JPP JOURNAL OF PROTEINS AND PROTEOMICS
7(3), 2016, pp. 213-220

Research Article

AIDBD: AUTOIMMUNE AND INFLAMMATORY DISEASES BIOMARKER DATABASE

Kulwinder Singh[1,*], Monika[2] and Neelam Verma[1]

[1]*Department of Biotechnology, Punjabi University, Patiala 147002, Punjab, India*
[2]*Department of Biotechnology, Mata Gujri College, Fatehgarh Sahib 140406, Punjab, India*

Abstract: One of the major challenges facing the healthcare industry is how to personalize, or tailor healthcare products and services to individuals' unique genetic and biomarker make-ups. Biomarkers provide information about normal or patho-physiological processes to detect or define disease progression or to predict or quantify therapeutic responses. Once these footprints have been identified and measured, they can then be used to personalize or tailor treatment plans, products and services to each individual's unique makeup and background. Autoimmune and Inflammatory Diseases Biomarker Database (AIDBD) is one of the first efforts to build an easily accessible and comprehensive literature-derived database covering information on known autoimmune and inflammatory diseases, biomarkers and available medications. It allows users to link autoimmune and inflammatory diseases to protein or gene biomarkers through its user interface. Currently, AIDBD integrates 206 biomarkers for 21 autoimmune and inflammatory diseases and data on 516 launched drugs for the treatment these diseases. The database is freely accessible at http://www.aidbd.in/.

Key words: Autoimmune diseases; Inflammatory diseases; Biomarker; Database; Drug development; Personalized medicine

Note: Coloured Figures available on Journal Website in "Archives" Section

Introduction

Autoimmune diseases are a family of more than 80 chronic and often disabling illnesses that develop when underlying defects in the immune system lead the body to attack its own organs, tissues and cells. Since cures are not yet available for most autoimmune diseases, patients face a lifetime of illness and treatment. They often endure debilitating symptoms, loss of organ function, reduced productivity at work, and high medical expenses (Jacobson *et al.*, 1997). And because most of these diseases disproportionately afflict women, and are among the leading causes of death for young and middle-aged women, they impose a heavy burden on patients' families and on society (Walsh and Rau, 2000; US Department of Health and Human Services Report to Congress, 2005).

Autoimmune diseases are commonly considered complex immune disorders. While many autoimmune diseases are rare, collectively these diseases afflict millions of patients. Despite their clinical diversity, they have one similarity, namely the dysfunction of the immune system (Hayter and Cook, 2012). It is suspected that genetic defects play a role in the etiology of these diseases. Modern high throughput technologies, like mRNA micro arrays have enabled researchers to investigate diseases at a genome-wide level (Cotsapas and Hafler, 2013). In contrast to classical inherited genetic diseases like sickle cell anemia, autoimmune diseases are not caused by the defect of a single gene but by the dysfunction of the complex interaction of a group of genes. Although no autoimmune disease has been completely

Corresponding Author: **Kulwinder Singh**
E-mail: kulwinder265@gmail.com

Received: June 12, 2016
Accepted: August 23, 2016
Published: August 30, 2016

analyzed, there has been tremendous success in recent years in identifying major players in the development of autoimmune diseases (Karopka *et al.*, 2006).

A biomarker, as defined by the Food and Drug Administration (FDA) of the United States, is any "characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (Biomarkers Definitions Working Group, 2001). Biomarkers are characteristics that can be objectively measured and evaluated. They provide information about normal or patho-physiological processes to detect or define disease progression or to predict or quantify therapeutic responses (Wilson *et al.*, 2007). With the recent explosion of high-performance 'omic' technologies – genomics, proteomics and metabolomics, among others – the rate at which biomarker candidates are being discovered is now faster than ever (Moore *et al.*, 2007).

In order to advance our understanding of biomarkers and their roles in early autoimmune and inflammatory processes, we have developed an integrated user-friendly database that catalogs putative and validated biomarkers and relates them to autoimmune and inflammatory disease processes. In addition, we have added information on approved medications for the treatment of autoimmune and inflammatory diseases. This freely accessible resource will be a valuable research tool and a contribution to improved public heath.

