THE LANCET

## Supplementary webappendix

This webappendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Ferreira MAR, Matheson MC, Duffy DL, et al, for the Australian Asthma Genetics Consortium. Identification of *IL6R* and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011; **378:** 1006–14.

# Identification of *IL6R* and chromosome 11q13.5 as risk loci for asthma

## Supplementary webappendix

Manuel A.R. Ferreira, Melanie C. Matheson, David L. Duffy, Guy B. Marks, Jennie Hui, Peter Le Souëf, Patrick Danoy, Svetlana Baltic, Dale R. Nyholt, Mark Jenkins, Catherine Hayden, Gonneke Willemsen, Wei Ang, Mikko Kuokkanen, John Beilby, Faang Cheah, Eco J. C. de Geus, Adaikalavan Ramasamy, Sailaja Vedantam, Veikko Salomaa, Désirée Mészáros, Mary Roberts, Melissa C. Southey, Euan R. Tovey, Nicole M. Warrington, Mathijs Kattenberg, Lyle J. Palmer, Loren Price, Margaret J. Wright, Scott D. Gordon, Li P. Chung, Anjali K. Henders, Graham Giles, Jouke-Jan Hottenga, Paul S. Thomas, Suzanna Temple, John B. Whitfield, Ian Feather, Anna-Liisa Hartikainen, Tari Haahtela, Tarja Laitinen, Pekka Jousilahti, Johan G. Eriksson, Elisabeth Widen, Olli T. Raitakari, Terho Lehtimäki, Mika Kähönen, Jorma Viikari, Aarno Palotie, Zofia K. Gajdos, Helen N. Lyon, George T. O'Connor, Stephen Morrison, Peter D. Sly, Chalermchai Mitrpant, Warwick J. Britton, David John, Pat G. Holt, Andrew S. Kemp, Pamela A. Madden, Andrew C. Heath, John L. Hopper, Peter M. Visscher, Bill Musk, Stephen R. Leeder, Marjo-Riitta Jarvelin, Craig Pennell, Dorret I. Boomsma, Joel N. Hirschhorn, Haydn Walters, Nicholas G. Martin, Alan James, Graham Jones, Michael J. Abramson, Colin F. Robertson, Shyamali C. Dharmage, Matthew A. Brown, Grant W. Montgomery* and Philip J. Thompson*

## Ascertainment of samples included in the Australian GWAS dataset

The primary GWAS described in this paper includes data from 7,197 Australian individuals from three cohorts (AAGC, Busselton, QIMR) which are described below. Clinical characteristics for these individuals are summarised in **Table S1**.

**Table S1. Demographics and clinical characteristics of the study participants.**

| Indicator | Asthmatics | Non-asthmatics | Asthma unknown[a] | Total |
|---|---|---|---|---|
| **AAGC** | | | | |
| N | 1695 | 115 | 0 | 1810 |
| Age (mean,sd,range) | 39,17.3,2-89 | 49,15.9,8-78 | - | 40,17.4,2-89 |
| Sex (% female) | 55.3 | 41.7 | - | 54.5 |
| Asthma onset (% ≤16, >16, unknown) | 64.5,29.5,6 | - | - | - |
| Atopy (% SPT+, SPT-, unknown) | 69.4,17.8,12.7 | 32.2,36.5,31.3 | - | 67.1,19,13.9 |
| Family history (% yes, no, unknown) | 69,30.4,0.6 | 0,73.9,26.1 | - | 64.6,33.2,2.2 |
| Smoker ever (% yes, no, unknown) | 43.1,48.8,8 | 40,27.8,32.2 | - | 42.9,47.5,9.6 |
| Medication ever (% yes, no, unknown) | 80.4,2.8,16.9 | 0,68.7,31.3 | - | 75.3,6.9,17.8 |
| Steroids ever (% yes, no, unknown) | 46.4,45.6,8 | 0,68.7,31.3 | - | 43.5,47.1,9.4 |
| Hospital ever (%yes, no, unknown) | 31.3,59.8,9 | 0,68.7,31.3 | - | 29.3,60.3,10.4 |
| **BUSSELTON** | | | | |
| N | 559 | 671 | 0 | 1230 |
| Age (mean,sd,range) | 49,17.5,18-89 | 58,16.3,17-91 | - | 54,17.3,17-91 |
| Sex (% female) | 59.7 | 57.4 | - | 58.5 |
| Asthma onset (% ≤16, >16, unknown) | 30.2,32.6,37.2 | - | - | - |
| Atopy (% SPT+, SPT-, unknown) | 60.5,24.7,14.8 | 31.7,49.6,18.6 | - | 44.8,38.3,16.9 |
| Family history (% yes, no, unknown) | 0,0,100 | 0,0,100 | - | 0,0,100 |
| Smoker ever (% yes, no, unknown) | 47.6,47.9,4.5 | 43.7,49.9,6.4 | - | 45.4,49,5.5 |
| Medication ever (% yes, no, unknown) | 0,0,100 | 0,0,100 | - | 0,0,100 |
| Steroids ever (% yes, no, unknown) | 0,0,100 | 0,0,100 | - | 0,0,100 |
| Hospital ever (%yes, no, unknown) | 0,0,100 | 0,0,100 | - | 0,0,100 |
| **QIMR_610K** | | | | |
| N | 387 | 1248 | 848 | 2483 |
| Age (mean,sd,range) | 21,9,12-65 | 31,14.5,12-89 | - | 29,14,12-89 |
| Sex (% female) | 47.3 | 62 | 55.9 | 57.6 |
| Asthma onset (% ≤16, >16, unknown) | 43.4,2.1,54.5 | - | - | - |
| Atopy (% SPT+, SPT-, unknown) | 8.3,1.8,89.9 | 0,0,100 | - | 1.3,0.3,98.4 |
| Family history (% yes, no, unknown) | 65.1,34.9,0 | 0,98.5,1.5 | - | 10.1,63.9,25.9 |
| Smoker ever (% yes, no, unknown) | 4.7,5.4,89.9 | 0,0,100 | - | 0.7,0.8,98.4 |
| Medication ever (% yes, no, unknown) | 6.5,3.6,89.9 | 0,0,100 | - | 1,0.6,98.4 |
| Steroids ever (% yes, no, unknown) | 1,9,89.9 | 0,0,100 | - | 0.2,1.4,98.4 |
| Hospital ever (%yes, no, unknown) | 0.5,0.8,98.7 | 0,0,100 | - | 0.1,0.1,99.8 |
| **QIMR_370K** | | | | |
| N | 28 | 667 | 979 | 1674 |

| | | | | |
|---|---|---|---|---|
| Age (mean,sd,range) | 33,8.6,23-54 | 33,13.4,17-92 | - | 33,13.2,17-92 |
| Sex (% female) | 53.6 | 64 | 41.7 | 50.8 |
| Asthma onset (% ≤16, >16, unknown) | 28.6,25,46.4 | - | - | - |
| Atopy (% SPT+, SPT-, unknown) | 82.1,17.9,0 | 0,0,100 | - | 1.4,0.3,98.3 |
| Family history (% yes, no, unknown) | 82.1,17.9,0 | 0,100,0 | - | 1.4,40.1,58.5 |
| Smoker ever (% yes, no, unknown) | 39.3,57.1,3.6 | 0,0,100 | - | 0.7,0.9,98.4 |
| Medication ever (% yes, no, unknown) | 78.6,21.4,0 | 0,0,100 | - | 1.3,0.4,98.3 |
| Steroids ever (% yes, no, unknown) | 7.1,92.9,0 | 0,0,100 | - | 0.1,1.5,98.3 |
| Hospital ever (%yes, no, unknown) | 0,7.1,92.9 | 0,0,100 | - | 0,0.1,99.9 |
| **TOTAL** | | | | |
| N | 2669 | 2701 | 1827 | 7197 |
| Age (mean,sd,range) | 39,18.3,2-89 | 39,18.6,8-92 | - | 39,18.5,2-92 |
| Sex (% female) | 55.1 | 60.5 | 48.2 | 55.4 |
| Asthma onset (% ≤16, >16, unknown) | 53.9,26.1,20 | - | - | - |
| Atopy (% SPT+, SPT-, unknown) | 58.8,16.9,24.2 | 9.2,13.8,77.1 | - | 25.2,11.5,63.3 |
| Family history (% yes, no, unknown) | 54.1,24.6,21.3 | 0,73.6,26.4 | - | 20,39.7,40.2 |
| Smoker ever (% yes, no, unknown) | 38.4,42.5,19.1 | 12.4,13.5,74.1 | - | 18.9,20.8,60.3 |
| Medication ever (% yes, no, unknown) | 52.8,2.5,44.7 | 0,2.9,97.1 | - | 19.6,2,78.4 |
| Steroids ever (% yes, no, unknown) | 29.7,31.2,39 | 0,2.9,97.1 | - | 11,12.7,76.3 |
| Hospital ever (%yes, no, unknown) | 19.9,38.1,41.9 | 0,2.9,97.1 | - | 7.4,15.2,77.4 |

[a] A total of 1,827 individuals provided no information about their asthma status and were included in the analysis as unselected controls to improve power (**Figure S2**). Family history was defined by the presence of one or more first-degree relatives with asthma. SPT=skin prick test.

Australian Asthma Genetics Consortium (AAGC) cohort (n=1,810)

In 2010, the AAGC was formed to promote a more rapid progress towards the identification of the genetic risk factors for asthma. As part of the AAGC, 2,030 subjects from Australia were genotyped using the Illumina 610K array, of which 1,810 (89%) passed genotyping quality control (QC) filters (**Table S2**), including 1,695 subjects with a lifetime physician diagnosis of asthma and 115 subjects without asthma. These individuals were ascertained and tested as part of five different studies which are summarised below: QIMR, TAHS, CAPS, MESCA and LIWA.

**Table S2. Sample QC filters applied to 2,160 samples selected for genotyping as part of the AAGC dataset.**

| Filter | QIMR | TAHS | CAPS | MESCA | LIWA | TOTAL |
|---|---|---|---|---|---|---|
| DNA selected | 613 | 419 | 121 | 195 | 812 | 2160 |
| DNA genotyped | 601 | 390 | 101 | 185 | 753 | 2030 |
| Call rate >0.98 | 601 | 388 | 100 | 184 | 748 | 2021 |
| DNA unique | 601 | 388 | 100 | 183 | 745 | 2017 |
| European ancestry[a] | 585 | 380 | 69 | 176 | 727 | 1937 |
| Unrelated[b] | 573 | 372 | 68 | 172 | 707 | 1892 |
| Sex check | 573 | 371 | 67 | 171 | 695 | 1877 |
| Phenotype check[c] | 571 | 371 | 53 | 170 | 645 | 1810 |
| | | | | | | |
| Asthmatics | 565 | 371 | 52 | 122 | 585 | 1695 |
| Non-asthmatics | 6 | 0 | 1 | 48 | 60 | 115 |

[a] Ancestry outliers were identified through multi-dimensional scaling analysis of identity-by-state (IBS) allele sharing, including 988 samples from the eleven HapMap3 populations (see also **Figure S1**).

[b] Pairs of relatives were identified through the analysis of genome-wide identical-by-descent (IBD) allele sharing and one individual from each pair was dropped from the analysis. For case-control pairs, the control individual was dropped. For case-case and control-control pairs, we dropped an individual based on the following study priority order: QIMR>MESCA>LIWA>CAPS>TAHS.
[c] Non-asthmatic individuals were removed from analysis if they reported a positive family history of disease.

*(A) Queensland Institute of Medical Research studies (QIMR, n=571).* Samples were drawn from three studies conducted at QIMR. The first and most significant source of participants was the 1995-1998 Asthma and Allergy study, which is described in detail elsewhere (1). Briefly, 3,073 subjects were recruited from 802 families that were registered on the Australian Twin Registry and had at least one twin that previously reported a history of wheezing in studies conducted at QIMR and by collaborators elsewhere in Australia. Participants completed a questionnaire that was designed to validate the diagnosis of asthma and to obtain data on respiratory symptoms, environmental exposures and family history of asthma. In addition, participants underwent clinical testing, including lung function and skin prick tests. For the present study, we selected one individual from each family that had available DNA and answered "Yes, told to me by a doctor" to the question "Have you ever had asthma?"; in families with multiple affected individuals, we selected the case with the highest asthma symptoms score (2). In total, 444 asthmatics were selected from this study for genotyping. Second, we supplemented this dataset with 156 unrelated asthmatic individuals that answered "Yes" to the question "Has a doctor ever diagnosed you as suffering from asthma?" in a study of atopic dermatitis conducted recently. Most of these (88%) also reported having had eczema diagnosed by a doctor. Lastly, we also selected 13 non-asthmatic individuals that answered "No" to the question "Has a doctor ever diagnosed you as suffering from Asthma?" and reported no asthma symptoms (including wheezing and chest tightness) in the Brisbane Adolescent Twin Study (BATS), which is described elsewhere (3). Briefly, twins were recruited in the context of an ongoing study of melanoma risk factors including benign melanocytic naevi, sun exposure time and pigmentation related variables. Twins were enlisted by contacting the principals of primary schools (first 7 years of education) in the greater Brisbane area, media appeals and by word of mouth. It is estimated that approximately 50 percent of the eligible birth cohort were recruited into the study. Twins were examined at age 12 years, and siblings at the same occasion if under 20 years of age. At the same time, twins and their parents completed questionnaires measuring melanoma risk factors, but that also included general health questions, such as asthma. Twins were again examined at age 14 and 16. Thus, a total of 613 individuals that participated in three QIMR studies were selected for genotyping; after QC, 571 samples were available for analysis.

