

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 He², Lihong Diao², Zhongyang Liu², Yang Li², Li Tang², Fuchu He², Dong Li^{2,*} and Shuzhen Guo^{1,*}

¹School of Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China and ²State 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 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

 $\ensuremath{\mathbb{C}}$ The Author(s) 2018. Published by Oxford University Press.

Page 1 of 7

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

anaphylaxis admission in the UK, with an estimated annual cost of \notin 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 literaturemining 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 allergyrelated 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-ofthe-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, 242066 sentences in 112979 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 allergyrelated keywords at single-sentence level were recognized and extracted from 42975 sentences in 27033 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,' '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

| | Page 4 | ot | 7 | | |
|--|--------|----|---|--|--|
|--|--------|----|---|--|--|

| | | Gene Name Nucleotide Security | Protein Sequence | Gene | D | Disease | Biomar | ker | |
|--|--|--|--|--|---|--|--|---|--|
| Sene Name ? | CCL22 | CAGGACAGAGCATGGCTCG | Incerchanger Inconcasactgeactectogetterecte Inteasgeactgeactectaege Inteasgeactgeatterectaege Internetiese | Gene 7 | • • | Disease ? | Biomarker 7 | in ascendir | |
| Disease Term ? | Drop-down list | atctacccctacacataat | GTCTGCTGCCGTGATTACGTCCGTTACC | COL22 | | | None | descending | gorde |
| E.g. Alopic dermat | | | | COLZE | | Allergic ministe Aspergiliosis | None | 1. | 8. 1 |
| | (Search) Reset | | Brenth Reset | CCL22 | | kathra | None | | 5 |
| | | tinue> | | CCL22 | | Mopio dermatilia | None | 20 | 21 |
| | Please specify the right | | n Uniprot. | CCL22 | | Nopic demetitis | Biomarker | 1 | 1 |
| | Gene Name | Identity Description | | CCL22 | | Sindress. | None | 1 | 1 |
| | CCL22_HUMAN | 100.00% C-C motif ch | emokine 22 | COL22 | | Dohn's disease | None | (2) | 23 |
| | CXD62_HUMAN | 75.00% Putative uncl | haracterized protein encoded by LINC00269 | COLES | | lye disease | None | 1 | K |
| Gene | Annotation | | ţ | | | | | | |
| Validated | * Increased expression of Th2-associated CCL22 in | CBMC preceded allergy developme | ent." * | | Ditte | erature Evide | | | 1 |
| evidence | Provided by Pubmed | | | | Gene | Disease PubMed II | nce | | vual Curatio |
| Gene name | CCL22 (HGNC Symbol) | →Literature E | vidence | | Gene | | | | |
| Synonyms | A-152E5.1, ABCD-1, DC/B-CK, MDC, MGC34554, | | | | | - PubMed | D Eivdence | Manu | al Curati |
| | | | | | CCL22 | Dermatilia 11876749 | Secure macropha | Yes | (0) No(0) |
| Biomarker ? | "Serum levels of Th2 chemokines, CCL17, CCL22, Provided by Pubmed | and CCL27, were the important ma | rkers of severity in infantile atopic dermatiti | s . 7 | | The user needs to | log in for the curation. Please re | ifier to our EAQ for | |
| Gene related | rs11865093 rs121565 rs170359 rs170360 | 0 m182668 m2074543 m223 | 816 m223817 m223818 m223819 | rs223820 rs223821 | - | detail. | Mopie dermatik | - | _ |
| SNP terms? | rs223823 rs223884 rs223885 rs223886 n rs4359426 rs4784799 rs57186204 rs574 | s223887 rs223888 rs223889 | rs3214179 rs34569362 rs34885482 | rs3760071 rs3859048 | CCL22 | Dermatitis 12110833 | Presence of high | | (0) No(0) |
| I | rs4309426 rs4784799 rs57186204 rs574 rs76720124 rs77239447 | | 12301 1572784894 1572784897 1573 | 007134 1576230833 | | | thymus and activation chemokine in platele | | |
| Generalated" | Count 20 💠 | →SNPs | | | | | plasma levels of thym | | |
| GO terms? | Coolin 10 | | | | | | activation-regulated c macrophage-derived | | |
