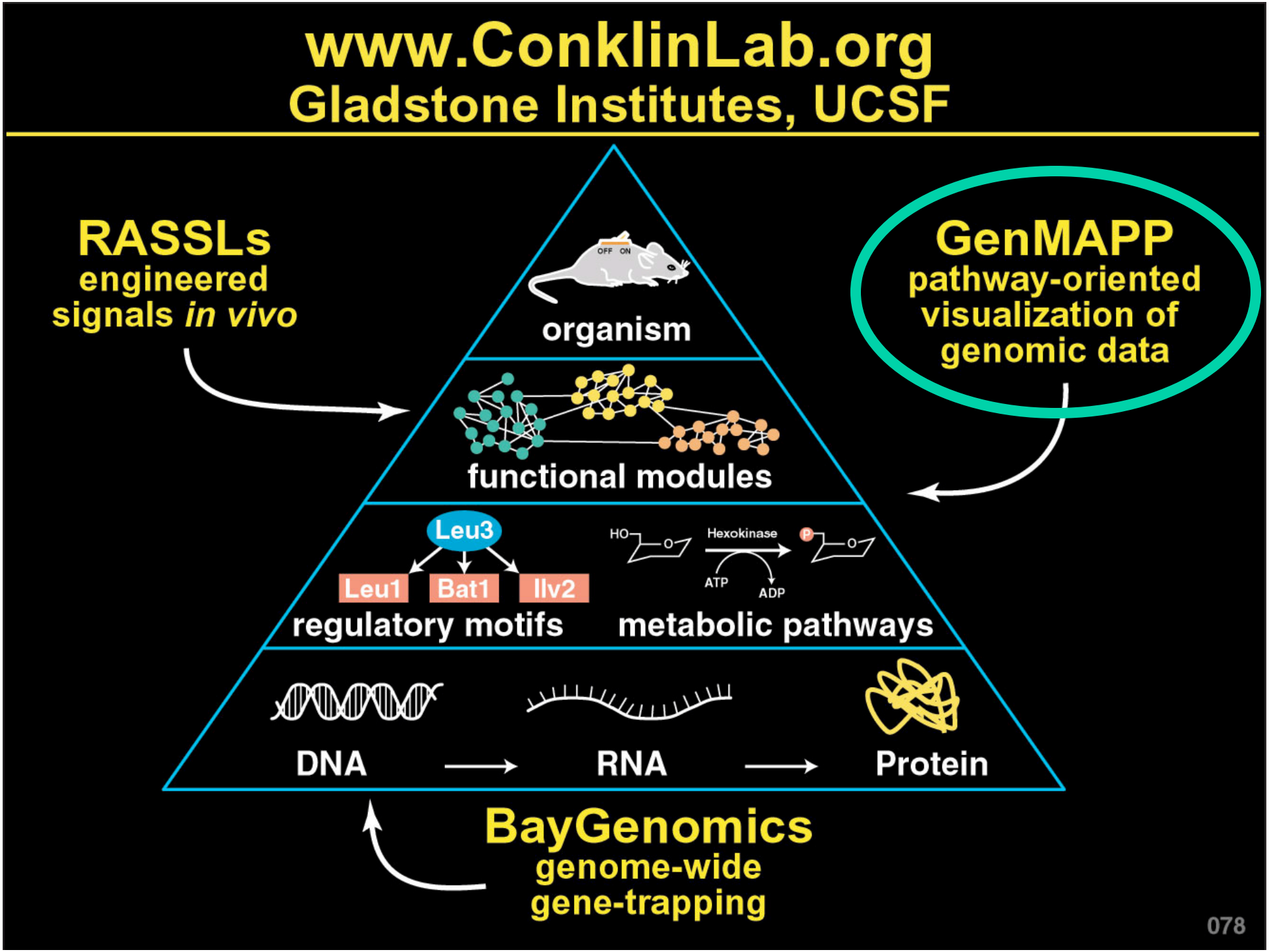
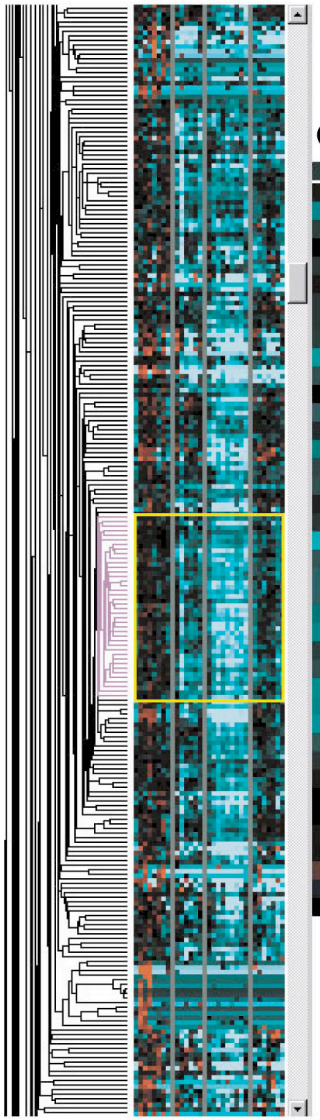
www.ConklinLab.org
Gladstone Institutes, UCSF

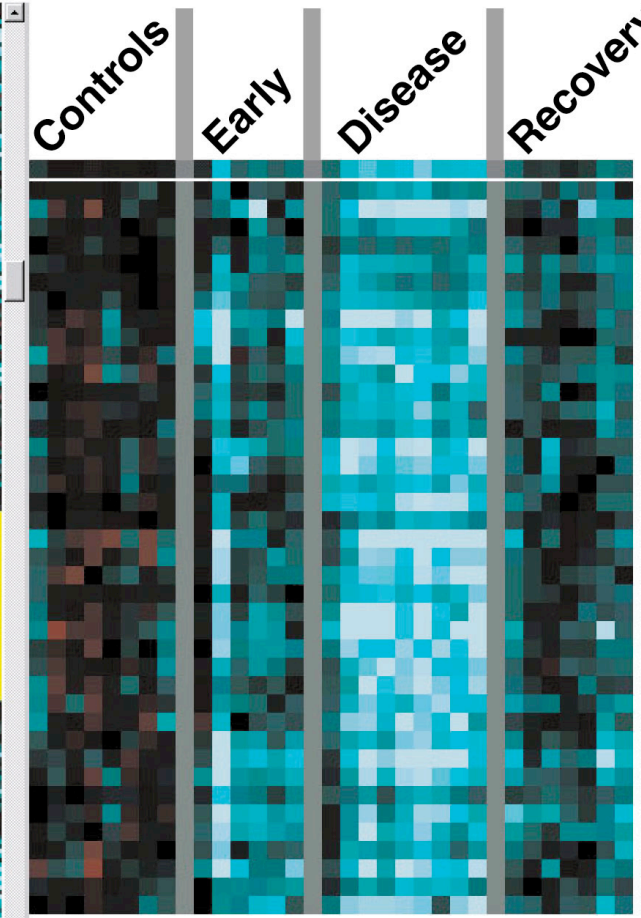


Challenges in Microarray Data Analysis

How can we combine microarray data with biological knowledge?