*(B) Tasmanian Health Study (TAHS; n=371).* The details of the TAHS methodology have been reported elsewhere (4-7). In brief, TAHS commenced in 1968 by recruiting 8,583 Tasmanian children born in 1961, who were surveyed for respiratory problems and underwent clinical examination and lung function measurements. Subsequent follow-up surveys were completed at the ages of 13 (in 1974), 20 (in 1981), 31 (in 1992) and most recently at 44 (2004). The age 44 follow-up involved a postal survey were 85.2% (n=7,312) of the original 1968 cohort were traced to an address and a response rate of 78.4% (n=5,729) to a postal survey was achieved (8). A subgroup of these respondents, selected based on their participation in the previous follow-ups, samples of which were enriched for asthma, were invited to participate in a more detailed laboratory study. Of 2,373 invited for the laboratory study, 1,389 (58.5%) took part in a full laboratory visit, 354 (15%) completed a telephone questionnaire only, and 630 (26.5%) withdrew. In total, 419 participants with asthma and a DNA sample were selected for genotyping as part of the AAGC GWAS. Doctor diagnosed asthma-ever was defined as a positive response to the question "Have you ever had asthma?", followed by "Was this confirmed by a doctor?".

*(C) Childhood Asthma Prevention Study (CAPS; n=53).* Between September 1997 and November 1999, we recruited pregnant women whose unborn children were at increased risk of developing asthma because one or more parents or siblings had asthma or wheezing. We excluded those with a pet cat at home, strict vegetarians, women with a non-singleton pregnancy, and infants born earlier than 36 weeks of gestation (9, 10). We assessed their asthma status by clinical assessment at ages 18 months, 3, 5 and 8 years (10-13). Between 2001 and 2004 blood was collected for DNA extraction from all available subjects whose parents gave consent. Cases ascertained for this study were those for whom DNA was stored and who had a physician-diagnosis of asthma at age 8 years.

*(D) Melbourne Epidemiological Study of Childhood Asthma (MESCA; n=170).* In 1964, a community-based study was initiated in which a cohort of 410 children was selected from 30,000 7-year-old children in Melbourne, Australia (14). The cohort was stratified to provide three groups of approximately equal size of children with a history of wheezing of varying frequency and a control group with no history of wheezing. Wheezing and asthmatic status were initially determined by parental response to a questionnaire and were subsequently confirmed by clinical examination. These children were surveyed again at age 10, at which time an additional 83 children from the same 1957 birth cohort were recruited to enrich the group with severe asthma (15). The subjects were followed and reviewed at ages 14, 21, 28, 35 and 42 years (16, 17). At each review, their atopic status was determined (18) and lung function parameters were measured. At the 42-year follow-up, blood was taken from each of the available 244 individuals for genotyping and

total IgE measurements, including the 195 individuals that are part of this study. Informed consent was obtained from all subjects, and the study was approved by the ethics committee of the Royal Children's Hospital, Melbourne.

*(E) Lung Institute of Western Australia (LIWA; n=645).* The LIWA cohort was recruited from patients and normal subjects resident in Perth, Western Australia. Asthma was defined as physician-diagnosed adult asthma (18 and over) and all had face to face clinical review. The non-asthmatic control subjects had no history of asthma as defined by medical staff interview. All participants were unrelated with self-reported European ancestry between 18 and 89 years of age. The control subjects were recruited by random mail-out using the telephone directory to generate random addresses and the asthma patients were similarly recruited but supplemented for the more severe patients through physician referrals. All subjects gave written informed consent and completed a comprehensive questionnaire which was used in the assessment of phenotype. Approximately 20ml of blood was obtained from each participant and lung function was assessed by spirometry. Assessment of atopic status was based on a positive skin prick reaction (weal diameter >3mm) to at least one of five common aeroallergens: cat, dog, house dust mite, mould mix (Alternaria tenuis, Aspergillus mix, Cladosporium, Penicillin mix) and grass pollen mix (Kentucky Blue, Orchard, Red Top, Timothy, Sweet Vernal, Meadow Fescue, Perennial Rye). The study protocol was approved by the Human Research Ethics Committee of the Sir Charles Gairdner Hospital. After QC, 645 samples from this cohort were available for analysis.

Busselton Health Study (n=1,230)

This cohort has been described in detail previously (19) and was included in the Moffatt et al. (20) asthma GWAS. Briefly, residents of the town of Busselton in the southwest of Western Australia have been involved in a series of health surveys since 1966. The population is predominantly of European origin. In 1994/95 there was a follow-up study involving a subset of those who had attended any of the previous surveys. Cases of asthma were defined as those who reported physician-diagnosed asthma at any survey that they attended from 1966 to 1994 (answer 'Yes' to 'Has your doctor ever told you that you had asthma?'). Controls are those who have consistently answered 'No' to 'Has your doctor ever told you that you had asthma?' at all previous surveys that they have attended from 1996 to 1994. After QC a total of 1,230 unrelated subjects were available for analysis.

QIMR GWAS cohort (n=4,157)

Between 2007 and 2009, a total of 15,259 DNA samples were genotyped with Illumina 370K or 610K arrays as part of six projects conducted by QIMR. This dataset, including QC checks, is described in detail in Medland et al. (21). For this study, we selected a subset of 2,483 samples genotyped with the 610K array and 1,674 with the 370K array that were confirmed to be unrelated to the samples included in the AAGC cohort. Of the 2,483 genotyped with the 610K array (QIMR_610K dataset), (1) 387 had answered "Yes" to the question "Has a doctor ever diagnosed you as suffering from asthma?" in the QIMR Asthma (n=39) or BATS (n=348) studies described above; (2) 1,248 answered "No" or "Never" to the questions "How often have you had asthma?", "Have you ever had asthma?" or "Has a doctor ever diagnosed you as suffering from asthma?" in the Canberra (22) (n=53), Alcohol (23) (n=527), Alcohol Relatives (24) (n =132), Aged (25) (n=165) or BATS (n=371) studies; and (3) 848 provided no information about their asthma status and were included in the analysis as unselected controls to improve power (**Figure S2**). Similarly, of the 1,674 selected samples genotyped with the 370K array (QIMR_370K dataset), (1) 28 answered "Yes, told to me by a doctor" to the question "Have you ever had asthma?" in the QIMR Asthma study; (2) 667 answered "No" or "Never" to the questions "How often have you had asthma?" or "Have you ever had asthma?"in the Canberra (n=60), Alcohol (n=491), Alcohol Relatives (n=26) or Aged (n=90) studies; and (3) 979 provided no information about their asthma status and were included in the analysis as unselected controls. Note that the partition of QIMR cases and controls between the AAGC GWAS cohort and this QIMR GWAS cohort is purely a function of whether the samples had been previously genotyped in the context of another study (QIMR GWAS) or not previously genotyped (AAGC GWAS).

Combined dataset

Thus, across the four datasets (AAGC, Busselton, QIMR_610K and QIMR_370K), genotype data were available for 7,197 subjects, including 2,669 physician-diagnosed asthmatics and 4,528 controls. Amongst the asthmatic cases were 1,438 (54%) subjects with childhood onset asthma (defined by an age-of-onset at or before age 16), 697 (26%) subjects with later onset asthma (age-of-onset after the age of 16) and 534 (20%) with unknown age-of-onset. All subjects were confirmed to be unrelated and of European ancestry (**Figure S1**) through the analysis of genome-wide allele sharing. This dataset includes 4,259 samples that have not been previously included in any asthma GWAS, 1,708 that were included in Ferreira et al. (26) and 1,230 samples from the Busselton cohort included in the GABRIEL study (20).

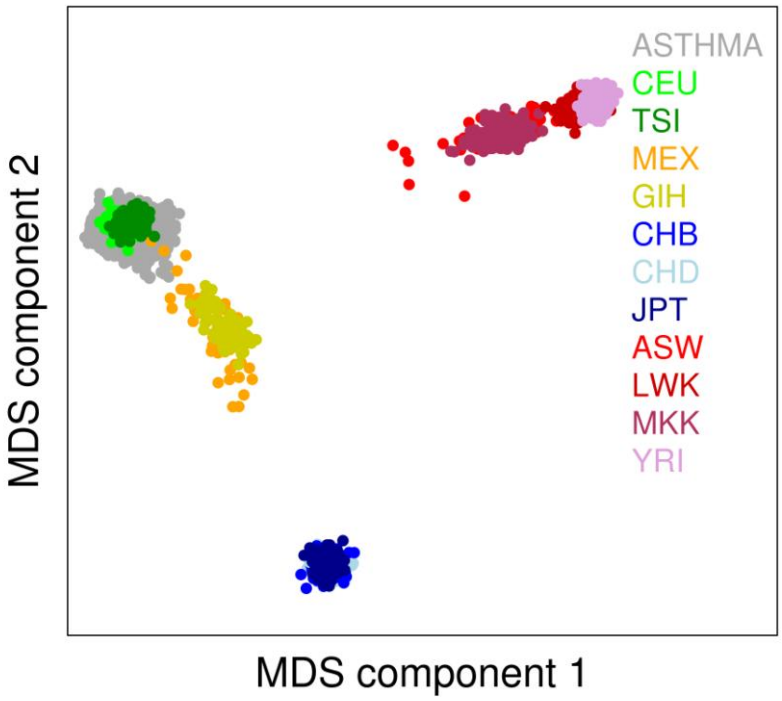


**Figure S1. Population substructure analyses.** Results from the multidimensional scaling (MDS) analysis of identity-by-state (IBS) distances calculated between all pairs of individuals that passed QC filters from the cohort analysed in this study (ASTHMA) and the eleven HapMap 3 populations. CEU=Utah residents with Northern and Western European ancestry from the CEPH collection. TSI=Toscani in Italia. MEX=Mexican ancestry in Los Angeles, California. GIH=Gujarati Indians in Houston, Texas. CHB=Han Chinese in Beijing, China. CHD=Chinese in Metropolitan Denver, Colorado. JPT=Japanese in Tokyo, Japan. ASW=African ancestry in Southwest USA. LWK=Luhya in Webuye, Kenya. MKK=Maasai in Kinyawa, Kenya. YRI=Yoruba in Ibadan, Nigeria.

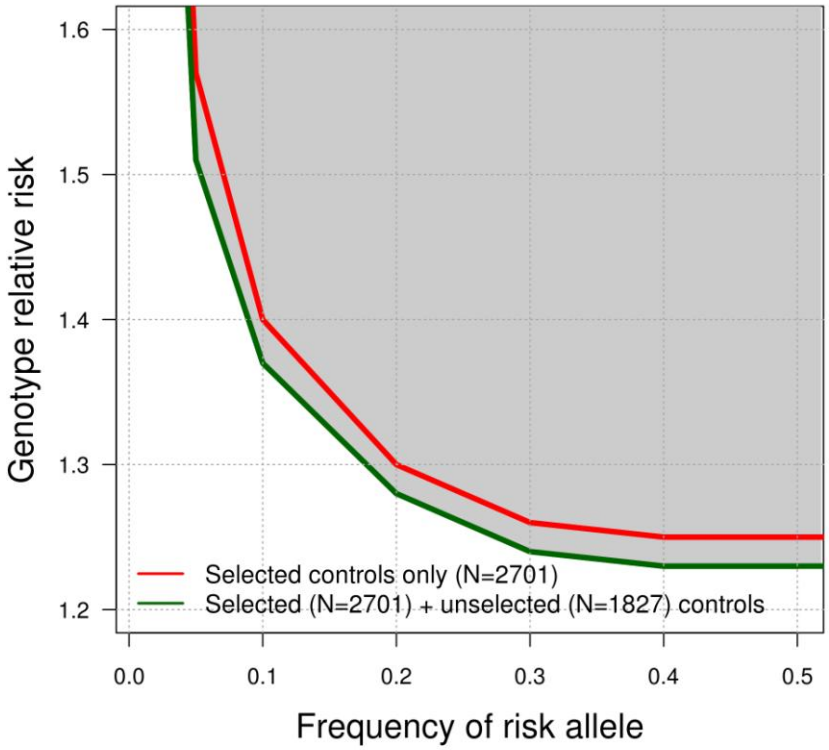


**Figure S2. Power to detect association for a range of risk allele frequencies.** The *y*-axis represents the smallest detectable genotype relative risk (GRR) with 80% power ($\alpha = 5x10^{-8}$) under a multiplicative model and assuming a disease prevalence of 10%. Two situations were considered. The first (green line below), which corresponds to the main analysis reported in the main text, represents the analysis of 2,669 asthma cases and 2,701 asthma-free plus 1,827 asthma-unknown controls. For comparison, we also show the smallest detectable GRR expected if the 1,827 individuals with an unknown asthma status were not included as controls in the analysis (red line above). For a higher disease prevalence (15% or 20%), the improvement in power by including the unselected controls is weaker (not shown).