| GO:0 02548 | histogical process monocule chemotaxis | →GO | | | | | patients with atopic of | dermatitis | |
| GO:0006935 | biological process chemotaxis | | | | CCL22 | Dermatitis 12642842 | | | (0) No(0) |
| GO:0006954 | biological process inflammatory response | | | | | | patients with AD sign with scoring atopic d | ermatitis | |
| | | | | | | | (SCORAD) scores st | en soluble IL-2 | |
| Gene related pathway | Count 6 🛨 | | | | | | (SCOPPE) SCORES, SC | | |
| pathway | Count: 6 🔸 | Dethursu | | | | | (SCORAD) scores, se receptor levels, serun selectin levels, serun | n soluble E- | |
| pathway terms? | Count 6 🛨 | →Pathway | | | | | selectin levels, serum activation-regulated | h thynus and hempkine levels, | |
| pathway terms? R3I(SA-162582 | | →Pathway | | | | | selectin levels, serum activation-regulated c and serum macrophy | h thynus and hempkine levels, | |
| | Signal Transduction | →Pathway | | | | | selectin levels, serum activation-regulated | h thynus and hempkine levels, | |
| R-HSA-162582 R-HSA-162582 R-HSA-372790 Disease section | Signal Transduction Signaling by GPCR | →Pathway | | | | | selectin levels, serum activation-regulated c and serum macrophy | h thynus and hempkine levels, | |
| pathway berms? R-JISA-162582 R-JISA-372790 Disease section Description | Signal Transduction Signaling by GPCI Photas Selected chicease Chemotive (C-C motif) Ngand 22 (HGMC Symbol) | | r q arm of chromosome 16. CytoRines are a | family of secreted protein | HIGHLIGHT | | selectin levels, serum activation-regulated and serum macrophi chemokine levels | h thynus and hempkine levels, | |
| pathway berms? R-JISA-162582 R-JISA-372790 Disease section Description | Signal Transduction Signaling by GPCI Photas Selected chicease Chemotive (C-C motif) Ngand 22 (HGMC Symbol) | | r gam of chromosome 16. Cytolines are a tios characterized by two alignet cytolines at the chromosome achieved it ymphicitytes | family of secreted protein The cytokine encoded by also pessays a mid | Highlight mate | ched sentence Highlight matche | selectin levels, serum activation-regulated c and serum macrophi chemokine levels | t thyn us and chempkine levets, age-cenived | |
| pathway berms? R-JISA-162582 R-JISA-372790 Disease section Description | Signal Transduction Signaling by GPCI Photas Selected chicease Chemotive (C-C motif) Ngand 22 (HGMC Symbol) | | rg arm of chromosome 16, Cystolines are a miss characterized by two alligenet cystomes miss characterized a chrome of the producties mission of activated T symphocyses to informeral of activated T symphocyses to informeral | family of secreted protein The cytowne encoded to It also deploys a mid (res. The product of this ony sites and other aspec | Highlight mate | ched sentence Highlight makine num cutaneous T cell-attracting cour | selectin levels, serum activation-regulated e and serum macrophi chemokine levels | t thyn us and shem kine levets, sge-cenved | sasis vulgaris. |
| pathway arms? R.HSA.112502 R.HSA.312709 Disease section Description Entrez gene summary | Signal Transferction Signaling by GPCI Penase select disease Penase select disease Commonities (C-C morth) Igand 22 (HGH4 (Symbol) This antiferentiating approximation of Several Cys-Cys (Symbol) activity for penasy activated for the penase select the temporation activity for penase activities of the penase select the penase select the penase select the penase for the temporative penase select the penase s | | e) anni of circumosone 16. Сурманев але а ms characterized by low allgered cyslower brochs, cosmophies and resing. T implacytes trophis, cosmophies and resing. T implacy and activated T symplecytes to industrial | Tamily of secreted protein The cytokine encoded by allow despite a mid table despite a mid ony sites and other aspec | Highlight mate Increased Ser BACKGROUT skin diseases | thed sentence Highlight matche num cutaneous T cel-attiacting (Net ND Boh atapis demastiss (AD) an a associated with various immunolog | selectin levels, serum activation-regulated à and serum macrophy chamoking levels d wms second (CCL27) levels in patients oponasis vigans (PAV) are cha canomaties. | t thyru us and thermakine levels, age-derived | ng inflammatory |
