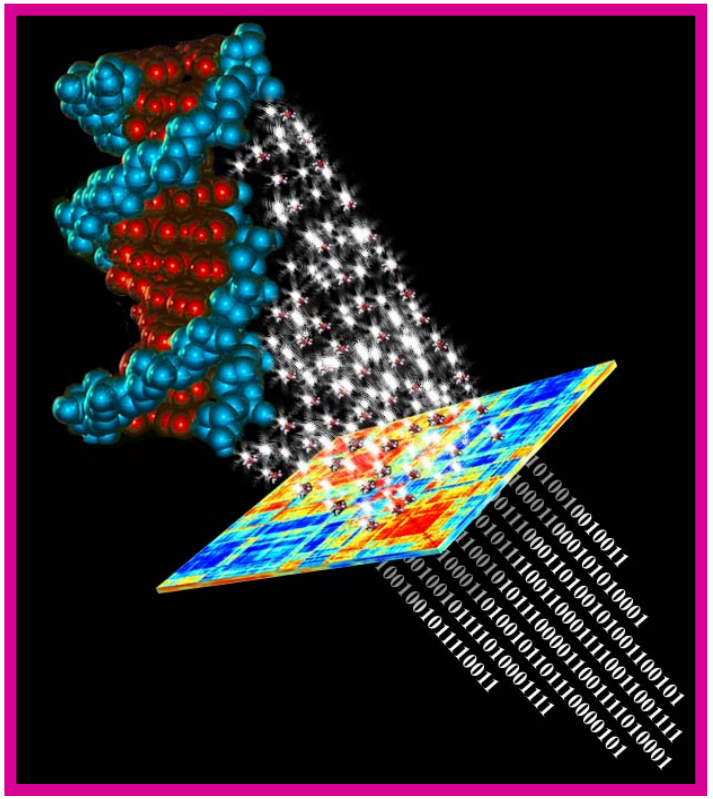
GoMiner and High-Throughput GoMiner: Leveraging the Gene Ontology for Microarrays and Other Omic Studies



Barry Zeeberg, *David Kane, John Weinstein, et al.

The Genomics & Bioinformatics Group, LMP, CCR, NCI

*** SRA International, Inc.**

GENOMICS AND BIOINFORMATICS GROUP

GoMiner

GoMiner batch-processes and organizes the gene lists from a microarray (or other genomic/proteomic platform) in the framework of the GO hierarchy.

It provides a number of outputs and visualizations to address the question, "Now that I've done my microarray experiment and analyzed the data statistically to identify a set of 'interesting' genes, what do those genes mean biologically?"

- GoMiner: Zeeberg, et al., Genome Biology 2003; 4:R28
- (MatchMiner: Bussey, et al., Genome Biology 2003; 4:R27)
- (MedMiner, CIMMiner, AbMiner)
- http://discover.nci.nih.gov/gominer

High-Throughput GoMiner is an extension that automatically processes an arbitrary number of arrays in a study and organizes the results in terms of a variety of outputs including "clustered image map" visuals.

GoMiner: Leverages the Gene Ontology

(Zeeberg, et al., Genome Biology 4: R28, 2003)

NCI/GA Tech - Emory Univ GoMiner

File Edit Data Source View

Input Genes(HUGO):D:\barry\gominer\BILL.GOOD

No.	Gene Name	Database	Changed
54	ATIC	mgi	G
55	ATM	mgi,rgd	G
56	ATP5J	mgi	G
57	ATR		↓
58	ATRX		G
59	AXL	mgi	↓
60	B2M	mgi,rgd	G
61	BAD	mgi	↓
62	BAG1	mgi	↑
63	BAK1	mgi	G
64	BCAT2	mgi	G
65	BCKDHA	mgi,rgd	G
66	BCL2	mgi,rgd	↑
67	BCL2A1		G
68	BCL2L1		G
69	BCL3		↑
70	BCL6	mgi	G
71	BCL7A		G
72	BCL7B		G
73	BCL9		G
74	BCR		G
75	BHLHB2		↑
76	BHMT	mgi	G
77	BIK		G
78	BIN1	mgi	G
79	BIRC2	mgi	↓
80	BIRC3	mgi	G