## Ascertainment of samples included in the follow-up phase

We carried out a replication study to validate in an independent panel of 25,358 individuals the association with seven SNPs identified with $P \leq 5x10^{-6}$ in the main association analysis (n=2) or in the meta-analysis of results from our study and Moffatt et al. (20) (n=5). This panel included 3,322 asthmatic individuals and 22,036 controls ascertained and previously genotyped through four different studies (Raine, QIMR, NTR, APCAT) as described below. SNPs were tested in each study separately and then combined by performing a fixed-effects meta-analysis with METAL (27).

Raine cohort (n=1,275)
Recruitment of the Western Australian Pregnancy (Raine) cohort has previously been described in detail (28-30). In brief, between 1989 and 1991, 2,900 pregnant women were recruited prior to 18-weeks gestation into a randomised controlled trial to evaluate the effects of repeated ultrasound in pregnancy. Children have been comprehensively phenotyped from birth to 21 years of age (average ages of one, two, three, six, eight, ten, fourteen, seventeen and twenty-one) by trained members the Raine research team. Data collection included questionnaires completed by the child's primary carer and by the adolescent from age 14, physical assessments by trained assessors at all follow-up years, DNA collection from year 14 follow-up. Cases of asthma (n=654) are defined as subjects who answered "Yes" to the question "Has a doctor (GP, paediatrician, respiratory specialist) ever told you that your child has asthma?" at least once in year 5, year 8, year 10, year 14 and year 17 follow-ups. Controls (n=621) answered "No" to the same question across all available time points. Only individuals of European ancestry were included in the analysis. The study was conducted with appropriate institutional ethics approval, and written informed consent was obtained from all mothers. DNA samples were genotyped using Illumina 660 Quad arrays; the seven SNPs considered for follow-up passed strict quality control filters.

QIMR follow-up cohort (n=2,808)
Amongst the 15,259 individuals previously genotyped by QIMR (see QIMR GWAS cohort above), there were 602 individuals with self-reported (but not physician-diagnosed) lifetime asthma that were not included in the Australian GWAS to ensure that results were based on a phenotype definition that matched that used by other asthma GWAS. However, we have previously noted that a lifetime asthma definition is highly correlated (0.99) with a lifetime physician-asthma definition (1) and that this self report is informative in mapping genuine asthma loci (26). We therefore selected these 602 asthmatics for the replication phase, including 306 samples that were genotyped with the 610K array and 296 with the 370K array. Of those 306 genotyped with the 610K array, (1) 235 answered "Sometimes", "Only as a child", "Often" or "Quite often" to the question "How often have you had asthma?" in the Canberra (22) (n=12), Alcohol (23) (n=144) or Alcohol Relatives (24) (n=79) studies; and (2) 71 answered "Yes", "Now", "Past" or "Now and Past" to the question "Have you ever had asthma?" in the Asthma (1) (n=45) or Aged (25) (n=26) studies. Similarly, of the 296 individuals genotyped with the 370K array, (1) 257 answered "Sometimes", "Only as a child", "Often" or "Quite often" to the question "How often have you had asthma?" in the Canberra (n=9), Alcohol (n=184) or Alcohol Relatives (n=64) studies; and (2) 39 answered "Yes", "Now", "Past" or "Now and Past" to the question "Have you ever had asthma?" in the Asthma (n=28) or Aged (n=11) studies. The average age of participants (62% female) was 35 years (range 18-80); there was no additional clinical information available for these participants. Controls for this cohort were drawn from the same set of 15,259 genotyped individuals from QIMR. In this case, we selected 2,206 individuals (all females) that were genotyped with the 610K array as part of a GWAS for endometriosis (31) and that had not been included in our GWAS discovery phase. No asthma information was available for these individuals and so they constitute a group of unselected controls. All individuals were confirmed to be of European ancestry through genome-wide IBS analyses.

The Netherlands Twin Registry cohort (NTR; n=2,671)
Data from two samples drawn from the NTR were included for analysis.

*(A) NTR1 sample (n=1,612).* Participants were drawn from the GAIN-MDD study, which is a case-control study of major depressive disorder in unrelated individuals. Average age of participants (65% female) was 43.8 years (SD = 13.6); participants gave informed consent to participation, and all studies were approved by appropriate ethics committees. Genotyping was conducted by Perlegen Sciences (Mountain View, CA, USA) with the use of a set of four proprietary, high-density oligonucleotide arrays. The SNP quality-control process is described in detail elsewhere (32). A total of 427,303 autosomal SNPs passed QC criteria. We inferred unmeasured HapMap SNPs in each panel separately using MACH and the phased haplotype data from the CEU HapMap samples (phase I+II, release22, build 36) as reference. Following imputation, data for 2.5 million Hapmap SNPs were available; of these, we dropped SNPs with an imputation score < 0.3, MAF < 0.01 or with a Hardy-Weinberg equilibrium test with $P < 10^{-6}$. Data were available for 205 and 1,407 individuals who respectively answered "Yes" and "No" to the question "Have you ever had asthma diagnosed by a doctor?".

*(B) NTR2 sample (n=1,059).* This sample was randomly drawn from the NTR-Biobank study (33). We optimized the number of unrelated individuals who were genotyped. Average age of participants (61% female) was 49.2 years (SD = 14.0); participants gave informed consent to participation, and all studies were approved by appropriate ethics committees. Genome-wide SNP-genotyping was performed with Illumina 660W arrays; QC excluded SNPs based on MAF < 0.01, missing genotype rate > 0.05 or a *P*-value < $10^{-5}$ in a test of Hardy-Weinberg equilibrium, leaving 515,781 SNPs available for analysis. Samples were excluded if they showed evidence for contamination by excessive allele sharing with multiple samples and excessive levels of heterozygosity (F < -0.10). Samples were also excluded if they had a higher than 90% genotype missing rate. Subsequently, genotypes of ~3.8 million SNPS were imputed with Impute (34), using the HapMap CEU data (release 24, NCBI build 36), available from the Impute website, as reference. Imputed SNPs were excluded if they had a minor allele frequency < 0.01 or a properinfo < 0.3, leaving 2,147,160 autosomal SNPs for analysis. Data were available for 145 and 914 individuals who respectively answered "Yes" and "No" to the question "Have you ever had asthma diagnosed by a doctor?".

Analysis in Population-based Cohorts of Asthma Traits (APCAT; n=18,604)

The APCAT consortium includes information for 1,716 physician-diagnosed asthmatics and 16,888 asthma-free controls (**Table S3**) with available genome-wide genotype data that participated in six population-based studies from Finland (Helsinki Birth Cohort [HBC]; Health 2000 [H2000]; Finrisk, including Finrisk 1992, 1997, 2002 and 2007; the Northern Finland Birth Cohort 1966 [NFBC66]; and the Young Finns Study) and the United States (Framingham Heart Study). These studies are described in detail in the references provided in **Table S3** and summarised below. Unmeasured HapMap SNPs were imputed with MACH. SNPs were tested for association with asthma using logistic regression and considering SNP additive effects. Age (H2000, Finrisk, Young Finns Study), sex and four ancestry-informative principal components were included as covariates. Three separate association analyses were conducted, corresponding to the following panels: (1) the Framingham Heart Study; (2) the NFBC66 study; and (3) the combined analysis of the HBC, H2000, Finrisk and Young Finns studies. Results from the three panels were then meta-analysed with METAL (27).

**Table S3. Studies that contributed data to the APCAT consortium.**

| Study | N cases | N controls | Reference |
|---|---|---|---|
| Helsinki Birth Cohort Study (HBCS) | 123 | 1533 | (35) |
| Health 2000 (H2000) | 153 | 1841 | (36) |
| Finrisk | 160 | 1705 | (37) |
| Northern Finland Birth Cohort 1966 (NFBC66) | 364 | 3502 | (38) |
| Young Finns Study | 119 | 1844 | (39) |
| Framingham Heart Study | 797 | 6463 | (40) |
| Total | 1716 | 16888 | |

*(A) Helsinki Birth Cohort Study (HBCS; n=1,656).* The HBCS includes 8,760 subjects born in Helsinki between 1934 and 1944 (35). Between 2001 and 2004, a representative subset of 928 males and 1,075 females participated in a clinical study focusing upon cardiovascular and metabolic outcomes and cognitive functions. Information on asthma, smoking and alcohol intake is available from questionnaires for 2,003 individuals who participated in the clinical study. Information on hospitalization due to alcohol abuse is available from the National Hospital Discharge Register. Psychological questionnaires have been used to assess personality characteristics including data on impulsivity. Genotyping was performed with the Illumina 670K array, with 1,656 individuals included in this analysis.

*(B) Health 2000 Study (H2000; n=1,994).* The protocol for this study is described in detail elsewhere (36) and included home interview, completion of several questionnaires, laboratory and anthropometrical measurements, spirometry with bronchodilator test and clinical examination by a physician. Further information was obtained by record linkage with the National Hospital Discharge Register and the National Social Insurance Institutions register data on reimbursement of asthma medication. Asthma diagnosis was based on all of the above mentioned data and confirmed by the physician. Samples were genotyped with Illumina 610K or 370K arrays, with 1,994 individuals included in this analysis.

*(C) Finrisk study (n=1,865).* This study is a population survey of risk factors for chronic diseases in Finland (37). The survey has been executed every five years from 1972 using independent, random and representative population samples from five geographical areas of the country. Participants have completed a health-related questionnaire and undergone a physical examination including measurement of anthropometric traits and blood draw. A total of 1,865 individuals with available physician-diagnosed asthma status were genotyped with Illumina 610K or Affymetrix 6.0 arrays.

*(D) Northern Finland Birth Cohort 1966 (NFBC66; n=3,866).* The NFBC66 study is an on-going follow-up study of people whose expected date of birth was in 1966 in the provinces of Oulu and Lapland (38). Primary clinical data collection on parents and the child occurred prenatally and at birth. Data collection on the child continued at several age points. At age 31, participants were sent a postal questionnaire and invited for clinical assessment where DNA was collected with informed consent. Information on asthma, hay fever and smoking used in this analysis was collected at this time point, with some incomplete information filled during clinical assessment. More details about this study can be found at http://kelo.oulu.fi/NFBC/. Genotyping was performed with Illumina Infinium for 339,629 SNPs.

*(E) The Young Finns Study (n=1,963).* The Young Finns cohort is a longitudinal population study sample on the evolution of cardiovascular risk factors from childhood to adulthood (39). The first cross-sectional survey was conducted in 1980 in five Finnish university cities and included 3,596 participants who were in the age groups of 3, 6, 9, 12, 15, and 18 years and were randomly chosen from the national population register; equal ratios of males and females were selected in each age group. In 2007, 2,204 subjects now aged 30 to 45 years participated in the latest follow-up study. Of these, 1,963 individuals with available physician-diagnosed asthma status were genotyped with the Illumina 670K array.

*(F) Framingham Heart Study (FHS; n=7,260).* The FHS is a collection of cohorts recruited to investigate cardiovascular disease and its risk factors as described in detail elsewhere (40). Asthma was classified based on self-report of a physician's diagnosis. Genotyping was performed using the Affymetrix 500K array and supplemented with the HuGeneFocused 50K array.

**Table S4. QC filters for SNP data in the four individual projects that contributed to the Australian GWAS.**

| QC step | 1. AAGC | 2. QIMR_610K | | | 3. QIMR_370K | | 4. Busselton |
|---|---|---|---|---|---|---|---|
| | | 2.1. MIG | 2.2. ADOL | 2.3. GL | 3.1. ALCO_C | 3.2. ALCO_D | |
| Illumina array used | 610K | 610K | 610K | 610K | 370K | 370K | 610K |
| N SNPs before QC | 598821 | 592385 | 592392 | 589296 | 343955 | 344962 | 582892 |
| GenCall < 0.7 | 39603 | 46931 | 47418 | 36877 | 24494 | 27459 | [a] |
| Call rate < 95% | 1244 | 8038 | 8447 | 12455 | 11584 | 7537 | 4034 |
| Hardy Weinberg equilibrium $P < 10^{-6}$ | 1498 | 1221 | 2841 | 15474 | 4318 | 1194 | 927 |
| MAF < 0.01 | [b] | 33347 | 33347 | 28607 | 7874 | 8976 | 32321 |
| N SNPs after dataset-specific QC | 556476 | 530922 | 529379 | 531042 | 323093 | 321267 | 545610 |
| N SNPs after merging 2.1, 2.2 and 2.3 | 556476 | 561815 | | | 323093 | 321267 | 545610 |
| N SNPs after merging 3.1 and 3.2 | 556476 | 561815 | | | 292083 | | 545610 |
| N SNPs after merging 1 and 2 | 508240 | | | | 292083 | | 545610 |
| MAF < 0.01 | 169 | | | | 50 | | NA |
| Hardy Weinberg equilibrium $P < 10^{-6}$ | 2 | | | | 0 | | NA |
| MAF difference between 1 and 2[b] | 1104 | | | | NA | | NA |
| N SNPs after QC | 506965 | | | | 292033 | | 545610 |

[a] GenCall score not available in this cohort.