| pathway serms? R.HSA.162502 R.HSA.372790 Disease section Description Entrez gene summary | Signal Transferction Signaling by GPCI Please select disease Please select disease Chemokine (C-C month tigand 22 (HOMAC Bymbol) Triss antimetrostatis gene to ord deveral Cys-Cyc Q- activation of the select of th | | ng ann of circumenous 16. Crystelines are a 09. enairestration by two adjustent crystelines frontiles. Cosmophiles and resting T symphocytes tropits, cosmophiles and resting T symphocy ing of activated T symphocytes to initiaeread | Tarrety of secreted protein The cytokine enclosed is free. The product of this free. The product of this ony sites and other aspec | Highlight mate Increased Ser BACKGROUT skin diseases | thed sentence Highlight matche num cutaneous T cell-attiacting Cher ND Both atagist dermatitist (AD) an a associated with various immunolog cell-attiacting Chernolitie (CTACH; | selectin levels, serum activation-regulated à and serum macrophy chamoking levels d wms second (CCL27) levels in patients oponasis vigans (PAV) are cha canomaties. | t thyru us and thermakine levels, age-derived | ng inflammatory |
| pathway serms? R.315A.112502 R.315A.312799 Disease section Description Envice gene summary Chromosome Cytoband Chromosome | Signal Transferction Signaling by GPCI Penase select disease Penase select disease Commonities (C-C morth) Igand 22 (HGH4 (Symbol) This antiferentiating approximation of Several Cys-Cys (Symbol) activity for penasy activated for the penase select the temporation activity for penase activities of the penase select the penase select the penase select the penase for the temporative penase select the penase s | | rg arm of christmosome 10. Cytotkines are a and for chromosy, artivates T ymphocyteo typothe, computing and resting T photo- topath, computing and resting T photo- rest, and the second transmitter of the second transmitter of activated T ymphocytes to interview | Tamily of secreted protein The cyclosise encoded to it show depays a mail only sizes and other agree | Heation mate Increased ser BACKGROUP skin diseases Cutaneous T chemoking in it is selective | this sentence Highlight makine man cutaneous T cel-attacting Chen MD Bohn Repair Kennatols (KA) was cel-attacting Chenologies cel-attacting Chenologies (CTACK: esceptor 10 y sepressed in skin and attracts CC | selectin levels, serum activation-equitate d and serum macrophi Chemokine levels | thyrdius and chemplane levels, age-perived | ng inflammatory Liligand for CC |
| pathway serms? R.315A.112502 R.315A.312799 Disease section Description Envice gene summary Chromosome Cytoband Chromosome | Signal Transferction Signaling by GPCI Petanies select clinesase Chemotheme (C-C month) liganed 22 (HGH4G Signated) Trans antiferentineating sparse is non of Serveral Cys-Cys of the select clinesase Chemotheme (C-C month) liganed 22 (HGH4G Signated) Trans antiferentineating activities of Serveral Cys-Cys of the select clinesase (C-C) and | CC) rytoknine genes i kussered on the deales. The CC cylokanes are prote to care and the cylokanes are protection to care care and the cylokanes and the termine may play a role in the traffick termine. The play 2014j | | Sensy of secreted protein The cyclester encoded in the company and the cyclest second other aspect | Highlight mate Increased ser BACKGROUT skin diseases Cutaneous T chemoking in it is selective The epidema OBJECTIVE: | thed sentence. Highlight makine num cutaneous T cel-attracting Electron Do Bon atoget deemators (AD) as a associated with vanisus immunolog cel-attracting Distanciates (CDAC); cooptor 10. If expressed in skin and attracts CD all explano(16 a main source of CDAC); the purpose of this investigation with | selectin levels, serum activation-equitate d and serum macrophi Chemokine levels | thyritus and chemplane levels, age-perived | ng intermatory Ligand for CC eases in the sera of |
