Shared genetic variants suggest common pathways in allergy and autoimmune diseases

Eskil Kreiner, MD, PhD[1,†]; Johannes Waage, PhD[1,†]; Marie Standl, PhD[2]; Susanne Brix, PhD[3]; Tune H Pers, PhD[3-5]; Alexessander Couto Alves, PhD[6]; Nicole M Warrington, PhD[7,8]; Carla MT Tiesler, MSc[2,9]; Elaine Fuertes, PhD[2]; Lude Franke, PhD[10]; AAGC, Joel N Hirschhorn, MD, PhD[4,5,11]; Alan James, MD[12-14]; Angela Simpson, MD, PhD[15]; Joyce Y Tung, PhD[16]; Gerard H Koppelman, MD, PhD[17]; Dirkje S Postma, MD, PhD[18]; Craig E Pennell, MD, PhD[8]; Marjo-Riitta Jarvelin, MD[19-23]; Adnan Custovic, MD, PhD[24]; Nicholas Timpson, PhD[25]; Manuel A Ferreira, PhD[26]; David P Strachan, MD[27]; John Henderson, MD[28]; David Hinds, PhD[16]; Hans Bisgaard, MD, DMSci[1]; Klaus Bønnelykke, MD, PhD[1]

1: Copenhagen Prospective Studies on Asthma in Childhood, Health Sciences, University of Copenhagen, Danish Pediatric Asthma Center, Copenhagen University Hospital, Gentofte, Denmark

2: Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

3: Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark

4: Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, USA

5: Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, USA

6: Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK

7: The University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Brisbane, Australia

8: School of Women's and Infants' Health, The University of Western Australia, Perth, Australia

9: Ludwig-Maximilians-University of Munich, Dr. von Hauner Children's Hospital, Munich, Germany

10: Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, 9711, The Netherlands

11: Department of Genetics, Harvard Medical School, Boston, 02115, USA

12: Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Perth, Western Australia, Australia

13: School of Medicine and Pharmacology, University of WA, Nedlands, Australia

14: Department of Pulmonary Physiology, West Australian Sleep Disorders Research Institute, Nedlands, Australia

15: University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester NHS Foundation Trust, Manchester, UK

16: 23andMe, Mountain View, Calif.

17: Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, GRIAC, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

18: University of Groningen, Department of Pulmonology, GRIAC, University Medical Center Groningen, Groningen, The Netherlands

19: Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment and Health, School of Public Health, Imperial College London, UK

20: Center for Life Course Epidemiology, Faculty of Medicine, P.O.Box 5000, FI-90014 University of Oulu, Finland

21: Biocenter Oulu, P.O.Box 5000, Aapistie 5A, FI-90014 University of Oulu, Finland

22: Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, P.O.Box 20, FI-90220 Oulu, 90029 OYS, Finland

23: Department of Children and Young People and Families, National Institute for Health and Welfare, Aapistie 1, Box 310, FI-90101 Oulu, Finland

24: University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester NHS Foundation Trust, Manchester, UK

25: The MRC Integrative Epidemiology Unit at the University of Bristol, UK

26: Queensland Institute of Medical Research, Brisbane, Australia

27: Population Health Research Institute, St George's, University of London, Cranmer Terrace, London SW17 0RE

28: School of Social and Community Medicine, University of Bristol, Bristol, UK

†: These authors contributed equally to this work.

Abstract

Background

The relationship between allergy and autoimmune disorders is complex and poorly understood.

Objective

To investigate commonalities in genetic loci and pathways between allergy and autoimmune diseases to elucidate shared disease mechanisms.

Methods

We meta-analyzed two GWAS on self-reported allergy and sensitization comprising a total of 62,330 individuals. These results were used to calculate enrichment for SNPs previously associated with autoimmune diseases. Furthermore, we probed for enrichment within genetic pathways and of transcription factor binding sites, and characterized commonalities in the variant burden on tissue-specific regulatory sites by calculating the enrichment of allergy SNPs falling in gene regulatory regions in various cells using Encode Roadmap DHS data, and compared the allergy data with all known diseases.

Results

Among 290 loci previously associated with 16 autoimmune diseases, we found a significant enrichment of loci also associated with allergy ($p=1.4e-17$) encompassing 29 loci at a false discovery rate<0.05. Such enrichment seemed to be a general characteristic for all

autoimmune diseases. Among the common loci, 48% had the same direction of effect for allergy and autoimmune diseases. Additionally, we observed an enrichment of allergy SNPs falling within immune pathways and regions of chromatin accessible in immune cells that was also represented in autoimmune diseases, but not in other diseases.

Conclusion

We identified shared susceptibility loci and commonalities in pathways between allergy and autoimmune diseases, suggesting shared diseases mechanisms. Further studies of these shared genetic mechanisms might help understanding the complex relationship between these diseases, including the parallel increase in disease prevalence.

Capsule Summary

We identified shared susceptibility loci and commonalities in pathways between allergy and autoimmune diseases. Further studies of these loci and related mechanisms might help understanding the complex relationship between allergy and autoimmunity.

Key messages

- Allergy and autoimmune diseases share genetic susceptibility loci.
- These results indicate commonalities in gene regulation and genetic pathways between allergy and autoimmune diseases.
- Further studies of common genetic loci and the related mechanisms might help understanding the complex relationship between allergy and autoimmunity.

Key words

Allergy, Single Nucleotide Polymorphism, Autoimmune Disease, Autoimmunity, Genetic Association Studies

Abbreviations

DEPICT: Data-driven Expression Prioritized Integration for Complex Traits gene set enrichment method

DHS: DNase Hypersensitive Sites

GWA(S): Genome Wide Association (Studies)

PCA: Principal Component Analysis

SNP: Single Nucleotide Polymorphism

Introduction

There has been a parallel increase in allergic and autoimmune disorders in recent decades in "westernized" countries[1] suggesting that these diseases may share disease mechanisms and etiologies. In line with this, allergy and autoimmune disorders seem to share environmental risk factors including birth by caesarian section[2]. This observation is in apparent contrast to the understanding of allergy and autoimmune diseases as representatives of distinct immunological disorders with counteracting underlying immune mechanisms. Autoimmune diseases are in general thought to act through a Th1/Th17-driven cell mediated immune response,[3] while allergy and asthma encompass a Th2-mediated response[4]. Counteracting immune mechanisms have been supported by some epidemiological studies of comorbidity suggesting a lower incidence of allergy among patients with autoimmune diseases, including rheumatoid arthritis[5,6] and multiples sclerosis[7], although the evidence of such inverse relationship is conflicting.[8]

Both allergy and autoimmune diseases are highly heritable diseases and an increasing number of susceptibility loci have been discovered.[9–11] We hypothesized that studying commonalities in the genetic architecture could provide a key to understanding the complex relationship between these diseases by pinpointing possible common disease mechanisms. In a wider perspective, this might help to explain the mechanisms and aetiologies responsible for the contemporaneous, dramatic increase in incidence of allergic and autoimmune disorders[1].

We studied commonalities between allergy and autoimmune disease in terms of susceptibility loci, genetic pathways and regulatory mechanisms. We meta-analyzed two

GWAS on allergic sensitization and allergic symptoms representing the largest GWAS on allergy to date. These data were then combined with publically available GWAS data on autoimmune diseases, as well as public data on molecular pathways, transcription factor binding sites and regulatory DNA regions. The primary analyses used a methodological approach that was agnostic to the direction of effect (including loci where the risk-allele was the same for allergy and autoimmune disease as well as loci where the risk allele for one disease was protective for the other). We hypothesized that loci with same direction of effect for different diseases might be involved in shared mechanisms and thereby help understanding the parallel epidemics of diseases, while loci with opposite direction of effect might help understanding counteracting (diverging) disease mechanisms. In analogue, gene variants in beta-adrenoreceptors are likely to show diverging effects with respect to risk of cardiac and respiratory diseases but still pinpoint an essential mechanism for both diseases. For allergy and autoimmune diseases such loci might be involved in “polarization” of the immune response. We therefore chose an approach that would capture common genetic loci with both same and opposite direction of effect since we believe that both might confer important information and thereby help to elucidate the complex mechanistic relationship between these diseases.

Methods

Study group

To obtain the largest possible GWAS data set for allergy, we meta-analyzed results from two recent GWAS on self-reported allergy[9] and sensitization[10], including ~2.2 million directly genotyped and imputed single nucleotide polymorphisms (SNPs), using a fixed effects model. The self-reported allergy dataset comprised 23,335 individuals with self-reported allergy for cat-, dust-mite and/or pollen-allergy and 26,311 control subjects without symptoms. The allergic sensitization dataset included 5,809 subjects with allergic sensitization defined by either skin prick test or specific IgE measurement and 9,875 control subjects without allergic sensitization with data from the following cohorts: AAGC, ALSPAC, B58C, COPSAC2000, LISA, MAAS, NFBC 1966, RAINE and PIAMA (**Supplementary Methods**).

Common genetic loci

To identify common genetic loci, we identified all SNPs previously shown to be associated with autoimmune diseases from the NHGRI GWAS catalog[12]. From each locus defined by genetic distance (**Supplementary Methods**) the most significant SNP associated with autoimmune diseases (the index disease, one of 16 different diseases) was chosen, leaving 290 candidate SNPs which where subsequently extracted from the allergy meta-analysis results (**Supplementary Methods, Supplementary Table 1 and Supplementary Figure 1**). For each SNP, the direction of effect of the risk allele for allergy was compared with the direction of effect for the index autoimmune disease as well as for other autoimmune diseases with reported loci in high LD.

