joanne@msl.ubc.ca

Bioinformatics

Common Tools & Tricks of the Trade

Welcome to Day 3

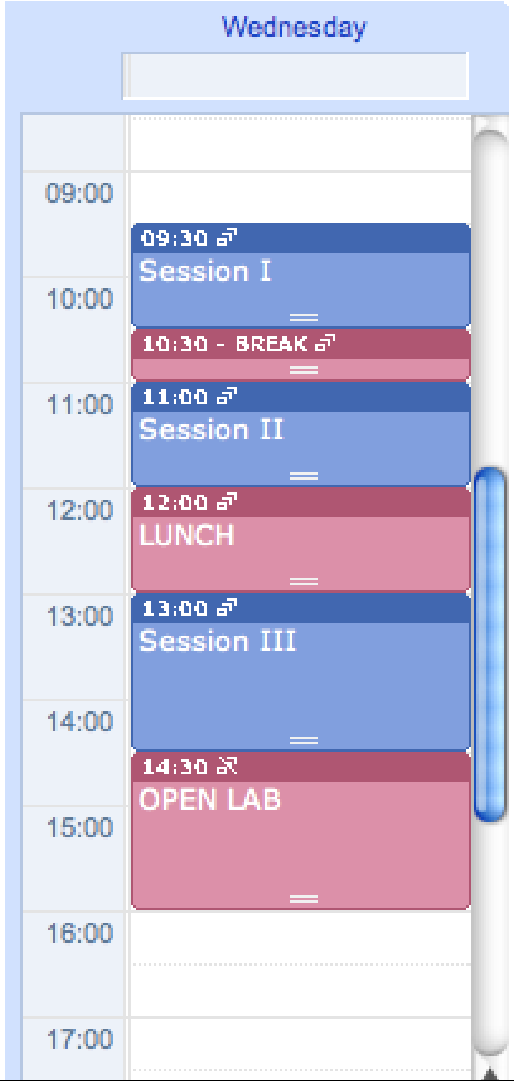
bioteach.ubc.ca/bioinfo2008

Workshop Schedule

- Laptops, available here for your use 9am - 4:30pm
- wireless login

  mslguest

  4myguest
- Vancouver guide books available



09:30 Session I

10:30 - BREAK

11:00 Session II

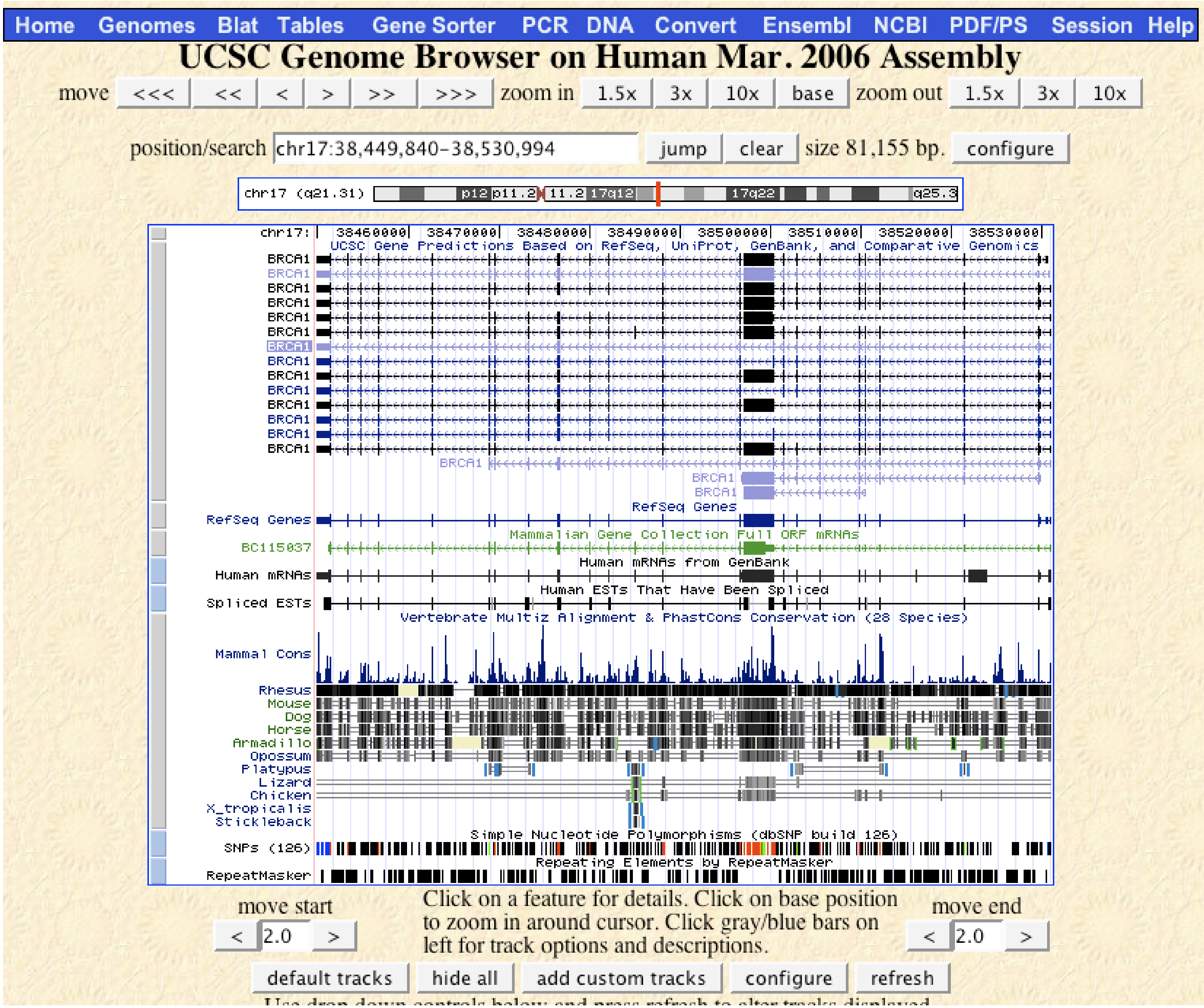
12:00 LUNCH

13:00 Session III

14:30 OPEN LAB

Today's Topics

- **PRACTICAL EXERCISES CONTINUED** Three different views of the BRCA1 gene
- **Discovering GEO**, the Gene Expression Omnibus.
- **Pathway Resources** for Systems Biology
- **Bioinformatics Links Directory**, Conducting Research on the Web
- **Let's go to the Pub!** ... and chat more informally about what you'd like to do next with bioinformatics



UCSC Genome Browser on Human Mar. 2006 Assembly

position/search chr17:38,449,840-38,530,994

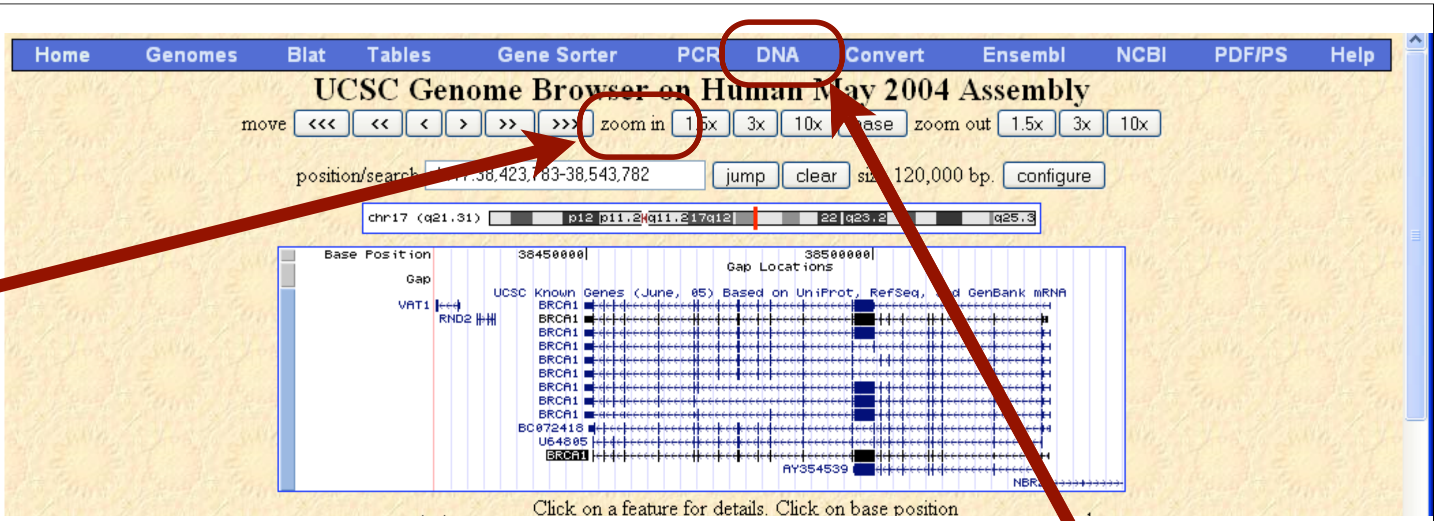
size 81,155 bp.

Click on a feature for details. Click on base position to zoom in around cursor. Click gray/blue bars on left for track options and descriptions.

Use drop down controls below and press refresh to alter tracks displayed.

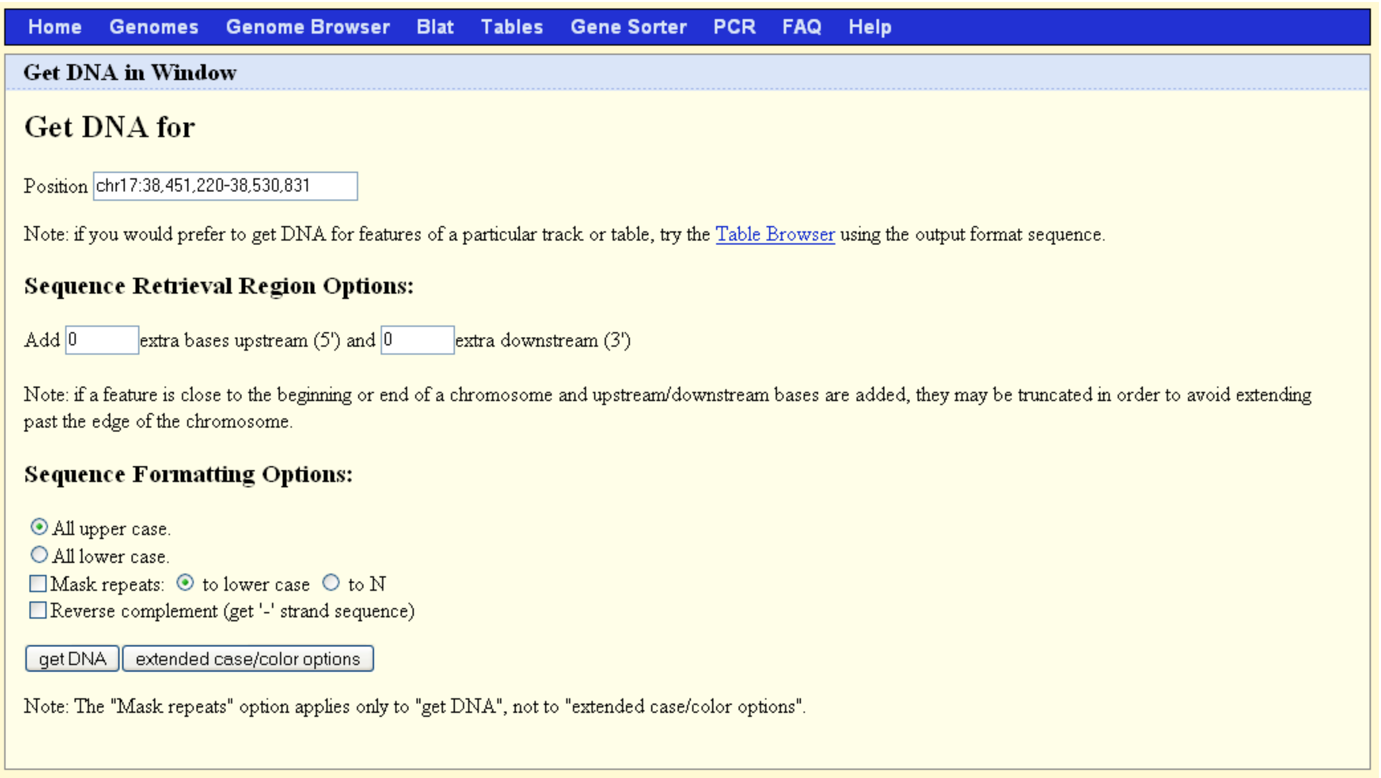
Two tasks

- What genes are on either side of BRCA1 on chr 17?
- Can you figure out how to download the genomic sequence for the BRCA1 region?



Zoom in
Zoom out

DNA link
Download
Sequence



Get DNA in Window

Get DNA for

Position chr17:38,451,220-38,530,831

Note: if you would prefer to get DNA for features of a particular track or table, try the Table Browser using the output format sequence.

Sequence Retrieval Region Options:

Add 0 extra bases upstream (5') and 0 extra downstream (3')

Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream bases are added, they may be truncated in order to avoid extending past the edge of the chromosome.

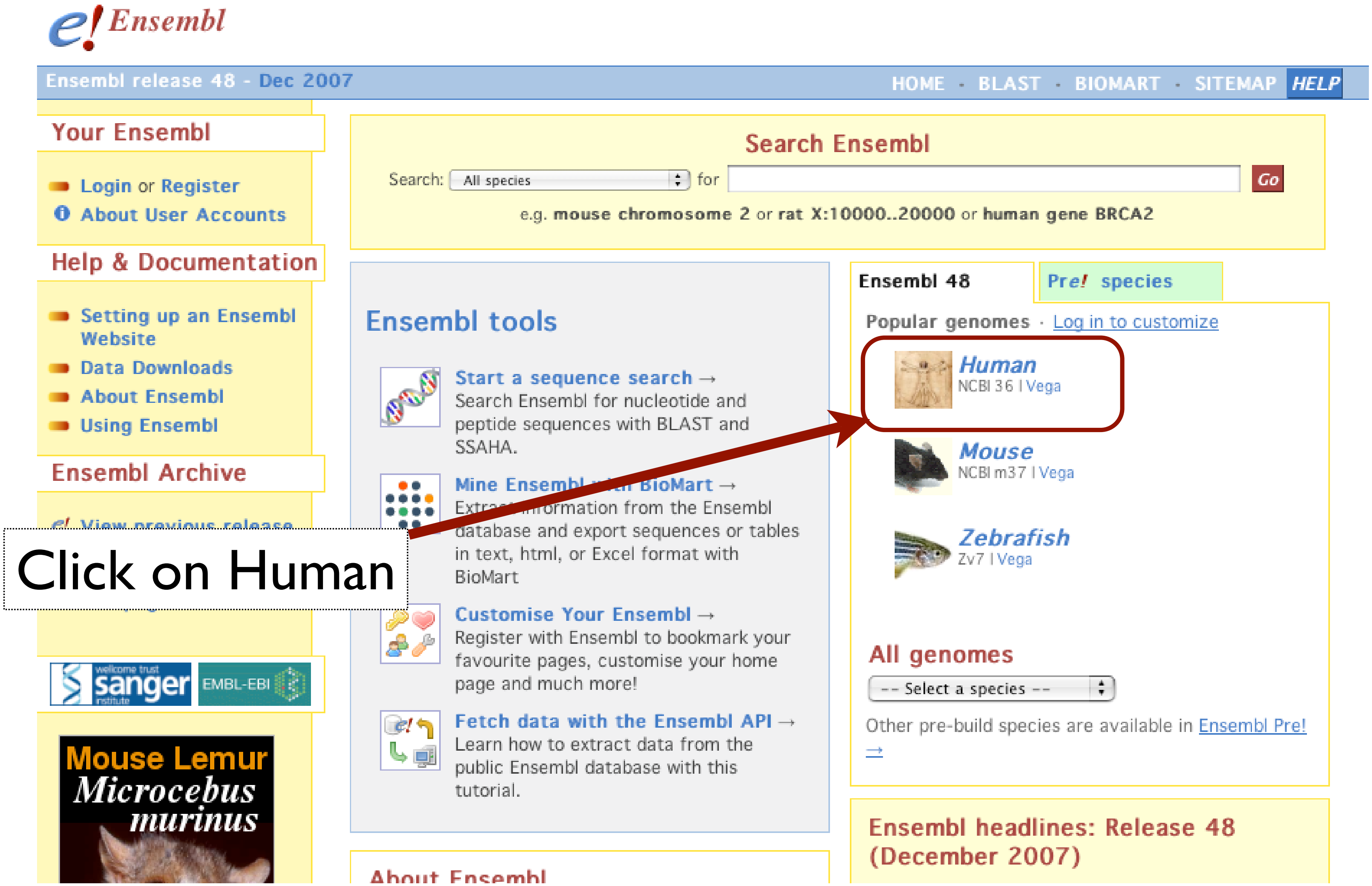
Sequence Formatting Options:

● All upper case.
○ All lower case.
☐ Mask repeats: ● to lower case ○ to N
☐ Reverse complement (get '-' strand sequence)

get DNA | extended case/color options

Note: The "Mask repeats" option applies only to "get DNA", not to "extended case/color options".

http://www.ensembl.org/



Your Ensembl

- Login or Register
- About User Accounts

Help & Documentation

- Setting up an Ensembl Website
- Data Downloads
- About Ensembl
- Using Ensembl

Ensembl Archive

- View previous release

Search Ensembl

Search: All species for Go

e.g. **mouse chromosome 2** or **rat X:10000..20000** or **human gene BRCA2**

Ensembl tools

Start a sequence search →
Search Ensembl for nucleotide and peptide sequences with BLAST and SSAHA.

Mine Ensembl with BioMart →
Extract information from the Ensembl database and export sequences or tables in text, html, or Excel format with BioMart

Customise Your Ensembl →
Register with Ensembl to bookmark your favourite pages, customise your home page and much more!

Fetch data with the Ensembl API →
Learn how to extract data from the public Ensembl database with this tutorial.

