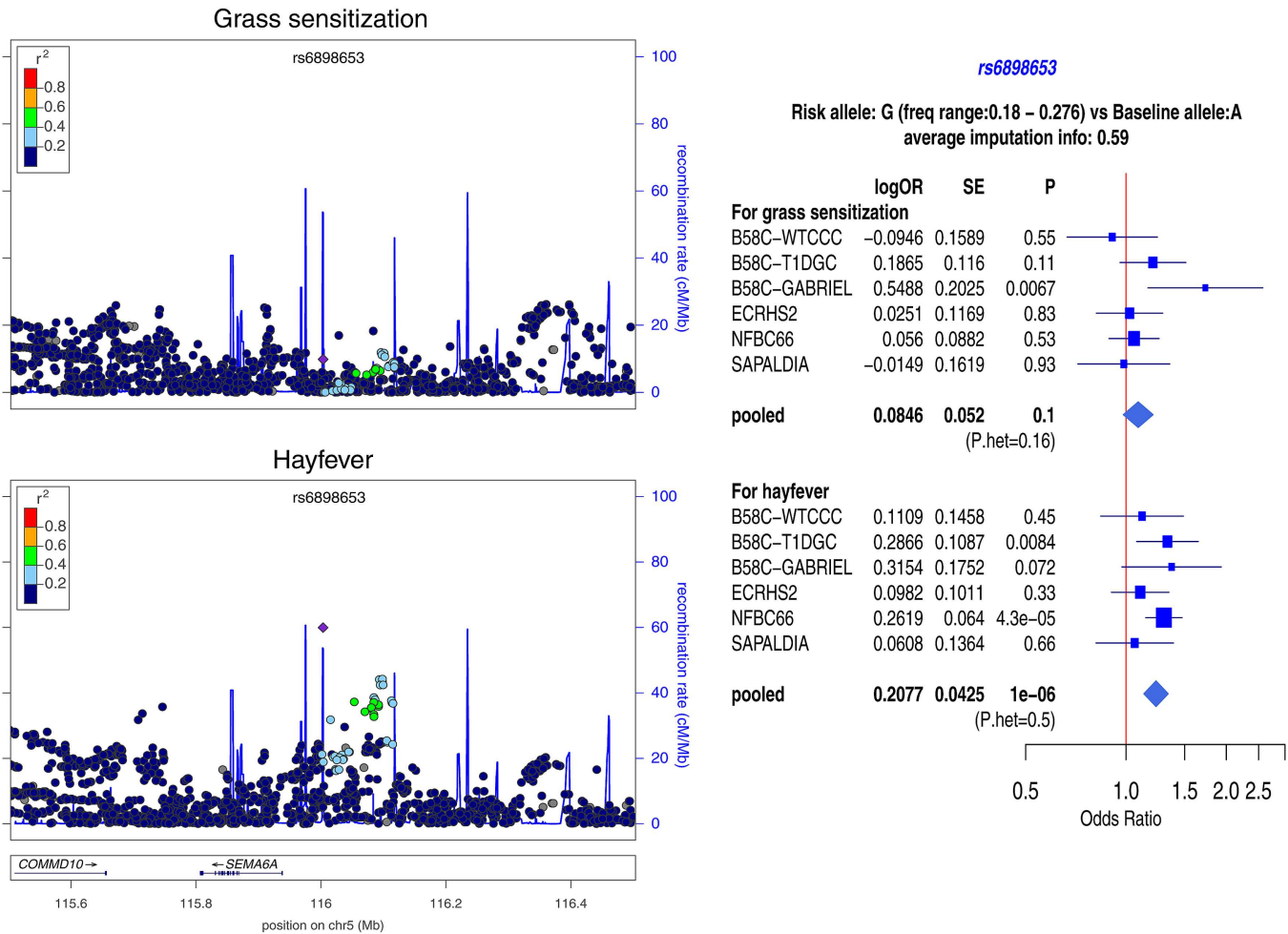


FIG E4. Regional association and forest plots for rs6898653, which is 70 kb to SEMA6A. OR, Odds ratio.

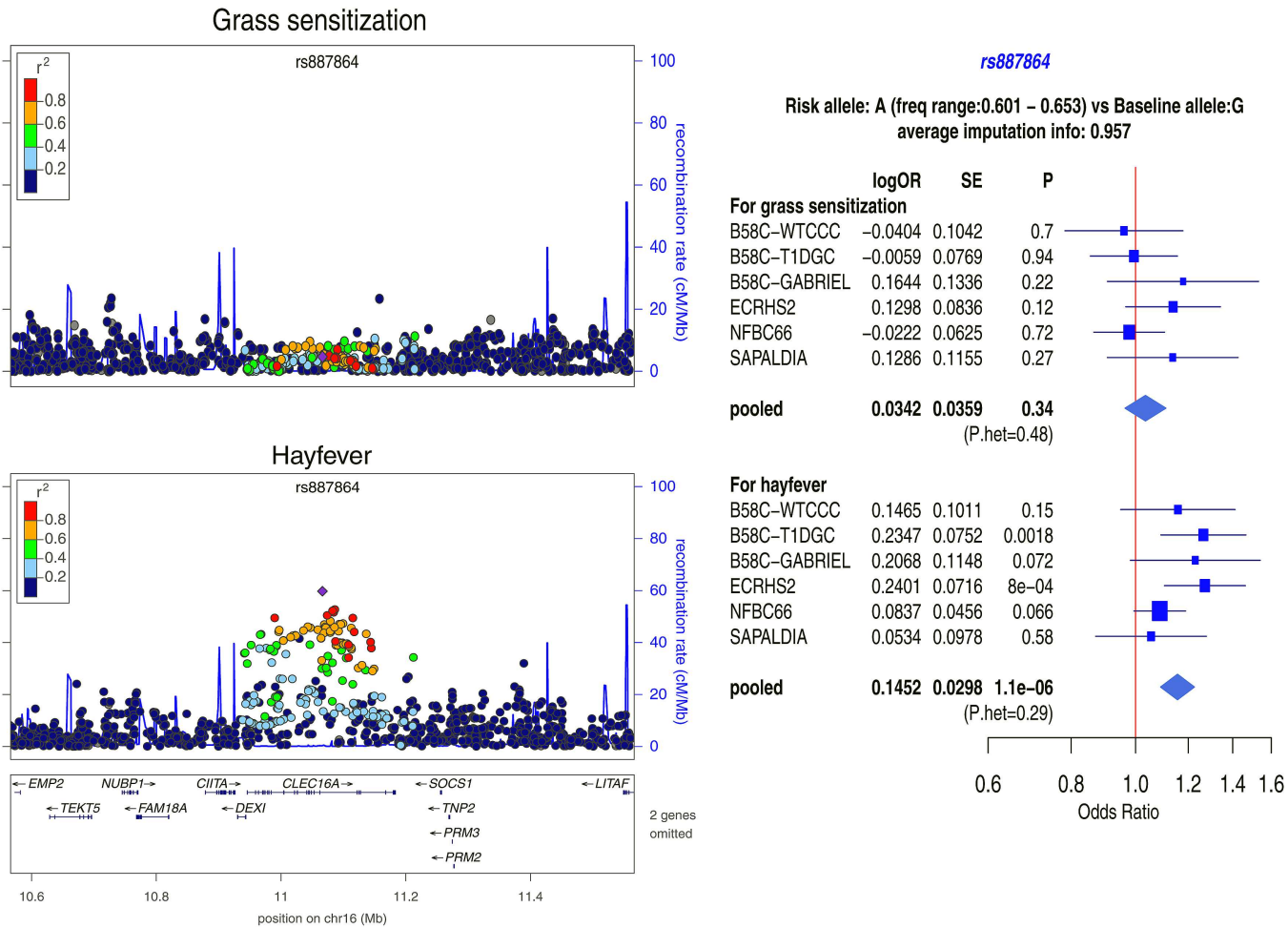


FIG E5. Regional association and forest plots for rs887864 in C-type lectin domain family 16, member A (CLEC16A). OR, Odds ratio.