[b] The AAGC dataset included predominantly (94%) asthmatic cases and so a MAF filter was only applied after merging this dataset with the QIMR_610K data. Furthermore, this over-representation of cases in one dataset could be a potential confounder for the analysis of SNPs whose genotype calling is affected by dataset-specific technical artefacts. To identify and remove such SNPs, we compared the allele frequency between AAGC cases (N=1,695) and QIMR_610K cases (N=397), as well as between AAGC controls (N=115) and QIMR_610K controls (N=1,267). The genomic inflation factor for these analyses was 1.001 and 1.014, respectively, indicating that there were no systematic allele frequency differences between the two datasets. We nonetheless removed a small subset of 1,104 SNPs (0.2%) that had significant ($P < 0.001$) allele frequency differences between the two datasets.

MAF=minor allele frequency. QC=quality control.

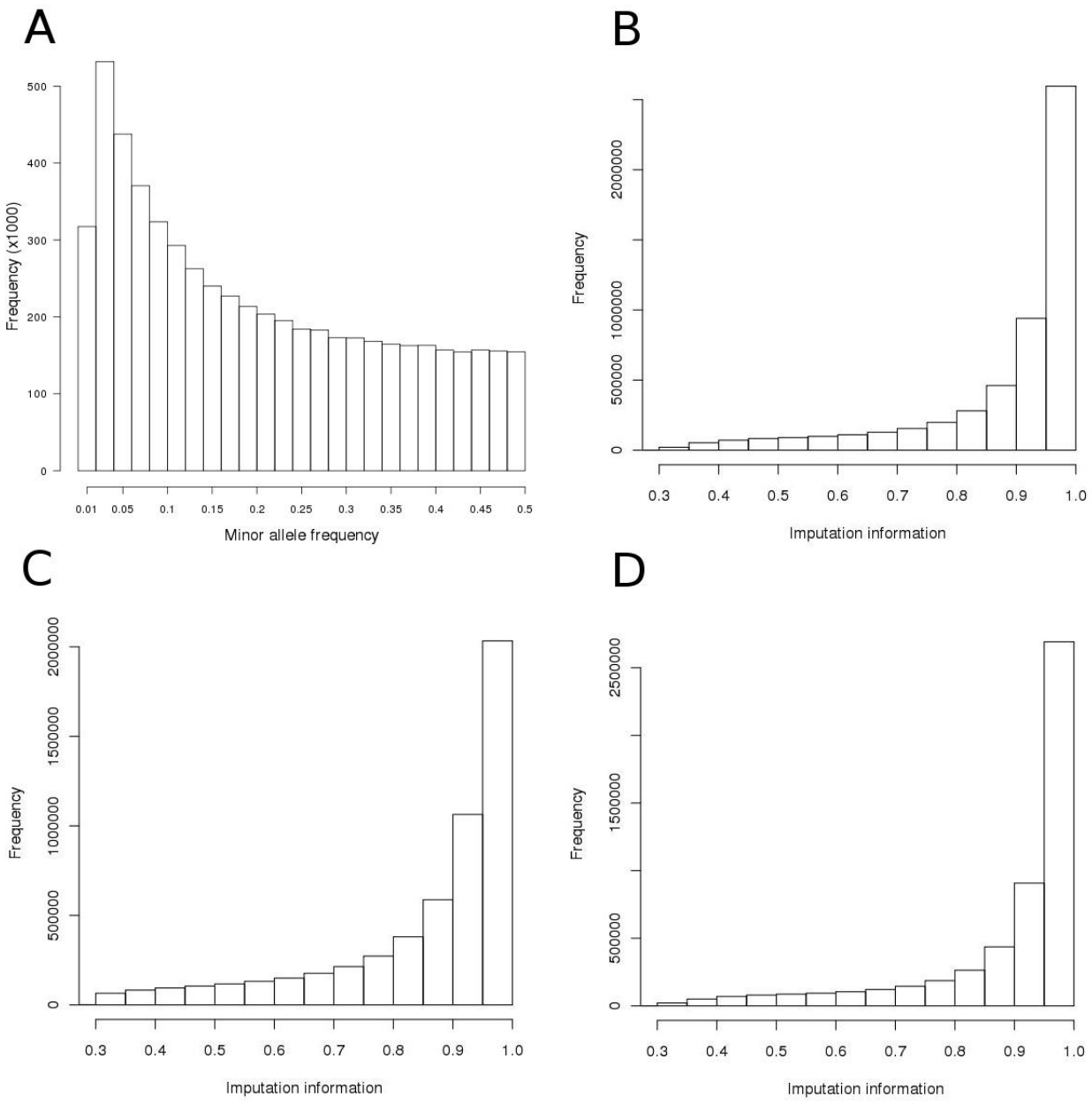


**Figure S3. Characteristics of 5.7 million SNPs analysed in this study.** (A) Minor allele frequency spectrum. (B – D) Imputation information for SNPs that were imputed in the AAGC+QIMR_610K datasets (B), QIMR_370K dataset (C) and Busselton dataset (D).

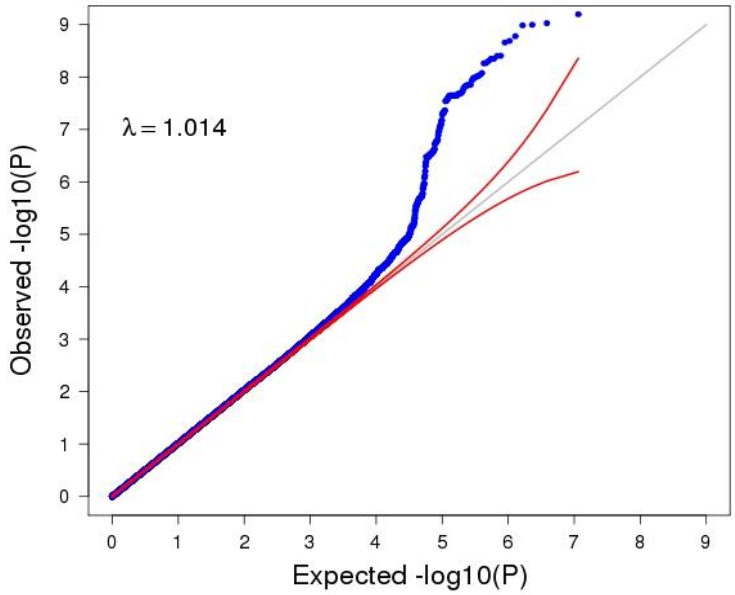


**Figure S4. Quantile-quantile plot for results from the Australian GWAS.** Observed (*y*-axis) and expected (*x*-axis) association results ($-\log_{10}P$-value) for 5.7 million SNPs included in our main association analysis of asthma in 2,669 cases and 4,528 controls. The genomic inflation factor ($\lambda$) of this analysis was 1.014.

**Table S5. Main association results ($P \leq 5x10^{-6}$) for the Australian GWAS (2,669 asthmatics and 4,528 controls).**

| Locus | Position[a], bp | SNP, Allele | Closest gene, bp distance | MAF | Odds Ratio (95% CI) | Association *P*-value | Heterogeneity test[b] *P*-value (*k*; $I^2$, 95% CI) |
|---|---|---|---|---|---|---|---|
| 17q21 | 35360125 | rs8071050, G | *GSDMA*,12626 | 0.48 | 1.27 (1.18-1.35) | $6.3x10^{-10}$ | 0.3069 (3; 15, 0-91) |
| 12q24 | 119562269 | chr12:119562269, T | *CABP1*,10484 | 0.02 | 1.85 (1.45-2.38) | $8.3x10^{-7}$ | 0.4918 (3; 0, 0-90) |
| 16q24 | 82819610 | rs7196274, A | *KCNG4*,6287* | 0.27 | 0.81 (0.74-0.88) | $1.3x10^{-6}$ | 0.0853 (3; 59, 0-88) |
| 8q22 | 95245191 | rs11776675, C | *CDH17*,36624* | 0.32 | 1.21 (1.12-1.30) | $2.7x10^{-6}$ | 0.3199 (3; 12, 0-91) |
| 5q22 | 110495398 | rs1438673,T | *WDR36*,1299 | 0.50 | 0.84 (0.78-0.90) | $3.2x10^{-6}$ | 0.9332 (3; 0, 0-90) |

[a] Base pair positions (bp) correspond to build 36.
[b] Results for the Breslow-Day test of heterogeneity.
The five SNPs were imputed with information > 0.8 for 610K array datasets (AAGC+QIMR_610K and Busselton) and > 0.5 for the QIMR_370K dataset.
MAF=minor allele frequency. *k*=number of analysis panels.

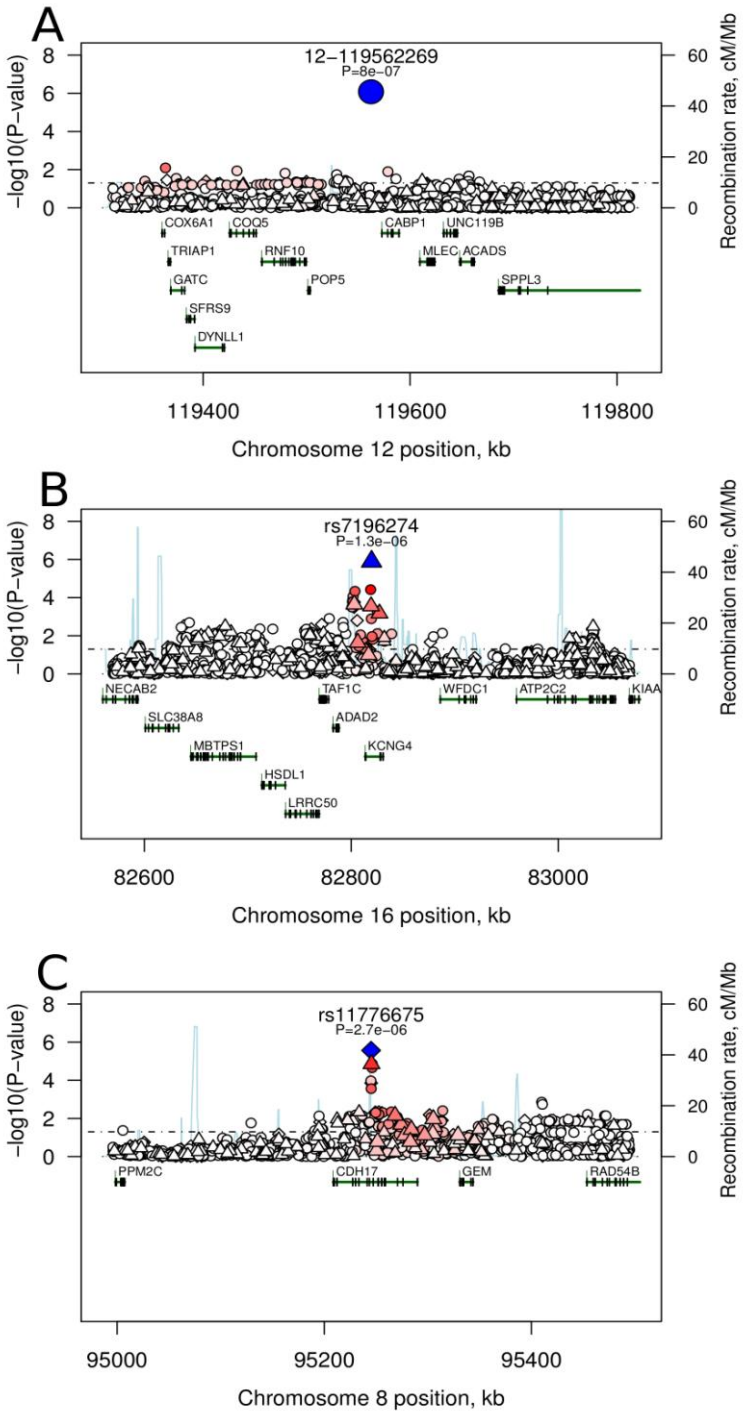


**Figure S5. Regional association results (-log$_{10}$*P*-value, *y*-axis) for chromosomes 12q24 (A), 16q24 (B) and 8q22 (C).** The most-associated SNP for each region is shown in blue, and the color of the remaining markers reflects the linkage disequilibrium ($r^2$) with the top SNP in each panel (increasing red hue associated with increasing $r^2$). SNPs directly genotyped in at least one of the three datasets (AAGC+QIMR_610K, QIMR_370K or BUSSELTON) are represented by triangles; SNPs imputed in all three datasets are represented by diamonds (HapMap 3 SNPs) or circles (1000 Genomes SNPs). The recombination rate (second *y*-axis) is plotted in light blue and is based on the CEU HapMap population. Exons for each gene are represented by vertical bars.