Among these autoimmune disease-associated SNPs (loci) we calculated an enrichment of significant SNPs in relation to allergy ($P<0.05$) as compared to expected under the null hypothesis. Enrichment was only calculated for diseases with at least ten loci associated with the index disease and therefore not calculated for systemic sclerosis, sarcoidosis, primary sclerosing cholangitis and myasthenia gravis. As a methodological negative control, SNPs related to two non-inflammatory phenotypes with reasonable GWA study sizes, migraine and the combination of bipolar disorder and schizophrenia from the GWAS catalog, were analyzed similarly.

Commonalities in genetic pathways

To investigate commonalities in functional origins, the Data-driven Expression Prioritized Integration for Complex Traits (DEPICT) gene set enrichment method[13] was applied to 123 diseases and traits in the GWAS catalog, as well as the current association data set on allergy, by analyzing gene set enrichments for 14,416 reconstituted gene sets capturing a wide spectrum of molecular pathways, functional annotations and mouse phenotypes. For visualization purposes the resulting enrichment matrix was reduced to fewer dimensions by principal component analysis (**Figure 1**, **Supplementary Methods**). Similarly, this approach was applied to the common loci between allergy and autoimmune diseases (**Figure 2**), and to allergy and Crohn's disease separately to identify common and disease specific networks (**Figure 3**, **Supplementary Methods**). Crohn's disease was used as a representative of autoimmune diseases as this represents the largest, with respect to sample size, public available GWAS dataset to date of an autoimmune disease[14]. Also, we consider this a good

representative of autoimmune diseases since it is considered to be Th1 driven immune like most autoimmune diseases.

The enrichment of transcription factor (TF) binding sites in loci common between allergy and autoimmune diseases was calculated using ENCODE data by assessing the overlap between loci and binding sites for 161 TFs compared to random expectations (**Supplementary Methods**).

Common disease-implicated cell types

To identify and visualize common disease-implicated cell types, the tendency of disease associated loci to fall in cell-type specific regulatory DNase Hypersensitive Sites (DHS) (a proxy for accessible and/or regulatory DNA) was calculated for all diseases in the GWAS catalog, as well as allergy based upon the current association data-set. This enrichment was computed for 168 cell types and cell lines (hereafter described as cells) from the ENCODE Roadmap repository[15]. Duplicates and directly redundant cell types were removed before analyses. As with pathway enrichments, the resulting enrichment matrix was reduced to fewer dimensions by PCA for visualization (**Supplementary Methods**). For allergy this was also compared for the full range of p-values within the allergy meta-analysis in bins of decreasing p-value thresholds, essentially as done by Maurano and colleagues[16]. This was similarly done for Crohn's disease for comparison[14] using data from a GWAS metanalysis[14] including data from 6 studies of European descent with in total 6,333 cases and 15,056 controls in the discovery imputed with Hapmap III or II with in total of 953,242 autosomal SNPs. To validate specific findings, the DHS enrichment was re-calculated for enhancer regions in data from the more conservative FANTOM5 set[17].

Results

Allergy meta-analysis

By use of the joint meta-analysis on combined data from allergic sensitization and self reported allergy we increased the number of allergy-associated SNPs compared to the previous GWAS (**Supplementary Figure 2**) resulting in a total number of 19 genome-wide significant loci (**Supplementary Figure 3**). One of these; rs11122898 near *ANAPC1/MERKT* (p=1.9e-8) has not previously been associated with allergy or any related trait. In addition there were 5 novel suggestive loci (p<e-5): rs7612543 near *ZBTB38* (p=1.0e-7); rs9790601 near *NFKB1* (P=7.4e-8); rs7072398 near *IL2RA* (p=5.2e-7); rs12365699 near *CXCR5* (p=6.5e-7) and rs12900122 near *RORA* (p=3.5e-7) (**Supplementary Figure 4 and Supplementary Table 2**). Notably, all of these suggestive loci have previously been described in relation to autoimmune disorders[11,18–30].

Commonalities in genetic loci, pathways and disease implicated cell types between allergy and autoimmune diseases

There was a significant enrichment of autoimmune-associated loci with low p-values among the allergy loci from the current meta-analysis compared with expected (enrichment OR=4.36 [3.2-5.9], p=1.4e-17) (**Figure 1A, Table 1 and supplementary table 3**). A similar enrichment was seen for the two allergy phenotypes separately (self reported allergy: OR=4.1 [3.0-5.5] p=1.1e-15, and allergic sensitization: OR=2.7 [1.6-4.0] p=1.8e-5) (**Supplementary Figure 5**). This enrichment was also seen for the individual autoimmune diseases (**Figure 1A and Supplementary Figure 6**), although not statistically significant for

Systemic Lupus Erythematosus and Ankylosing Spondylitis and for Psoriasis and Graves Disease after adjusting for multiple testing (**Table 1**). We a priori chose a P-value threshold for the calculations of enrichment of <0.05. Using lower thresholds (P <0.01 or 0.001) generally resulted in higher enrichment and significant enrichment for all autoimmune diseases.

As a methodological negative control, SNPs related to two non-inflammatory phenotypes with a large number of known genome-wide significant loci, migraine and the combination of bipolar disorder and schizophrenia from the GWAS catalog were extracted from the allergy meta-analysis. These showed no significant overlap with allergy loci (**Supplementary Figures 7 and 8**).

Using DEPICT to identify significantly enriched reconstituted gene sets for allergy (**Supplementary Table 4**) and all diseases in the GWAS catalog, we identified a strong separation of autoimmune diseases, allergy and asthma from other traits on the first PCA component (Wilcoxon rank sum test p_{PC1}=8.142e-08, Wilcoxon rank sum test) (**Figure 1B and Supplementary Figures 9 and 10**). Notably, no other disease groups seemed to cluster strongly together (**Supplementary Figures 9 and 10**).

Analysis of all published SNP-to-trait associations and the tendency of these SNPs to fall within regions of open chromatin (represented by DHS) in specific cell types[31] likewise revealed that autoimmune diseases, allergy and asthma clustered together, and differentiated from other non-immune diseases on the first two PCA components (**Supplementary Figures 11 and 12**, p_{PC1}=0.0035, p_{PC2}=3.856e-07, Wilcoxon rank sum test). Loadings indicated that immune cells are responsible for this partitioning (**Supplementary**

Figure 13). Hierarchical clustering of the DHS sites within immune cells similarly showed the tendency of co-clustering of autoimmune diseases with allergy and asthma (**Supplementary Figure 14**). Specific DHS-analyses of allergy and Crohn's disease showed similar enrichment in variants falling within DHS regions of immune cells (**Supplementary Figure 15**). For allergy, these findings were validated in the FANTOM5 high confidence enhancer data set, which showed comparable enrichment among immune cells (**Supplementary Figure 16**)"

The specific common genetic loci

Out of the 290 autoimmune disease SNPs investigated, 29 were significantly associated with allergy at a FDR <0.05 (**Table 2**). Eleven of these common loci (*C11ORF30*, *LPP*, *PLCL1*, *HLA-B*, *SMAD3*, *IKZF3*, *MYC*, *CLEC16A, NDFIP1, BACH2* and *IL2RA/IL15RA*) were already reported to be associated with allergy[9,10,32,33] (**Table 2 and Supplementary Table 5 and 6**). The remaining 18 overlapping loci included the mapped genes: *NFKB1*, *SH2B3*, *AKAP11/TNFSF11*, *ABO*, *C12ORF30*, *CD247*, *RADSNB*, *EP300*, *RORC*, *PSMG1*, *HDAC7/VDR*, *HLA-DPB2*, *BACH2*, *TNFAIP3*, *KIF1B*, *RPS6KA4*, *ERBB3*, *THADA*, and *GLB1/CCR4*. Look up of the 29 common loci in the separate results from the 2 different allergy-phenotypes (allergic symptoms and allergic sensitization) generally showed similar results, including same direction of effect for the two allergy-phenotypes for 28 out of 29 loci (**Supplementary Table 5**). For the majority of loci the effect size was higher for allergic sensitization compared to allergic symptoms, as would be expected due to the more homogenous phenotype, and for the majority of loci with evidence of heterogeneity between phenotypes, heterogeneity was due to this. The NDFIP1 locus seemed only to be associated with self reported allergy.

Comparison of the direction of effect between allergy and autoimmune disease was possible for 27 of the 29 common loci compared with the index autoimmune disease. For 13 of these (48%), we found the same direction of effect meaning that the allele increasing the risk of allergy also increases the risk of the autoimmune disease (**Table 2**). For the remaining 14 loci (52%), we found opposite direction of effect with the risk allele for allergy being associated with reduced risk of the index autoimmune disease (**Table 2 and Supplementary Table 5**). The majority of autoimmune diseases showed examples of both same and opposite direction of effect compared to allergy. For some of the common loci where several autoimmune diseases have reported association, the different autoimmune diseases showed differences in direction of effect relative to allergy (*HLA-B*, *BACH2* and *RPS6KA4*). After gene set enrichment analyses of these 29 common loci, the resulting significant hits were all ontological terms relating to immune function, including immunoglobulin diversification and production, T-and B-cell activation, signaling by nuclear receptors and abnormal immune cell physiology. (**Figure 2**, **Supplementary Figure 17 and Supplementary Table 7**). Coloring for loci with same vs. opposite direction of effect for allergy and autoimmune diseases showed no clear systematic differences (Figure 2).