Methodology

Asthma, rhinitis, arthritis, diabetes, transplantation, biomarker, target, inhibitor, antagonist, agonist etc. were used as keywords in PubMed Medline Database to search for the research papers. All the results were screened at the abstract level to segregate the false positive papers from the hit list. All potential published studies on drugs, candidate protein and gene biomarkers were evaluated. The true positive papers were collected to perform the manual data curation process on diseases, biomarkers and drugs. Information on the proteins, taxonomy ID, lineage, amino acid length, catalytic activity, molecular function, and role in diseases was retrieved from UniProt, NCBI, OMIM, HGNC, RefSeq, PIR and other biological databases. Information on the genes, its gene type, taxonomy, identifiers of other databases, its molecular function, nucleotide length, nucleotide sequence, amino acid length etc. was retrieved from NCBI, KEGG, HGNC, RefSeq, OMIM, MGI and GenBank databases. Information on drugs, its mechanism of action, mode of administration was retrieved from literature, Google, PubChem and clinical trials databases. The collected information was used to create the database using HTML, CSS, JavaScript languages and tools. The front end of the database was designed with HTML for creating web pages and CSS, a style sheet language, for enriching the look and format of the web pages. An open source JavaScript library was used to generate visual effects on the web pages that display results. PHP, a server-side scripting language for web development was used for the interface and MySQL for the backend coding of the database.

Discussion

Early response to autoimmune and inflammatory diseases depends on rapid clinical diagnosis and detection, which, if in place, are able to ameliorate suffering and economic loss. Biomarkers, molecules that can be sensitively measured in the human body, are by definition potentially diagnostic. The efficacy of biomarkers to autoimmune and inflammatory diseases lies in their capability to provide early detection, establish highly specific diagnosis, determine accurate prognosis, direct molecular-based therapy and monitor disease progression (Baker, 2005). They are increasingly important in both diagnostic and therapeutic processes, with high potential to guide preventive interventions. Vast resources have been devoted to identifying and developing biomarkers that can help determine the treatments for patients. There is growing consensus that multiple markers will be required for most diagnoses, while single markers may serve in only selected cases. Despite intensified interest and research, however, the rate of development of novel biomarkers has been falling (Rifai *et al.*, 2006), suggesting that a resource that leverages existing data is overdue. At present the

databases containing information about biomarkers are focused predominantly on cancer: gastric cancer knowledgebase (Lee *et al.*, 2006), early detection research network (Srinivas *et al.*, 2001), integrated cancer biomarker information system (Feng *et al.*, 2005) etc. Furthermore, there are very good resources available with respect to information on drugs like ChEBI; a freely available dictionary of molecular entities focused on 'small' chemical compounds (Hastings *et al.*, 2013), DrugBank, a bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target information (Wishart *et al.*, 2006) etc. but no systematic effort has been described for easily accessible integrating information from the disease specific biomarker domains and therapies especially for autoimmune and inflammatory diseases. AIDBD introduces a community annotation database of biomarkers, with interfaces for users to directly explore the information on autoimmune and inflammatory diseases, putative and validated biomarkers and known therapies. It was designed to collect, store and display information about biomarkers, conjoined to identifiers of research tools for sequence and structural analyses of the data. AIDBD currently includes information on 201 biomarkers from 21 autoimmune and inflammatory diseases and 516 launched drugs for these diseases. Diseases covered in AIDBD includes autoimmune and inflammatory diseases of the respiratory tract i.e. Asthma, Chronic Obstructive Pulmonary Disease (COPD), Allergic Rhinitis & Cystic Fibrosis; rheumatic joint disorders i.e. Rheumatoid Arthritis, Osteoarthritis, Psoriatic Arthritis & Ankylosing Spondylitis; autoimmune diseases affecting multiple organs i.e. Systemic Lupus Erythmatosus & Sjogren's Syndrome; skin disorders: Psoriasis & Atopic Dermatitis; autoimmune disease due to the destruction of insulin-producing beta cells in the pancreas: Diabetes Type 1; demyelinating disease of central nervous system: Multiple sclerosis; Digestive disorders & Gastrointestinal diseases: Crohn's disease, Ulcerative colitis, Irritable bowel syndrome; inflammatory disease of the blood vessels: Vasculitis and autoimmune disease of the liver i.e. Primary Biliary Cirrhosis. Home page of AIDBD is illustrated in Figure 1.

