**INSILICO SYSTEMS BIOLOGY PRACTICALS**

**EXPERIMENT -19**

**AIM**- To explore David Database to interpret the biological mechanisms associated with large gene lists.

**Theory-**

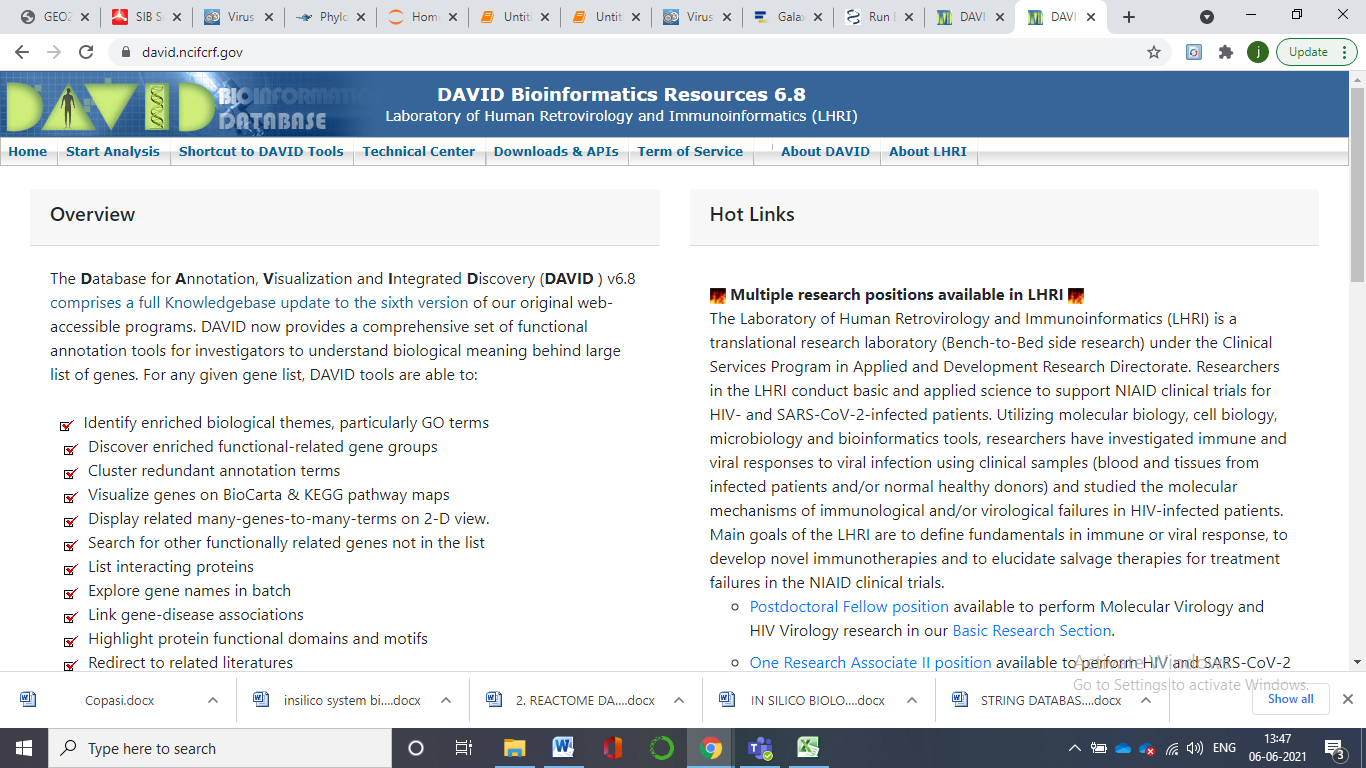
DAVID bioinformatics resources is an integrated biological knowledgebase and analytical tool aimed at extracting biological meaning from large gene/protein lists generated from a variety of High throughput genomic experiments. The Database for Annotation, Visualization and Integrated Discovery (DAVID ) v6.8 comprises a full Knowledgebase update to the sixth version of our original web-accessible programs. DAVID now provides a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes. The expanded DAVID Knowledgebase now integrates almost all major and well-known public bioinformatics resources centralized by the DAVID Gene Concept, a single-linkage method to agglomerate tens of millions of diverse gene/protein identifiers and annotation terms from a variety of public bioinformatics databases. For any uploaded gene list, the DAVID Resources now provides not only the typical gene-term enrichment analysis, but also new tools and functions that allow users to condense large gene lists into gene functional groups, convert between gene/protein identifiers, visualize many-genes-to-many-terms relationships, cluster redundant and heterogeneous terms into groups, search for interesting and related genes or terms, dynamically view genes from their lists on bio-pathways and more. For any given gene list, DAVID tools are able to:

