Operational Text and Data Mining

- Perspective from a Research Scientist-

Dr. ADG de Roos
Brussels, 26 March 2015
WG Intellectual Property Rights and Copyright Reform

The scientist's perspective

- Albert de Roos
  - PhD in Life Sciences (Nijmegen, 1997)
  - Laboratory, Computer Research, Text and data mining
  - Roles in TDM pipeline as contributor and end user

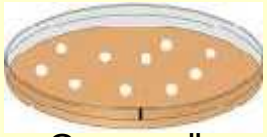
- Goal is to show TDM from a scientist's perspective
  - Using state-of-the-art genomics tool (DAVID)
  - How it is built by the scientific community
  - How it used by research scientists

A typical research question in cancer research

Why are some cells resistant to chemotherapy?

Laboratory → **Gene Analysis** → **Clinical Relevance**

Normal cells



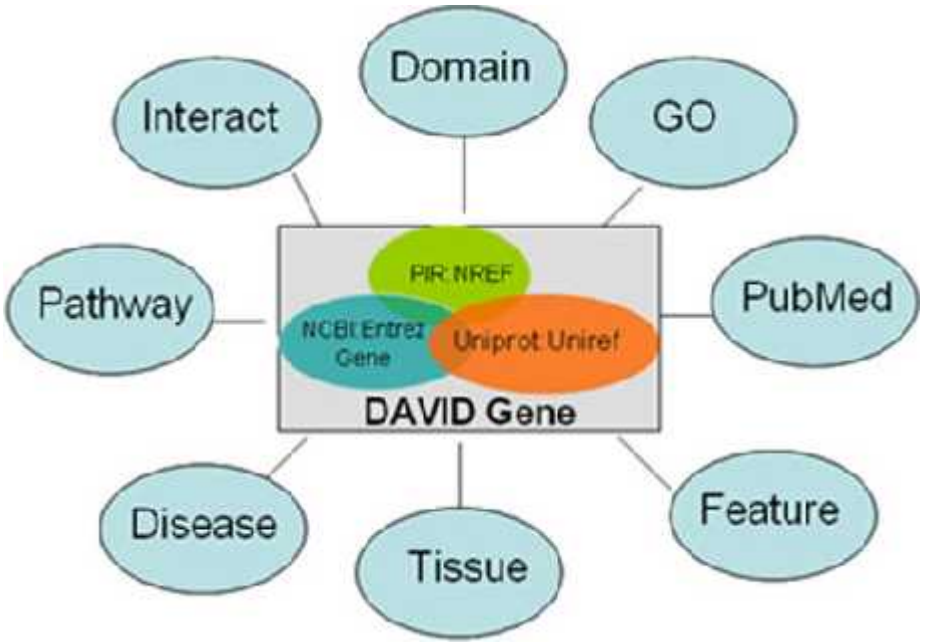
Cancer cells

Genes involved in multidrug resistance