The record entry in the database contains the following information about the protein targets: recommended name, its brief description, gene name, organism information, taxonomy ID, lineage, amino acid length, protein existence, its function, catalytic activity (if any); enzyme regulation (if any); involvement in disease, family to which it belongs, its molecular function, identifiers of UniProt, NCBI, Ensembl, OMIM, HGNC, RefSeq, PIR, UNIGENE databases and tools, AA sequence, molecular weight, summary, inhibitors developed or under development (if any) and references. An example of information stored in AIDBD on 'Interleukin-4' as a biomarker of respiratory diseases is listed in Table 1. In the gene biomarker section, AIDBD provides information on gene name, official symbol, full name, synonyms (if any), summary, gene type, organism, taxonomy, identifiers of NCBI, KEGG, GENECARDS, HGNC, RefSeq, OMIM, Ensembl, MGI, HOMOLOGENE databases and tools, its molecular function, involvement in biological process, nucleotide length, nucleotide FASTA identifier, nucleotide sequence, amino acid length, amino acid FASTA identifier, amino acid sequence, homologous genes and references. An example of information stored in AIDBD on 'RNASE3' as a gene biomarker is listed in Table 2. Drug section of AIDBD provides information on launched drugs for the treatment of autoimmune and inflammatory diseases. Currently, drugs section of AIDBD provides information on drug names including brand names, drug class, mechanism of action and mode of administration. Information stored in AIDBD on some of the drugs is listed in Table 3.

Utility and future directions

AIDBD has been developed as a new resource to help the scientific and medical community. Currently, AIDBD provide useful targets or biomarkers relevant for clinical diagnosis of autoimmune and inflammatory diseases. It helps in accelerating the research as it presents the underlying molecular mechanism of the disease, underpinning the targets. It also provides information on drugs as well as their mechanism

Table 1
An example of information stored in AIDBD on 'Interleukin-4' as a biomarker of respiratory diseases

Protein Name	Interleukin-4
Recommended name	Interleukin-4
Description	Protein is a pleiotropic cytokine produced by activated T cells. The interleukin 4 receptor also binds to IL13, which may contribute to many overlapping functions of this cytokine and IL13.
Gene Name	IL4
Organism	Homo sapiens
Taxanomy ID	9606
Lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo
AA Length	153 AA
Protein existence	Evidence at protein level.
Function	Participates in at least several B-cell activation processes as well as of other cell types. It is a costimulator of DNA-synthesis. It induces the expression of class II MHC molecules on resting B-cells.
Catalytic activity	none
Enzyme regulation	none
Involvement in disease	The TH2-like cytokine interleukin (IL)-4 play a pivotal role in airway wall inflammation in asthma and these cytokines are increased in peripheral blood and bronchoalveolar lavage fluid from asthmatic patients.
Family	cytokine
Molecular Function	cytokine activity,growth factor activity,interleukin-4 receptor binding
UniProtID	P05112
NCBIID	P05112
Ensembl	ENST00000231449; ENSP00000231449; ENSG00000113520; ENST00000350025; ENSP00000325190; ENSG00000113520;
OMIM	147780. gene.,601367. phenotype.
HGNC	6014
RefSeq	NP_000580.1.,NP_758858.1.
PIR	A25946, A30546.
UNIGENE	Hs.73917
AA Length	153 AA.
AA Sequence	MGLTSQLLPPLFFLLACAGNFVHGHKCDITLQEIIKTLNSLTEQKTLCTELTVTDIFA ASKNTTEKETFCRAATVLRQFYSHHEKDTRCLGATAQQFHRHKQLIRFLKRLDRNLWGLAGLN SCPVKEANQSTLENFLERLKTIMREKYSKCSS
MolWeight	17492.2
Summary	Interleukin-4 is a cytokine that induces differentiation of naive helper T cells (Th0 cells) to Th2 cells. Upon activation by IL-4, Th2 cells subsequently produce additional IL-4.
Inhibitors	SIL-4R Nuvance
Reference	J Allergy Clin Immunol.;107(6):963-70. Jun 2001Borish LC, Nelson HS, Corren J, Bensch G, Busse WW, Whitmore JB, Agosti JM; IL-4R Asthma Study Group.PMID: 11398072

