Database, 2018, 1–7
doi: 10.1093/database/bay010
Original article

Original article

AllerGAtlas 1.0: a human allergy-related genes database

Jinying Liu[1,†], Yuan Liu[2,†], Dan Wang[2,†], Mengqi He[2], Lihong Diao[2], Zhongyang Liu[2], Yang Li[2], Li Tang[2], Fuchu He[2], Dong Li[2,*] and Shuzhen Guo[1,*]

[1]School of Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China and [2]State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences-Beijing (PHOENIX Center), Beijing Institute of Lifeomics, Beijing, China

*Corresponding author: Tel: +86 10 61777057; Fax: +86 10 61777057; Email: lidong.bprc@foxmail.com

Correspondence may also be addressed to Shuzhen Guo. Tel: 86-10-6428-6153; Fax: 86-10-6428-6871; Email: guoshz@bucm.edu.cn

[†]These authors contributed equally to this work.

Citation details: Liu,J., Liu,Y., Wang,D. *et al.* AllerGAtlas 1.0: a human allergy-related genes database. *Database* (2018) Vol. 2018: article ID bay010; doi:10.1093/database/bay010

Received 9 August 2017; Revised 13 December 2017; Accepted 15 January 2018

Abstract

Allergy is a detrimental hypersensitive response to innocuous environmental antigen, which is caused by the effect of interaction between environmental factors and multiple genetic pre-disposition. In the past decades, hundreds of allergy-related genes have been identified to illustrate the epidemiology and pathogenesis of allergic diseases, which are associated with better endophenotype, novel biomarkers, early-life risk factors and individual differences in treatment responses. However, the information of all these allergy-related genes is dispersed in thousands of publications. Here, we present a manually curated human allergy-related gene database of AllerGAtlas, which contained 1195 well-annotated human allergy-related genes, determined by text-mining and manual curation. AllerGAtlas will be a valuable bioinformatics resource to search human allergy-related genes and explore their functions in allergy for experimental research.

Database URL: http://biokb.ncpsb.org/AlleRGatlas/

Introduction

Allergy is a detrimental immunological hypersensitive response to innocuous environmental antigen (1). By definition, allergy includes in form of various heterogeneous conditions such as anaphylaxis, allergic asthma, atopic dermatitis (AD) and the colorful spectrum of food- or drug–induced hypersensitivity reactions (2, 3). Allergy is characterized as the new epidemic of the 21st century due to the continuous rise in the prevalence and severity (4, 5). According to *The World Allergy Organization White Book on Allergy*, up to 40% of the population has suffered from one or more type of allergy with significant associated medical and financial burden (6). From 1992 to 2012, there was a 615% increase in the rate of hospital anaphylaxis admission in the UK, with an estimated annual cost of €900 million (7, 8).

Allergies are clinically and genetically heterogeneous diseases with a variable clinical course and with important divergences in the response to therapy, which can lead to significant challenges for the correct diagnosis and proper treatment. Although allergies can be classified into distinct disorder categories, they show clinical overlap and share common genetic risk factors. For example, allergic asthma or AD, is unlikely to be a single disease but rather a series of overlapping individual clinical features or phenotypes with unique genetic and environmental contributors (9, 10). There is a growing consensus that allergy is caused by the effect of interaction between environmental factors and multiple genetic pre-disposition (11, 12). Therefore, it is important to understand how associated genetic and environmental factors increase the complexity of allergic disease.

Hundreds of allergy-related genes have been discovered, which are involved in better endophenotype, identification of at-risk individuals in early life, novel biomarkers and individual treatment responses. For example, FLG (filament aggregating protein) has already been proposed as a robust screening biomarker for early-onset severe AD, which also has highlighted the importance of epidermal barrier dysfunction in the development of allergic diseases (13, 14). The gene polymorphisms of cluster of differentiation 14, toll-like receptors, Glutathione S-transferase Mu 1 and Glutathione S-Transferase Pi 1 (15) have shed light on the importance of gene–environment interaction for allergic diseases (16–18). Serum levels of Interleukin 16, Interleukin 31, eosinophil cationic protein and High mobility group box 1 can reflect and stratify diseases severity and have been used as reliable markers (19–21). Studies of associations between variants in Adrenoceptor Beta 2, Arachidonate 5-lipoxygenase, IL-4 Receptor Subunit Alpha

and response to related pharmaceuticals are crucial for precision medicine (22–24). Recently, large numbers of genome wide association studies aiming to explore genetic susceptibility have accelerated the search for novel and interesting genes for human allergy-related disease. For example, Bønnelykke *et al.* (25) increased the number of susceptibility loci from 3 to 10 with replication in 6114 case individuals and 9920 controls by the first large-scale genome-wide association study, including single nucleotide polymorphisms (SNPs) in or near TLR6, C11orf30, STAT6, SLC25A46, HLA-DQB1, IL1RL1, LPP, MYC, IL2 and HLA-B. In a large-scale GWAS on 11 025 AD cases and 40 398 controls, Paternoster *et al.* (26) identified and replicated two novel SNPs related to epidermal proliferation and differentiation (rs479844 upstream of OVOL1 and rs2164983 near ACTL9). Ramasamy *et al.* (27) identified three common genetic variants (rs7775228, rs2155219, rs17513503) associated with prevalent allergic rhinitis and grass sensitization, by using approximately 2.2 million genotyped or imputed SNPs in four large European adult cohorts for allergic rhinitis (3933 self-reported cases vs 8965 control subjects) and grass sensitization (2315 cases vs 10 032 control subjects). Studies also identified new genetic susceptibility factors and suggested previously unidentified biological pathways associated with allergic diseases in ethnically different populations, such as rs7701890and rs6010620 in the Chinese sample, rs6010620 in the German sample, as well as rs2243250 and rs2227284 in Pakistani sample (28, 29).