- Cell Biology
- Tumour Biology
- Drug targets
- Diseases

Protein p150 as a new target for multidrug-resistant cells

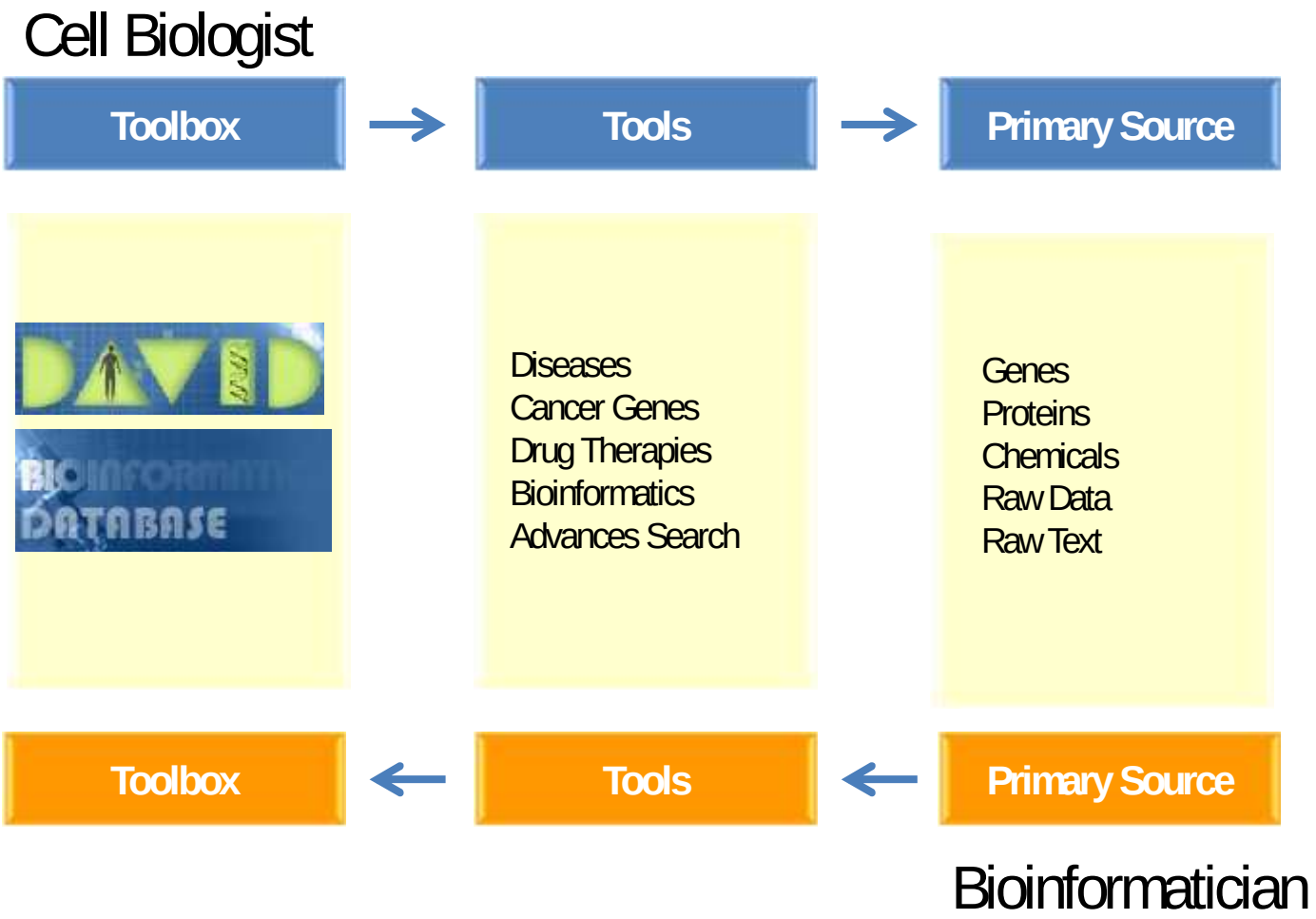
DAVID*: From Gene to Disease



- There is too much information to be analyzed by the individual researcher
- Set of tools for investigators to understand biological meaning behind large list of genes.
- How can I relate the difference in gene patterns to multidrug cancer resistance

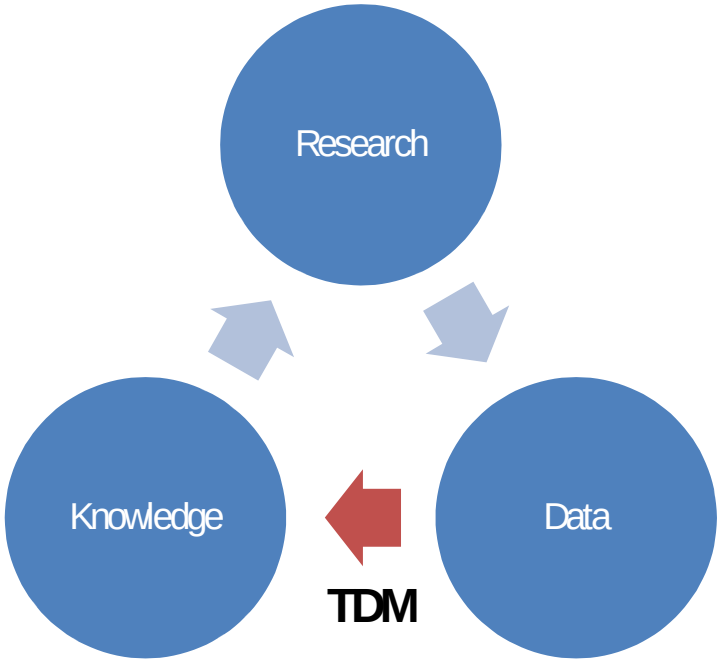
*The Database for Annotation, Visualization and Integrated Discovery (DAVID)

Cell Biology and Bioinformatics



Next research steps

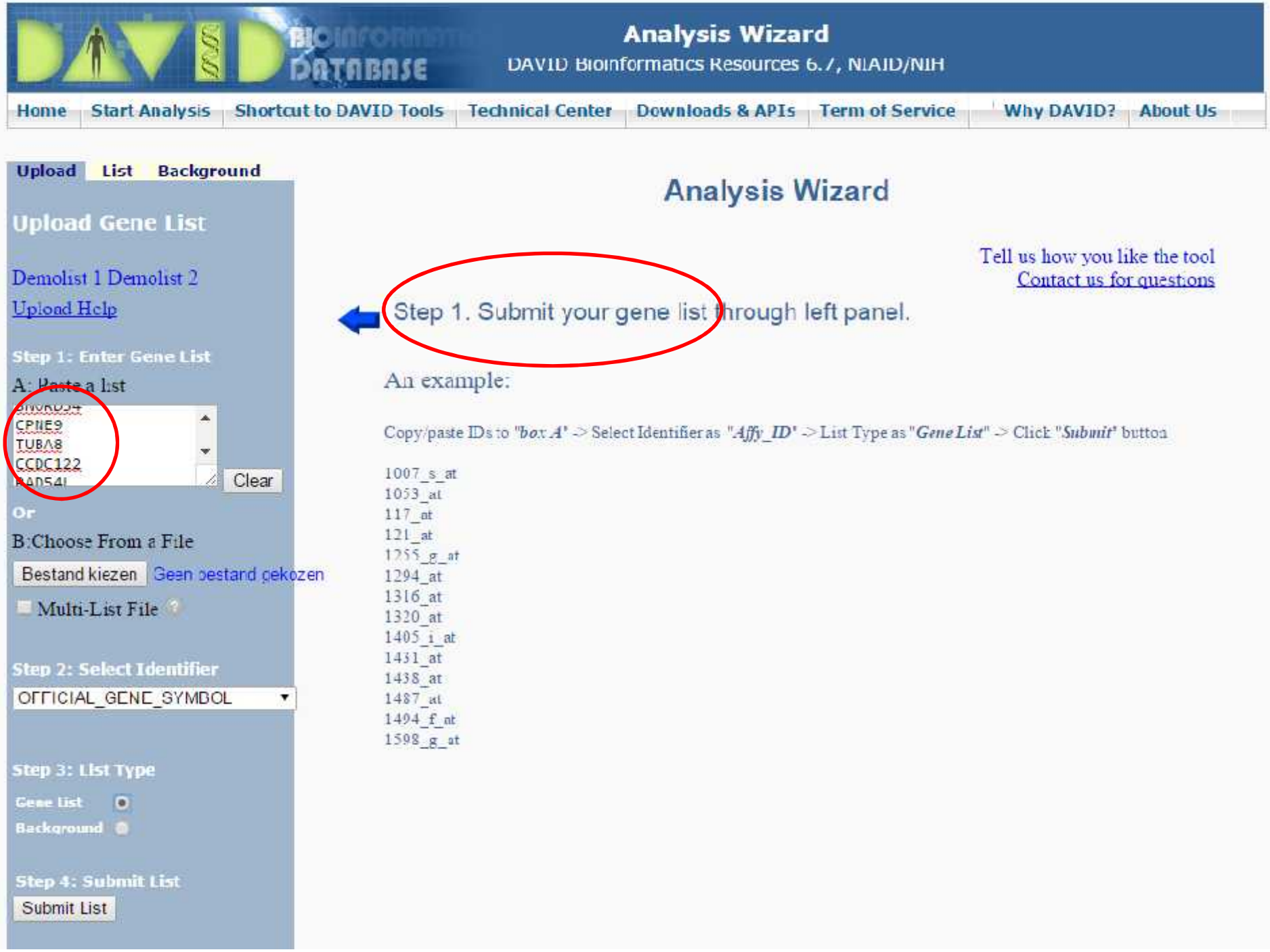
- Design new experiments based on results using additional tools and applications
- Write an article or create data sets that becomes part of the primary source of the used tools