Genes Mapped On GO

Gene_Ontology (923 1.0 1.0 1.0 p= -)
- biological_process (788 1.0 1.0 1.0 p=0.886)
- cellular_component (644 1.0 1.0 1.0 p=0.824)
- molecular_function (837 1.0 1.0 1.0 p=1.000)
  - apoptosis regulator (14 3.2 2.0 2.4 p=0.070)
    - ↓ BAD(mgi)
    - ↑ BAG1(mgi)
    - G BAK1(mgi)
    - ↑ BCL2(mgi,rgd)
    - ↓ BIRC2(mgi)
    - G BIRC3(mgi)
    - G CASP1(mgi,rgd)
    - G CFLAR(mgi)
    - G CRADD(mgi)
    - G DFFA(mgi)
    - G FADD(mgi)
    - G PEA15(mgi)
  - apoptosis activator (1 0 0 0.0 p=1.000)

Selected Single Gene View

Gene_Ontology (923 1.0 1.0 1.0 p= -)
- cellular_component (644 1.0 1.0 1.0 p=0.824)
  - cell (583 1.0 1.0 1.0 p=1.000)
    - intracellular (486 1.1 1.0 1.1 p=0.414)
      - cytoplasm (124 1.3 1.0 1.1 p=0.653)
        - mitochondrion (22 0 2.5 1.6 p=0.313)
          - ↑ BCL2(mgi,rgd)
    - membrane (244 0.9 0.8 0.8 p=1.000)
      - integral membrane protein (217 0.8 0.8 0.8 p=1.000)
        - ↑ BCL2(mgi,rgd)
    - cell fraction (49 1.4 1.7 1.6 p=0.166)
      - membrane fraction (49 1.4 1.7 1.6 p=0.166)
        - ↑ BCL2(mgi,rgd)
- biological_process (788 1.0 1.0 1.0 p=0.886)
  - death (47 2.4 1.8 2.0 p=0.018)
    - cell death (44 2.6 1.9 2.1 p=0.013)
      - programmed cell death (44 2.6 1.9 2.1 p=0.013)

Operations

File Name D:\barry\gominer\BILL.GOOD\under.over.2

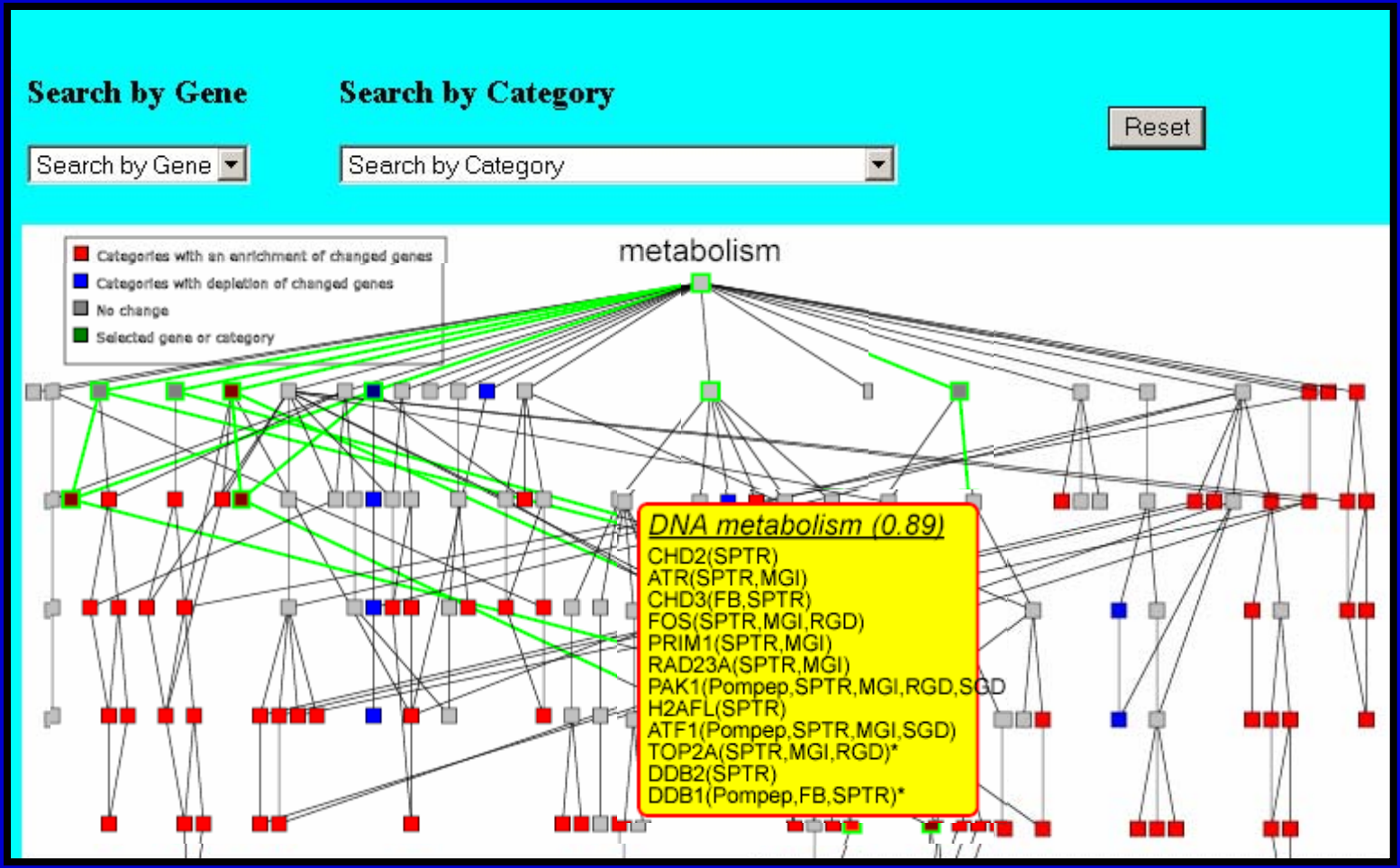
Query Gene File | Reset All | Query Changed Ge