However, the rich scientific information about previously identified allergy-related genes is dispersed in thousands of publications. There is still no database focusing on allergy-related genes so far, which presents the obstacle to understand the genetic architecture of allergic diseases. To address this need, we build the AllerGAtlas database 1.0 (http://biokb.ncpsb.org/AllerGAtlas/) that collects 1195 identified allergy-related genes by the literature-mining and manual curation. AllerGAtlas will lead to new insights into the pathogenesis and epidemiology of individual cases, novel diagnostic and prognostic biomarker, individual treatment responses and precision medicine.

Literature mining and manual curation

To obtain a complete list of allergy-related genes, we performed a comprehensive search for allergy-related literature abstracts in PubMed. Gene-nomenclature recognition and extraction from these abstracts for human allergy-related gene candidates were performed by self-developed ontology-based bio-entity recognizer, which has the precision, recall, F-measure of 0.810, 0.883, 0.845 against the CRAFT corpus for gene/protein recognition based on Protein Ontology (PR) and is on par with current state-of-the-art biomedical annotation systems like BeCAS (30).

A list of human allergy-related genes together with their related diseases and evidence from PubMed abstract was compiled in the following three steps. First, 242 066 sentences in 112 979 PubMed abstracts containing the keywords of ‘allergy,’ ‘allergic,’ ‘anaphylaxis,’ ‘allergic reaction,’ ‘allergic response,’ ‘hypersensitivity,’ ‘atopic,’ ‘atopic’ or their lexical variants were collected. Second, a list of 3150 human genes co-occured with the allergy-related keywords at single-sentence level were recognized and extracted from 42 975 sentences in 27 033 PubMed abstracts by our bio-entity recognizer based on PR. Third, all these 3150 candidates were manually curated by our experts and 1195 genes were finally identified as human allergy-related genes.

The human allergy related disease terms were identified from PubMed abstracts by bio-entity recognizer based on Human Disease Ontology (DO) (31). Associations between allergy-related genes/proteins and human disease terms were obtained based on sentence-level co-occurrence. Furthermore, the biomarker of certain genes/proteins are recognized and marked with the keywords of either 'biomarker' or 'marker' or their lexical variants like 'biomarkers,' 'markers,' 'mark' and 'biochemical markers.'

Genes annotation

With the purpose to facilitate deeper interpretations of the relations with allergy, we provided detailed annotations for each gene. NCBI Entrez Gene ID or gene symbol were used as the central ID for cross-linking and annotation. The basic gene annotation files were downloaded from NCBI FTP site. The gene_info and gene2refseq files were parsed to extract the basic gene information such as gene symbol, synonyms, full name, genetic location, gene type, chromosome, chromosomal location and reference sequence information. The gene ontology (GO) annotations for each gene were obtained from the GOA database (32) and the gene-pathway mapping relations were downloaded from Reactome database (33). The public database dbSNP (34) was utilized to map SNPs to genes by the literature's PMIDs (PubMed Unique Identifier). Public databases including Ensembl (35), Entrez gene (36), UniProt (37), neXtProt (38), Anti-bodypedia (39) also were used for mapping and annotating.

