

CDRH: A Database of Complex Disease-related Haplotypes in Human

Ruijie Zhang✉, Yongshuai Jiang✉, Hongchao Lv✉, Xuehong Zhang✉, Peng Sun✉, Yan Zhang✉, Jin Li✉, Mingming Zhang✉, Zhenwei Shang✉, Xia Li✉

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150086, China

✉ Corresponding Author email: zhangruijie2013@gmail.com; ✉ Author

Computational Molecular Biology, 2011, Vol.1, No.3 doi: 10.5376/cmb.2011.01.0003

Copyright © 2013 Zhang et al. This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract Many common variations in DNA sequences and their specific combinations (haplotypes) may be the underlying causes of differences in individual susceptibility to complex diseases. Great progress has been made in accumulating abundant resources relating to complex disease-related haplotypes. However, these resources are scattered among different literature, resulting in reduced utilization of the information. Therefore, we developed a database of complex disease-related haplotypes in human (CDRH). To date, a total of 1 125 haplotypes involved in 114 complex diseases, such as breast cancer, type 2 diabetes, and rheumatoid arthritis, have been manually extracted from 274 papers. After careful review of the literature, we obtained detailed information on haplotypes and diseases. Furthermore, we integrated gene- and SNP- (and/or microsatellite-) related information from external databases to facilitate further analysis. Via a user-friendly interface, users can query the CDRH by disease name, gene name, chromosome number, or SNP ID (rs#). We hope that CDRH will enrich our knowledge of haplotypes and promote research into the relationship between haplotypes and heritable risk for complex diseases. The CDRH database is freely available at <http://bioinfo.hrbmu.edu.cn/cdrh>.

Keywords CDRH; Haplotypes; Complex diseases

Introduction

A haplotype comprises a specific allele set observed on a single chromosome, or part of a chromosome (HapMap, 2003; LIN and ZENG, 2006). Haplotypes can provide critical insights into complex traits, population histories, and natural selection (Tishkoff et al., 2000; Daly et al., 2001; Gao et al., 2009). Importantly, there is increasing evidence from empirical and simulation studies that, in some circumstances, haplotypes in a chromosomal region of interest can be more powerful than using individual markers in the identification of complex disease susceptibility (Zhao et al., 2003; Gabriel et al., 2002). Many studies based on haplotypes have successfully detected genetic susceptibilities to complex human diseases (Berger et al., 2008; Soma et al., 2008), such as prostate cancer (Yaspan et al., 2008), breast cancer (Slattery et al., 2008), type 1 diabetes (Santiago et al., 2008), and rheumatoid arthritis (Hung et al., 2007).

With the exponential increases in the scale and density of genetic variation data sets, haplotype analyses have become more important in genetic studies of human diseases, and large amounts of haplotype data have been accumulated. Some haplotype-related databases have been developed for collecting and preserving haplotype information in past decades. D-HaploDB (Higasa et al., 2007) is a genome-wide definitive haplotypes database constructed by a collection of completely genotyped hydatidiform mole samples. YHRD (Kayser et al., 2002) aims to deposit Y-chromosomal short tandem repeat haplotypes for U.S. populations. mtDB (Ingman and Gyllensten, 2006) provides mitochondrial haplotypes search functions for medical and human population genetic researchers. However, there is no specific database compiling studies of haplotypes associated with complex diseases.

Preferred citation for this article:

Zhang et al., 2011, CDRH: A Database of Complex Disease-related Haplotypes in Human, Computational Molecular Biology, Vol.1, No.3 12-19 (doi: 10.5376/cmb.2011.01.0003)

Received: 11 Oct., 2011 | Accepted: 09 Nov., 2011 | Published: 28 Nov., 2011

To satisfy the requirements of molecular biologists, geneticists, and pathologists, we developed a manually curated database of complex disease-related human haplotypes (CDRH, <http://bioinfo.hrbmu.edu.cn/cdrh>.) by integrating information on haplotypes and diseases scattered in a large number of articles. CDRH is a comprehensive and well-annotated database, and is a useful resource for researchers to understand complex diseases at the haplotype level.