* Identify enriched biological themes, particularly GO terms
* Discover enriched functional-related gene groups
* Cluster redundant annotation terms
* Visualize genes on BioCarta & KEGG pathway maps
* Display related many-genes-to-many-terms on 2-D view.
* Search for other functionally related genes not in the list
* List interacting proteins
* Explore gene names in batch
* Link gene-disease associations
* Highlight protein functional domains and motifs
* Redirect to related literatures
* Convert gene identifiers from one type to another.

And more

**PROTOCOL-**

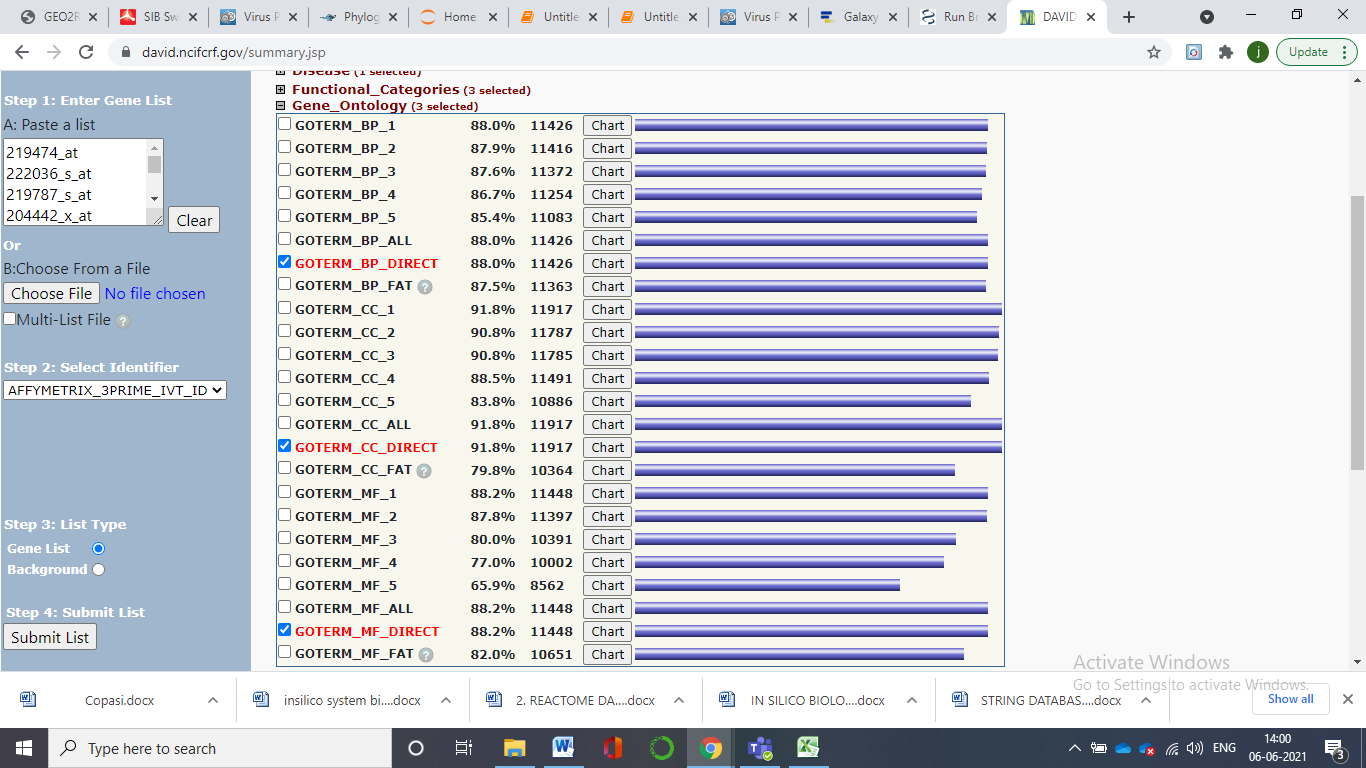
1. Open the home page for the DAVID resources site using <http://david.abcc.ncifcrf.gov>