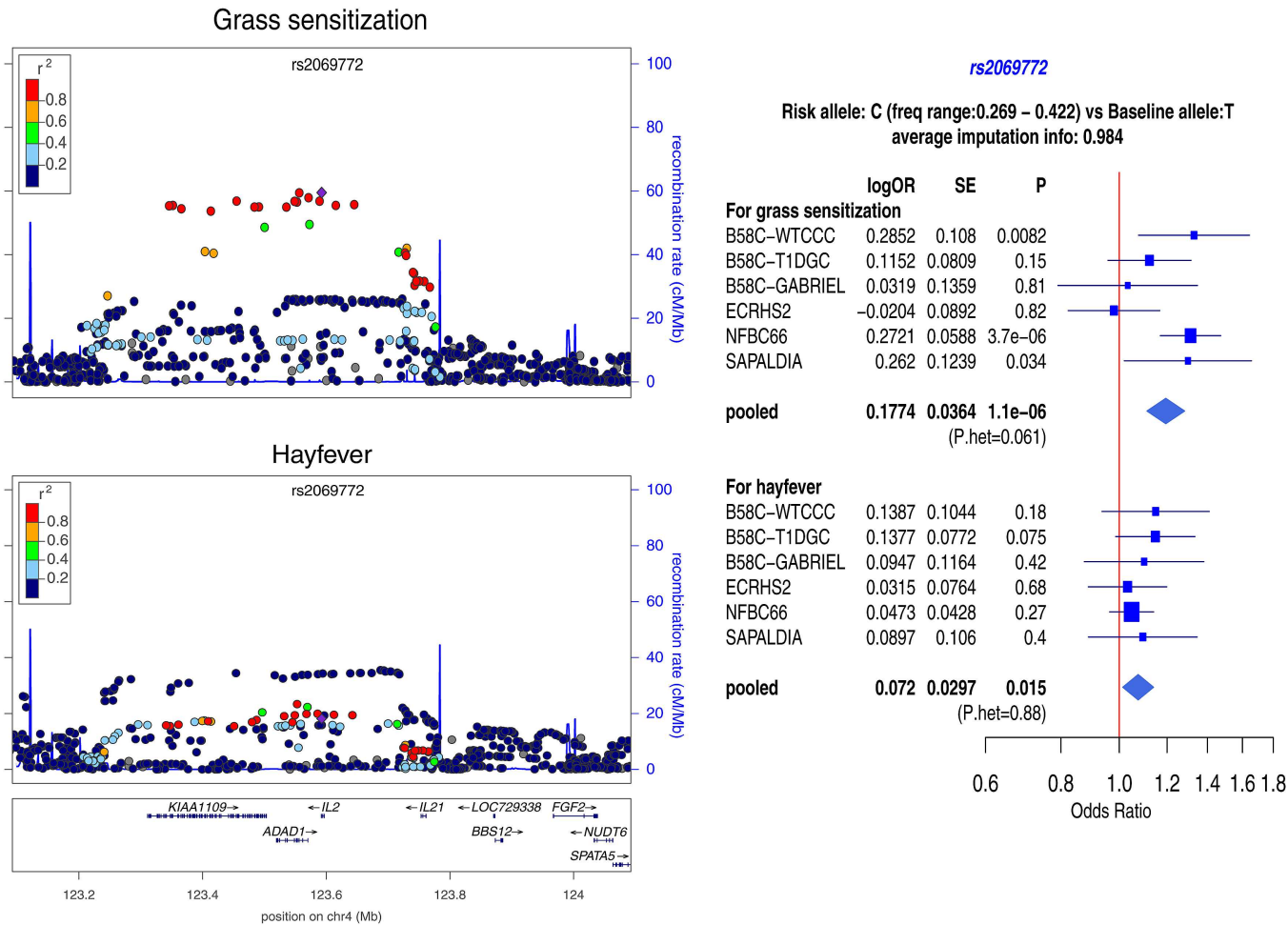


FIG E6. Regional association and forest plots for rs2069772 in IL2. OR, Odds ratio.