About Ensembl

Ensembl 48 | Pre! species

Popular genomes · Log in to customize

Human NCBI 36 | Vega

Mouse NCBI m37 | Vega

Zebrafish Zv7 | Vega

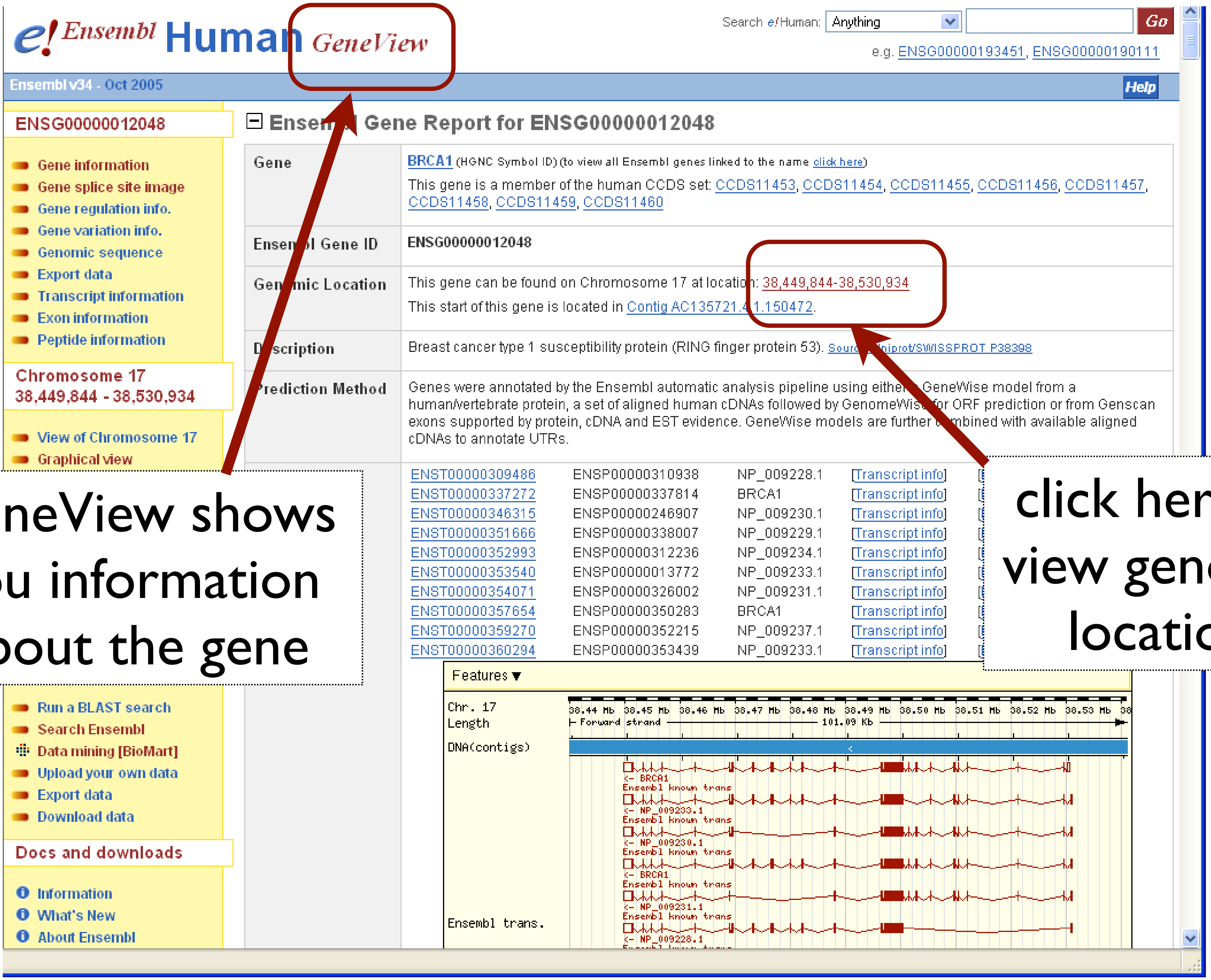
All genomes

-- Select a species --

Other pre-build species are available in Ensembl Pre! →

Ensembl headlines: Release 48 (December 2007)

Click on Human



e.g. ENSG00000193451, ENSG00000190111

Ensembl v34 - Oct 2005

ENSG00000012048

- Gene information
- Gene splice site image
- Gene regulation info.
- Gene variation info.
- Genomic sequence
- Export data
- Transcript information
- Exon information
- Peptide information

Chromosome 17
38,449,844 - 38,530,934

- View of Chromosome 17
- Graphical view

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Export data
- Download data

Docs and downloads

- Information
- What's New
- About Ensembl

⊟ Ensembl Gene Report for ENSG00000012048

Gene	BRCA1 (HGNC Symbol ID) (to view all Ensembl genes linked to the name click here) This gene is a member of the human CCDS set: CCDS11453, CCDS11454, CCDS11455, CCDS11456, CCDS11457, CCDS11458, CCDS11459, CCDS11460
Ensembl Gene ID	ENSG00000012048
Genomic Location	This gene can be found on Chromosome 17 at location: 38,449,844-38,530,934 This start of this gene is located in Contig AC135721.4 1.150472.
Description	Breast cancer type 1 susceptibility protein (RING finger protein 53). Source: Uniprot/SWISSPROT P38398
Prediction Method	Genes were annotated by the Ensembl automatic analysis pipeline using either a GeneWise model from a human/vertebrate protein, a set of aligned human cDNAs followed by GenomeWise for ORF prediction or from Genscan exons supported by protein, cDNA and EST evidence. GeneWise models are further combined with available aligned cDNAs to annotate UTRs.

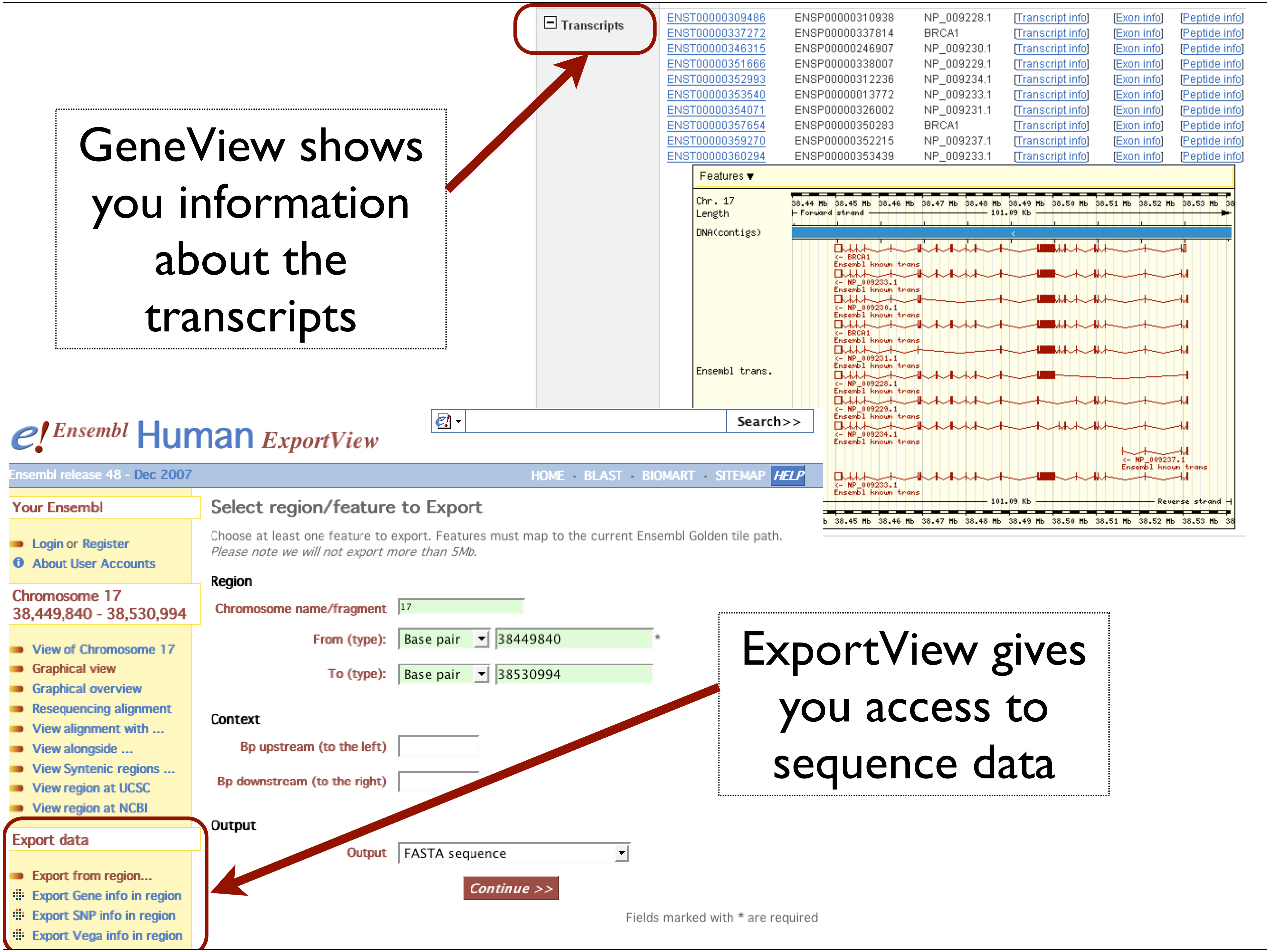
ENST00000309486	ENSP00000310938	NP_009228.1	[Transcript info]
ENST00000337272	ENSP00000337814	BRCA1	[Transcript info]
ENST00000346315	ENSP00000246907	NP_009230.1	[Transcript info]
ENST00000351666	ENSP00000338007	NP_009229.1	[Transcript info]
ENST00000352993	ENSP00000312236	NP_009234.1	[Transcript info]
ENST00000353540	ENSP00000013772	NP_009233.1	[Transcript info]
ENST00000354071	ENSP00000326002	NP_009231.1	[Transcript info]
ENST00000357654	ENSP00000350283	BRCA1	[Transcript info]
ENST00000359270	ENSP00000352215	NP_009237.1	[Transcript info]
ENST00000360294	ENSP00000353439	NP_009233.1	[Transcript info]

GeneView shows you information about the gene

click here to view genomic location

Two tasks

- Using GeneView, can you figure out how many different alternatively spliced isoforms exist for BRCA1?
- Using ContigView, can you figure out how to download the genomic sequence for the BRCA1 region?



GeneView shows you information about the transcripts

Transcripts

ENST00000309486	ENSP00000310938	NP_009228.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000337272	ENSP00000337814	BRCA1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000346315	ENSP00000246907	NP_009230.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000351666	ENSP00000338007	NP_009229.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000352993	ENSP00000312236	NP_009234.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000353540	ENSP00000013772	NP_009233.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000354071	ENSP00000326002	NP_009231.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000357654	ENSP00000350283	BRCA1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000359270	ENSP00000352215	NP_009237.1	[Transcript info]	[Exon info]	[Peptide info]
ENST00000360294	ENSP00000353439	NP_009233.1	[Transcript info]	[Exon info]	[Peptide info]

Select region/feature to Export

Choose at least one feature to export. Features must map to the current Ensembl Golden tile path.
Please note we will not export more than 5Mb.

Region

Chromosome name/fragment: 17

From (type): Base pair 38449840 *

To (type): Base pair 38530994

Context

Bp upstream (to the left):

Bp downstream (to the right):

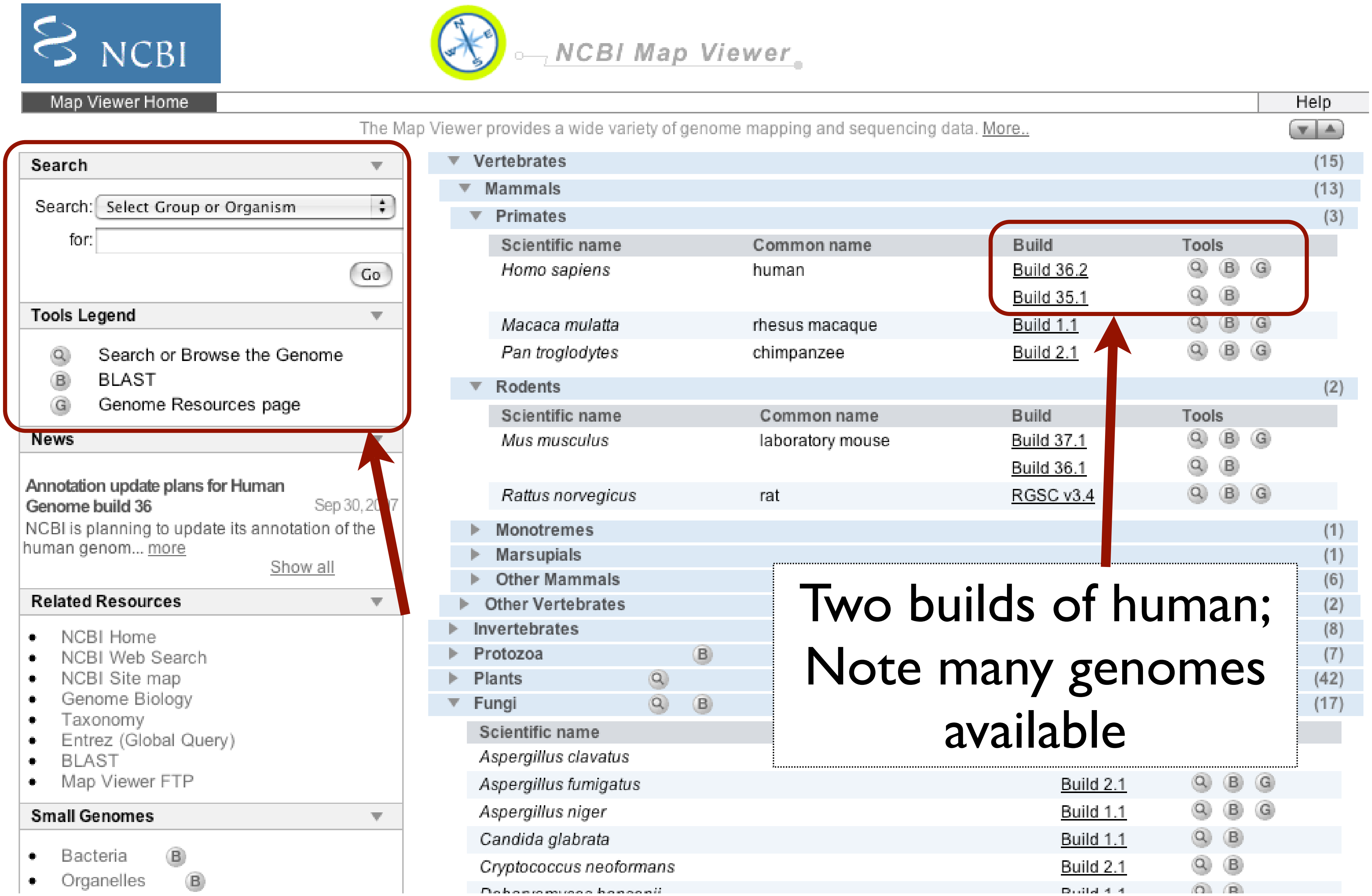
Output

Output: FASTA sequence

Fields marked with * are required

ExportView gives you access to sequence data

http://www.ncbi.nlm.nih.gov/mapview/



The Map Viewer provides a wide variety of genome mapping and sequencing data. More..

Search

Search: Select Group or Organism

for:

Go

Tools Legend

🔍	Search or Browse the Genome
B	BLAST
G	Genome Resources page

News

Annotation update plans for Human Genome build 36 Sep 30, 20[illegible]7

NCBI is planning to update its annotation of the human genom... more

Show all

Related Resources

- NCBI Home
- NCBI Web Search
- NCBI Site map
- Genome Biology
- Taxonomy
- Entrez (Global Query)
- BLAST
- Map Viewer FTP

Small Genomes

- Bacteria B
- Organelles B

Vertebrates (15)

Mammals (13)

Primates (3)

Scientific name	Common name	Build	Tools
Homo sapiens	human	Build 36.2	🔍 B G
		Build 35.1	🔍 B
Macaca mulatta	rhesus macaque	Build 1.1	🔍 B G
Pan troglodytes	chimpanzee	Build 2.1	🔍 B G

Rodents (2)

Scientific name	Common name	Build	Tools
Mus musculus	laboratory mouse	Build 37.1	🔍 B G
		Build 36.1	🔍 B
Rattus norvegicus	rat	RGSC v3.4	🔍 B G

Monotremes (1)

Marsupials (1)

Other Mammals (6)

Other Vertebrates (2)

Invertebrates (8)

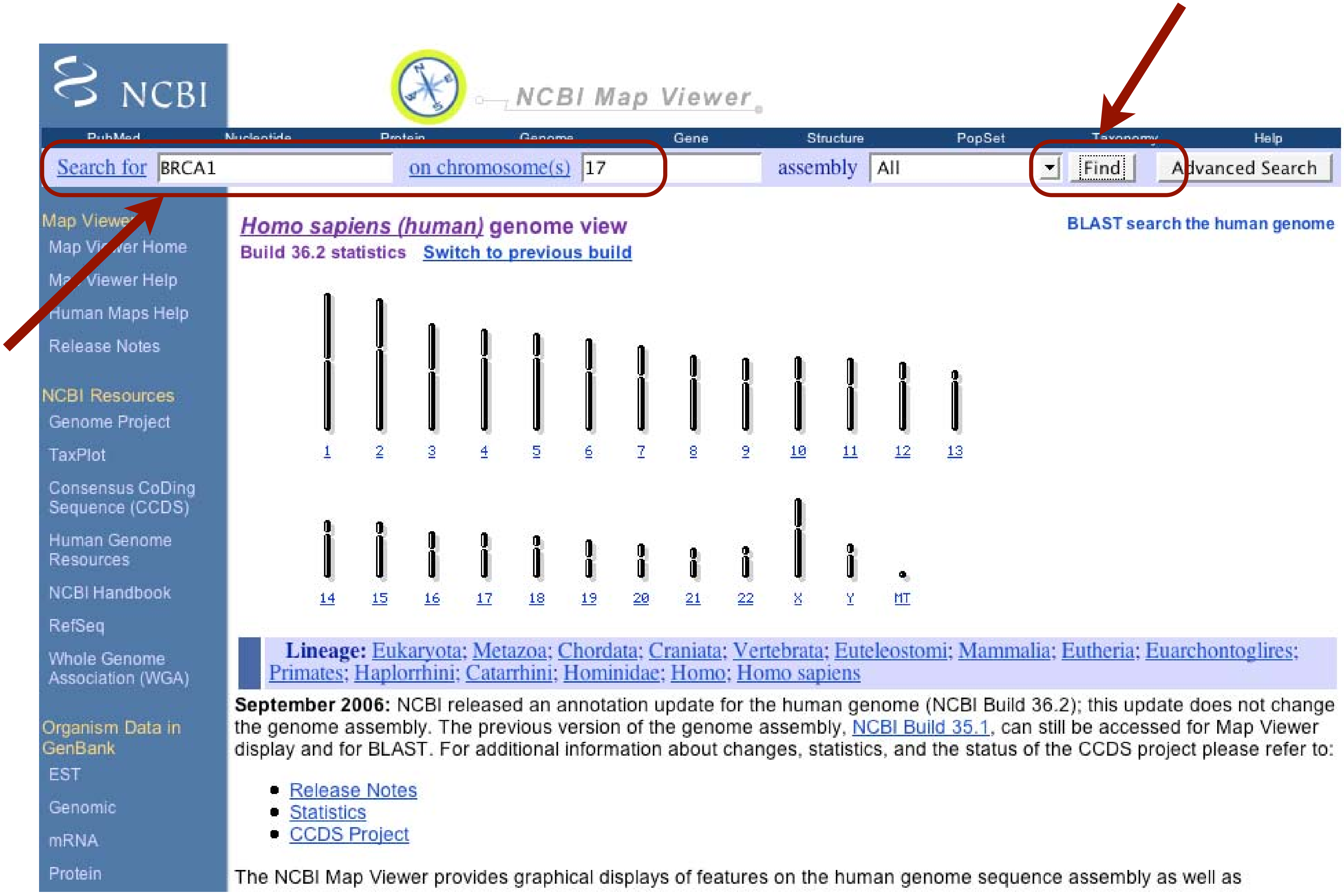
Protozoa B (7)