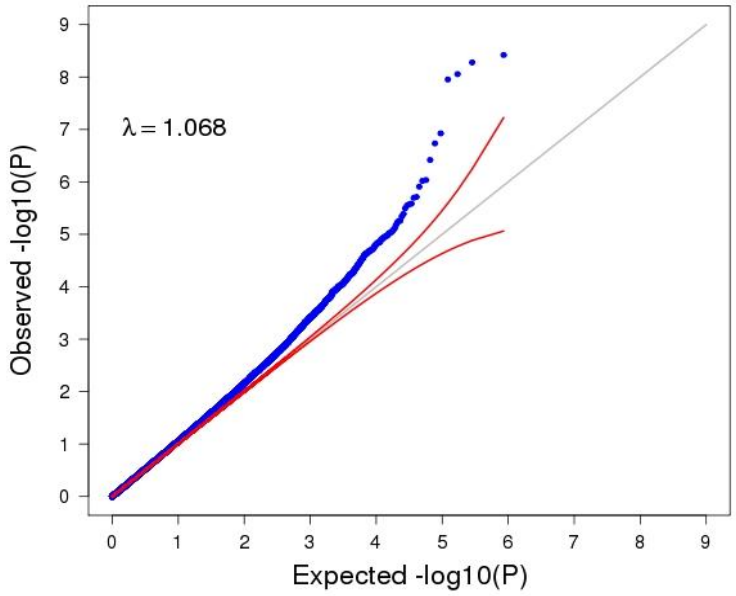


**Figure S6. Meta-analysis quantile-quantile plot.** Observed (*y*-axis) and expected (*x*-axis) association results ($-\log_{10}P$-value) for 421,334 autosomal SNPs included in a meta-analysis of results from the Australian GWAS and the GABRIEL (20), comprising 12,475 asthmatics and 19,967 controls (after excluding overlapping samples). SNPs with $P<5\times10^{-8}$ in the GABRIEL (20) are not shown. The genomic inflation factor ($\lambda$) of this analysis was 1.068.

**Table S6. SNPs most associated with asthma ($P \leq 5x10^{-6}$) in the meta-analysis of results from the Australian GWAS and GABRIEL (20).**

| Locus | SNP, allele | Nearest gene | Australian GWAS[a] | | | | | GABRIEL[b] | | | | Meta-analysis | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | MAF | OR | SE | Association *P*-value | Heterogeneity test[c] *P*-value (*k*; $I^2$, 95% CI) | OR | SE | Association *P*-value | Heterogeneity test[c] *P*-value (*k*; $I^2$, 95% CI) | OR (95% CI) | Association *P*-value | Heterogeneity test[c] *P*-value (*k*; $I^2$, 95% CI) |
| 17q21 | rs8079416,C | *ORMDL3* | 0.44 | 1.28 | 0.043 | $1.1x10^{-08}$ | 0.4890 (2;0,0-100) | 1.17 | 0.020 | $7.1x10^{-16}$ | 0.0480 (36;30,0-54) | 1.19 (1.15-1.23) | $2.4x10^{-22}$ | 0.0246 (38;34,2-56) |
| 2q12.1 | rs3771166,A | *IL1RL1* | 0.39 | 0.86 | 0.044 | 0.0005 | 0.5630 (2;0,0-100) | 0.87 | 0.021 | $3.5x10^{-12}$ | 0.1808 (36;18,0-46) | 0.86 (0.83-0.90) | $7.9x10^{-15}$ | 0.1809 (38;17,0-45) |
| 9p24.1 | rs1342326,C | *IL33* | 0.16 | 1.20 | 0.056 | 0.0010 | 0.7232 (2;0,0-100) | 1.20 | 0.026 | $8.7x10^{-12}$ | 0.2169 (36;15,0-44) | 1.20 (1.14-1.26) | $3.5x10^{-14}$ | 0.2127 (38;15,0-43) |
| 17q12 | rs2271308,T | *STARD3* | 0.26 | 1.12 | 0.048 | 0.0217 | 0.9535 (2;0,0-100) | 1.14 | 0.022 | $2.4x10^{-09}$ | 0.4752 (36;0,0-38) | 1.14 (1.09-1.18) | $1.7x10^{-10}$ | 0.4509 (38;1,0-38) |
| 22q12.3 | rs2284033,A | *IL2RB* | 0.43 | 0.90 | 0.043 | 0.0132 | 0.6578 (2;0,0-100) | 0.89 | 0.020 | $1.2x10^{-08}$ | 0.9228 (36;0,0-38) | 0.89 (0.86-0.92) | $5.0x10^{-10}$ | 0.9230 (38;0,0-37) |
| 5q31.1 | rs6871536,C | *RAD50* | 0.19 | 1.11 | 0.053 | 0.0407 | 0.0624 (2;71,0-100) | 1.14 | 0.024 | $1.9x10^{-08}$ | 0.0921 (36;25,0-50) | 1.14 (1.09-1.19) | $2.4x10^{-09}$ | 0.0748 (38;26,0-50) |
| 15q22.33 | rs744910,G | *SMAD3* | 0.49 | 1.06 | 0.043 | 0.1452 | 0.3074 (2;4,0-100) | 1.12 | 0.020 | $3.9x10^{-09}$ | 0.8515 (36;0,0-38) | 1.11 (1.07-1.15) | $2.7x10^{-09}$ | 0.8477 (38;0,0-37) |
| 15q22.2 | rs11071559,T | *RORA* | 0.14 | 0.85 | 0.065 | 0.0103 | 0.2650 (2;20,0-100) | 0.85 | 0.030 | $1.1x10^{-07}$ | 0.8429 (36;0,0-38) | 0.85 (0.81-0.90) | $3.8x10^{-09}$ | 0.7652 (38;0,0-37) |
| 5q22.1 | rs1043828,C | *WDR36* | 0.35 | 1.17 | 0.044 | 0.0004 | 0.8681 (2;0,0-100) | 1.10 | 0.021 | $3.1x10^{-06}$ | 0.1808 (36;18,0-46) | 1.11 (1.07-1.15) | $1.1x10^{-08}$ | 0.1488 (38;19,0-47) |
| 11q13.5 | rs7130588,G | *LRRC32* | 0.36 | 1.12 | 0.044 | 0.0111 | 0.3523 (2;0,0-100) | 1.10 | 0.021 | $3.2x10^{-06}$ | 0.1602 (36;19,0-47) | 1.10 (1.06-1.15) | $1.2x10^{-07}$ | 0.1894 (38;17,0-44) |
| 6p21.32 | rs3763309,A | *BTNL2* | 0.22 | 1.07 | 0.050 | 0.1931 | 0.7730 (2;0,0-100) | 1.13 | 0.024 | $4.0x10^{-07}$ | 0.1255 (36;22,0-48) | 1.12 (1.07-1.17) | $2.8x10^{-07}$ | 0.0760 (38;26,0-50) |
| 9p24.1 | rs340908,T | *IL33* | 0.18 | 0.98 | 0.056 | 0.7735 | 0.2951 (2;9,0-100) | 0.87 | 0.026 | $3.5x10^{-08}$ | 0.0662 (36;28,0-52) | 0.89 (0.85-0.93) | $3.2x10^{-07}$ | 0.0648 (38;27,0-51) |
| 13q21.31 | rs3119939,C | *PCDH20* | 0.49 | 0.91 | 0.043 | 0.0254 | 0.6933 (2;0,0-100) | 0.92 | 0.020 | $1.3x10^{-05}$ | 0.3354 (36;8,0-38) | 0.92 (0.88-0.95) | $9.6x10^{-07}$ | 0.5472 (38;0,0-37) |
| 6p21.32 | rs241425,A | *TAP2* | 0.44 | 0.89 | 0.043 | 0.0079 | 0.8348 (2;0,0-100) | 0.92 | 0.020 | $3.2x10^{-05}$ | 0.6721 (36;0,0-38) | 0.91 (0.88-0.95) | $9.9x10^{-07}$ | 0.7766 (38;0,0-37) |
| 10q21.1 | rs7922491,A | *PRKG1* | 0.11 | 1.15 | 0.067 | 0.0350 | 0.0148 (2;83,0-100) | 1.15 | 0.032 | $1.2x10^{-05}$ | 0.8763 (36;0,0-38) | 1.15 (1.09-1.22) | $1.2x10^{-06}$ | 0.6991 (38;0,0-37) |
| 1q21.3 | rs4129267,T | *IL6R* | 0.40 | 1.09 | 0.043 | 0.0372 | 0.0417 (2;76,0-100) | 1.09 | 0.020 | $1.9x10^{-05}$ | 0.1734 (36;18,0-46) | 1.09 (1.05-1.13) | $2.0x10^{-06}$ | 0.1839 (38;17,0-44) |
| 7q22.3 | rs6967330,A | *CDH28* | 0.16 | 1.11 | 0.058 | 0.0692 | 0.4156 (2;0,0-100) | 1.12 | 0.027 | $1.8x10^{-05}$ | 0.0266 (36;34,0-56) | 1.12 (1.07-1.17) | $3.2x10^{-06}$ | 0.0431 (38;30,0-53) |
| 11q13.3 | rs10896379,T | *IGHMBP2* | 0.19 | 0.84 | 0.055 | 0.0018 | 0.2909 (2;10,0-100) | 0.91 | 0.026 | $2.9x10^{-04}$ | 0.2702 (36;12,0-41) | 0.90 (0.86-0.94) | $4.1x10^{-06}$ | 0.1934 (38;16,0-44) |
| 5q31.3 | rs11167764,A | *NDFIP1* | 0.20 | 0.95 | 0.053 | 0.2894 | 0.9028 (2;0,0-100) | 0.89 | 0.025 | $4.9x10^{-06}$ | 0.0562 (36;29,0-53) | 0.90 (0.86-0.94) | $4.6x10^{-06}$ | 0.0731 (38;26,0-51) |

[a] Analyses were performed after excluding overlapping samples between this study and the GABRIEL (20).
[b] The association *P*-value corresponds to the fixed-effects meta-analysis conducted by the GABRIEL (20).
[c] Results for a test of heterogeneity (Breslow-Day test for the Australian GWAS; Cochran's Q test for all other analyses).
The table highlights: (1) in yellow, 12 SNPs located in regions reported to associate with asthma in previous GWAS (20, 41); (2) in blue, two SNPs with $P>0.05$ in the Australian GWAS and not reported to associate with asthma in previous GWAS; and (3) in green, five SNPs with $P\leq0.05$ in the Australian GWAS and not reported to associate with asthma in previous GWAS – only these five SNPs were selected for follow-up.
MAF=minor allele frequency. OR=odds ratio. *k*=number of analysis panels.

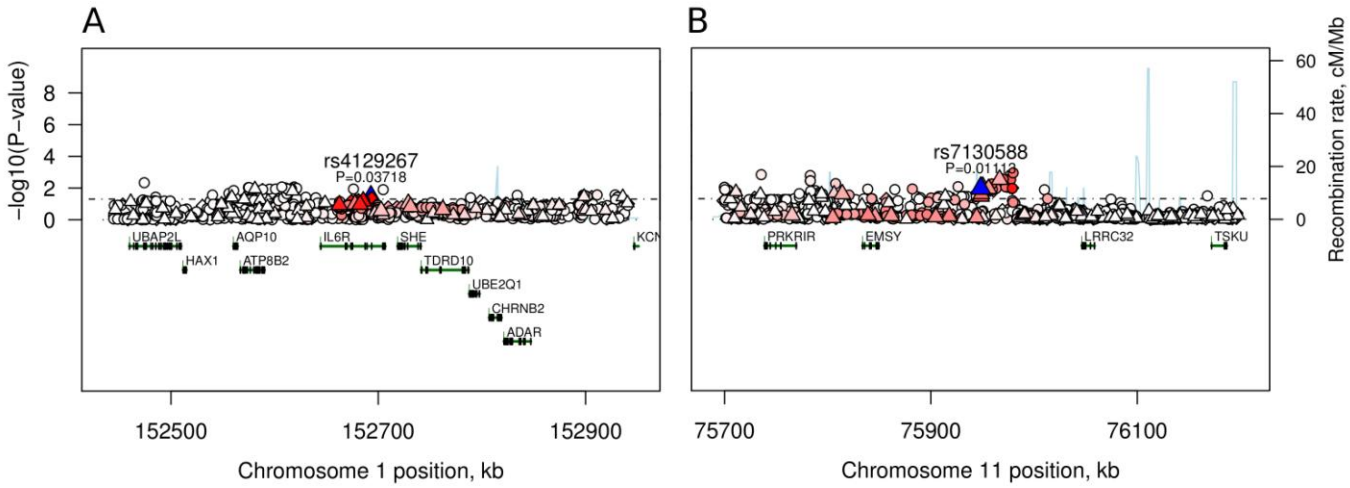


**Figure S7. Regional association plots for chromosomes 1q21.3 (A) and 11q13.5 (B) in the Australian GWAS.** Association results ($-\log_{10}P$-value, *y*-axis) are shown for genotyped (triangles) and imputed (diamonds, HapMap 3; circles, 1000 Genomes Project) SNPs located within 250 kb of rs4129267 and rs7130588 (blue triangles in each panel), and tested in 2,110 asthma cases and 3,857 controls from our study. The color of the remaining markers reflects the linkage disequilibrium ($r^2$) with rs4129267 or rs7130588 (increasing red hue associated with increasing $r^2$). Samples from the Busselton cohort that was included in the GABRIEL study (20) were excluded from this analysis. The recombination rate (second *y*-axis) is plotted in light blue and is based on the CEU HapMap population. Exons for each gene are represented by vertical bars.