To investigate the potential enrichment for certain transcription factors in mediating the effect at common loci, public transcription factor binding data from ENCODE were analyzed for overlap within loci common between allergy and autoimmune diseases. This revealed significant enrichment for several immune related transcription factors (**Supplementary Figure 18**).

In order to investigate if the common loci seem to tag the same genetic variant, we performed regional association plots showing that, for almost all loci, the autoimmune disease tagging SNP is also within the LD-block with strongest regional association with allergy (**Supplementary Figure 19**). Tagging of the same genetic variation at common loci was further supported by paired comparison of regional association plots for Crohn's disease[14] and allergy, respectively, for the most statistically significant common loci (**Supplementary Figure 20**).

Shared and differential genetic pathways for Allergy and Crohn's disease

A direct comparison of pathways targeted by allergy-related loci vs. Crohn's disease loci revealed that a large proportion of shared as well as disease-specific pathways, with a strong predominance of T cell signaling modules within the shared grouping (**Figure 3**).

Discussion

Our study demonstrated substantial commonalities between allergy and autoimmune diseases in terms of susceptibility loci, genetic pathways and genomic regulatory sites (DHS). This overlap in genetic mechanisms seemed to be a general phenomenon for allergy autoimmune diseases and distinct from other diseases. Our study identifies a substantial number of novel overlapping loci for allergy and autoimmune diseases suggesting both shared (increasing risk of both autoimmune and allergic disease) as well as diverging genetic mechanisms.

Strengths and limitations

By combining two GWAS on allergic sensitization and allergic symptoms respectively, we were able to obtain the most powerful GWAS dataset on allergy to date, which allowed a systematic analysis of the genetic commonalities of allergy and autoimmune diseases. The clinical phenotypes are not identical, with allergic sensitization often being present without symptoms and vice versa, which can also be seen as a limitation of the study. However, the phenotypes are closely correlated genetically as suggested from the initial publications of the individual GWAS' showing highly consistent mutual replication of top-SNPs between studies[9,10]. The improved number of significant SNPs and a higher number of genome-wide significant loci in the combined meta-analysis performed here compared with the individual GWAS results also underscores the validity of the combined meta-analysis approach. Furthermore, the common loci between allergy and autoimmunity generally showed similar results for the two allergy-related phenotypes (**Supplementary Table 5**), and

the enrichment of autoimmune loci was also similar for both phenotypes (**Supplementary Figure 5**), strongly suggesting that the conclusions of this study is not affected by the combination of allergy-related phenotypes.

In the primary analyses we combined GWAS loci from several autoimmune diseases. This was done in order to assess commonalities between allergy and autoimmune diseases in general and to obtain enough autoimmune loci and thereby statistical power to perform more systematic analyses. However, it should be noted that autoimmune diseases represent a heterogenous group of diseases, and the genetic architecture of autoimmune genes seem to include shared, but also differential and opposing genetic mechanisms[34-36]. We therefore also performed separate enrichment analyses for the different autoimmune diseases showing that enrichment for allergy-related loci seem to be the case for the majority of autoimmune diseases (**Supplementary table 8**).

It is a limitation that we did not have access to GWAS data for all autoimmune diseases. For the regional analyses of shared loci, a publicly available GWAS dataset on Crohn's disease was therefore used as a representative of autoimmune disease. It is a limitation of the analyses based on GWAS-catalog markers, that genotype chips and imputation panels and procedures differ between studies, adding marker coverage heterogeneity.

Our study is also limited by differences in study size of GWAS on autoimmune diseases, meaning that the diseases where the largest GWAS have been performed, and most loci have been discovered, will have a relatively larger impact on the results of the analysis combining all autoimmune diseases. The number of reported loci for each autoimmune disease is shown in Table 1.

Interpretation

We found substantial overlap of susceptibility loci for allergy and autoimmune diseases suggesting commonalities in the genetic background and hence the possibility of shared pathogenetic mechanisms. The possible common etiology of allergy and autoimmune diseases was further supported by co-clustering of autoimmune diseases, asthma and allergy in terms of genetic pathways and regulatory sites implying shared disease pathways beyond overlap of associated SNPs.

To our knowledge, no previous study has had sufficient statistical power to systematically explore commonalities in the genetic architecture between allergy and autoimmunity. One previous study found no association between susceptibility loci for type 1 diabetes and total IgE levels[37] arguing against a shared genetic background. However, that study had much lower statistical power than the present study, and the genetics of total IgE has been shown to be different from the genetics of allergen-specific IgE[10]. Another study compared results from a GWAS on asthma with published GWAS results on autoimmune diseases and found evidence of 7 overlapping susceptibility loci, both showing examples of opposite and same direction of effect for asthma and autoimmune diseases[38].

Importantly, our study demonstrates common loci with both the same and opposite direction of effects, potentially pointing towards both converging (shared) and diverging (counteracting) mechanisms, either increasing the risk of both allergy and autoimmune diseases, or increasing the risk for one of the diseases while protecting against the other. Our study thereby provides further understanding of the complex genetic relationship between allergy and autoimmune diseases. We hypothesize that the common loci with same

394 direction of effect may be involved in the mechanisms causing the contemporaneous
395 epidemics of allergy and autoimmune diseases[39] by increasing the susceptibility to immune
396 disorders in general, probably by mechanisms involving gene-environment interaction.
397 Complementarily, common loci with opposite direction of effect may be involved in
398 mechanisms determining the type of immune disorder developing in the individual, resulting
399 in the inverse relationship observed between allergy and some autoimmune diseases[5–7].
400 Understanding the mechanisms of these common genetic loci may improve understanding
401 of the epidemics of allergy and autoimmune diseases. It may also help predict how targeting
402 specific disease mechanisms could have the unintended consequence of increasing the risk
403 of other diseases. We expect the loci with *same direction of effect* to be of particular
404 importance in the search for common mechanisms driving the parallel increases in disease
405 incidence. Of specific interest are the loci at *C11orf30*, *PLCL1* and *SMAD3,* which were
406 strongly associated with allergy as well as several autoimmune diseases. Of these, *C11orf30*
407 and *SMAD3* were also identified in a previous study comparing asthma and autoimmune
408 diseases[38]. Some of the loci with same direction of effect were linked to immune-related
409 transcription factors (*RORC*, *SMAD3*), transcriptional co-factors (*EP300*), cell-cycle regulators
410 (*THADA*) or regulators of transcription (*C11ORF30*, *PLCL1*, *AKAP11*, *NDFIP1*). Several are
411 directly implicated in regulation of regulatory T cells (Treg, *SMAD3*[40], *EP300*[41]) and Th17 cells
412 (*RORC*[42], *SMAD3*[40]), or regulation of immune activity *(IL2RA*). This supports a commonality of
413 autoimmune and allergic diseases based on defects in immune suppressive functions. Other
414 loci with same direction of effect included *CLEC16A*, *TNFSF11*, *CCR4* and *GLB1*, all involved in
415 immune function by different means.

Loci with *opposite direction of effect* might be involved in "polarization" of the immune response. Since allergy is acknowledged to involve Th2-mediated pathology, while most autoimmune diseases involve Th1 cells as well as pathogenic Th17 cells, any genetic factors that perturb Th differentiation ability or the immunosuppressive Treg function could potentially influence the risk of specific disease development. The loci with opposite direction of effect included *LPP*, *NFKB1, TNFAIP3,* all involved in immune cell signaling and activation/deactivation processes. Several of the other loci with opposite direction of effect function as possible regulators of immune cell differentiation (*HDAC7*/*VDR*[43,44], *IKZF3*[45], *SH2B3*, *BACH2*, *RPS6KA4* (*MSK2*)[46]) and cell-cycle regulation (*RAD51L1*), while the remaining loci (*C12ORF30*, *PSMG1*, *HLA-DPB2*, *KIF1B*) are more general regulators of cellular function. It should however be noted that given the complexity of genetic regulation and the fact that the causal genetic variant is unknown for most susceptibility loci, the finding of common genetic loci with apparent opposite direction of effect should be interpreted with caution. It is possible that this is the result of different underlying causal variants in the region, although comparison of regional association plots suggested that the association signals for allergy and autoimmune disease did tag the same genetic variation. Potential polarizing mechanisms associated with these loci should be addressed in future studies, e.g. by demonstrating opposite disease relationships on the level of gene expression and/or protein levels.

Gene set enrichment analyses of the 29 shared loci (**Figure 2**, **Supplementary Figure 17 and Supplementary Table 7**) indicate that immune functioning and activation status specifically within T cells represents a shared focal point for allergy and autoimmune diseases.

Involvement of shared immunological paths was further supported based on the type of transcription factors found to be enriched within the common loci where generic immune regulatory transcription factors including MTA3, WRNIP1 and IKZF1 (Ikaros) turned out as central players (**Supplementary Figure 18**). MTA3 has been shown to be involved in B-cell and T helper cell differentiation[47,48] and WRNIP1 has been reported to regulate expression of transcription factors involved in Treg functioning[49]. IKZF1 activates extensive transcriptional programs involving especially regulation of T, B and NK cells with effects on differentiation, proliferation and apoptotic programs in these cell types[50]. Moreover, several other directly immune-related transcription factors were identified amongst the top 10 hits, including NFATC1, STAT2, MEF2C, RELA, SP2 and ZEB1[51–56].