Plants 🔍 (42)

Fungi 🔍 B (17)

Scientific name	Build	Tools
Aspergillus clavatus		
Aspergillus fumigatus	Build 2.1	🔍 B G
Aspergillus niger	Build 1.1	🔍 B G
Candida glabrata	Build 1.1	🔍 B
Cryptococcus neoformans	Build 2.1	🔍 B
Debaryomyces hansenii	Build 1.1	🔍 B

Two builds of human; Note many genomes available



Search for: BRCA1

on chromosome(s): 17

assembly: All

Homo sapiens (human) genome view

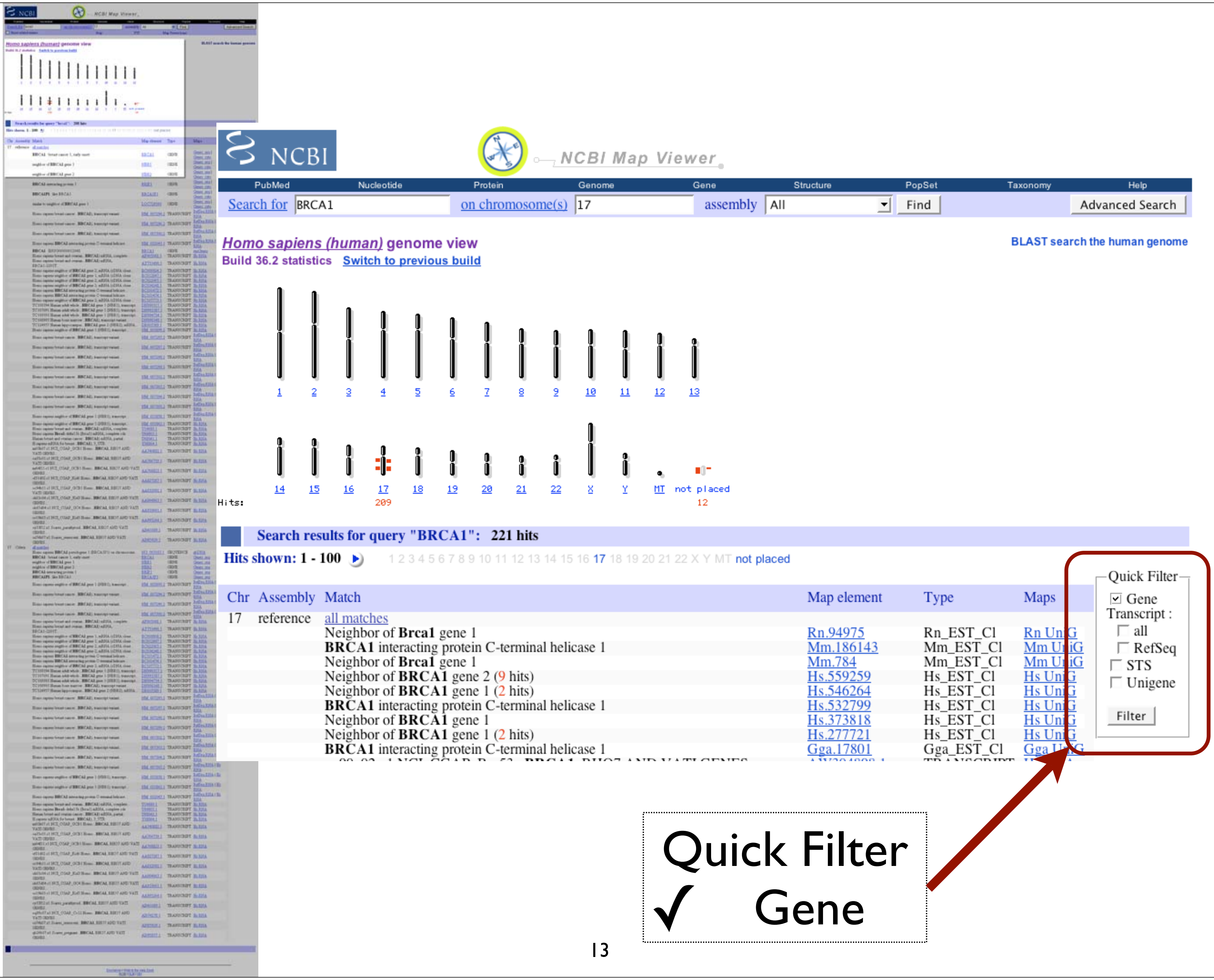
Build 36.2 statistics   Switch to previous build

Lineage: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo; Homo sapiens

September 2006: NCBI released an annotation update for the human genome (NCBI Build 36.2); this update does not change the genome assembly. The previous version of the genome assembly, NCBI Build 35.1, can still be accessed for Map Viewer display and for BLAST. For additional information about changes, statistics, and the status of the CCDS project please refer to:

- Release Notes
- Statistics
- CCDS Project

The NCBI Map Viewer provides graphical displays of features on the human genome sequence assembly as well as



Search for: BRCA1 on chromosome(s): 17 assembly: All

Homo sapiens (human) genome view

Build 36.2 statistics Switch to previous build

BLAST search the human genome

Search results for query "BRCA1": 221 hits

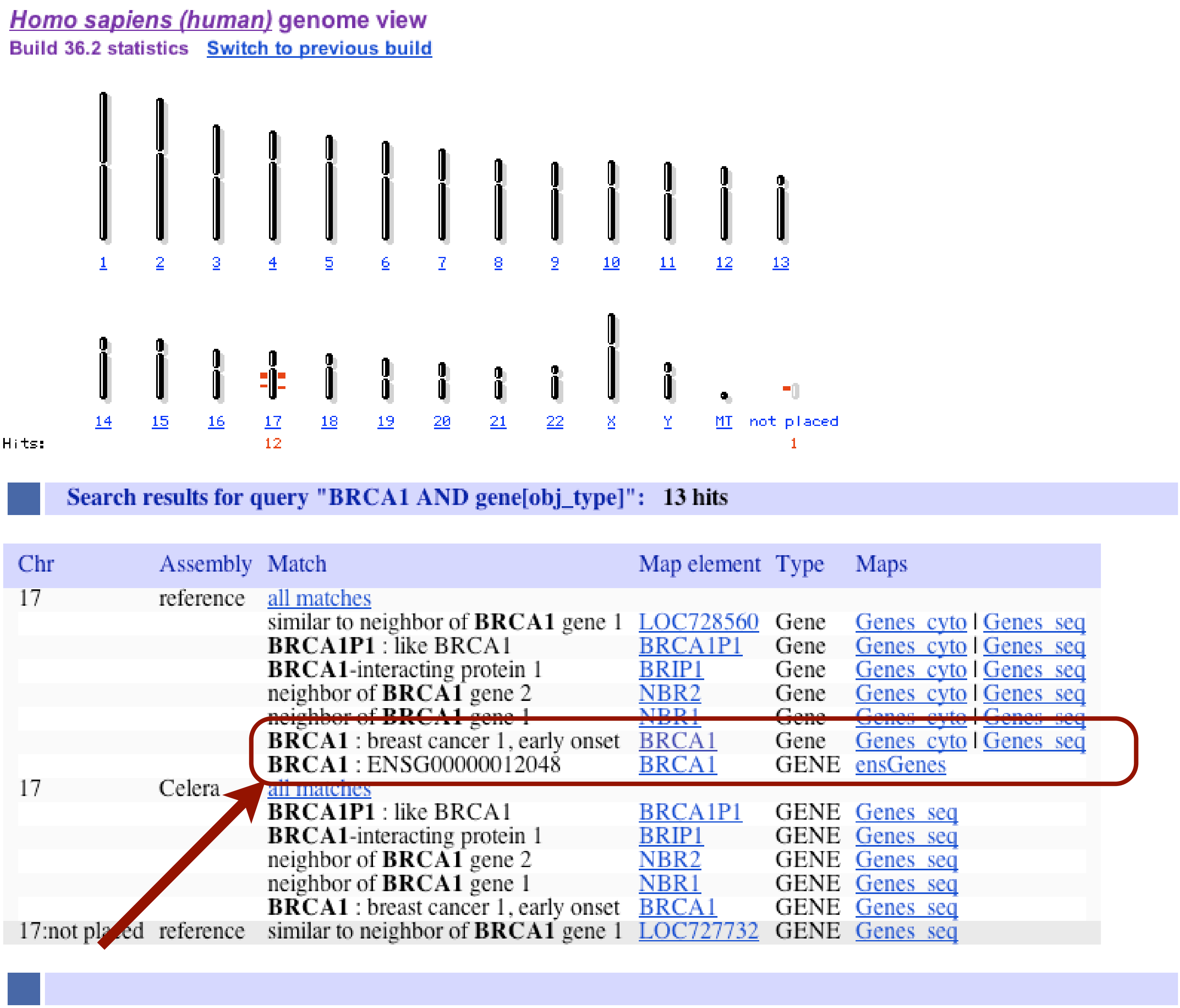
Hits shown: 1 - 100

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT not placed

Chr	Assembly	Match	Map element	Type	Maps
17	reference	all matches			
		Neighbor of **Brca1** gene 1	Rn.94975	Rn_EST_Cl	Rn Un G
		BRCA1 interacting protein C-terminal helicase 1	Mm.186143	Mm_EST_Cl	Mm U iG
		Neighbor of **Brca1** gene 1	Mm.784	Mm_EST_Cl	Mm U iG
		Neighbor of **BRCA1** gene 2 (9 hits)	Hs.559259	Hs_EST_Cl	Hs Uni G
		Neighbor of **BRCA1** gene 1 (2 hits)	Hs.546264	Hs_EST_Cl	Hs Uni G
		BRCA1 interacting protein C-terminal helicase 1	Hs.532799	Hs_EST_Cl	Hs Uni G
		Neighbor of **BRCA1** gene 1	Hs.373818	Hs_EST_Cl	Hs Uni G
		Neighbor of **BRCA1** gene 1 (2 hits)	Hs.277721	Hs_EST_Cl	Hs Uni G
		BRCA1 interacting protein C-terminal helicase 1	Gga.17801	Gga_EST_Cl	Gga U

Quick Filter
☑ Gene
Transcript :
☐ all
☐ RefSeq
☐ STS
☐ Unigene

Quick Filter
✓ Gene

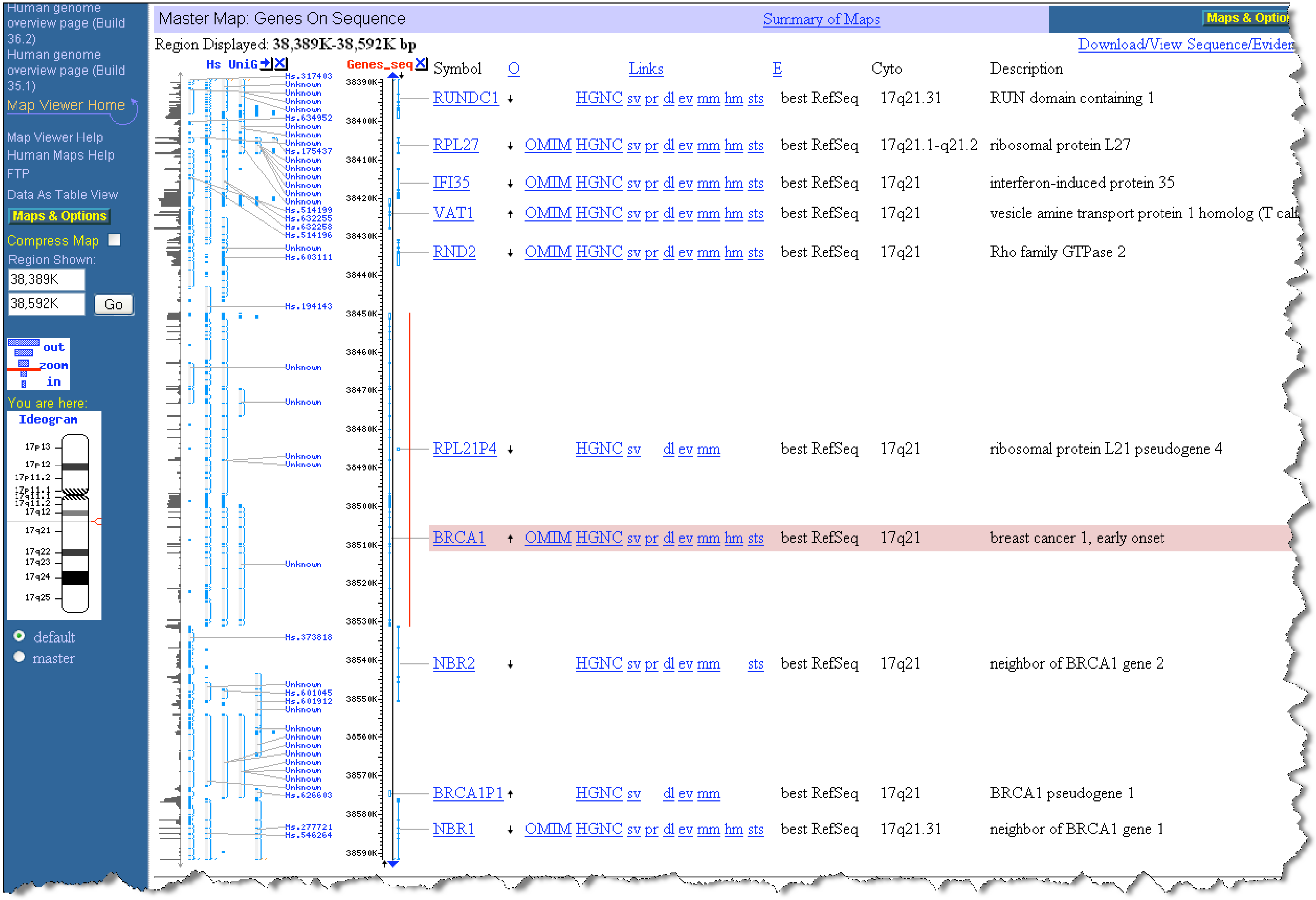


Homo sapiens (human) genome view

Build 36.2 statistics Switch to previous build

Search results for query "BRCA1 AND gene[obj_type]": 13 hits

Chr	Assembly	Match	Map element	Type	Maps
17	reference	all matches			
		similar to neighbor of **BRCA1** gene 1	LOC728560	Gene	Genes_cyto | Genes_seq
		BRCA1P1 : like BRCA1	BRCA1P1	Gene	Genes_cyto | Genes_seq
		BRCA1-interacting protein 1	BRIP1	Gene	Genes_cyto | Genes_seq
		neighbor of **BRCA1** gene 2	NBR2	Gene	Genes_cyto | Genes_seq
		neighbor of **BRCA1** gene 1	NBR1	Gene	Genes_cyto | Genes_seq
		BRCA1 : breast cancer 1, early onset	BRCA1	Gene	Genes_cyto | Genes_seq
		BRCA1 : ENSG00000012048	BRCA1	GENE	ensGenes
17	Celera	all matches			
		BRCA1P1 : like BRCA1	BRCA1P1	GENE	Genes_seq
		BRCA1-interacting protein 1	BRIP1	GENE	Genes_seq
		neighbor of **BRCA1** gene 2	NBR2	GENE	Genes_seq
		neighbor of **BRCA1** gene 1	NBR1	GENE	Genes_seq
		BRCA1 : breast cancer 1, early onset	BRCA1	GENE	Genes_seq
17:not placed	reference	similar to neighbor of **BRCA1** gene 1	LOC727732	GENE	Genes_seq



Master Map: Genes On Sequence

Region Displayed: 38,389K-38,592K bp

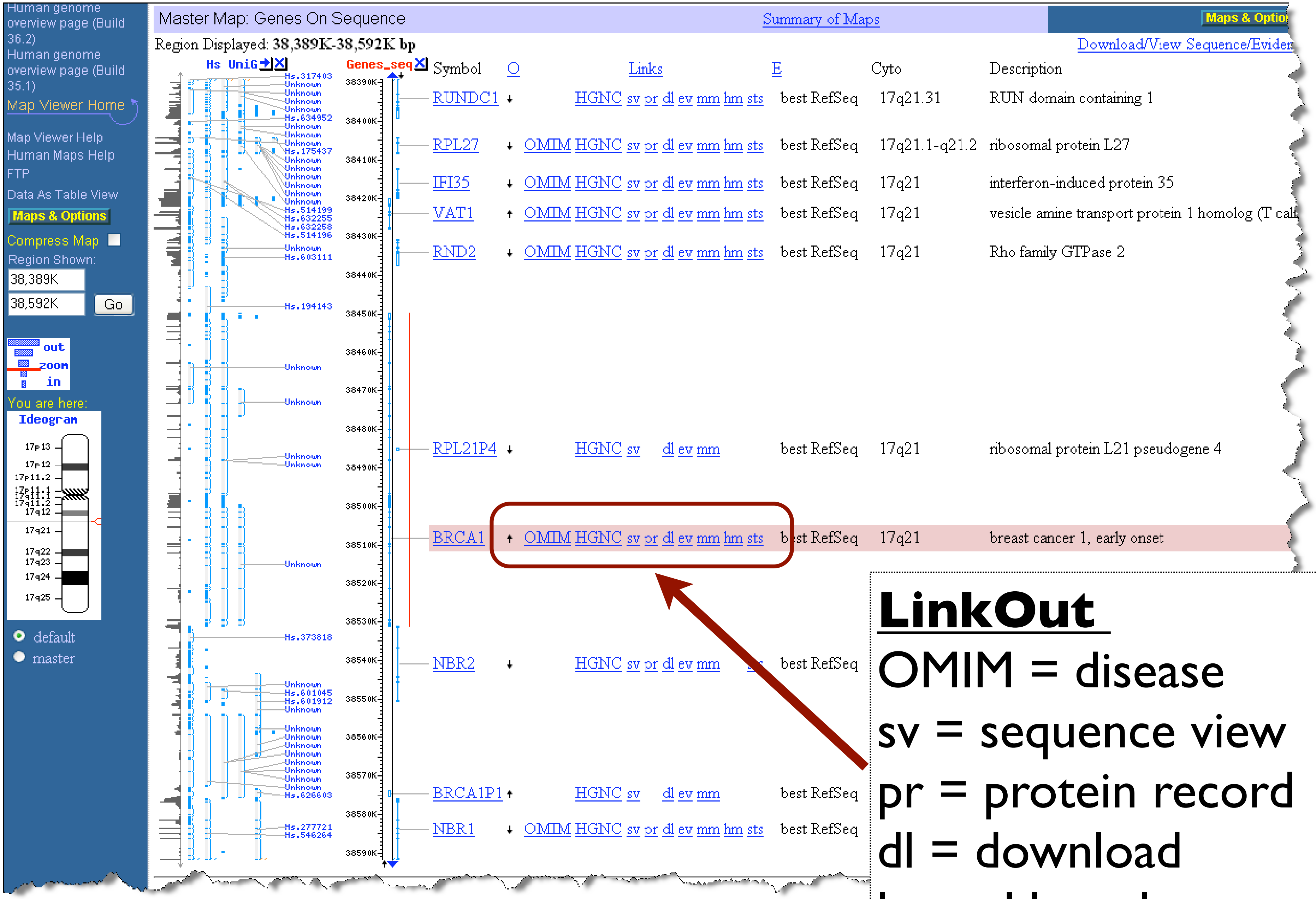
Region Shown: 38,389K – 38,592K

● default
○ master

Symbol	O	Links	E	Cyto	Description
RUNDC1	↓	HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.31	RUN domain containing 1
RPL27	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.1-q21.2	ribosomal protein L27
IFI35	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	interferon-induced protein 35
VAT1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	vesicle amine transport protein 1 homolog (T cali
RND2	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	Rho family GTPase 2
RPL21P4	↓	HGNC sv dl ev mm	best RefSeq	17q21	ribosomal protein L21 pseudogene 4
BRCA1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	breast cancer 1, early onset
NBR2	↓	HGNC sv pr dl ev mm sts	best RefSeq	17q21	neighbor of BRCA1 gene 2
BRCA1P1	↑	HGNC sv dl ev mm	best RefSeq	17q21	BRCA1 pseudogene 1
NBR1	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.31	neighbor of BRCA1 gene 1

Two tasks

- Can you figure out how to LinkOut to the OMIM and/or Homologene entries for BRCA1?
- Can you figure out how to download the genomic sequence for the BRCA1 region?