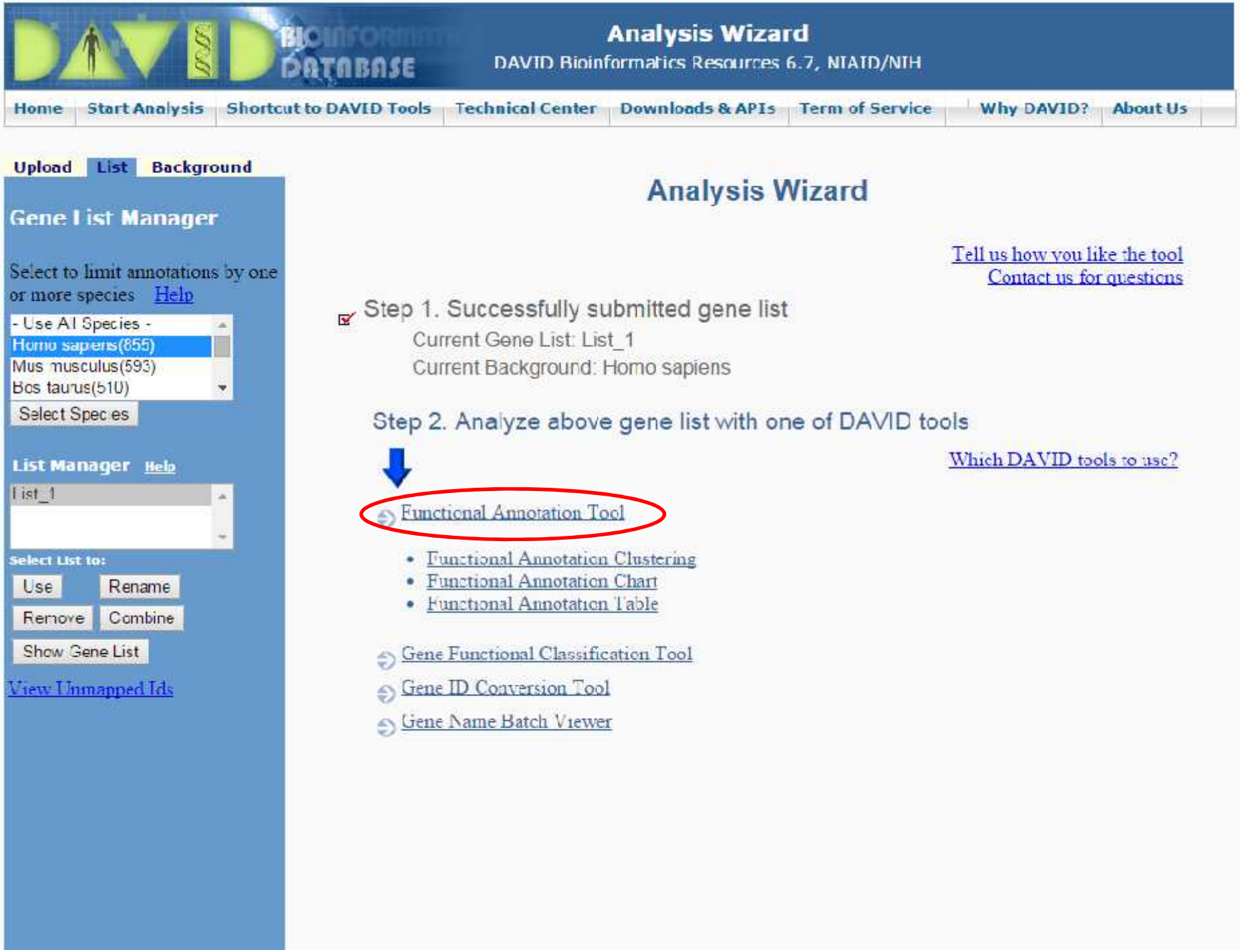


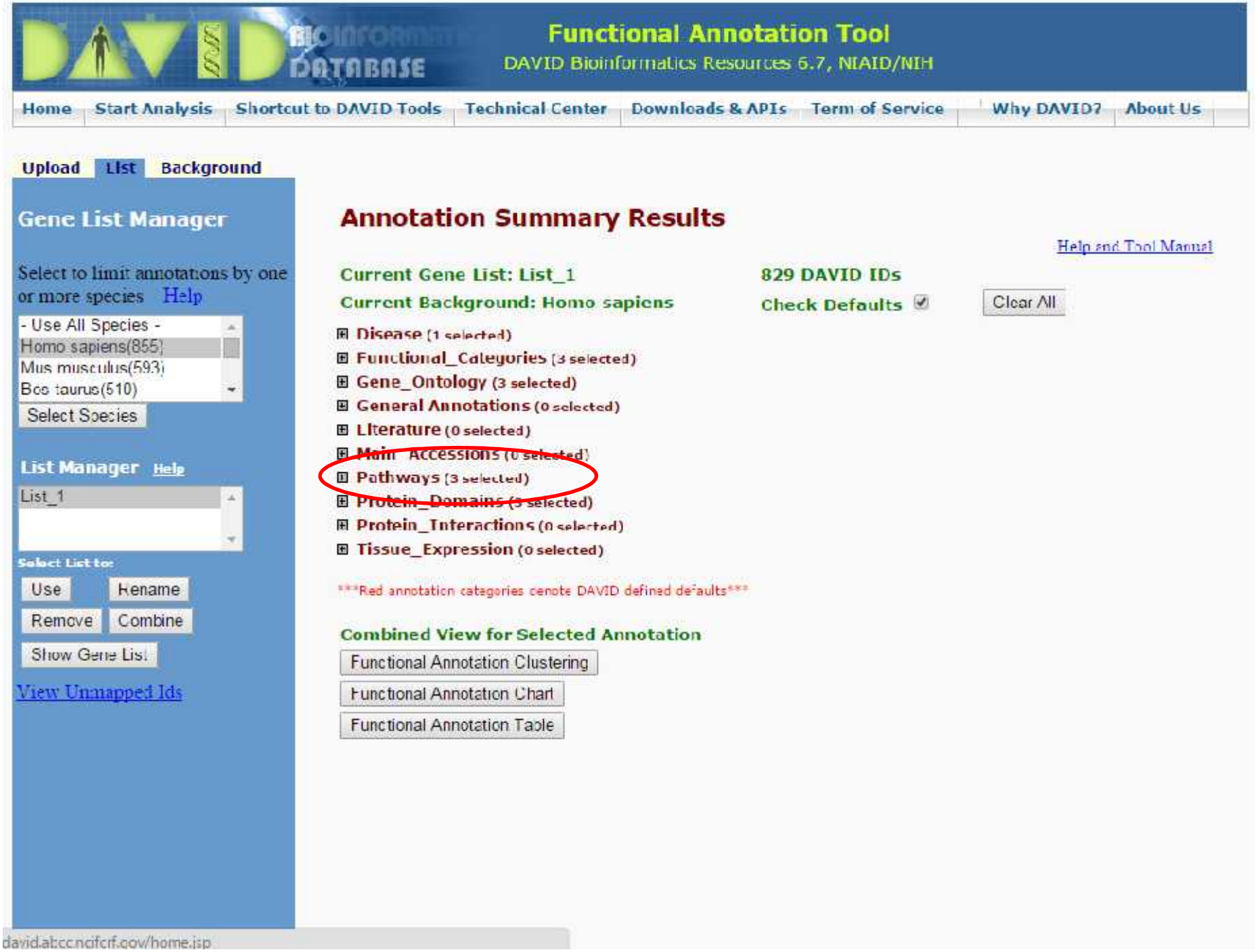
Demo of DAVID as an Open Source Genomic Tool Set