molecular_function (837 1.0 1.0 1.0 p=1.000)
apoptosis regulator (14 3.2 2.0 2.4 p=0.070)
BAD(mgi)
BAG1(mgi)
BAK1(mgi)
BCL2(mgi,rgd)

(We thank M.D. Wang, G. Wang, and W. Feng, GA Tech/Emory U. for collaboration in the early development of GoMiner)

NFORMATICS GROUP

GoMiner: SVG Representation of DAG



Grab File Edit Capture Window Help Wed 15:50

NCI/GA Tech - Emory Univ GoMiner

File Edit Data Source Organisms View LookupSettings

Input Genes

Input	Gene	Source	Status
101F6	101F6		
182-FIP	182-FIP		
24432	24432		
76P	76P		
A2LP	A2LP		
AAMP	AAMP	SPTR	
AARS	AARS	SPTR	
AASD...	AASD...		
AATF	AATF		
ABCA7	ABCA7	SPTR	
ABCB6	ABCB6	SPTR	
ABCC10	ABCC10		
ABCC3	ABCC3	SPTR	
ABCE1	ABCE1		
ABCF1	ABCF1		

Genes Mapped On GO

Gene_Ontology (2203 1.00 p=1.00 1.00 p=1.00 1.00 p=1
- biological_process (2025 1.00 p=0.58 1.06 p=0.21
  - behavior (10 1.12 p=0.61 0.00 p=1.00 0.95 p=0
  - biological_process unknown (91 0.86 p=0.72 0.
  - cellular process (1092 1.11 p=0.07 1.50 p=0.00
    - cell communication (513 1.59 p=0.00 2.09 p
      - cell adhesion (96 2.33 p=0.00 1.97 p=0.
      - cell-cell signaling (61 2.20 p=0.01 9.29 p
      - signal transduction (435 1.49 p=0.00 2.1
        - ALCAM(SPTR)
        - C5R1(SPTR)
        - CASP1(SPTR)
        - CCL2(SPTR)
        - CCL20(SPTR)

Summary View | Selected Gene View

Category Name	P-Chng	P-Undr	P-Ovr	Tot	Chng	Undr	Ovr	Category ID
immunoglobulin b...	0.0000	0.0000	1.0000	7	6	6	0	GO:0019865
chemoattractant a...	0.0000	0.0701	0.0000	18	11	4	7	GO:0042056
di-, tri-valent inor...	0.0000	0.0002	0.1204	8	6	5	1	GO:0030005
cell-cell signaling	0.0000	0.0064	0.0000	61	21	12	9	GO:0007267
cell surface recept...	0.0000	0.0003	0.0005	154	36	27	9	GO:0007166
signal transduction	0.0000	0.0004	0.0014	435	73	58	15	GO:0007165
cell communication	0.0000	0.0000	0.0009	513	90	73	17	GO:0007154
innate immune res...	0.0000	0.0001	0.0000	62	24	16	8	GO:0045087
integral to plasma ...	0.0000	0.0000	0.0385	205	50	43	7	GO:0005887
plasma membrane	0.0000	0.0000	0.1613	287	69	62	7	GO:0005886
vacuole	0.0000	0.0000	1.0000	48	21	21	0	GO:0005773
lysosome	0.0000	0.0000	1.0000	47	21	21	0	GO:0005764
extracellular space	0.0000	0.0026	0.0000	48	20	11	9	GO:0005615
extracellular	0.0000	0.0018	0.0000	105	33	19	14	GO:0005576
sugar binding	0.0000	0.0005	0.0504	23	10	8	2	GO:0005529