Master Map: Genes On Sequence

Region Displayed: 38,389K-38,592K bp

Symbol	O	Links	E	Cyto	Description
RUNDC1	↓	HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.31	RUN domain containing 1
RPL27	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21.1-q21.2	ribosomal protein L27
IFI35	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	interferon-induced protein 35
VAT1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	vesicle amine transport protein 1 homolog (T cali
RND2	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	Rho family GTPase 2
RPL21P4	↓	HGNC sv dl ev mm	best RefSeq	17q21	ribosomal protein L21 pseudogene 4
BRCA1	↑	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq	17q21	breast cancer 1, early onset
NBR2	↓	HGNC sv pr dl ev mm	best RefSeq		
BRCA1P1	↑	HGNC sv dl ev mm	best RefSeq		
NBR1	↓	OMIM HGNC sv pr dl ev mm hm sts	best RefSeq		

LinkOut

OMIM = disease

sv = sequence view

pr = protein record

dl = download

hm = Homologene

Credits

- UCSC Genome Browser

  http://genome.cse.ucsc.edu/

- Ensembl Genome Browser

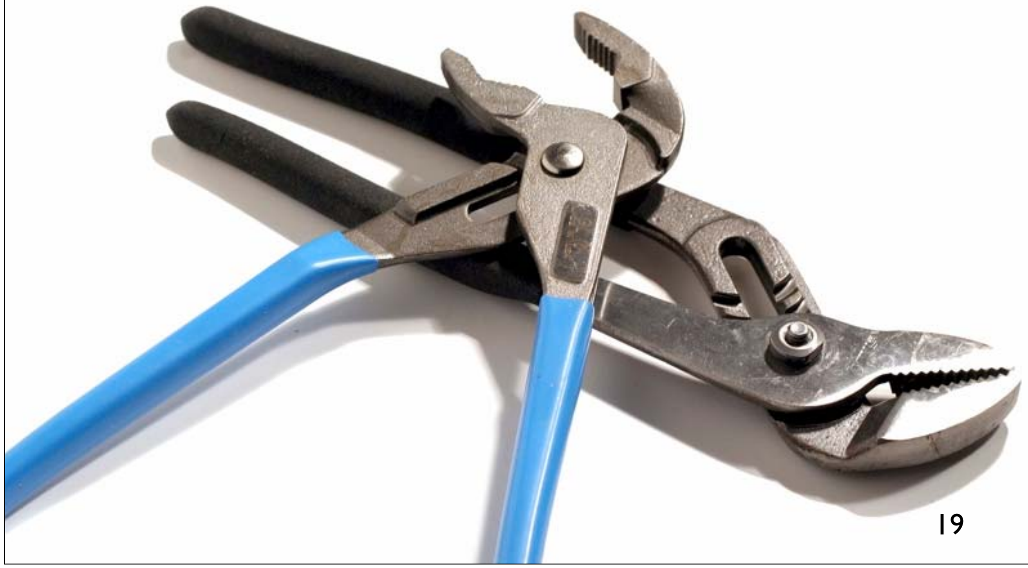
  http://www.ensembl.org/index.html

- NCBI MapViewer

  http://www.ncbi.nlm.nih.gov/mapview/index.html

Bioinformatics

Session 3.1 - Discovering GEO, the Gene Expression Omnibus.



Functional Genomics

- What kinds of questions can you ask with microarray data?

✓ basic research

✓ drug target discovery

✓ biomarker discovery

✓ pharmacology & toxicogenomics

✓ clinical diagnosis - prognosis, diagnosis, & disease classification

✓ gene regulatory networks

✓ protein-DNA binding

✓ + more

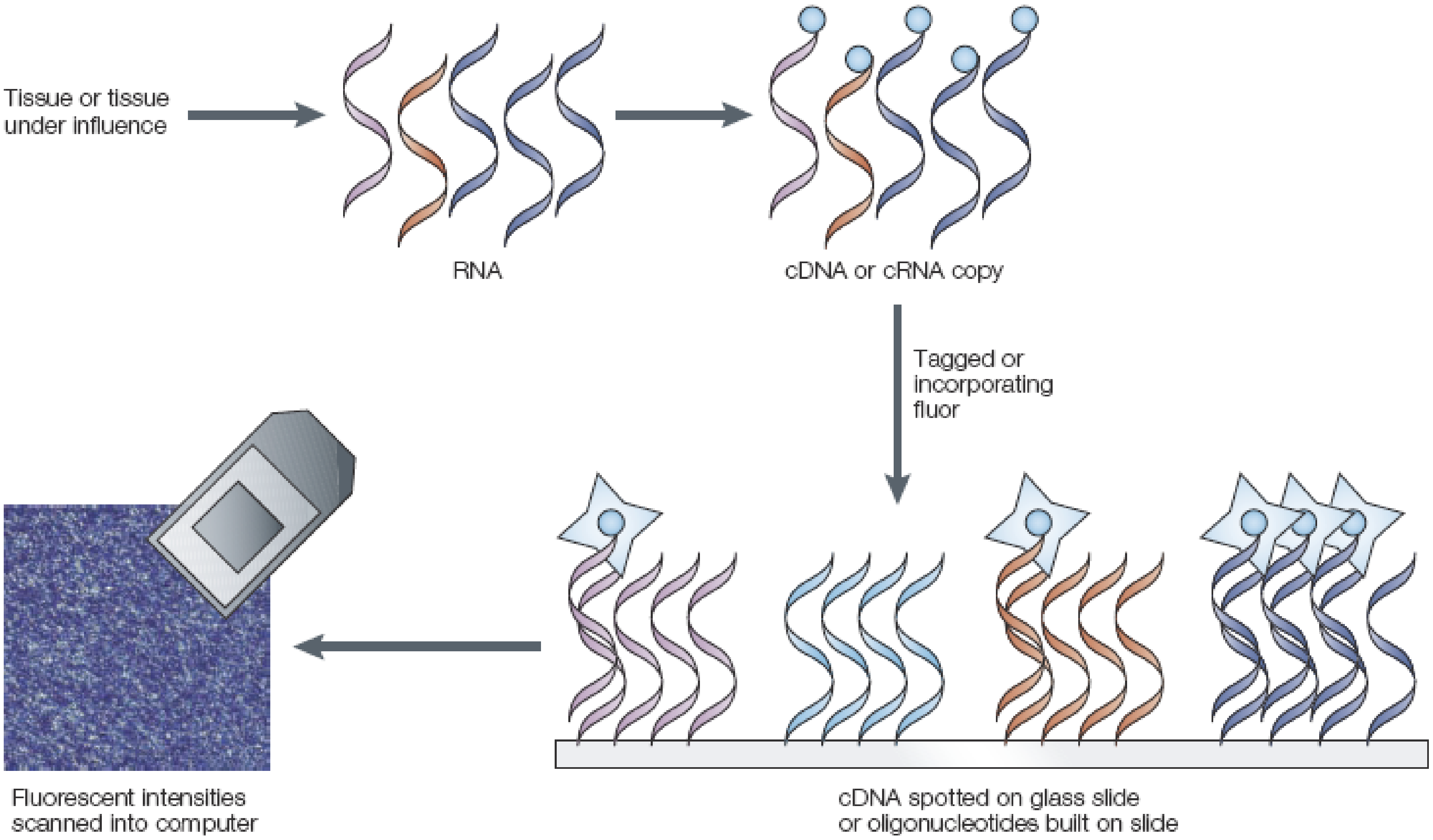


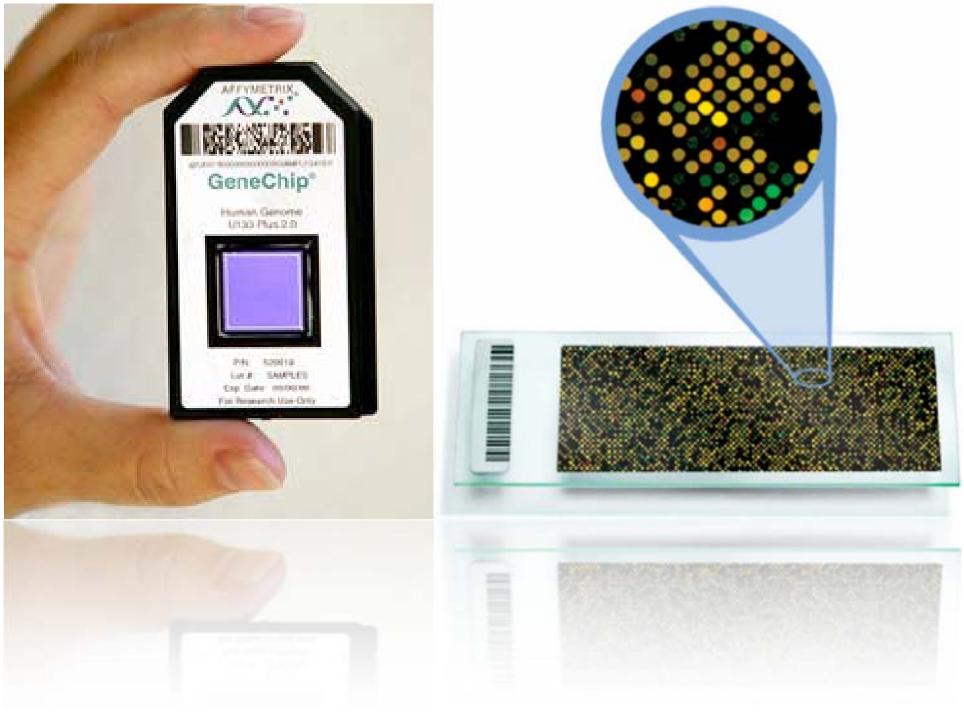
Figure 1 | **Schematized experimental process using a microarray.** Although the specific protocols differ, the microarray approach first involves isolating RNA or messenger RNA from appropriate biological samples, making the RNA (or a copy of it) fluorescent, hybridizing it to the microarray, washing off the excess and scanning the microarray under laser light.

source: Butte, A Nature Reviews 2002

Different Platforms

in situ oligonucleotide
single sample, absolute levels

spotted DNA/cDNA
two samples, relative levels



Microarray Experiment

- Design
- Collect
- Pre-Process
- Analyze
- Interpret
- Submit
- Publish

Example = Normalization

Examples = Distance measures, data classification, clustering, + more

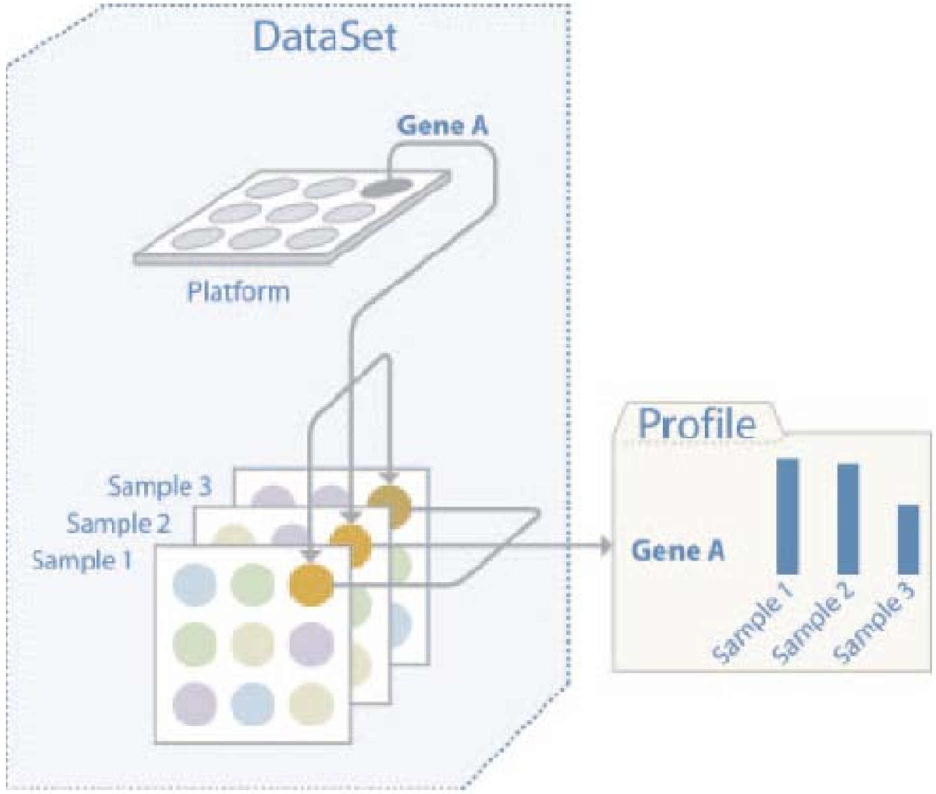
Rate Limiting Step = What do these results actually mean?

Public Microarray Data

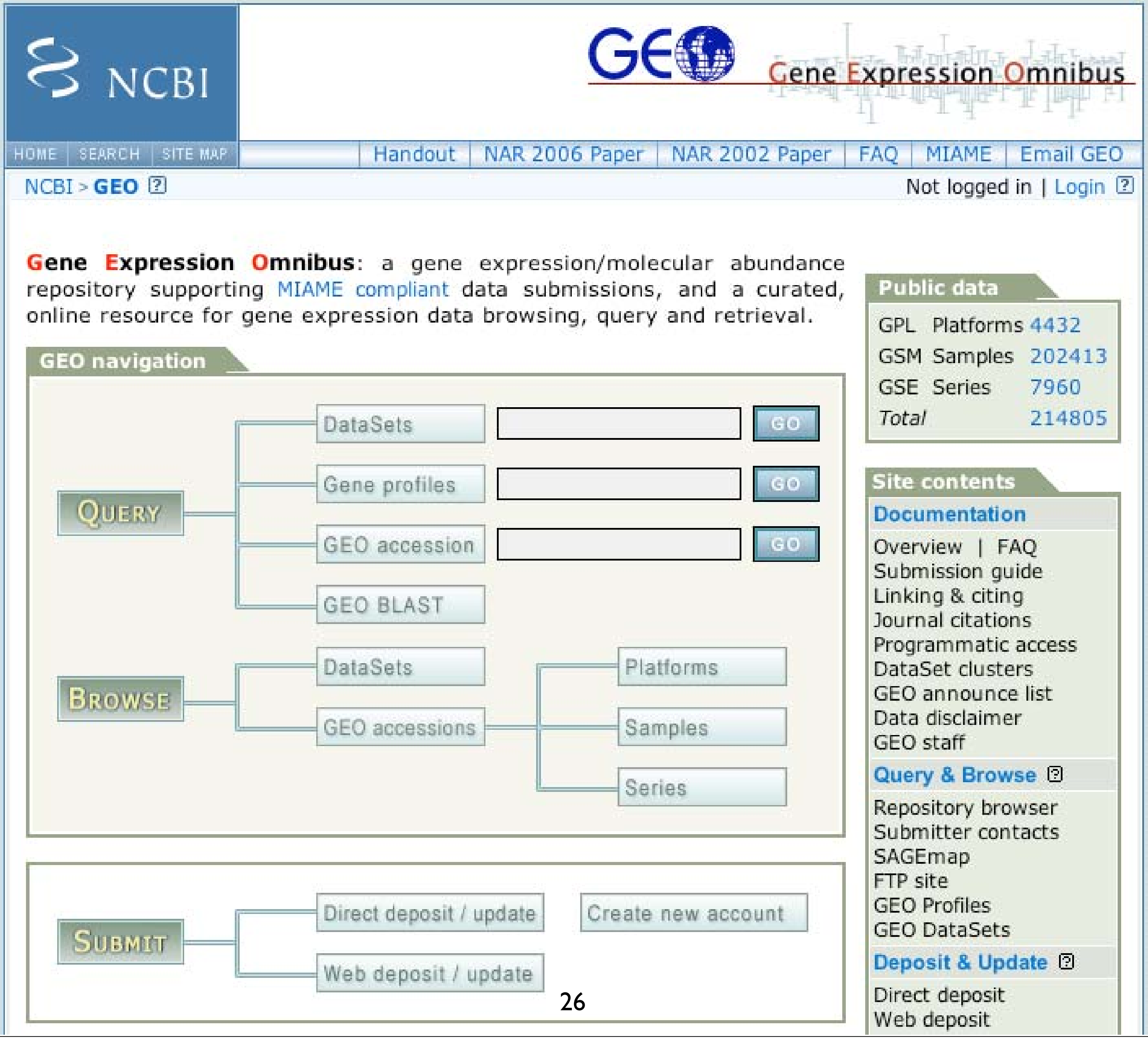
- The Gene Expression Omnibus (GEO)
  - ⦿ repository/archive gene expression data
- data submitted by the research community in fulfillment of journal requirements
- this public data represents an untapped resource; potential discovery from existing data sets is at your fingertips