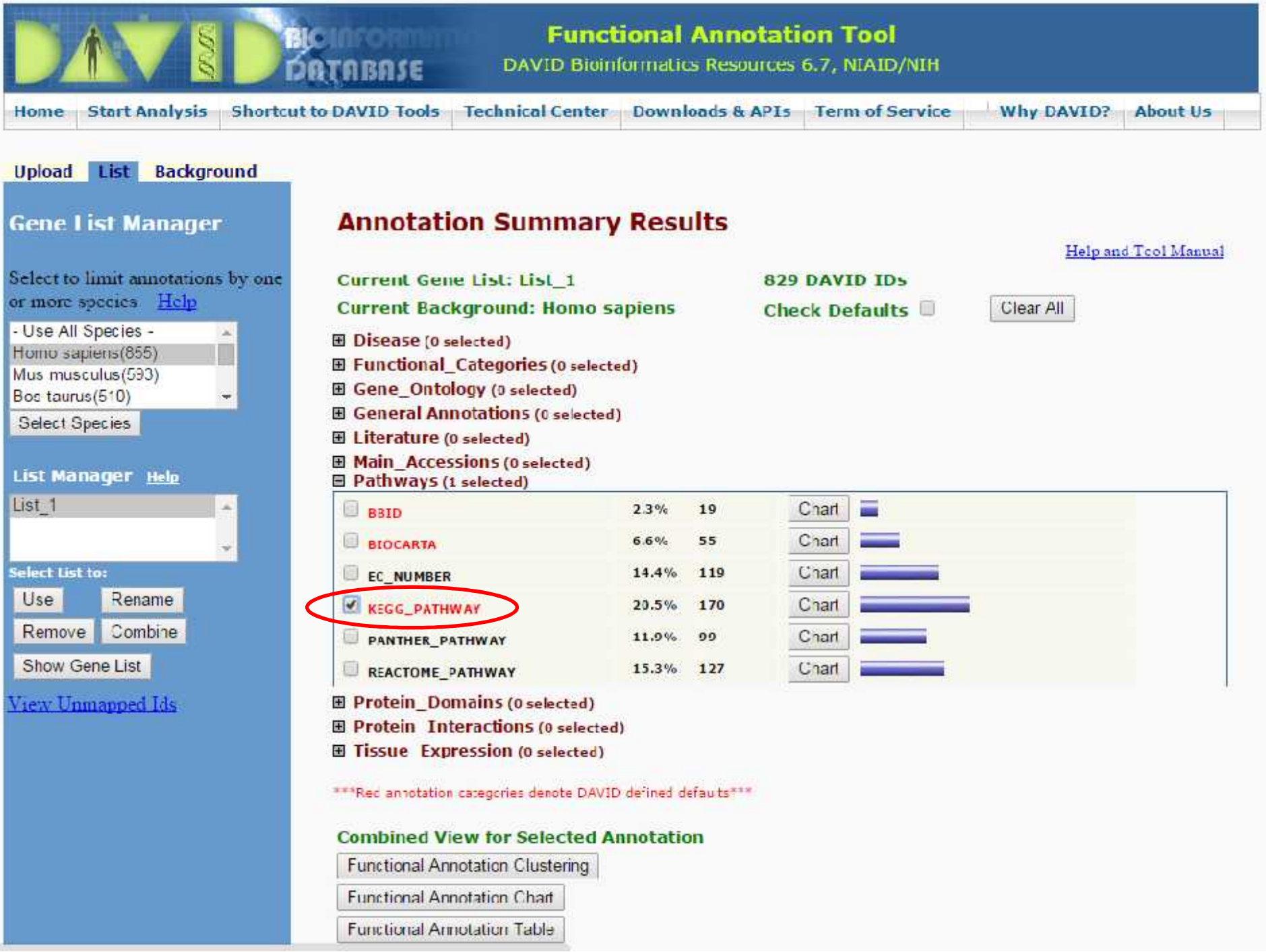
Scientist's view on operational aspects of Text and Data Mining

Dr. Albert de Roos, 26 March 2015


Analysis Wizard
DAVID Bioinformatics Resources 6.7, NIAID/NIH
Home
Start Analysis
Shortcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Why DAVID?
About Us
Upload
List
Background
Upload Gene List
Demolist 1 Demolist 2
Upload Help
Step 1: Enter Gene List
A: Paste a list
CPNE9
TUBA8
CCDC122
Clear
Or
B:Choose From a File
Bestand kiezen
Geen bestand gekozen
Multi-List File
Step 2: Select Identifier
OFFICIAL_GENE_SYMBOL
Step 3: List Type
Gene List
Background
Step 4: Submit List
Submit List
Analysis Wizard
Tell us how you like the tool
Contact us for questions
Step 1. Submit your gene list through left panel.
An example:
Copy/paste IDs to "box A" -> Select Identifier as "Affy_ID" -> List Type as "Gene List" -> Click "Submit" button
1007_s_at
1053_at
117_at
121_at
1255_g_at
1294_at
1316_at
1320_at
1405_i_at
1431_at
1438_at
1487_at
1494_f_at
1598_g_at


Analysis Wizard
DAVID Bioinformatics Resources 6.7, NIAID/NIH
Home
Start Analysis
Shortcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Why DAVID?
About Us
Upload
List
Background
Gene List Manager
Select to limit annotations by one or more species Help
- Use All Species -
Homo sapiens(855)
Mus musculus(593)
Bos taurus(510)
Select Species
List Manager Help
List_1
Select List to:
Use
Rename
Remove
Combine
Show Gene List
View Unmapped Ids
Analysis Wizard
Tell us how you like the tool
Contact us for questions
Step 1. Successfully submitted gene list
Current Gene List: List_1
Current Background: Homo sapiens
Step 2. Analyze above gene list with one of DAVID tools
Which DAVID tools to use?
Functional Annotation Tool
Functional Annotation Clustering
Functional Annotation Chart
Functional Annotation Table
Gene Functional Classification Tool
Gene ID Conversion Tool
Gene Name Batch Viewer