Category Name	P-Chng	P-Undr	P-Ovr	Tot	Chng	Undr	O
immunoglobulin b...	0.0000	0.0000	1.0000	7	6	6	0
chemoattractant a...	0.0000	0.0701	0.0000	18	11	4	7
di-, tri-valent inor...	0.0000	0.0002	0.1204	8	6	5	1
cell-cell signaling	0.0000	0.0064	0.0000	61	21	12	9
cell surface recept...	0.0000	0.0003	0.0005	154	36	27	9
signal transduction	0.0000	0.0004	0.0014	435	73	58	15
cell communication	0.0000	0.0000	0.0009	513	90	73	17
innate immune res...	0.0000	0.0001	0.0000	62	24	16	8
integral to plasma ...	0.0000	0.0000	0.0385	205	50	43	7
plasma membrane	0.0000	0.0000	0.1613	287	69	62	7
vacuole	0.0000	0.0000	1.0000	48	21	21	0
lysosome	0.0000	0.0000	1.0000	47	21	21	0
extracellular space	0.0000	0.0026	0.0000	48	20	11	9

Query Gene File | Reset All | Query Changed Gene File | Reset Changed

Build:117

GENOMICS AND BIOINFORMATICS GROUP

High-Throughput Gominer

GoMiner Wrapper - Microsoft Internet Explorer provided by SRA International

File | Back Search Favorites Media

Links SRA Portal SRA Web ERP SRA Web Site SRA Webmail SRA Webtime

Address Go

High-Throughput Home | Web Interface | Command Line Interface | Input Files | Output Files | Error Codes | Paper

This is the web interface for High-Throughput GoMiner. You need to upload two files to the server, a total file with all of the genes in your experiment, and one of two types of changed files. Detailed descriptions of the input files are available. You will receive an email message with instructions on how to download your results once they are complete. Documentation for both the output files generated and possible error codes are also online.

Step 1: Select the total file Browse...

Input should be a text file with '.txt' extension

Step 2: Select the changed file Browse...

Accepts zip file with '.zip' or text file with '.txt' extension

Step 3: Select DataSource(s)

- All
- UniProt (H. sapiens et al.)
- CGEN (H. sapiens et al.)
- ENSEMBL (H. sapiens et al.)
- FB (D. melanogaster)