GEO Database

<u>Organized by:</u>
Platform
Sample
Series/DataSet
Profile



http://www.ncbi.nlm.nih.gov/geo/



Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Public data

GPL	Platforms	4432
GSM	Samples	202413
GSE	Series	7960
Total		214805

Site contents

Documentation

- Overview | FAQ
- Submission guide
- Linking & citing
- Journal citations
- Programmatic access
- DataSet clusters
- GEO announce list
- Data disclaimer
- GEO staff

Query & Browse

- Repository browser
- Submitter contacts
- SAGEmap
- FTP site
- GEO Profiles
- GEO DataSets

Deposit & Update

- Direct deposit
- Web deposit

Searching GEO

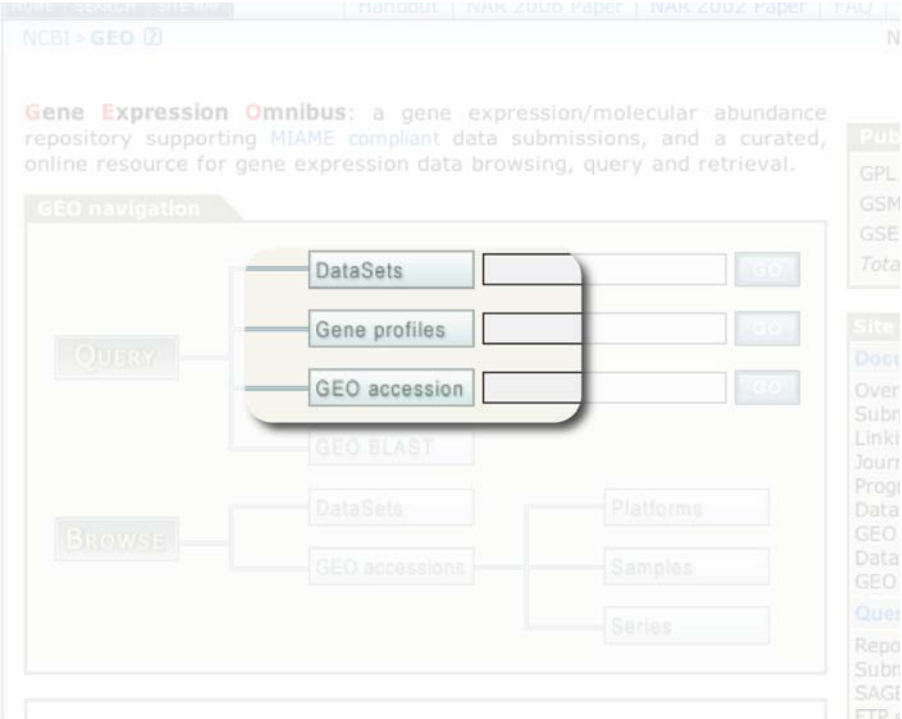
- Are you interested in a particular type of expt?

✓ GEO DataSets

- Are you looking for your favorite gene?

✓ GEO Profiles

Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

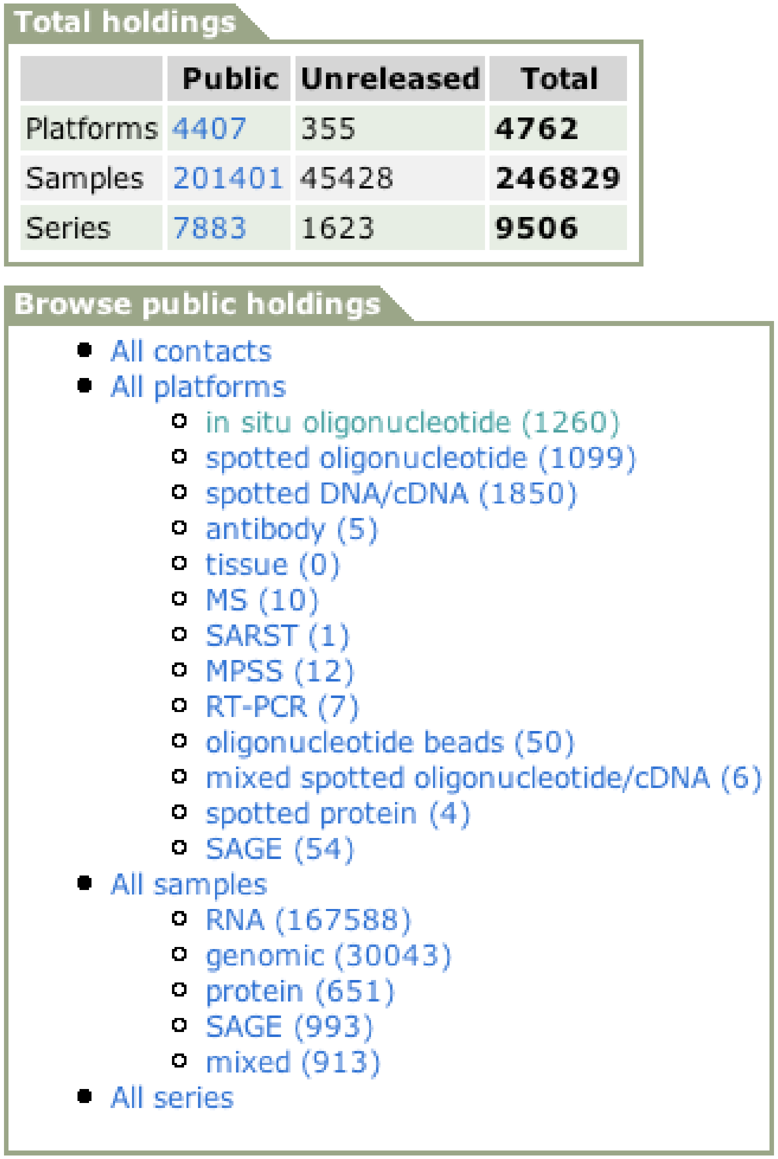


Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Data in GEO

>120,000 samples
>3.2 billion measurement
200+ organisms
from >2000 labs

freely available online
ftp downloads

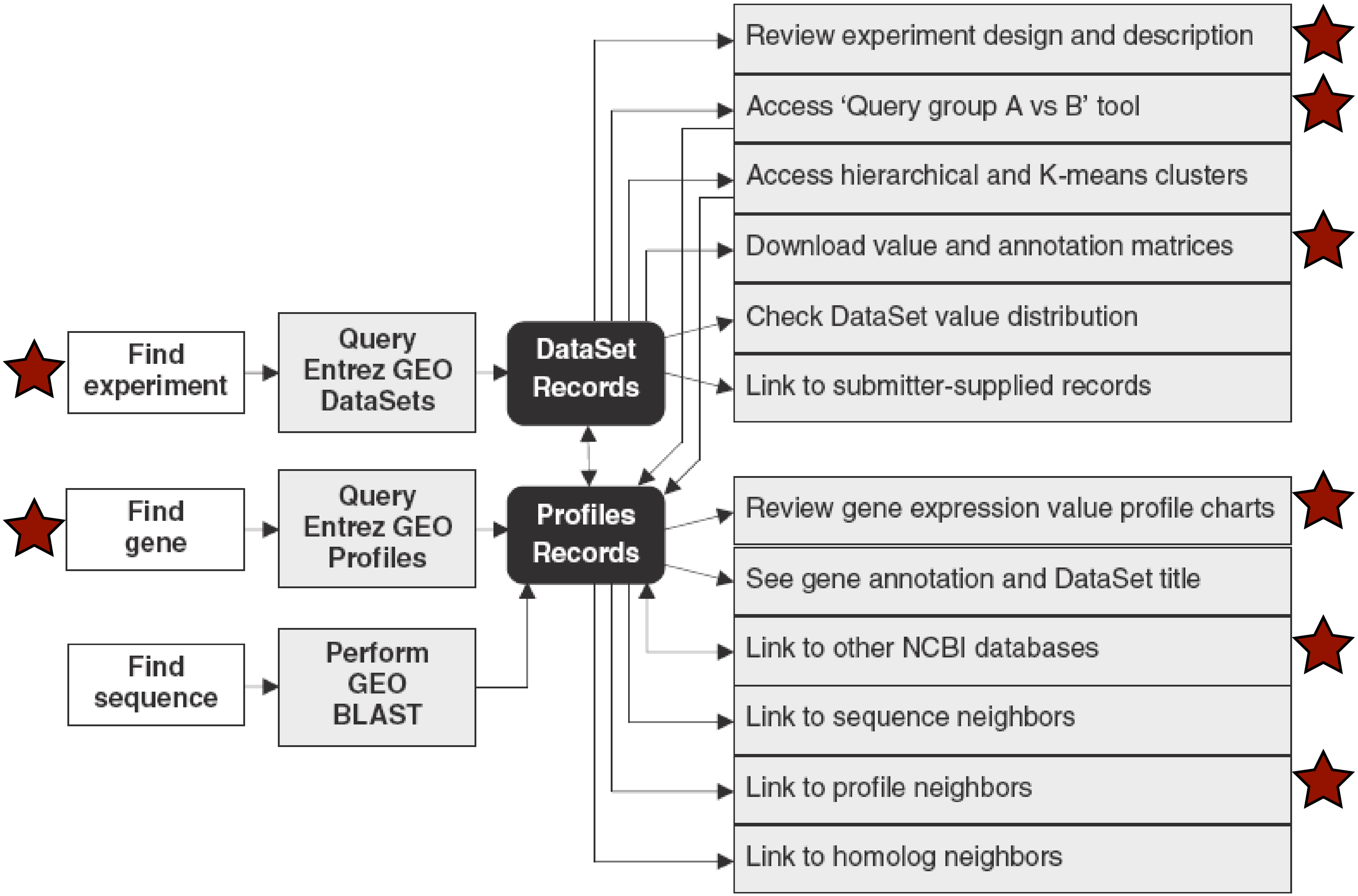


Total holdings

	Public	Unreleased	Total
Platforms	4407	355	**4762**
Samples	201401	45428	**246829**
Series	7883	1623	**9506**

Browse public holdings

- All contacts
- All platforms
    - in situ oligonucleotide (1260)
    - spotted oligonucleotide (1099)
    - spotted DNA/cDNA (1850)
    - antibody (5)
    - tissue (0)
    - MS (10)
    - SARST (1)
    - MPSS (12)
    - RT-PCR (7)
    - oligonucleotide beads (50)
    - mixed spotted oligonucleotide/cDNA (6)
    - spotted protein (4)
    - SAGE (54)
- All samples
    - RNA (167588)
    - genomic (30043)
    - protein (651)
    - SAGE (993)
    - mixed (913)
- All series



★ Find experiment → Query Entrez GEO DataSets → DataSet Records

★ Find gene → Query Entrez GEO Profiles → Profiles Records

Find sequence → Perform GEO BLAST → Profiles Records

DataSet Records:

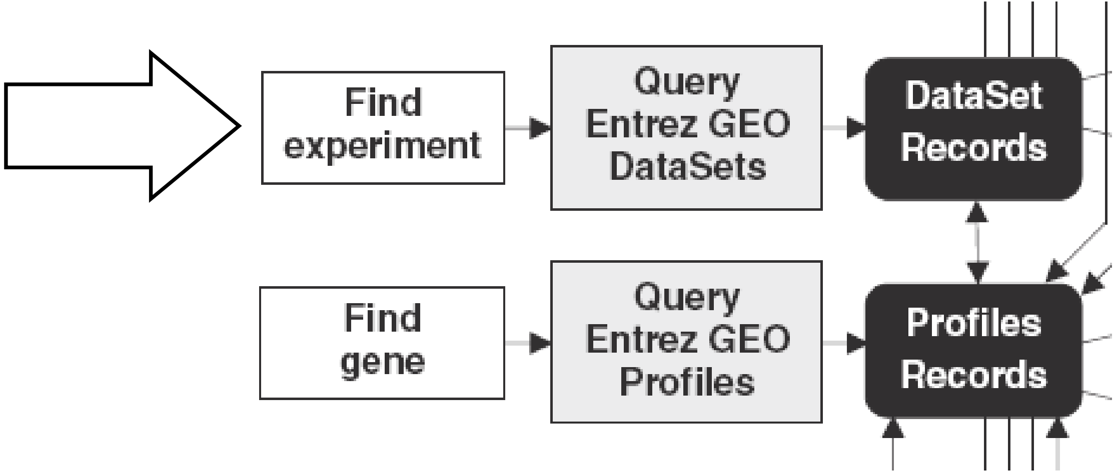
- Review experiment design and description ★
- Access 'Query group A vs B' tool ★
- Access hierarchical and K-means clusters
- Download value and annotation matrices ★
- Check DataSet value distribution
- Link to submitter-supplied records

Profiles Records:

- Review gene expression value profile charts ★
- See gene annotation and DataSet title
- Link to other NCBI databases ★
- Link to sequence neighbors
- Link to profile neighbors ★
- Link to homolog neighbors

An Example

- Find microarray experiments that look at the expression of genes in cancer



Find experiment

Query Entrez GEO DataSets

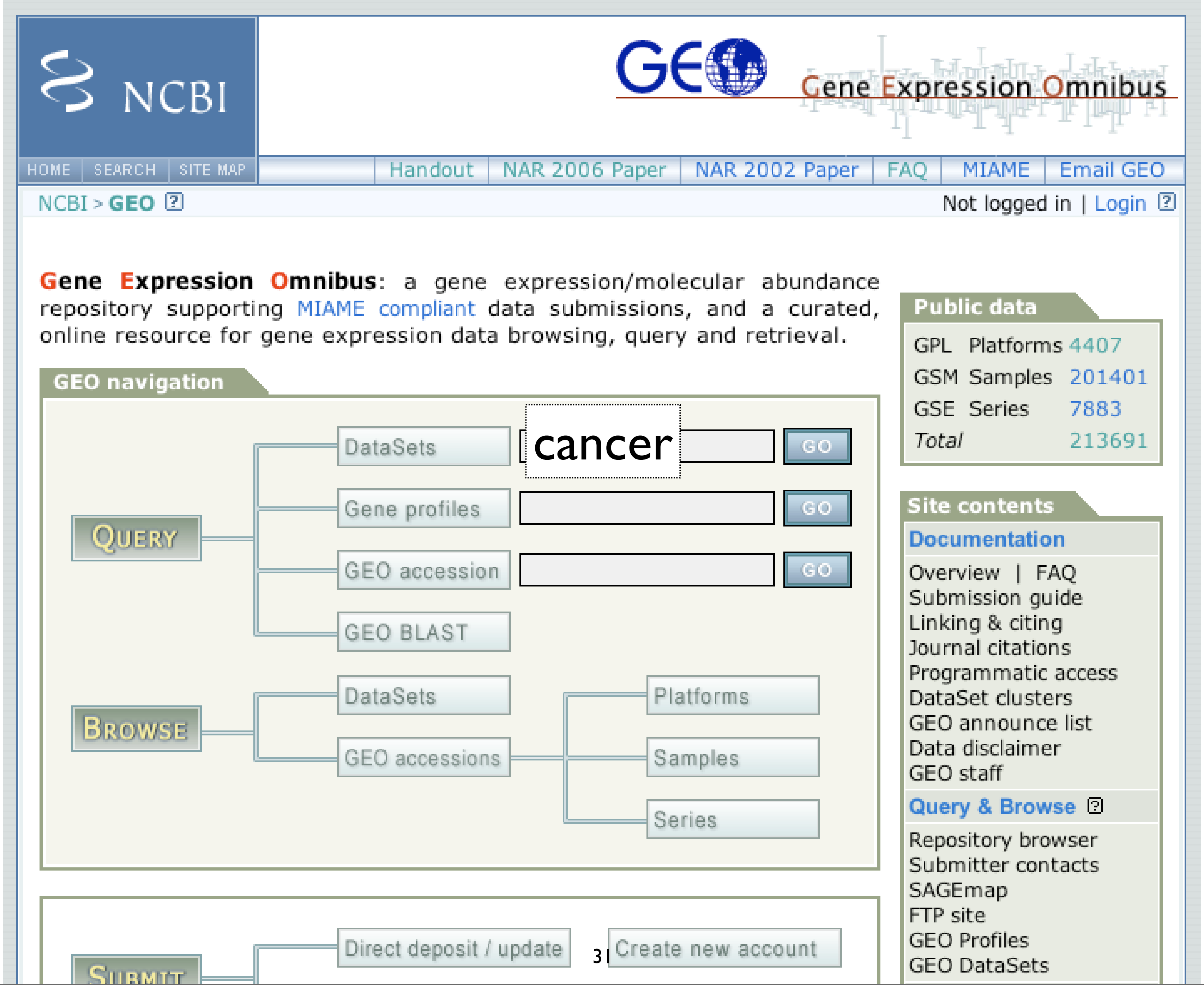
DataSet Records

Find gene

Query Entrez GEO Profiles

Profiles Records

You can use these GEO data mining tools for quick and easy identification of relevant & noteworthy data sets. For serious analyses, you should download the data and use a microarray data analysis software suite.