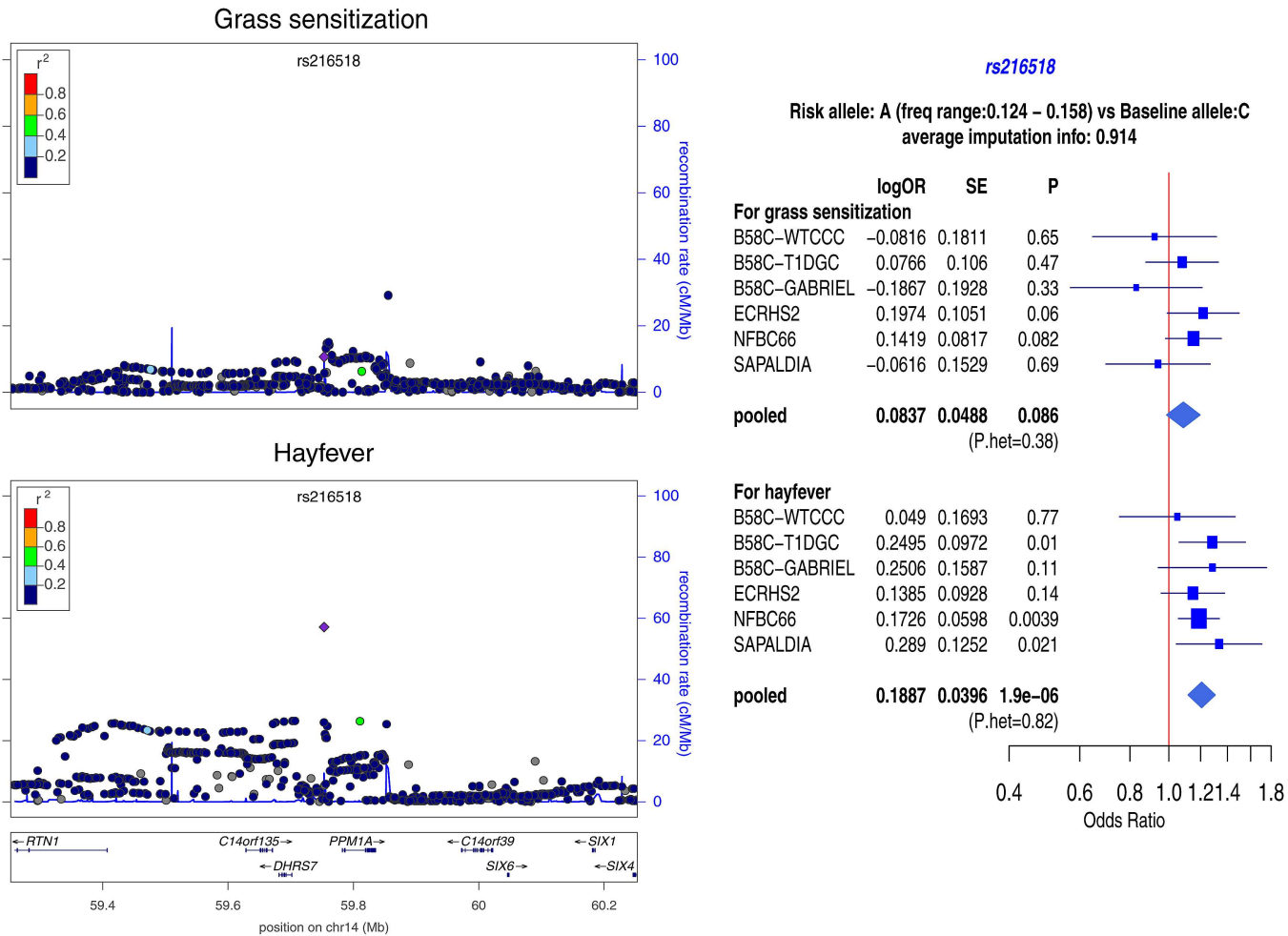


FIG E7. Regional association and forest plots for rs216518 near *PPM1A* and *DHR7* (dehydrogenase/reductase). OR, Odds ratio.

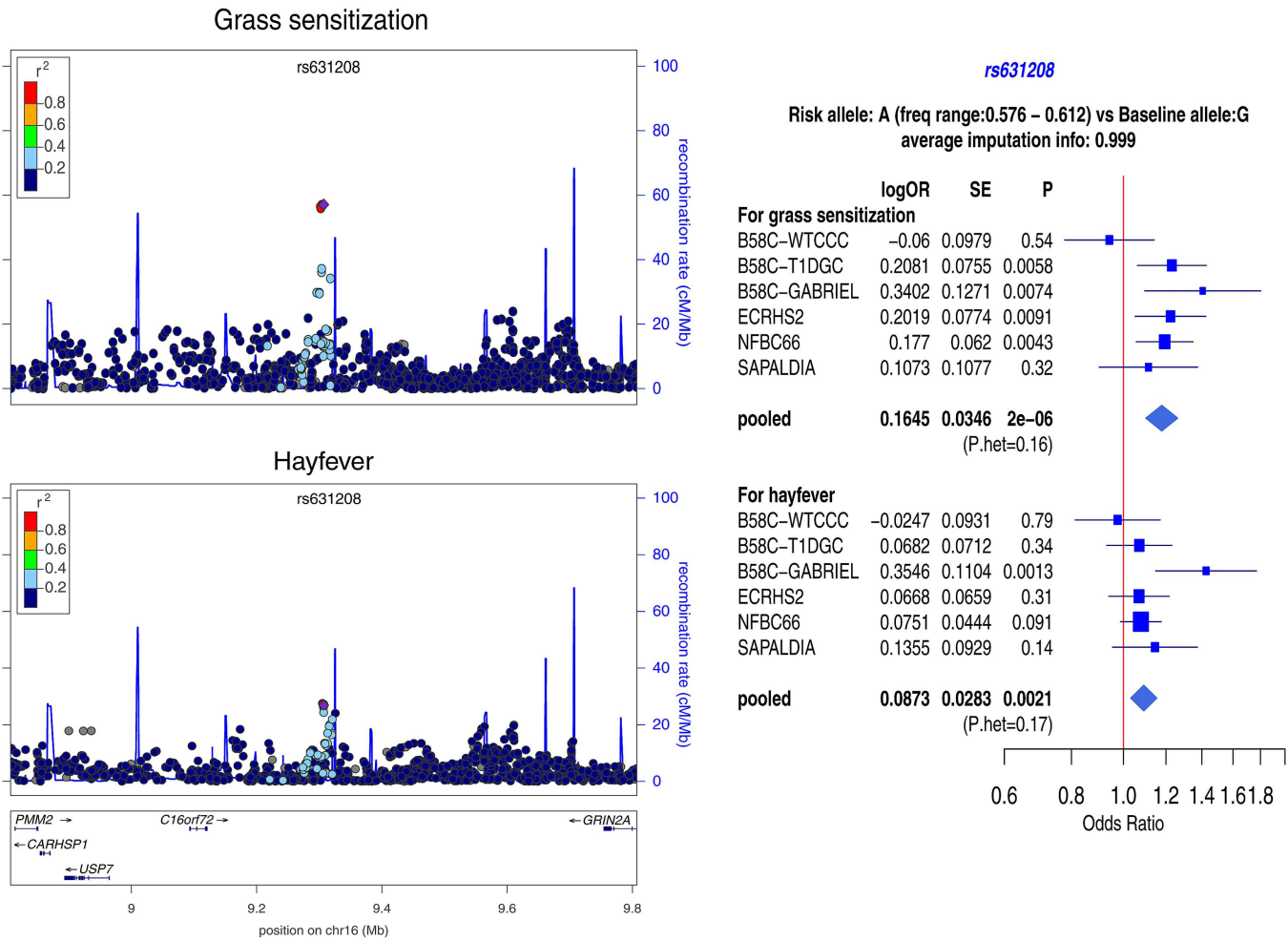


FIG E8. Regional association and forest plots for rs631208 (intergenic region). OR, Odds ratio.

