Genetic heterogeneity of asthma phenotypes identified by a clustering approach

Valérie Siroux[1,2,23], Juan R. González[3,4,5,23], Emmanuelle Bouzigon[6,7], Ivan Curjuric[8,9], Anne Boudier[1,2], Medea Imboden[8,9], Josep Maria Anto[3,5,10,11], Ivo Gut[12,13], Deborah Jarvis[14], Mark Lathrop[7,12], Ernst Reidar Omenaas[15,16], Isabelle Pin[1,2,17], Mathias Wjst[18,19], Florence Demenais[6,7], Nicole Probst-Hensch[8,9], Manolis Kogevinas[3,5,11,20] and Francine Kauffmann[21,22]

Affiliations: [1]Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Inserm, U823, Grenoble, [2]Université Joseph Fourier, Grenoble, [6]Inserm, UMRS-946, Paris, [7]Université Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, Paris, [12]CEA-Centre National de Genotypage, Evry, [17]Dept of Pediatrics, CHU Grenoble, Grenoble, [21]Inserm, U1018, CESP Centre for Research in Epidemiology and Population Health, Respiratory and Environmental Epidemiology Team, Villejuif, and [22]Université Paris Sud, UMRS 1018, Villejuif, France. [3]Centre for Research in Environmental Epidemiology (CREAL), Barcelona, [4]Dept of Mathematics, Autonomous University of Barcelona, Barcelona, [5]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, [10]Universitat Pompeu Fabra, Dept de Ciències Experimentals i de la Salut, Barcelona, [11]IMIM (Hospital del Mar Medical Research Institute), Barcelona, and [13]Centro Nacional de Analisis Genomico, Barcelona, Spain. [8]Swiss Tropical and Public Health Institute, Basel, and [9]University of Basel, Basel, Switzerland. [14]Respiratory Epidemiology and Public Health, Imperial College, London, and MRC-HPA Centre for Environment and Health, London, UK. [15]Bergen Respiratory Research Group, Institute of Medicine, University of Bergen, Bergen, and [16]Center for Clinical Research, Haukeland University Hospital, Bergen, Norway. [18]Comprehensive Pneumology Center (CPC), Helmholtz Zentrum Muenchen, German Research Center for Environmental Health (GmbH), Neuherberg, and [19]Institute of Medical Statistics and Epidemiology, Klinikum rechts der Isar der TU Muenchen, Munchen, Germany. [20]National School of Public Health, Athens, Greece. [23]Both authors contributed equally.

Correspondence: V. Siroux, Centre de Recherche INSERM/UJF U823, Institut Albert Bonniot, BP 170, 38042 Grenoble Cedex 9, France. E-mail: valerie.siroux@ujf-grenoble.fr

ABSTRACT The aim of the study was to identify genetic variants associated with refined asthma phenotypes enabling multiple features of the disease to be taken into account.

Latent class analysis (LCA) was applied in 3001 adults ever having asthma recruited in the frame of three epidemiological surveys (the European Community Respiratory Health Survey (ECRHS), the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA) and the Epidemiological Study on the Genetics and Environment of Asthma (EGEA)). 14 personal and phenotypic characteristics, gathered from questionnaires and clinical examination, were used. A genome-wide association study was conducted for each LCA-derived asthma phenotype, compared to subjects without asthma (n=3474).

The LCA identified four adult asthma phenotypes, mainly characterised by disease activity, age of asthma onset and atopic status. Associations of genome-wide significance ($p<1.25\times10^{-7}$) were observed between "active adult-onset nonallergic asthma" and rs9851461 flanking *CD200* (3q13.2) and between "inactive/mild nonallergic asthma" and rs2579931 flanking *GRIK2* (6q16.3). Borderline significant results ($2.5\times10^{-7}<p<8.2\times10^{-7}$) were observed between three single nucleotide polymorphisms (SNPs) in the *ALCAM* region (3q13.11) and "active adult-onset nonallergic asthma". These results were consistent across studies. 15 SNPs identified in previous genome-wide association studies of asthma have been replicated with at least one asthma phenotype, most of them with the "active allergic asthma" phenotype.

Our results provide evidence that a better understanding of asthma phenotypic heterogeneity helps to disentangle the genetic heterogeneity of asthma.

@ERSpublications
Our data show that the genetic heterogeneity of asthma may be elucidated by clarifying asthma phenotypic heterogeneity http://ow.ly/r5ljh

Introduction

Recent genetic studies, including meta-analyses of large-scale genome-wide association studies (GWAS), have successfully identified several genetic loci that influence asthma susceptibility, providing a better understanding of the pathogenesis of this complex disorder [1, 2]. However, only a small proportion of heritability can be explained by the previously identified single nucleotide polymorphisms (SNPs) associated with asthma [3, 4]. The missing heritability could partly reside in the phenotypic heterogeneity of asthma, not taken into account in genetic studies.

Asthma is a heterogeneous disease constituting separate overlapping syndromes probably with different, but as yet undefined, aetiologies and natural histories [5]. Childhood- and adult-onset asthma are among the most commonly accepted phenotypes. Interestingly, 17q21 genetic variants were specifically associated with childhood onset asthma in the French Epidemiological Study on the Genetics and Environment of Asthma (EGEA), a result further confirmed by a large GWAS conducted within the GABRIEL consortium [1, 6]. Such results provided the first evidence for a genetic heterogeneity of asthma phenotypes. Unsupervised models aiming to identify homogeneous subgroups of subjects have been applied to unravel the phenotypic heterogeneity of asthma [7]. In adult asthma, these studies led to the identification of asthma phenotypes that exhibited differences in clinical response to treatment, in clinical, physiological and inflammatory parameters and in health-related quality of life [5, 8–10]. Although such refinement of asthma characterisation may shed light on asthma genetics, no genetic association studies have been conducted to date on asthma phenotypes defined by a clustering approach.

We aimed to identify genetic variants associated with cluster-derived asthma phenotypes in a large set of subjects recruited in three large epidemiological studies: the European Community Respiratory Health Survey (ECRHS), the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA) and the EGEA (comprising the ESE consortium).

Methods

Further information is provided in the online supplementary material.

Study populations

The ECRHS study is a European population-based study of young adults with an 8-year follow-up (ECRHSI (1991–1993) n=18 356; ECRHSII (1999–2002), n=10 933) [11]. SAPALDIA is a cohort study in the Swiss population initiated in 1991 (SAPALDIA1, n=9651) with a follow-up assessment in 2002 (SAPALDIA2, n=8047) [12]. The EGEA is a French case–control and family-based study with a 12-year follow-up investigation (EGEA1 (1991–1995) n=2047; EGEA2 (2003–2007), n=1601) [13]. Similar protocols, questionnaires and clinical examination were used in the three studies.

Cluster analysis in adult subjects with ever asthma

In ECRHSII and SAPALDIA2, subjects with asthma answered positively to ''Have you ever had asthma?''. In EGEA2, asthma was defined by a positive answer to ''Have you ever had attacks of breathlessness at rest with wheezing?'' or ''Have you ever had asthma attacks?'' or being recruited as an asthma case in chest clinics.

We first performed a latent class analysis (LCA) in 3001 adults who had ever had asthma (ECRHSII, n=1895; SAPALDIA2, n=465; EGEA2, n=641), irrespective of the availability of genotypes, to define asthma phenotypes. 14 variables covering personal characteristics (age and sex), asthma characteristics (age at asthma onset and asthma exacerbations), respiratory symptoms over the past 12 months, allergic characteristics, lung function and bronchial hyperresponsiveness (BHR) were considered in the LCA model. Asthma treatment was not included because of the lack of detailed information in the SAPALDIA survey, but sensitivity analyses conducted in ECRHS and EGEA showed that the model including treatment leads to similar clusters.

Genotypic data

Next, we conducted genetic analyses on subjects with genotypic information: 1689 subjects with asthma and 3452 controls without asthma. Genotyped data were available for almost the whole EGEA population.

This article has supplementary material available from www.erj.ersjournals.com

Received: Feb 21 2013 | Accepted after revision: July 30 2013 | First published online: Dec 05 2013

Support statement: A list of supporting bodies of the EGEA, ECRHS II and SAPALDIA can be found in the acknowledgements section.

Conflict of interest: Disclosures can be found alongside the online version of this article at www.erj.ersjournals.com

DOI: 10.1183/09031936.00032713

In SAPALDIA and ECRHS the sample included in the genetic analysis represents a nested asthma case–control sample from the cohort (all subjects with asthma at baseline or follow-up with DNA and a random sample of controls).

The subjects were genotyped within the framework of the European GABRIEL consortium. Genotyping was carried out using the Human610 quad array (Illumina; San Diego, CA, USA), at the French national genotyping centre (Centre National de Genotypage, Evry, France). After quality control of genotyping, as previously described, the number of SNPs analysed was 499 138 [14]. The 39 candidate genetic loci included were those identified in previously published GWAS for asthma [1, 2, 15–24]. If the reported SNP was not genotyped in our data, the closest proxy (among SNPs in strong linkage disequilibrium (LD), assessed using a web-based tool (SNP Annotation and Proxy Search; www.broadinstitute.org/mpg/snap) with r^2 in the European Ancestry (CEU) panel of HapMap project) was used [25].