Functional Annotation Tool
DAVID Bioinformatics Resources 6.7, NIAID/NIH
Home
Start Analysis
Shortcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Why DAVID?
About Us
Upload
List
Background
Gene List Manager
Select to limit annotations by one or more species
Help
- Use All Species -
Homo sapiens(855)
Mus musculus(593)
Bos taurus(510)
Select Species
List Manager
Help
List_1
Select List to:
Use
Rename
Remove
Combine
Show Gene List
View Unmapped Ids
Annotation Summary Results
Help and Tool Manual
Current Gene List: List_1
Current Background: Homo sapiens
829 DAVID IDs
Check Defaults
Clear All
Disease (1 selected)
Functional_Categories (3 selected)
Gene_Ontology (3 selected)
General Annotations (0 selected)
Literature (0 selected)
Main_Accessions (0 selected)
Pathways (3 selected)
Protein_Domains (3 selected)
Protein_Interactions (0 selected)
Tissue_Expression (0 selected)
Red annotation categories denote DAVID defined defaults
Combined View for Selected Annotation
Functional Annotation Clustering
Functional Annotation Chart
Functional Annotation Table
david.abcc.ncifcrf.gov/home.jsp


Functional Annotation Tool
DAVID Bioinformatics Resources 6.7, NIAID/NIH
Home
Start Analysis
Shortcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Why DAVID?
About Us
Upload
List
Background
Gene List Manager
Select to limit annotations by one or more species
Help
- Use All Species -
Homo sapiens(855)
Mus musculus(593)
Bos taurus(510)
Select Species
List Manager
Help
List_1
Select List to:
Use
Rename
Remove
Combine
Show Gene List
View Unmapped Ids
Annotation Summary Results
Help and Tool Manual
Current Gene List: List_1
829 DAVID IDs
Current Background: Homo sapiens
Check Defaults
Clear All
Disease (0 selected)
Functional_Categories (0 selected)
Gene_Ontology (0 selected)
General Annotations (0 selected)
Literature (0 selected)
Main_Accessions (0 selected)
Pathways (1 selected)
BBID 2.3% 19 Chart
BIOCARTA 6.6% 55 Chart
EC_NUMBER 14.4% 119 Chart
KEGG_PATHWAY 20.5% 170 Chart
PANTHER_PATHWAY 11.9% 99 Chart
REACTOME_PATHWAY 15.3% 127 Chart
Protein_Domains (0 selected)
Protein Interactions (0 selected)
Tissue Expression (0 selected)
Red annotation categories denote DAVID defined defaults
Combined View for Selected Annotation
Functional Annotation Clustering
Functional Annotation Chart
Functional Annotation Table