1 Results

1.1 Data collection and database content

Text mining was used to collect complex disease-related haplotypes and other detailed information for database construction. We searched the PubMed database (<http://www.ncbi.nlm.nih.gov/pubmed>) with a series of keywords, such as ‘complex disease haplotype’, ‘cancer haplotype’, ‘diabetes haplotype’, limiting the results to publications before May 2010 for the current version of CDRH. For systematic and reliable data collection, we checked the important information manually and implemented the following criteria: (i) the article must propose and elaborate a relationship between a complex disease and susceptible (or protective) haplotypes; and (ii) the susceptible (or protective) haplotypes must have a certain threshold or p-value of statistical tests. Ultimately, a total of 1 125 haplotypes associated with 114 complex diseases were deposited and maintained in the current CDRH database. Most of the archived information in the database is for the SNP haplotypes, and the rest consists of microsatellites.

In the CDRH database, each entry contains detailed information regarding haplotypes and diseases. The information collected includes the disease name, haplotypes associated with the disease, haplotype frequencies, the risk status of the haplotypes, the p-value of the statistical tests, the chromosome upon which the haplotypes are located, the gene symbol with which the haplotypes are associated, SNPs (or microsatellites) that make up the haplotype, and the bibliographical information from the cited literature. We not only collected a wide range of risk haplotypes, but also considered protective haplotypes, both of

which provide valuable information for future genetic studies of complex diseases.

We also integrated certain biological annotations from external databases to complement and extend the literature information. Basic information on the genes that were identified by the related haplotypes was retrieved from NCBI, including Entrez Gene ID, Unigene ID, full gene name, chromosome location of the gene, and a brief description of the gene function. Most of the haplotypes in CDRH comprised a series of SNPs; therefore, we collected information on haplotype-related SNPs from dbSNP, including SNP ID, physical position, and alleles for each SNP. In addition, many convenient links were also provided to external databases, such as dbSNP, PubMed, D-HaploDB, and HapMap, which will facilitate the future investigation of complex disease-related haplotypes. Table 1 illustrates the statistical information in the CDRH database.

Table 1 Summary of the data in CDRH

Resources	Number
Haplotype	1125
Disease	114
SNP	954
Microsatellite	288
Popoulation	66
Chromosome	22 autosomes, X Y and mitochondrion
Gene	218

1.2 Database implementation and web interface

The CDRH database uses MySQL 5.0 to store and manage the data, and implements it in PHP scripts running in an Apache/PHP environment.

1.3 Search page

The CDRH database is accessible online and allows users to retrieve detailed information pertaining to complex disease-related human haplotypes by disease name, gene name, chromosome number, or SNP ID (rs#). We first introduce the search by disease name, which is sorted alphabetically in a drop-down list box. For example, if a user selects ‘colorectal cancers’ as a query term (Figure 1a), search and browse results will

be displayed in a new page (Figure 1c). The detailed information consists of three sections: disease, literature, and haplotype. The disease section focuses on a brief summary of the pathogenesis and clinical characteristics of colorectal cancers. If users desire more comprehensive knowledge of the disease and its effects, they can enter the web site of Patient UK or Wikipedia by an included hyperlink. The literature section lists all documents concerning susceptible (or protective) haplotypes for colorectal cancers, including PubMed ID, publication date, title, and the abstract. This information provides a preliminary insight into progress in the detection and treatment of colorectal cancers based on haplotype analysis. The haplotype section presents all colorectal cancer related haplotypes, haplotype frequencies, related chromosome number, and gene symbol, SNPs (or microsatellites) that comprise a haplotype, the risk status of haplotypes, the p-value of statistical tests, and study populations (Figure 1f). For more detailed information about genes or haplotypes, users are able to click on relevant links and a new page will appear, as shown in Figure 1e and Figure 1g. An image showing the haplotype location on chromosome bands is displayed on the left, which gives users visual indication of the haplotype location. In addition to disease related haplotypes, we provide all the other haplotypes defined by the same SNPs (or microsatellites) in the same study populations and their frequencies to users (Figure 1h). Users can also query CDRH by using combinations of disease names and chromosome numbers (Figure 1b). The results are the same as searching only by disease name.