Strategy of analysis

First, we aimed at identifying distinct adult asthma phenotypes by applying LCA, a latent variable model that serves to cluster subjects into classes, as previously used in ECRHS and EGEA [10]. Models with different numbers of latent classes were compared using the Bayesian information criterion (BIC) and when BIC were of similar magnitude on the phenotypes' prevalence (to avoid low-prevalent phenotypes in the prospective GWAS analysis). Each subject was assigned to the latent class for which they had the highest membership probability. To better characterise the phenotypes observed, smoking, treatment (in ECRHS and EGEA) and blood eosinophil and neutrophil counts (in EGEA) were compared between LCA-derived phenotypes.

Next, in order to identify genetic variants associated with specific asthma phenotypes, a genome-wide association analysis of each LCA-derived asthma phenotype compared to nonasthma controls was conducted. Genetic associations under a genetic additive model were assessed using a logistic regression model using robust sandwich estimation of the variance to model clustering of family genotypes, with adjustments for sex, study (ECRHS, SAPALDIA and EGEA) and informative principal components for within-Europe diversity (snpMatrix R package). The quantile–quantile plots are shown in online supplementary fig. E1. λ-values for phenotypes A, B, C and D were 1.10, 1.07, 1.11 and 0.95, respectively. On the basis of the effective number of independent markers for the adjustment of multiple testing, $p<1.25\times10^{-7}$ was considered as significant in the GWAS [26]. We also reported all SNPs indicating association signals, defined with two consecutive p-values $<10^{-6}$. We further investigated regions of ~20 kb upstream and downstream of these loci using the imputed genomic data (estimated by the MACH software (www.sph.umich.edu/csg/abecasis/MACH) and the HapMap2 Release 22 CEU reference sample). The regional association plot for each region was performed using LocusZoom (http://csg.sph.umich.edu/locuszoom/). In addition, nonparametric Fisher tests are presented to account for the effect of low minor allele frequency in some SNPs (table E1). To statistically compare the SNP effects across phenotypes, a test for heterogeneity was conducted using a multinomial regression model described in MORRIS *et al.* [27]. To better interpret our GWAS findings on specific asthma phenotypes in light of the largest asthma GWAS conducted so far on a simple asthma outcome, we provided the association observed in the meta-analysis in the GABRIEL study, after exclusion of the three ESE consortium studies, for each SNP identified in our GWAS. As a sensitivity analysis, we conducted the GWAS analysis using LCA probabilities (continuous outcomes), to address the robustness of our results to the outcomes definition. We also investigated the contribution of genetic loci identified for asthma by previous GWASs. In this candidate loci analysis, $p<0.01$ was considered significant, since only *a priori* defined candidate genetic loci were tested [28].

Results

The population under study includes 3001 subjects with asthma (mean age 42.9 years; 44% male) and 3452 subjects without asthma (mean age 46.2 years; 48.9% male) (table 1). Half of the population reported childhood-onset asthma (age <16 years). Current asthma treatment was reported by 45.2% of the population. In ECRHSII and EGEA2, 19% reported daily inhaled corticosteroid use at the time of the survey. The population participating in the genetic analysis comprised older subjects, more males, more asthmatics with early childhood-onset asthma and BHR compared to the population excluded from genetic analysis. A description of the asthmatic population for each survey is presented in table E1.

Asthma phenotypes identified by LCA

The four-class model was retained (fig. E2). The mean highest posterior probability was high (82%), varying from 80% for phenotype D to 83% for phenotypes A and B, indicating that participants were assigned to classes with a fairly high probability. Phenotype A (18% of subjects), labelled "inactive/mild nonallergic asthma" was characterised by individuals with no or few asthma symptoms at the time of examination, low allergic disorders and BHR and high forced expiratory volume in 1 s (FEV_1) (fig. 1 and

TABLE 1 Description of the population included in the present analysis

	Asthmatics included in the LCA	Asthmatics with GWAS data	Asthmatics without GWAS data
Subjects n	3001	1689	1312
Age years	42.9±11.1	43.8±13.1	41.8±7.7*
Male	44.3	46.2	41.8*
Age of asthma onset years			
⩽4	19.0	21.0	16.4*
4–16	30.7	29.4	32.3
>16	50.3	49.6	51.3
Asthma attack in the past 12 months	39.7	37.8	42.1*
Atopy#	65.5	65.5	65.4
FEV1 <80% predicted	13.9	14.6	12.9
BHR PD20 ⩽1 mg	44.6	42.3	48.5*

Data are presented as mean±SD or %, unless otherwise stated. LCA: latent class analysis; GWAS: genome-wide association study; FEV1: forced expiratory volume in 1 s; BHR: bronchial hyperresponsiveness; PD20: provocative dose causing a 20% fall in FEV1. #: assessed by skin prick tests or specific IgE. *: $p<0.05$ comparing subjects with and without GWAS data.

table 2). Phenotype B (37% of subjects), labelled "inactive/mild allergic asthma" was characterised by individuals with no or few symptoms at the time of examination but presenting with atopy and allergic disorders. Phenotype C (27% of subjects), labelled "active allergic asthma", was composed of younger individuals with childhood-onset asthma, atopy, asthma symptoms and BHR at examination. Phenotype D (18% of subjects), labelled "active adult-onset nonallergic asthma", was characterised by subjects reporting adult-onset asthma and asthma symptoms at examination, of whom few had atopy. Subjects belonging to phenotype D had an FEV1<80% predicted more often than the other groups. In the EGEA dataset, this last phenotype was significantly associated with higher blood neutrophil counts (geometric mean neutrophil counts were 3801 cells·mm^{-3}, 3584 cells·mm^{-3}, 3956 cells·mm^{-3} and 4626 cells·mm^{-3} for phenotypes A, B, C and D, respectively; $p<0.0001$).

The frequency of current smoking did not vary strongly across phenotypes (22.2–25.8%); however, subjects included in phenotypes B and C, characterised by a younger age than phenotypes A and D, were more often never-smokers (table E2). Subjects belonging to phenotypes A and B used asthma treatment in the past 3 months less often than others (table E2).

The total overall agreement between the latent classes identified and a simple classification defined by atopy status and asthma attacks in the past 12 months (two highly discriminative variables in the LCA, often

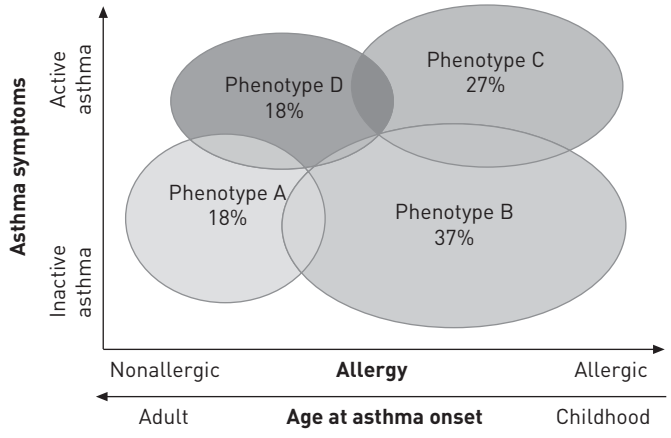

FIGURE 1 Summary of the asthma phenotypes identified using the latent class analysis. The phenotypes are plotted according to the characteristics that play a major role in their classification. The overlaps between the clusters are proportional to the estimated membership probabilities. As an example, subjects assigned to phenotype B had a mean posterior probability to belong to phenotype A, C and D of 7%, 7% and 2%, respectively.

DOI: 10.1183/09031936.00032713

TABLE 2 Characteristics of the population and probability of individuals presenting with characteristics in each of the phenotypes identified by latent class analysis

	Frequency of each variable in the whole sample	Phenotype			
		A	B	C	D
Subjects n (%)					
ECRHS	1895 (100)	328 (18)	648 (34)	533 (28)	386 (20)
SAPALDIA	465 (100)	143 (31)	176 (38)	62 (13)	84 (18)
EGEA	641 (100)	83 (13)	276 (43)	215 (34)	67 (10)
ESE consortium	3001 (100)	554 (18)	1100 (37)	810 (27)	537 (18)
Age ⩾40 years	0.59	0.76	0.53	0.42	0.81
Male	0.44	0.33	0.54	0.48	0.29
Age of asthma onset years					
⩽4	0.19	0.17	0.21	0.28	0.03
4–16	0.31	0.18	0.40	0.41	0.10
>16	0.50	0.65	0.39	0.31	0.87
Woken by coughing past 12 months	0.43	0.37	0.25	0.50	0.75
Asthma symptoms past 12 months					
0	0.26	0.48	0.47	0.00	0.01
1 or 2	0.41	0.48	0.50	0.28	0.33
⩾3	0.33	0.04	0.03	0.72	0.66
Chronic cough or phlegm	0.20	0.15	0.09	0.23	0.45
Asthma attack past 12 months	0.40	0.06	0.09	0.81	0.75
Exacerbation[#] past 12 months	0.11	0.02	0.02	0.19	0.23
Eczema	0.55	0.51	0.56	0.63	0.45
Rhinitis	0.64	0.32	0.74	0.80	0.54
Atopy[¶]	0.65	0.01	0.98	0.98	0.18
IgE ⩾100 IU·mL^{-1}	0.48	0.09	0.59	0.77	0.24
FEV_1 <80% predicted	0.14	0.08	0.09	0.17	0.25
BHR PD_{20} ⩽1 mg	0.45	0.19	0.39	0.72	0.47