Functional Annotation Chart

Help and Manual

Current Gene List: List_1
Current Background: Homo sapiens
829 DAVID IDs

⊞ **Options**

Rerun Using Options | Create Sublist

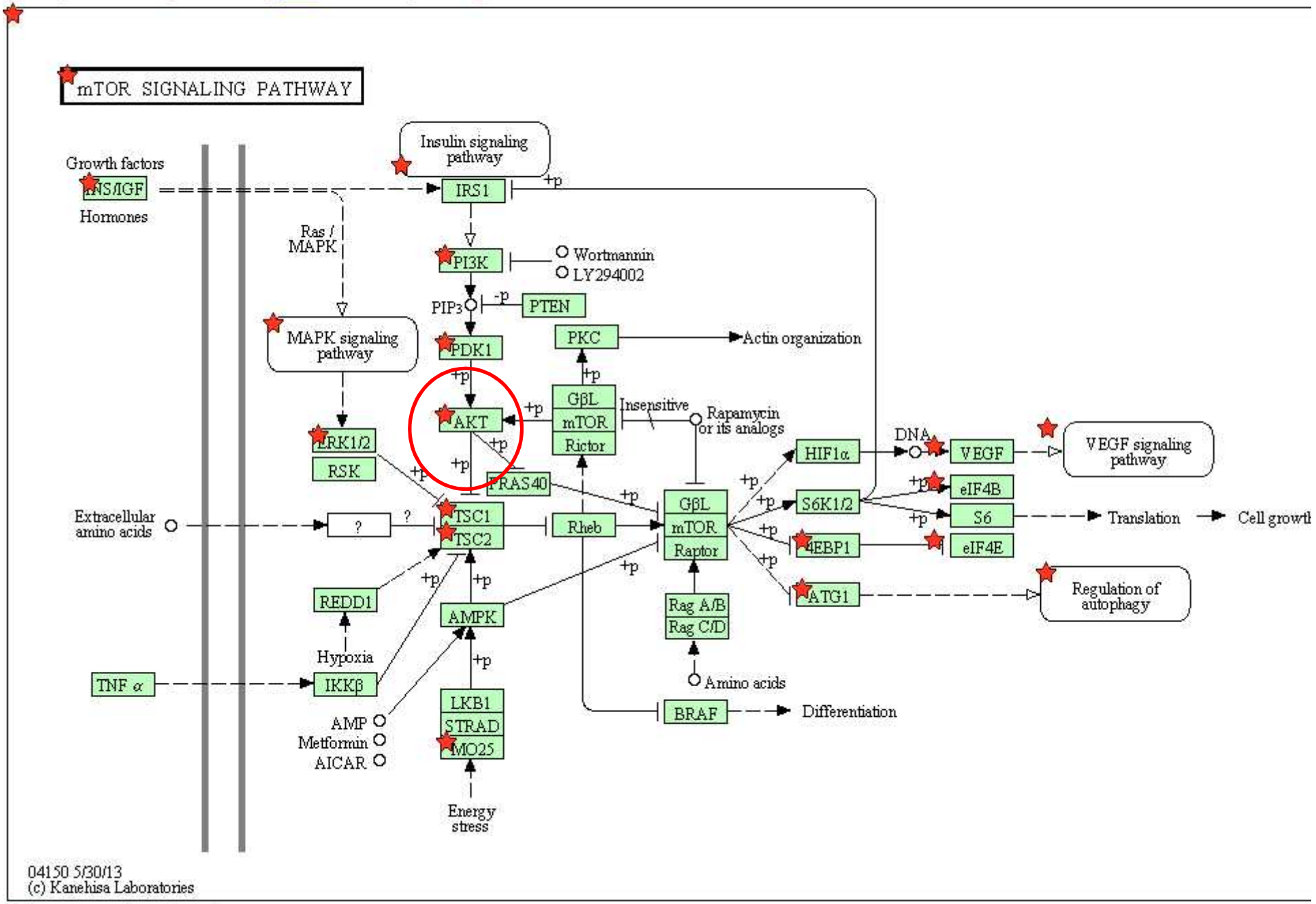
44 chart records

Download File

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
☐	KEGG_PATHWAY	mTOR signaling pathway	RT		24	2,9	3.3E-21	4,7E-19
☐	KEGG_PATHWAY	Progesterone-mediated oocyte maturation	RT		18	2,2	1.8E-9	1,3E-7
☐	KEGG_PATHWAY	Insulin signaling pathway	RT		21	2,5	1.2E-8	5,8E-7
☐	KEGG_PATHWAY	Aldosterone-regulated sodium reabsorption	RT		12	1,4	5.3E-8	1,9E-6
☐	KEGG_PATHWAY	Type II diabetes mellitus	RT		12	1,4	2.4E-7	6,9E-6
☐	KEGG_PATHWAY	Colorectal cancer	RT		15	1,8	4.9E-7	1,7E-5
☐	KEGG_PATHWAY	Renal cell carcinoma	RT		13	1,6	2.4E-6	5,0E-5
☐	KEGG_PATHWAY	Pancreatic cancer	RT		13	1,6	3.3E-6	5,9E-5
☐	KEGG_PATHWAY	Systemic lupus erythematosus	RT		15	1,8	3.8E-6	6,0E-5
☐	KEGG_PATHWAY	Glioma	RT		12	1,4	5.4E-6	7,7E-5
☐	KEGG_PATHWAY	Endometrial cancer	RT		11	1,3	6.0E-6	7,7E-5
☐	KEGG_PATHWAY	Non-small cell lung cancer	RT		11	1,3	8.5E-6	1,0E-4
☐	KEGG_PATHWAY	Melanoma	RT		12	1,4	1.8E-5	2,0E-4
☐	KEGG_PATHWAY	Prostate cancer	RT		13	1,6	3.1E-5	3,2E-4
☐	KEGG_PATHWAY	Acute myeloid leukemia	RT		10	1,2	1.0E-4	9,9E-4
☐	KEGG_PATHWAY	ErbB signaling pathway	RT		12	1,4	1.2E-4	1,1E-3

Pathway:mTOR signaling pathway

Pathway information generated by KEGG. ☑ **Stop Blinking**



List genes are shown in red

Gene Report

Help and Manua

Download File

List Id: v-akt murine thymoma viral oncogene homolog 1	v-akt murine thymoma viral oncogene homolog 1	Related Genes	Homo sapiens
CHROMOSOME	14,		
CYTOBAND	14q32.32, 14q32.32|14q32.32,		
ENSEMBL_GENE_ID	ENSG00000142208,		
ENTREZ_GENE_ID	207,		
GENERIF_SUMMARY	Binding of CTMP to PKBalpha reduces its activity by inhibiting phosphorylation on serine 473 and threonine 308., regulated in platelets by collage receptor glycoprotein VI, Absence of mutations in the pleckstrin homology (PH) domain of protein kinase B (PKB/Akt) in malignant melanoma., Immunohistochemical localization of phosphorylated AKT/PKB in multiple myeloma cells., Akt enhances Mdm2-mediated ubiquitination and degradation of p53., Identification of 14-3-3zeta as a protein kinase B/Akt substrate., Phosphorylation of HDM2 by Akt, and protein binding, IGF-I protects the cells from apoptosis by blocking the activation of caspases, which may be responsible for the loss of FAK and Akt., Akt promotes cell-cycle progression through the mechanisms of phosphorylation-dependent 14-3-3 binding to p27(Kip1) and cytoplasmic localization., AKT activation delays radiation-induced apoptosis, allowing the DNA repair mechanism more time to remove cyclobutane thymine dimers, Different cellular localization, translocation, and insulin-induced phosphorylation of PKBalpha in HepG2 cells and hepatocytes, This study shows that activation of Akt by pervanadate or serum is associated with tyrosine phosphorylation of Akt., 3' phosphoinositide