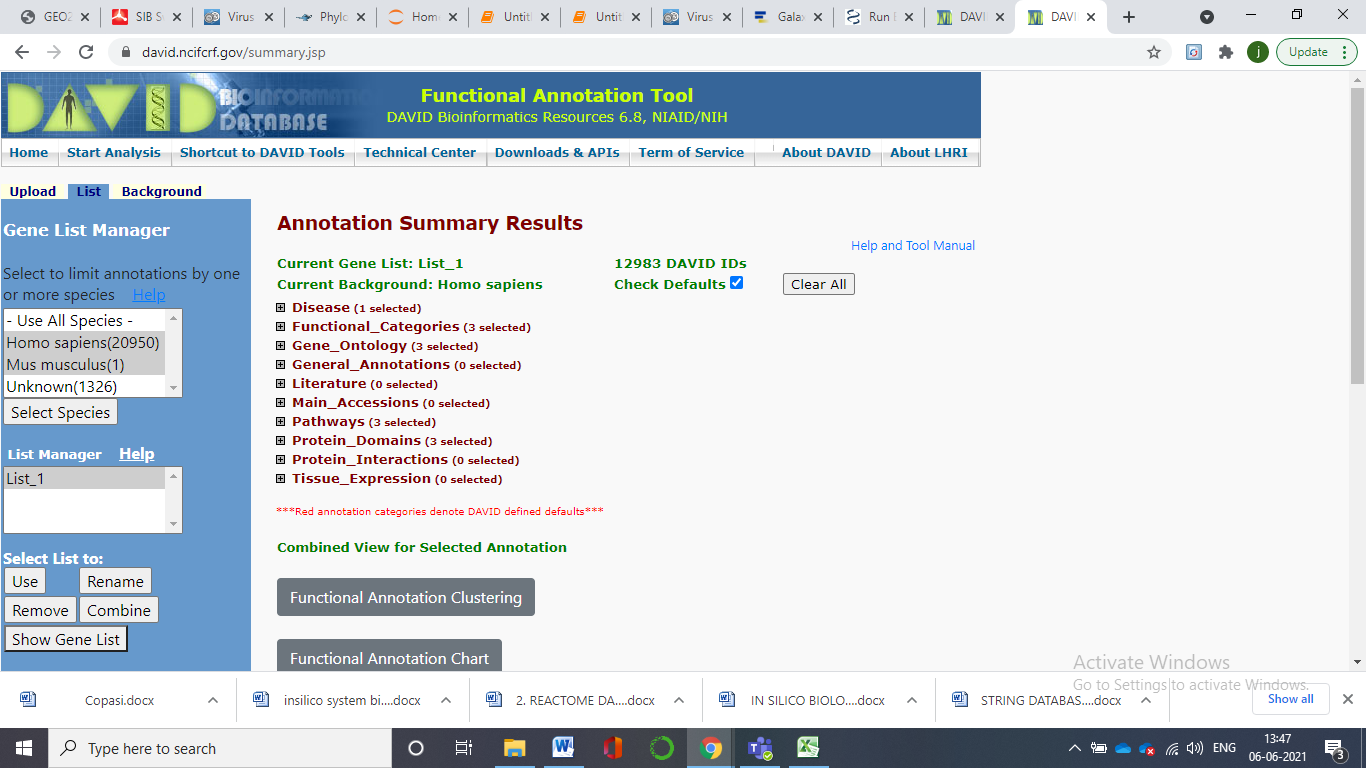
2. Go to Functional Annotation

3. Paste gene sequence in the box on the left side or upload a gene sequence from the device folder. Select the type of sequence submitted and type of list. Further submit the sequence. (NOTE : One gene could have several different identifiers within one or more database(s). Similarly, the biological terms associated with different gene identifiers for the same gene could be collected in different levels across different databases. Due to these issues, most high-throughput annotation tools rely on one, or at most a few, resource(s), which limits the analytic comprehensiveness and the level of throughput. The DAVID Knowledgebase is now built around the ‘DAVID Gene Concept’, a single linkage method to agglomerate