Signal Dependent Genes - Downregulated

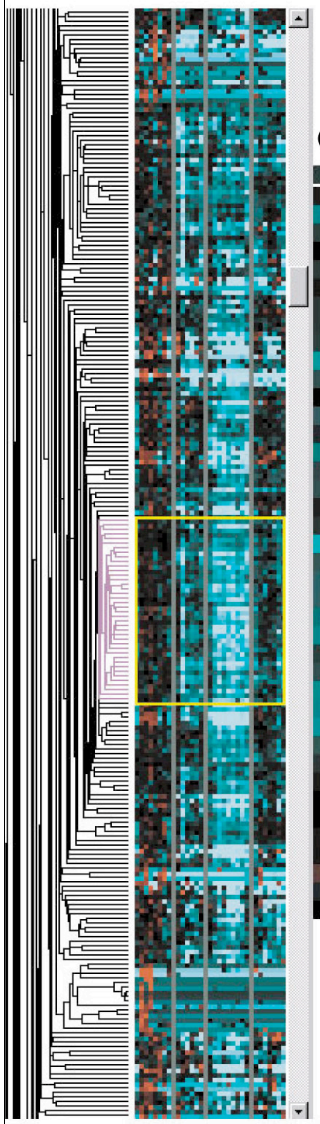


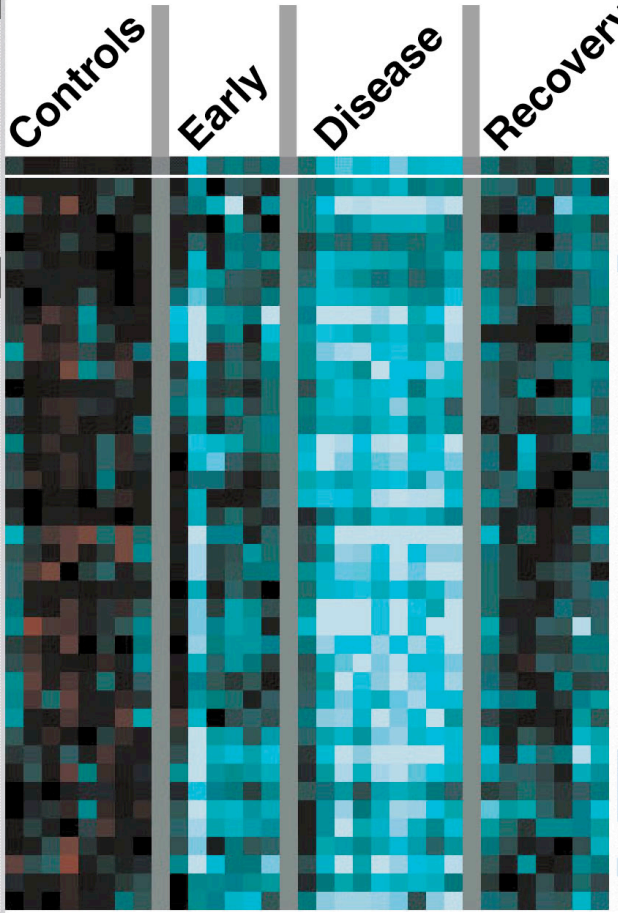


BLAST Definitions

"Homologous to sp O07021: PRE-MRNA SPLICING FACTOR SF2, P32 SUBUNIT PRECURSOR (GCIQ-R PROTEIN)
"Mus musculus proteasome activator PA28 alpha subunit mRNA, complete cds"
"Mus musculus cdc37 homolog mRNA, complete cds"
"Mus musculus ornithine decarboxylase antizyme gene, complete cds"
M.musculus mRNA for carnitine acetyltransferase
"Homologous to sp O00779: CALCIUM-TRANSPORTING ATPASE SARCOPLASMIC RETICULUM TYPE (EC 3.6.1.3
"Homologous to sp P11507: CALCIUM-TRANSPORTING ATPASE ENDOPLASMIC RETICULUM TYPE (EC 3.6.1.38)
M.musculus ENO3 mRNA for enolase beta subunit
"Homologous to sp P47858: 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (PHOSPHOFRUCTOKINAS
Homologous to sp P23327: SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BINDING PROTEIN PRECURS
"Mouse AE3 mRNA, complete cds"
"M.musculus glucose transporter 2 mRNA, complete cds"
Mus musculus aspartate aminotransferase gene 5'-flank and exon 1
"Mus musculus thioredoxin-dependent peroxide reductase (tpx) mRNA, complete cds"
"Homologous to sp P47858: 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (PHOSPHOFRUCTOKINAS
"Homologous to sp P11508: CALCIUM-TRANSPORTING ATPASE SARCOPLASMIC RETICULUM TYPE (EC 3.6.1.3
"Homologous to sp P35434: ATP SYNTHASE DELTA CHAIN, MITOCHONDRIAL PRECURSOR (EC 3.6.1.34)."
"Mus musculus F1F0ATP synthase complex E subunit (Atp5k) gene, complete cds"
"Mus musculus NAD(H)-specific isocitrate dehydrogenase gamma subunit precursor, mRNA, complete
"M.musculus gene for dodecenoyl-CoA delta-isomerase, exons 1 and 2"
"Mus musculus cytochrome c oxidase subunit VIII-H precursor (COX8H) mRNA, complete cds"
"Homologous to sp P35745: ACYLPHOSPHATASE, MUSCLE TYPE ISOZYME (EC 3.6.1.7) (ACYLPHOSPHATE PH
"Mus musculus CD-1 cardiac troponin I mRNA, complete cds"
"Homologous to sp P00566: CREATINE KINASE, M CHAIN (EC 2.7.3.2) (MU-2 PROTEIN)."
Mouse mRNA for protein with homology to transition protein 2 (TP2)
Mus musculus Selenium-binding liver protein mRNA
"Mus musculus (clone MAR1) aldose reductase mRNA, complete cds"
"Mus musculus vascular endothelial growth factor B 186 (VEGF-B) precursor, mRNA, complete cds"
M.musculus mRNA for NADP transhydrogenase
"Mus musculus aldehyde dehydrogenase (ALDH2) mRNA, nuclear gene encoding mitochondrial protein
"Mouse cytosolic epoxide hydrolase mRNA, complete cds"
"Mus musculus 129SV carnitine palmitoyltransferase II mRNA, complete cds"
"Mus musculus medium-chain acyl-CoA dehydrogenase mRNA, complete cds"
"Mus musculus long-chain acyl-CoA dehydrogenase mRNA, complete cds"
"Mus musculus very-long chain acyl-CoA dehydrogenase mRNA, partial cds"
Mouse muscle creatine kinase mRNA (EC 2.7.3.2)
"Mus musculus isocitrate dehydrogenase mRNA, complete cds"
"Mus musculus long chain fatty acyl CoA synthetase mRNA, complete cds"
"Mus musculus sterol carrier protein-2 (SCP-2) gene, complete cds"
"Mouse alpha-tubulin isotype M-alpha-4 mRNA, complete cds"

