**GenGen**

GenGen is a suite of free software tools to facilitate the analysis of high-throughput genomics data sets. The package is currently a work-in-progress and infrequently updated.

Some of the most useful and more stable programs in this package are:

* **calculate\_association.pl**: perform genome-wide association analysis for cohorts with case-control design and trio design, with ability to handle PLINK-formatted TPED/TFAM files and handle phenotype permutations.
* **calculate\_gsea.pl** combine\_gsea.pl: use pathway-based approaches to test the significane of genes within a pathway for genome-wide association studies (see reference below).
* scan\_region.pl: scan two sets of genomic regions to find overlap (for example, scan most conserved genomic elements, scan transcription factor binding sites, scan segmental duplication regions, scan predicted microRNA sites, scan pseudogenes, scan EvoFold regions, etc.
* convert\_mach.pl, convert\_mach2snptest.pl, combine\_snptest.pl: convert MACH-imputed genotype data sets into standard PED/MAP format or SNPTEST format for subsequent association test.
* convert\_bim\_allele.pl: convert SNP allele coding between Illumina A/B alleles, Illumina 1/2 alleles, Illumina TOP strand alleles and dbSNP forward strand alleles.
* retrieve\_ensembl.pl, retrieve\_variation.pl, retrieve\_compara.pl: query the Ensembl databases to get the desired information on genes, transcripts, variations, cross-references, domains, predicted regulatory elements and so on.

**Reference**

* Wang K, Li M, Bucan M. [Pathway-based approaches for analysis of genome-wide association studies](http://www.cell.com/ajhg/abstract/S0002-9297(07)63775-6). ***American Journal of Human Genetics***, 81:1278-1283, 2007
* Wang K, Li M, Hakonarson H. [Analysing biological pathways from genome-wide association studies](http://www.nature.com/nrg/journal/v11/n12/full/nrg2884.html). ***Nature Reviews Genetics***, 11:843-854, 2010