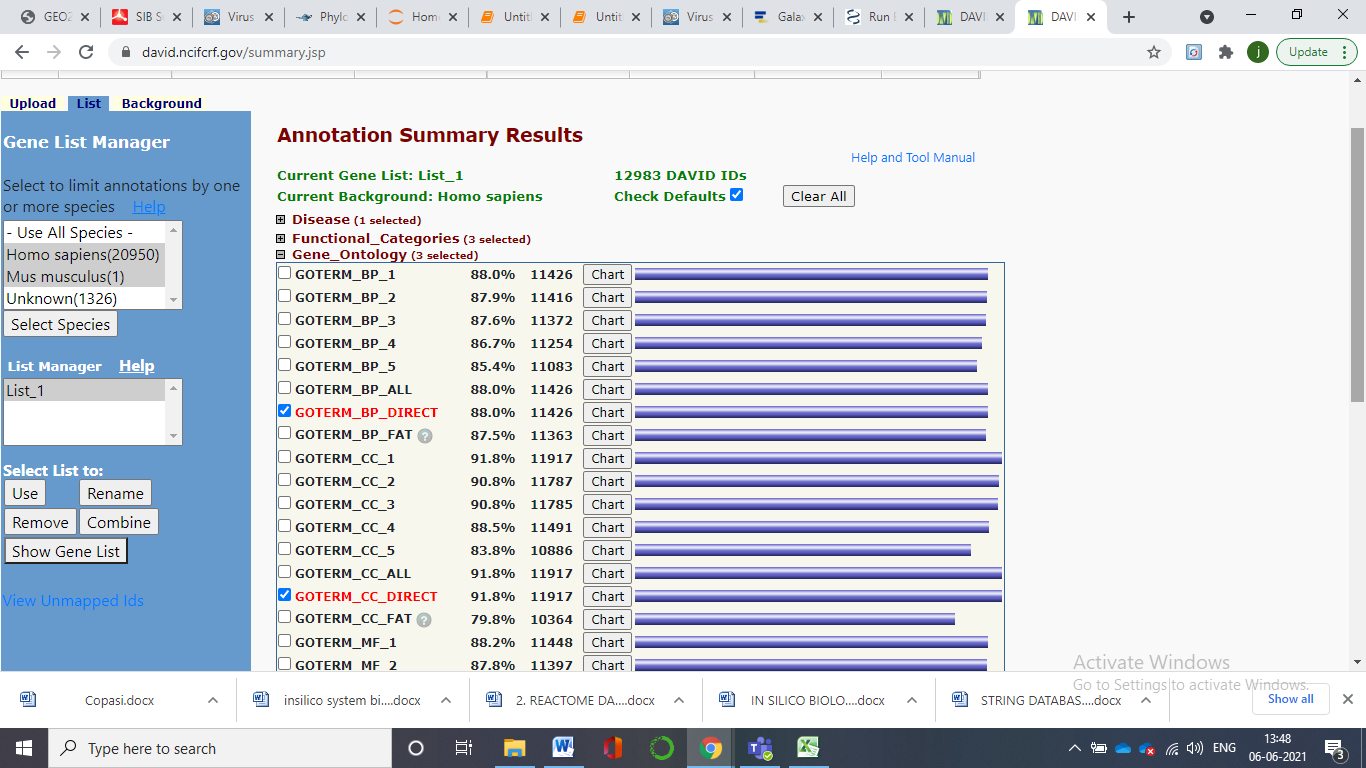
tens of millions of gene/protein identifiers from a variety of public genomic resources including NCBI, PIR and UniProt.) Over 22 types of gene identifiers integrated by the DAVID Gene Concept within the DAVID Knowledgebase.



4. Results are displayed consisting of all the categorizations.. Now chose the organism you need to work with. This function uses a novel algorithm to measure relationships among the annotation terms based on the degrees of their co-association genes to group the similar, redundant and heterogeneous annotation contents from the same or different resources into annotation groups. This reduces the burden of associating similar redundant terms and makes the biological interpretation more focused in a group level. (The gene names are hot links to pages with information about the genes. The RG links will display in a new tab, a list of other genes within the submitted list that are related to the gene you clicked on. These relationships are based on a statistical measure of functional similarity. )