Signal Dependent Genes - Downregulated



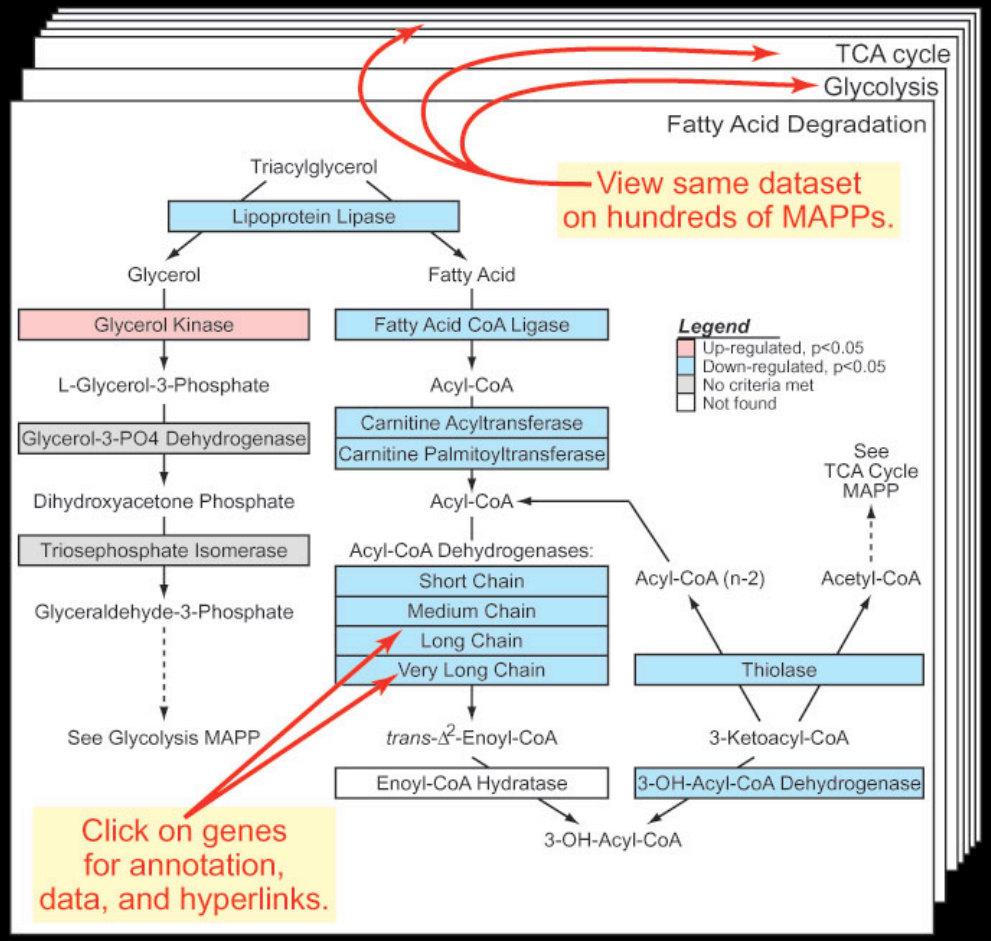


BLAST Definitions

"Homologous to sp O07021: PRE-MRNA SPLICING FACTOR SF2, P32 SUBUNIT PRECURSOR (GCIQ-R PROTEIN)
"Mus musculus proteasome activator PA28 alpha subunit mRNA, complete cds"
"Mus musculus cdc37 homolog mRNA, complete cds"
"Mus musculus ornithine decarboxylase antizyme gene, complete cds"
M.musculus mRNA for carnitine acetyltransferase
"Homologous to sp O00779: CALCIUM-TRANSPORTING ATPASE SARCOPLASMIC RETICULUM TYPE (EC 3.6.1.3
"Homologous to sp P11507: CALCIUM-TRANSPORTING ATPASE ENDOPLASMIC RETICULUM TYPE (EC 3.6.1.38)
M.musculus ENO3 mRNA for enolase beta subunit
"Homologous to sp P47858: 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (PHOSPHOFRUCTOKINAS
Homologous to sp P23327: SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BINDING PROTEIN PRECURS
"Mouse AE3 mRNA, complete cds"
"M.musculus glucose transporter 2 mRNA, complete cds"
Mus musculus aspartate aminotransferase gene 5'-flank and exon 1
"Mus musculus thioredoxin-dependent peroxide reductase (tpx) mRNA, complete cds"
"Homologous to sp P47858: 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (PHOSPHOFRUCTOKINAS
"Homologous to sp P11508: CALCIUM-TRANSPORTING ATPASE SARCOPLASMIC RETICULUM TYPE (EC 3.6.1.3
"Homologous to sp P35434: ATP SYNTHASE DELTA CHAIN, MITOCHONDRIAL PRECURSOR (EC 3.6.1.34)."
"Mus musculus F1F0ATP synthase complex E subunit (Atp5k) gene, complete cds"
"Mus musculus NAD(H)-specific isocitrate dehydrogenase gamma subunit precursor, mRNA, complete
"M.musculus gene for dodecenoyl-CoA delta-isomerase, exons 1 and 2"
"Mus musculus cytochrome c oxidase subunit VIII-H precursor (COX8H) mRNA, complete cds"
"Homologous to sp P35745: ACYLPHOSPHATASE, MUSCLE TYPE ISOZYME (EC 3.6.1.7) (ACYLPHOSPHATE PH
"Mus musculus CD-1 cardiac troponin I mRNA, complete cds"
"Homologous to sp P00566: CREATINE KINASE, M CHAIN (EC 2.7.3.2) (MU-2 PROTEIN)."
Mouse mRNA for protein with homology to transition protein 2 (TP2)
Mus musculus Selenium-binding liver protein mRNA
"Mus musculus (clone MAR1) aldose reductase mRNA, complete cds"
"Mus musculus vascular endothelial growth factor B 186 (VEGF-B) precursor, mRNA, complete cds"
M.musculus mRNA for NADP transhydrogenase
"Mus musculus aldehyde dehydrogenase (ALDH2) mRNA, nuclear gene encoding mitochondrial protein
"Mouse cytosolic epoxide hydrolase mRNA, complete cds"
"Mus musculus 129SV carnitine palmitoyltransferase II mRNA, complete cds"
"Mus musculus medium-chain acyl-CoA dehydrogenase mRNA, complete cds"
"Mus musculus long-chain acyl-CoA dehydrogenase mRNA, complete cds"
"Mus musculus very-long chain acyl-CoA dehydrogenase, partial cds"
Mouse muscle creatine kinase mRNA (EC 2.7.3.2)
"Mus musculus isocitrate dehydrogenase mRNA, complete cds"
"Mus musculus long chain fatty acyl CoA synthetase mRNA, complete cds"
"Mus musculus sterol carrier protein-2 (SCP-2) gene, complete cds"
"Mouse alpha-tubulin isotype M-alpha-4 mRNA, complete cds"

Components of Fatty Acid Degradation

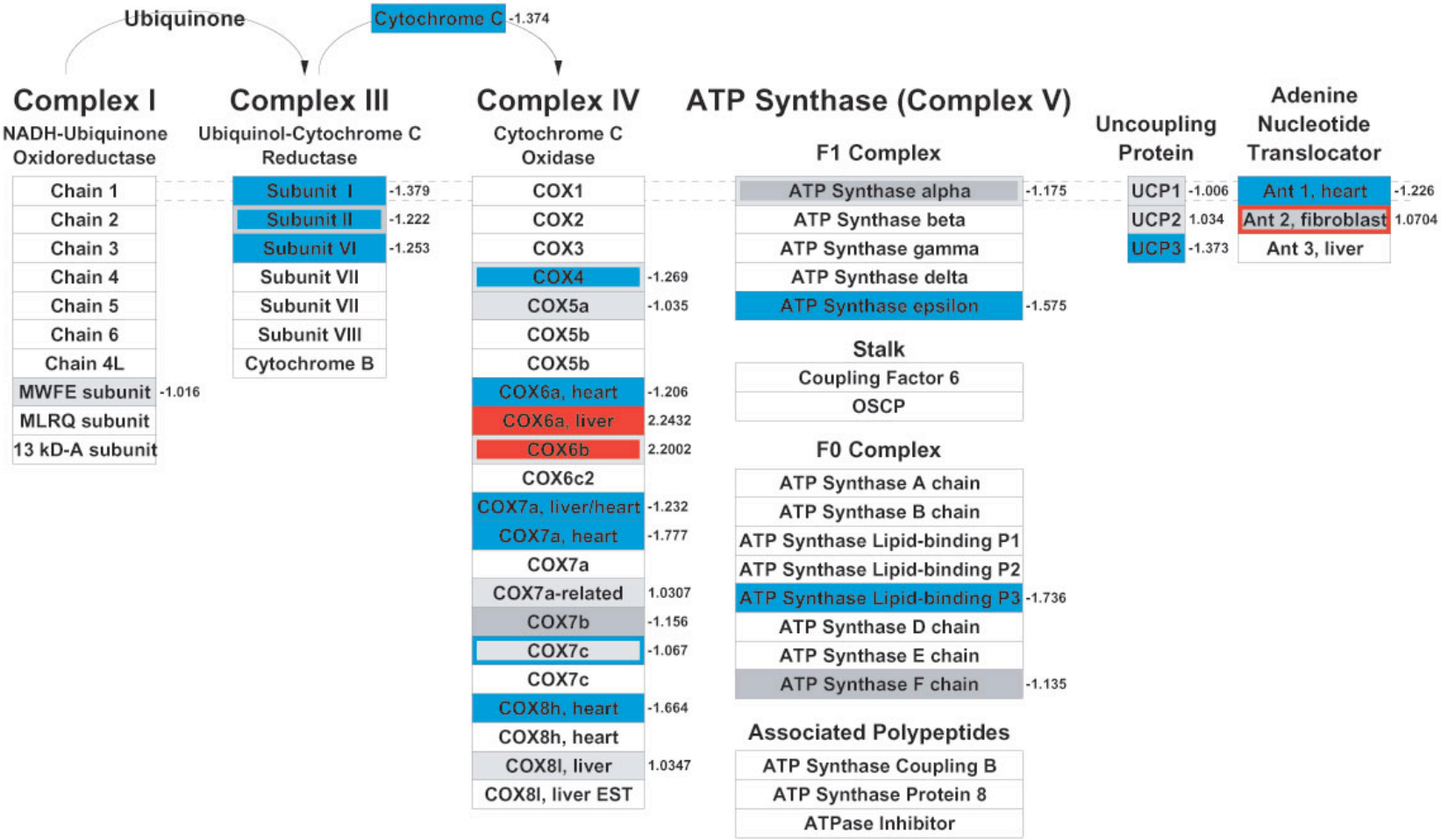
GenMAPP
Gene MicroArray Pathway Profiler
Dahlquist et al., Nature Genetics, May 2002



- Graphics tools to draw MAPPs
- Underlying gene database
- Import microarray data and choose coloring criteria
- MAPPs shared with anyone, program is free