The comparative pathway analysis between allergy and Crohn's disease highlights that the common pathways of these diseases are founded on (de)regulation of adaptive immune signaling, involving TCR, CD28, IL2/3/5/6, IFNg, GM-CSF, JAK/STAT and IL receptor signaling as well as of apoptosis and IgA production. The Crohn's disease-specific network primarily included pathways associated with innate immune activation involving TLR and NOD signaling, while allergy-specific pathways were associated with signaling from receptor tyrosine kinases such as EGFR. (**Figure 3**)

A strong co-clustering and separation of allergy, asthma and autoimmune diseases in terms of enrichment of SNPs in cell type specific DHS sites, points towards common immune modulatory mechanisms facilitated by the effect of the genetic SNP burden in immune cell-specific regulatory regions (**Supplementary Figures 11 and 13**). For Allergy and Crohn's disease, these related primarily to T and B cell functions (**Supplementary Figure 15**). In

accordance with our findings of coinciding immune cell involvement in several immune-mediated diseases, one previous study on Multiple Sclerosis reported several genetic markers in relevant DHS sites in immune cells.[57] Accordingly and in line with our findings a recent paper focusing on autoimmune disease variants show how these immune mediated diseases are correlated and cluster in tendency for associated variants to be enriched in specifically immune cell enhancers.[58] Moreover, the enrichment of immune-cell specific DHS sites specifically within promoter regions (**Supplementary Methods, Supplementary Figure 21**) may be a result of the adaptability and immediate early gene response requirements of the immune system, and since their positioning is essential for gene expression levels this might also explain why the identified specific SNP variants pose increased risk for disease penetrance.

In addition to pinpointing common loci between allergy and autoimmune diseases, our identification of common loci also suggests a large number of novel allergy loci; 18 of the 29 common loci have not previously been associated with allergy (**Table 2**). Our study suggests that these are all susceptibility loci for allergy that did not reach the criteria for genome-wide significance and was therefore not discovered in previous GWAS.

Furthermore, our combined meta-analysis on allergy and allergic symptoms identified suggestive allergy loci with several previously related to immune function (*NFKB1*[59] and *CXCR5*[60], **Supplementary Table 2**). One novel suggestive allergy locus was discovered at rs11122898 and reached the genome-wide significance threshold. This locus is close to *ANAPC1* but is most likely affecting the upstream *MERKT* gene (**Supplementary Figure 22**), which is involved in regulation of dendritic cells via B-cell activating factor (BAFF)[61]. One

484 additional suggestive locus was at *ZBTB38*, and genes within this family have been shown to
485 have important function in naive B-cell differentiation[62], and have been associated to
486 eczema[63] and asthma with hay fever[33] although it has not been shown that these genes are
487 the causal genes within these loci.
488
489 In conclusion, we performed the first large study on commonalities in the genetic origins of
490 allergy and autoimmune diseases and documented substantial genetic overlap between
491 these diseases. The recent availability of a vast array of genomics data from the ENCODE and
492 other consortia provides a solid foundation for systems biology analysis in disease settings.
493 Exploiting this approach, we identified common molecular pathways between allergy and
494 autoimmune diseases, identical patterns of overlaps with open chromatin specifically within
495 immune cell-specific regulatory regions, and overlap in transcription factor binding sites,
496 emphasizing common characteristics in gene regulation. Further investigation of these
497 commonalities in genetic mechanisms might improve understanding of important biological
498 pathways that increases the risk of allergy and autoimmune diseases as well as mechanisms
499 differentiating these diverse diseases. Understanding of potential shared genetic origins of
500 allergy and autoimmune diseases, maybe particularly related to the loci with the same
501 direction of effect, could point to vulnerable “hot” points in immune system pathways that
502 may also be affected by other modifiers such as epigenetics and environmental exposures.
503 This insight might provide important clues for understanding the parallel epidemics of these
504 diseases and thereby enforce future disease prevention.
505
506

Acknowledgements and funding

Please see supplement.

Author Contributions

Statistical and functional analysis: E.K-M., J.W., M.S, T.H.P, A.C.A, N.M.W., L.F., J.N.H, A.S., G.H.K., D.S.P., A.C., N.T., D.P.S, D.H.,

Study design: E.K-M., J.W., A.S., J.Y.T., G.H.K., D.S.P., A.C., N.T., M.A.F, D.P.S, A.J.H., D.H., H.B., K.B.

Writing paper: E.K-M., J.W., M.S, S.B., T.H.P, H.B., K.B.

Data collection: E.K.-M., M.S, E.F., A.J., A.S., G.H.K., D.S.P., C.E.P., A.C., D.P.S, A.J.H., H.B., K.B.

Genotyping: C.M.T.T., A.S., G.H.K., C.E.P., A.C., H.B.

Revising and reviewing paper: E.K.-M., J.W., M.S, S.B., T.H.P, A.C.A, N.M.W., C.M.T.T., E.F., L.F., J.N.H, A.J., A.S., J.Y.T., G.H.K., D.S.P., C.E.P., M.R.J., A.C., N.T., M.A.F, D.P.S, A.J.H., D.H., H.B., K.B.

Competing financial interests

D.H. and J.Y.T. are employees of 23andMe. The remaining authors declared no competing interests of relevance to this paper.

Figure Legends

Figure 1. Commonalities in susceptibility loci and genetic pathways between allergy and autoimmune diseases

A) Autoimmune disease-associated loci and their association with allergy. Quantile-quantile plots representing the observed vs. expected p-values in the combined allergy GWAS for all 290 candidate autoimmune disease-associated loci (large panel), and separately for selected autoimmune diseases (smaller panels). Solid line reflects the p-value distribution under the null, while the dashed line is the distribution of all SNPs from the allergy meta-analysis. T1D: Type 1 diabetes.

B) Commonalities in genetic pathways. Principal component analysis of DEPICT gene set enrichment results based on trait- associated variants for 123 traits from the GWAS Catalog and allergy. The blue area represents the shared minimal ellipsoid area of allergy and autoimmune diseases.

Figure 2. Pathway-based analysis of common loci

Principal component analysis of DEPICT gene set enrichment results based on the 29 common loci between allergy and autoimmune diseases, with each dot representing a single enriched gene set. Loadings for the index genes are illustrated by arrows. Genes denoted in blue and orange represent same and opposite direction of effect for allergy, respectively, a*s compared to index autoimmune disease*. For genes denoted in gray, index autoimmune disease did not report effect allele. The blue areas represent cohesive clusters of gene sets

549 with similar immune function. The individual gene set names are shown in **Supplementary**

550 **Figure 15.**

551

552 **Figure 3. Common and disease-specific genetic pathways**

553 DEPICT gene-set enrichment map of common (blue), allergy-specific (red) and Crohn's

554 Disease-specific (green) pathways. The correlation between pathways is depicted by the line

555 width. Sets with a correlation < 0.4 and singletons are not shown.

556

557

558

559 **Table 1 – Enrichment of significant allergy loci among autoimmune disease-associated loci**

Phenotype	# Loci	# Sig. Loci*	Proportion Sig.	eOR	CI	P
All Autoimmune Loci	290	57	0.20	4.36	[3.20-5.85]	**1.40E-17§**
Inflammatory Bowel Disease	178	42	0.24	5.51	[3.80-7.84]	**2.60E-16§**
Crohn's Disease	97	26	0.27	6.53	[4.00-10.36]	**5.00E-12§**
Ulcerative Colitis	60	16	0.27	6.48	[3.41-11.72]	**6.40E-08§**
Multiple Sclerosis	54	12	0.22	5.09	[2.44-9.86]	**2.10E-05§**
Type 1 Diabetes	43	8	0.19	4.07	[1.63-8.94]	**1.70E-03§**
Arthritis	40	9	0.23	5.18	[2.17-11.14]	**2.00E-04§**
Celiac Disease	36	11	0.31	7.85	[3.48-16.53]	**1.70E-06§**
Systemic Lupus Erythematosus	27	4	0.15	3.10	[0.78-9.07]	5.30E-02
Primary Biliary Cirrosis	23	7	0.30	7.80	[2.71-20.03]	**1.40E-04§**
Psoriasis	19	4	0.21	4.75	[1.15-14.93]	**1.62E-02**
Graves Disease	15	3	0.20	4.46	[0.81-16.52]	**4.20E-02**
Ankylosing Spondylitis	11	2	0.18	3.96	[0.42-19.14]	1.10E-01

* With P<0.05 to allergy

Bold: Nominal significant. §: Significant at Bonferroni (0.05/13)

eOR: enrichment Odds-Ratio calculated for autoimmune diseases with at least 10 loci