lipid-dependent translocation of PKB to the plasma membrane promotes serine 473 phosphorylation, which is, in turn, necessary for PDK1-mediated phosphorylation of threonine 308 and, consequentially, full PKB activation., determination of high-resolution structure of the pleckstrin homology domain of bound to phosphatidylinositol (3,4,5)-trisphosphate, connective tissue growth factor induced fibronectin production, cell migration, and cytoskeletal rearrangement are associated with recruitment of Src and phosphorylation of p42/44 MAPK and protein kinase B, These data indicate that Akt may contribute to tumor-cell proliferation by phosphorylation and cytosolic retention of p27(kip1), thus relieving CDK2 from p27-induced inhibition., Data show that activation of protein kinase B (PKB)/Akt, contributes to resistance to antiproliferative signal and breast cancer progression in part by impairing the nuclear import and action of p27., Data show that cytoplasmic relocalization of p27(kip1), secondary to Akt-mediated phosphorylation, inactivates the growth inhibitory properties of p27(kip1) and sustains the proliferation of breast cancer cells., PECAM-1 involvement through Akt/PKB activation in starvation-induced transendothelial migration of CD34+CD14+ circulating precursors, data demonstrate that Rho/ROCK pathway negatively regulates eNOS phosphorylation through inhibition of protein kinase B (PKB), whereas it downregulates eNOS expression independent of PKB, chemotherapeutic drugs exhibited their cytotoxic effects in part by down-regulating Akt signaling following TRADD expression, We conclude that normal HERG function in HEK293 cells requires basal activity of PKB. Our data represent the first evidence that PKB phosphorylation regulates K(+) channels., Decreased phosphorylation of protein kinase B and erk1/erk2 in neutrophils from patients with myelodysplastic syndrome, This protein protects HL60 leukemia cells from TRAIL-induced apoptosis through a mechanism involving NF-kappaB activation and cFLIP(L) up-regulation., our data suggest that HRG-beta1, bound to the ErbB2 ErbB3 heterodimer, in the presence of membrane ER-alpha, interacts with and activates PI 3-K/Akt., Increased phosphorylation of this protein was observed in A431 clonal variants., The protein kinase Akt induces epithelial mesenchymal transition and promotes enhanced motility and invasiveness of squamous cell carcinoma lines., activation of Notch1 signaling mediates p53 function in HPV16 E6 and E7 cell transformation via phosphatidylinositol(PI3K)-PKB/Akt pathway, Akt i activated by adrenomedullin, AKT1 is regulated by JIP1, results suggest that TRB3 promotes glucose output from liver under fasting conditions by binding to and interfering with Akt phosphorylation in response to residual insulin signaling, Akt regulates basic helix-loop helix transcription factor-coactivator complex formation and activity during neuronal differentiation, phosphorylation by Akt regulates the		