068B

Electron Transport Chain

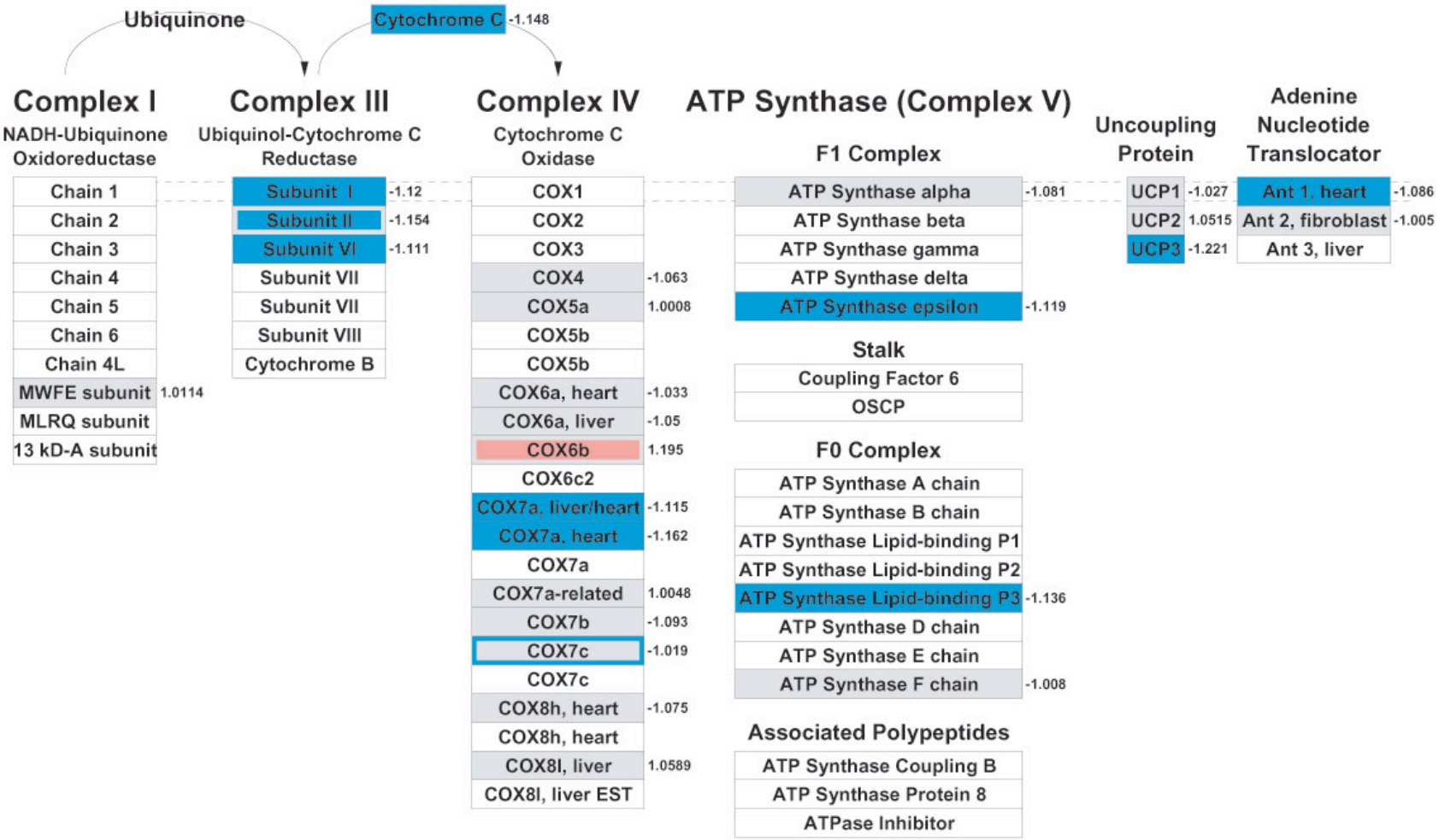


Red=Up
Blue=Down

Cardiomyopathy
Gi-coupled GPCR

076A

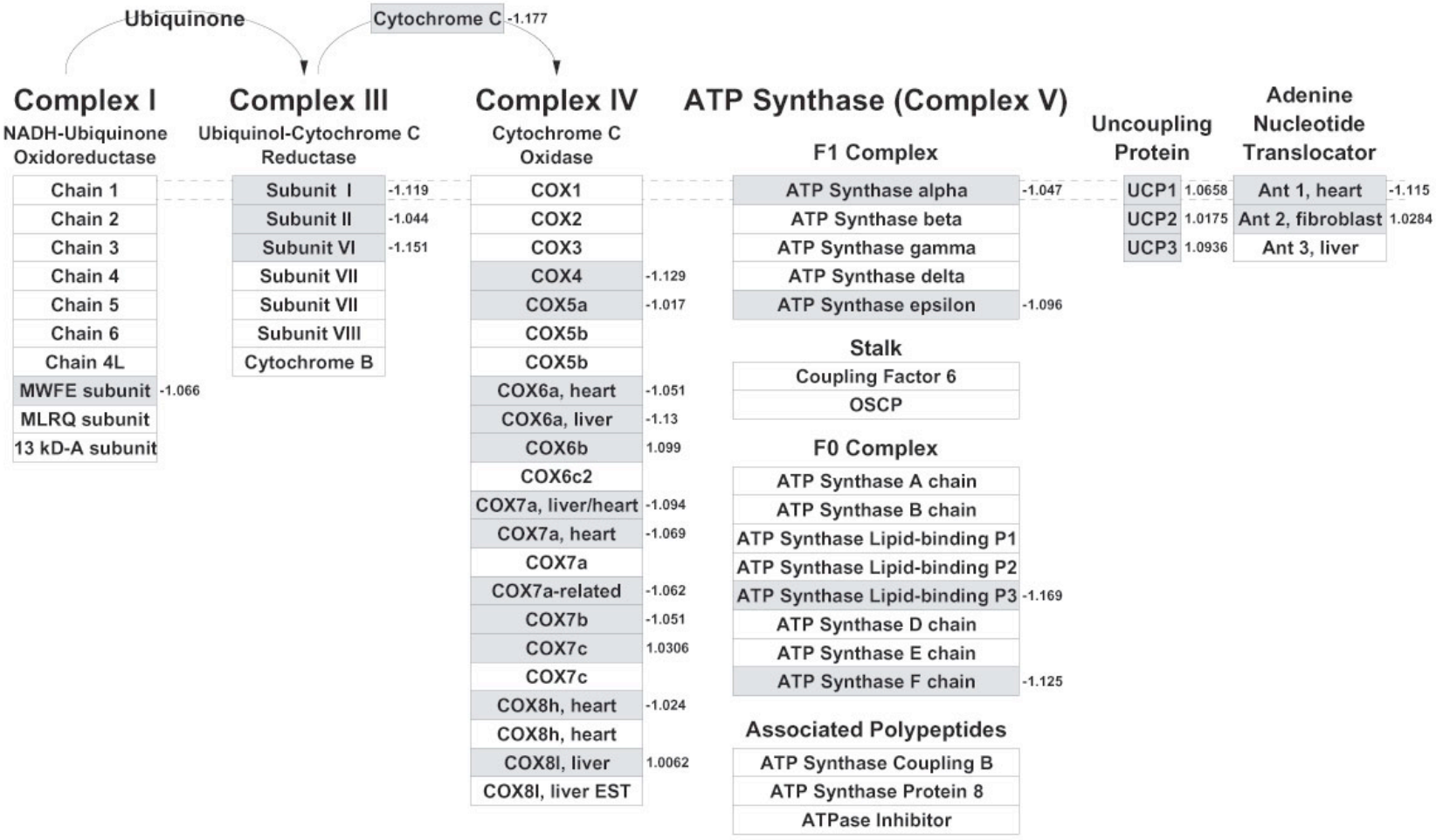
Electron Transport Chain



Red=Up
Blue=Down

Cardiomyopathy Activated Gq

Electron Transport Chain



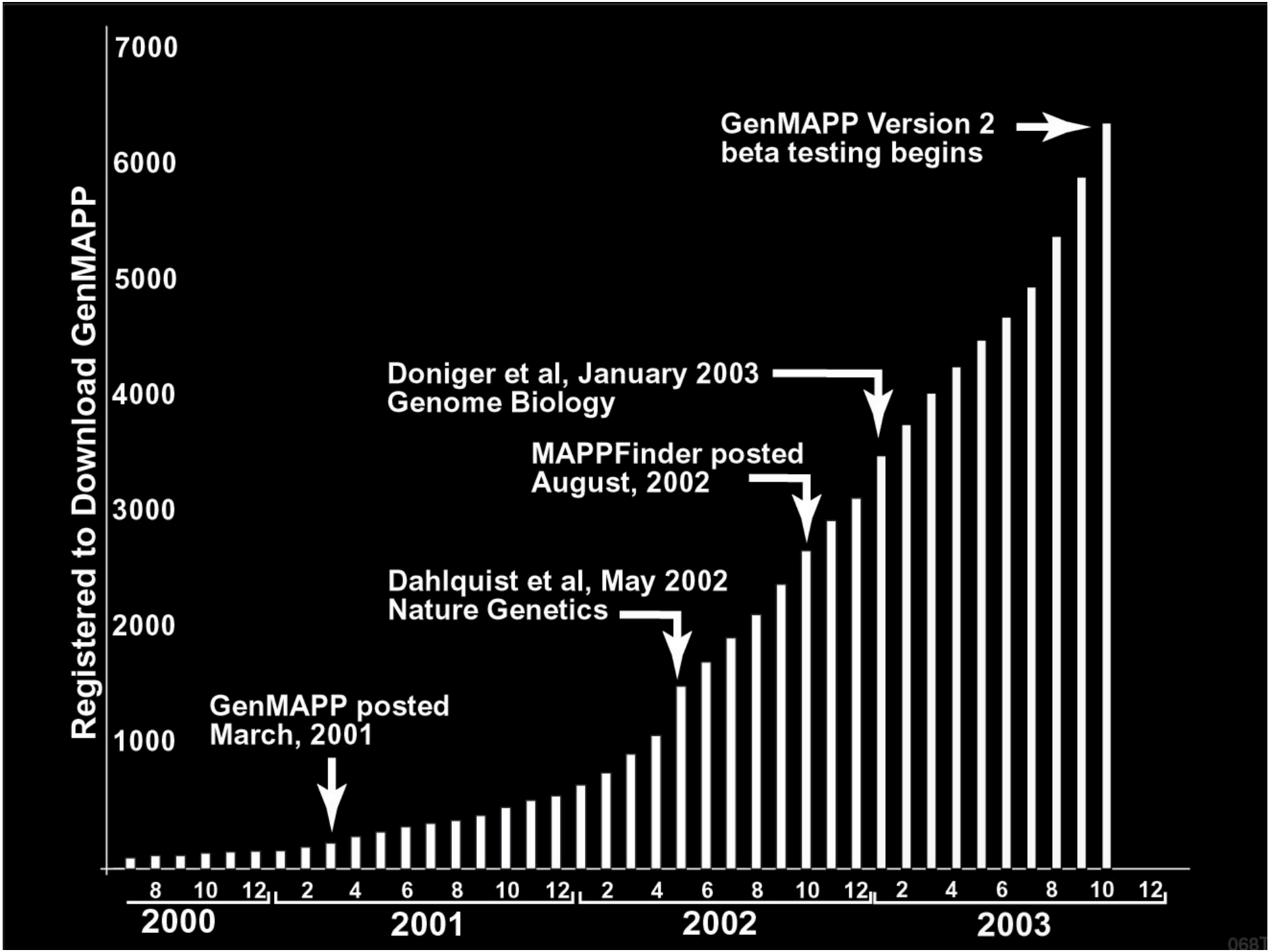
Red=Up
Blue=Down

Cardiomyopathy Recovery

www.GenMAPP.org

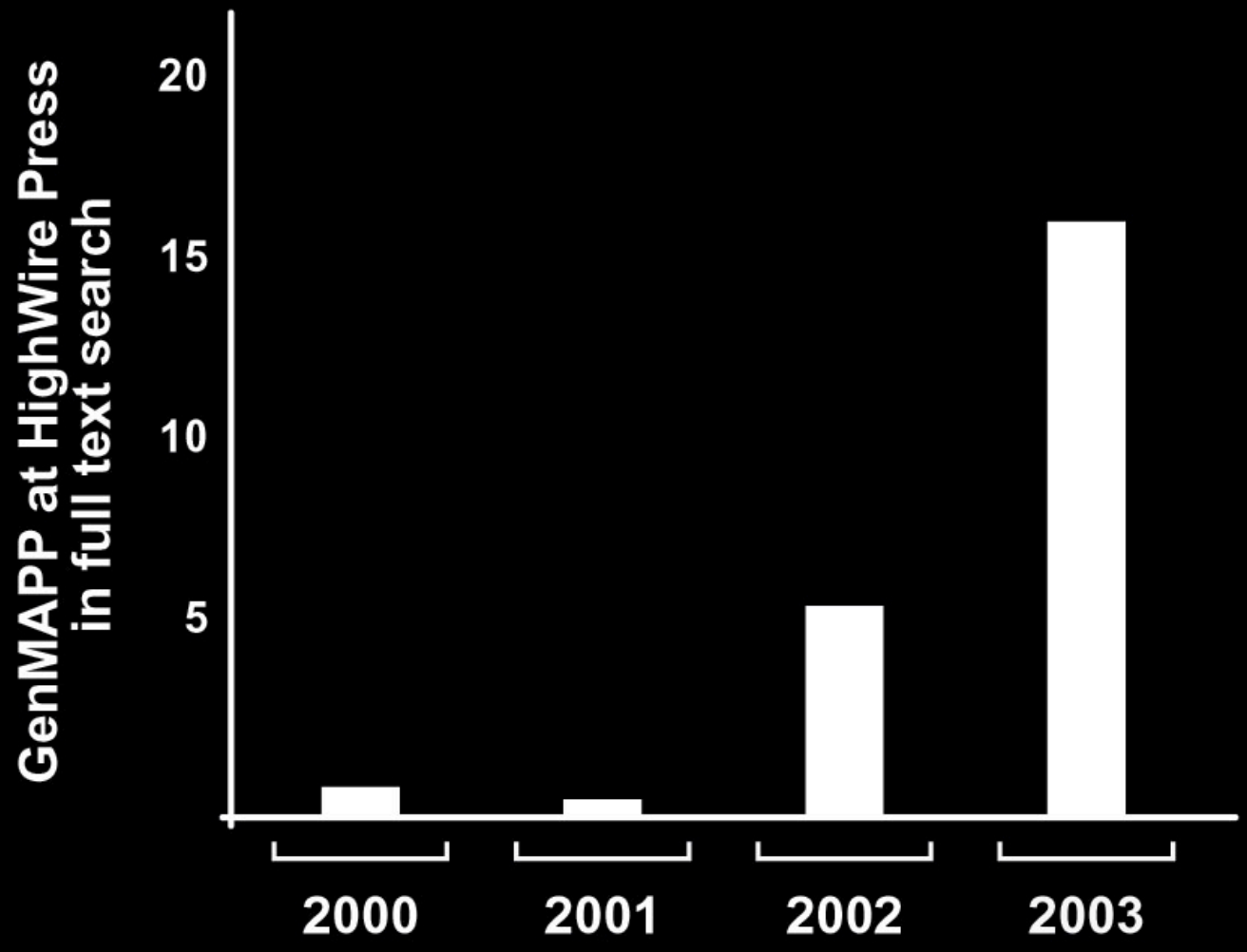
- **GenMAPP is freely distributed to all researchers.**

  released April 2001, major update April 2002
  see Dahlquist et al, Nature Genetics, May 2002
  see Doniger et al, Genome Biology, Janurary, 2003


Registered to Download GenMAPP
7000
6000
5000
4000
3000
2000
1000
GenMAPP Version 2
beta testing begins
Doniger et al, January 2003
Genome Biology
MAPPFinder posted
August, 2002
Dahlquist et al, May 2002
Nature Genetics
GenMAPP posted
March, 2001
8 10 12 2 4 6 8 10 12 2 4 6 8 10 12 2 4 6 8 10 12
2000
2001
2002
2003

GenMAPP in Published Papers

source: HighWire Press



www.GenMAPP.org

- **GenMAPP is freely distributed to all researchers.**
  - *released April 2001, major update April 2002*
