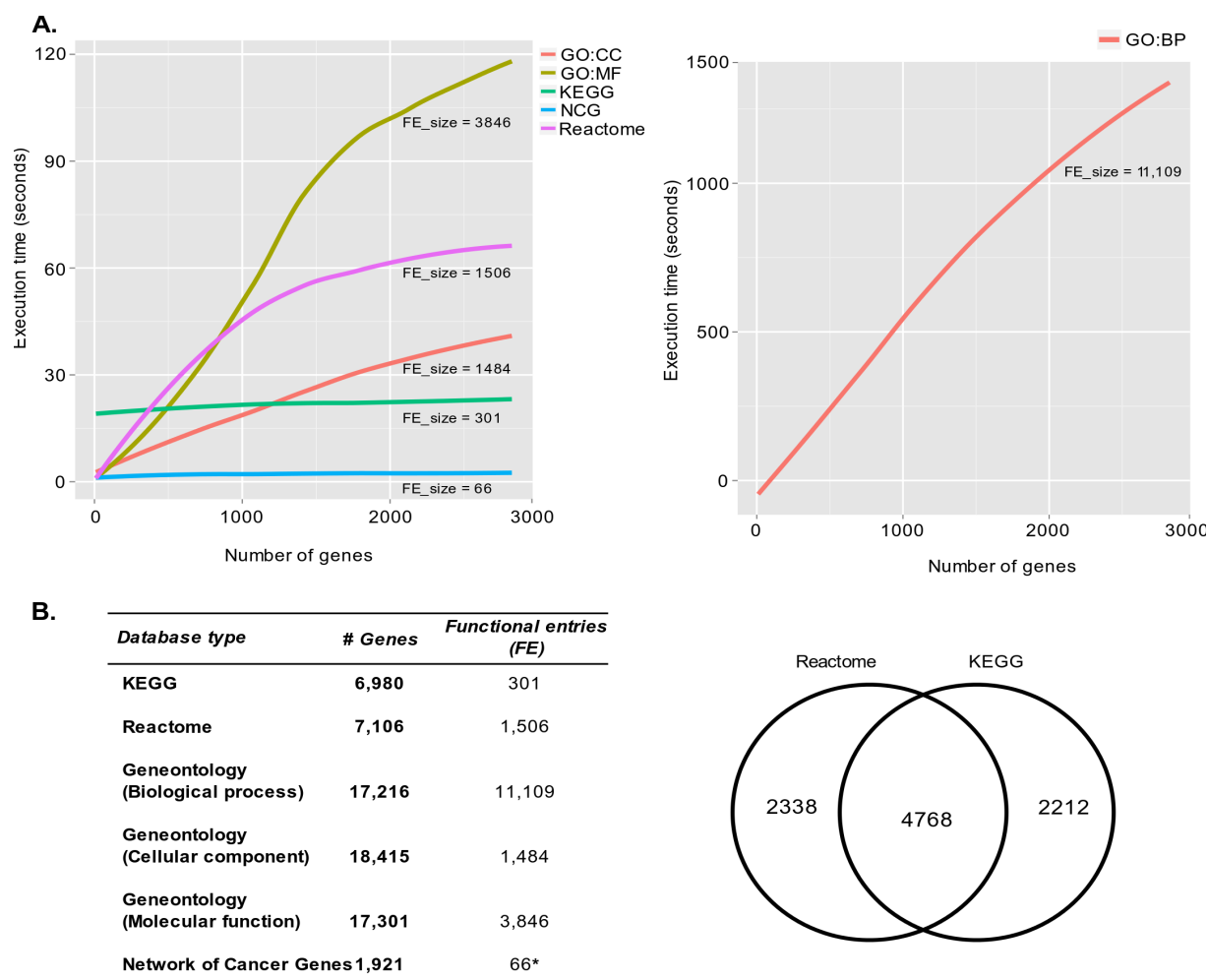


- Subhash, S. & Kanduri, C. GeneSCF: a real-time based functional enrichment tool with support for multiple organisms. *BMC Bioinformatics* **17**, 365 (2016).



Performance and database size¹

The above case studies have been performed with Intel(R) Core(TM) i7-4600U CPU at 2.10GHz machine (GeneSCF executed on single core every time). GeneSCF was also simulated with different number of genes using different databases and the execution time was monitored. For this purpose we used the same gene list of 2,824 genes from the first case study (**Additional file 3, see link below**) on CLL. This simulation concludes that the GeneSCF program execution time depends on the number of functions in the database and number of genes in the input list but not the number of genes covered in the database (**Figure A**). Since both KEGG and Reactome covers similar kind of functional information and has equal number of genes (**Figure B**), we compared the execution time between them. The KEGG out performed well due to its nature of the database which has concise functional terms compared to the narrow verbose terms in Reactome. Even though Reactome has similar number of genes but due to its narrow terms makes the database larger compared to KEGG and thus the execution time increases (**Figure B**).

Access input genelist: https://static-content.springer.com/esm/art%3A10.1186%2Fs12859-016-1250-z/MediaObjects/12859_2016_1250_MOESM3_ESM.xls