Phenotype A: inactive/mild nonallergic asthma; phenotype B: inactive/mild allergic asthma; phenotype C: active allergic asthma, more often childhood onset and bronchial hyperresponsiveness (BHR); phenotype D: active adult onset nonallergic asthma, more often in females; ECRHS: European Community Respiratory Health Survey; SAPALDIA: Study on Air Pollution and Lung Disease in Adults; EGEA: Epidemiological Study on the Genetics and Environment of Asthma; ESE: ECRHS, SAPALDIA and EGEA; FEV_1: forced expiratory volume in 1 s; PD_{20}: provocative dose causing a 20% fall in FEV_1. [#]: defined as either hospitalisation for asthma or the use of oral steroids in the past 12 months; [¶]: assessed by skin prick test or specific IgE.

collected in epidemiological surveys) was 82.4%, but varied widely between phenotypes (from 92.8% for phenotype A to 61.1% for phenotype D).

GWAS results

Manhattan plots of association results for each asthma phenotype are presented in figure 2. We detected two genome-wide significant associations between ''active adult-onset nonallergic asthma'' (phenotype D) and rs9851461 on chromosome 3, flanking *CD200* ($p=9.4\times10^{-9}$), and between ''inactive/mild nonallergic asthma'' (phenotype A) and rs2579931 on chromosome 6 in *GRIK2* ($p=2.7\times10^{-9}$) (table 3). Based on the criteria of two consecutive SNPs at $p<10^{-6}$, four chromosomal locations were detected. Pairwise LD measures between SNPs of each of these genes showing multiple signals are presented in table E4. Of the 11 SNPs, three belonging or flanking the *ALCAM* gene located on chromosome 3 were associated with phenotype D (active adult-onset nonallergic asthma; rs9842772, rs9288812 and rs1051124 with p-values of 2.5×10^{-7}, 6.6×10^{-7} and 8.2×10^{-7}, respectively). The two latter SNPs were not in strong LD with rs9842772 ($r^2=0.39$). Association signals were detected between phenotype A (inactive/mild non-allergic asthma) and four SNPs in *LOC401410* on chromosome 7 (rs10264996, rs10259042, rs10230811 and rs17162196), two SNPs in *LRRC6* on chromosome 8 (rs7834760 and rs13272108) and two SNPs in *SBF2* on chromosome 11 (rs4576815 and rs7938647), with p-values ranging from 2.5×10^{-7} to 9.6×10^{-7}. All SNPs detected at $p<10^{-6}$ exhibited statistically significant heterogeneity of effects observed across the asthma phenotypes ($p<0.002$). GWAS on the other two asthma phenotypes did not provide prominent association signals (table 3). None of the SNPs identified in our GWAS showed any trend for association with asthma in the GABRIEL

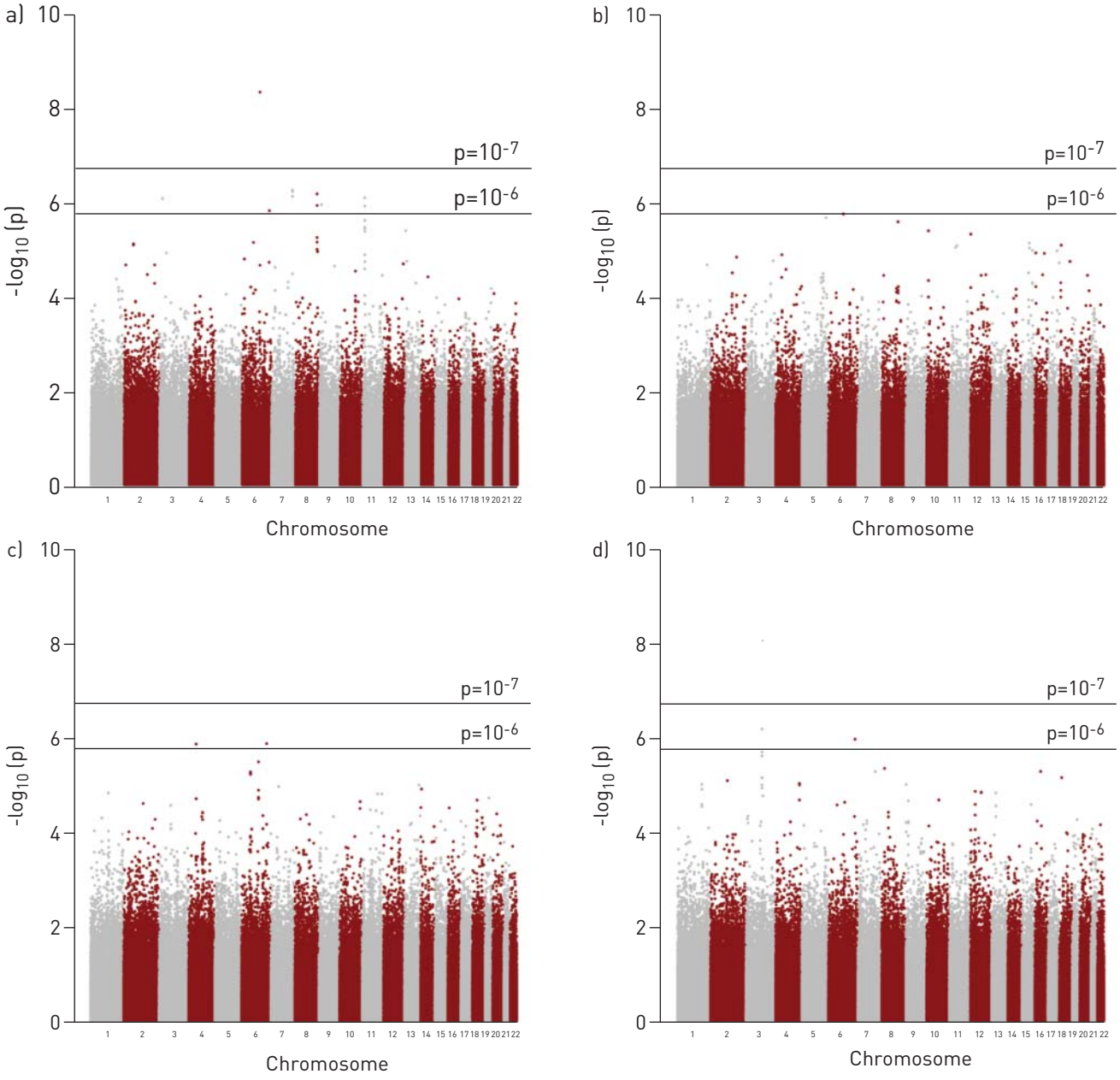

FIGURE 2 Manhattan plots of association results for each asthma phenotype. a) Phenotype A, b) phenotype B, c) phenotype C and d) phenotype D.

meta-analysis after exclusion of the ESE studies ($p>0.50$) (table 3). GWAS analyses using the LCA probabilities (continuous outcomes) led to the same conclusions (table E5).

The six loci detected at $p<10^{-6}$ in the present GWAS were investigated using imputed data (fig. E3) and the analyses were conducted in the pooled sample as well as separately in each study (fig. E4). Association signals were consistent across studies for all SNPs, except for rs9851461 flanking *CD200* and the active adult-onset nonallergic asthma phenotype, showing stronger association in EGEA.

The two SNPs flanking *ALCAM* (rs9288812 and rs1051124) were significantly related to blood neutrophil counts, with the alleles associated with a greater risk for active adult-onset asthma (phenotype D) also being associated with an increased level of neutrophils (n=533, adjusted p-values=0.01). No association was detected between these SNPs and blood neutrophils among subjects without asthma and between blood neutrophils and the two other SNPs reported for phenotype D.