560

561

562

563 **Table 2 – Common loci between allergy and autoimmune diseases**

Chr:BP	Gene	SNP	MAF*	Beta	P, FDR	Het. P**	Traits§	Direction £	Known allergy Locus	Other genes in LD > 0.5
11:75976842	*C11ORF30*	rs2155219[T]	0.47	0.12	6.10E-23	5.20E-03	CrD(+), **IBD(+)**, UC(+)	Same	Yes$	-
3:189595248	*LPP*	rs1464510[A]	0.48	-0.07	4.20E-07	1.40E-01	**CeD(-)**	Opposite	Yes$	-
2:198605140	*PLCL1*	rs6738825[A]	0.48	0.06	9.20E-07	1.90E-01	**CrD**(+)	Same	Yes$	ANKRD44, BOLL, COQ10B, HSPD1, HSPE1, MARS2, MOB4,, RFTN2, SF3B1
6:31444079	*HLA-B*	rs7743761[A]	0.26	-0.07	1.90E-06	4.00E-01	**AS(na)**, GD(-), SLE(+), UC(+)	na	Yes$	-
15:65229650	*SMAD3*	rs17293632[T]	0.27	0.07	2.00E-06	1.90E-01	**CrD(+)**, IBD(+)	Same	Yes$	-
17:35175785	*IKZF3*	rs907092[A]	0.49	-0.06	4.10E-05	8.30E-01	CrD(-), **IBD(-)**, PBC(-), T1D(-), UC(-)	Opposite	Yes$	GSDMB, ORMDL3, ZPBP2
8:128884211	*MYC*	rs4410871[T]	0.28	0.06	4.10E-05	8.60E-04	**MS(-)**	Opposite	Yes$	-
12:110368991	*SH2B3*	rs3184504[T]	0.45	-0.06	5.30E-05	7.50E-02	A(-), CeD(-), **T1D(-)**	Opposite	Novel	ACAD10, ADAM1, ALDH2, ATXN2, BRAP, C12orf47, C12orf51, ERP29, MAPKAP5, NAA25, TMEM116, TRAFD1
13:41950880	*AKAP11, TNFSF11*	rs2062305[A]	0.43	-0.05	2.00E-03	8.50E-01	**CrD(+)**	Same	Novel	-
16:11087374	*CLEC16A*	rs12708716[A]	0.35	0.05	2.00E-03	8.60E-02	MS(+), PBC(+), **T1D(+)**	Same	Yes$	-
9:135139050	*ABO*	rs505922[T]	0.37	-0.05	2.00E-03	2.90E-01	**GD(+)**	Opposite	Novel	-
12:110971201	*C12ORF30*	rs17696736[A]	0.42	0.05	2.20E-03	1.50E-01	**T1D(-)**	Opposite	Novel	-
4:103770651	*NFKB1€*	rs7665090[A]	0.49	0.04	2.70E-03	7.00E-01	**PBC(-)**	Opposite	Novel	MANBA
1:165678008	*CD247*	rs864537[A]	0.38	0.04	4.80E-03	6.10E-01	**A(+)**, CeD(+), SC(+)	Same	Novel	-
14:67823346	*RADSNB*	rs911263[T]	0.29	-0.04	6.40E-03	1.80E-02	**PBC(-)**	Opposite	Novel	-
22:39761288	*EP300*	rs4820425[A]	0.29	0.05	6.60E-03	8.60E-01	**CrD**(+)	Same	Novel	CHADL, DNAJB7, L3MBTL2, RANGAP1, ST13, XPNPEP3, ZC3H7B
1:150068	*RORC*	rs484560	0.14	-0.06	7.50E-03	7.40E-01	**IBD(+)**	Same	Novel	-

304		4[A]								
10:61095 67	*IL2RA, IL15RA*	rs127225 63[A]	0.12	0.06	7.50E-03	1.00E+00	**T1D(na)**	na	Yes$$	-
21:39387 404	*PSMG1*	rs283687 8[A]	0.26	0.05	7.50E-03	3.10E-01	AS(-), **IBD(-)**, UC(-)	Opposite	Novel	-
12:46494 635	*HDAC7, VDR*	rs111682 49[T]	0.48	0.04	8.70E-03	4.60E-01	**IBD(-)**	Opposite	Novel	-
6:331680 96	*HLA-DPB2*	rs228138 8[A]	0.02	-0.13	8.80E-03	3.10E-01	**GD(-)**	Opposite	Novel	-
6:910149 52	*BACH2*	rs117555 27[C]	0.42	0.04	8.80E-03	3.90E-02	CrD(+), **T1D(-)**	Opposite	Yes$$$	-
6:138048 197	*TNFAIP3*	rs692022 0[A]	0.22	-0.05	1.00E-02	3.20E-01	A(-), CeD(-), **IBD(-)**, UC(-)	Opposite	Novel	-
5:141493 388	*NDFIP1*	rs686341 1[A]	0.37	-0.04	1.00E-02	4.90E-02	**IBD(+)**	Same	Yes$$$	-
1:102756 99	*KIF1B*	rs104929 72[T]	0.34	0.04	1.40E-02	5.00E-01	**MS(-)**	Opposite	Novel	PGD, UBE4B
11:63864 311	*RPS6KA4*	rs663743[A]	0.36	0.04	2.50E-02	2.40E-01	CrD(-), **S(+)**	Same	Novel	CCDC88B, ESRRA, GPR137, PRDX5, TRMT112
12:56482 180	*ERBB3*	rs229223 9[T]	0.32	0.04	2.50E-02	2.10E-01	**T1D(+)**	Same	Novel	IKZF4, SUOX
2:436604 22	*THADA*	rs104959 03[T]	0.12	0.05	3.40E-02	5.50E-01	**CrD(+)**, IBD(+)	Same	Novel	-
3:329904 73	*GLB1, CCR4*	rs133149 93[T]	0.43	-0.03	4.70E-02	8.00E-02	**CeD(+)**	Same	Novel	-

** As reported by Bønnelykke et al.[5]*

*** Test for heterogeneity within the allergy meta-analysis*

§ Reported traits associated with locus (autoimmune trait marker LD with allergy marker > 0.6 and allergy marker p < 0.05) A: Arthritis, AS: Ankylosing Spondylitis, CeD: Celiac Disease, CrD: Crohn's Disease, GD: Graves Disease, IBD: Inflammatory Bowel Disease, MS: Multiple Sclerosis, P:Psoriasis, PBC: Primary Biliary Cirrhosis, SLE: Systemic lupus erythematosus, SS: Systemic Sclerosis UC: Ulcerative Colitis, T1D: Type 1 Diabetes. The index disease used in the comparison of effect direction marked by bold and underlined. +/- denotes direction of effect for the autoimmune disease compared to the direction allergy.

£ As compared to index autoimmune disease (underlined and bold in the "Traits" column)-SNP effect direction. Na: Some diseases have not reported the effect allele

$ Genome-wide significant association to allergy/allergic sesnsitization reported by Hinds et al, and/or Bønnelykke et al.[4,5]

$$ Suggestive association to asthma with hayfever reported by Ferreira et al.[33]

$$$ Suggestive association to asthma reported by Ferreira et al.[32]