  - *see Dahlquist et al, Nature Genetics, May 2002*
  - *see Doniger et al, Genome Biology, Janurary, 2003*

- **MAPPs representing pathways will be written and reviewed by experts in each field.**
  - *~50 pathway MAPPs created by the Conklin Lab*
  - *public domain datasets in GenMAPP form*

www.GenMAPP.org

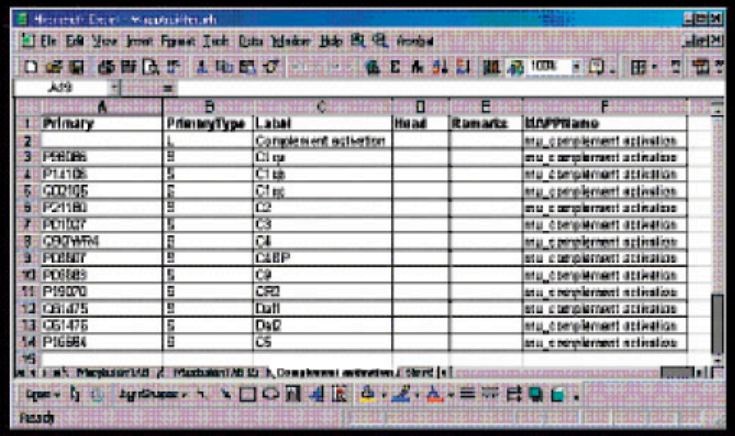
- **GenMAPP is freely distributed to all researchers.**
  - *released April 2001, major update April 2002*
  - *see Dahlquist et al, Nature Genetics, May 2002*
  - *see Doniger et al, Genome Biology, Janurary, 2003*

- **MAPPs representing pathways will be written and reviewed by experts in each field.**
  - *~50 pathway MAPPs created by the Conklin Lab*
  - *public domain datasets in GenMAPP form*

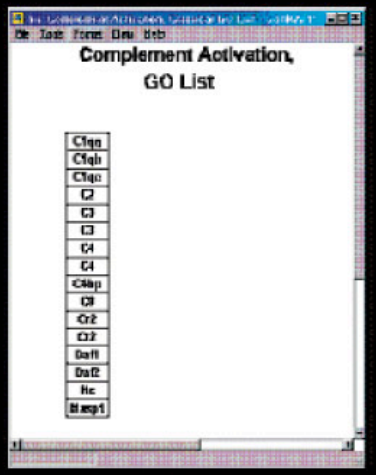
- **Gene lists from the Gene Ontology Project are converted to MAPPs automatically.**
  - *1594 mouse GO MAPPs*
  - *2446 human GO MAPPs (over 10,000 genes)*

GO Provides Content for GenMAPP

Lists of genes annotated by GO can be automatically transformed into MAPPs with MAPPBuilder