DOI: 10.1183/09031936.00032713

TABLE 3 Results from the genome-wide analysis

Chromosome	Gene	rs number	Position	Reference *versus* alternative allele	Alternative allele frequency	Phenotype								p-value for heterogeneity[f]	Association with asthma in the GABRIEL study after exclusion of ESE studies [1]	
						A[#]		B[¶]		C[+]		D[§]				
						OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value		OR (95% CI)	p-value
3	ALCAM	rs9842772	106735892	G/A	0.86	1.01 (0.81–1.28)	0.98	1.09 (0.91–1.29)	0.41	1.08 (0.89–1.33)	0.42	1.98 (1.45–2.71)	2.5×10^{-7}	3.3×10^{-4}	0.98 (0.92–1.04)	0.50
3	*ALCAM*[##]	*rs9288812*	*106783827*	A/G	0.80	0.99 (0.82–1.19)	0.71	1.06 (0.91–1.23)	0.71	1.02 (0.86–1.21)	0.91	1.63 (1.3–2.05)	8.2×10^{-7}	0.005	1.00 (0.94–1.05)	0.88
3	*ALCAM*[##]	*rs10511245*	*106783905*	G/A	0.80	0.99 (0.82–1.2)	0.74	1.05 (0.9–1.21)	0.81	1.03 (0.86–1.22)	0.86	1.64 (1.31–2.06)	6.6×10^{-7}	0.004	1.00 (0.95–1.05)	0.94
3	CD200[##]	rs9851461	113593771	C/T	0.94	1.05 (0.75–1.45)	0.86	0.88 (0.69–1.12)	0.32	0.93 (0.7–1.24)	0.71	**2.95 (1.7–5.13)**	$\mathbf{9.4\times10^{-9}}$	0.002	0.99 (0.90–1.09)	0.89
6	GRIK2	rs2579931	101961354	A/G	0.90	**2.28 (1.59–3.27)**	$\mathbf{2.7\times10^{-9}}$	1 (0.82–1.22)	0.92	1.19 (0.93–1.53)	0.18	1.02 (0.76–1.37)	0.89	3.4×10^{-4}	0.99 (0.92–1.08)	0.84
7	LOC401410[##]	rs10264996	140731438	A/G	0.95	2.71 (1.49–4.93)	4.3×10^{-7}	1.08 (0.83–1.41)	0.68	1.02 (0.76–1.37)	0.83	1.36 (0.91–2.04)	0.06	2.2×10^{-4}	1.00 (0.91–1.10)	0.98
7	LOC401410[##]	rs10259042	140734967	A/G	0.95	2.73 (1.5–4.96)	3.4×10^{-7}	1.11 (0.85–1.45)	0.56	0.97 (0.73–1.29)	0.58	1.32 (0.89–1.96)	0.08	1.7×10^{-4}	0.99 (0.88–1.10)	0.82
7	LOC401410[##]	rs10230811	140735196	T/C	0.95	2.73 (1.5–4.96)	3.1×10^{-7}	1.11 (0.85–1.45)	0.57	0.99 (0.74–1.32)	0.68	1.32 (0.89–1.97)	0.08	1.9×10^{-4}	1.00 (0.91–1.10)	0.97
7	LOC401410[##]	rs17162196	140739352	T/C	0.95	2.73 (1.5–4.96)	3.1×10^{-7}	1.12 (0.86–1.46)	0.53	0.98 (0.73–1.3)	0.63	1.38 (0.92–2.07)	0.05	1.4×10^{-4}	1.00 (0.91–1.10)	0.97
8	*LRRC6*	*rs7834760*	*133714562*	C/T	0.90	1.97 (1.39–2.79)	6.5×10^{-7}	1.19 (0.96–1.47)	0.08	1.04 (0.83–1.31)	0.76	1.09 (0.83–1.44)	0.52	1.3×10^{-5}	1.01 (0.93–1.11)	0.78
8	LRRC6	rs13272108	133736255	G/A	0.92	2.11 (1.43–3.12)	3.7×10^{-7}	1.15 (0.91–1.44)	0.22	0.9 (0.72–1.14)	0.30	0.98 (0.74–1.31)	0.96	5.4×10^{-5}	1.02 (0.92–1.12)	0.74
11	*SBF2*	*rs4576815*	*9943804*	C/T	0.71	1.52 (1.27–1.82)	9.6×10^{-7}	1.02 (0.89–1.16)	0.77	1.03 (0.88–1.2)	0.76	0.91 (0.75–1.11)	0.37	4.9×10^{-5}	0.99 (0.92–1.06)	0.72
11	*SBF2*	*rs7938647*	*10017999*	C/T	0.72	1.54 (1.28–1.85)	6.1×10^{-7}	1.02 (0.89–1.17)	0.74	1.01 (0.87–1.18)	0.94	0.9 (0.74–1.1)	0.35	4.1×10^{-5}	0.99 (0.92–1.06)	0.78

All single nucleotide polymorphisms (SNPs) showing one p-value $<1.25\times10^{-7}$ or two consecutive p-values $<1.0\times10^{-6}$ with at least one asthma phenotype are shown. Bold indicates significant association at $p<1.25\times10^{-7}$. rs: reference SNP; ESE: European Community Respiratory Health Survey (ECRHS), Study on Air Pollution and Lung Disease in Adults (SAPALDIA) and the Epidemiological Study on the Genetics and Environment of Asthma (EGEA). [#]: 331 *versus* 3452 controls; [¶]: 618 *versus* 3452 controls; [+]: 453 *versus* 3452 controls; [§]: 287 *versus* 3452 controls; [f]: p-value for the heterogeneity of the association observed between the asthma phenotypes, assessed using the multinomial model, as described in MORRIS *et al.* [27]; [##]: not localised in a gene from the SNP database, but near the gene.

Replication of SNPs identified in previous asthma GWAS

13 SNPs located in *IL1RL1*, *IL18R1*, *DPP10*, *TSLP*, *RAD50-IL13*, *HLA-DQ*, *IL33*, *RORA*, *ORMDL3/GSDMB* and *IL12RB* were replicated ($p<0.01$) with active allergic asthma (phenotype C), and heterogeneity of association across phenotypes was significant (p-value for heterogeneity $\leqslant 0.01$) for eight SNPs located in *IL1RL1*, *IL18R1*, *DPP10*, *TSLP*, *HLA-DQ* and *IL33* (table 4). Six SNPs located in *IL1RL1*, *HLA-DQ*, *IL33* and *SMAD3* were significantly associated with inactive/mild allergic asthma (phenotye B). A single SNP in *SRP9* (rs4653433) was associated with active adult-onset nonallergic asthma (phenotype D) (p-value for heterogeneity=0.003). None of the 39 SNPs were associated with inactive/mild nonallergic asthma (phenotype A).

Discussion

Applying LCA, a model-based clustering approach, in a large sample of well-characterised subjects with asthma led to the identification of four asthma phenotypes, mainly characterised by disease activity, allergic status and age of asthma onset. Beside these main characteristics, sex, FEV_1 and BHR also played a role in the classification. GWAS on each asthma phenotype revealed a gene of potential interest in active adult-onset nonallergic asthma, *ALCAM* (activated leukocyte cell adhesion molecule), with evidence of heterogeneity of SNP effect across asthma phenotypes. All replication of asthma SNPs identified by previous GWAS (located in *IL1RL1*, *SMAD3*, *RORA*, *ORMDL3/GSDMB*, *DPP10*, *TSLP*, *RAD50-IL13*, *HLA-DQ*, *IL33* and *IL12RB*) were observed with the allergic asthma phenotypes, except one belonging to the *SRP9* gene found to be associated with active adult-onset nonallergic asthma. Taken together, our results support the hypothesis that a better understanding of the phenotypic heterogeneity of asthma may help to disentangle the genetic heterogeneity of asthma.

One strength of the study relates to the large sample of well-characterised adults with asthma, recruited in three epidemiological settings using standardised protocols and clinical examination, allowing identification of subgroups of subjects with shared characteristics of multiple disease features. As the results of the mostly population-based study designs, the population includes both persistent and remittent asthma, and the prevalence of severe asthma is low in this population. Interestingly, the asthma phenotypes identified by LCA conducted on the pooled EGEA2, ECRHSII and SAPALDIAII population with asthma were highly consistent with phenotypes previously identified separately in ECRHSII and EGEA2 [10]. As previously discussed [5, 10], there are similarities with previous cluster-derived adult asthma phenotypes [8, 9] in the identification of a group of subjects with early-onset atopic asthma and groups of subjects with benign (mild) asthma. Furthermore, our phenotype D, mainly characterised by adult-onset nonatopic asthma, shows similarities to phenotype 5 described by MOORE *et al.* [8] (groups showing higher airflow limitation and exacerbation rate compared to the other phenotypes). Nevertheless, the phenotypes were defined at one time-point and further work is needed in the context of longitudinal data to also account for disease expression variability over time.

Although the phenotypic heterogeneity has been considered as a major limitation in understanding the genetic determinants of asthma, few studies have examined to what extent a better phenotypic resolution leads to identify new genetic determinants [3]. MOFFAT *et al.* [1] previously applied this approach with a single asthma characteristic, age at asthma onset (childhood *versus* adult onset). LI *et al.* [17] performed a GWAS on a population of well-characterised patients with severe or difficult-to-treat asthma and identified the RAD50-IL13 region and the HLA DR/DQ region. These studies used stratification on a limited number of traditional phenotypic traits and thus may have had little opportunity to unravel new associations. Our study is the first to perform genetic analysis in a large population-based sample in which asthma phenotypes were obtained in an unsupervised manner by means of LCA.

The lack of formal replication of the genetic association signals detected by this GWAS approach is a limitation of our study. Nevertheless, the findings were supported by the association patterns observed within the region and the consistency of the association across studies. Replication is particularly challenging here, as there are very few large adult studies with similarly detailed phenotypic information that would enable generation of the phenotypes we have used. The low agreement between the LCA-derived asthma phenotypes and a simple classification based on atopy and the presence of asthma attacks in the past 12 months for phenotype D (active adult-onset nonallergic asthma) indicates that these two characteristics are not sufficient to suitably define this phenotype of major interest, given our GWAS results. Overall our GWAS association findings with specific asthma phenotypes have to be interpreted cautiously.