| pathway Perms? R-JISA-112502 R-JISA-112502 R-JISA-112502 Basea pasteon Description Description Description Chromosome fortomosome fortomosome fortomosome | Signal Transferction Signating by GPCI Persons select choicease Commonities (C-C) month ligand 222 (HGH4G Symbol Trans and transmission of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the gener displayer seminative activities of Several Cys-Cys of the generative activities of Several Cys of the generative activities of Several Cys of Several Cys of the generative activities of Several Cys of the generative activities of Several Cys of the generative activities of the generative activ | CC) (1)tokano genes chatered on the cesses. The CC cytokanes are prote to, denoting cares, hadrard and real notations may play a cale in the static hereine, slep 2014) External Daba | tase (Ensembl, | Samp of secreted protects The cytaker encoded at the same secret | Highlight mate Increased ser BACKGROUT skin diseases Cutaneous T chemoking in it is selective The epidema OBJECTIVE: | Highlight manne num cutaneous T cel-attacting tilen ND Boh attapti Breasous menunolog cell-attacting Unempilities (CTACK, ceoptor 10 ly expressed in skin and attracts CC is learning to a man source of C | selectin levels, serum activation-equitate d and serum macrophi Chemokine levels | thyritus and chemplane levels, age-perived | ng intermatory Ligand for CC eases in the sera of |
| pathway Jarma ? R. JISA. H2SB2 R. JISA. H2SB2 R. JISA. H2SB2 Disease action Description Extense Summary Chromosome Gytoband Gytoband Chromosome Description Ensembl | Signal Transferction Signaling by GPCI Persons bester of discusse Persons bester of discusse Commonwere (C-C) more) higher 22 (HGH4G Bymolog) This and reference on of Several Cys-Cys (a) for general cys-Cys (b) for general cys (b) for | CO) rytoman genes clustered on the societies. The CC cystames are proto is, denoted cares, hadrard atter cares motioner may play and in the static territory, sep 2014) External Daba ICBI Gene, Ur | tase (Ensembl, niProtKB, | Terminy of Accreted andrese . The criterian encodes and the factor of th | Highlight mate Increased set BACKGROUP skit diseases Cutaneous T BRERKING in R is selective The epidema OBJECTIVE patients with J and PsV, METHODS: V | tend determine "High-byth Histone municataneous" Tech-stratucting Data NIC Bein Bargel demantes (AD) en associated with virtual immunological extensional strategies (CEDAC), ecceptor 10, y expressed in skin and attracts OC all immunological strategies of this investigation with AD and PaV and Detamine the cor We measured the serum CTACK with | selectin levels, serum activation-regulated c and serum macrophis chemoking levels | thyrd us and chemp blank elvels, age-centved | I ligand for CC eases. In the sera of etients with AD trol subjects. |