**Table S7. Results for five loci with $P > 5x10^{-8}$ in the combined analysis of discovery and follow-up panels.**

| Panel | | N cases | N controls | Odds Ratio (95% CI) | Association *P*-value | Heterogeneity test[a] *P*-value (*k*; $I^2$, 95% CI) |
|---|---|---|---|---|---|---|
| **Locus=13q21, SNP:allele=rs3119939:C, nearest gene=*PCDH20*** | | | | | | |
| Discovery | Australian GWAS[b] | 2110 | 3857 | 0.91 (0.84-0.99) | 0.0254 | 0.6933 (2; 0, 0-100) |
| | GABRIEL | 10365 | 16110 | 0.92 (0.88-0.95) | $1.3x10^{-5}$ | 0.3354 (36; 8, 0-38) |
| | Combined | 12475 | 19967 | 0.92 (0.88-0.95) | $9.6x10^{-7}$ | 0.5472 (38; 0, 0-37) |
| Follow-up | APCAT | 1716 | 16888 | 0.97 (0.90-1.05) | 0.4911 | 0.0126 (3; 77, 26-93) |
| | Raine | 654 | 621 | 0.88 (0.75-1.02) | 0.0902 | - |
| | QIMR | 602 | 2206 | 1.08 (0.95-1.23) | 0.2368 | - |
| | NTR | 350 | 2321 | 1.00 (0.85-1.17) | 0.9928 | 0.3127 (2; 2, 0-100) |
| | Combined | 3322 | 22036 | 0.98 (0.93-1.04) | 0.5403 | 0.0279 (7; 58, 2-82) |
| All samples | | 15797 | 42003 | 0.93 (0.91-0.96) | $7.7x10^{-6}$ | 0.1533 (45; 18, 0-43) |
| **Locus=10q11, SNP:allele=rs7922491:A, nearest gene=*PRKG1*** | | | | | | |
| Discovery | Australian GWAS[b] | 2110 | 3857 | 1.15 (1.01-1.31) | 0.0350 | 0.0148 (2; 83, 0-100) |
| | GABRIEL | 10365 | 16110 | 1.15 (1.08-1.22) | $1.2x10^{-5}$ | 0.8763 (36; 0, 0-38) |
| | Combined | 12475 | 19967 | 1.15 (1.09-1.22) | $1.2x10^{-6}$ | 0.6991 (38; 0, 0-37) |
| Follow-up | APCAT | 1716 | 16888 | 1.10 (0.94-1.28) | 0.2316 | 0.1013 (3; 56, 0-88) |
| | Raine | 654 | 621 | 1.30 (1.03-1.64) | 0.0280 | - |
| | QIMR | 602 | 2206 | 1.03 (0.85-1.25) | 0.7979 | - |
| | NTR | 350 | 2321 | 0.92 (0.69-1.23) | 0.5805 | 0.5651 (2; 0, 0-100) |
| | Combined | 3322 | 22036 | 1.09 (0.99-1.20) | 0.0974 | 0.1875 (7; 32, 0-71) |
| All samples | | 15797 | 42003 | 1.13 (1.08-1.19) | $4.7x10^{-7}$ | 0.5674 (45; 0, 0-35) |
| **Locus=11q13.3, SNP:allele=rs10896379:T, nearest gene=*IGHMBP2*** | | | | | | |
| Discovery | Australian GWAS[b] | 2110 | 3857 | 0.84 (0.76-0.94) | 0.0018 | 0.2909 (2; 10, 0-100) |
| | GABRIEL | 10365 | 16110 | 0.91 (0.87-0.96) | $2.9x10^{-4}$ | 0.2702 (36; 12, 0-41) |
| | Combined | 12475 | 19967 | 0.90 (0.86-0.94) | $4.1x10^{-6}$ | 0.1934 (38; 16, 0-44) |
| Follow-up | APCAT | 1716 | 16888 | 0.96 (0.87-1.07) | 0.4845 | 0.4293 (3; 0, 0-90) |
| | Raine | 654 | 621 | 0.98 (0.81-1.19) | 0.8733 | - |
| | QIMR | 602 | 2206 | 1.01 (0.86-1.18) | 0.9191 | - |
| | NTR | 350 | 2321 | 1.00 (0.82-1.21) | 0.9624 | 0.3250 (2; 0, 0-100) |
| | Combined | 3322 | 22036 | 0.98 (0.91-1.05) | 0.5902 | 0.8177 (7; 0, 0-71) |
| All samples | | 15797 | 42003 | 0.92 (0.89-0.96) | $3.0x10^{-5}$ | 0.2177 (45; 14, 0-41) |
| **Locus=16q24, SNP:allele=rs7196274:A, nearest gene=*KCNG4*** | | | | | | |
| Discovery | Australian GWAS | 2669 | 4528 | 0.81 (0.74-0.88) | $1.3x10^{-6}$ | 0.0853 (3; 59, 0-88) |
| Follow-up | APCAT | 1716 | 16888 | 1.03 (0.94-1.14) | 0.5225 | 0.2668 (3; 24, 0-92) |
| | Raine | 654 | 621 | 0.89 (0.75-1.05) | 0.1705 | - |
| | QIMR | 602 | 2206 | 0.84 (0.73-0.98) | 0.0233 | - |
| | NTR | 350 | 2321 | 1.01 (0.85-1.21) | 0.8681 | 0.6664 (2; 0, 0-100) |
| | Combined | 3322 | 22036 | 0.96 (0.90-1.03) | 0.2863 | 0.1584 (7; 35, 0-73) |
| All samples | | 5991 | 26564 | 0.90 (0.86-0.95) | $1.2x10^{-4}$ | 0.0040 (10; 63, 26-81) |
| **Locus=8q22, SNP:allele=rs11776675:C, nearest gene=*CDH17*** | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Discovery | Australian GWAS | 2669 | 4528 | 1.21 (1.12-1.30) | $2.7x10^{-6}$ | 0.3199 (3; 12, 0-91) |
| Follow-up | APCAT | 1716 | 16888 | 1.02 (0.94-1.10) | 0.6490 | 0.2640 (3; 25, 0-92) |
| | Raine | 654 | 621 | 1.01 (0.86-1.20) | 0.8879 | - |
| | QIMR | 602 | 2206 | 1.06 (0.93-1.20) | 0.4139 | - |
| | NTR | 350 | 2321 | 0.93 (0.78-1.10) | 0.3791 | 0.2006 (2; 39, 0-100) |
| | Combined | 3322 | 22036 | 1.01 (0.96-1.07) | 0.6435 | 0.4515 (7; 0, 0-71) |
| All samples | | 5991 | 26564 | 1.08 (1.03-1.13) | 0.0016 | 0.0160 (10; 56, 9-78) |

[a] Results for a test of heterogeneity (Breslow-Day test for the Australian GWAS and NTR analyses; Cochran's Q test for all other analyses) are provided for analyses that incorporated multiple cohorts/panels.
[b] Analyses were performed after excluding overlapping samples between this study and the GABRIEL (20).
*k*=number of analysis panels.

**Table S8. Association between *IL6R* and 11q13.5 and nine asthma subphenotypes in asthmatics ($N_{max}$=2,669).**

| | N | Effect[a] | SE | *P*-value |
|---|---|---|---|---|
| ***IL6R* (rs4129267:T)** | | | | |
| Childhood asthmatics *vs* later onset asthmatics | 1262 *vs* 682 | 1.065 | 0.069 | 0.3571 |
| Atopic[b] asthmatics *vs* non-atopic asthmatics | 1515 *vs* 440 | 1.067 | 0.078 | 0.4050 |
| Asthmatics with eczema[c] *vs* asthmatics with no eczema | 335 *vs* 1224 | 0.973 | 0.088 | 0.7563 |
| Clinical- *vs* questionnaire-based diagnosed asthmatics | 759 *vs* 1910 | 1.005 | 0.061 | 0.9333 |
| Forced Expiratory Volume in 1s ($FEV_1$) | 1478 | -0.001 | 0.028 | 0.9779 |
| Forced Vital Capacity (FVC) | 1477 | -0.009 | 0.034 | 0.7965 |
| $FEV_1$/FVC | 1476 | 0.003 | 0.004 | 0.5068 |
| Peripheral blood eosinophil counts | 400 | 0.082 | 0.055 | 0.1371 |
| Total serum IgE levels | 902 | 0.099 | 0.068 | 0.1476 |
| **11q13.5 (rs7130588:G)** | | | | |
| Childhood asthmatics *vs* later onset asthmatics | 1262 *vs* 682 | 1.024 | 0.070 | 0.7385 |
| Atopic[b] asthmatics *vs* non-atopic asthmatics | 1515 *vs* 440 | 1.322 | 0.082 | 0.0007 |
| Asthmatics with eczema[c] *vs* asthmatics with no eczema | 335 *vs* 1224 | 1.170 | 0.089 | 0.0782 |
| Clinical- *vs* questionnaire-based diagnosed asthmatics | 759 *vs* 1910 | 0.986 | 0.063 | 0.8258 |
| Forced Expiratory Volume in 1s ($FEV_1$) | 1477 | 0.047 | 0.029 | 0.1007 |
| Forced Vital Capacity (FVC) | 1476 | 0.069 | 0.034 | 0.0460 |
| $FEV_1$/FVC | 1475 | -0.001 | 0.004 | 0.8040 |
| Peripheral blood eosinophil counts | 399 | 0.009 | 0.056 | 0.8647 |
| Total serum IgE levels | 901 | 0.147 | 0.071 | 0.0377 |

[a] Effect corresponds to the odds ratio for the minor allele for the four binary traits or the slope from a linear regression (beta) for the five continuous traits, the latter including age, sex, height, ever and current smoker as covariates. Eosinophil and total IgE levels were log transformed prior to the analysis.
[b] Atopy was defined by a positive skin prick test to at least one common allergen.
[c] Eczema was defined by a self-reported physician diagnosis of the disease.

**Table S9. Frequency of the rs7130588:G predisposing allele as a function of asthma, eczema and atopy status in the Australian GWAS.**

| | | Case definition | | | Control definition | | | N cases | N controls | MAF cases | MAF controls | Association test[a] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Asthma | Eczema | Atopy | Asthma | Eczema | Atopy | | | | | OR (95% CI) | *P*-value |
| Atopic asthmatics *vs* | Atopic non-asthmatics | Yes | - | Yes | No | - | Yes | 1570 | 250 | 0.385 | 0.354 | 1.21 (0.97-1.52) | 0.0916 |
| | Non-atopic non-asthmatics | Yes | - | Yes | No | - | No | 1570 | 375 | 0.385 | 0.328 | 1.29 (1.06-1.58) | 0.0124 |
| | Non-asthmatics | Yes | - | Yes | No | - | Yes or No | 1570 | 625 | 0.385 | 0.338 | 1.26 (1.06-1.49) | 0.0091 |
| Non-atopic asthmatics *vs* | Atopic non-asthmatics | Yes | - | No | No | - | Yes | 452 | 250 | 0.319 | 0.354 | 0.86 (0.66-1.13) | 0.2914 |
| | Non-atopic non-asthmatics | Yes | - | No | No | - | No | 452 | 375 | 0.319 | 0.328 | 0.87 (0.67-1.12) | 0.2667 |
| | Non-asthmatics | Yes | - | No | No | - | Yes or No | 452 | 625 | 0.319 | 0.338 | 0.88 (0.70-1.10) | 0.2532 |
| Eczema *vs* | Atopics, no eczema | - | Yes | - | - | No | Yes | 368 | 914 | 0.416 | 0.380 | 1.14 (0.95-1.37) | 0.1539 |
| | Non-atopics, no eczema | - | Yes | - | - | No | No | 368 | 581 | 0.416 | 0.325 | 1.41 (1.11-1.78) | 0.0049 |
| | No eczema | - | Yes | - | - | No | Yes or No | 368 | 1495 | 0.416 | 0.358 | 1.21 (1.01-1.43) | 0.0356 |
| Atopics *vs* Non-atopics | | - | - | Yes | - | - | No | 1871 | 848 | 0.380 | 0.325 | 1.27 (1.12-1.44) | 0.0002 |
| Atopics, no eczema *vs* Non-atopics, no eczema | | - | No | Yes | - | No | No | 914 | 581 | 0.380 | 0.325 | 1.24 (1.05-1.46) | 0.0111 |
| Atopics, no eczema, no asthma *vs* Non-atopics, no eczema, no asthma | | No | No | Yes | No | No | No | 239 | 372 | 0.360 | 0.329 | 1.15 (0.90-1.46) | 0.2675 |
| Asthmatics, no eczema *vs* Non-asthmatics, no eczema | | Yes | No | - | No | No | - | 1224 | 4451 | 0.375 | 0.355 | 1.10 (0.99-1.21) | 0.0834 |

[a] Corresponds to a Cochran-Mantel-Haenszel test of association with three strata (QIMR_610K, QIMR_370K and BUSSELTON datasets).
Asthma=physician-diagnosed asthma. Eczema=physician-diagnosed eczema. Atopy=positive skin prick test to at least one common allergen. MAF=minor allele frequency. OR=odds ratio. A "-" indicates "Yes, No or Unknown".