5. Click on the Plus symbol and expand the categorization results.

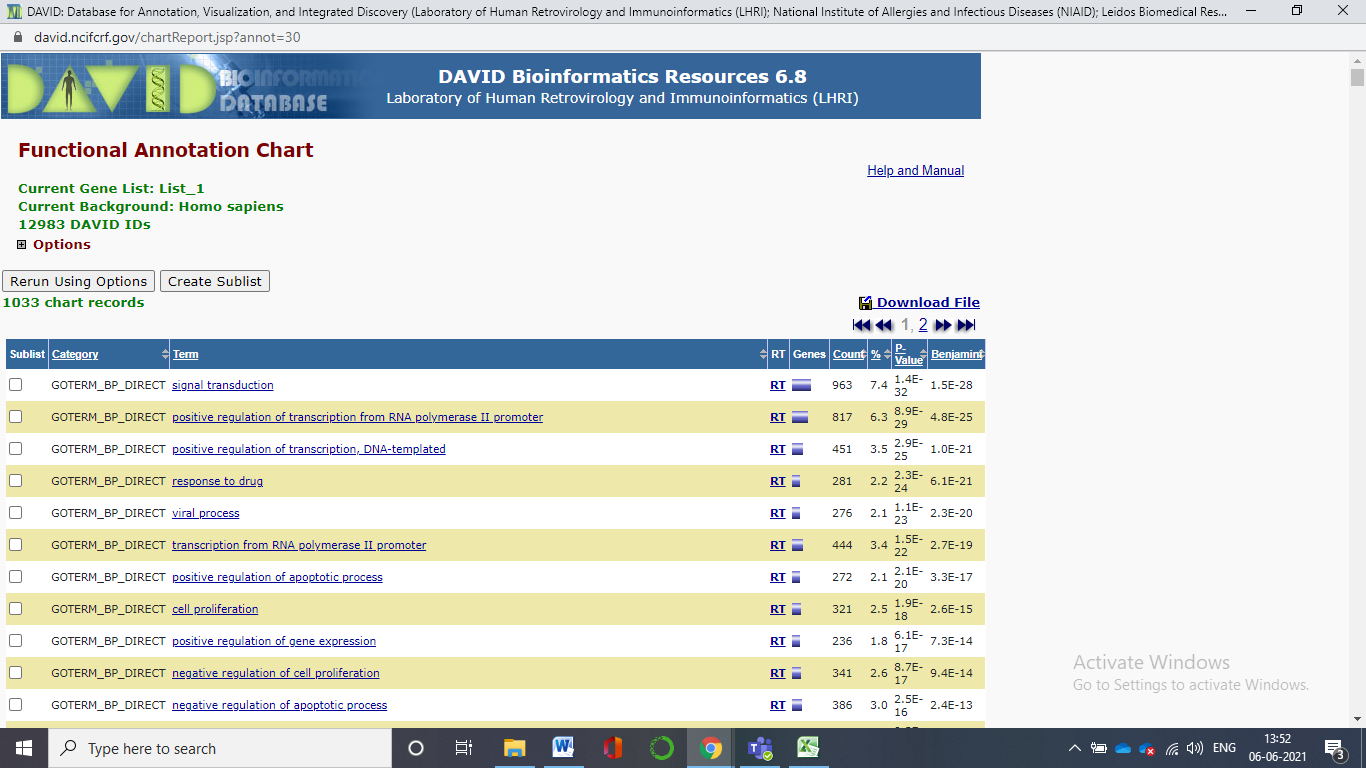


Here we can see that in the Genetic Association Database (GAD), there are 88 % (11426 of them) of our chainlist that overlap or map with genetic association database.