| pathway Jarma ? R. JISA. H2SB2 R. JISA. H2SB2 R. JISA. H2SB2 Disease action Description Extense Summary Chromosome Gytoband Gytoband Chromosome Description Ensembl | Signal Transferction Signaling by GPCI Persons bester of discusse Persons bester of discusse Commonwere (C-C) more) higher 22 (HGH4G Bymolog) This and reference on of Several Cys-Cys (a) for general cys-Cys (b) for general cys (b) for | CO) rytoman genes clustered on the societies. The CC cystames are proto is, denoted cares, hadrard atter cares motioner may play and in the static territory, sep 2014) External Daba ICBI Gene, Ur | tase (Ensembl, | Remay of secreted protein The secreted protein the secreted protein and the secreted protein and the secreted protein the secreted protein and the secreted protein and the secreted protein the secreted protein and the secreted protein and | Highlight mild Increased ser BACKERS Stati diseases Cutaneous 7: diseases Cutaneous 7: diseases Cutaneous 7: diseased Cutaneous 7: d | Not determine Hydrayfe mat no han cutaneous T cel-attracting Stell tro too material acrossing (A) and associated with version immunologi cel-attracting transmission (CTAC); de oppressed in sith and attracting (CTAC); de oppressed the setture (CTAC); de do Spatients with AD into 3 groups | selectin levels, serum activation-regulated c and serum macrophic chemoking levels | thyrd us and chemploine levels, age-centived | I ligand for CC eases. In the sera of etients with AD trol subjects. |
| pathway Jarms? R. HSA. H2SB2 R. HSA. H2SB2 R. HSA. H2SB2 R. HSA. H2SB2 Disease section Description Enversione Cytoband Cytoband Cytoband Cytoband Chromosome focation (bp) Ensembl | Signal Transferction Signaling by GPCI Persons bester of discusse Persons bester of discusse Commonwere (C-C) more) higher 22 (HGH4G Bymolog) This and reference on of Several Cys-Cys (a) for general cys-Cys (b) for general cys (b) for | CO) rytoman genes clustered on the societies. The CC cystames are proto is, denoted cares, hadrard atter cares motioner may play and in the static territory, sep 2014) External Daba ICBI Gene, Ur | tase (Ensembl, niProtKB, | family of accreted proton The cyclobian encoded to the cyclobian encoded to the The product of the any sites and other agree where and other agree | Highlight mater Increased set BACKGROUT Skin diseases Cutaneous T- Cessolocities The epidemia OBJECTIVE: patients with J and PsV METHODS: V We also divid calingorities Morsover, we | Inter Setterice. Playlogit matches mun cataneous T cel-attactory EMM ND Bon Barget Kennesse (Annue cataneous T cel-attactory EMM cel-attactory Characteristics) (CTAC) cel-attactory Barget Characteristics) (Parpersonal Mathematics) (CTAC) exceptor 10 W proprised Mathematics) and attactory of this Investigation via AD and PVV and to examine the co- AD and PVV and to examine the co- We measured the serum CTACX level is compared the serum CTACX level | selectin levels, serum activation-equated c and serum macrophi chamoking levels | t thryfu sand chemp Sunn elevels, sge certwed s win Blook accustors and relepsi termeleted as chrone and relepsi termeleted as chrone and relepsi termeleted as chrone and relepsi termeleter as chrone and relepsi termeleter as a chron | eases. In the sena of etients with AD throl subjects. d them among 3 |
| pathway arms? Raisa.12522 Raisa.2522 Raisa.2529 Chromosome Summary Chromosome Golor | Signal Transferction Signaling by GPCI Persons bester of discusse Persons bester of discusse Commonwere (C-C) more) higher 22 (HGH4G Bymolog) This and reference on of Several Cys-Cys (a) for general cys-Cys (b) for general cys (b) for | CO) rytoman genes clustered on the societies. The CC cystames are proto is, denoted cares, hadrard atter cares motioner may play and in the static territory, sep 2014) External Daba ICBI Gene, Ur | tase (Ensembl, niProtKB, | Sensy of secreted protein The cyclester encodes and the cyclester second protein cyclester second other aspect | Highlight meter Increased ser BACKGROUT skit diseases Cutaneous T. Breassonn in it is selective patients with a data psy METHODS V We also clinic categories Moreover, we Immunohisto patients with a | Not determine "High byfe matt no- num cutaneous." T cell-attracting Filled ND Bom Beget extendes (Ad) as associated with versions immunologi cell-attracting benefative (CTAC), ceopter 10. If verpressed histis and attracting to excepter 10. The purpose of this investigation wi we measured the seture (CTAC), where the assisted the seture (CTAC) is over the measured the seture (CTAC) is over a compared the seture (CTAC) is over compared the seture (CTAC) is over the seture (CTAC) is o | selectin levels, serum activation-equated c and serum macrophi- chemoking levels | t thryfu sand chemp Sume levels, sge certwell | eases in the sera of etients with AD trol subjects d them among 3 estional skin of |