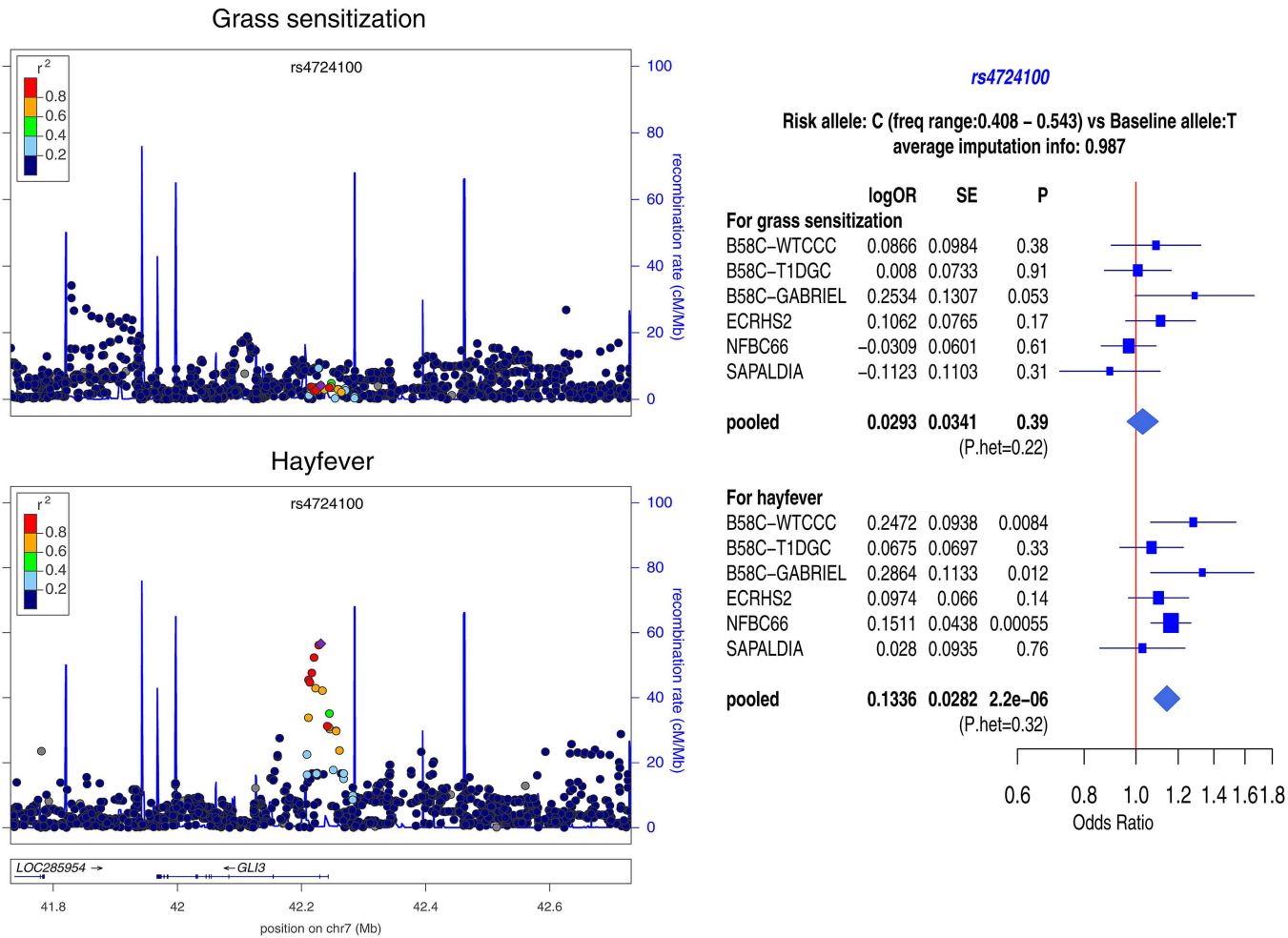


FIG E9. Regional association and forest plots for rs4724100 in GLI family zinc finger 3 (*GLI3*). OR, Odds ratio.

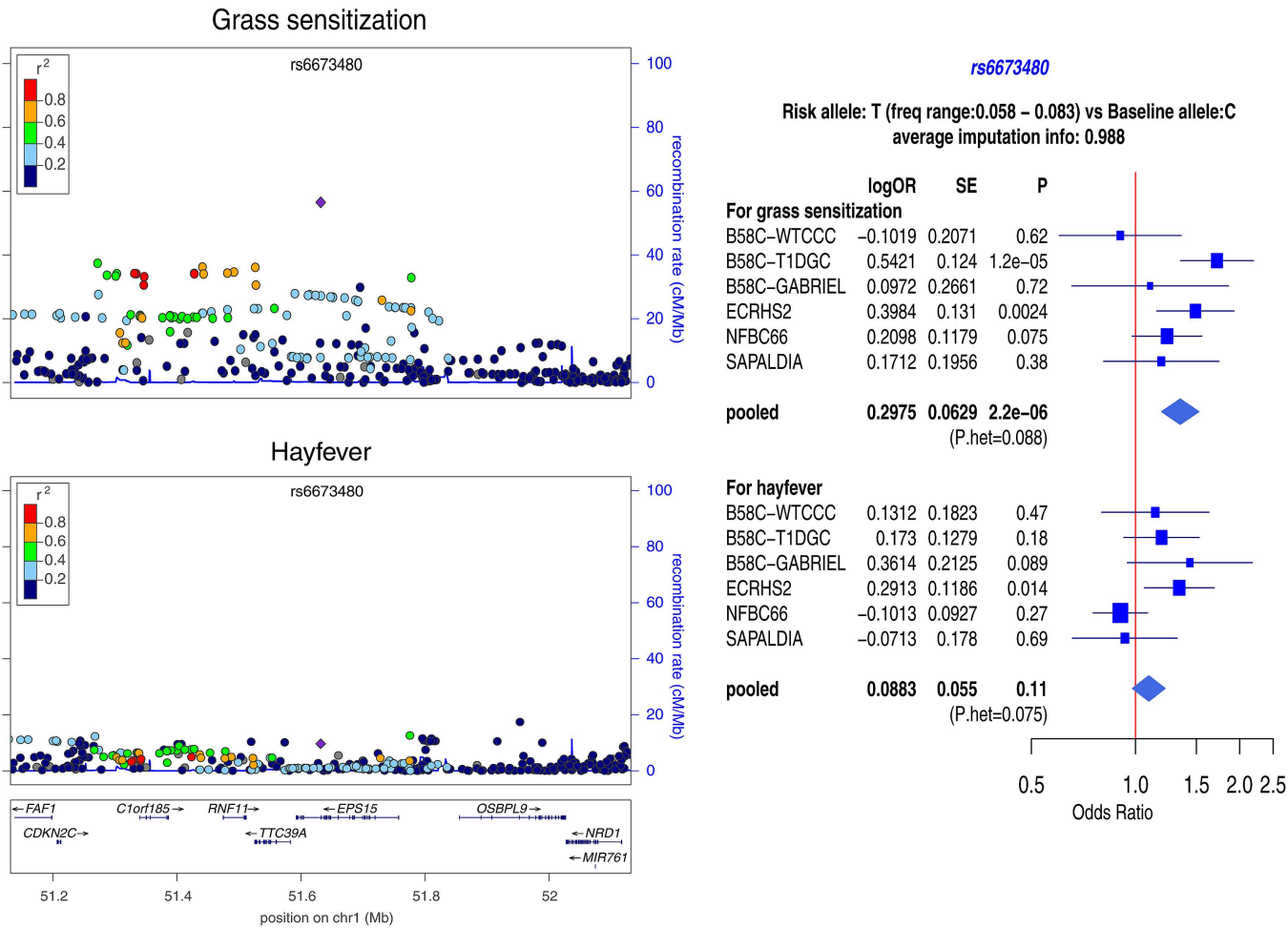


FIG E10. Regional association and forest plots for rs6673480 in epidermal growth factor receptor pathway substrate 15 (*EPS15*). *OR*, Odds ratio.