Table 2
An example of information stored in AIDBD on 'RNASE3' as a gene biomarker

Gene Name	RNASE3
Official Symbol	RNASE3
Full Name	ribonuclease, RNase A family, 3 (eosinophil cationic protein)
Synonyms	ECP, RNS3
Summary	Cytotoxin and helminthotoxin with low-efficiency ribonuclease activity. Possesses a widevariety of biological activities. Exhibits antibacterial activity.
Gene Type	Protein coding
Organism	Homo sapiens
Taxanomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
NCBIGI	6037
KEGGID	hsa:6037.
GENECARDS	GC14P020429
HGNCID	10046
RefSeqID	NM_002935
OMIM	131398
Ensembl	ENSG00000169397
MGI	1858598
HOMOLOGENE	110844
Chromosome	Chromosome: 14
Molecular Function	nucleic acid binding, pancreatic ribonuclease activity
Biological Process	RNA catabolic process, defence response to bacterium
NT Length	716 bp
NT FASTA	>gi|45243506|ref|NM_002935.2| Homo sapiens ribonuclease, RNase A family, 3 (eosinophil cationic protein) (RNASE3), mRNA
NT Sequence	GAACAACCAGCTGGATCAGTTCTCACAGGAGCCACAGCTCAGAGACTGGGAAACATG GTTCCAAAACTGTTCACTTCCCAAATTTGTCTGCTTCTTCTGTTGGGGCTTATGGGTGT GGAGGGCTCACTCCATGCCAGACCCCCACAGTTTACGAGGGCTCAGTGGTTTGCCATC CAGCACATCAGTCTGAACCCCCCTCGATGCACCATTGCAATGCGGGCAATTAACAATT ATCGATGGCGTTGCAAAAACCAAAATACTTTTCTTCGTACAACTTTTGCTAATGTAGTT AATGTTTGTGGTAACCAAAGTATACGCTGCCCTCATAACAGAACTCTCAACAATTGTC ATCGGAGTAGATTCCGGGTGCCTTTACTCCACTGTGACCTCATAAATCCAGGTGCACA GAATATTTCAAACTGCACGTATGCAGACAGACCAGGAAGGAGGTTCTATGTAGTTGCA TGTGACAACAGAGATCCACGGGATTCTCCACGGTATCCTGTGGTTCCAGTTCACCTGG ATACCACCATCTAAGCTCCTGTATCAGCAGTCCTCATCATCACTCATCTGCCAAGCTCC TCAATCATAGCCAAGATCCCATCCCTCCATGTACTCTGGGTATCAGCAACTGTCCTCAT CAGTCTCCATACCCCTTCAGCTTTCCTGAGCTGAAGTCCCTTGTGAACCCTGCAATAAA CTGCTTTGCAAATTCA
AA Length	160 AA
AA FASTA	>gi|147744558|sp|P12724.2|ECP_HUMAN RecName: Full=Eosinophil cationic protein; Short=ECP; AltName: Full=Ribonuclease 3; Short=RNase 3; Flags: Precursor
AA Sequence	MVPKLFTSQICLLLLLGLMGVEGSLHARPPQFTRAQWFAIQHISLNPPRCTIAMR AINNYRWRCKNQNTFLRTTFANVVNVCGNQSIRCPHNRTLNNCHRSRFRVPLLHCDLINP GAQNISNCTYADRPGRRFYVVACDNRDPRDSPRYPVVPVHLDTTI
Metabolic pathway	Asthma
Homologous gene	Ear5, Mus musculus- GeneID: 54159 ,Rnase2, Rattus norvegicus-GeneID: 474169,Rnase17, Rattus norvegicus-GeneID: 497195,Ear11, Rattus norvegicus-GeneID: 192264
References	Eosinophil cationic protein in serum from adults with asthma and with chronic obstructive pulmonary disease.Yoshizawa A, Kamimura M, Sugiyama H, Kudo K, Kabe J.Nihon Kyobu Shikkan Gakkai Zasshi. 34(1):24-9. Jan 1996.PMID: 8717287