| pathway arms? A 1154-12522 A 1154-12522 A 1154-12522 Diseases Dise | Signal Transferction Signating by GPCI Persons select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol Trans, selected choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 (HGH4G Symbol select by select choicease Commonitore (C-C month) signed 22 | CO roteman genes clustered on the costs. The CC cystames are proved in denoting care, hadrard where can hadrard where care, hadrard where care hadrard where care hadrard are care and the static static care. The cost of the static static care and the static care and the static care and the static static care and the static care and the static care and the static static care and the static care and the static care and t | tase (Ensembl, niProtKB, | Sensity of secreted proteins . The cycleanse enclosed is the second seco | Horizon mate Increased series (accessed) (ac | Not determine "High light matching mum cataneous" T cell-attracting IMM ND Born Biggs etermating (Arg) as associated bit versions immunologi cell-attracting Bisenational (CTAC), cell-attracting Bisenational (CTAC), exception 10. The purpose of this trendingtorio with and PioV and to examine the corr AD and PioV and to examine the corr AD and PioV and to examine the corr and pioV and to examine the corr et compared the serum CTACX level compared the serum CTACX level compared the serum CTACX level compared the serum CTACX level compared the serum CTACX levels compared the serum CTACX levels as an examination of CTACX and IP AD and PIoV. | selectin levels, serum activation-equated c and serum macrophi- chemoking levels | t thryfu sand chem bizne levels, sge certwell | rigand for CC eases. In the sets of etients with AD trol subjects. It was a stand of estimation of the subjects. es, servin solub |
| pathway parms? R154-16282 R155-17292 R155-17292 Desessor Description Extract gene Chromosome Cytoband Chromosome Cytoband Chromosome Cytoband Chromosome Cytoband Chromosome Chr | Signal Transferction Signaling by GPCI Persons select clinesas Commonities (C-C month) sparsed 22 (PGAPAG Symmetry) Trans-metrimentarial generes and out of several Crys-Cryst as select as premiserial sectoralistic in charge to the memory to select as premiserial sectoralistic in the memory to select as a select of the memory to the memory to se | CO) rotomer genes characted on the costs. The CC cystames are prote to determine the CC systames are prote- tion of the CC systames are protected to the CC systames are protected to the CC systames are pro- tected to the CC systames are pro- sent to the CC | tase (Ensembl, hiProtKB, antibodypedia) | Samp, of secretary protection, The cylindrate encloses of the same second seco | Hostori inste BACKGROUP six diseases Cutaneous 1. diseases Cutaneous 1. diseases The epidemia OBJECTIVE Distribution The epidemia OBJECTIVE Vie also divid calingomis METHODS Vie also divid calingomis Vie also divid calingomis Vie also divid calingomis Vie also divid calingomis Vie also divid calingomis Vie also divid vie also divi | Not location: I field high matters must clanerout. T cell-attracting Shell ND tions large acrossing (A) and cell-attracting transmission cell-attracting transmission of C execution 1 and attracting (C execution 1) and | selectin levels, serum activation-equated c and serum macrophi- chemoking levels | thyrd tus and chemp blank elvels, age-centived | rigand for CC eases. In the sets of etients with AD trol subjects. It was a stand of estimation of the subjects. es, servin solub |