cancer

Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Public data

GPL	Platforms	4407
GSM	Samples	201401
GSE	Series	7883
Total		213691

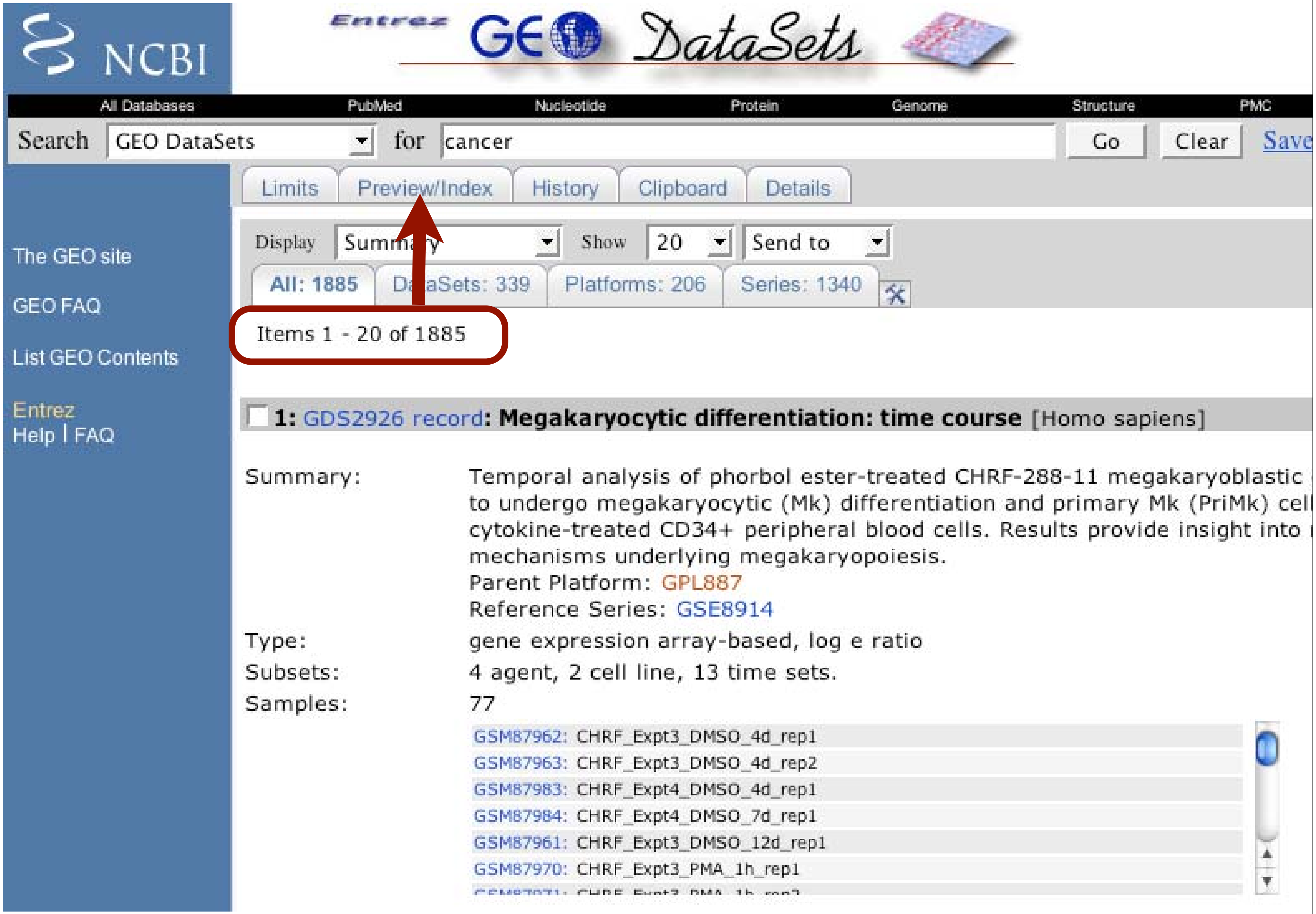
Site contents

Documentation

Overview | FAQ
Submission guide
Linking & citing
Journal citations
Programmatic access
DataSet clusters
GEO announce list
Data disclaimer
GEO staff

Query & Browse

Repository browser
Submitter contacts
SAGEmap
FTP site
GEO Profiles
GEO DataSets



Search: GEO DataSets for: cancer

Display: Summary Show: 20 Send to

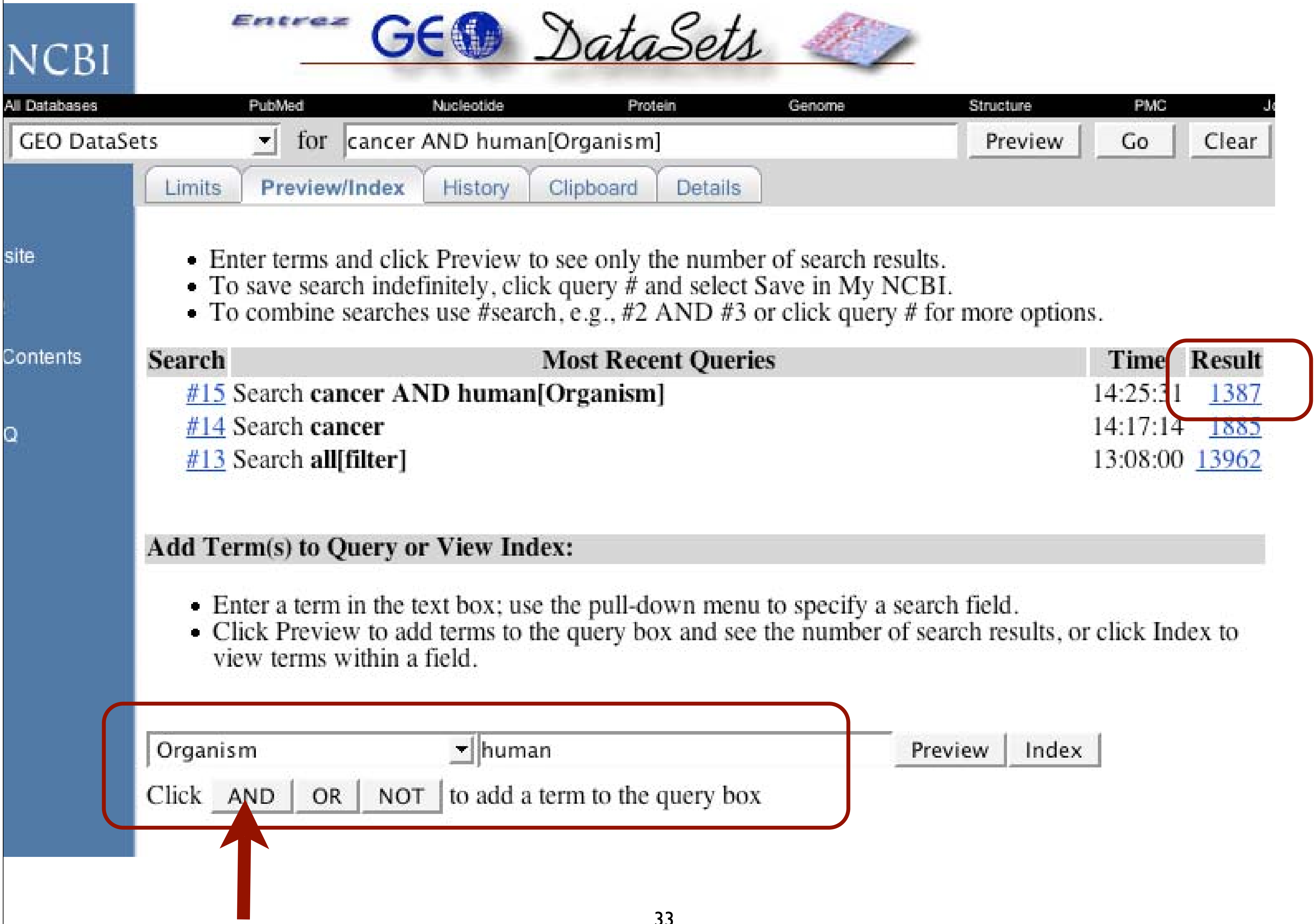
All: 1885 DataSets: 339 Platforms: 206 Series: 1340

Items 1 - 20 of 1885

☐ **1: GDS2926 record: Megakaryocytic differentiation: time course** [Homo sapiens]

Summary:	Temporal analysis of phorbol ester-treated CHRF-288-11 megakaryoblastic to undergo megakaryocytic (Mk) differentiation and primary Mk (PriMk) cel cytokine-treated CD34+ peripheral blood cells. Results provide insight into mechanisms underlying megakaryopoiesis. Parent Platform: GPL887 Reference Series: GSE8914
Type:	gene expression array-based, log e ratio
Subsets:	4 agent, 2 cell line, 13 time sets.
Samples:	77

GSM87962: CHRF_Expt3_DMSO_4d_rep1
GSM87963: CHRF_Expt3_DMSO_4d_rep2
GSM87983: CHRF_Expt4_DMSO_4d_rep1
GSM87984: CHRF_Expt4_DMSO_7d_rep1
GSM87961: CHRF_Expt3_DMSO_12d_rep1
GSM87970: CHRF_Expt3_PMA_1h_rep1



GEO DataSets for cancer AND human[Organism]

- Enter terms and click Preview to see only the number of search results.
- To save search indefinitely, click query # and select Save in My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

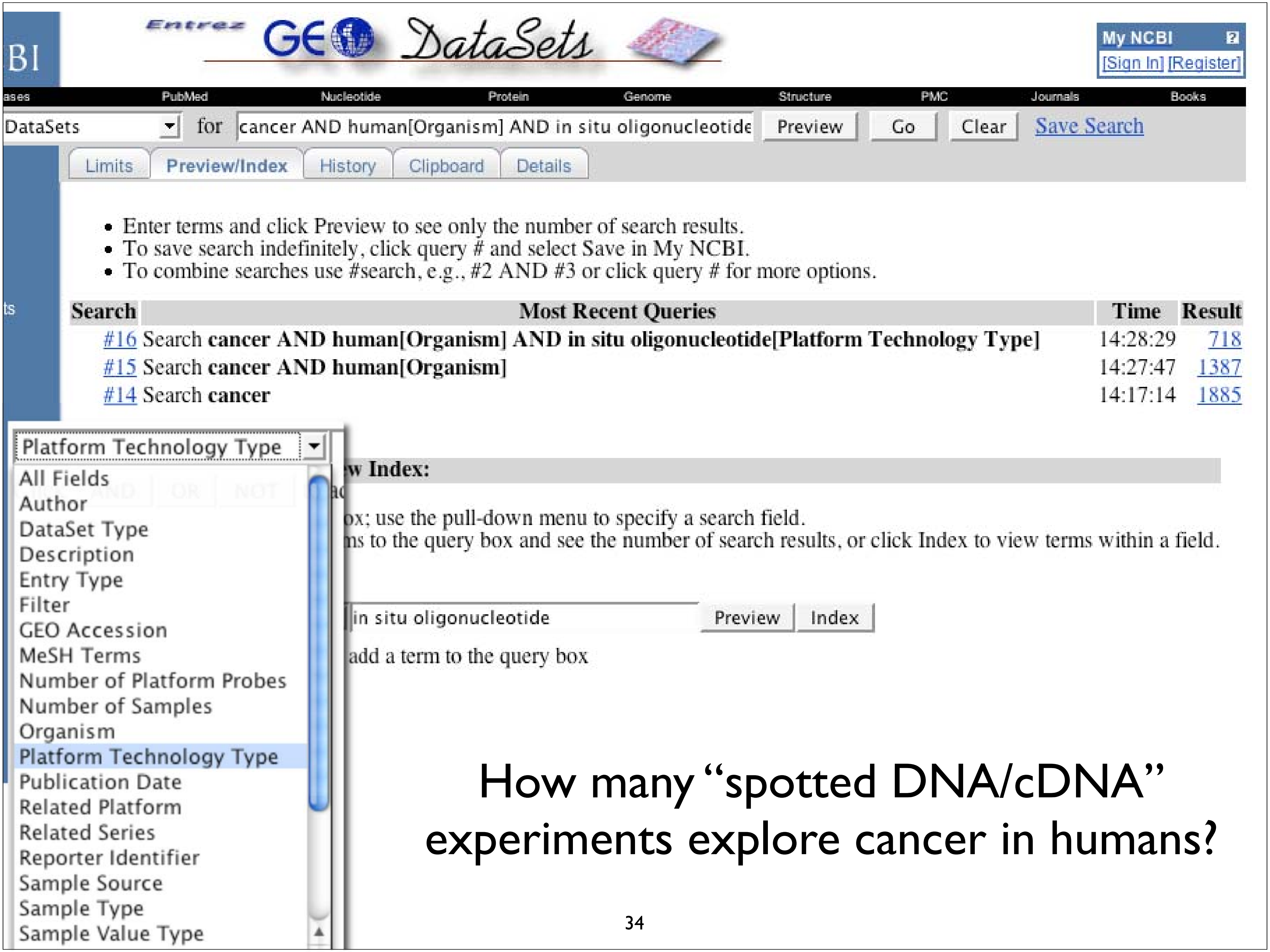
Search	Most Recent Queries	Time	Result
#15	Search **cancer AND human[Organism]**	14:25:31	1387
#14	Search **cancer**	14:17:14	1885
#13	Search **all[filter]**	13:08:00	13962

Add Term(s) to Query or View Index:

- Enter a term in the text box; use the pull-down menu to specify a search field.
- Click Preview to add terms to the query box and see the number of search results, or click Index to view terms within a field.

Organism: human

Click AND OR NOT to add a term to the query box



cancer AND human[Organism] AND in situ oligonucleotide

- Enter terms and click Preview to see only the number of search results.
- To save search indefinitely, click query # and select Save in My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

Search	Most Recent Queries	Time	Result
#16	Search **cancer AND human[Organism] AND in situ oligonucleotide[Platform Technology Type]**	14:28:29	718
#15	Search **cancer AND human[Organism]**	14:27:47	1387
#14	Search **cancer**	14:17:14	1885

w Index:

ox; use the pull-down menu to specify a search field.
ns to the query box and see the number of search results, or click Index to view terms within a field.

in situ oligonucleotide

add a term to the query box

How many "spotted DNA/cDNA" experiments explore cancer in humans?

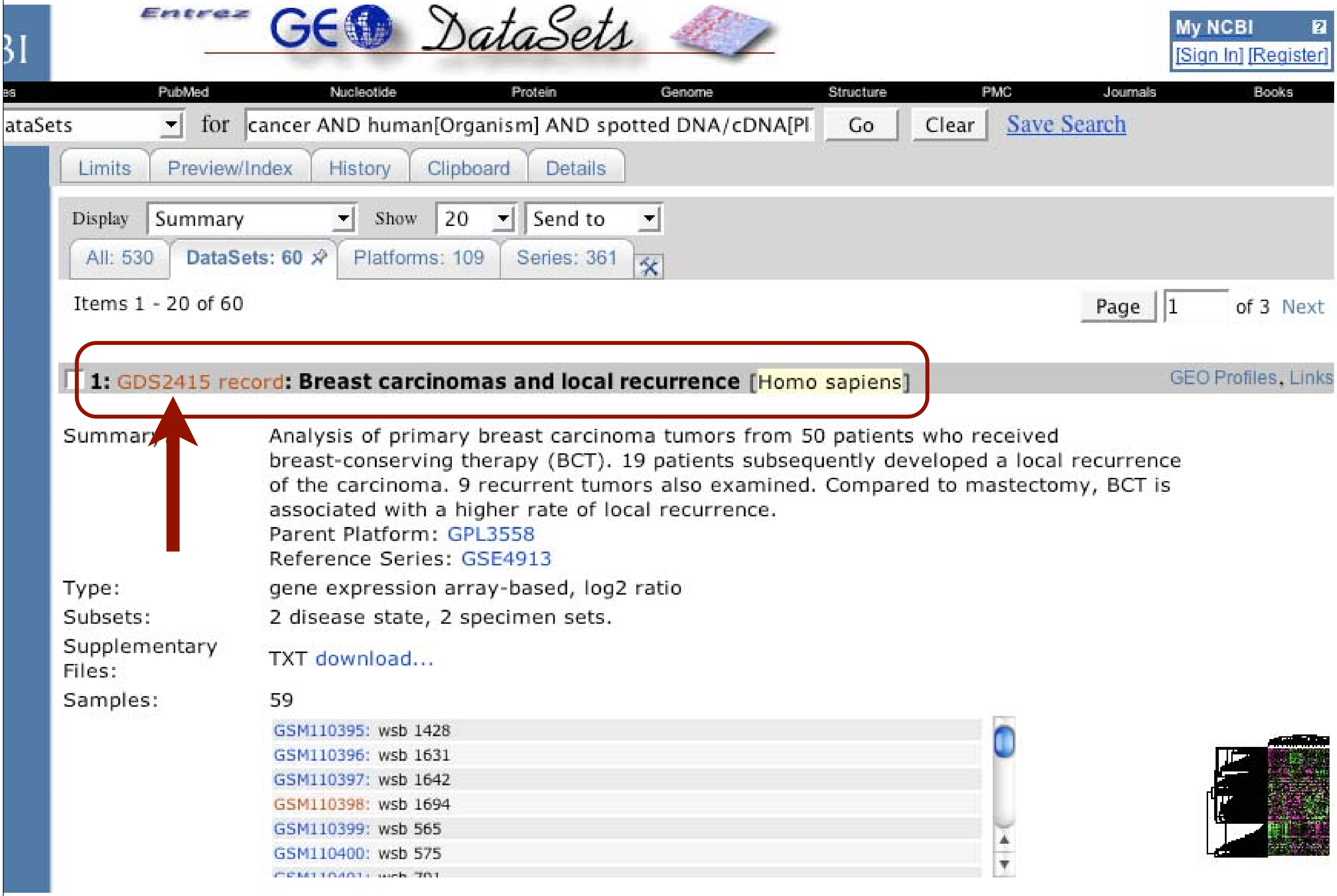
cancer AND human[Organism] AND spotted DNA/cDNA[Pl

Items 1 - 20 of 530

☐ **1:** GDS2415 record: **Breast carcinomas and local recurrence** [Homo sapiens]

Summary:	Analysis of primary breast carcinoma tumors from 50 patients who received breast-conserving therapy (BCT). 19 patients subsequently developed a local recurrence of the carcinoma. 9 recurrent tumors also examined. Compared to mastectomy, BCT is associated with a higher rate of local recurrence. Parent Platform: GPL3558 Reference Series: GSE4913
Type:	gene expression array-based, log2 ratio
Subsets:	2 disease state, 2 specimen sets.
Supplementary Files:	TXT download...
Samples:	59

GSM110395: wsb 1428
GSM110396: wsb 1631
GSM110397: wsb 1642
GSM110398: wsb 1694
GSM110399: wsb 565
GSM110400: wsb 575