Database implement and service status

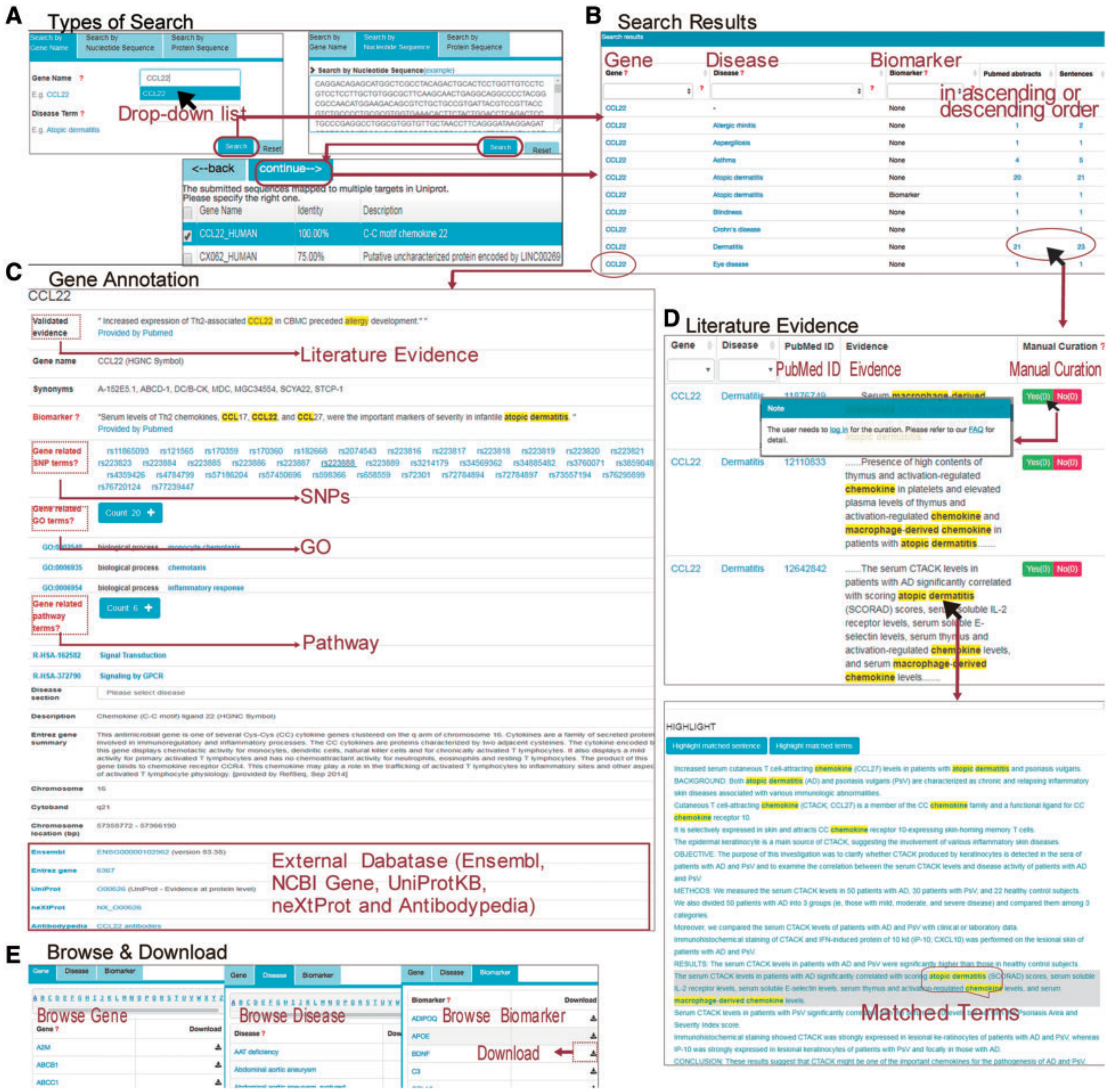
All extracted allergy-related genes/proteins, human disease terms as well as their biomarker roles were loaded into a local MySQL database. PHP was used to build the website of AllerGAtlas on a Windows server. The web service is available at http://biokb.ncpsb.org/AllerGAtlas/. All the data of AllerGAtlas is available to all users without login or registration.

Database search and navigation

AllerGAtlas provides a user-friendly web interface that facilitates searching and browsing database (http://biokb.ncpsb.org/aagatlas/), which comprises five sections including 'Home,' 'Browse & Download,' 'Feedback,' 'FAQ' and 'Contact' (Figure 1). In the page of 'Home,' users can search AllerGAtlas 1.0 database by three types of queries: protein name, nucleotide sequence and protein sequence. For the gene name query, the user can input a gene name in the search box of 'Gene Symbol,' and a drop down list with auto-completed gene symbols will be present in the AllerGAtlas. After selecting one of them and clicking the 'Search' button, the search engine will run and return the queried results containing the queried gene associated human disease terms and supporting literature evidence. If you search the gene by nucleotide sequence or protein sequence, the sequence identity score from BLAST will be listed. Users can specify the matched gene symbol and click 'continue' for result page (Figure 1A). On the result page, a table containing the queried gene, associated human disease terms and number of the supporting abstracts or sentences is displayed (Figure 1B). The hyperlink of the queried gene can lead to the gene info page with plenty of information, including a list of SNP terms from dbSNP, GO terms from GOA, pathway terms from Reactome, the protein description from UniProtKB, the gene expression info from the Expression Atlas, the protein expression info from Human Protein Atlas and the related disease info from Human Disease Ontology (Figure 1C). The hyperlink of number of the evidence abstracts or sentences can lead to a table containing the gene, the disease, the PubMed ID, the evidence sentence and the manual validation information. In addition, the hyperlink of an individual interested evidence sentence leads to the whole abstract with the supporting sentence and various types of extracted entity terms highlighted, i.e. gene name and disease terms (Figure 1D). Our website supports three different approaches for browsing by the page of 'Browse & Download.' All the information can be downloaded (Figure 1E).

Application case of the database

AD is the most common skin disease, affecting up to 30% of children and 3% of adults worldwide (40). We searched our database with the disease name of 'AD' and found a list of 538 expert curated genes with detailed annotations. To obtain more detailed functional annotation of each gene, we can click the hyperlink of the related genes which can lead to the gene info page with plenty of information, such as gene related SNP, gene related pathway and gene related GO terms. Further, we can perform the functional and pathway analyses on the list of AD related genes. Analysis with Reactome (http://www.reactome.org/) reveals that these genes tend to participate in the pathways of immune system, signal transduction, gene expression (transcription), metabolism of proteins, developmental biology, hemostasis, cell–cell communication, extracellular



Gene Name ? CCL22
E.g. CCL22
Disease Term ?
E.g. Atopic dermatitis

Search by Nucleotide Sequence(example)

```
CAGGACAGAGCATGGCTCGCCTACAGACTGCACTCCTGGTTGTCCTC
GTCCTCCTTGCTGTGGCGCTTCAAGCAACTGAGGCAGGCCCCTACGG
CGCCAACATGGAAGACAGCGTCTGCTGCCGTGATTACGTCCGTTACC
GTCTGCCCCTGCGCGTGGTGAAACACTTCTACTGGACCTCAGACTCC
TGCCCGAGGCCTGGCGTGGTGTTGCTAACCTTCAGGGATAAGGAGAT
```

The submitted sequences mapped to multiple targets in Uniprot.
Please specify the right one.