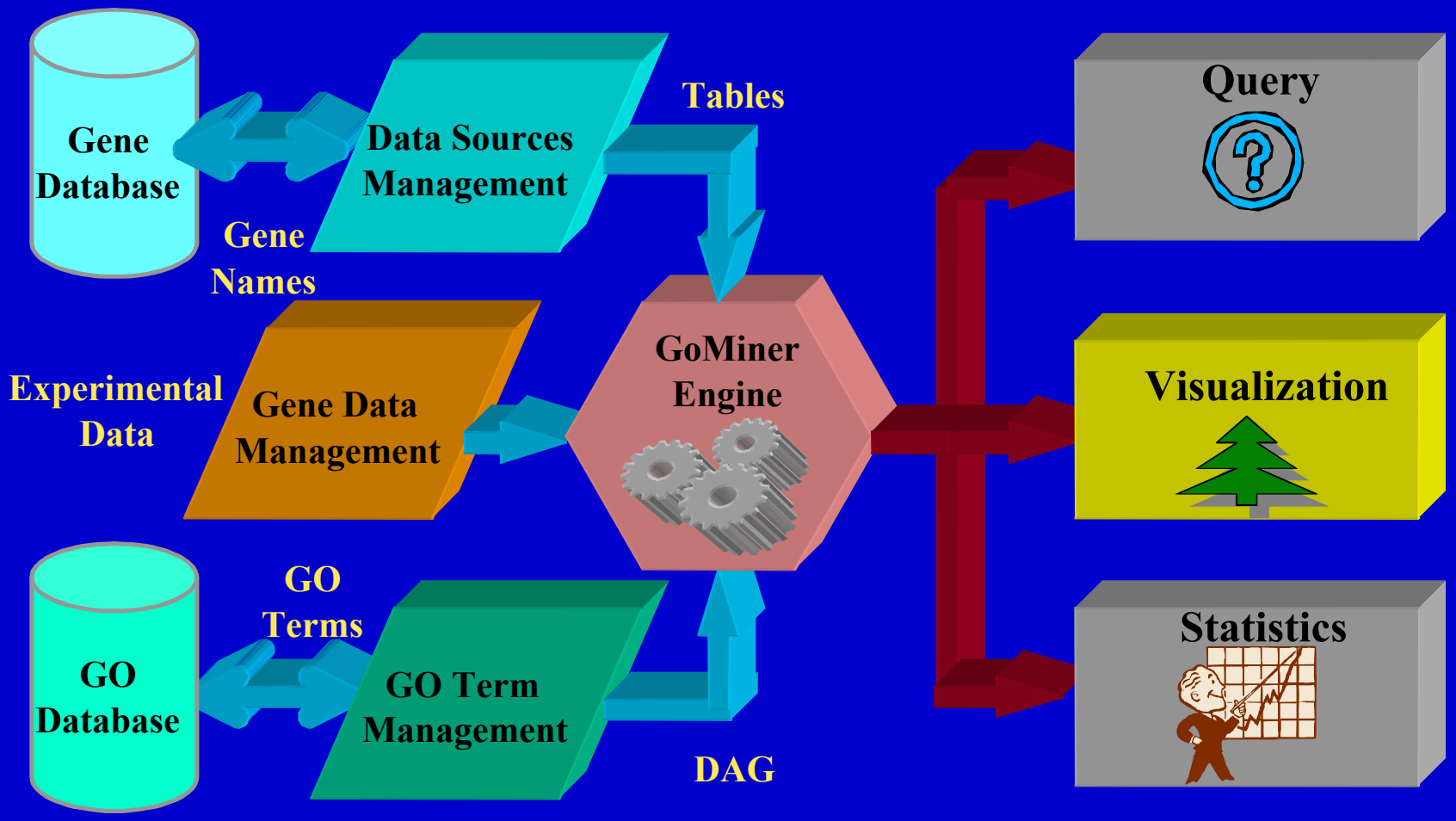
Step 4: Select Organism(s)