Figure 1c shows the row called 'risk status' of the query results. It has four different values: 'risk' and 'protection' stand for haplotypes that increase or decrease, respectively, the disease risk as described in the literature; 'statistical inference risk' and 'statistical inference protection' stand for haplotypes that increase or decrease the disease risk, respectively, which were only present in the results table of an association test.

Similar to the search by disease name, users can search the database by gene name (currently supports Entrez Gene ID and Gene Symbol). This is effective in helping users directly identify haplotypes related to

a gene of interest. Users can also search the database by chromosome number. Complex disease-related haplotype-centered information is shown in the order of the online publication date of the articles. Users can track developments in the design and analysis of haplotype studies for complex human diseases on this chromosome. In addition, users can retrieve information by SNP ID (rs#). If the query SNP has been identified as being part of a haplotype in our database, the search result will be returned in a new page. The basic SNP information and the concise description of relevant references will help users better understand genetic susceptibility to complex diseases. Users can view the details of interesting items by clicking on hypertext links. Our database also preserves the search history records for each query model, which allows users to recall previous search results.

The query results obtained in different ways can be directly downloaded as an Excel file by the download link at the top of view page (Figure 1d). Furthermore, all data for complex disease-related haplotypes, as well as the corresponding analysis software, are freely available on the download page.

1.4 Submit page

We encourage users to submit information concerning complex disease-related haplotypes that are not documented. Data can be directly submitted to CDRH via the Submit Web page. Required submission information includes disease name, population, chromosome number, gene symbol, haplotype, PubMed ID, and the correspondence details of the submitters. All submissions will receive a systematic quality assurance review.

The submitted records, and other essential information, will be added in the CDRH as soon as possible if the submissions pass the above checks. The data contained in CDRH is updated regularly by manual extraction of relevant information from publications retrieved from the literature databases of PubMed. The collection of new and improved items will be displayed in the top of the browse page after each update.

2 Discusson

Understanding the relationship between genetic variation and heritable risk for complex human

diseases is a formidable challenge for modern human genetics. This is also an important step towards the discovery of genes that influence complex human diseases. To provide a central resource for molecular biologists and geneticists who study complex disease-related haplotypes, we have collected a considerable amount of information, which was scattered in existing studies, and have developed a database of complex disease-related haplotypes, CDRH. It not only offers an easy-to-use interface to query the valuable reference information concerning haplotypes and diseases extracted from the literature, but also integrates vast quantities of complementary biological annotations from external database. The CDRH database clearly reflects the relationships between haplotypes and complex diseases. Thus, it facilitates the gathering of more comprehensive information on complex disease-related haplotypes, and at the same time, saves researchers the trouble of searching multiple databases and large quantities of literatures.

Currently, 1 125 haplotypes are documented in the CDRH database, referring to 22 autosomes, the chromosome X, the chromosome Y, and the mitochondrion. Figure 2a represents a histogram of the number of complex disease-related haplotypes on each chromosome. Figure 2b represents a histogram of the number of complex disease-related genes on each chromosome. As is evident from Figure 2, the overwhelming majority of haplotypes (431 haplotypes) and genes (39 genes) are located on chromosome 6. In particular, these haplotypes and genes are mainly concentrated in the 6p21.3 (74.36%) region. Some previous studies indicated that this region is associated with many complex immune diseases, such as type 1 diabetes (Noble et al., 1996; Hermann et al., 2003), rheumatoid arthritis (Newton et al., 2004), rheumatic heart disease (Hernandez-Pacheco et al., 2003), and systemic lupus erythematosus (Vargas-Alarcon et al., 2001). These results imply that certain complex diseases share some common biomarkers and might have underlying functional interaction among predisposing genes. In the future, more studies will give us a deeper comprehension of the 6p21.3 region.

Figure 2a also indicates that there are no complex disease-related haplotypes located on chromosome 21. This phenomenon is attributable to there being no exact haplotype information for chromosome 21 in the literatures.