	Gene Name	Identity	Description
☑	CCL22_HUMAN	100.00%	C-C motif chemokine 22
□	CX062_HUMAN	75.00%	Putative uncharacterized protein encoded by LINC00269

Search results

Gene ?	Disease ?	Biomarker ?	Pubmed abstracts	Sentences
CCL22	-	None		
CCL22	Allergic rhinitis	None	1	2
CCL22	Aspergillosis	None	1	1
CCL22	Asthma	None	4	5
CCL22	Atopic dermatitis	None	20	21
CCL22	Atopic dermatitis	Biomarker	1	1
CCL22	Blindness	None	1	1
CCL22	Crohn's disease	None	1	1
CCL22	Dermatitis	None	21	23
CCL22	Eye disease	None	1	1

CCL22

Validated evidence	" Increased expression of Th2-associated CCL22 in CBMC preceded allergy development." " Provided by Pubmed
Gene name	CCL22 (HGNC Symbol)
Synonyms	A-152E5.1, ABCD-1, DC/B-CK, MDC, MGC34554, SCYA22, STCP-1
Biomarker ?	"Serum levels of Th2 chemokines, CCL17, CCL22, and CCL27, were the important markers of severity in infantile atopic dermatitis. " Provided by Pubmed
Gene related SNP terms?	rs11865093 rs121565 rs170359 rs170360 rs182668 rs2074543 rs223816 rs223817 rs223818 rs223819 rs223820 rs223821 rs223823 rs223884 rs223885 rs223886 rs223887 rs223888 rs223889 rs3214179 rs34569362 rs34885482 rs3760071 rs3859048 rs4359426 rs4784799 rs57186204 rs57450696 rs598366 rs658559 rs72301 rs72784894 rs72784897 rs73557194 rs76295899 rs76720124 rs77239447
Gene related GO terms?	Count 20 +
GO:0002548	biological process monocyte chemotaxis
GO:0006935	biological process chemotaxis
GO:0006954	biological process inflammatory response
Gene related pathway terms?	Count 6 +
R-HSA-162582	Signal Transduction
R-HSA-372790	Signaling by GPCR
Disease section	Please select disease
Description	Chemokine (C-C motif) ligand 22 (HGNC Symbol)
Entrez gene summary	This antimicrobial gene is one of several Cys-Cys (CC) cytokine genes clustered on the q arm of chromosome 16. Cytokines are a family of secreted protein involved in immunoregulatory and inflammatory processes. The CC cytokines are proteins characterized by two adjacent cysteines. The cytokine encoded b this gene displays chemotactic activity for monocytes, dendritic cells, natural killer cells and for chronically activated T lymphocytes. It also displays a mild activity for primary activated T lymphocytes and has no chemoattractant activity for neutrophils, eosinophils and resting T lymphocytes. The product of this gene binds to chemokine receptor CCR4. This chemokine may play a role in the trafficking of activated T lymphocytes to inflammatory sites and other aspec of activated T lymphocyte physiology. [provided by RefSeq, Sep 2014]
Chromosome	16
Cytoband	q21
Chromosome location (bp)	57358772 - 57366190
Ensembl	ENSG00000102962 (version 83.38)
Entrez gene	6367
UniProt	O00626 (UniProt - Evidence at protein level)
neXtProt	NX_O00626
Antibodypedia	CCL22 antibodies

Gene	Disease	PubMed ID	Evidence	Manual Curation ?
CCL22	Dermatitis	11876749	Serum macrophage-derived chemokine ... atopic dermatitis......	Yes(0) No(0)
CCL22	Dermatitis	12110833	......Presence of high contents of thymus and activation-regulated chemokine in platelets and elevated plasma levels of thymus and activation-regulated chemokine and macrophage-derived chemokine in patients with atopic dermatitis.......	Yes(0) No(0)
CCL22	Dermatitis	12642842	......The serum CTACK levels in patients with AD significantly correlated with scoring atopic dermatitis (SCORAD) scores, serum soluble IL-2 receptor levels, serum soluble E-selectin levels, serum thymus and activation-regulated chemokine levels, and serum macrophage-derived chemokine levels.......	Yes(0) No(0)

Note
The user needs to log in for the curation. Please refer to our FAQ for detail.