- All

Internet

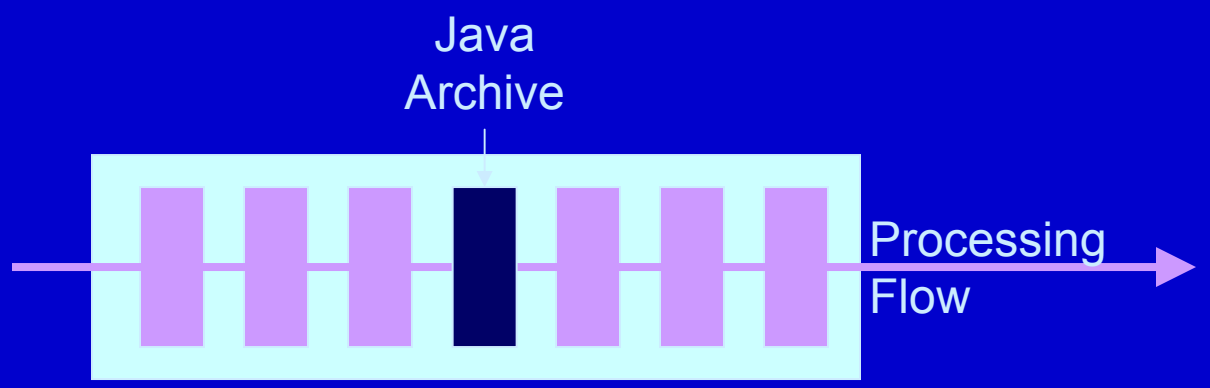
Start 2:27 PM

GoMiner Processing Flow

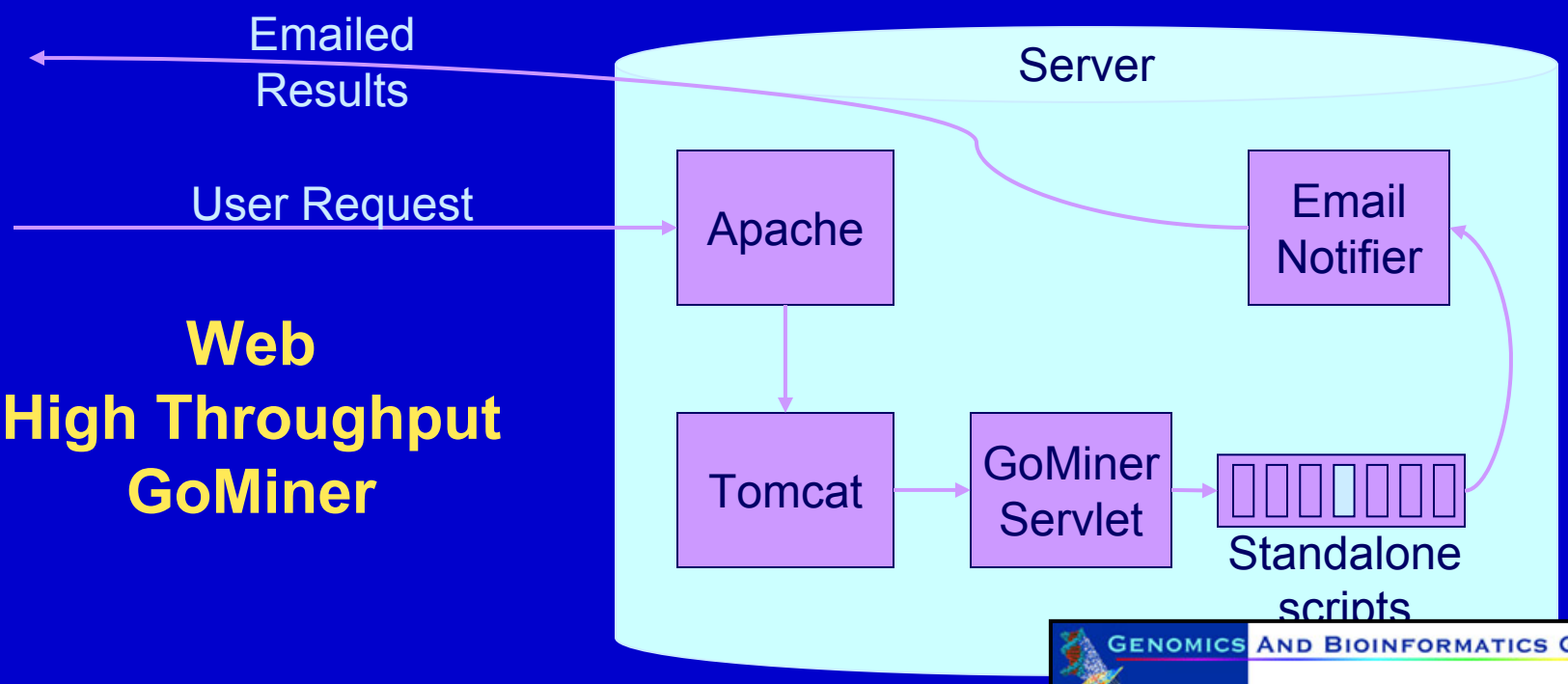


HT-GoMiner Processing Flow

Standalone High-Throughput GoMiner



Web High Throughput GoMiner



Leveraging caBIG Resources

- We expect GoMiner to use a number of caBIG Resources
  - Array Repositories
    - Browse available data sets
    - Select and format
    - Import into GoMiner
  - Alternative Ontologies
    - Analyze classifications in addition to GO
- At first, we expect to leverage existing interfaces
- Move to grid-based API's as they become available

Providing caBIG Resources

- HT-GoMiner's web-based interface provides a starting point for providing grid-enabled GoMiner web services
- GoMiner annotation analysis and annotation services can be provided to the grid as a service
  - Collaborating with NCI's ABCC to provide computational horsepower to offer such services

GoMiner and Vocabulary

- Many of the terms and concepts in GoMiner correspond directly with those used by the GO Consortium
  - There are some additional terms and concepts
- We hope this will make conceptual interoperability with caBIG easier to achieve