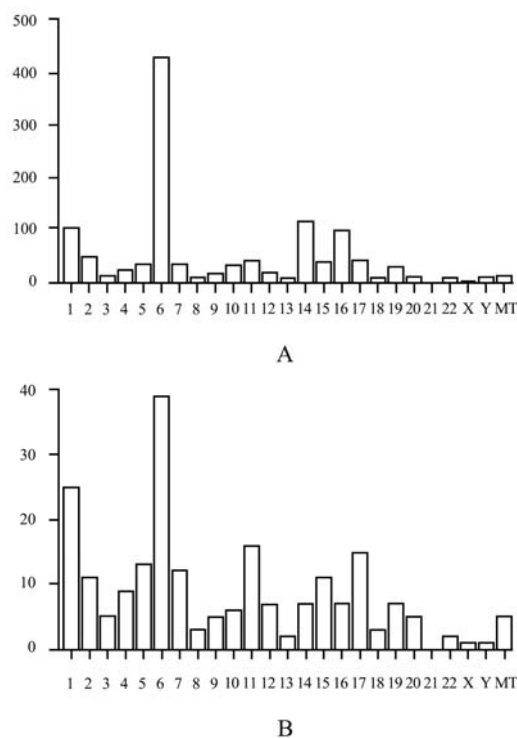


Figure 2 The chromosomal distribution of complex disease-related haplotypes and genes in the CDRH database
Note: MT; mitochondrion. (a) Histogram of the number of complex disease-related haplotypes on each chromosome. (b) Histogram of the number of complex disease-related genes on each chromosome

To date, the CDRH database has records of 114 complex diseases. Table 2 shows the statistical information of the top six complex diseases, in order of the number of haplotypes. These diseases involve at least two populations, and more than one chromosome and gene, which implies that these diseases are more common compared with the others and may be caused by multiple genes. Multiple sclerosis (Rosati, 2001) and rheumatoid arthritis (Harris, 1990) each have at least two studies in the literature in our database, which might imply that researchers should pay more attention to these diseases.

Table 2 The statistical information of the top six complex diseases in the CDRH database

Disease name	Haplotypes (n)	SNP/micro ^a (n)	References (n)	Populations (n)	Genes (n)	Chr. num. ^b
Type 1 diabetes mellitus	247	147	15	15	17	4
Multiple sclerosis	64	81	16	15	15	4
Bipolar disorder	58	30	2	2	2	2
Schizophrenia	57	49	8	9	8	8
Age-related macular degeneration	48	61	12	13	13	5
Rheumatoid arthritis	40	58	11	12	12	6

Note: SNP/micro: SNPs and microsatellites; Chr. Num: chromosome number

Haplotypes can contain more information than a single marker, and can reveal synergistic effects among SNPs. Thus, haplotypes that are responsible for some genetic disorders are being developed for molecular diagnosis of genetic disorders (especially for autosomal recessive genetic disorders). Some studies (Basel et al., 2004; Sossenheimer et al., 1997, Repiso et al., 2005, Lian et al., 2004) have indicated that haplotype analysis is highly informative for molecular disease diagnosis and carrier status. Consequently, by offering detailed information about complex disease-related haplotypes, CDRH may help in the design of future experimental and computational biology studies.

3 Conclusion

CDRH is the first database to emphasize complex human diseases at the haplotype level by collecting and cataloguing a great variety of literature. It provides a user-friendly interface to search for detailed information concerning haplotypes and diseases. We encourage researchers to submit interesting new data and offer a download function. We are committed to the maintenance and update of the CDRH database, and hope that it will guide researchers to a fuller understanding of complex human diseases.

4 Future Perspective

With the rapid improvement in SNP genotyping technology and haplotype analysis methods, we can conveniently obtain genome-wide SNP data. Thus, genome-wide association studies based on haplotypes might be an efficient way to identify genetic regions or genes that are implicated in complex diseases. Our group will closely follow the future developments in

haplotype studies of complex human diseases, and provide users with timely information. We believe that the CDRH database will provide deeper insights into the relationships between haplotypes and complex diseases.