| A TO C C C C C C C C C C C C C C C C C C | Signal Transferiction Signaling by GPCI Parase select disease Chemothem (C-C moth) ligand 22 (cHCMe University of the select disease d | CO) 1915/2000 geness chardered on the constant. The ICC systemes are proto to constant the ICC systemes are proto to constraint of the ICC systemes are not constraint of the ICC systemes are not constraint of the ICC system hereines, sep 2014 External Daba ICCBI Gene, Ur leXtProt and A Borner | tase (Ensembl, hiProtKB, antibodypedia) | Dewrload | Hostori note Strataged an BACKGROUP skill disease Cutaneous 1 disease Cutaneous 1 disease Cutaneous 1 disease Cutaneous 1 disease The epidemia OBJECTIVE: patients whith Acceptories Monover, see semantuchable patients whith Categories La Conceptori manophologi Bernio TCAC | Not locknow. High high muttoo must calaneous T cel- attraction (2014) ND tions large atmoster (2014) and (2014) cel-attraction immunologic end-attraction immunologic end-attraction (2014) and attraction (2014) if expensed in site and attraction (2014) the page of this investigations with the immorphism is an associated of the page of this investigations with eld 00 patients with AO with 3 pages compared the serum CTACK level is compared the serum CTACK level is compared the serum CTACK level is exprared the serum CTACK level is compared the serum CTACK level is compared the serum CTACK level is patients with AO with 3 patients TACK level is patients with AO set is patients attraction of the celesticn level is compared to the serum the serue the serue the serum compared to weak. | selectin levels, serum activation-regulated c and serum macrophil chamoking levels | t thryfu sa and chem biane levels, sge certwell | I ligand for CC eases, in the sera of effects with AD trol subjects is them among 3 estonal skin of hrul subjects es, sonum solubli serum |
| pathway parms? R115A.15202 R115A.15202 R115A.15202 Disease Dis | Signal Transferiction Signaling by GPCI Parase select disease Chemothem (C-C moth) ligand 22 (cHCMe University of the select disease d | CO) rotomer genes characted on the costs. The CC cystames are prote to determine the CC systames are prote- tion of the CC systames are protected to the CC systames are protected to the CC systames are pro- tected to the CC systames are pro- sent to the CC | tase (Ensembl, hiProtKB, antibodypedia) | Dewrload | Topologi made Increased set BACKGROU Standieses Curaneous T Standieses Curaneous T Standieses patrens with AMETHOOS // We also delot patrens with AMETHOOS // METHOOS // METHOOS // Standieses Patrens with AMETHOOS // AMETHOOS // AMETH | Not locknow. High high muttoo must calaneous T cel- attraction (2014) ND tions large atmoster (2014) and (2014) cel-attraction immunologic end-attraction immunologic end-attraction (2014) and attraction (2014) if expensed in site and attraction (2014) the page of this investigations with the immorphism is an associated of the page of this investigations with eld 00 patients with AO with 3 pages compared the serum CTACK level is compared the serum CTACK level is compared the serum CTACK level is exprared the serum CTACK level is compared the serum CTACK level is compared the serum CTACK level is patients with AO with 3 patients TACK level is patients with AO set is patients attraction of the celesticn level is compared to the serum the serue the serue the serum compared to weak. | selectin levels, serum activation-regulated d and serum macropath chemoking levels | thyrd us and chemplane levels, sge certwell | I ligand for CC eases. In the sora of atterts with AD trof subjects. at them among 3 estonal skin of troi subjects. es, sorium solubl serum |