564

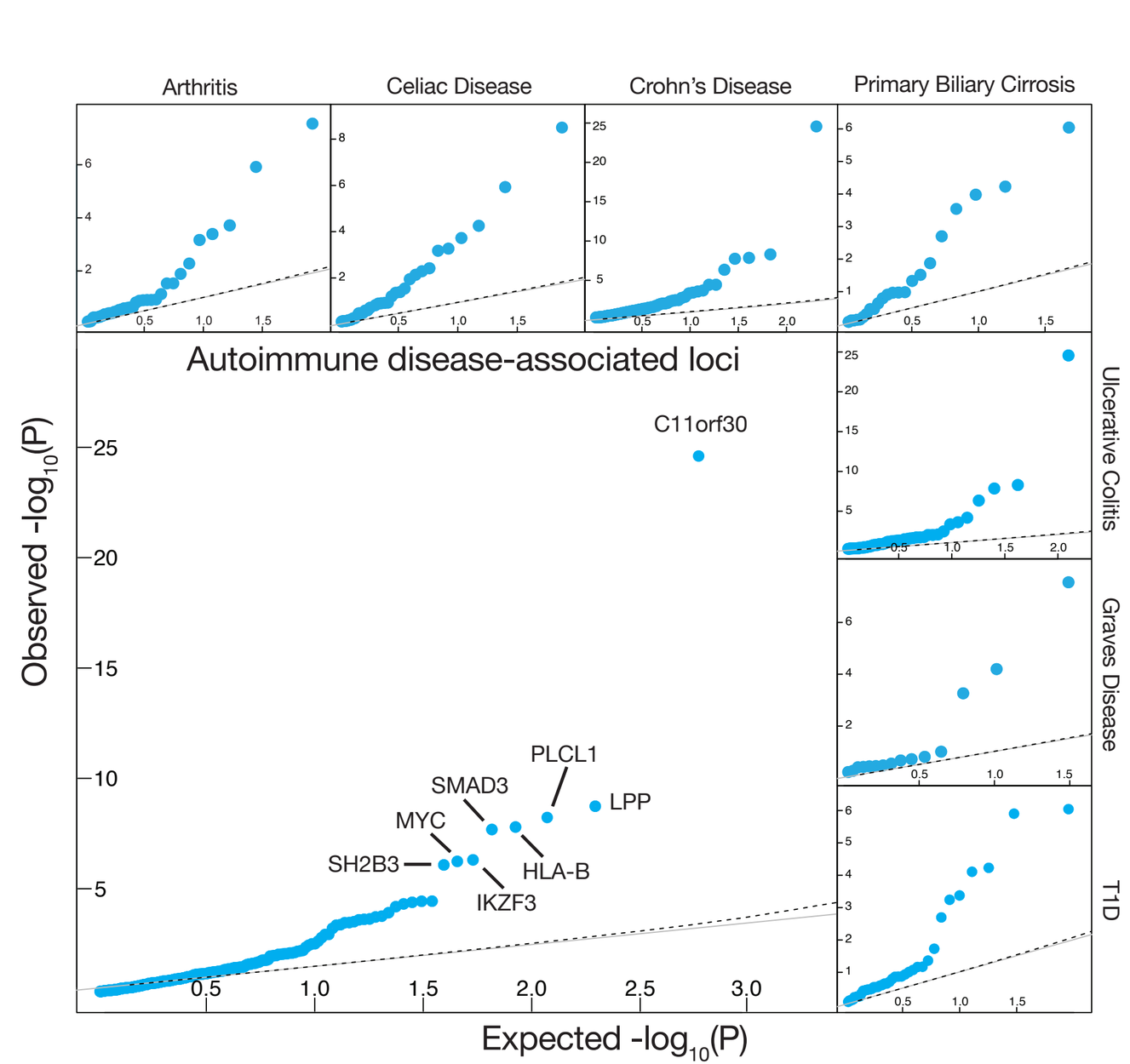
References

1. Bach J-F. The Effect of Infections on Susceptibility to Autoimmune and Allergic Diseases. N Engl J Med. 2002;347(12):911–20.
2. Sevelsted A, Stokholm J, Bønnelykke K, Bisgaard H. Cesarean section and chronic immune disorders. Pediatrics. 2015 Jan;135(1):e92–8.
3. Steinman L. A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell–mediated tissue damage. Nat Med. 2007 Jan;13(1):139–45.
4. Lloyd CM, Hessel EM. Functions of T cells in asthma: more than just TH2 cells. Nat Rev Immunol. 2010 Dec;10(12):838–48.
5. Rudwaleit M, Andermann B, Alten R, Sörensen H, Listing J, Zink A, et al. Atopic disorders in ankylosing spondylitis and rheumatoid arthritis. Ann Rheum Dis. 2002 Nov;61(11):968–74.
6. Verhoef CM, van Roon JA, Vianen ME, Bruijnzeel-Koomen CA, Lafeber FP, Bijlsma JW. Mutual antagonism of rheumatoid arthritis and hay fever; a role for type 1/type 2 T cell balance. Ann Rheum Dis. 1998 May;57(5):275–80.
7. Bergamaschi R, Villani S, Crabbio M, Ponzio M, Romani A, Verri A, et al. Inverse relationship between multiple sclerosis and allergic respiratory diseases. Neurol Sci Off J Ital Neurol Soc Ital Soc Clin Neurophysiol. 2009 Apr;30(2):115–8.
8. Rabin RL, Levinson AI. The nexus between atopic disease and autoimmunity: a review of the epidemiological and mechanistic literature. Clin Exp Immunol. 2008 Jul;153(1):19–30.
9. Hinds DA, McMahon G, Kiefer AK, Do CB, Eriksson N, Evans DM, et al. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. Nat Genet. 2013 Aug;45(8):907–11.
10. Bønnelykke K, Matheson MC, Pers TH, Granell R, Strachan DP, Alves AC, et al. Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. Nat Genet. 2013 Aug;45(8):902–6.
11. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012 Nov 1;491(7422):119–24.
12. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014 Jan;42(Database issue):D1001–6.
13. Pers TH, Karjalainen JM, Chan Y, Westra H-J, Wood AR, Yang J, et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6:5890.
14. Franke A, McGovern DPB, 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.
15. Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 Sep 6;489(7414):57–74.
16. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012 Sep 7;337(6099):1190–5.
17. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of active enhancers across human cell types and tissues. Nature. 2014 Mar

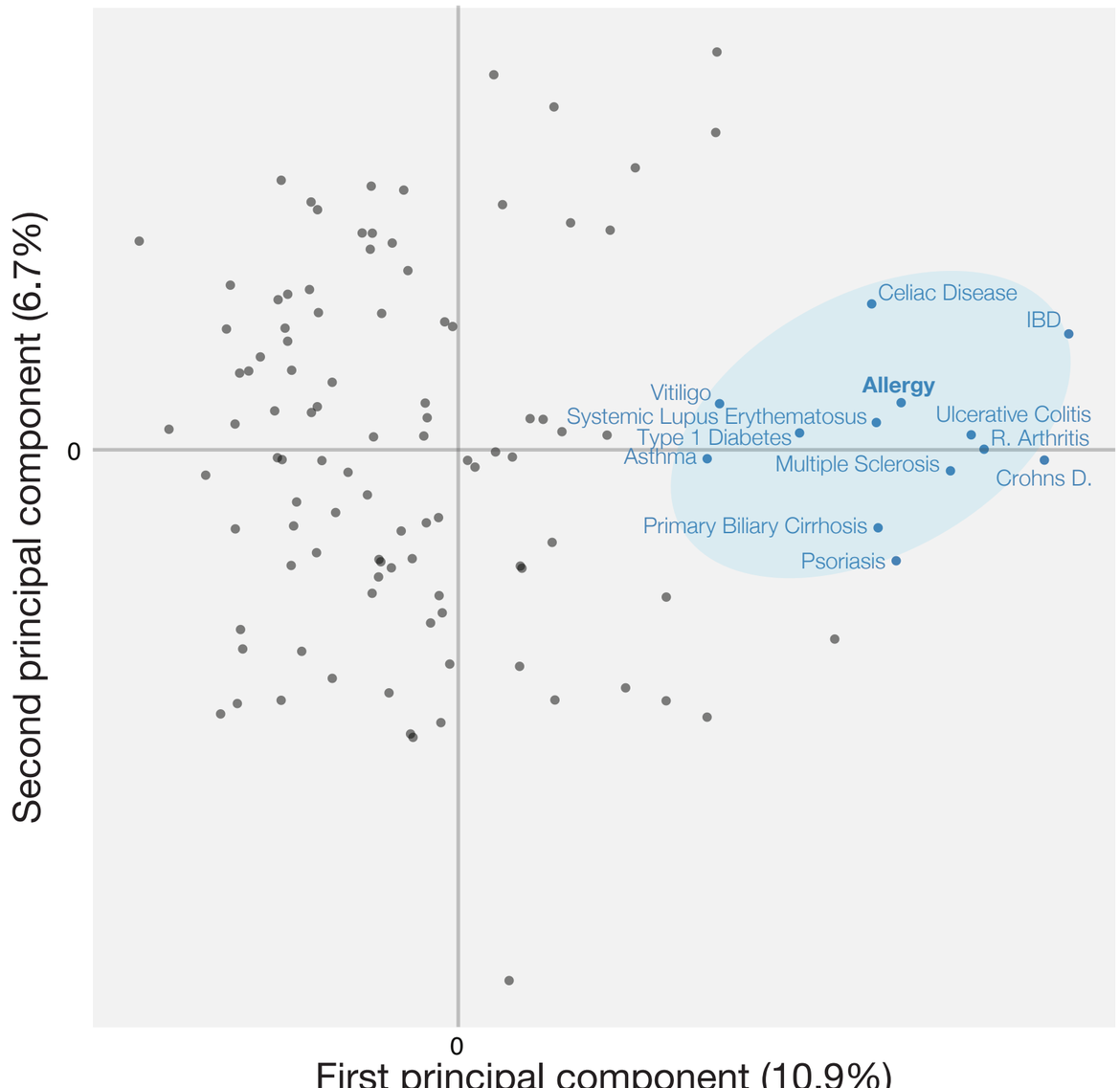
611 27;507(7493):455–61.
612 18. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M, et al. Genome-
613 wide association study identifies eight new susceptibility loci for atopic dermatitis in the
614 Japanese population. Nat Genet. 2012 Nov;44(11):1222–6.
615 19. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-
616 scale, consortium-based genomewide association study of asthma. N Engl J Med. 2010 Sep
617 23;363(13):1211–21.
618 20. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM,
619 et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial
620 infarction. Nat Genet. 2009 Mar;41(3):342–7.
621 21. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genome-wide scan
622 reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat Genet. 2009
623 Feb;41(2):199–204.
624 22. Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A, et al.
625 Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus
626 status-defined subgroups. J Natl Cancer Inst. 2012 Feb 8;104(3):240–53.
627 23. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-
628 wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat
629 Genet. 2010 Jun;42(6):508–14.
630 24. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control
631 Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, et al. Genetic risk and a
632 primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011 Aug
633 11;476(7359):214–9.
634 25. Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, et al.
635 Genome-wide association study in alopecia areata implicates both innate and adaptive
636 immunity. Nature. 2010 Jul 1;466(7302):113–7.
637 26. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al.
638 Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1
639 diabetes. Nat Genet. 2009 Jun;41(6):703–7.
640 27. Han J-W, Zheng H-F, Cui Y, Sun L-D, Ye D-Q, Hu Z, et al. Genome-wide association
641 study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus
642 erythematosus. Nat Genet. 2009 Nov;41(11):1234–7.
643 28. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al.
644 Hundreds of variants clustered in genomic loci and biological pathways affect human height.
645 Nature. 2010 Oct 14;467(7317):832–8.
646 29. Zhernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, et al.
647 Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis
648 identifies fourteen non-HLA shared loci. PLoS Genet. 2011 Feb;7(2):e1002004.
649 30. Mells GF, Floyd JAB, Morley KI, Cordell HJ, Franklin CS, Shin S-Y, et al. Genome-
650 wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat
651 Genet. 2011 Apr;43(4):329–32.
652 31. Gusev A, Lee SH, Trynka G, Finucane H, Vilhjálmsson BJ, Xu H, et al. Partitioning
653 heritability of regulatory and cell-type-specific variants across 11 common diseases. Am J
654 Hum Genet. 2014 Nov 6;95(5):535–52.
655 32. Ferreira MAR, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souëf P, et al.
656 Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet. 2011 Sep
657 10;378(9795):1006–14.