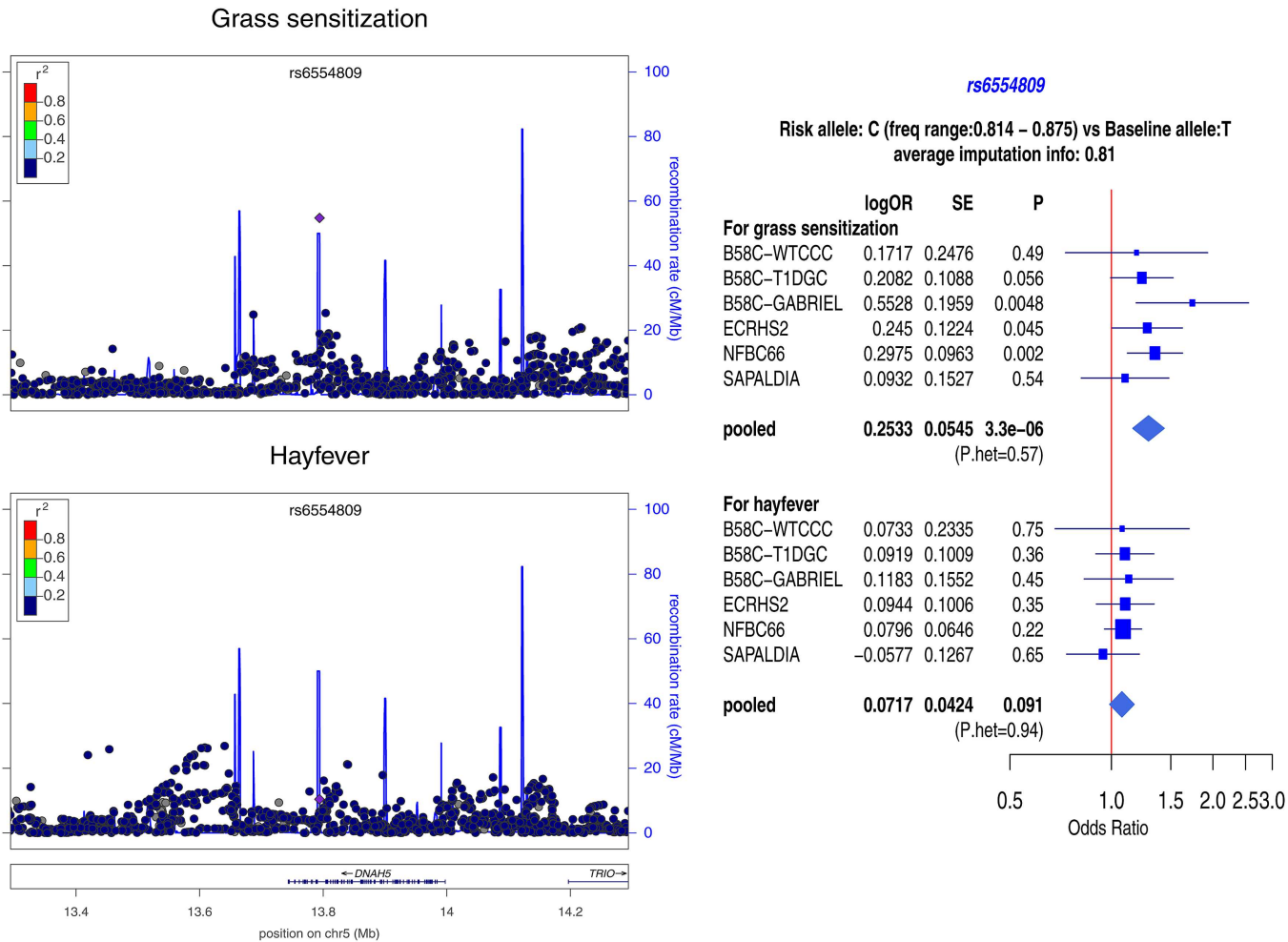


FIG E11. Regional association and forest plots for rs6554809 in *DNAH5* (force generating of respiratory cilia). *OR*, Odds ratio.

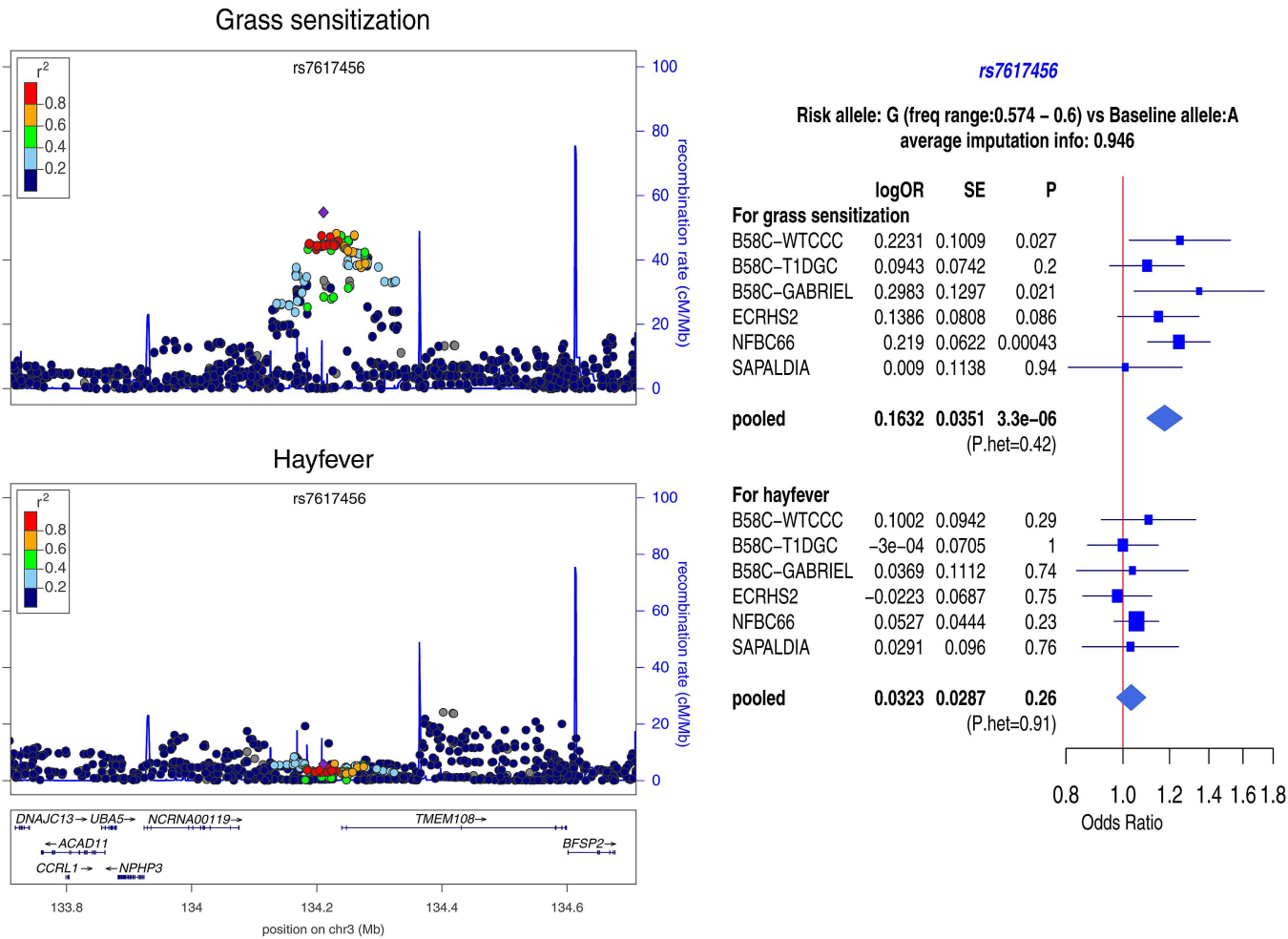


FIG E12. Regional association and forest plots for rs7617456, which is 30 kb to transmembrane protein 108 (*TMEM108*). OR, Odds ratio.

