



GWAS Central: a database for the integration and comparative interrogation of genome-wide association study datasets

URL: <http://www.gwascentral.org>

Presenting author: Sirisha Gollapudi

Co-authors: Robert C. Free, Robert K. Hastings, Tim Beck, Gudmundur A. Thorisson, Anthony J. Brookes

How GWAS Central benefits biologists:

Genome-wide association studies (GWAS) are an increasingly popular method for establishing associations between genotypes and common diseases, drug responses and other complex phenotypes. In most cases, only a small percentage of summary-level GWAS data is made publicly available, which does not provide a comprehensive picture of potential associations, and the problem is compounded by the lack of sophisticated bioinformatics platforms for the storage and interrogation of these data sets.

We have therefore created GWAS Central, a database of genotype-to-phenotype relationships built upon a basal layer of markers (all known SNPs and other variants). It contains summary-level genetic association significance findings from over 700 GWAS, sourced from public repositories, researchers and consortia, and through direct submission. Experimental biologists interested in a particular phenotype (which in GWAS Central are on a broad spectrum from disease descriptions to individual signs and symptoms), gene, genomic region or marker can take advantage of an advanced toolkit for searching and analysing relevant data sets, including: powerful graphical meta-analysis for interrogation of positive and negative association results; custom uploads; search result and report exports.

Comparison to similar tools:

1. NHGRI GWAS catalog - <http://www.genome.gov/26525384>
2. dbGaP - <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>
3. GWAS DB - https://gwas.lifesciencedb.jp/cgi-bin/gwasdb/gwas_top.cgi

The data available in GWAS Central are significantly more comprehensive than those available via the three resources listed above, with over 21 million p-values. We do not set a significance threshold for p-values, and store all the available summary-level data obtained as part of each study, rather than a minimal number of 'top' p-values. The storage and integration of seemingly insignificant results is very important when evaluating the biological significance of a GWAS. Besides the extensive data content, GWAS Central has a complement of powerful graphical meta-analysis tools, advanced search and data mining features. Our sophisticated genome browser also offers significant advantages when compared to the browsers of other applications, with fully integrated genome and region views, and also provides users with the facility to temporarily upload and view their own data in the context of existing studies using the 'Custom Track' function - a feature not provided by any other GWAS browser. Once the user has selected and/or uploaded data sets of interest, signals can be filtered by significance, and viewed alongside RefSeq genes, HapMap SNPs and pair-wise LD plots, and Human Genetic Mutation Database (HGMD) variants.