The sample size may be seen as a limitation of the study. However, this consideration of power has to be discussed taking into account the improved phenotypic characterisation accounting for the disease heterogeneity [30]. Large consortia on asthma genetics have been set up based on ''poor'' asthma phenotype definition, which, in the context of a highly heterogeneous disease, may explain part of the missing

DOI: 10.1183/09031936.00032713

TABLE 4 Genetic association with candidate single nucleotide polymorphisms (SNPs) for asthma

First author [ref.]	Chromosome	Gene	SNP previously reported	SNP used#	LD r^2	Position	Reference *versus* alternative allele	Alternative allele frequency	Phenotype A		Phenotype B		Phenotype C		Phenotype D		p-value for heterogeneity¶
									OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	
Torgerson [2]	1	*CRCT1*	rs4845783	rs1053590	0.73	150755052	T/C	0.63	1.01 (0.85–1.20)	0.98	1.03 (0.91–1.17)	0.60	0.86 (0.74–0.99)	0.06	1.13 (0.94–1.35)	0.22	0.34
Ferreira [24]	1	*IL6R*	rs4129267			152692888	T/C	0.58	0.97 (0.83–1.13)	0.64	0.88 (0.78–0.99)	0.06	0.90 (0.78–1.04)	0.25	0.99 (0.83–1.18)	0.76	0.29
Sleiman [16]	1	*DENND1B*	rs2786098			195592531	T/G	0.77	0.89 (0.73–1.07)	0.29	0.95 (0.82–1.09)	0.61	1.00 (0.85–1.18)	0.99	1.10 (0.89–1.36)	0.43	0.61
Ober [29]	1	*CHI3L1*	rs880633	rs762625	0.17	201419424	T/C	0.82	0.85 (0.69 – 1.04)	0.15	1.03 (0.88–1.21)	0.83	1.04 (0.85–1.28)	0.76	0.98 (0.78–1.22)	0.92	0.29
Torgerson [2]	1	*SRP9*	rs4653433			224041154	A/G	0.56	0.87 (0.74–1.02)	0.09	0.95 (0.84–1.08)	0.57	1.01 (0.87–1.16)	0.79	**0.75 (0.64–0.89)**	$\mathbf{8.7\times10^{-4}}$	**0.003**
Torgerson [2]	2	*IL1RL1*	rs10173081	rs13431828	1	102321085	T/C	0.86	1.25 (0.98–1.61)	0.04	**1.37 (1.14–1.65)**	$\mathbf{1.6\times10^{-4}}$	**1.46 (1.16–1.84)**	$\mathbf{4.4\times10^{-4}}$	1.16 (0.90–1.5)	0.23	$\mathbf{4.6\times10^{-4}}$
Gudbjartsson [18]	2	*IL1RL1*	rs1420101	rs17026974	0.85	102318792	A/G	0.73	0.95 (0.80–1.13)	0.47	0.93 (0.81–1.07)	0.26	0.86 (0.74–1.00)	0.07	1.02 (0.85–1.24)	0.82	0.71
Moffatt [1]	2	*IL18R1*	rs3771166+ $p=6.2\times10^{-6}$			102352654	A/G	0.63	1.16 (0.98–1.36)	0.05	1.14 (1.01–1.29)	0.04	**1.31 (1.13–1.52)**	$\mathbf{5.2\times10^{-4}}$	1.04 (0.87–1.24)	0.61	**0.001**
Mathias [20]	2	*DPP10*	rs1435879			115209357	G/A	0.88	0.98 (0.77–1.26)	0.95	0.84 (0.70–1.00)	0.05	**0.76 (0.62–0.93)**	**0.01**	0.99 (0.75–1.30)	0.97	**0.01**
Torgerson [2]	3	*RTP2*	rs2017908	rs7616923	1	188903913	T/G	0.87	1.05 (0.83–1.33)	0.58	0.99 (0.82–1.18)	0.89	0.99 (0.80–1.23)	0.87	0.95 (0.73–1.24)	0.74	0.32
Torgerson [2]	4	*EPHA5*	rs11735820	rs7697951	1	66169736	C/T	0.70	1 (0.85–1.19)	0.86	0.92 (0.81–1.04)	0.13	1.09 (0.94–1.27)	0.19	0.91 (0.76–1.08)	0.29	0.26
Himes [19]	5	*PDE4D*	rs1588265			59405551	G/A	0.70	1.02 (0.86–1.22)	0.94	0.95 (0.83–1.08)	0.40	0.88 (0.76–1.03)	0.14	1.00 (0.83–1.21)	0.99	0.20
Torgerson [2]	5	*TSLP*	rs1837253			110429771	T/C	0.76	0.86 (0.72–1.03)	0.12	1.15 (1.00–1.33)	0.07	**1.29 (1.10–1.52)**	**0.002**	1.25 (1.02–1.53)	0.02	$\mathbf{2.0\times10^{-4}}$
Gudbjartsson [18]	5	*WDR36*	rs2416257			110463389	T/C	0.87	0.83 (0.66–1.04)	0.13	1.12 (0.92–1.35)	0.24	1.22 (0.98–1.51)	0.08	1.12 (0.87–1.45)	0.34	**0.006**
Moffatt [1]	5	*SLC22A5*	rs2073643+ $p=2.4\times10^{-6}$			131751187	T/C	0.57	0.92 (0.79–1.08)	0.32	0.91 (0.80–1.03)	0.08	0.90 (0.78–1.04)	0.11	1.03 (0.87–1.23)	0.65	0.15
Li [17]	5	*RAD50-IL13*	rs2897443			131957493	T/G	0.80	1.04 (0.85–1.26)	0.78	0.96 (0.83–1.12)	0.57	**0.77 (0.66–0.91)**	**0.006**	0.96 (0.78–1.17)	0.67	0.22
Moffatt [1]	5	*IL13*	rs1295686+ $p=1.8\times10^{-6}$			132023742	T/C	0.80	1.10 (0.89–1.36)	0.50	0.99 (0.85–1.16)	0.99	0.87 (0.74–1.03)	0.18	1.00 (0.81–1.24)	0.89	0.10
Torgerson [2]	5	*GALNT10*	rs10064618			153752482	A/G	0.61	1.00 (0.85–1.18)	0.91	1.04 (0.92–1.17)	0.66	0.97 (0.84–1.11)	0.52	1.05 (0.88–1.25)	0.51	0.26
Mathias [20]	5	*ADRA1B*	rs10515807			159297576	A/G	0.80	1.17 (0.95–1.43)	0.15	1.01 (0.87–1.17)	0.83	1.10 (0.92–1.31)	0.23	1.17 (0.95–1.45)	0.16	0.28
Moffatt [1]	6	*HLA-DQ*	rs9273349+ $p=4.4\times10^{-8}$			32733847	T/C	0.59	1.15 (0.97–1.36)	0.12	**1.28 (1.13–1.45)**	$\mathbf{7.3\times10^{-5}}$	**1.33 (1.14–1.54)**	$\mathbf{1.7\times10^{-4}}$	1.19 (0.99–1.43)	0.06	$\mathbf{1.3\times10^{-5}}$
Noguchi [22]	6	*HLADP*	rs987870			33150858	G/A	0.83	1.16 (0.93–1.46)	0.13	0.87 (0.74–1.02)	0.18	0.79 (0.66–0.94)	0.03	0.9 (0.72–1.13)	0.33	**0.01**
Torgerson [2]	8	*FBXO43*	rs2453626			101207073	T/C	0.55	1.02 (0.87–1.2)	0.59	0.95 (0.84–1.07)	0.41	1.10 (0.95–1.28)	0.23	1.14 (0.96–1.35)	0.13	0.27
Moffatt [1]	9	*IL33*	rs1342326+ $p=1.2\times10^{-7}$			6180076	C/A	0.81	0.83 (0.69–1.01)	0.07	**0.78 (0.68–0.90)**	**0.003**	**0.76 (0.64–0.90)**	**0.006**	0.91 (0.73–1.13)	0.36	**0.005**
Torgerson [2]	9	*IL33*	rs2381416	rs928413	1	6203387	G/A	0.71	0.81 (0.68–0.97)	0.02	**0.78 (0.68–0.89)**	$\mathbf{7.2\times10^{-4}}$	**0.73 (0.63–0.84)**	$\mathbf{1.6\times10^{-4}}$	0.89 (0.74–1.08)	0.22	$\mathbf{8.1\times10^{-5}}$