Table 3
Information stored in AIDBD on drugs

Drug Name	*Drug Class*	*Mechanism of Action*	*Mode of administration*
(-)-cetirizine; Sepracor; levocetirizine; levocetirizine dihydrochloride; Xazal; Xusal; Xyzal; Xyzala; Xyzall	Antiallergic, non-asthma; Antiasthma; Antipruritic/ inflamm, allergic	Histamine H1 receptor antagonist; Histamine receptor antagonist	Oral; Oral, swallowed
131I-rituximab; anti-CD20 MAb, Genentech; anti-CD20 MAb, IDEC; anti-CD20 MAb, Roche; anti-CD20 MAb, Zenyaku; IDEC C2B8 (SC); IDEC-C2B8 (IV); MabThera; MabThera (IV); MabThera (SC)	Anti-inflammatory; Antiarthritic, immunological; Anticancer, immunological; Antiparkinsonian; Haematological; Immunosuppressant; Monoclonal antibody, chimaeric; Multiple sclerosis treatment; Ophthalmological; Urological	CD20 antagonist	Injectable; Injectable, intravenous; Injectable, subcutaneous
alemtuzumab; alemtuzumab (IP); alemtuzumab (IV); alemtuzumab (SC); alentuzumab; alentuzumab (IP); alentuzumab (IV); alentuzumab (SC); Campath; Campath (IP); Campath (IV);Campath (SC); Campath-1H; Campath-1H (IV); Campath-1H (SC); LDP-03; LDP-03 (IP	Antianaemic; Antiarthritic, immunological; Anticancer, immunological; Immunosuppressant; Monoclonal antibody, humanized; Multiple sclerosis treatment	CD52 antagonist; Immunosuppressant; Lymphocyte inhibitor	Injectable; Injectable, intraperitoneal; Injectable, intravenous; Injectable, subcutaneous
CNTO 148; CNTO-148; CNTO-148 (intravenous); CNTO-148 (subcutaneous); golimumab; golimumab (intravenous); golimumab (subcutaneous); Simponi; Simponi (intravenous); Simponi (subcutaneous)	Anti-inflammatory; Antiarthritic, immunological; Antipsoriasis; Cardiovascular; GI inflammatory/ bowel disorders; Monoclonal antibody, human; Neuroprotective	Tumour necrosis factor alpha antagonist	Injectable; Injectable, intravenous; Injectable, subcutaneous
Emlucast; Kipres; L-706631; Lukair; Lukasm; MK-0476; MK-0476 (granules); MK-476; MK0476; montelukast; montelukast (chewable); montelukast (granules); montelukast (IV); montelukast sodium; MR-4524; Romilast; Singulair	Antiallergic, non-asthma; Antiasthma	Leucotriene D4 antagonist; Leucotriene receptor antagonist	Injectable; Injectable, intravenous; Oral; Oral, swallowed

of action for better understanding on how these drugs are involved in the biological processes for treatment of autoimmune and inflammatory diseases. Its disease section also provides brief overviews on autoimmune and inflammatory diseases, symptoms, diagnosis methods, severity levels and available treatment options. The database content is carefully maintained and updated. Repeated literature searches and curation are ongoing for identification and periodic update of new data into the database. Development of a search engine is also planned by accommodating the search based on gene identifiers, disease name, protein name, drug name etc. Inclusion of data for other related diseases, to broaden the scope of the database to a larger audience is also under consideration. AIDBD can be publicly accessed from any Web browser at http://www.aidbd.in/.

Acknowledgement

Authors are grateful to University Grants Commission (UGC), Government of India for financial support under the major research project scheme (File No. 39-290/2010 (SR)). The authors report no conflict of interest.

Abbreviations

AIDBD, Autoimmune and Inflammatory Diseases Biomarker Database; mRNA, messenger Ribonucleic Acid; FDA, Food and Drug Administration; NCBI, National Center for Biotechnology Information; OMIM, Online Mendelian Inheritance in Man; HGNC, The HUGO Gene Nomenclature Committee; RefSeq, NCBI Reference