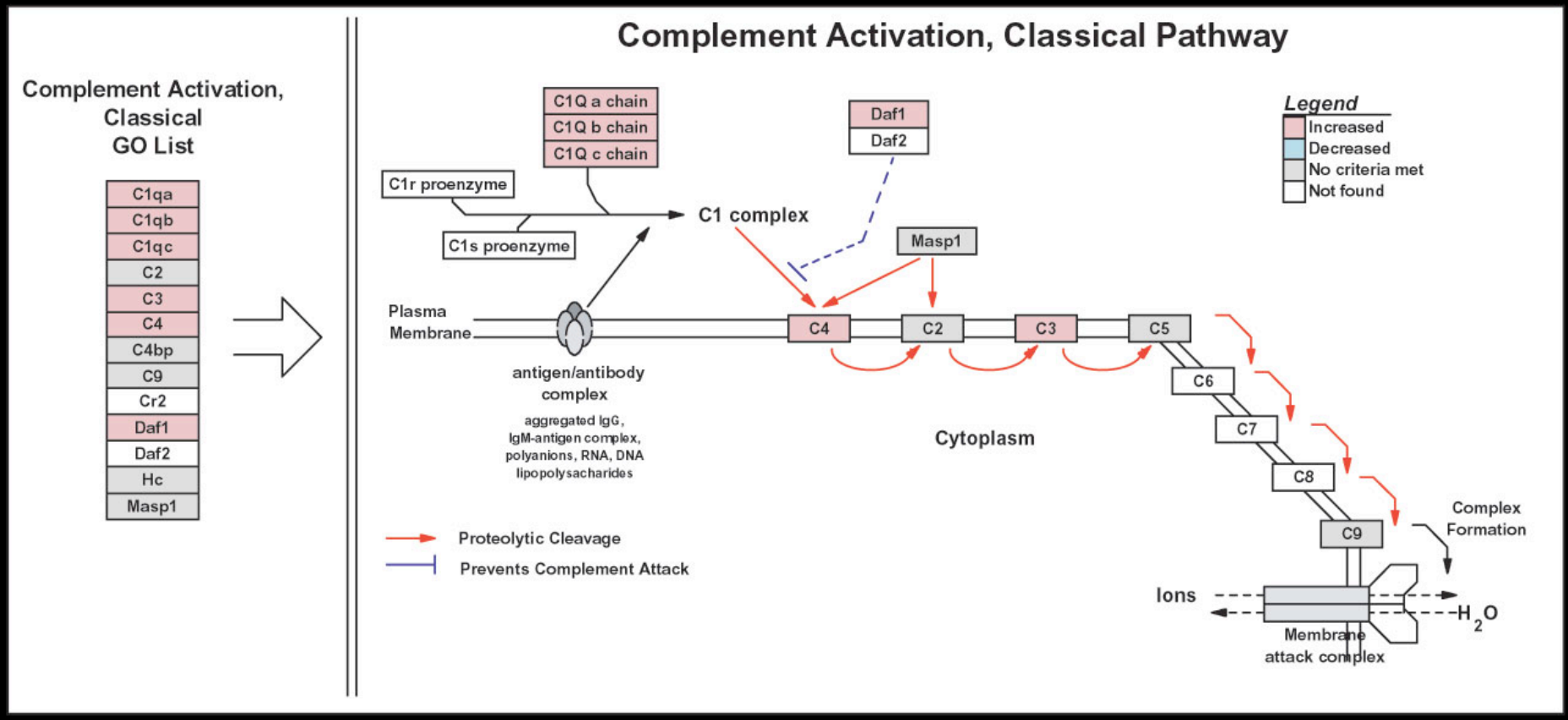


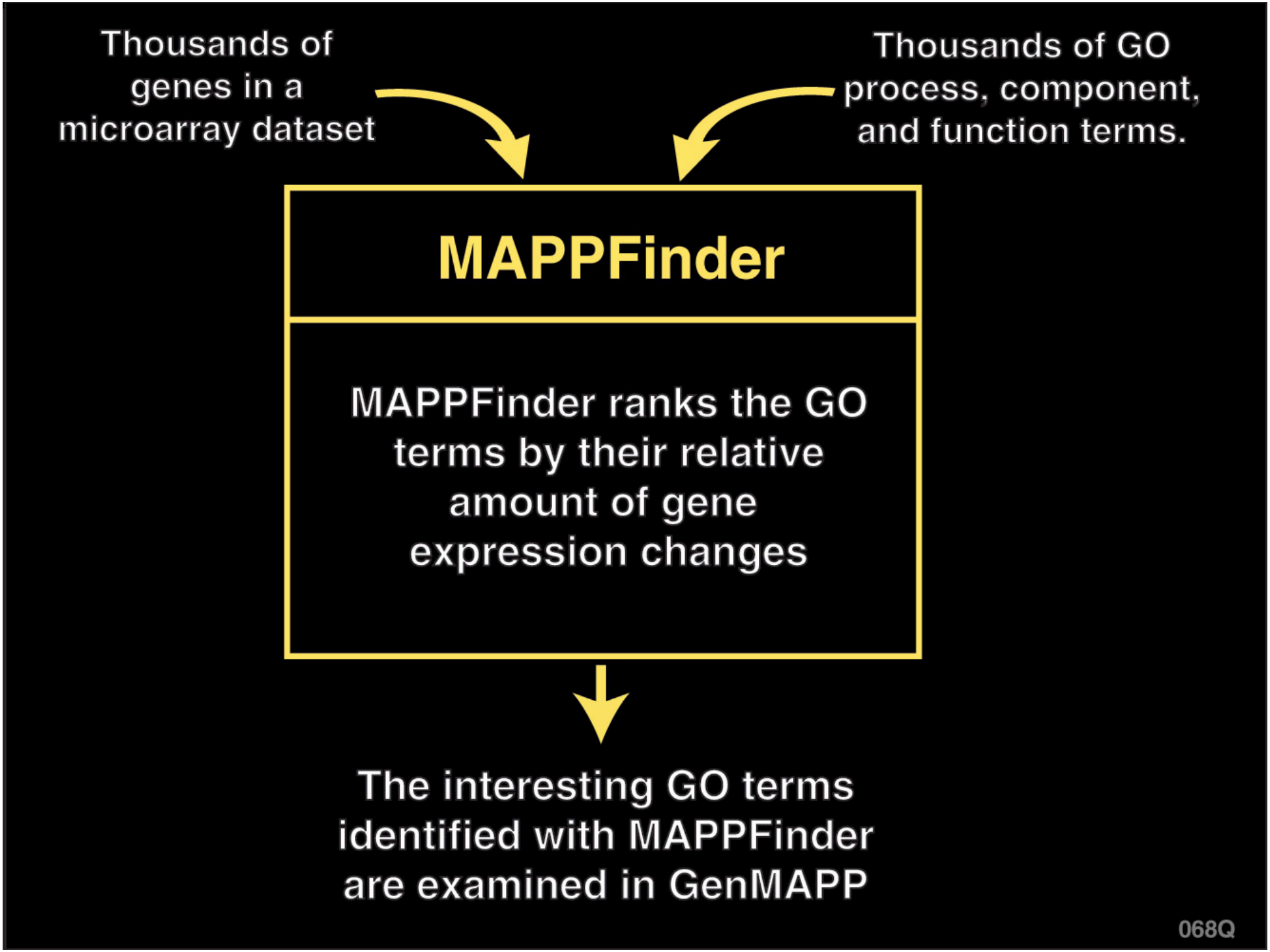
MAPPBuilder



The GO classifications are a great starting point for building detailed pathway MAPPs

Pathways Can be Drawn in GenMAPP From GO Lists




Thousands of genes in a microarray dataset
Thousands of GO process, component, and function terms.
MAPPFinder
MAPPFinder ranks the GO terms by their relative amount of gene expression changes
The interesting GO terms identified with MAPPFinder are examined in GenMAPP
068Q

MAPPFinder and GO Reveal Biological Groups

Prostate Cancer

chromosome organization
nuclear organization
regulation of chromatin
transmembrane Rec Tyr kinase
amine biosynthesis
TGF beta signaling
transcription co-activator
cell growth and maintenance
heterotrimeric G proteins
glutathione conjugation

Cardiomyopathy

fatty acid metabolism
mitochondria
cell cycle arrest
ribosome structure
primary active transporter
muscle development
respiratory chain complex
protein biosynthesis
GTP binding protein
ion channels

GenMAPP: Future Development

Version 2

- **Expansion and integration of gene database**

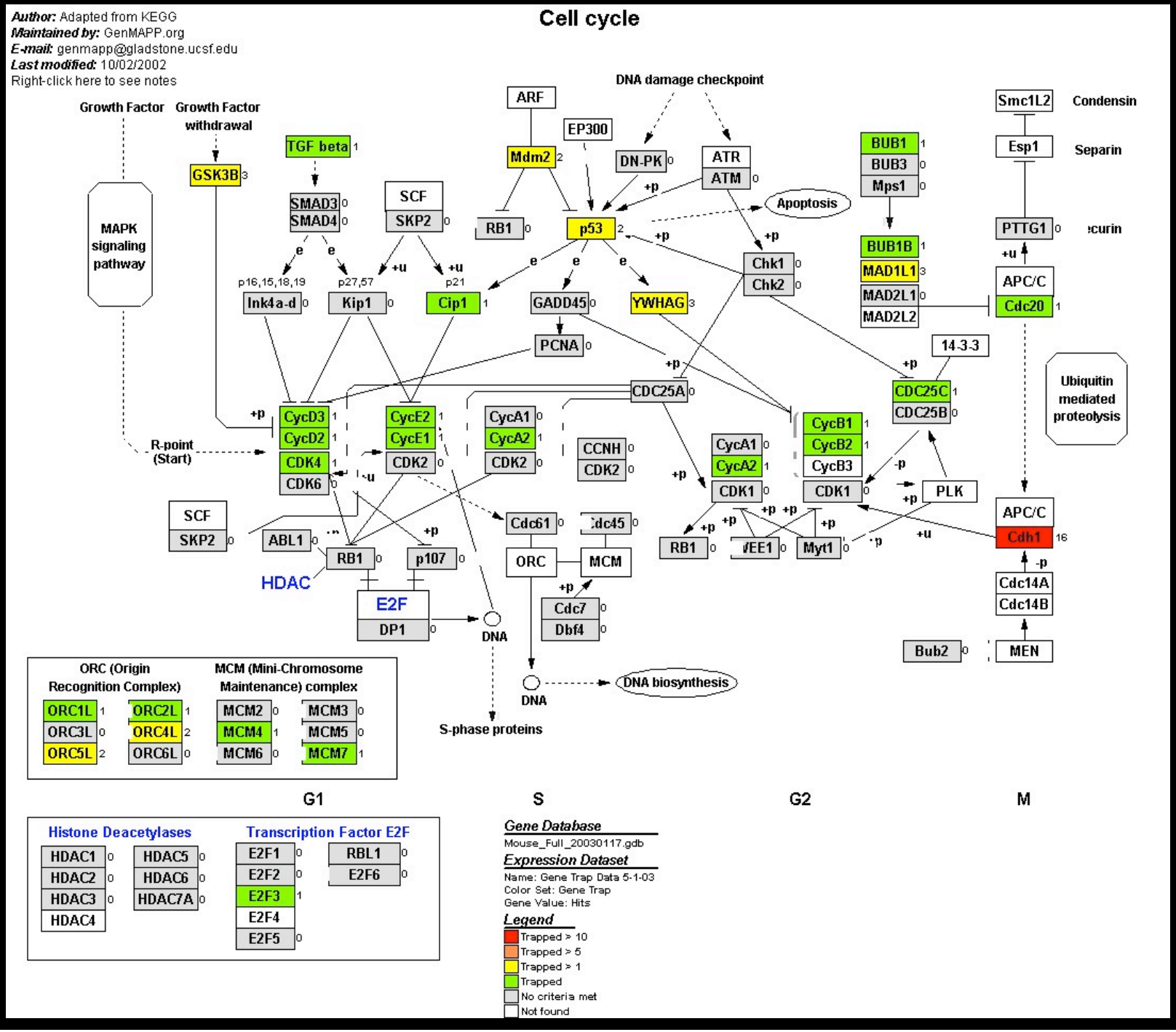
SWISS-PROT	*H. sapiens*
GenBank	*M. musculus*
LocusLink	*R. norvegicus*
RefSeq	*S. cerevisiae*
Unigene	*C. elegans*
InterPro	*D. melanogaster*
Gene Ontology	*D. rerio*
Genome-specific databases	*A. thaliana*
	E. coli