6. Click on the blue bar. It will take you to the 11426 disease database.



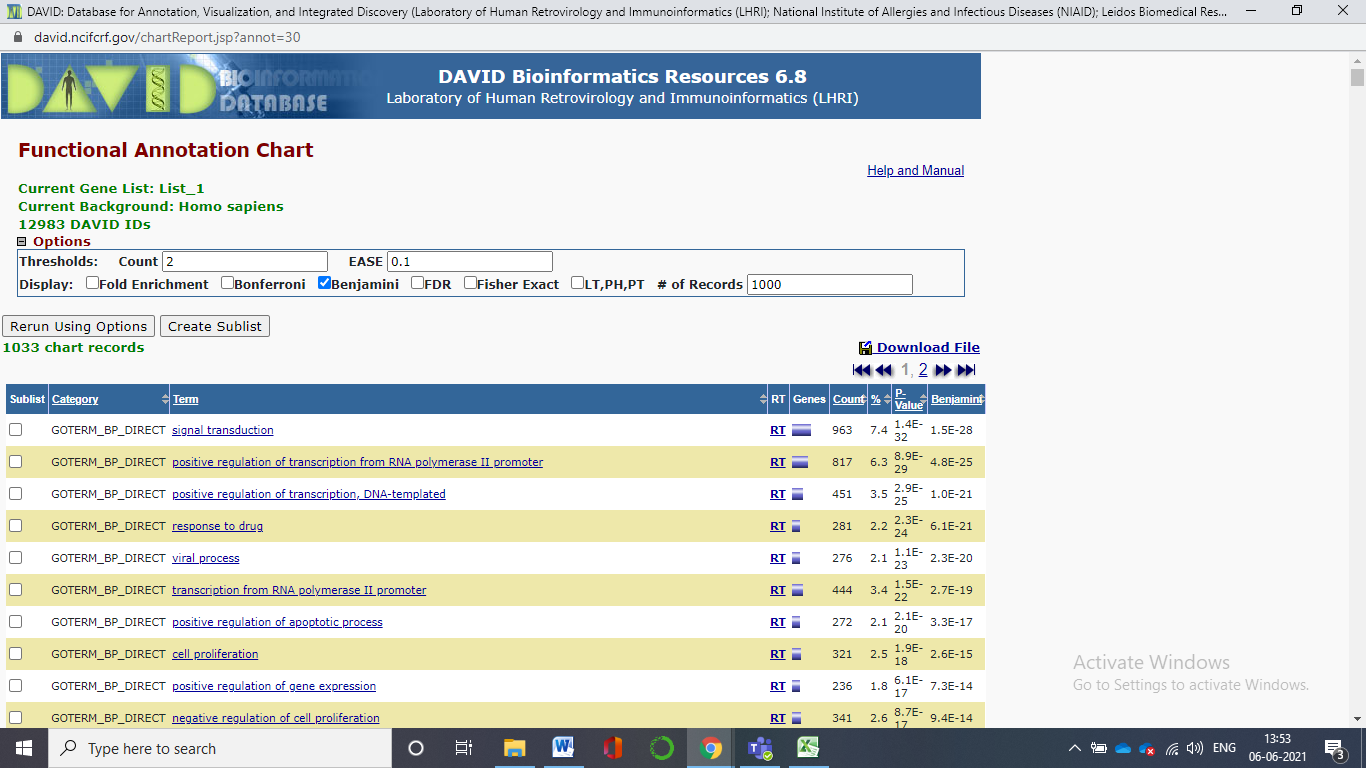
7. Clicking on the chart button given us the enrichment analysis results.



8. The plus symbol in the options can help us edit the requirements. We can adjust the threshold value, fold enrichment, number of records etc. The new DAVID Gene Name Batch Viewer is designed to simply list the gene names for all given genes. In addition, hyperlinks are provided on each gene entry, allowing users to explore in depth other functional information around the gene. Thus, this tool provides users with a first glance and initial ideas about their interesting genes before proceeding to analysis by other more comprehensive analytic tool. Moreover, hyperlinks, labeled as ‘RT’, are provided for each

gene in order to search other functionally related genes in user's gene list or the entire genome.

(instead of a long list of single terms, you see clusters of terms that are related to each other. The clusters are shown in order of decreasing enrichment scores. If you click on the red G button located on the blue bar at the top of each cluster, you should get a list of the genes in that cluster that are annotated with one or more of the terms.)



9. In the first mentioned Biological Process is signal transduction -The cellular process in which a signal is conveyed to trigger a change in the activity or state of a cell. Signal transduction begins with reception of a signal (e.g. a ligand binding to a receptor or receptor activation by a stimulus such as light), or for signal transduction in the absence of ligand, signal-withdrawal or the activity of a constitutively active receptor. Signal transduction ends with regulation of a downstream cellular process, e.g. regulation of transcription or regulation of a metabolic process. Signal transduction covers signalling from receptors located on the surface of the cell and signalling via molecules located within the cell. For signalling between cells, signal transduction is restricted to events at and within the receiving cell. By clicking on the list another tab. The download option in that file help us to download the results.

