

### 3. Entrez Gene search tool

Annotation databases such as DAVID and ConceptGen provide an excellent resource for performing functional enrichment on a large group of genes. However, they do not provide details on how a particular gene is related to a specific term. For example, if a transcription factor is annotated with a term like “Glucose Metabolism”, it is unclear if the transcription factor activates or inhibits glucose metabolism genes, or which genes are the target genes, and what their role in glucose metabolism is.

A custom script was developed to provide more detailed information on the link between a gene and a biological function using NCBI Entrez Utils (<http://eutils.ncbi.nlm.nih.gov>). The script searches all of the fields of the Entrez Gene entries (i.e. name, description, GeneRIF, etc.) of a set of genes for certain user-specified search terms, for example “glucose metabolism”, and returns a list of genes whose entries contain that term. However, there may be many false positives in that list, so individual Entrez Gene entries and the literature must be manually examined to determine the actual relationship between the gene and the search terms. This method does not automatically decipher the relationship between a gene and a function, but only provides a shortened list of genes for which it is now feasible to manually read the entries to determine the relationship between the gene and the term.