- **Users choose which catalogs and species to download**
- **Add custom gene catalogs and relationships**

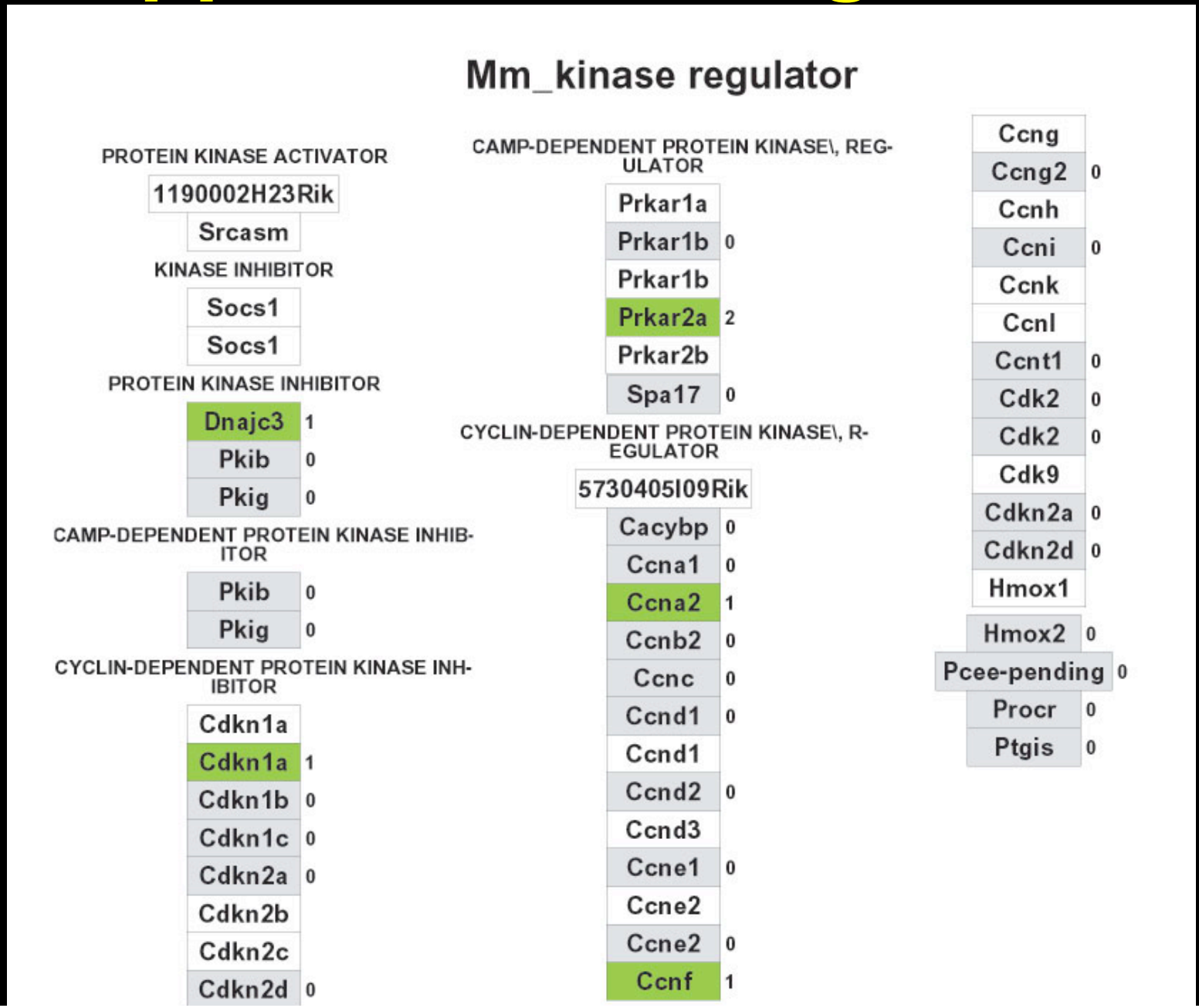
Using the Genome Knowledgebase with GenMAPP

GenMAPP for displaying any type of genomic data
RNAi, plasmids, antibodies…

Example: Gene trapping at BayGenomics


Author: Adapted from KEGG
Maintained by: GenMAPP.org
E-mail: genmapp@gladstone.ucsf.edu
Last modified: 10/02/2002
Right-click here to see notes
Cell cycle
DNA damage checkpoint
Growth Factor
Growth Factor withdrawal
MAPK signaling pathway
GSK3B
TGF beta
SMAD3
SMAD4
SCF
SKP2
ARF
Mdm2
EP300
DN-PK
ATR
ATM
Apoptosis
RB1
p53
Chk1
Chk2
BUB1
BUB3
Mps1
BUB1B
MAD1L1
MAD2L1
MAD2L2
Smc1L2
Condensin
Esp1
Separin
PTTG1
securin
APC/C
Cdc20
Ubiquitin mediated proteolysis
p16,15,18,19
Ink4a-d
p27,57
Kip1
p21
Cip1
GADD45
YWHAG
PCNA
14-3-3
CDC25A
CDC25C
CDC25B
R-point (Start)
CycD3
CycD2
CDK4
CDK6
CycE2
CycE1
CDK2
CycA1
CycA2
CDK2
CCNH
CDK2
CycA1
CycA2
CDK1
CycB1
CycB2
CycB3
CDK1
PLK
APC/C
Cdh1
Cdc14A
Cdc14B
MEN
Bub2
SCF
SKP2
ABL1
RB1
p107
HDAC
E2F
DP1
Cdc61
Cdc45
ORC
MCM
Cdc7
Dbf4
DNA
DNA biosynthesis
S-phase proteins
ORC (Origin Recognition Complex)
ORC1L
ORC2L
ORC3L
ORC4L
ORC5L
ORC6L
MCM (Mini-Chromosome Maintenance) complex
MCM2
MCM3
MCM4
MCM5
MCM6
MCM7
Histone Deacetylases
HDAC1
HDAC2
HDAC3
HDAC4
HDAC5
HDAC6
HDAC7A
Transcription Factor E2F
E2F1
E2F2
E2F3
E2F4
E2F5
RBL1
E2F6
Gene Database
Mouse_Full_20030117.gdb
Expression Dataset
Name: Gene Trap Data 5-1-03
Color Set: Gene Trap
Gene Value: Hits
Legend
Trapped > 10
Trapped > 5
Trapped > 1
Trapped
No criteria met
Not found
G1
S
G2
M

Trapped Kinase Regulators



GenMAPP Future Directions Map

DNA-microarray data

↓

Run MAPPFinder
on >1,000 GO groups
and local MAPPs

↓

View MAPPs,
Add MAPP annotation,
pathway organization

↓

Insights

↓

Publish or exchange
PDF, HTML, XML

GenMAPP Future Directions Map

DNA-microarray data

↓

Run MAPPFinder
on >1,000 GO groups
and local MAPPs

↓

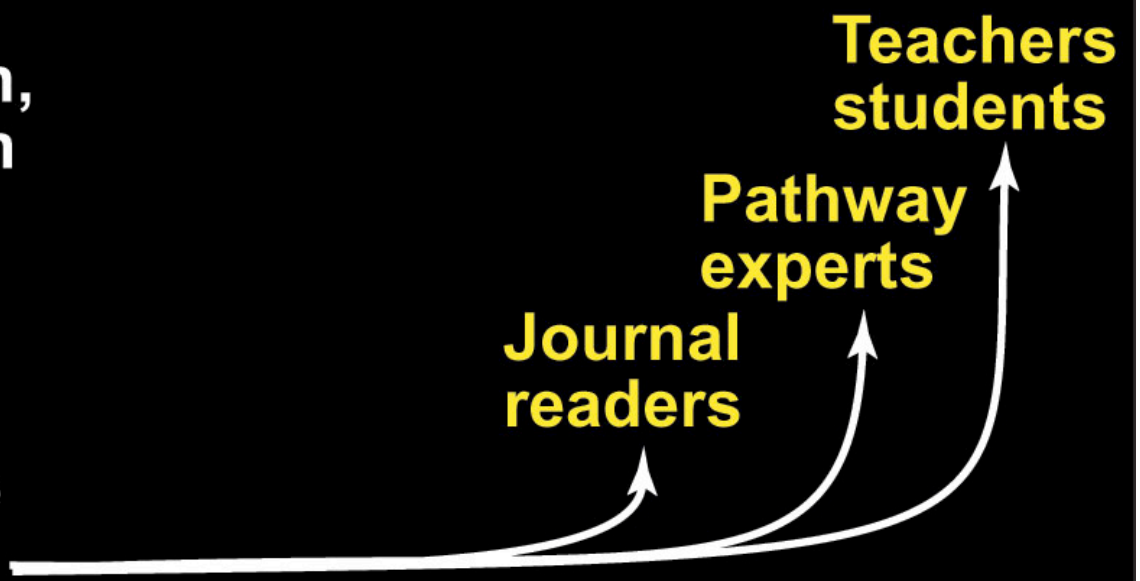
View MAPPs,
Add MAPP annotation,
pathway organization