HIGHLIGHT
Highlight matched sentence | Highlight matched terms

Increased serum cutaneous T cell-attracting chemokine (CCL27) levels in patients with atopic dermatitis and psoriasis vulgaris.
BACKGROUND: Both atopic dermatitis (AD) and psoriasis vulgaris (PsV) are characterized as chronic and relapsing inflammatory skin diseases associated with various immunologic abnormalities.
Cutaneous T cell-attracting chemokine (CTACK; CCL27) is a member of the CC chemokine family and a functional ligand for CC chemokine receptor 10.
It is selectively expressed in skin and attracts CC chemokine receptor 10-expressing skin-homing memory T cells.
The epidermal keratinocyte is a main source of CTACK, suggesting the involvement of various inflammatory skin diseases.
OBJECTIVE: The purpose of this investigation was to clarify whether CTACK produced by keratinocytes is detected in the sera of patients with AD and PsV and to examine the correlation between the serum CTACK levels and disease activity of patients with AD and PsV.
METHODS: We measured the serum CTACK levels in 50 patients with AD, 30 patients with PsV, and 22 healthy control subjects.
We also divided 50 patients with AD into 3 groups (ie, those with mild, moderate, and severe disease) and compared them among 3 categories.
Moreover, we compared the serum CTACK levels of patients with AD and PsV with clinical or laboratory data.
Immunohistochemical staining of CTACK and IFN-induced protein of 10 kd (IP-10; CXCL10) was performed on the lesional skin of patients with AD and PsV.
RESULTS: The serum CTACK levels in patients with AD and PsV were significantly higher than those in healthy control subjects.
The serum CTACK levels in patients with AD significantly correlated with scoring atopic dermatitis (SCORAD) scores, serum soluble IL-2 receptor levels, serum soluble E-selectin levels, serum thymus and activation-regulated chemokine levels, and serum macrophage-derived chemokine levels.
Serum CTACK levels in patients with PsV significantly corre... ...levels ... Psoriasis Area and Severity Index score.
Immunohistochemical staining showed CTACK was strongly expressed in lesional ke-ratinocytes of patients with AD and PsV, whereas IP-10 was strongly expressed in lesional keratinocytes of patients with PsV and focally in those with AD.
CONCLUSION: These results suggest that CTACK might be one of the important chemokines for the pathogenesis of AD and PsV.

Gene | Disease | Biomarker
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Gene ?	Download
A2M	⭳
ABCB1	⭳
ABCC1	⭳

Disease ?	Download
AAT deficiency	
Abdominal aortic aneurysm	
Abdominal aortic aneurysm, ruptured	

Biomarker ?	Download
ADIPOQ	⭳
APOE	⭳
BDNF	⭳
C3	⭳

Figure 1. (**A**) The page of 'Home' supports three types of queries: search by gene name, search by nucleotide or protein sequence. Users can submit the gene name of 'CCL22' in the search box. Users can also search the gene by sequence, and the sequence identity score from BLAST will be listed. Users can specify the matched gene symbol and click 'continue' for result page. (**B**) On the result page, a table containing the queried gene associated human disease terms and supporting evidence is displayed. (**C**) After clicking the gene name of 'CCL22' in the page of 'result,' users can see detailed annotation about this gene and cross references to external databases. (**D**) After clicking the number of the evidence abstracts or sentences in the page of 'result,' users can see a table containing the gene, associated disease term, the PubMed ID, the evidence sentence and the manual validation information. Users can also click on the link of evidence in this page to see the original abstract highlighted with the key words. (**E**) The page of 'Browse & Download' presents three different approaches for browsing. All the information can be downloaded.

matrix organization, cellular responses to external stimuli and programmed cell death (Figure 2A). The protein class analysis with Panther (41) (http://pantherdb.org/) indicates that these genes tend to be with the function of signaling molecule, receptor, defense/immunity protein, nucleic acid binding, hydrolase, transcription factor, cell adhesion molecule, transferase, etc. (Figure 2B). This result suggests that the interaction between innate/adaptive immune responses and skin epithelial function play a major role in the development of AD. The above speculation can be validated by the literature (42).

Discussion

The comprehensive collection of AllerGAtlas 1.0 database allows us to have an overview of human