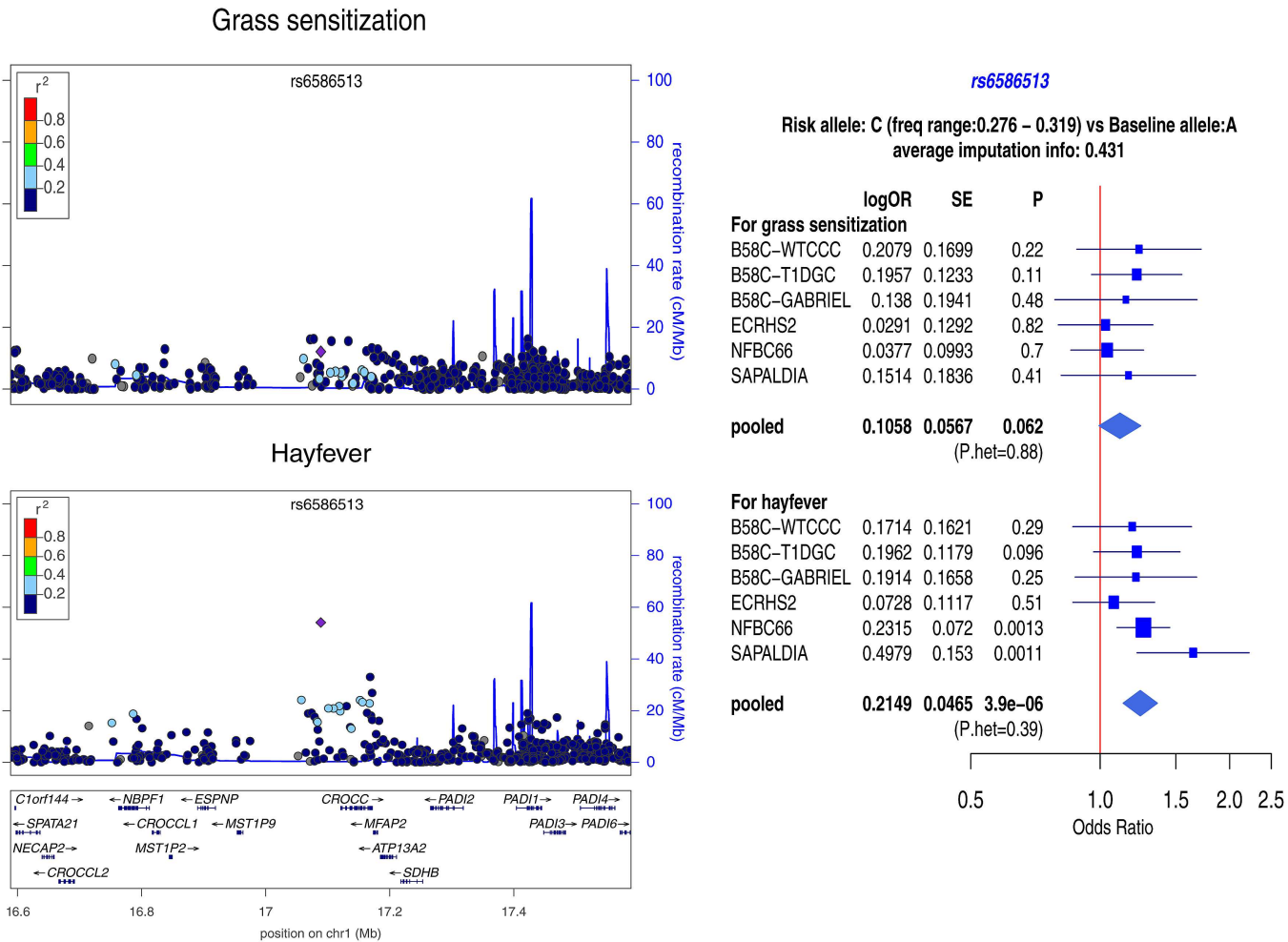


FIG E13. Regional association and forest plots for rs6586513, which is 7 kb to ciliary rootlet coiled-coil, rootletin (*CROCC*). OR, Odds ratio.