**Table S10. Results for specific SNPs reported to associate with asthma risk in previous GWAS.**

| Locus[a] | Literature SNP, allele | Original report | | Australian GWAS[b] | | | | | | Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Total sample | | Childhood onset | | Later onset | | |
| | | OR | *P*-value | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | |
| *GSDMB* | rs2305480, A | 0.85 | $9.6x10^{-8}$ | 0.77 (0.70-0.83) | $5.7x10^{-10}$ | 0.71 (0.64-0.78) | $6.9x10^{-12}$ | 0.94 (0.82-1.08) | 0.3658 | (20) |
| *GSDMA* | rs3894194,A | 1.17 | $4.6x10^{-9}$ | 1.26 (1.16-1.37) | $7.1x10^{-8}$ | 1.37 (1.24-1.51) | $3.8x10^{-10}$ | 0.99 (0.86-1.13) | 0.8586 | (20) |
| *IL18R1* | rs3771166,A | 0.87 | $3.4x10^{-9}$ | 0.86 (0.79-0.94) | 0.0005 | 0.81 (0.73-0.90) | $4.8x10^{-5}$ | 0.92 (0.80-1.06) | 0.2683 | (20) |
| *IL33* | rs1342326,C | 1.20 | $9.2x10^{-10}$ | 1.20 (1.08-1.34) | 0.0010 | 1.25 (1.10-1.42) | 0.0005 | 1.16 (0.97-1.38) | 0.1023 | (20) |
| *RORA* | rs11071559,T | 0.85 | $1.1x10^{-7}$ | 0.85 (0.75-0.96) | 0.0102 | 0.88 (0.76-1.02) | 0.1036 | 0.71 (0.57-0.88) | 0.0022 | (20) |
| *IL2RB* | rs2284033,A | 0.89 | $1.2x10^{-8}$ | 0.90 (0.83-0.98) | 0.0132 | 0.92 (0.84-1.02) | 0.1078 | 0.84 (0.74-0.97) | 0.0169 | (20) |
| *HLA-DQ* | rs9273349,T | 0.85 | $7.0x10^{-14}$ | 0.93 (0.85-1.01) | 0.0824 | 0.88 (0.79-0.97) | 0.0103 | 0.99 (0.86-1.14) | 0.8818 | (20) |
| *SMAD3* | rs744910,G | 1.12 | $3.9x10^{-9}$ | 1.06 (0.98-1.16) | 0.1460 | 1.04 (0.95-1.15) | 0.3718 | 1.12 (0.98-1.29) | 0.0904 | (20) |
| *PDE4D* | rs1588265,G | 0.85 | $2.5x10^{-8}$ | 1.06 (0.97-1.16) | 0.2151 | 1.03 (0.93-1.13) | 0.5931 | 1.05 (0.92-1.19) | 0.4871 | (42) |
| *SLC22A5* | rs2073643,T | 0.90 | $2.2x10^{-7}$ | 0.95 (0.88-1.04) | 0.2739 | 0.94 (0.85-1.04) | 0.2408 | 0.98 (0.86-1.12) | 0.7894 | (20) |
| *IL13* | rs1295686,T | 0.87 | $1.4x10^{-7}$ | 1.03 (0.93-1.15) | 0.5548 | 1.05 (0.93-1.18) | 0.4616 | 1.07 (0.90-1.26) | 0.4379 | (20) |
| *DENND1B* | rs1775456,G | 0.75 | $3.0x10^{-7}$ | 1.00 (0.90-1.10) | 0.9285 | 1.00 (0.90-1.12) | 0.9357 | 0.99 (0.86-1.15) | 0.9183 | (43) |

[a] For completeness, we also included the three loci reported by the GABRIEL consortium (20) that did not reach genome-wide significance in that study, namely *RORA*, *SLC22A5* and *IL13*.

[b] SNPs were directly genotyped in all cohorts with the exception of rs3771166 and rs2073643 (imputed in the QIMR_370K dataset only, information = 0.99 and 0.92, respectively), rs2284033 and rs9273349 (imputed in all cohorts, information = 0.84 and 0.98, respectively). Analyses were performed after excluding overlapping samples between this study and the GABRIEL (20). OR=odds ratio. Ref=reference.

**Table S11. Results for the most significant SNP in the Australian GWAS[a] for 12 loci reported to associate with asthma in previous GWAS.**

| Locus[b] | Best SNP, allele | N SNPs tested[c] | LD with literature SNP[d], $r^2$ | Allele frequency | OR (95% CI) | *P*-value | |
|---|---|---|---|---|---|---|---|
| | | | | | | Uncorrected | Corrected[e] |
| *GSDMB* | rs4795405,T | 189 | 0.83 | 0.46 | 0.76 (0.70-0.83) | $1.6x10^{-10}$ | $<10^{-4}$ |
| *GSDMA* | rs4795405,T | 287 | 0.51 | 0.46 | 0.76 (0.70-0.83) | $1.6x10^{-10}$ | $<10^{-4}$ |
| *IL18R1* | 2-102348465,T | 625 | 0.30 | 0.15 | 0.78 (0.69-0.89) | 0.0001 | 0.0163 |
| *IL33* | 9-6165855,T | 289 | 0.97 | 0.16 | 1.20 (1.08-1.34) | 0.0008 | 0.0545 |
| *RORA* | 15-58604480,C | 338 | 0.00 | 0.10 | 1.22 (1.06-1.40) | 0.0062 | 0.3522 |
| *IL2RB* | rs228953,A | 326 | 0.96 | 0.43 | 0.90 (0.82-0.97) | 0.0104 | 0.5619 |
| *HLA-DQ* | rs9273148, A | 4,436 | 0.13 | 0.17 | 1.20 (1.08-1.34) | 0.0009 | 0.1888 |
| *SMAD3* | rs266335,G | 523 | 0.00 | 0.32 | 1.15 (1.06-1.26) | 0.0016 | 0.1988 |
| *PDE4D* | 5-58522048,T | 3,679 | 0.00 | 0.08 | 1.32 (1.13-1.54) | 0.0005 | 0.3047 |
| *SLC22A5* | 5-131756744,T | 312 | 0.00 | 0.01 | 1.31 (0.96-1.77) | 0.0857 | 0.9471 |
| *IL13* | rs1881457,C | 156 | 0.35 | 0.18 | 1.14 (1.03-1.27) | 0.0137 | 0.3846 |
| *DENND1B* | rs12118513,A | 420 | 0.06 | 0.22 | 0.85 (0.77-0.95) | 0.0029 | 0.2328 |

[a] Analyses were performed after excluding overlapping samples between this study and the GABRIEL (20).
[b] For completeness, we also included the three loci reported by the GABRIEL (20) that did not reach genome-wide significance in that study, namely *RORA*, *SLC22A5* and *IL13*.
[c] Number of SNPs located in or within 50 kb of each gene.
[d] cf. **Table S10**.
[e] Significance of the best SNP after accounting for the number of, and LD between, SNPs tested in the respective gene, estimated from 10,000 permutations.
LD=linkage disequilibrium.

**Table S12. Fifty-four inflammatory- or immune-related traits with one or more locus reported in the catalog of GWAS with $P \leq 5x10^{-8}$.**

| | |
|---|---|
| AIDS | Multiple sclerosis (severity) |
| AIDS progression | Neutrophil count |
| Alopecia areata | Plasma eosinophil count |
| Ankylosing spondylitis | Plasma E-selectin levels |
| Atopic dermatitis | Plasma level of vitamin B12 |
| Atopy | Plasma levels of Protein C |
| CD4:CD8 lymphocyte ratio | Primary biliary cirrhosis |
| Celiac disease | Primary sclerosing cholangitis |
| Chronic Hepatitis C infection | Protein quantitative trait loci |
| Chronic obstructive pulmonary disease | Psoriasis |
| C-reactive protein | Psoriatic arthritis |
| Crohn's disease | Pulmonary function |
| Eosinophilic esophagitis (pediatric) | Pulmonary function measures |
| Factor VII | Pulmonary function traits (other) |
| Fibrinogen | Rheumatoid arthritis |
| Hepatitis B | Serologic markers in systemic lupus erythematosus |
| HIV-1 control | Serum IgE levels |
| HIV-1 susceptibility | Serum matrix metalloproteinase |
| HIV1 viral setpoint | Serum soluble E-selectin |
| HIV (mother-to-child transmission) | Soluble ICAM-1 |
| Immunoglobulin A | Soluble levels of adhesion molecules |
| Inflammatory bowel disease | Systemic lupus erythematosus |
| Inflammatory bowel disease (early onset) | Type 1 diabetes |
| Interleukin-18 levels | Ulcerative colitis |
| Lupus | Vitamin D insufficiency |
| Mean forced vital capacity from 2 exams | Vitamin D levels |
| Multiple sclerosis | YKL-40 levels |

**Table S13. Results for 16 SNPs previously reported to associate with inflammatory- or immune-related traits ($P \leq 5x10^{-8}$) and that associated with asthma risk ($P \leq 0.01$) in the meta-analysis of our study and the GABRIEL (20).**

| Chr | Position[a], bp | Literature SNP | SNP proxy ($r^2 > 0.8$) | Nearest gene | Allele | OR | Asthma *P*-value | Literature disease | Literature *P*-value | Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 141459249 | rs11167764 | - | *NDFIP1* | C | 1.11 | $4.6x10^{-6}$ | CrD | $2x10^{-9}$ | (44) |
| 6 | 31463297 | rs3134792 | rs2596560 | *HLA-B* | T | 0.92 | $6.5x10^{-5}$ | PSOR | $10^{-9}$ | (45) |
| 3 | 189571948 | rs1464510 | rs13076312 | *LPP* | T | 0.93 | 0.00016 | CeD | $3x10^{-40}$ | (46) |
| 6 | 91029880 | rs1847472 | - | *BACH2* | C | 1.07 | 0.00023 | CrD | $5x10^{-9}$ | (44) |
| 1 | 25176163 | rs10903122 | - | *RUNX3* | G | 1.07 | 0.00035 | CeD | $2x10^{-10}$ | (46) |
| 5 | 110433574 | rs3806932 | - | *TSLP* | G | 0.94 | 0.00036 | EE | $3x10^{-9}$ | (47) |
| 6 | 128320491 | rs802734 | - | *PTPRK* | G | 1.07 | 0.00080 | CeD | $3x10^{-14}$ | (46) |
| 21 | 15735083 | rs1736135 | rs1736148 | *USP25* | T | 1.06 | 0.00096 | CrD | $7x10^{-9}$ | (48) |
| 3 | 32990473 | rs13314993 | - | *CCR4* | T | 0.95 | 0.00283 | CeD | $3x10^{-9}$ | (46) |
| 4 | 123351942 | rs4505848 | - | *KIAA1109* | G | 1.06 | 0.00337 | T1D | $5x10^{-13}$ | (49) |
| 9 | 122692719 | rs881375 | - | *TRAF1* | T | 1.06 | 0.00383 | RA | $4x10^{-8}$ | (50) |
| 1 | 20012623 | rs1317209 | - | *RNF186* | G | 1.07 | 0.00417 | UC | $2x10^{-10}$ | (51) |
| 1 | 157966663 | rs3093059 | rs11265260 | *CRP* | G | 1.11 | 0.00505 | CRP | $4x10^{-21}$ | (52) |
| 14 | 87547635 | rs8005161 | rs3742704 | *GPR65* | C | 1.09 | 0.00552 | CrD | $4x10^{-18}$ | (44) |
| 2 | 127886369 | rs1158867 | rs6753288 | *PROC* | G | 1.05 | 0.00793 | PROC | $4x10^{-36}$ | (53) |
| 10 | 64140681 | rs224136 | - | *ZNF365* | T | 1.07 | 0.00824 | CrD | $10^{-10}$ | (54) |

[a] Base pair positions (bp) correspond to build 36.

Chr=chromosome. bp=base pair. OR=odds ratio. CrD=Crohn's disease. CRP=c-reactive protein. PSOR=psoriasis. CeD=celiac disease. RA=reumathoid arthritis. EE=eosinophilic esophagitis. T1D=type-1 diabetes. UC=ulcerative colitis. PROC=protein C plasma levels. Ref=reference.

## Acknowledgments

QIMR. We thank the twins and their families for their participation; Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney (sample collection); Lisa Bowdler, Steven Crooks (DNA processing); David Smyth, Harry Beeby, Daniel Park (IT support).

APCAT. *NFBC66:* The DNA extractions, sample QC, biobank up-keeping and aliquotting was performed in the National Institute for Health and Welfare, Biomedicum Helsinki, Finland. We thank Professor Paula Rantakallio (launch of NFBC1966 and 1986), Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). *Framingham/Children's*: The research on the Framingham Heart Study cohort was conducted using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine based on analyses by Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project.

Raine. We are grateful to the Raine Study participants and their families; the Raine Study research staff for cohort coordination and data collection; and the assistance of the Western Australian DNA Bank (NHMRC National Enabling Facility).

Busselton. The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia and the numerous Busselton community volunteers who assisted with data collection and the study participants from the Shire of Busselton.