TABLE 4 Continued

First author [ref.]	Chromosome	Gene	SNP previously reported	SNP used#	LD r^2	Position	Reference *versus* alternative allele	Alternative allele frequency	Phenotype A		Phenotype B		Phenotype C		Phenotype D		p-value for heterogeneity¶
									OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	
Gudbjartsson [18]	9	*IL33*	rs3939286			6200099	T/C	0.72	0.81 (0.68–0.96)	0.02	**0.79 (0.69–0.91)**	**0.002**	**0.74 (0.64–0.86)**	**3.1×10^{-4}**	0.90 (0.74–1.08)	0.23	**2.6×10^{-4}**
Hancock [21]	9	*TLE4*	rs2378383			81229182	G/A	0.89	1.00 (0.77–1.30)	0.94	0.89 (0.74–1.08)	0.16	1.11 (0.88–1.39)	0.52	0.88 (0.68–1.14)	0.44	0.66
Ferreira [24]	11	*LRRC32_GARP*	rs7130588			75948331	G/A	0.64	0.99 (0.84–1.16)	0.89	0.88 (0.78–1.00)	0.08	0.89 (0.76–1.03)	0.16	1.00 (0.84–1.20)	0.94	0.08
Torgerson [2]	11	*C11orf71*	rs11214966	rs12223585	0.83	113740900	T/C	0.94	1.31 (0.93–1.85)	0.11	1.28 (0.99–1.66)	0.05	1.04 (0.78–1.38)	0.80	1.01 (0.72–1.41)	0.87	0.62
Torgerson [2]	12	*RASSF8*	rs16929496	rs10842635	0.96	25924553	A/G	0.78	1.04 (0.85–1.26)	0.52	1.11 (0.95–1.29)	0.16	1.16 (0.97–1.39)	0.17	1.19 (0.96–1.49)	0.08	0.41
Moffatt [1]	15	*RORA*	rs11071559+ $p=5.9\times10^{-6}$			58857280	T/C	0.87	0.98 (0.77–1.25)	0.87	1.22 (1.00–1.49)	0.03	**1.48 (1.17–1.86)**	**3.4×10^{-4}**	1.26 (0.95–1.65)	0.08	0.02
Moffatt [1]	15	*SMAD3*	rs744910+ $p=3.4\times10^{-7}$			65233839	G/A	0.50	0.98 (0.84–1.15)	0.71	**0.83 (0.74–0.94)**	**0.002**	0.85 (0.74–0.97)	0.03	0.82 (0.69–0.97)	0.02	**0.002**
Torgerson [2]	17	*AURKB*	rs9891949	rs7503353	0.43	8048704	G/T	0.50	1.01 (0.87–1.18)	0.90	1.00 (0.89–1.12)	0.90	0.98 (0.85–1.12)	0.78	1.01 (0.85–1.19)	0.99	0.92
Moffatt [1]	17	*GSDMB*	rs2305480+ $p=1.2\times10^{-7}$			35315722	A/G	0.56	1.00 (0.86–1.18)	0.86	1.12 (0.99–1.26)	0.11	**1.26 (1.09–1.46)**	**0.004**	1.14 (0.95–1.36)	0.12	0.03
Moffatt [15]	17	*ORMDL3/GSDMB*	rs7216389			35323475	C/T	0.52	1.00 (0.85–1.18)	0.97	1.09 (0.97–1.23)	0.25	**1.25 (1.08–1.44)**	**0.006**	1.13 (0.95–1.35)	0.12	0.05
Moffatt [1]	17	*GSDMA*	rs3894194+ $p=2.0\times10^{-8}$			35375519	A/G	0.53	0.95 (0.80–1.12)	0.49	0.94 (0.83–1.06)	0.46	0.85 (0.74–0.98)	0.06	0.86 (0.72–1.03)	0.06	0.32
Mathias [20]	17	*GBNA13*	rs3972219			60448995	G/A	0.95	1.27 (0.87–1.85)	0.30	1.07 (0.82–1.40)	0.72	1.06 (0.78–1.44)	0.69	1.00 (0.70–1.42)	0.99	0.72
Torgerson [2]	19	*C19orf2*	rs335016	rs34707	1	35068558	T/G	0.75	0.87 (0.73–1.05)	0.18	0.93 (0.81–1.07)	0.37	0.96 (0.82–1.13)	0.68	1.24 (1.01–1.53)	0.03	0.40
Mathias [20]	20	*PRNP*	rs6052761			4605017	C/T	0.89	1.05 (0.81–1.36)	0.61	0.79 (0.65–0.95)	0.02	0.85 (0.69–1.05)	0.12	1.25 (0.92–1.68)	0.12	0.04
Moffatt [1]	22	*IL12RB*	rs2284033+ $p=5.5\times10^{-7}$			35863980	A/G	0.57	1.00 (0.85–1.17)	0.94	1.04 (0.92–1.17)	0.51	**1.21 (1.05–1.40)**	**0.01**	1.16 (0.98–1.37)	0.08	0.06

Bold represents associations with $p\leqslant0.001$. LD: linkage disequilibrium. #: if the SNP reported in the literature was not available in our genotyped data, we reported association with the available SNP showing the strongest LD with the initial reported SNP; ¶: p-value for the heterogeneity of the association observed between the asthma phenotypes, assessed using the multinomial model, as described in Morris *et al.* [27]; +: because the GABRIEL study (Moffatt *et al.* [1]) includes participants from European Community Respiratory Health Survey (ECHRS), the Study on Air Pollution and Lung Disease in Adults (SAPALDIA) and the Epidemiological Study on the Genetics and Environment of Asthma (EGEA), the association from the GABRIEL study after excluding these three studies in the meta-analysis is presented.

DOI: 10.1183/09031936.00032713

heritability as some genetic effects might be diluted as the result of phenotype missclassification [3]. The approach used in the present analysis, favouring phenotypic characterisation over sample size, could therefore provide complementary insights to large asthma genetic studies. This is, in part, supported by the lack of overlap between our GWAS results and the GABRIEL results. One limitation on the approach used lies in the difficulty to directly compare findings across studies, since cluster-based phenotypes differ between studies.

ALCAM, a member of the Ig superfamily, is a good asthma candidate gene. The *ALCAM* gene encodes the CD166 antigen and was originally identified as a transmembrane receptor that is involved in T-cell activation and may play a role in the binding of T- and B-cells to activated leukocytes. Altered expression of ALCAM has been associated with differentiation state and progression in many tumours [31]. More interestingly, *ALCAM* was identified as a common gene in three inflammatory diseases: Crohn's disease, rheumatoid arthritis and type 1 diabetes [32]. Furthermore, *ALCAM* interacts with *ADAM17* (a disintegrin and metalloproteinase 17), which is implicated in immune cell development and function and has been shown to play a role in the epidermal barrier [33, 34]. Interestingly, a further SNP in the *CD200* gene (CD200 molecule), located in the same genomic region as *ALCAM* and encoding a protein also belonging to the Ig superfamily, was exclusively associated with the active adult-onset asthma phenotype. Nevertheless, this result in CD200 should be interpreted cautiously because of the low minor allele frequency of the identified SNP. To date, there is no strong biological evidence to support the role of the four genes showing an association with inactive/mild nonallergic asthma (*GRIK2* (glutamate receptor, ionotropic, kainate 2, 6q16.3) involved in neurophysiological processes; *LOC401410* (7q34); *LRRC6* (leucine-rich repeat containing 6, 8q24.22), possibly involved in spermatocytogenesis; and *SBF2* (SET-binding factor, 11p15.4), possibly involved in biological processes related to bone and muscle growth [35]). The lack of hits with phenotype C might be explained by limited statistical power.

None of the variants identified in previous asthma GWAS meet the GWAS significance level in the present study. This might be explained by a lack of power of our study to identify shared genetic variants between phenotypes compared to previous asthma GWAS, which considered the whole group of subjects with asthma. Most of the replications (using the 0.01 threshold) were observed for the active allergic asthma phenotype, including SNPs in or near *IL1RL1*, *IL18R1*, *DPP10*, *RAD50-IL13*, *HLA-DQ*, *IL33*, *RORA*, *ORMDL3/GSDMB* and *IL12RB*. This may have occurred because childhood-onset asthma is more prone to association with allergic phenotypes, and has been more intensively investigated in previous asthma GWAS. Our results are consistent with a specific role of SNPs in *ORMDL3/GSDMB* with childhood-onset asthma (stronger odds ratio observed with phenotype C), although the p-value for heterogeneity did not indicate marked differences between the described phenotypes (possibly because of our smaller sample size). Our results indicate a specific role of SNP rs1837253 in *TSLP*, a gene involved in the T-helper type-2 cell immune processes in the airways of subjects with asthma, in active asthma as supported by the heterogeneity in allelic odds ratios between phenotypes. This suggests that phenotypic heterogeneity may partly explain the genetic heterogeneity previously identified by the GABRIEL study for this SNP.

Our results support the hypothesis that a better understanding of the asthma phenotypic heterogeneity helps to disentangle the genetic heterogeneity of asthma. The genetic and environmental components of the aetiology of asthma may be clarified by considering specific asthma phenotypes.