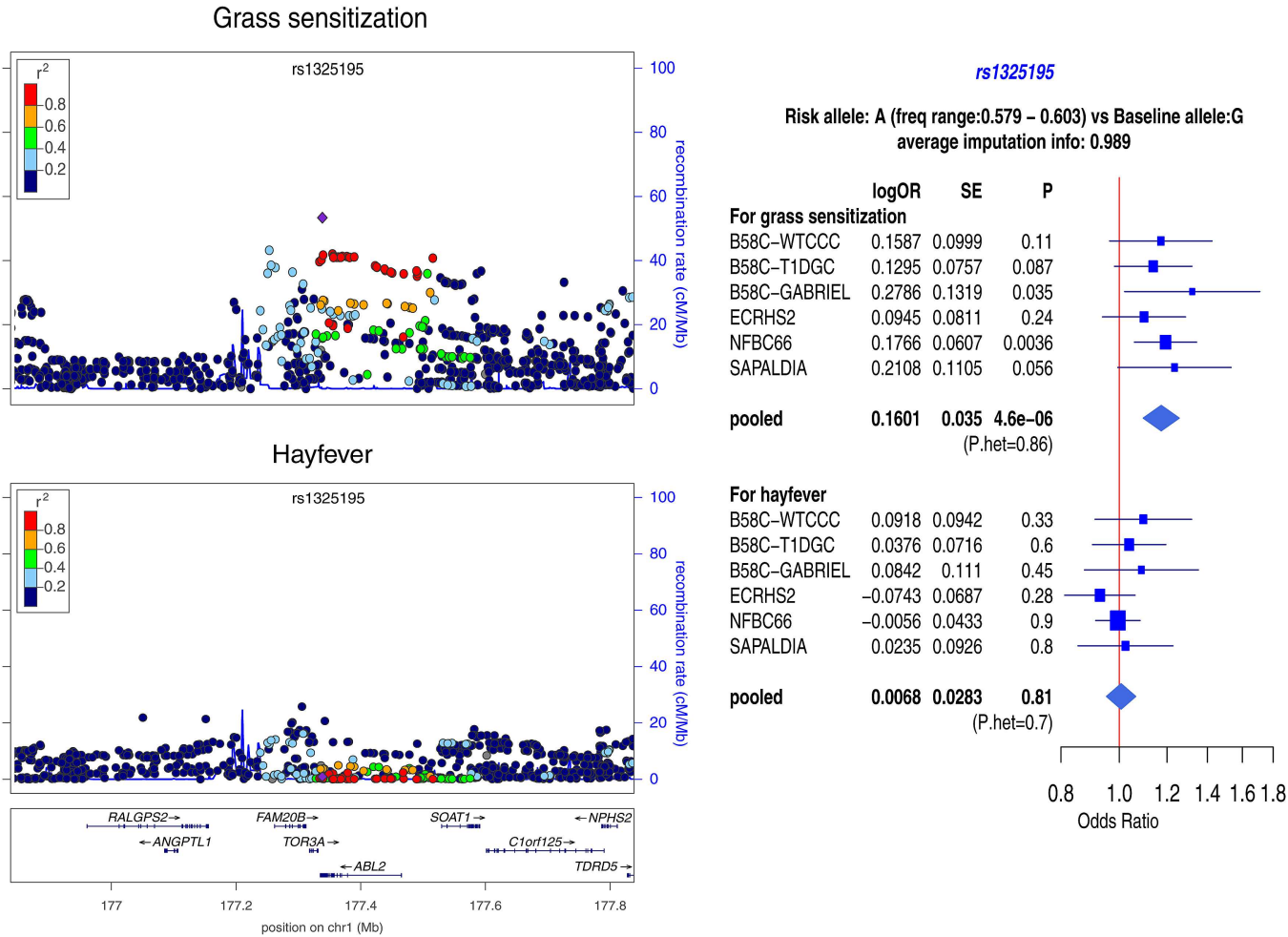


FIG E14. Regional association and forest plots for rs1325195 in v-abl Abelson murine leukemia viral oncogene homolog 2 (*ABL2*). OR, Odds ratio.

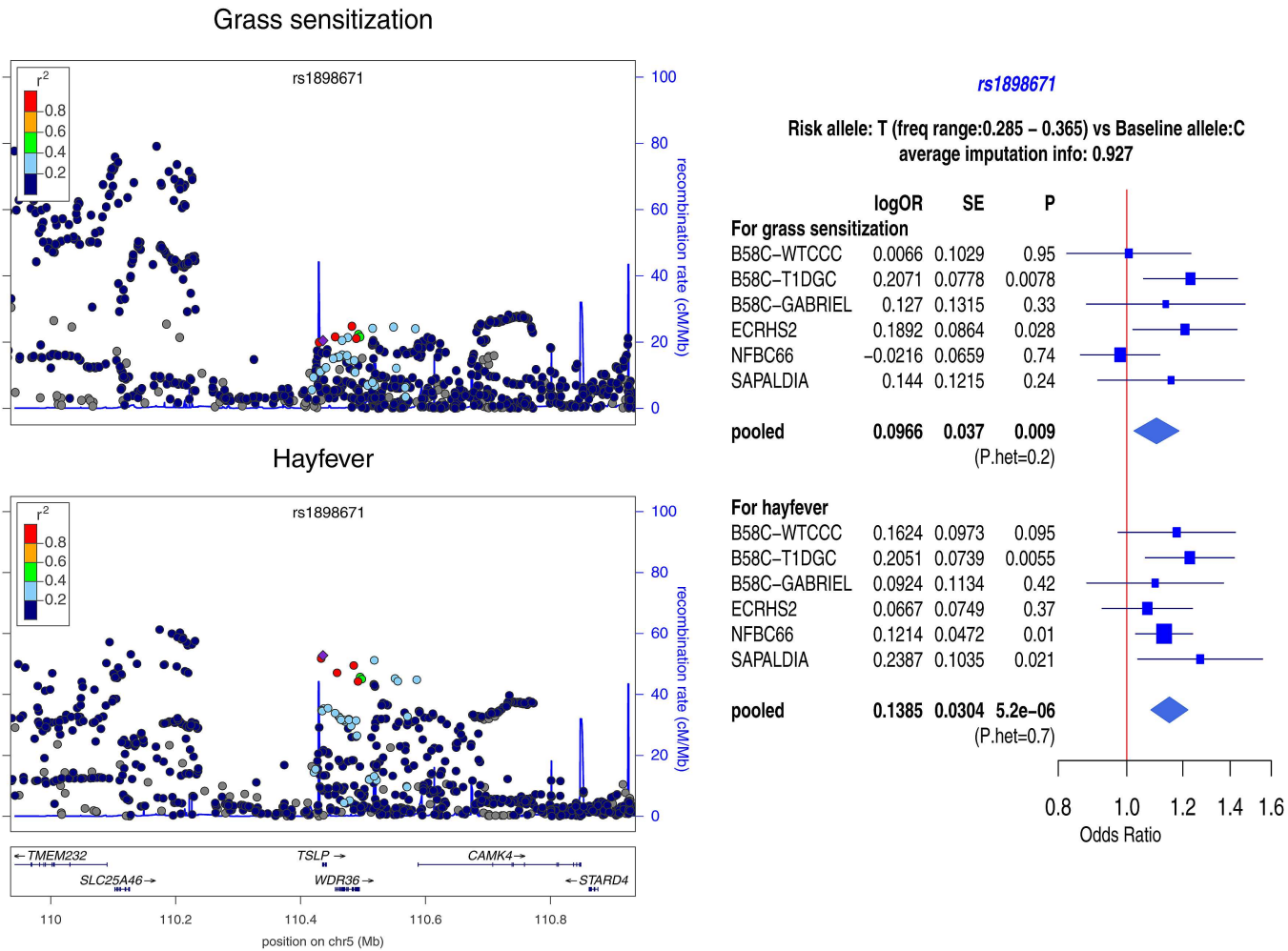


FIG E15. Regional association and forest plots for rs1898671 in *TSLP*. OR, Odds ratio.