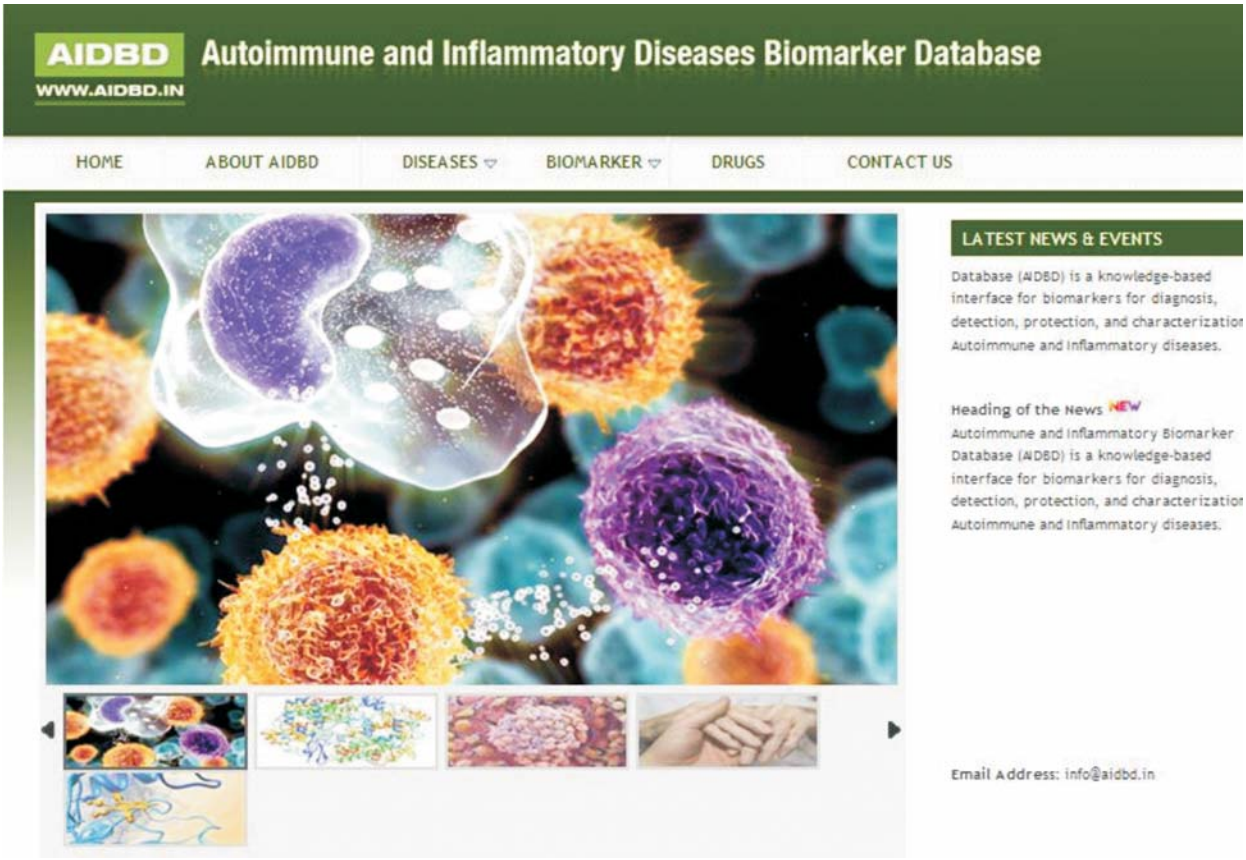

Figure 1: **Home page of AIDBD displaying various sections of database: Disease, Biomarker and Drugs section.**

Sequence Database; PIR, The Protein Information Resource; KEGG, Kyoto Encyclopedia of Genes and Genomes; MGI, Mouse Genome Informatics; COPD, Chronic Obstructive Pulmonary Disease; UniProt, Universal Protein Resource; GENECARDS, Database of Human Genes; HOMOLOGENE, NCBI tool for automated detection of homologs among the annotated genes of several completely sequenced eukaryotic genomes; FASTA, Fast Alignment; RNASE3, Ribonuclease 3; HTML, Hyper Text Markup Language; CSS, Cascading Style Sheets; ChEBI, Chemical Entities of Biological Interest.

References

Baker, M. (2005). In biomarker we trust? Nat. Biotechnol. 23, 297–304.

Biomarkers Definitions Working Group (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin. Pharmacol. Ther. 69, 89-95.

Cotsapas, C. and Hafler, D. A. (2013). Immune-mediated disease genetics: the shared basis of pathogenesis. Trends Immunol. 34, 22–26.