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

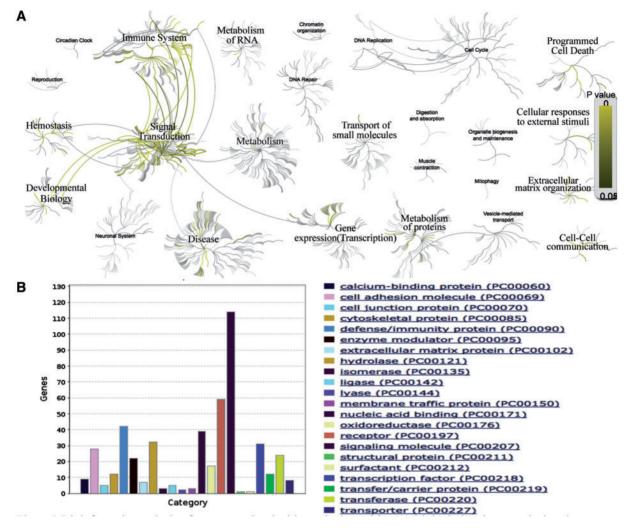


Figure 2. Bioinformatics analysis of genes associated with AD. (A) Biological pathway analysis using Reactome (http://www.reactome.org/). (B) Protein class analysis using PANTHER (http://pantherdb.org/).

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 allergyrelated 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.

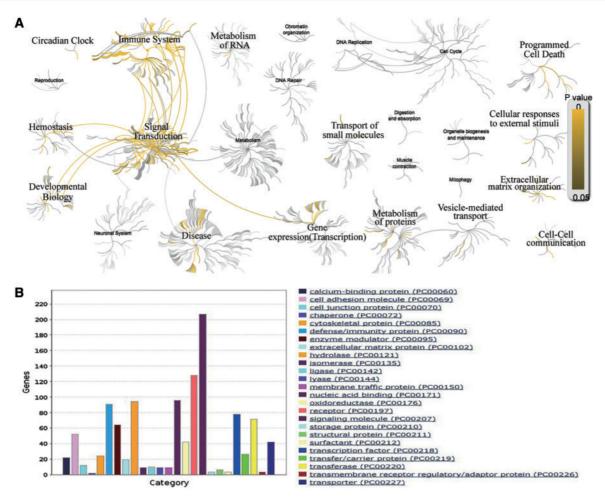


Figure 3. Bioinformatics analysis of the human allergy-related genes. (A) Biological pathway analysis using Reactome. (B) Protein class analysis using PANTHER.

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.
- Ring, J.T., Plattsmills (2005) Allergy in Practice. Journal of the American Medical Association. 133, 864–865.
- 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.

- 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.
- Pawankar, R., Canonica, G.W., St Holgate, S.T. *et al.* (Update. 2013) *The WAO White Book on Allergy*. World Allergy Organization, Wisconsin.
- 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.
- 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.
- Meng, J.F. and Rosenwasser, L.J. (2010) Unraveling the genetic basis of asthma and allergic diseases. *Allergy Asthma Immunol. Res.*, 2, 215–227.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.

- 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.
- 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.
- 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.
- 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.
- Nunes, T., Campos, D., Matos, S. *et al.* (2013) BeCAS: biomedical concept recognition services and visualization. *Bioinformatics*, 29, 1915–1916.
- Schriml,L.M., Arze,C., Nadendla,S. *et al.* (2011) Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Res.*, 40, D940–D946.
- 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.
- Fabregat,A., Sidiropoulos,K., Garapati,P. et al. (2016) The Reactome pathway knowledgebase. Nucleic Acids Res., 44, D481–D487.
- 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.
- 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.
- Bairoch, A., Apweiler, R., Wu, C.H. *et al.* (2005) The universal protein resource (UniProt). *Nucleic Acids Res.*, 33(suppl_1), D154–D159.
- Lane,L., Argoud-Puy,G., Britan,A. *et al.* (2012) neXtProt: a knowledge platform for human proteins. *Nucleic Acids Res.*, 40, D76–D83.
- Björling, E. and Uhlén, M. (2008) Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol. Cell. Proteomics*, 7, 2028–2037.
- Weidinger, S. and Novak, N. (2016) Atopic dermatitis. *Lancet*, 387, 1109–1122.
- 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**.