Acknowledgments

This work was supported in part by grants from the National Natural Science Foundation of China (Grant Nos. 81172842, 31200934) and the Natural Science Foundation of Heilongjiang Province (Grant No. C201206). We thank all members of the statistical genetics workshop at the College of Bioinformatics Science and Technology, Harbin Medical University.

References

- Basel D., Kilpatrick M.W., and Tsiouras P., 2004, Haplotype analysis enables the diagnosis of Marfan syndrome, *Conn Med*, 68(6): 363-366
- Berger M., Moscatelli H., Kulle B., Luxembourg B., Blouin K., Spannagl M., Lindhoff-Last E., and Schambeck C.M., 2008, Association of ADAMDEC1 haplotype with high factor VIII levels in venous thromboembolism, *Thromb Haemost*, 99(5): 905-908
- Daly M.J., Rioux J.D., Schaffner S.F., Hudson T.J., and Lander E.S., 2001, High-resolution haplotype structure in the human genome, *Nat Genet.*, 29(2): 229-232
<http://dx.doi.org/10.1038/ng1001-229>
- Gabriel S.B., Schaffner S.F., Nguyen H., Moore J.M., Roy J., Blumenstiel B., Higgins J., Defelice M., Lochner A., Faggart M., Liu-Cordero S.N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander E.S., Daly M.J., and Altshuler D., 2002, The structure of haplotype blocks in the human genome, *Science*, 296(5576): 2225-2229
<http://dx.doi.org/10.1126/science.1069424>
PMid:12029063
- Gao G., Allison D.B., and Hoeschele I., 2009, Haplotyping methods for pedigrees, *Hum Hered*, 67(4): 248-266