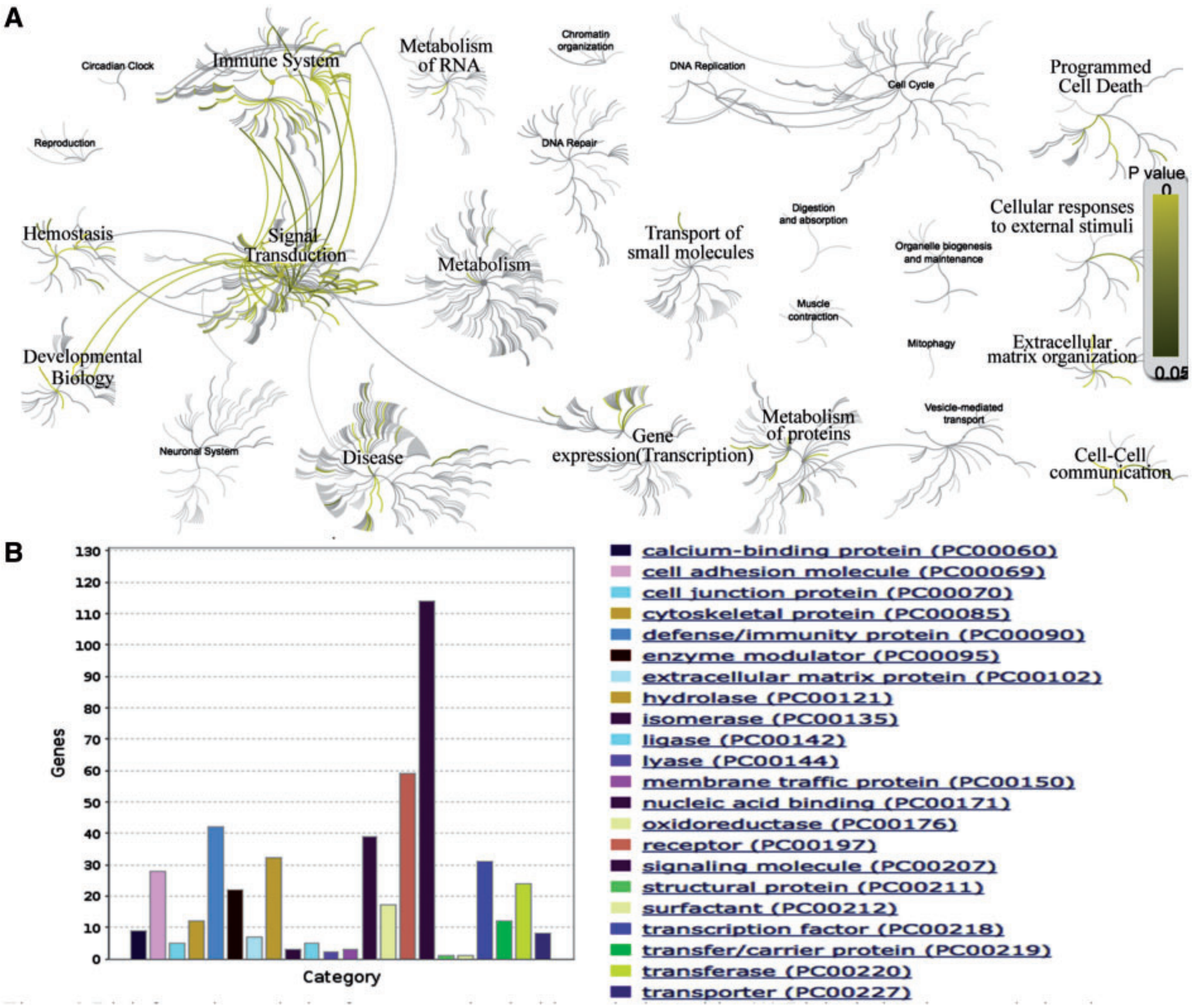


Figure 2. Bioinformatics analysis of genes associated with AD. (**A**) Biological pathway analysis using Reactome (http://www.reactome.org/). (**B**) Protein class analysis using PANTHER (http://pantherdb.org/).

allergy-related genes and their related diseases. The analysis with Reactome reveals that these genes participate in the immune system, signal transduction, transcription, metabolism of proteins, developmental biology, hemostasis, vesicle-mediated transport, cell–cell communication, extracellular matrix organization, cellular responses to external stimuli, programmed cell death, transport of small molecules, metabolism of RNA and circadian clock (Figure 3A). The GO analysis with Panther reveals that the largest class of allergy-related gene proteins is signaling molecule, followed by receptor, nucleic acid binding, hydrolase, defense/immunity protein, transcription factors, transferase, enzyme modulator, cell adhesion molecule, etc (Figure 3B). All these results show the importance and the value of our collection, as well as demonstrate that AllerGAtlas 1.0 database will greatly facilitate allergist to explore the pathogenesis of allergies.

The web service of AllerGAtlas 1.0 also supports the function of community curation. All logged in users can provide their feedback by simply clicking the 'Yes' or 'No' button to confirm or reject the evidence phrases. Our database will be updated periodically according to these feedbacks.

In summary, AllerGAtlas 1.0 is the first attempt to provide a comprehensive non-redundant catalog of allergy-related genes along with supporting evidence from published literature. The availability and use of AllerGAtlas 1.0 will be expected to be a unique value-added resource, which can help scientists and clinicians to search the literature on allergy-related genes and their involvement in diseases.