Acknowledgements

We thank the Epidemiological Study on the Genetics and Environment of Asthma (EGEA) cooperative group members as follows. Coordination: V. Siroux (epidemiology, PI since 2013; Inserm U823, Grenoble, France); F. Demenais (genetics; Inserm U946, Paris, France); I. Pin (clinical aspects; CHU Grenoble, Grenoble); R. Nadif (biology; Inserm CESP/U 1018, Villejuif, France); and F. Kauffmann (PI 1992-2012; Inserm CESP/U 1018, Villejuif). Respiratory epidemiology: M. Korobaeff (Egea1), F. Neukirch (Egea1) (both Inserm, U700, Paris); I. Annesi-Maesano (Egea1-2; Inserm U 707, Paris); F. Kauffmann, N. Le Moual, R. Nadif, M.P. Oryszczyn (Egea1-2), R. Varraso (all Inserm CESP/U 1018, Villejuif); V. Siroux (Inserm U 823, Grenoble). Genetics: J. Feingold (Inserm U 393, Paris); E. Bouzigon, F. Demenais, M.H. Dizier (all Inserm U 946, Paris); I. Gut (CNG, Evry; now CNAG, Barcelona, Spain), M. Lathrop (CNG, Evry; now University McGill, Montreal, Canada). Clinical centres: I. Pin, C. Pison (both CHU Grenoble, Grenoble); D. Ecochard (Egea1), F. Gormand, Y. Pacheco (all CHU Lyon, Lyon, France); D. Charpin (Egea1), D. Vervloet (Egea12) (both CHU Marseille, Marseille, France); J. Bousquet (CHU Montpellier, Montpellier, France); A. Lockhart (Egea1), R. Matran (both Hopital Cochin, Paris); E. Paty (Egea12), P. Scheinmann (Egea12) (both Hopital Necker, Paris); A. Grimfeld (Egea12), J. Just (both Hopital Trousseau, Paris). Data and quality management: J. Hochez (EGEA1, Inserm ex-U155, Paris); N. Le Moual (Inserm CESP/U 1018, Villejuif); C. Ravault (Egea12) (Inserm ex-U780, Villejuif); N. Chateigner (Egea12) (Inserm ex-U794, Paris); J. Quentin-Ferran (Egea1-2) (CHU Grenoble, Grenoble). The authors thank all those who participated in the setting of the study and on the various aspects of the examinations involved: interviewers, technicians for lung function testing, coders, those involved in quality control and data management, and all those who supervised the study in all centres. The authors are grateful to the three CIC-Inserm of Necker, Grenoble and Marseille who supported the study and in which subjects were examined. They are indebted to all the subjects who participated, without whom that study would not have been possible.

We thank the European Community Respiratory Health Survey (ECHRS) study team, a collaboration of European research groups, many of whom also agreed to provide blood samples for genotying as part of the GABRIEL initiative. Investigators in the collaborating centers are: D. Jarvis (Imperial College London, London, UK), M. Wjst (Helmholtz Zentrum Munchen, Munich, Germany), M. Kogevinas (CREAL, Barcelona, Spain), R. Jogi (University Hospital Tartu, Tartu, Estonia), C. Janson (University of Uppsala, Uppsala, Sweden), K. Franklin (University of Umea, Umea, Sweden), E. Omenaas (University of Bergen, Bergen, Norway), B. Leynaert (Inserm U700, Paris, France), I. Pin (CHU Grenoble, Grenoble, France), J. Heinrich (Helmholtz Zentrum Munchen, Munich), N. Kuenzli (Swiss Tropical Institute, Basel, Switzerland), N.M. Probst-Hensch (Swiss Tropical Institute, Basel), J.M. Anto (CREAL, Barcelona), J. Sunyer (CREAL, Barcelona), J-A. Maldonado (Hospital Juan Ramón Jiménez, Huelva, Spain), J. Martinez-Moratalla (Complejo Hospitalario Universitario de Albacet, Albacete, Spain), I. Urritia (Hospital Galdakao-Usansolo-Osakidetza, Galdakao, Spain), and F. Payo (Hospital Universitario Central de Asturias, Oviedo, Spain).

We thank the SAPALDIA study team members. Study directorate: T. Rochat (pneumology; University Hospital, Geneva, Switzerland), J.M. Gaspoz (cardiology; University Hospital, Geneva), N. Künzli (epidemiology/exposure; SwissTPH, Basel, Switzerland), N.M. Probst Hensch (epidemiology/genetic and molecular biology; SwissTPH, Basel), C. Schindler (statistics; SwissTPH, Basel). Scientific team: J.C. Barthélémy (cardiology; University Hospital, St Etienne, France), W. Berger (genetic and molecular biology; University of Zürich, Zürich, Switzerland), R. Bettschart (pneumology; Hirslanden Klinik, Aarau, Switzerland), A. Bircher (allergology; University Hospital, Basel, Switzerland), G. Bolognini (pneumology; Regional Hospital, Mendrisio, Switzerland), O. Brändli (pneumology), C. Brombach (nutrition, FHS Wädenswil, Switzerland), M. Brutsche (pneumology; Cantonal Hospital, St Gallen, Switzerland), L. Burdet (pneumology; Regional Hospital, Payerne, Switzerland), M. Frey (pneumology; Klinik Barmelweid, Aarau), U. Frey (paediatrics, University Children's Hospital, Basel), MW. Gerbase (pneumology; University Hospital, Geneva), D. Gold (epidemiology/cardiology/pneumology; Harvard School of Public Health, Boston, MA, USA), E. de Groot (cardiology, Academic Medical Center, Amsterdam, the Netherlands), W. Karrer (pneumology; Luzerner Höhenklinik,Crans Montana, Switzerland), R. Keller (pneumology; Aarau, Switzerland), B. Knöpfli (pneumology; Davos, Switzerland), B. Martin (physical activity; University of Zürich, Zürich), D. Miedinger, U. Neu (exposure; ProClim, Bern, Switzerland), L. Nicod (pneumology; University Hospital, Lausanne, Switzerland), M. Pons (pneumology; Regional Hospital, Lugano, Switzerland), F. Roche (cardiology; University Hospital, St Etienne), T. Rothe (pneumology; Zürcher Höhenklinik, Davos, Switzerland), E. Russi (pneumology; University Hospital, Zürich), P. Schmid-Grendelmeyer (allergology; University Hospital, Zürich), A. Schmidt-Trucksäss (physical activity; University Basel, Basel, Switzerland), A. Turk (Höhenklinik, Wald, Switzerland), J. Schwartz (epidemiology; Harvard School of Public Health, Boston), D. Stolz (pneumology, University Hospital, Basel), P. Straehl (exposure; BAFU, Bern, Switzerland), J.M. Tschopp (pneumology; Centre Valaisan de Pneumologie, Crans Montana), A. von Eckardstein (clinical chemistry; University Hospital, Zürich), E. Zemp Stutz (epidemiology; SwissTPH, Basel, Switzerland). Scientific team at coordinating centers: M. Adam (epidemiology; SwissTPH, Basel), E. Boes (genetic and molecular biology; SwissTPH, Basel), P.O. Bridevaux (pneumology; University Hospital, Geneva), D. Carballo (cardiology; University Hospital, Geneva), E. Corradi (epidemiology; University Hospital, Geneva), I. Curjuric (epidemiology; SwissTPH, Basel), J. Dratva (epidemiology; SwissTPH, Basel), A. Di Pasquale (statistics; SwissTPH, Basel), L. Grize (statistics; SwissTPH, Basel), D. Keidel (statistics; SwissTPH, Basel), S. Kriemler (physical activity; SwissTPH, Basel), A. Kumar (genetic and molecular biology; SwissTPH, Basel), M. Imboden (genetic and molecular biology; SwissTPH, Basel), N. Maire, A. Mehta (epidemiology; SwissTPH, Basel), F. Meier (epidemiology; SwissTPH, Basel), H. Phuleria (exposure; SwissTPH, Basel), E. Schaffner (statistics; SwissTPH, Basel), G.A. Thun (epidemiology; SwissTPH, Basel) A. Ineichen (exposure; SwissTPH, Basel), M. Ragettli (epidemiology; SwissTPH, Basel), M. Ritter (exposure; SwissTPH, Basel), T. Schikowski (epidemiology; SwissTPH, Basel), G. Stern (paediatrics, University Children's Hospital, Basel), M. Tarantino (statistics; SwissTPH, Basel), M. Tsai (exposure; SwissTPH, Basel), M. Wanner (physical activity; University of Zürich, Zürich). 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: S. Brun, G. Giger, M. Sperisen and M. Stahel (Aarau); C. Bürli, C. Dahler, N. Oertli, I. Harreh, F. Karrer, G. Novicic and N. Wyttenbacher (Basel); A. Saner, P. Senn and R. Winzeler (Davos); F. Bonfils, B. Blicharz, C. Landolt and J. Rochat (Geneva); S. Boccia, E. Gehrig, M.T. Mandia, G. Solari and B. Viscardi (Lugano, Switzerland); A.P. Bieri, C. Darioly and M. Maire (Montana); F. Ding, P. Danieli andA. Vonnez (Payerne); D. Bodmer, E. Hochstrasser, R. Kunz, C. Meier, J. Rakic, U. Schafroth and A. Walder (Wald). Administrative staff: C. Gabriel (University Hospital, Geneva) and R. Gutknecht (SwissTPH, Basel).

EGEA, SAPPALDIA and ECRHS were part of the GABRIEL Consortium, a European 6th Framework Research project on asthma genetics (HEALTH-LSH-2005-018996), which allowed us to obtain the genotype information used in this analysis.