- <http://dx.doi.org/10.1159/000194978>
PMid:19172084 PMCID:PMC2692835
- Hapmap, 2003, The International HapMap Project, *Nature*, 426(6968): 789-796
- Harris E.D. Jr., 1990, Rheumatoid arthritis. Pathophysiology and implications for therapy, *N. Engl. J. Med.*, 322(18): 1277-1289
<http://dx.doi.org/10.1056/NEJM199005033221805>
PMid:2271017
- Hermann R., Turpeinen H., Laine A.P., Veijola R., Knip M., Simell O., Sipila I., Akerblom H.K., and Ilonen J., 2003, HLA DR-DQ-encoded genetic determinants of childhood-onset type 1 diabetes in Finland: an analysis of 622 nuclear families, *Tissue Antigens*, 62(2): 162-169
<http://dx.doi.org/10.1034/j.1399-0039.2003.00071.x>
PMid:12889996
- Hernandez-Pacheco G., Aguilar-Garcia J., Flores-Dominguez C., Rodriguez-Perez J.M., Perez-Hernandez N., Alvarez-Leon E., Reyes P.A., and Vargas-Alarcon G., 2003, MHC class II alleles in Mexican patients with rheumatic heart disease, *Int. J. Cardiol.*, 92: 49-54
[http://dx.doi.org/10.1016/S0167-5273\(03\)00040-8](http://dx.doi.org/10.1016/S0167-5273(03)00040-8)
- Higasa K., Miyatake K., Kukita Y., Tahira T., and Hayashi K., 2007, D-HaploDB: a database of definitive haplotypes determined by genotyping complete hydatidiform mole samples, *Nucleic Acids Res.*, 35: D685-689
<http://dx.doi.org/10.1093/nar/gkl848>
PMid:17166862 PMCID:PMC1781173
- Hung H.C., Lin C.Y., Liao Y.F., Hsu P.C., Tsay G.J., and Liu G.Y., 2007, The functional haplotype of peptidylarginine deiminase IV (S55G, A82V and A112G) associated with susceptibility to rheumatoid arthritis dominates apoptosis of acute T leukemia Jurkat cells, *Apoptosis*, 12(3): 475-487
<http://dx.doi.org/10.1007/s10495-006-0005-0>
PMid:17216583
- Ingman M., and Gyllensten U., 2006, mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences, *Nucleic Acids Res.*, 34: D749-751
<http://dx.doi.org/10.1093/nar/gkj010> PMid:16381973
- Kayser M., Brauer S., Willuweit S., Schadlich H., Batzer M.A., Zawacki J., Prinz M., Roewer L., and Stoneking M., 2002, Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations, *J. Forensic Sci.*, 47: 513-519
- Lian J.F., Cui C.C., Xue X.L., Huang C., Cui H.B., and Zhang H.Z., 2004, Long QT syndrome gene diagnosis by haplotype analysis, *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, 21: 272-273
- Lin D.Y., and Zeng D., 2006, Likelihood-Based Inference on Haplotype Effects in Genetic Association Studies, *Journal of the American Statistical Association*, 101: 104-106
<http://dx.doi.org/10.1198/016214505000000808>
- Newton J.L., Harney S.M., Timms A.E., Sims A.M., Rockett K., Darke C., Wordsworth B.P., Kwiatkowski D., and Brown M.A., 2004, Dissection of class III major histocompatibility complex haplotypes associated with rheumatoid arthritis, *Arthritis Rheum*, 50: 2122-2129
<http://dx.doi.org/10.1002/art.20358>
- Noble J.A., Valdes A.M., Cook M., Klitz W., Thomson G., and Erlich H.A., 1996, The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families, *Am. J. Hum. Genet.*, 59(5): 1134-1148
PMid:8900244
- Repiso A., Corrons J.L., Vulliamy T., Killeen N., Layton M., Carreras J., and Climent F., 2005, New haplotype for the Glu104Asp mutation in triose-phosphate isomerase deficiency and prenatal diagnosis in a Spanish family, *J. Inherit. Metab. Dis.*, 28(5): 807-809
<http://dx.doi.org/10.1007/s10545-005-0098-6>
PMid:16151918
- Rosati G., 2001, The prevalence of multiple sclerosis in the world: an update, *Neurol Sci*, 22(2): 117-139
<http://dx.doi.org/10.1007/s100720170011>
PMid:11603614
- Santiago J.L., Martinez A., Nunez C., De La Calle H., Fernandez-Arquero M., De La Concha E.G., and Urcelay E., 2008, Association of MYO9B haplotype with type 1 diabetes, *Hum. Immunol.*, 69(2): 112-115
<http://dx.doi.org/10.1016/j.humimm.2008.01.003>
- Slattery M.L., Curtin K., Sweeney C., Wolff R.K., Baumgartner R.N., Baumgartner K.B., Giuliano A.R., and Byers T., 2008, Modifying effects of IL-6 polymorphisms on body size-associated breast cancer risk, *Obesity (Silver Spring)*, 16(2): 339-347
<http://dx.doi.org/10.1038/oby.2007.44>
PMCID:PMC2925502
- Soma H., Yabe I., Takei A., Fujiki N., Yanagihara T., and Sasaki H., 2008, Associations between multiple system atrophy and polymorphisms of SLC1A4, SQSTM1, and EIF4EBP1 genes, *Mov. Disord*, 23(8): 1161-1167
<http://dx.doi.org/10.1002/mds.22046>
- Sossenheimer M.J., Aston C.E., Preston R.A., Gates L.K., Jr., Ulrich C.D., Martin S.P., Zhang Y., Gorry M.C., Ehrlich