J Biol Chem. 2002 Oct 11;277(41):38021-8. Epub 2002 Jul 30.

Direct identification of tyrosine 474 as a regulatory phosphorylation site for the Akt protein kinase.

Conus NM[1], Hannan KM, Cristiano BE, Hemmings BA, Pearson RB.

Author information

Abstract

Understanding the regulation of Akt has been of major interest for elucidating the control of normal cellular physiology as well as malignant transformation. The paradigm for activation of Akt involves phosphatidylinositol 3-kinase-dependent membrane localization followed by activating phosphorylation of Thr-308 and Ser-473. Many of the activating signals for Akt involve the stimulation of receptor and non-receptor tyrosine kinases, and the most potent activator known is the tyrosine phosphatase inhibitor pervanadate, highlighting a possible role for tyrosine phosphorylation in the regulation of the enzyme. In this study we show that activation of Akt by pervanadate or serum is associated with tyrosine phosphorylation of Akt. In addition, in SKOV3 ovarian carcinoma cells that exhibit high basal levels of Akt activity, Akt was tyrosine-phosphorylated in the basal state, and this phosphorylation was further enhanced by both pervanadate and insulin-like growth factor-1. We have used NH(2)-terminal sequencing and phosphate release analysis to directly identify Tyr-474 as the site of tyrosine phosphorylation. Substitution of Tyr-474 with phenylalanine abolished tyrosine phosphorylation of Akt and resulted in up to 55% inhibition of Akt activation, indicating phosphorylation at Tyr-474 is required for full activation of the kinase. Our data identifies a novel regulatory mechanism for this pleiotropic enzyme that may be applicable to the AGC family of protein kinases given the conserved nature of the COOH-terminal hydrophobic motif containing Tyr-474.

PMID: 12149249 [PubMed - indexed for MEDLINE] Free full text

jbc THE JOURNAL OF BIOLOGICAL CHEMISTRY

Direct Identification of Tyrosine 474 as a Regulatory Phosphorylation Site for the Akt Protein Kinase*,[210]

Nelly Marmy Conus, Katherine M. Hannan, Briony E. Cristiano, Brian A. Hemmings[‡] and Richard B. Pearson[§]

+ Author Affiliations

Abstract

Understanding the regulation of Akt has been of major interest for elucidating the control of normal cellular physiology as well as malignant transformation. The paradigm for activation of Akt involves phosphatidylinositol 3-kinase-dependent membrane localization followed by activating phosphorylation of Thr-308 and Ser-473. Many of the activating signals for Akt involve the stimulation of receptor and non-receptor tyrosine kinases, and the most potent activator known is the tyrosine phosphatase inhibitor pervanadate, highlighting a possible role for tyrosine phosphorylation in the regulation of the enzyme. In this study we show that activation of Akt by pervanadate or serum is associated with tyrosine phosphorylation of Akt. In addition, in SKOV3 ovarian carcinoma cells that exhibit high basal levels of Akt activity, Akt was tyrosine-phosphorylated in the basal state, and this phosphorylation was further enhanced by both pervanadate and insulin-like growth factor-1. We have used NH_2-terminal sequencing and

« Previous | Next Article »
Table of Contents

This Article

First Published on July 30, 2002, doi: 10.1074/jbc.M203387200

October 11, 2002 The Journal of Biological Chemistry, 277, **38021-38028.**

Abstract *Free*
» **Full Text** *Free*
Full Text (PDF) *Free*
PDF including Supp Data
Supplemental Data

All Versions of this Article:
M203387200v1

End of Presentation

- Operational Text and Data Mining-