The Epidemiological Study on the Genetics and Environment of Asthma (EGEA) was supported, in part, by grants from Merck Sharp and Dohme; the Hospital Programme of Clinical Research (Paris, France); National Research Agency (ANR) (ANR 06-CEBS) Health Environment, Health-Work Programme (ANR 05-SEST-020-02/05-9-97); ANR Biological Collections for Health Programme; ANR Public Health Research Programme (ANR 10-PRSP); French Agency for Environmental and Occupational Health Safety (Afsset-APR-SE-2004); the Scientific Committee of the Medico-Technology Support at Home (Agir à dom) and Isere Committee Against Respiratory Diseases.

The coordination of the European Community Respiratory Health Survey (ECRHS) II was supported by the European Commission, as part of their quality of life programme. The following bodies funded the local studies in ECRHSII: Albacete, Spain: Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01 and 99/0034-02), Hospital Universitario de Albacete and Consejeria de Sanidad; Barcelona, Spain: SEPAR, Public Health Service (grant R01 HL62633-01), Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01 and 99/0034-02), CIRIT (grant 1999SGR 00241) and Red Respira ISCII; CIBER Epidemiologia y Salud Publica (CIBERESP); Basel, Switzerland: Swiss National Science Foundation, Swiss Federal Office for Education and Science, Swiss National Accident Insurnce Fund (SUVA) and USC NIEHS Center (grant 5P30 ES07048); Bergen, Norway: Norwegian Research Council, Norwegian Asthma and Allergy Association (NAAF), Glaxo Wellcome AS and Norway Research Fund; Erfurt, Germany: GSF-National Research Centre for Environment and Health and Deutsche Forschungsgemeinschaft (DFG) (grant FR 1526/1-1); Galdakao, Spain: Basque Health Dept; Grenoble, France: 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

 DOI: 10.1183/09031936.00032713

Solidarite, Direction Generale de la Sante and Comite des Maladies Respiratoires de l'Isere; Hamburg, Germany: GSF-National Reasearch Centre for Environment and Health and Deutsche Forschungsgemeinschaft (DFG) (grant MA 711/4-1); Ipswich and Norwich, UK: Asthma UK (formerly known as National Asthma Campaign); Huelva, Spain: Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01 and 99/0034-02); Oviedo, Spain: Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01 and 99/0034-02); Paris, France: 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, Estonia: Estonian Science Foundation; Umea, Sweden: Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences and Allergy Research, Swedish Asthma and Allergy Foundation, Swedish Cancer and Allergy Foundation; Uppsala, Sweden: Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences and Allergy Research, Swedish Asthma and Allergy Foundation, Swedish Cancer and Allergy Foundation. Financial support for ECRHSI for centres in ECRHSII was provided by: Ministere de la Sante, Glaxo France, Insitut Pneumologique d'Aquitaine, Contrat de Plan Etat-Region Languedoc-Rousillon, CNMATS, CNMRT (90MR/10, 91AF/6), Ministre delegue de la sante, RNSP, France; GSF, and the Bundesminister fur Forschung und Technologie, Bonn, Germany; Norwegian Research Council project no. 101422/310; Ministero Sanidad y Consumo FIS (grants #91/0016060/00E-05E and #93/0393) and grants from Hospital General de Albacete, Hospital General Juan Ramon Jimenenz, Consejeria 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, UK Department of Health, South Thames Regional Health Authority, UK. A. Ramasamy was supported by the Department of Health, UK and the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1.

The Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) was supported by the Swiss National Science Foundation (grants no 33CS30_134276/1, 33CSCO-108796, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, 3233-054996 and PDFMP3-123171), 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-Land, Geneva, Luzern, Ticino, Valais, Zurich, the Swiss Lung League, the canton's Lung League of Basel Stadt/Basel Landschaft, Geneva, Ticino, Valais and Zurich, Schweizerische Unfallversicherungsanstalt (SUVA), Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH and Abbott Diagnostics.

References

1 Moffatt MF, Gut IG, Demenais F, *et al.* A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010; 363: 1211–1221.
2 Torgerson DG, Ampleford EJ, Chiu GY, *et al.* Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011; 43: 887–892.
3 Manolio TA, Collins FS, Cox NJ, *et al.* Finding the missing heritability of complex diseases. *Nature* 2009; 461: 747–753.
4 Wjst M, Sargurupremraj M, Arnold M. Genome-wide association studies in asthma: what they really told us about pathogenesis. *Curr Opin Allergy Clin Immunol* 2013; 13: 112–118.
5 Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012; 18: 716–725.
6 Bouzigon E, Corda E, Aschard H, *et al.* Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008; 359: 1985–1994.
7 Siroux V, Garcia-Aymerich J. The investigation of asthma phenotypes. *Curr Opin Allergy Clin Immunol* 2011; 11: 393–399.
8 Moore WC, Meyers DA, Wenzel SE, *et al.* Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010; 181: 315–323.
9 Haldar P, Pavord ID, Shaw DE, *et al.* Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008; 178: 218–224.
10 Siroux V, Basagaña X, Boudier A, *et al.* Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J* 2011; 38: 310–317.
11 The European Community Respiratory Health Survey II Steering Committee. The European Community Respiratory Health Survey II. *Eur Respir J* 2002; 20: 1071–1079.
12 Ackermann-Liebrich U, Kuna-Dibbert B, Probst-Hensch NM, *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 Praventivmed* 2005; 50: 245–263.
13 Kauffmann F, Dizier MH, Pin I, *et al.* Epidemiological study of the genetics and environment of asthma, bronchial hyperresponsiveness, and atopy: phenotype issues. *Am J Respir Crit Care Med* 1997; 156: S123–S129.
14 Imboden M, Bouzigon E, Curjuric I, *et al.* Genome-wide association study of lung function decline in adults with and without asthma. *J Allergy Clin Immunol* 2012; 129: 1218–1228.
15 Moffatt MF, Kabesch M, Liang L, *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448: 470–473.
16 Sleiman PM, Flory J, Imielinski M, *et al.* Variants of *DENND1B* associated with asthma in children. *N Engl J Med* 2010; 362: 36–44.
17 Li X, Howard TD, Zheng SL, *et al.* Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol* 2010; 125: 328–335.
18 Gudbjartsson DF, Bjornsdottir US, Halapi E, *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009; 41: 342–347.
19 Himes BE, Hunninghake GM, Baurley JW, *et al.* Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet* 2009; 84: 581–593.
20 Mathias RA, Grant AV, Rafaels N, *et al.* A genome-wide association study on African-ancestry populations for asthma. *J Allergy Clin Immunol* 2009; 125: 336–346.

21 Hancock DB, Romieu I, Shi M, *et al.* Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in Mexican children. *PLoS Genet* 2009; 5: e1000623.
22 Noguchi E, Sakamoto H, Hirota T, *et al.* Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011; 7: e1002170.
23 Ober C, Cox NJ, Abney M, *et al.* Genome-wide search for asthma susceptibility loci in a founder population. The Collaborative Study on the Genetics of Asthma. *Hum Mol Genet* 1998; 7: 1393–1398.
24 Ferreira MAR, Matheson MC, Duffy DL, *et al.* Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011; 378: 1006–1014.
25 Johnson AD, Handsaker RE, Pulit SL, *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008; 24: 2938–2939.
26 Li MX, Yeung JM, Cherny SS, *et al.* Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet* 2012; 131: 747–756.
27 Morris AP, Lindgren CM, Zeggini E, *et al.* A powerful approach to sub-phenotype analysis in population-based genetic association studies. *Genet Epidemiol* 2010; 34: 335–343.
28 Ege MJ, Strachan DP, Cookson WO, *et al.* Gene–environment interaction for childhood asthma and exposure to farming in Central Europe. *J Allergy Clin Immunol* 2011; 127: 138–144.
29 Ober C, Tan Z, Sun Y, *et al.* Effect of variation in *CHI3L1* on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008; 358: 1682–1691.
30 Bennett SN, Caporaso N, Fitzpatrick AL, *et al.* Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience. *Genet Epidemiol* 2011; 35: 159–173.
31 Weidle UH, Eggle D, Klostermann S, *et al.* ALCAM/CD166: cancer-related issues. *Cancer Genomics Proteomics* 2010; 7: 231–243.
32 Eleftherohorinou H, Wright V, Hoggart C, *et al.* Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. *PLoS One* 2009; 4: e8068.
33 Miccichè F, Da Riva L, Fabbi M, *et al.* Activated leukocyte cell adhesion molecule expression and shedding in thyroid tumors. *PLoS One* 2011; 6: e17141.
34 Murthy A, Shao YW, Narala SR, *et al.* Notch activation by the metalloproteinase ADAM17 regulates myeloproliferation and atopic barrier immunity by suppressing epithelial cytokine synthesis. *Immunity* 2012; 36: 105–119.
35 Lei SF, Tan LJ, Liu XG, *et al.* Genome-wide association study identifies two novel loci containing *FLNB* and *SBF2* genes underlying stature variation. *Hum Mol Genet* 2009; 18: 1661–1669.

DOI: 10.1183/09031936.00032713