

Supplementary Figure 2: ToppGene Suite workflow

A

The screenshot shows the ToppGene Suite web interface. The main heading is "ToppGene Suite" with a subtitle: "A one-stop portal for gene list enrichment analysis and candidate gene prioritization based on functional annotations and protein interactions network". Below this, it specifies "ToppFun: Transcriptome, ontology, phenotype, proteome, and pharmacome annotations based gene list functional enrichment analysis".

The input section is titled "Select your gene identifier type, paste your sets below or select example set, then submit." It includes a dropdown for "Entry Type" (set to "HGNC Symbol") and "Example gene sets" (HGNC Symbol, Entrez ID). A large text area labeled "Training Gene Set" is highlighted with a red box and an arrow pointing to it, with the text "Paste the genes here". Below the text area are "Clear" and "Submit" buttons.

B

The screenshot shows the "Results" page of the ToppGene Suite. It displays enrichment analysis results for four categories: Molecular Function, Biological Process, Cellular Component, and Human Phenotype. Each category includes a table of results with columns for ID, Name, Source, pValue, FDR BSM, FDR BSV, Bonferroni, Genes from Input, and Genes in Annotation.

1: GO: Molecular Function [Display Chart] 3360 input genes in category / 2708 annotations before applied cutoff / 18091 genes in category

ID	Name	Source	pValue	FDR BSM	FDR BSV	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0003700 DNA-binding transcription factor activity	6.559E-4	1.171E-4	1.457E-3	1.834E-4	305	1294	
2	GO:0046873 metal ion transmembrane transporter activity	1.223E-7	9.171E-4	1.457E-3	3.423E-4	120	426	
3	GO:0043585 sequence-specific DNA binding	7.866E-6	4.801E-3	4.088E-2	1.977E-2	253	1096	
4	GO:0019787 ubiquitin-like protein transferase activity	7.721E-6	4.801E-3	4.088E-2	2.160E-2	116	441	
5	GO:0011667 regulatory region nucleic acid binding	1.827E-5	4.801E-3	4.088E-2	2.875E-2	205	886	

2: GO: Biological Process [Display Chart] 3365 input genes in category / 8995 annotations before applied cutoff / 19623 genes in category

ID	Name	Source	pValue	FDR BSM	FDR BSV	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0048738 cardiac muscle tissue development	3.709E-6	3.709E-4	3.629E-3	3.709E-4	72	215	
2	GO:0048729 tissue morphogenesis	2.122E-7	9.171E-4	8.976E-3	2.121E-3	183	782	
3	GO:0005937 muscle tissue development	3.689E-7	9.171E-4	8.976E-3	3.691E-3	119	415	
4	GO:0007165 enzyme linked receptor protein signaling pathway	3.996E-7	9.171E-4	8.976E-3	3.996E-3	240	1016	
5	GO:0007007 heart development	4.208E-7	9.171E-4	8.976E-3	4.208E-3	155	580	

3: GO: Cellular Component [Display Chart] 3410 input genes in category / 1185 annotations before applied cutoff / 15961 genes in category

ID	Name	Source	pValue	FDR BSM	FDR BSV	Bonferroni	Genes from Input	Genes in Annotation
1	GO:0005756 Golgi stack	5.698E-6	5.453E-2	4.175E-2	6.752E-3	45	132	
2	GO:0005768 endosome	9.204E-6	5.453E-2	4.175E-2	1.091E-2	166	526	
3	GO:0044431 Golgi apparatus part	2.803E-6	1.623E-2	7.035E-2	3.072E-2	214	528	
4	GO:0000139 Golgi membrane	3.452E-6	1.623E-2	7.035E-2	4.809E-2	170	716	
5	GO:0011884 organelle subcompartment	5.322E-6	1.261E-2	8.016E-2	6.109E-2	90	341	

4: Human Phenotype [Display Chart] 385 input genes in category / 1818 annotations before applied cutoff / 4377 genes in category

ID	Name	Source	pValue	FDR BSM	FDR BSV	Bonferroni	Genes from Input	Genes in Annotation
1	HP:0100687 Brachydactyly (hand)	7.737E-6	1.680E-2	1.690E-1	4.764E-2	91	319	
2	HP:0001056 Brachydactyly	6.278E-6	1.680E-2	1.690E-1	6.016E-2	127	480	
3	HP:0011927 Short digit	6.278E-6	1.680E-2	1.690E-1	6.016E-2	127	480	
4	HP:0100632 Pulmonary sequestration	2.044E-6	2.384E-2	2.219E-1	1.205E-1	19	39	
5	HP:0000756 Delayed speech and language development	2.255E-6	2.384E-2	2.219E-1	1.365E-1	109	407	