## Funding

QIMR. Funding was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498, 613627), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254) and the U.S. National Institutes of Health (NIH grants AA07728, AA07535, AA10248, AA11998, AA13320, AA13321, AA13326, AA14041, AA17688, DA12854, MH66206). A portion of the genotyping on which this study was based (Illumina 370k scans on 4300 individuals) was carried out at the Center for Inherited Disease Research, Baltimore (CIDR) through an access award to our late colleague Dr. Richard Todd (Psychiatry, Washington University School of Medicine, St Louis). DRN, GWM and PMV are supported by the NHMRC Fellowship Scheme.

TAHS. The Tasmanian Longitudinal Health Study has been supported by the NHMRC, Asthma Foundations in Tasmania, Queensland and Victoria, The Clifford Craig Trust in Northern Tasmania, Lew Carty Foundation, Royal Hobart Research Foundation and the University of Melbourne. MCM and SCD are supported by the NHMRC.

CAPS. National Health and Medical Research Council of Australia, Cooperative Research Centre for Asthma, New South Wales Department of Health, Children's Hospital Westmead, University of Sydney, Faculty of Medicine. Contributions of goods and services were made by Allergopharma Joachim Ganzer KG Germany, John Sands Australia, Hasbro, Toll refrigerated, AstraZeneca Australia, and Nu-Mega Ingredients Pty Ltd. Goods were provided at reduced cost by Auspharm, Allersearch and Goodman Fielder Foods. University of Sydney, Faculty of Medicine, Strategic Research Grant 2008.

APCAT. *NFBC66:* 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 QLG1-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). The DNA extractions, sample QC, biobank up-keeping and aliquotting was supported financially by the Academy of Finland and Biocentrum Helsinki. A.R. was supported by the Department of Health UK and 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. *FINRISK:* Academy of Finland grants (including 129494 and 139635), Foundation of the Finnish Anti-Tuberculosis and Allergy Associations. *HBCS*: financial support was received from the Academy of Finland (including 129287 and134839), Samfundet Folkhälsan, Finnish Diabetes Research Foundation, Finska Läkaresällskapet, Finnish Foundation for Cardiovascular Research; Yrjö Jahnsson Foundation, Foundation Liv och Hälsa. *YSF*: financial support received by the Academy of Finland (126925, 121584, 124282, 129398 and 117941), the Social Insurance Institution of Finland, Kuopio, Tampere and Turku University Hospital Medical Funds (9L054 and 9M048), Juho Vainio Foundation, Paavo Nurmi Foundation, Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation, Emil Aaltonen Foundation, Tampere Tuberculosis Foundation. *SANGER*: SYNSYS - Synaptic Systems: dissecting brain function in health and disease, Grant agreement: HEALTH-2009-2.1.2-1, The Center of Excellence in Complex Disease Genetics of the Finnish Academy, BioSHaRE - Biobank standardisation and harmonisation for research excellence in the European Union, Grant agreement: HEALTH-2010.1.1-1. *Framingham/Children's*: Financial support was provided by the American Asthma Foundation (grant 07-0282). This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

Raine. The University of Western Australia (UWA), Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, The Telethon Institute for Child Health Research and Women and Infants Research Foundation for core management of the Raine Study; NHMRC grants 403981 and 003209; the Canadian Institutes of Health Research (MOP-82893).

University of Queensland Diamantina Institute. MAB is supported by an NHMRC Principal Research Fellowship (455836). This work was also supported by the Australian Cancer Research Foundation and Rebecca Cooper Foundation.

Busselton. The Busselton Health Study is supported by The Great Wine Estates of the Margaret River region of Western Australia.

NTR1 and NTR2. Funding was obtained from the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW): Twin family database for behavior genomics studies (480-04-004); Genotype/phenotype database for behavior genetic and genetic epidemiological studies (40-0056-98-9032); Spinozapremie (SPI 56-464-14192); CMSB: Center for Medical Systems Biology (NWO Genomics); NBIC/BioAssist/RK/2008.024; BBMRI –NL: Biobanking and Biomolecular Resources Research Infrastructure; the VU University: Institute for Health and Care Research (EMGO+ ) and Neuroscience Campus Amsterdam (NCA); ENGAGE (HEALTH-F4-2007-201413); the European Science Council (ERC) Genetics of Mental Illness (230374); Collaborative study of the genetics of DZ twinning (NIH R01D0042157-01A); the Genetic Association Information Network, a public–private partnership between the NIH and Pfizer Inc., Affymetrix Inc. and Abbott Laboratories; the Netherlands Scientific Organization (904-61-090, 904-61-193, 480-04-004, 400-05-717).

## References


1. Ferreira MA, O'Gorman L, Le Souef P, Burton PR, Toelle BG, Robertson CF, et al. Variance components analyses of multiple asthma traits in a large sample of Australian families ascertained through a twin proband. Allergy. 2006 Feb;61(2):245-53.
2. Pekkanen J, Sunyer J, Anto JM, Burney P. Operational definitions of asthma in studies on its aetiology. Eur Respir J. 2005 Jul;26(1):28-35.
3. Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, O'Gorman L, et al. A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. Am J Hum Genet. 1999 Aug;65(2):483-92.
4. Gibson HB, Silverstone H, Gandevia B, Hall GJ. Respiratory disorders in seven-year-old children in Tasmania. Aims, methods and administration of the survey. Med J Aust. 1969 Jul 26;2(4):201-5.
5. Giles GG, Lickiss N, Gibson HB, Shaw K. Respiratory symptoms in Tasmanian adolescents: a follow up of the 1961 birth cohort. Aust N Z J Med. 1984 Oct;14(5):631-7.
6. Jenkins MA, Hopper JL, Flander LB, Carlin JB, Giles GG. The associations between childhood asthma and atopy, and parental asthma, hay fever and smoking. Paediatr Perinat Epidemiol. 1993 Jan;7(1):67-76.
7. Jenkins MA, Hopper JL, Bowes G, Carlin JB, Flander LB, Giles GG. Factors in childhood as predictors of asthma in adult life. BMJ. 1994 Jul 9;309(6947):90-3.
8. Wharton C, Dharmage S, Jenkins M, Dite G, Hopper J, Giles G, et al. Tracing 8,600 participants 36 years after recruitment at age seven for the Tasmanian Asthma Study. Aust N Z J Public Health. 2006 Apr;30(2):105-10.
9. Mihrshahi S, Peat JK, Webb K, Tovey ER, Marks GB, Mellis CM, et al. The childhood asthma prevention study (CAPS): design and research protocol of a randomized trial for the primary prevention of asthma. Control Clin Trials. 2001 Jun;22(3):333-54.
10. Marks GB, Mihrshahi S, Kemp AS, Tovey ER, Webb K, Almqvist C, et al. Prevention of asthma during the first 5 years of life: a randomized controlled trial. J Allergy Clin Immunol. 2006 Jul;118(1):53-61.
11. Mihrshahi S, Peat JK, Marks GB, Mellis CM, Tovey ER, Webb K, et al. Eighteen-month outcomes of house dust mite avoidance and dietary fatty acid modification in the Childhood Asthma Prevention Study (CAPS). J Allergy Clin Immunol. 2003 Jan;111(1):162-8.
12. Peat JK, Mihrshahi S, Kemp AS, Marks GB, Tovey ER, Webb K, et al. Three-year outcomes of dietary fatty acid modification and house dust mite reduction in the Childhood Asthma Prevention Study. J Allergy Clin Immunol. 2004 Oct;114(4):807-13.
13. Toelle BG, Ng KK, Crisafulli D, Belousova EG, Almqvist C, Webb K, et al. Eight-year outcomes of the Childhood Asthma Prevention Study. J Allergy Clin Immunol. 2010 Aug;126(2):388-9, 9 e1-3.
14. Williams H, McNicol KN. Prevalence, natural history, and relationship of wheezy bronchitis and asthma in children. An epidemiological study. Br Med J. 1969 Nov 8;4(5679):321-5.
15. McNicol KN, Macnicol KN, Williams HB. Spectrum of asthma in children. I. Clinical and physiological components. Br Med J. 1973 Oct 6;4(5883):7-11.
16. Oswald H, Phelan PD, Lanigan A, Hibbert M, Bowes G, Olinsky A. Outcome of childhood asthma in mid-adult life. BMJ. 1994 Jul 9;309(6947):95-6.
17. Phelan PD, Robertson CF, Olinsky A. The Melbourne Asthma Study: 1964-1999. J Allergy Clin Immunol. 2002 Feb;109(2):189-94.
18. Wolfe R, Carlin JB, Oswald H, Olinsky A, Phelan PD, Robertson CF. Association between allergy and asthma from childhood to middle adulthood in an Australian cohort study. Am J Respir Crit Care Med. 2000 Dec;162(6):2177-81.
19. James AL, Knuiman MW, Divitini ML, Hui J, Hunter M, Palmer LJ, et al. Changes in the prevalence of asthma in adults since 1966: the Busselton health study. Eur Respir J. 2010 Feb;35(2):273-8.
20. 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 Sep 23;363(13):1211-21.
21. Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G, et al. Common variants in the trichohyalin gene are associated with straight hair in Europeans. Am J Hum Genet. 2009 Nov;85(5):750-5.
22. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. Am Rev Respir Dis. 1990 Dec;142(6 Pt 1):1351-8.
23. Heath AC, Martin NG. Genetic influences on alcohol consumption patterns and problem drinking: results from the Australian NH&MRC twin panel follow-up survey. Ann N Y Acad Sci. 1994 Feb 28;708:72-85.
24. Lake RI, Eaves LJ, Maes HH, Heath AC, Martin NG. Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. Behav Genet. 2000 May;30(3):223-33.
25. Kirk KM, Martin NG. The Short Interpersonal Reactions Inventory, Self-regulation and differentiation scales in an older Australian twin sample. Personality and Individual Differences. 1998;25(3):591-604.
26. Ferreira MA, McRae AF, Medland SE, Nyholt DR, Gordon SD, Wright MJ, et al. Association between ORMDL3, IL1RL1 and a deletion on chromosome 17q21 with asthma risk in Australia. Eur J Hum Genet. 2011

Apr; 19(4):458-64.
27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 Sep 1;26(17):2190-1.
28. Newnham JP, Evans SF, Michael CA, Stanley FJ, Landau LI. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. Lancet. 1993 Oct 9;342(8876):887-91.
29. Williams LA, Evans SF, Newnham JP. Prospective cohort study of factors influencing the relative weights of the placenta and the newborn infant. BMJ. 1997 Jun 28;314(7098):1864-8.
30. Evans S, Newnham J, MacDonald W, Hall C. Characterisation of the possible effect on birthweight following frequent prenatal ultrasound examinations. Early Hum Dev. 1996 Jul 19;45(3):203-14.
31. Painter JN, Anderson CA, Nyholt DR, Macgregor S, Lin J, Lee SH, et al. Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. Nat Genet. 2010 Jan;43(1):51-4.
32. Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry. 2009 Apr;14(4):359-75.
33. Willemsen G, de Geus EJ, Bartels M, van Beijsterveldt CE, Brooks AI, Estourgie-van Burk GF, et al. The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Res Hum Genet. 2010 Jun;13(3):231-45.
34. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009 Jun;5(6):e1000529.
35. Eriksson JG, Forsen T, Tuomilehto J, Osmond C, Barker DJ. Early growth and coronary heart disease in later life: longitudinal study. BMJ. 2001 Apr 21;322(7292):949-53.
36. Aromaa A, Koskinen S. Health and functional capacity in Finland: baseline results of the Health 2000 health examination survey. Helsinki: National Public Health Institute; 2004.
37. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol. 2010 Apr;39(2):504-18.
38. Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. Acta Paediatr Scand. 1969;193:Suppl 193:1+.
39. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, et al. Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol. 2008 Dec;37(6):1220-6.
40. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol. 2007 Jun 1;165(11):1328-35.
41. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet. 2009 Mar;41(3):342-7.
42. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. Am J Hum Genet. 2009 May;84(5):581-93.
43. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. N Engl J Med. 2010 Jan 7;362(1):36-44.
44. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010 Dec;42(12):1118-25.
45. Capon F, Bijlmakers MJ, Wolf N, Quaranta M, Huffmeier U, Allen M, et al. Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. Hum Mol Genet. 2008 Jul 1;17(13):1938-45.
46. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet. 2010 Apr;42(4):295-302.
47. Rothenberg ME, Spergel JM, Sherrill JD, Annaiah K, Martin LJ, Cianferoni A, et al. Common variants at 5q22 associate with pediatric eosinophilic esophagitis. Nat Genet. Apr;42(4):289-91.
48. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet. 2008 Aug;40(8):955-62.
49. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009 Jun;41(6):703-7.
50. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, et al. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat Genet. 2009 Jul;41(7):820-3.
51. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet. Apr;42(4):332-7.
52. Okada Y, Takahashi A, Ohmiya H, Kumasaka N, Kamatani Y, Hosono N, et al. Genome-wide association study for C-reactive protein levels identified pleiotropic associations in the IL6 locus. Hum Mol Genet. Jan 10.
53. Tang W, Basu S, Kong X, Pankow JS, Aleksic N, Tan A, et al. Genome-wide association study identifies novel

loci for plasma levels of protein C: the ARIC study. Blood. Dec 2;116(23):5032-6.
54. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet. 2007 May;39(5):596-604.