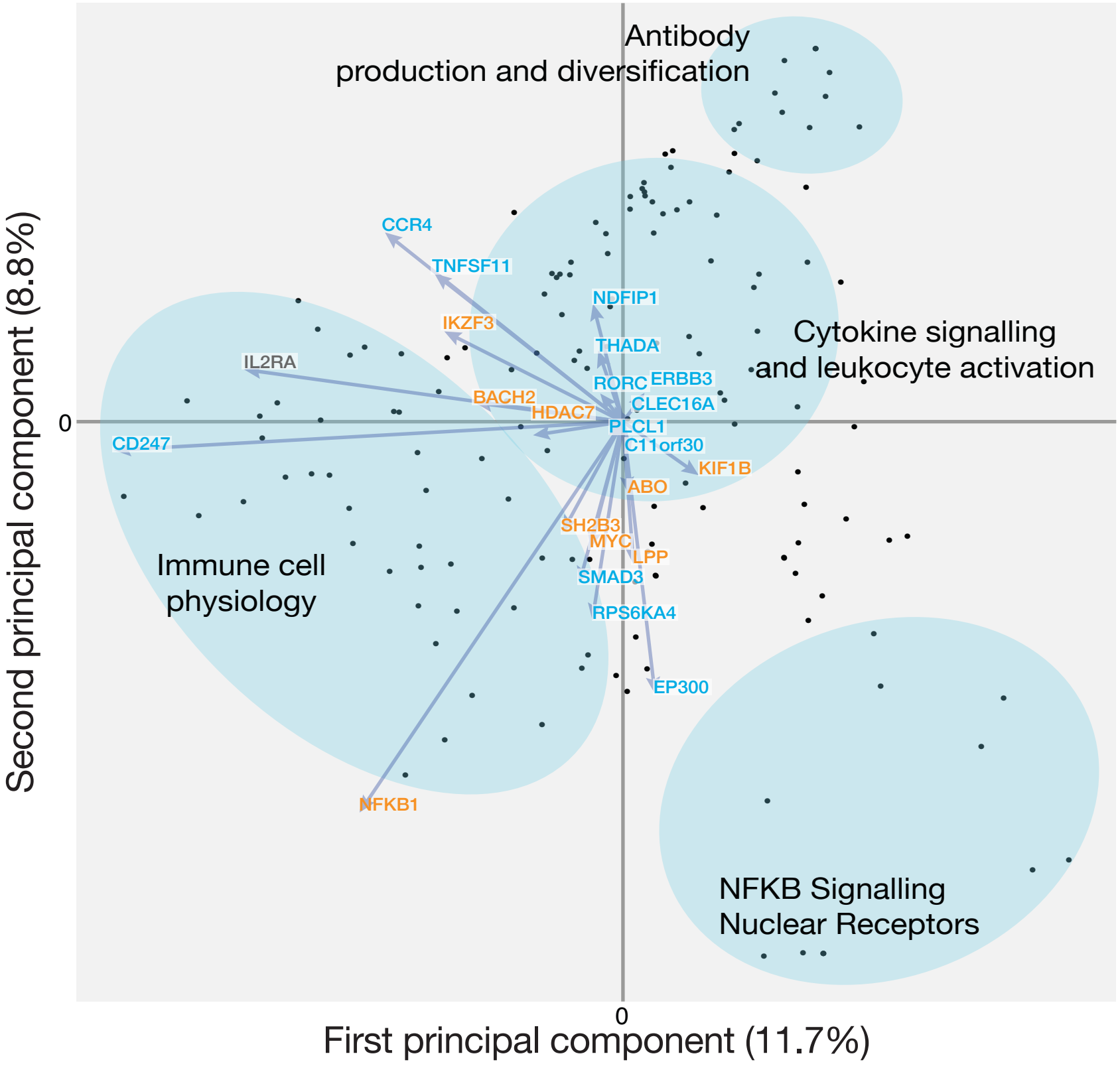
33. Ferreira MAR, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. J Allergy Clin Immunol. 2014 Jun;133(6):1564–71.
34. Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, et al. Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 2011 Aug;7(8):e1002254
35. Sirota M, Schaub MA, Batzoglou S, Robinson WH, Butte AJ. Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet. 2009 Dec;5(12):e1000792
36. Cotsapas C, Hafler DA. Immune-mediated disease genetics: the shared basis of pathogenesis. Trends Immunol. 2013 Jan;34(1):22-6
37. Maier LM, Howson JMM, Walker N, Spickett GP, Jones RW, Ring SM, et al. Association of IL13 with total IgE: evidence against an inverse association of atopy and diabetes. J Allergy Clin Immunol. 2006 Jun;117(6):1306–13.
38. Li X, Ampleford EJ, Howard TD, Moore WC, Torgerson DG, et al. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. J Allergy Clin Immunol. 2012 Oct;130(4):861-8.e7
39. Bach J-F. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med. 2002 Sep 19;347(12):911–20.
40. Takimoto T, Wakabayashi Y, Sekiya T, Inoue N, Morita R, Ichiyama K, et al. Smad2 and Smad3 are redundantly essential for the TGF-beta-mediated regulation of regulatory T plasticity and Th1 development. J Immunol Baltim Md 1950. 2010 Jul 15;185(2):842–55.
41. Bhatt D, Ghosh S. Regulation of the NF-κB-Mediated Transcription of Inflammatory Genes. Front Immunol. 2014;5:71.
42. Kofler DM, Marson A, Dominguez-Villar M, Xiao S, Kuchroo VK, Hafler DA. Decreased RORC-dependent silencing of prostaglandin receptor EP2 induces autoimmune Th17 cells. J Clin Invest. 2014 Jun 2;124(6):2513–22.
43. Jeffery LE, Burke F, Mura M, Zheng Y, Qureshi OS, Hewison M, et al. 1,25-Dihydroxyvitamin D3 and IL-2 combine to inhibit T cell production of inflammatory cytokines and promote development of regulatory T cells expressing CTLA-4 and FoxP3. J Immunol Baltim Md 1950. 2009 Nov 1;183(9):5458–67.
44. Palmer MT, Lee YK, Maynard CL, Oliver JR, Bikle DD, Jetten AM, et al. Lineage-specific effects of 1,25-dihydroxyvitamin D(3) on the development of effector CD4 T cells. J Biol Chem. 2011 Jan 14;286(2):997–1004.
45. Liao W, Lin J-X, Wang L, Li P, Leonard WJ. Modulation of cytokine receptors by IL-2 broadly regulates differentiation into helper T cell lineages. Nat Immunol. 2011 Jun;12(6):551–9.
46. Ananieva O, Darragh J, Johansen C, Carr JM, McIlrath J, Park JM, et al. The kinases MSK1 and MSK2 act as negative regulators of Toll-like receptor signaling. Nat Immunol. 2008 Sep;9(9):1028–36.
47. Fujita N, Jaye DL, Geigerman C, Akyildiz A, Mooney MR, Boss JM, et al. MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. Cell. 2004 Oct 1;119(1):75–86.
48. Chen R, Bélanger S, Frederick MA, Li B, Johnston RJ, Xiao N, et al. In vivo RNA interference screens identify regulators of antiviral CD4(+) and CD8(+) T cell differentiation. Immunity. 2014 Aug 21;41(2):325–38.
49. Du W, Wang D, Shen Y-W, Lee W-H, Riggs A, Liu C-P. WRNIP1 is a key regulator of antigen-specific Treg cells. (IRC4P.480). J Immunol. 2014 May 1;192(1 Supplement):60.7–60.7.