- G.D., and Whitcomb D.C., 1997, Clinical characteristics of hereditary pancreatitis in a large family, based on high-risk haplotype. The Midwest Multicenter Pancreatic Study Group (MMPSTG), *Am. J. Gastroenterol.*, 92(7): 1113-1116
- Tishkoff S.A., Pakstis A.J., Ruano G., and Kidd K.K., 2000, The accuracy of statistical methods for estimation of haplotype frequencies: an example from the CD4 locus, *Am. J. Hum. Genet.*, 67(2): 518-522
<http://dx.doi.org/10.1086/303000> PMID:PMC1287198
- Vargas-Alarcon G., Salgado N., Granados J., Gomez-Casado E., Martinez-Laso J., Alcocer-Varela J., Arnaiz-Villena A., and Alarcon-Segovia D., 2001, Class II allele and haplotype frequencies in Mexican systemic lupus erythematosus patients: the relevance of considering homologous chromosomes in determining susceptibility, *Hum. Immunol.*, 62(8): 814-820
[http://dx.doi.org/10.1016/S0198-8859\(01\)00267-1](http://dx.doi.org/10.1016/S0198-8859(01)00267-1)
- Yaspan B.L., McCreynolds K.M., Elmore J.B., Breyer J.P., Bradley K.M., and Smith J.R., 2008, A haplotype at chromosome Xq27.2 confers susceptibility to prostate cancer, *Hum. Genet.*, 123(4): 379-386
<http://dx.doi.org/10.1007/s00439-008-0486-8> PMID:PMC2811403
- Zhao H., Pfeiffer R., and Gail M.H., 2003, Haplotype analysis in population genetics and association studies, *Pharmacogenomics*, 4(2): 171-178
<http://dx.doi.org/10.1517/phgs.4.2.171.22636>

(a) Search by disease name

Complex disease-related haplotypes can be searched by the disease name, which sorted by alphabetical order in the drop-down list box. After submission of the disease name, haplotypes centered information will be displayed in a new page, including the brief description of disease, all complex disease-related haplotypes, corresponding literature references etc.

(b) Search by chromosome & disease

Complex disease-related haplotypes can be retrieved by the combinations according to chromosome number. The output results are the same as searching only by disease name.

(c) Search by disease name

Please select the disease name you're interested in!

Disease name: Colorectal cancers

submit

History records: Colorectal cancers

(d) Download page

The download link of Colorectal cancers Query Results

(e) Annotation information about: ABCB1

Entrez ID: 5243
 Gene symbol: ABCB1
 Gene fullname: ATP-binding cassette, sub-family B (MDR/TAP), member 1
 UniGene: Hs.489033
 RefseqRNA: NM_000927
 Chromosome number: 7
 Location: 7q21.1
 Start position: 86970884
 End position: 87180506
 OMIM: 120080 171050

(f) Summary:

The membrane-associated protein encoded by this gene is a member of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes. ABC genes are divided into seven distinct subfamilies (ABC1, MDR/TAP, MRP, ALD, QABP, GCM20, White). This protein is a member of the MDR/TAP subfamily. Members of the MDR/TAP subfamily are involved in multidrug resistance. The protein encoded by this gene is an ATP-dependent drug efflux pump for xenobiotic compounds with broad substrate specificity. It is responsible for decreased drug accumulation in multidrug-resistant cells and often mediates the development of resistance to anticancer drugs. This protein also functions as a transporter in the blood-brain barrier.

(g) Haplotype information about: TT

Haplotype: TT
 Location: 7q21.12
 Related gene: ABCB1
 Related marks: rs1128503 (Build 37.1 Physical position: 87017537)
 rs2032582 (Build 37.1 Physical position: 86998554)
 Risk/protective: protection
 P-value/OR: 0.0064
 Haplotype frequency: Haplotype frequency (male cohort): 4.7% More details

(h) Table 4. P values, and odds ratios for disease association study between MSI CRC patients and haplotypes consisted of rs1128503, rs2032582.

Haplotypes	Controls	MSI-H tumors	P-value	Odds ratio
C-G	0.589	0.431	0.012	1.896
T-T	0.374	0.556	0.003	0.477
C-T	0.023	0.014	1.000	1.671

Figure 1 The results of searching by 'colorectal cancer'

Note: (a) Searching with 'colorectal cancer'. (b) Searching with 'colorectal cancer' and 'chromosome 7'. (c) The list of results retrieved by searching with 'colorectal cancer'. (d) The download page of retrieved results. (e) The detailed information obtained by clicking the hyperlink of the gene symbol 'ABCB1'. (f) The description of the 'Slovenia population'. (g) Detailed information obtained by clicking the hyperlink for haplotype 'TT'. (h) All the other haplotypes defined by the same SNPs in the same study populations and their frequencies.