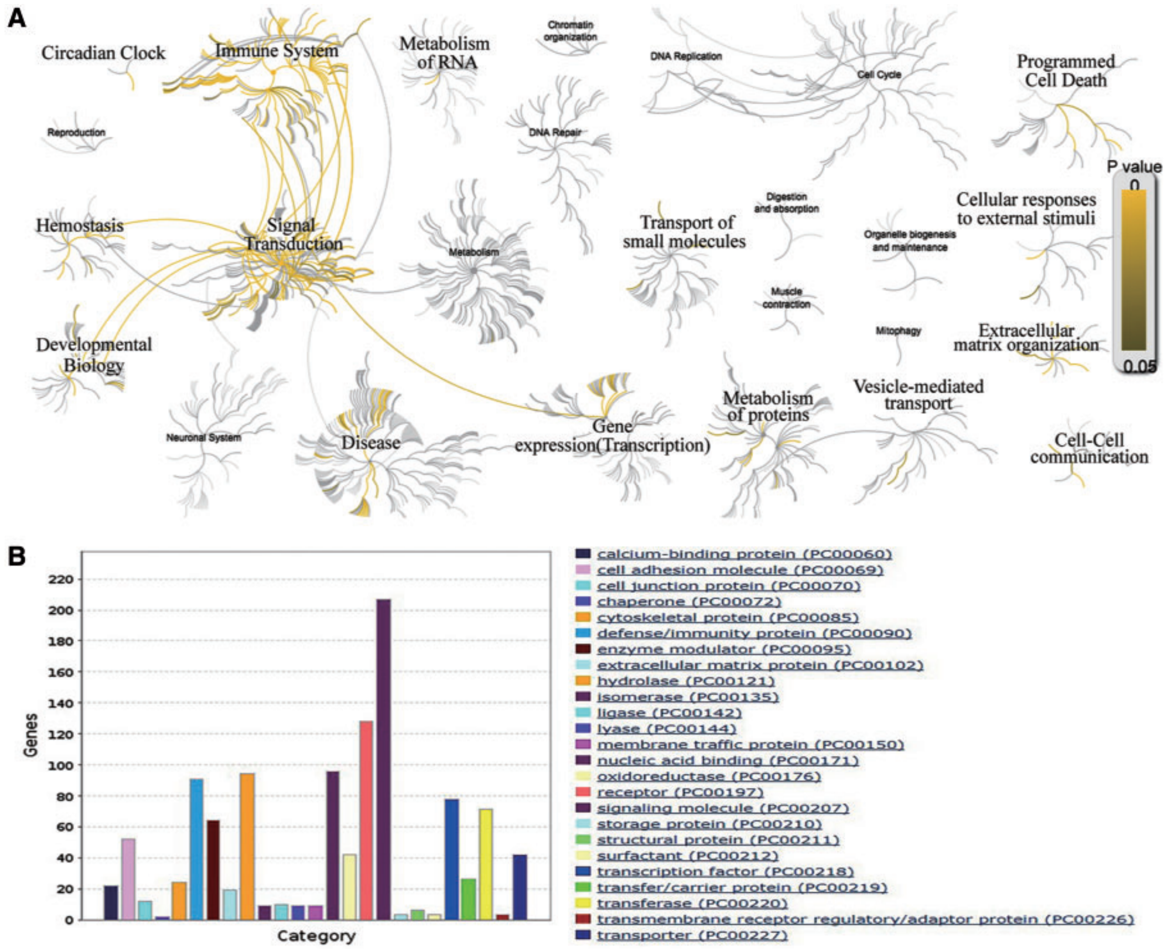


Figure 3. Bioinformatics analysis of the human allergy-related genes. (**A**) Biological pathway analysis using Reactome. (**B**) Protein class analysis using PANTHER.

Funding

This work was supported by the Program of the National Key Research and Development Program (2017YFC1700105), Program of International S&T Cooperation (2014DFB30020), Innovation Project (16CXZ027), Beijing Nova Program (xx2014009), National Natural Science Foundation of China (31601064) and the Beijing Nova Program (Z171100001117117), Program of Precision Medicine (2016YFC0901905).

Conflict of interest. None declared.

References

1. Akdis,C.A., Agache,I. (eds.) (2014) *Global Atlas of allergy*. European Academy of Allergy and Clinical Immunology, Munich.
2. Ring,J.T., Plattsmills (2005) Allergy in Practice. *Journal of the American Medical Association*. **133**, 864–865.
3. Demoly,P., Tanno,L.K., Akdis,C.A. *et al*. (2014) Global classification and coding of hypersensitivity diseases–An EAACI–WAO survey, strategic paper and review. *Allergy*, **69**, 559–570.
4. Lai,C., Beasley,R., Crane,J. *et al*. (2009) Global variation in the prevalence and severity of asthma symptoms: phase three of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax*, **64**, 476–483.
5. Shaw,R.A., Crane,J., O'Donnell,T.V. *et al*. (1990) Increasing asthma prevalence in a rural New Zealand adolescent population: 1975-89. *Arch. Dis. Childhood*, **65**, 1319–1323.
6. Pawankar,R., Canonica,G.W., St Holgate,S.T. *et al*. (Update. 2013) *The WAO White Book on Allergy*. World Allergy Organization, Wisconsin.
7. Turner,P.J., Gowland,M.H., Sharma,V. *et al*. (2015) Increase in anaphylaxis-related hospitalizations but no increase in fatalities: an analysis of United Kingdom national anaphylaxis data, 1992-2012. *J. Allergy Clin. Immunol*., **135**, 956–963.
8. Holgate,S.T. and Ewan,P. (2003) Allergy: the unmet need. A blueprint for better patient care. In: *Royal College of Physicians*. Lavenham Press Ltd., Sudbury, Suffolk, Great Britain.
9. Anderson,G.P. (2008) Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet*, **372**, 1107–1119.
10. Meng,J.F. and Rosenwasser,L.J. (2010) Unraveling the genetic basis of asthma and allergic diseases. *Allergy Asthma Immunol. Res*., **2**, 215–227.
11. Ober,C. and Yao,T.C. (2011) The genetics of asthma and allergic disease: a 21st century perspective. *Immunol. Rev*., **242**, 10–30.