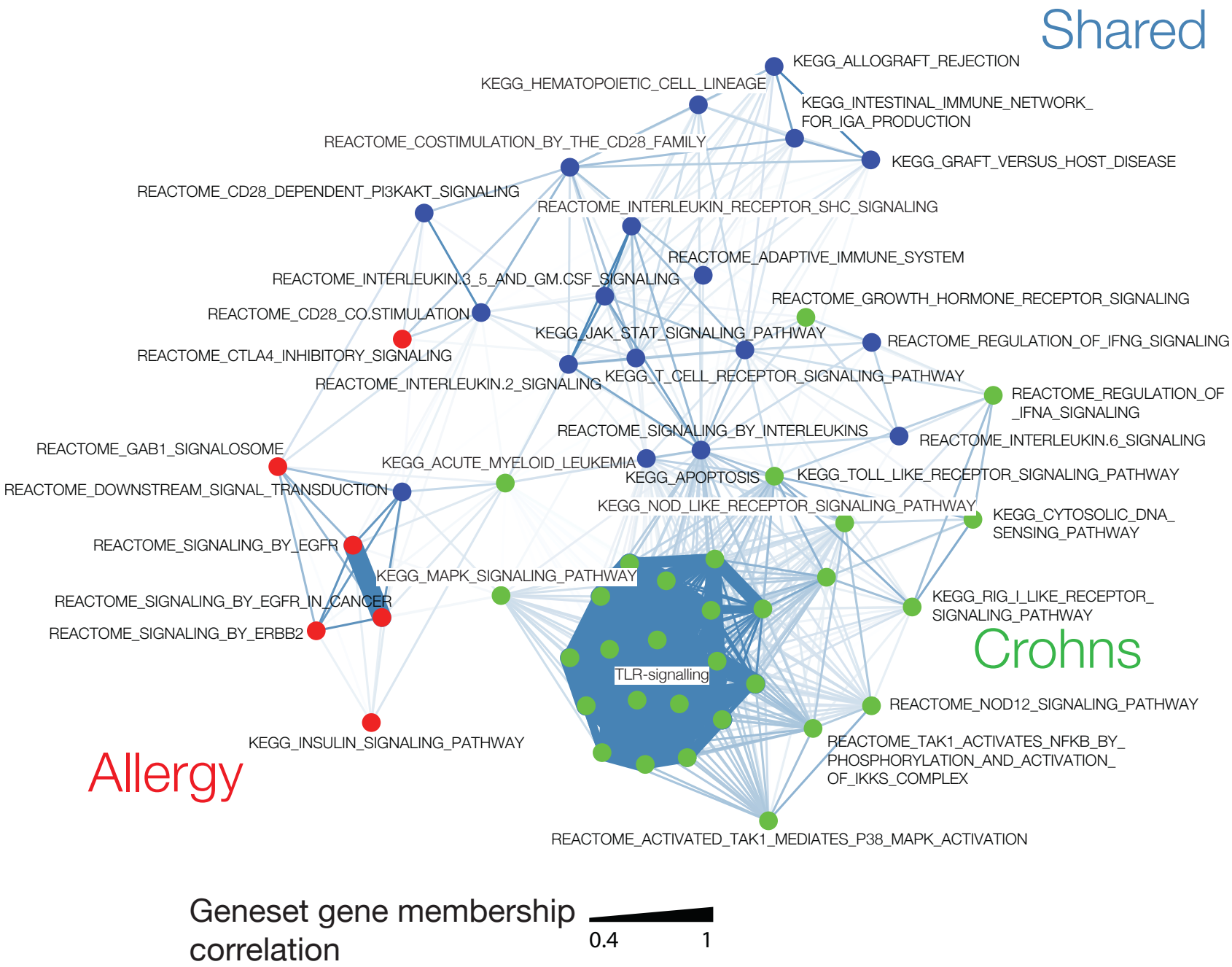
50. Yoshida T, Ng SY-M, Georgopoulos K. Awakening lineage potential by Ikaros-mediated transcriptional priming. Curr Opin Immunol. 2010 Apr;22(2):154–60.
51. Koch S, Reppert S, Finotto S. NFATc1 deletion in T lymphocytes inhibits the allergic trait in a murine model of asthma. Clin Exp Allergy J Br Soc Allergy Clin Immunol. 2015 Feb 1;
52. Yan H, Krishnan K, Greenlund AC, Gupta S, Lim JT, Schreiber RD, et al. Phosphorylated interferon-alpha receptor 1 subunit (IFNaR1) acts as a docking site for the latent form of the 113 kDa STAT2 protein. EMBO J. 1996 Mar 1;15(5):1064–74.
53. Clark RI, Tan SWS, Péan CB, Roostalu U, Vivancos V, Bronda K, et al. MEF2 is an in vivo immune-metabolic switch. Cell. 2013 Oct 10;155(2):435–47.
54. Pittet LA, Quinton LJ, Yamamoto K, Robson BE, Ferrari JD, Algül H, et al. Earliest innate immune responses require macrophage RelA during pneumococcal pneumonia. Am J Respir Cell Mol Biol. 2011 Sep;45(3):573–81.
55. Terrados G, Finkernagel F, Stielow B, Sadic D, Neubert J, Herdt O, et al. Genome-wide localization and expression profiling establish Sp2 as a sequence-specific transcription factor regulating vitally important genes. Nucleic Acids Res. 2012 Sep;40(16):7844–57.
56. Wang J, Lee S, Teh CE-Y, Bunting K, Ma L, Shannon MF. The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells. Int Immunol. 2009 Mar;21(3):227–35.
57. Elangovan RI, Disanto G, Berlanga-Taylor AJ, Ramagopalan SV, Handunnetthi L. Regulatory genomic regions active in immune cell types explain a large proportion of the genetic risk of multiple sclerosis. J Hum Genet. 2014 Apr;59(4):211–5.
58. Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015 Feb 19;518(7539):337–43.
59. Tak PP, Firestein GS. NF-kappaB: a key role in inflammatory diseases. J Clin Invest. 2001 Jan;107(1):7–11.
60. Chevalier N, Jarrossay D, Ho E, Avery DT, Ma CS, Yu D, et al. CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses. J Immunol Baltim Md 1950. 2011 May 15;186(10):5556–68.
61. Gohlke PR, Williams JC, Vilen BJ, Dillon SR, Tisch R, Matsushima GK. The receptor tyrosine kinase MerTK regulates dendritic cell production of BAFF. Autoimmunity. 2009 Mar;42(3):183–97.
62. Lee S-U, Maeda T. POK/ZBTB proteins: an emerging family of proteins that regulate lymphoid development and function. Immunol Rev. 2012 May;247(1):107–19.
63. Paternoster L, Standl M, Chen C-M, Ramasamy A, Bønnelykke K, Duijts L, et al. Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. Nat Genet. 2011 Dec 25;44(2):187–92.


Arthritis
Celiac Disease
Crohn's Disease
Primary Biliary Cirrosis
Autoimmune disease-associated loci
Observed -log10(P)
Expected -log10(P)
C11orf30
PLCL1
SMAD3
LPP
MYC
SH2B3
HLA-B
IKZF3
Ulcerative Colitis
Graves Disease
T1D

B


Second principal component (6.7%)
First principal component (10.9%)
0
0
Celiac Disease
IBD
Allergy
Vitiligo
Systemic Lupus Erythematosus
Ulcerative Colitis
Type 1 Diabetes
R. Arthritis
Asthma
Multiple Sclerosis
Crohns D.
Primary Biliary Cirrhosis
Psoriasis


Antibody
production and diversification
Cytokine signalling
and leukocyte activation
Immune cell
physiology
NFKB Signalling
Nuclear Receptors
CCR4
TNFSF11
NDFIP1
IKZF3
THADA
IL2RA
RORC
ERBB3
BACH2
HDAC7
CLEC16A
PLCL1
CD247
C11orf30
KIF1B
ABO
SH2B3
MYC
LPP
SMAD3
RPS6KA4
EP300
NFKB1
0
0
Second principal component (8.8%)
First principal component (11.7%)


Shared
KEGG_ALLOGRAFT_REJECTION
KEGG_HEMATOPOIETIC_CELL_LINEAGE
KEGG_INTESTINAL_IMMUNE_NETWORK_
FOR_IGA_PRODUCTION
REACTOME_COSTIMULATION_BY_THE_CD28_FAMILY
KEGG_GRAFT_VERSUS_HOST_DISEASE
REACTOME_CD28_DEPENDENT_PI3KAKT_SIGNALING
REACTOME_INTERLEUKIN_RECEPTOR_SHC_SIGNALING
REACTOME_ADAPTIVE_IMMUNE_SYSTEM
REACTOME_INTERLEUKIN.3_5_AND_GM.CSF_SIGNALING
REACTOME_GROWTH_HORMONE_RECEPTOR_SIGNALING
REACTOME_CD28_CO.STIMULATION
KEGG_JAK_STAT_SIGNALING_PATHWAY
REACTOME_REGULATION_OF_IFNG_SIGNALING
REACTOME_CTLA4_INHIBITORY_SIGNALING
REACTOME_INTERLEUKIN.2_SIGNALING
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY
REACTOME_REGULATION_OF
_IFNA_SIGNALING
REACTOME_SIGNALING_BY_INTERLEUKINS
REACTOME_INTERLEUKIN.6_SIGNALING
REACTOME_GAB1_SIGNALOSOME
KEGG_ACUTE_MYELOID_LEUKEMIA
KEGG_APOPTOSIS
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
REACTOME_DOWNSTREAM_SIGNAL_TRANSDUCTION
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_CYTOSOLIC_DNA_
SENSING_PATHWAY
REACTOME_SIGNALING_BY_EGFR
KEGG_MAPK_SIGNALING_PATHWAY
REACTOME_SIGNALING_BY_EGFR_IN_CANCER
KEGG_RIG_I_LIKE_RECEPTOR_
SIGNALING_PATHWAY
REACTOME_SIGNALING_BY_ERBB2
Crohns
TLR-signalling
REACTOME_NOD12_SIGNALING_PATHWAY
KEGG_INSULIN_SIGNALING_PATHWAY
REACTOME_TAK1_ACTIVATES_NFKB_BY_
PHOSPHORYLATION_AND_ACTIVATION_
OF_IKKS_COMPLEX
Allergy
REACTOME_ACTIVATED_TAK1_MEDIATES_P38_MAPK_ACTIVATION
Geneset gene membership
correlation
0.4
1