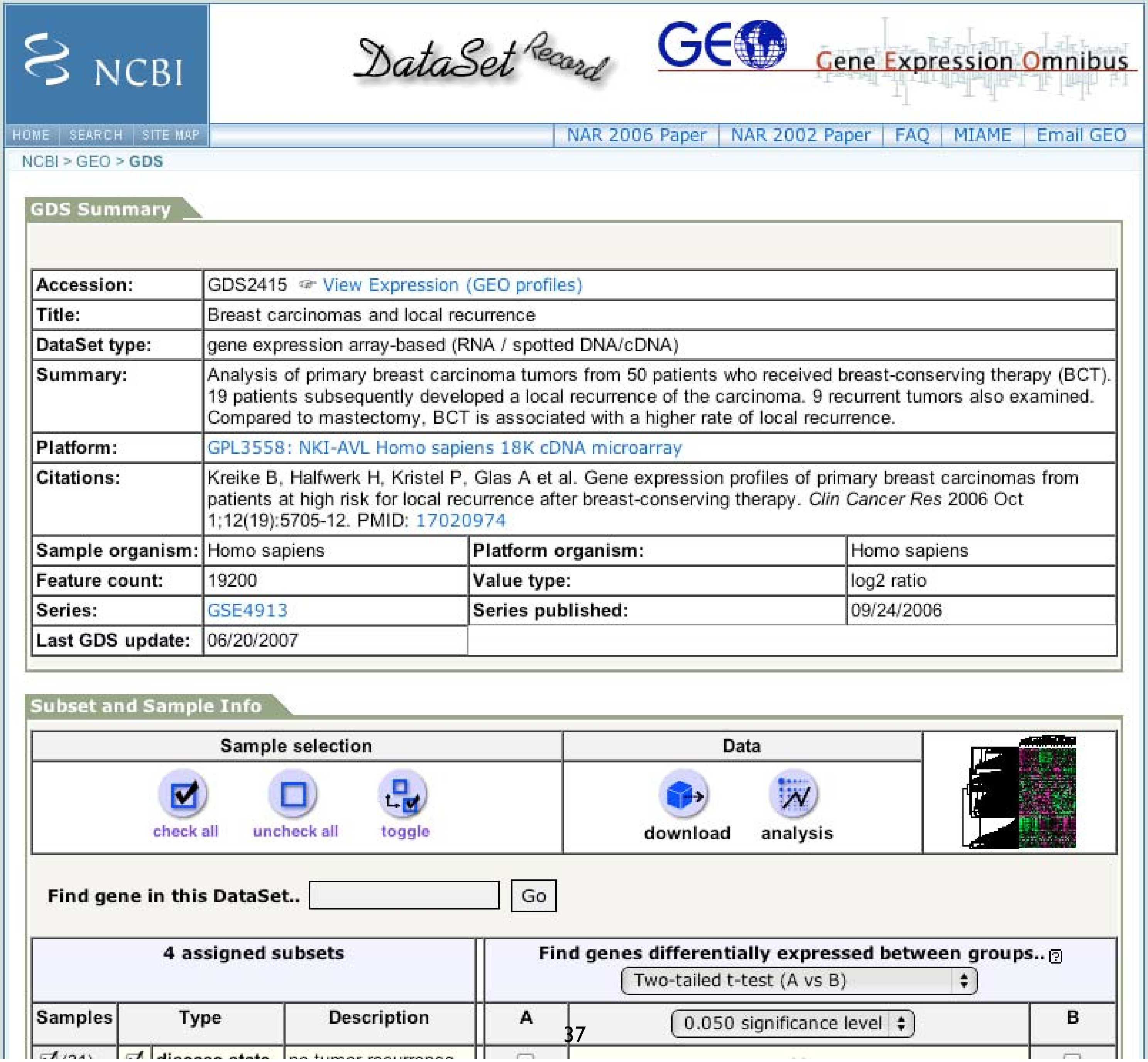
cancer AND human[Organism] AND spotted DNA/cDNA[Pl

Items 1 - 20 of 60

☐ **1: GDS2415 record: Breast carcinomas and local recurrence** [Homo sapiens]

Summary:	Analysis of primary breast carcinoma tumors from 50 patients who received breast-conserving therapy (BCT). 19 patients subsequently developed a local recurrence of the carcinoma. 9 recurrent tumors also examined. Compared to mastectomy, BCT is associated with a higher rate of local recurrence. Parent Platform: GPL3558 Reference Series: GSE4913
Type:	gene expression array-based, log2 ratio
Subsets:	2 disease state, 2 specimen sets.
Supplementary Files:	TXT download...
Samples:	59

GSM110395: wsb 1428
GSM110396: wsb 1631
GSM110397: wsb 1642
GSM110398: wsb 1694
GSM110399: wsb 565
GSM110400: wsb 575



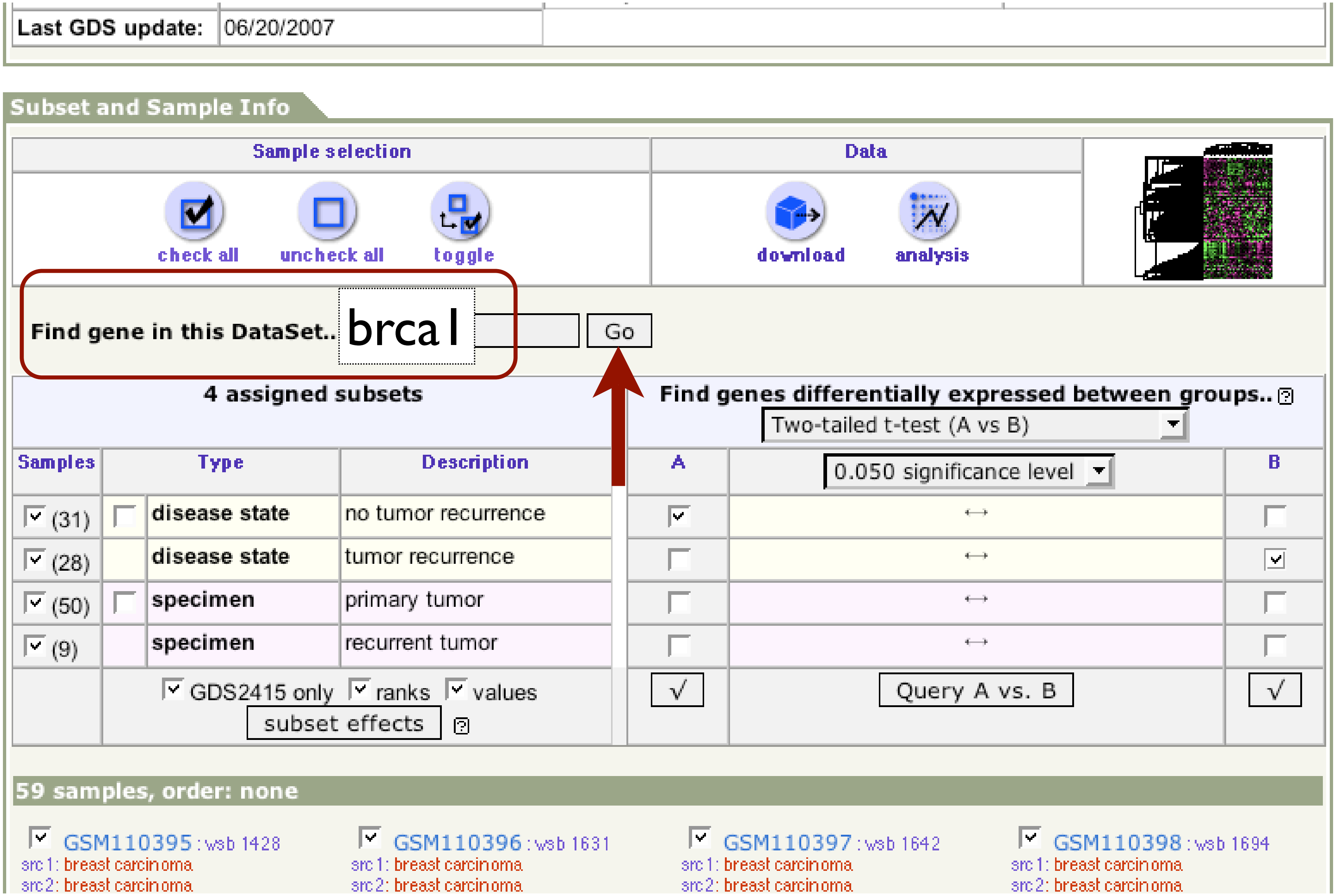
GDS Summary

Accession:	GDS2415 ☞ View Expression (GEO profiles)		
Title:	Breast carcinomas and local recurrence		
DataSet type:	gene expression array-based (RNA / spotted DNA/cDNA)		
Summary:	Analysis of primary breast carcinoma tumors from 50 patients who received breast-conserving therapy (BCT). 19 patients subsequently developed a local recurrence of the carcinoma. 9 recurrent tumors also examined. Compared to mastectomy, BCT is associated with a higher rate of local recurrence.		
Platform:	GPL3558: NKI-AVL Homo sapiens 18K cDNA microarray		
Citations:	Kreike B, Halfwerk H, Kristel P, Glas A et al. Gene expression profiles of primary breast carcinomas from patients at high risk for local recurrence after breast-conserving therapy. *Clin Cancer Res* 2006 Oct 1;12(19):5705-12. PMID: 17020974		
Sample organism:	Homo sapiens	**Platform organism:**	Homo sapiens
Feature count:	19200	**Value type:**	log2 ratio
Series:	GSE4913	**Series published:**	09/24/2006
Last GDS update:	06/20/2007		

Subset and Sample Info

Find gene in this DataSet.. Go

4 assigned subsets			Find genes differentially expressed between groups..		
			Two-tailed t-test (A vs B)		
Samples	**Type**	**Description**	**A**	0.050 significance level	**B**



Last GDS update: 06/20/2007

Subset and Sample Info

Find gene in this DataSet.. brca1 Go

4 assigned subsets

Find genes differentially expressed between groups..

Two-tailed t-test (A vs B)

Samples		Type	Description	A	0.050 significance level	B
☑ (31)	☐	**disease state**	no tumor recurrence	☑	↔	☐
☑ (28)		**disease state**	tumor recurrence	☐	↔	☑
☑ (50)	☐	**specimen**	primary tumor	☐	↔	☐
☑ (9)		**specimen**	recurrent tumor	☐	↔	☐

☑ GDS2415 only ☑ ranks ☑ values

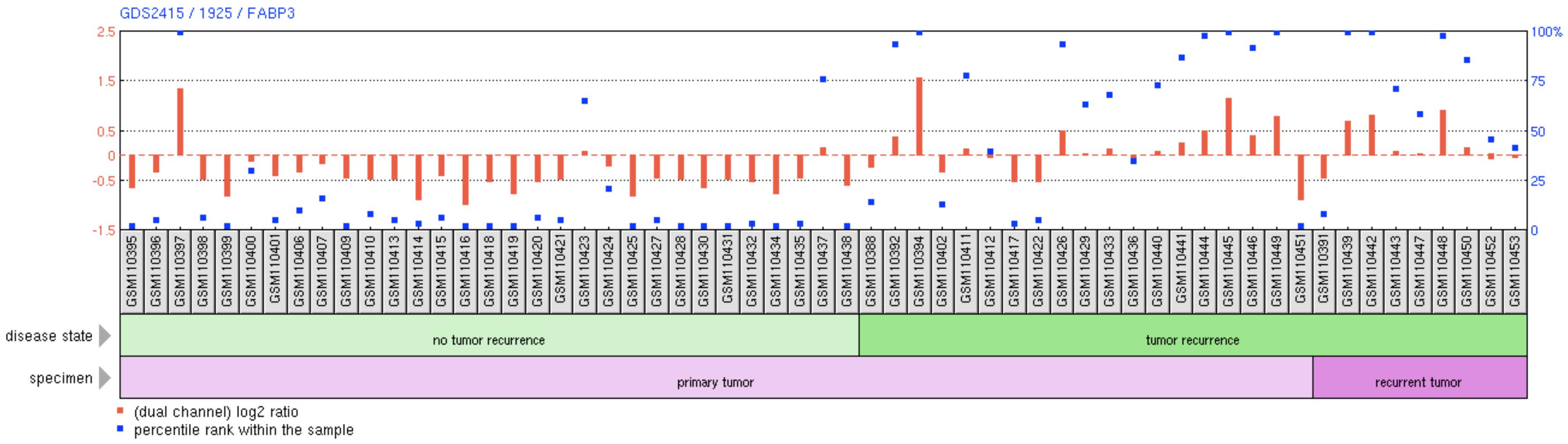
59 samples, order: none

☑ GSM110395 : wsb 1428
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110396 : wsb 1631
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110397 : wsb 1642
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110398 : wsb 1694
src 1: breast carcinoma
src 2: breast carcinoma

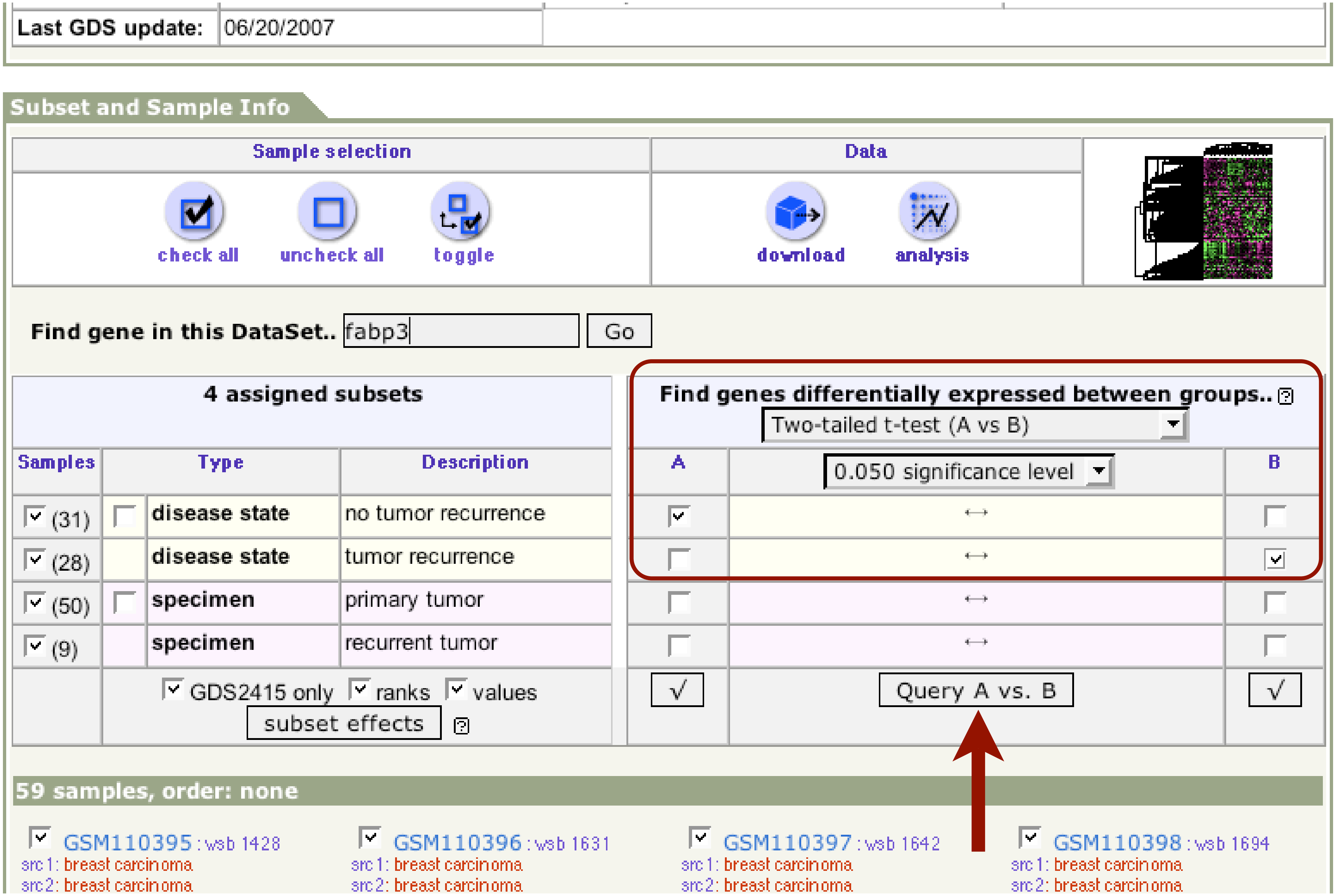


- thumbnail image represents the abundance profile for an individual gene across each Sample in a DataSet

- bars at the bottom of the chart represent experimental subsets within the DataSet.

- Red bar: measured level of abundance

- Blue square: indication of where the expression of that gene falls with respect to all other genes on that array



Last GDS update:	06/20/2007

Subset and Sample Info

Sample selection

Data

Find gene in this DataSet.. fabp3

4 assigned subsets

Find genes differentially expressed between groups..