12. Cookson,W. (1999) The alliance of genes and environment in asthma and allergy. *Nature*, **402**, 5–11.
13. Palmer,C.N., Irvine,A.D., Terron-Kwiatkowski,A. *et al.* (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat. Genet.*, **38**, 441–446.
14. van den Oord,R.A. and Sheikh,A. (2009) Filaggrin gene defects and risk of developing allergic sensitisation and allergic disorders: systematic review and meta-analysis. *BMJ*, **339**, b2433.
15. Gilliland,F.D., Li,Y.F., Saxon,A. *et al.* (2004) Effect of Glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement of allergic responses: randomised, placebo-controlled crossover study. *Lancet*, **363**, 119–125.
16. Lau,M., Dharmage,S.C., Burgess,J.A. *et al.* (2014) CD14 polymorphisms, microbial exposure and allergic diseases: a systematic review of gene–environment interactions. *Allergy*, **69**, 1440–1453.
17. Smit,L.A., Siroux,V., Bouzigon,E. *et al.* (2009) CD14 and toll-like receptor gene polymorphisms, country living, and asthma in adults. *Am. J. Respiratory Crit. Care Med.*, **179**, 363–368.
18. Smit,L.A., Bongers,S.I., Ruven,H.J. *et al.* (2007) Atopy and new-onset asthma in young Danish farmers and CD14, TLR2, and TLR4 genetic polymorphisms: a nested case-control study. *Clin. Experimental Allergy*, **37**, 1602–1608.
19. Angelova-Fischer,I., Hipler,U.C., Bauer,A. *et al.* (2006) Significance of interleukin-16, macrophage- derived chemokine, eosinophil cationic protein and soluble E-selectin in reflecting disease activity of atopic dermatitis–from laboratory parameters to clinical scores. *Br. J. Dermatol.*, **154**, 1112–1117.
20. Ezzat,M.H., Hasan,Z.E. and Shaheen,K.Y. (2011) Serum measurement of interleukin-31 (IL-31) in paediatric atopic dermatitis: elevated levels correlate with severity scoring. *J. Eur. Acad. Dermatol. Venereol.*, **25**, 334–339.
21. Cuppari,C., Manti,S., Chirico,V. *et al.* (2015) Sputum high mobility group box-1 in asthmatic children: a noninvasive sensitive biomarker reflecting disease status. *Ann. Allergy Asthma Immunol.*, **115**, 103–107.
22. Ortega,V.E., Hawkins,G.A., Moore,W.C. *et al.* (2014) Effect of rare variants in ADRB2 on risk of severe exacerbations and symptom control during longacting *β* agonist treatment in a multiethnic asthma population: a genetic study. *Lancet Respiratory Med.*, **2**, 204–213.
23. Drazen,J.M., Yandava,C.N., Dubé,L. *et al.* (1999) Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. *Nat. Genet.*, **22**, 168–170.
24. Slager,R.E., Otulana,B.A., Hawkins,G.A. *et al.* (2012) IL-4 receptor polymorphisms predict reduction in asthma exacerbations during response to an anti–IL-4 receptor α antagonist. *J. Allergy Clin. Immunol.*, **130**, 516–522.
25. Bønnelykke,K., Matheson,M.C., Pers,T.H. *et al.* (2013) Meta-analysis of genome-wide association studies identifies ten loci influencing allergic sensitization. *Nat. Genet.*, **45**, 902–906.
26. Paternoster,L., Standl,M., Chen,C.M. *et al.* (2012) Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. *Nat. Genet.*, **44**, 187–192.
27. Ramasamy,A., Curjuric,I., Coin,L.J. *et al.* (2011) A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *J. Allergy Clin. Immunol.*, **128**, 996–1005.
28. Sun,L.-D., Xiao,F.-L., Li,Y. *et al.* (2011) Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. *Nat. Genet.*, **43**, 690–694.
29. Micheal,S., Minhas,K., Ishaque,M. *et al.* (2013) IL4 gene polymorphisms and their association with atopic asthma and allergic rhinitis in Pakistani patients. *J. Investig. Allergol. Clin. Immunol.*, **23**, 107–111.
30. Nunes,T., Campos,D., Matos,S. *et al.* (2013) BeCAS: biomedical concept recognition services and visualization. *Bioinformatics*, **29**, 1915–1916.
31. Schriml,L.M., Arze,C., Nadendla,S. *et al.* (2011) Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Res.*, **40**, D940–D946.
32. Camon,E., Magrane,M., Barrell,D. *et al.* (2004) The gene ontology annotation (GOA) database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res.*, **32**, D262–D266.
33. Fabregat,A., Sidiropoulos,K., Garapati,P. *et al.* (2016) The Reactome pathway knowledgebase. *Nucleic Acids Res.*, **44**, D481–D487.
34. Sherry,S.T., Ward,M.H., Kholodov,M. *et al.* (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
35. Flicek,P., Amode,M.R., Barrell,D. *et al.* (2011) Ensembl. *Nucleic Acids Res.*, **39**, D800–D806.
36. Maglott,D., Ostell,J., Pruitt,K.D. *et al.* (2005) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **33(suppl_1)**, D54–D58.
37. Bairoch,A., Apweiler,R., Wu,C.H. *et al.* (2005) The universal protein resource (UniProt). *Nucleic Acids Res.*, **33(suppl_1)**, D154–D159.
38. Lane,L., Argoud-Puy,G., Britan,A. *et al.* (2012) neXtProt: a knowledge platform for human proteins. *Nucleic Acids Res.*, **40**, D76–D83.
39. Björling,E. and Uhlén,M. (2008) Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol. Cell. Proteomics*, **7**, 2028–2037.
40. Weidinger,S. and Novak,N. (2016) Atopic dermatitis. *Lancet*, **387**, 1109–1122.
41. Mi,H., Poudel,S., Muruganujan,A. *et al.* (2016) PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.*, **44**, D336–D342.
42. Bin,L. and Donald,Y.M. (2016) Genetic and epigenetic studies of atopic dermatitis. *Allergy Asthma Clin. Immunol.*, **12**, 52.