↓

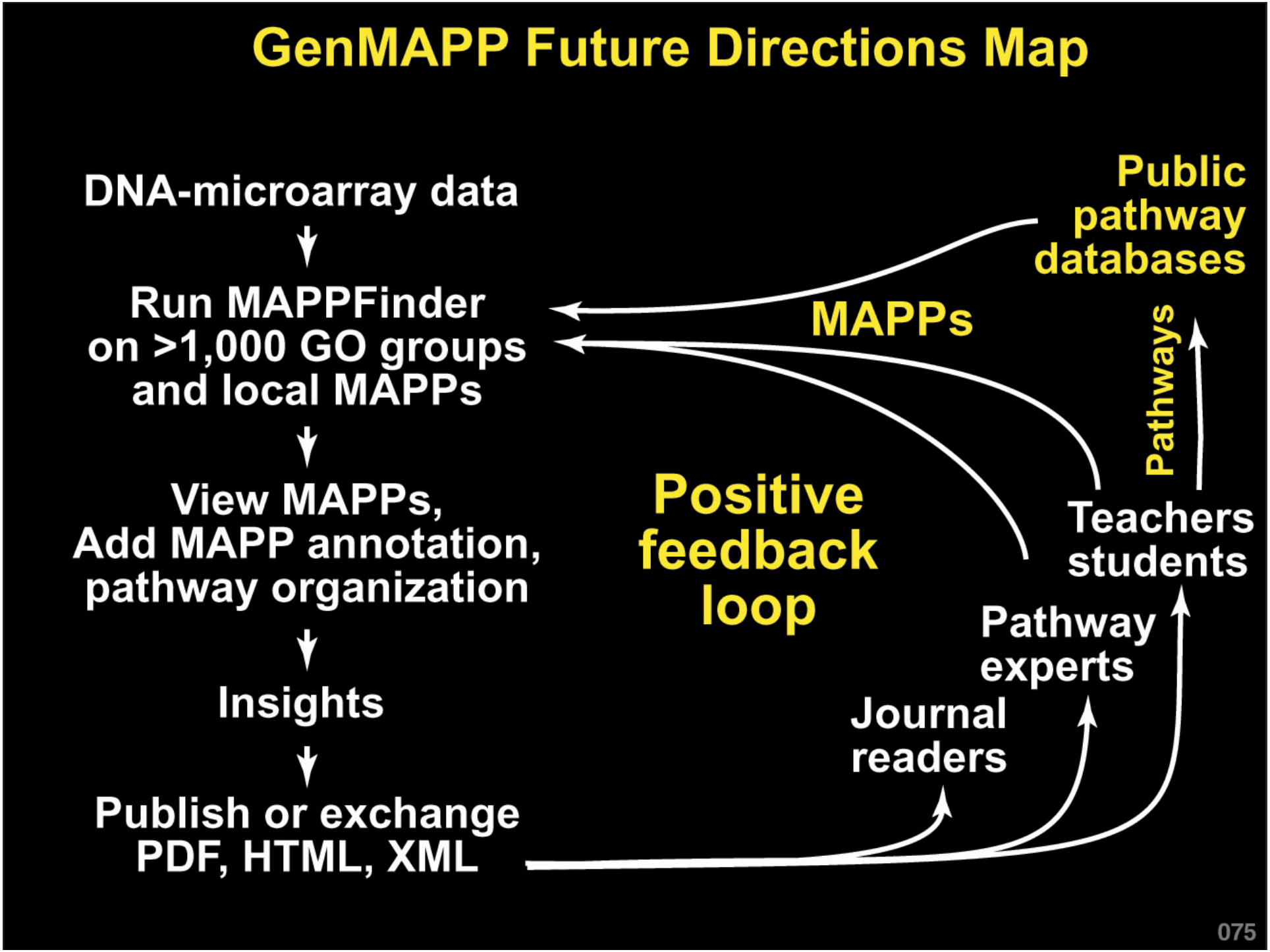
Insights

↓

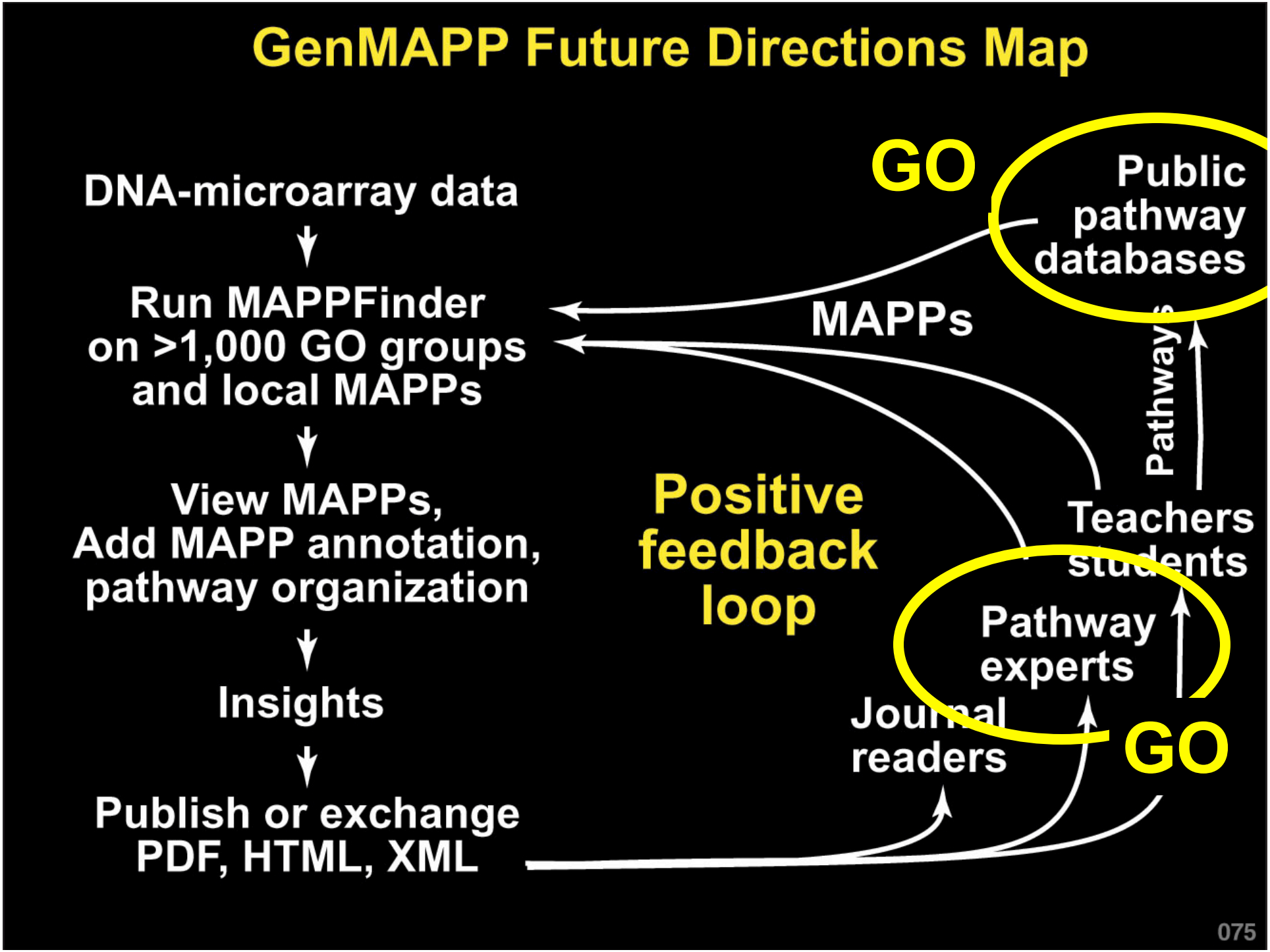
Publish or exchange
PDF, HTML, XML



GenMAPP Future Directions Map



GenMAPP Future Directions Map



Gladstone Institutes UC San Francisco

www.ConklinLab.org

RASSLs & ES cells

Peter Chang
Taeryn Kim
Kimberly Scearce-Levie
Supriya Srinivasan
Whit Tingley
Alex Zambon

Nathalie Cotte*
Mike Lieberman*
Robert Nissenson**
Christian Vaisse**

GenMAPP

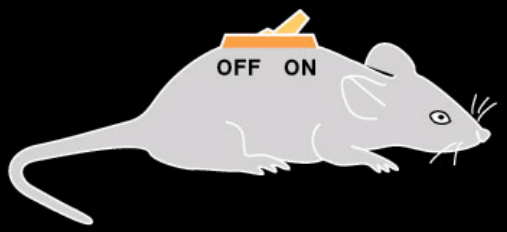
Lynn Ferrante
Kristina Hanspers
Steven Lawlor
Nathan Salomonis
Karen Vranizan

Kam Dahlquist*
Scott Doniger*
Katie Pollard**

BayGenomics

Bill Skarnes**
Stephen Young**
Pao-Tien Chuang**
Dean Sheppard**
Patsy Babbitt**
Tom Ferrin**

*recent alumni
**major collaborators



Funding

Gladstone
NIH, NHLBI
NHGRI