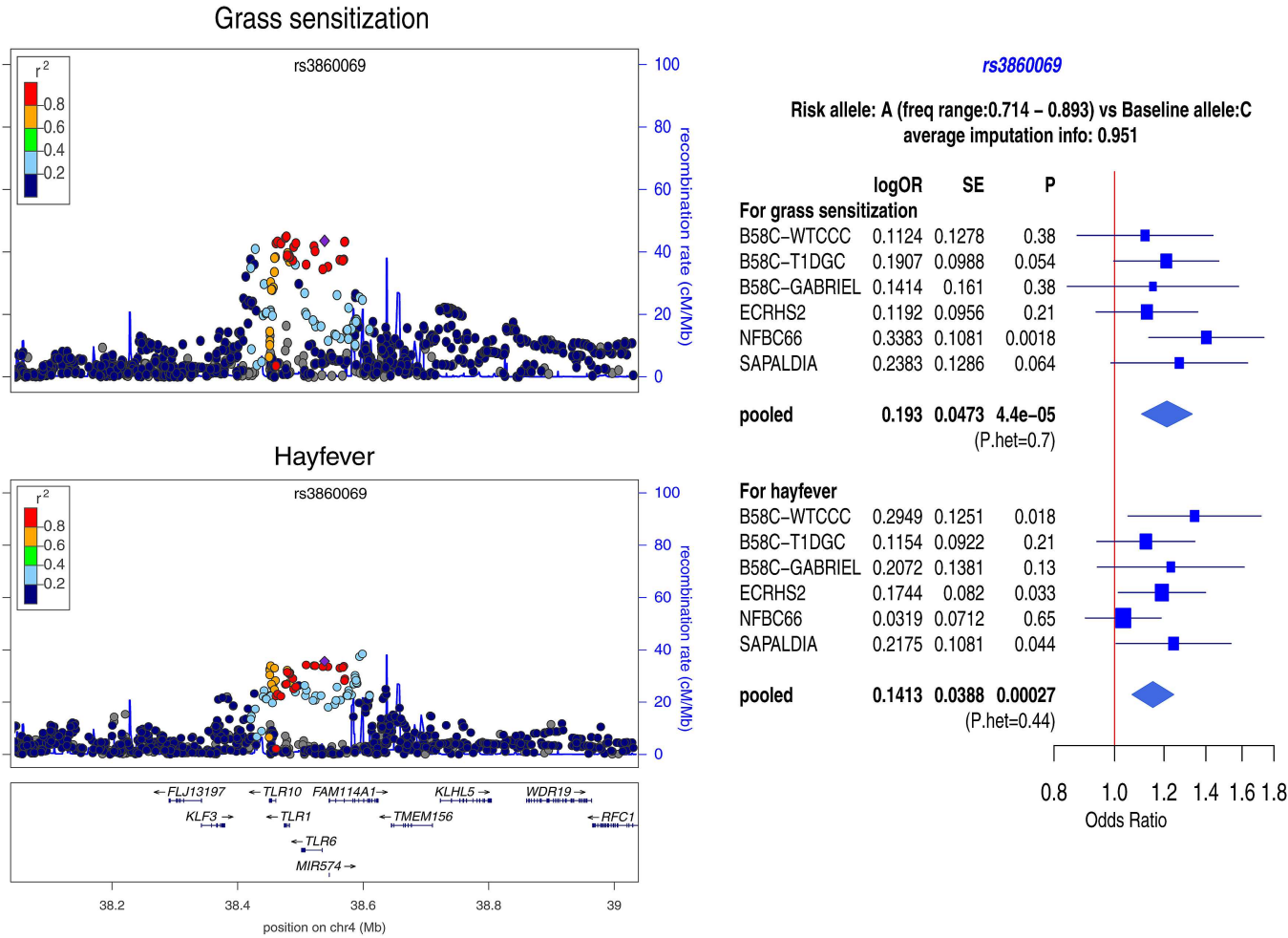


FIG E16. Regional association and forest plots for rs3860069 in *TLR6*. OR, Odds ratio.

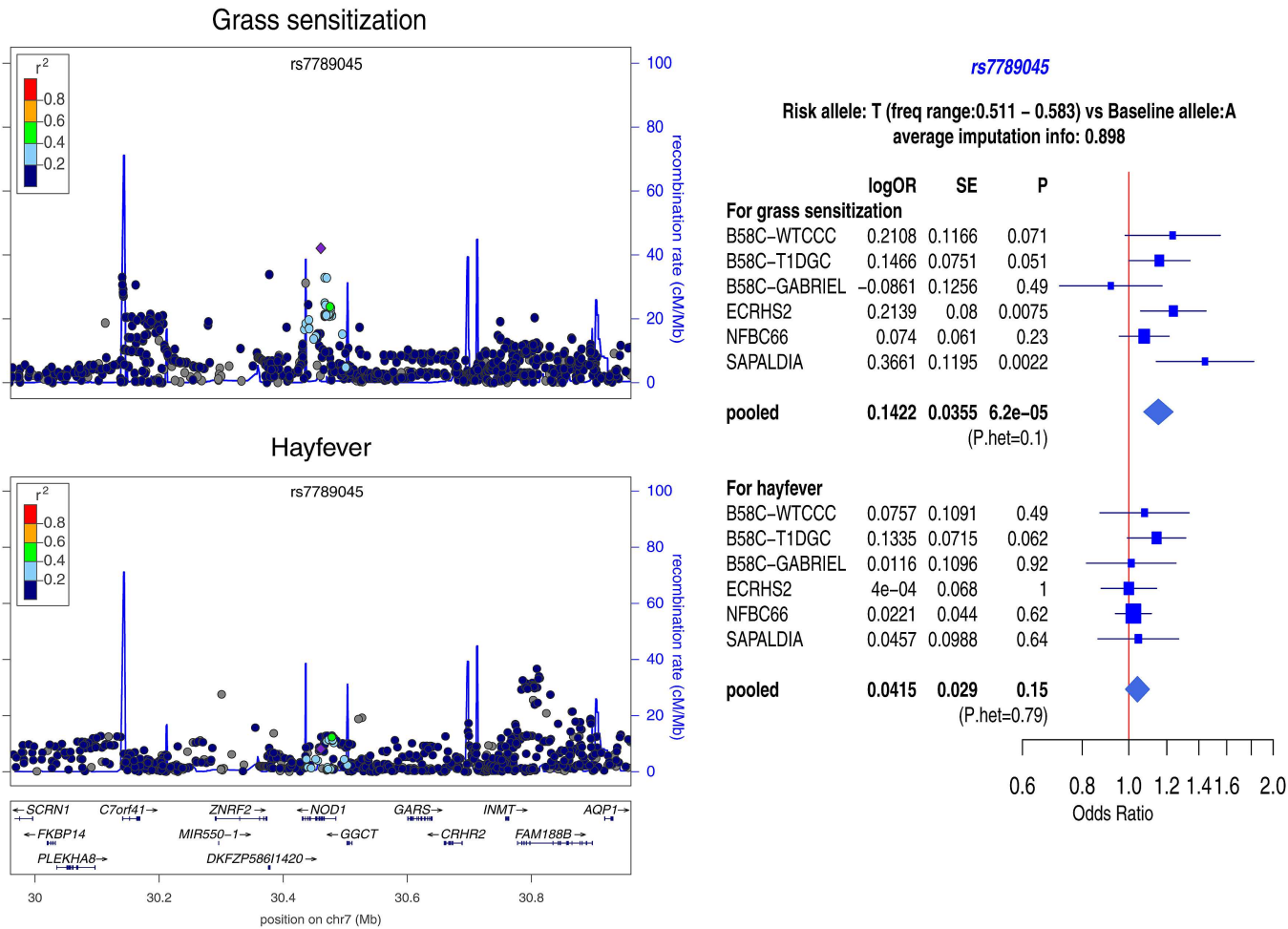


FIG E17. Regional association and forest plots for rs7789045 in *NOD1*, previously *CARD4*. OR, Odds ratio.