Feng, W., Wu, B., Phan, J., Dale, J., Young, A. N. and Wang, M. D. (2005). An integrated cancer biomarker information system. IEEE Eng. Med. Biol. Soc. 3, 2851–2854.

Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., Muthukrishnan, V., Owen, G., Turner, S., Williams, M. and Steinbeck, C. (2013). The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Res. 41(Database issue), D456-D463.

Hayter, S. M. and Cook, M. C. (2012). Updated assessment of the prevalence, spectrum and case definition of autoimmune disease. Autoimmun. Rev. 11, 754–65.

Jacobson, D. L, Gange, S. J, Rose, N. R and Graham, N. M (1997). Epidemiology and estimated population burden of selected autoimmune diseases in the United States. Clin. Immunol. Immunopathol. 84, 223–243.

Karopka, T., Fluck, J., Mevissen, H.T. and Glass, A. (2006). The Autoimmune Disease Database: a dynamically compiled literature-derived database. BMC Bioinform. 27, 325-341.

Lee, B. T. K, Song, C. M., Yeo, B. H., Chung, C. W, Chan, Y. L., Lim, T. T, Chua, Y. B., Loh, M. C. S., Ang, B. K. Vijayakumar, P., Liew, L., Lim, J., Lim, Y. P., Wong, C. H., Chuon, D., Rajagopal, G. and Hill, J. (2006). Gastric cancer (biomarkers) knowledgebase (GCBKB): a curated and fully integrated knowledgebase of putative biomarkers related to gastric cancer. Biomarker Insights 2, 135–141.

Moore, R.E., Kirwan, J., Doherty, M.K. and Whitfield, P.D. (2007). Biomarker Discovery in Animal Health and Disease: The Application of Post-Genomic Technologies. Biomarker Insights 2, 185-196.

Rifai, N., Gillette, M. A. and Carr, S. A. (2006). Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat. Biotechnol. 24, 971–983.

Srinivas, P. R., Kramer, B. S. and Srivastava, S. (2001). Trends in biomarker research for cancer detection. Lancet Oncol. 2, 698-704.

The Autoimmune Diseases Coordinating Committee. Progress in Autoimmune Diseases Research: Report to Congress. US Department of Health and Human Services. 2005.

Walsh, S. J, Rau, L. M. (2000). Autoimmune diseases: a leading cause of death among young and middle-aged women in the United States. Am. J. Public Health 90, 1463–1466.

Wilson, C.L., Schultz, S. and Waldman, S. A. (2007). Where Medicine, Business, And Public Policy Intersect. Biotechnol. Healthc. 4, 33-42.

Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z. and Woolsey, J. (2006). DrugBank: a comprehensive resource for *in silico* drug discovery and exploration. Nucleic Acids Res. 34, D668-D672.

Web References

Annotated Protein Sequence Database, available at: *http://pir.georgetown.edu/*

Chemical Entities of Biological Interest, available at: *https://www.ebi.ac.uk/chebi/*

Clinical Trials Database, availble at: *https://clinicaltrials.gov/*

Collection of databases dealing with genomes, biological pathways, diseases, drugs, and chemical substances, available at: *http://www.genome.jp/kegg/*

Database resource for the laboratory mouse, providing integrated genetic, genomic, and biological data to facilitate the study of human health and disease, available at: *http://www.informatics.jax.org/*

DrugBank, available at: *http://www.drugbank.ca/*

Early Detection Research Network, available at: *https://edrn.nci.nih.gov/*

Gastric Cancer Knowledgebase, available at: *http://biomarkers.bii.a-star.edu.sg/*

GenBank Database, available at: *http://www.ncbi.nlm.nih.gov/genbank/*

Gene Nomenclature & Gene Families Database, available at: *http://www.genenames.org/*

Nucleotide Sequence Database, available at: *http://www.ncbi.nlm.nih.gov/refseq/*

Online Mendelian Inheritance in Man: An Online Catalog of Human Genes and Genetic Disorders, available at: *http://www.omim.org/*

PubChem Database, available at: *https://pubchem.ncbi.nlm.nih.gov/*

PubMed Literature Database, available at: *http://www.ncbi.nlm.nih.gov/pubmed*

Universal Protein Resource Database, available at: *http://www.uniprot.org/*