Two-tailed t-test (A vs B)

Samples		Type	Description	A	0.050 significance level	B
☑ (31)	☐	**disease state**	no tumor recurrence	☑	↔	☐
☑ (28)		**disease state**	tumor recurrence	☐	↔	☑
☑ (50)	☐	**specimen**	primary tumor	☐	↔	☐
☑ (9)		**specimen**	recurrent tumor	☐	↔	☐
	☑ GDS2415 only ☑ ranks ☑ values subset effects			√	Query A vs. B	√

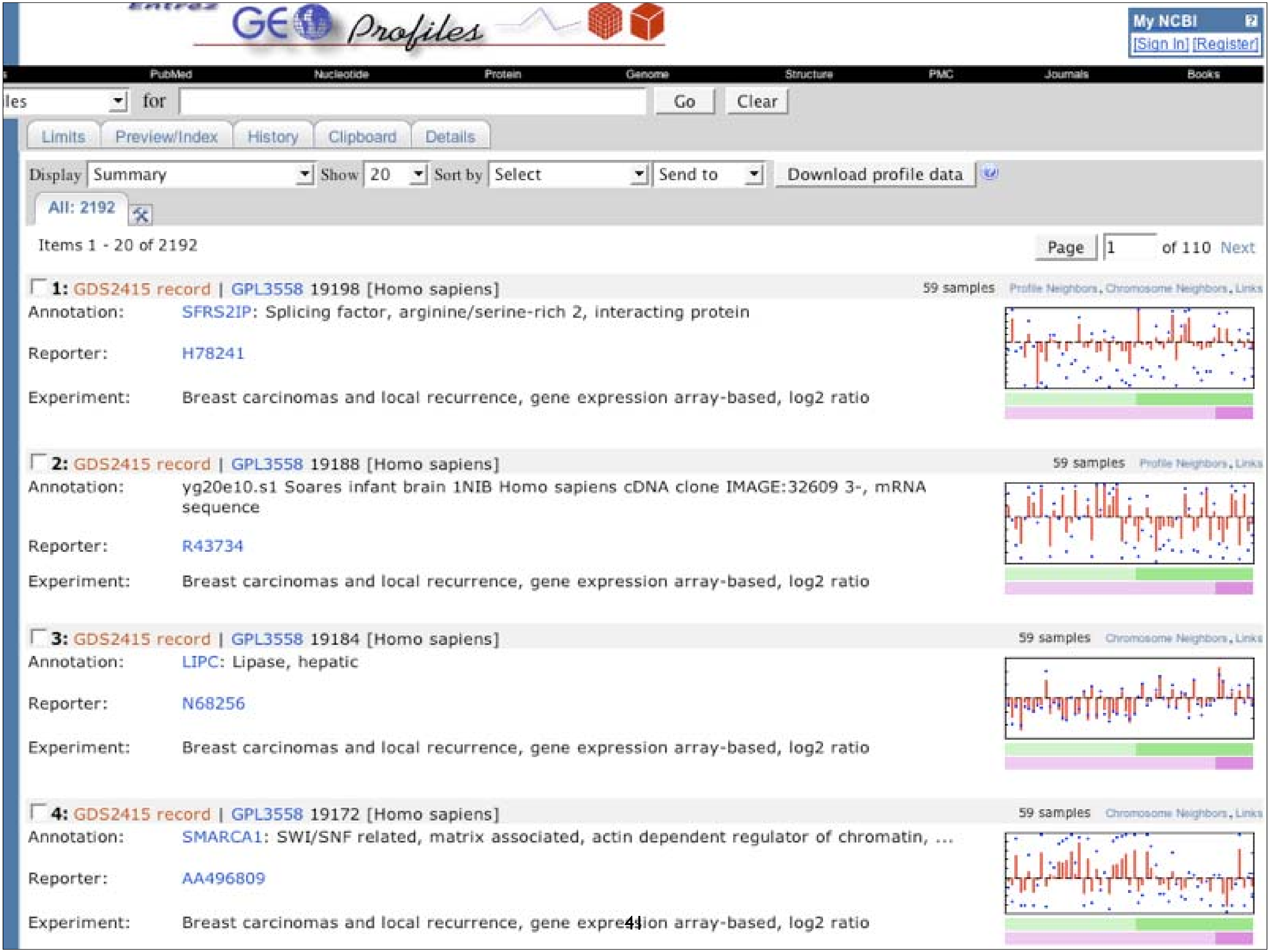
59 samples, order: none

☑ GSM110395 : wsb 1428
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110396 : wsb 1631
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110397 : wsb 1642
src 1: breast carcinoma
src 2: breast carcinoma

☑ GSM110398 : wsb 1694
src 1: breast carcinoma
src 2: breast carcinoma



Items 1 - 20 of 2192

☐ **1:** GDS2415 record | GPL3558 19198 [Homo sapiens] 59 samples Profile Neighbors, Chromosome Neighbors, Links

Annotation:	SFRS2IP: Splicing factor, arginine/serine-rich 2, interacting protein
Reporter:	H78241
Experiment:	Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **2:** GDS2415 record | GPL3558 19188 [Homo sapiens] 59 samples Profile Neighbors, Links

Annotation:	yg20e10.s1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:32609 3-, mRNA sequence
Reporter:	R43734
Experiment:	Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **3:** GDS2415 record | GPL3558 19184 [Homo sapiens] 59 samples Chromosome Neighbors, Links

Annotation:	LIPC: Lipase, hepatic
Reporter:	N68256
Experiment:	Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **4:** GDS2415 record | GPL3558 19172 [Homo sapiens] 59 samples Chromosome Neighbors, Links

Annotation:	SMARCA1: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, ...
Reporter:	AA496809
Experiment:	Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

A vs B Query Tool

Take home message: GEO data analysis tools are great for quick identification of interesting leads; you download the data to carry out more robust statistical analyses

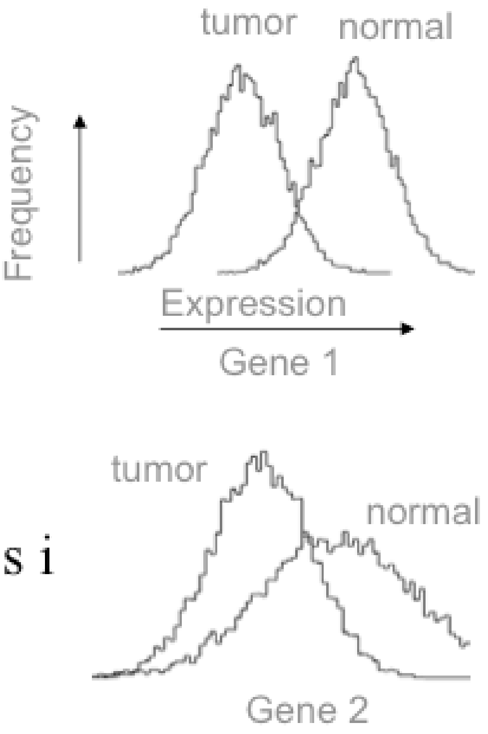
- **Purpose:** To help identify gene profiles that display marked differences in expression level between two subsets of experimental factors (e.g. tissue, strain, time, dose, etc).
- **Caveats:** The "mean group A vs B" is perhaps the most rudimentary means of filtering data; t-test is well established but comes with a set of basic assumptions.

A Simple Test

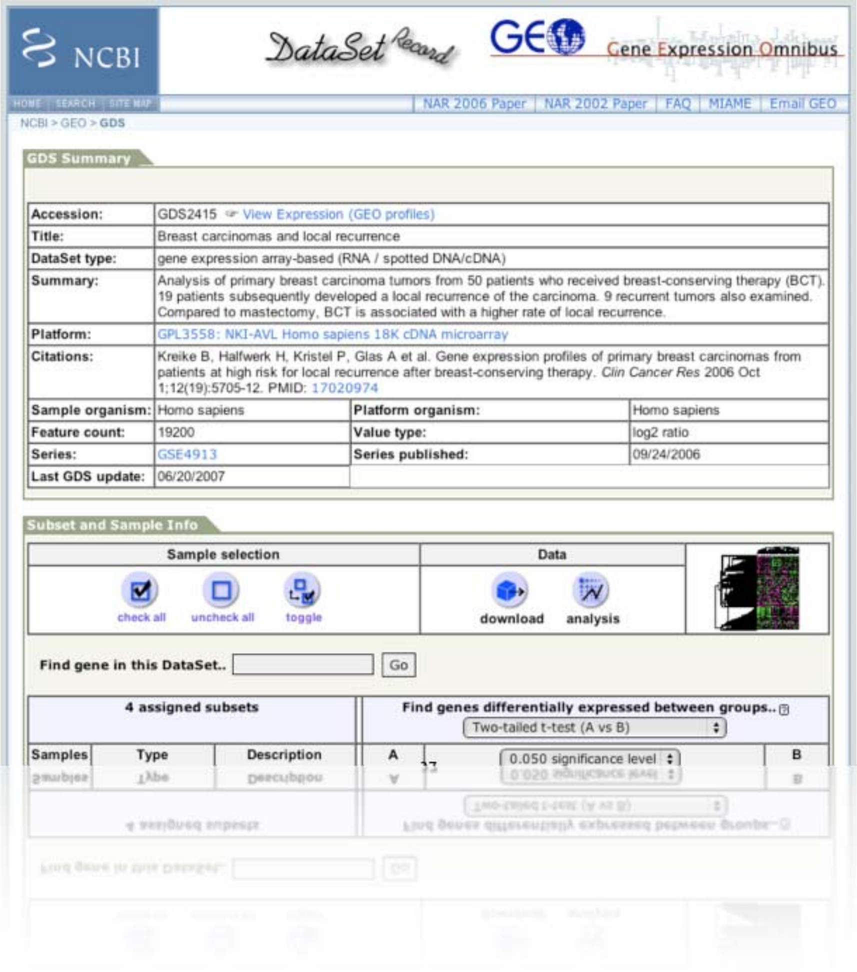
- Student's t-test
  - Assumptions: Normality, equal variance

$$t = \frac{m_{tumor} - m_{normal}}{\sqrt{\frac{s_{tumor}^{2}}{N_{tumor}} + \frac{s_{normal}^{2}}{N_{normal}}}}$$

m_i = mean expression value in class i
N_i = number of examples in class i
s_i^2 = variance



Using GEO for differential expression



Accession:	GDS2415 View Expression (GEO profiles)		
Title:	Breast carcinomas and local recurrence		
DataSet type:	gene expression array-based (RNA / spotted DNA/cDNA)		
Summary:	Analysis of primary breast carcinoma tumors from 50 patients who received breast-conserving therapy (BCT). 19 patients subsequently developed a local recurrence of the carcinoma. 9 recurrent tumors also examined. Compared to mastectomy, BCT is associated with a higher rate of local recurrence.		
Platform:	GPL3558: NKI-AVL Homo sapiens 18K cDNA microarray		
Citations:	Kreike B, Halfwerk H, Kristel P, Glas A et al. Gene expression profiles of primary breast carcinomas from patients at high risk for local recurrence after breast-conserving therapy. *Clin Cancer Res* 2006 Oct 1;12(19):5705-12. PMID: 17020974		
Sample organism:	Homo sapiens	**Platform organism:**	Homo sapiens
Feature count:	19200	**Value type:**	log2 ratio
Series:	GSE4913	**Series published:**	09/24/2006
Last GDS update:	06/20/2007		

http://www.ncbi.nlm.nih.gov/projects/geo/gds/gds_browse.cgi?gds=2415

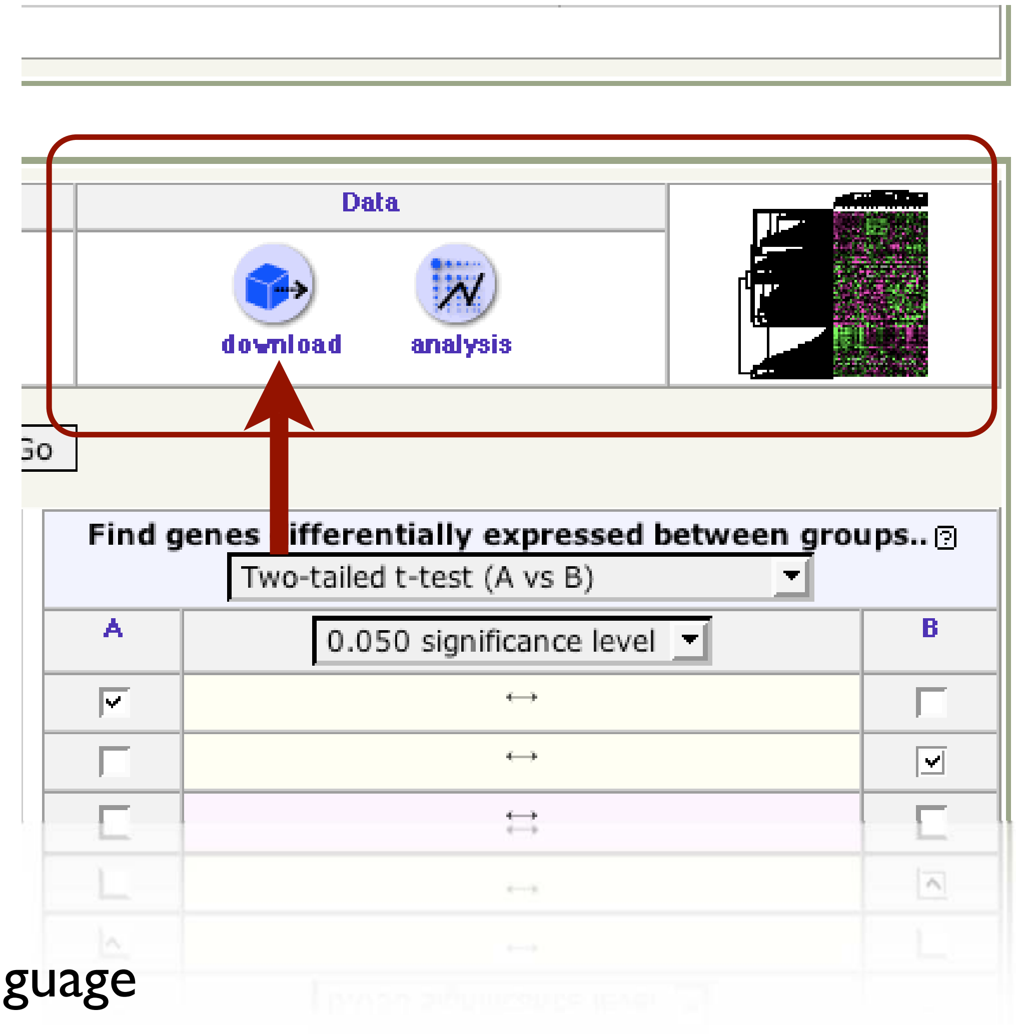
GEO limitations

- Differential expression can only be done for “Datasets”(GDS****)
- T-tests only
- Very little control over parameters
- Output is not that easy to use

Be careful with $p < 0.05$

- In the example, 2912 genes met $p<0.05$
- This is 11% of the genes on the array
- Expect 5% by chance (this is what $p < 0.05$ means)
- Probably ~1/2 of the selected 2912 are false positives
- This is the "multiple testing" problem

Download Data Types



SOFT - text based
MiNiML* - xml based

*MIAME Notation in Markup Language

all GEO data are available for bulk download:
ftp://ftp.ncbi.nih.gov/pub/geo/DATA

More Serious Tools

- Free
  - R + Bioconductor
  - TIGR MultiExperimentViewer (MeV)

  …
- Commerical
  - Genespring–ArrayAssist
  - Rosetta Resolver

  …

Imaging, Diagnosis, Prognosis

Gene Expression Profiles of Primary Breast Carcinomas from Patients at High Risk for Local Recurrence after Breast-Conserving Therapy

Bas Kreike,[1,3] Hans Halfwerk,[2,3] Petra Kristel,[2,3] Annuska Glas,[2] Hans Peterse,[2]
Harry Bartelink,[1] and Marc J. van de Vijver[2]

Abstract **Purpose:** Several risk factors for local recurrence of breast cancer after breast-conserving therapy (BCT) have been identified. The identification of additional risk factors would be very useful in guiding optimal therapy and also in improving understanding of the mechanisms underlying local recurrence. We used cDNA microarray analysis to identify gene expression profiles associated with local recurrence.

Experimental Design: Using 18K cDNA microarrays, gene expression profiles were obtained from 50 patients who underwent BCT. Of these 50 patients, 19 developed a local recurrence; the remaining 31 patients were selected as controls as they were free of local recurrence at least 11 years after treatment. For 9 of 19 patients, the local recurrence was also available for gene expression profiling. Unsupervised and supervised methods of classification were used to separate patients in groups corresponding to disease outcome and to study the overall gene expression pattern of primary tumors and their recurrences.

Results: Hierarchical clustering of patients did not show any grouping reflecting local recurrence status. Supervised analysis revealed no significant set of genes that was able to distinguish recurring tumors from nonrecurring tumors. Paired-data analysis of primary tumors and local recurrences showed a remarkable similarity in gene expression profile between primary tumors and their recurrences.

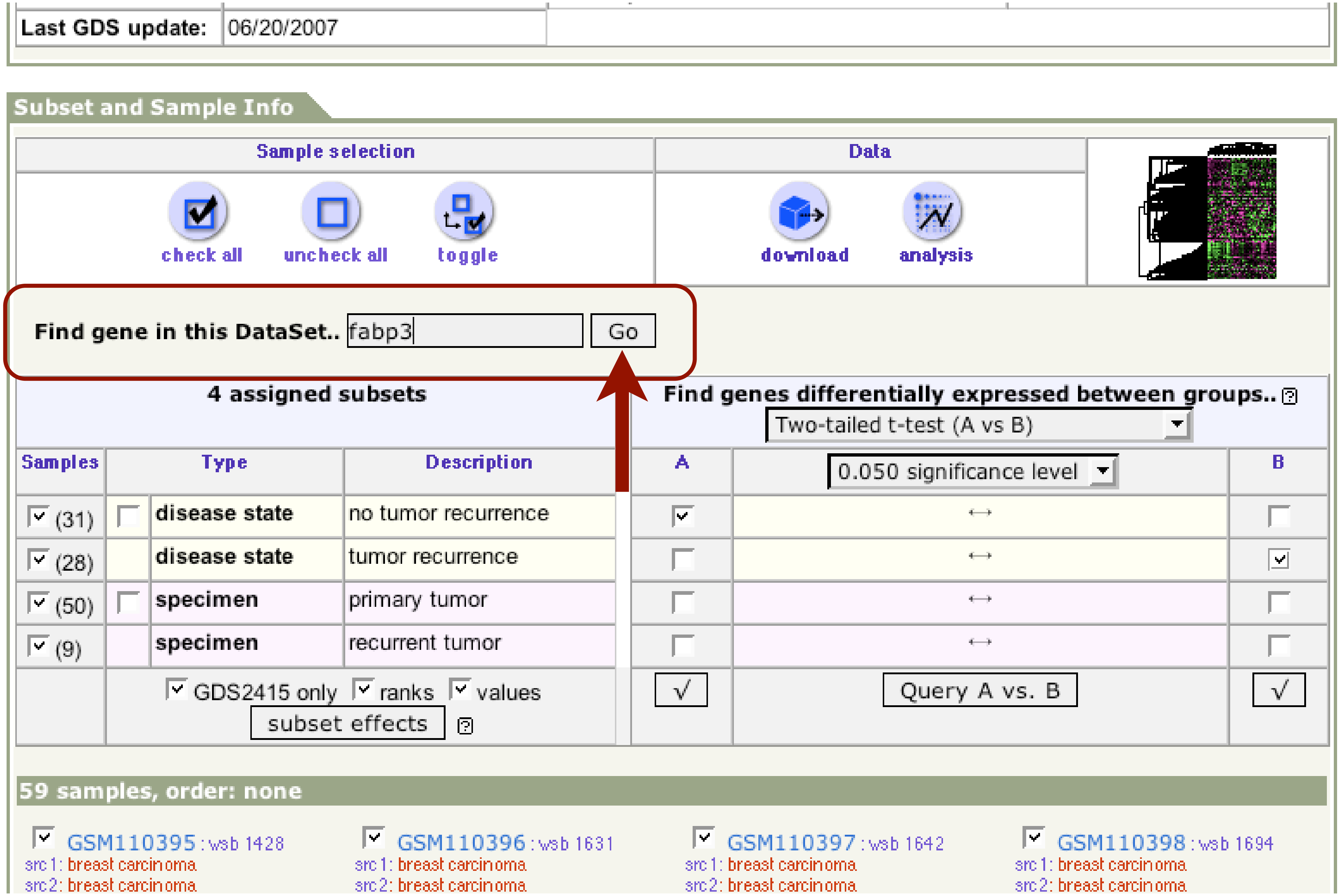
Conclusions: No significant differences in gene expression between primary breast cancer tumors in patients with or without local recurrence after BCT were identified. Furthermore, analyses of primary tumors and local recurrences show a preservation of the overall gene expression pattern in the local recurrence, even after radiotherapy.

Breast-conserving therapy (BCT) has become the therapy of choice for a large proportion of breast cancer patients. Several randomized controlled trials have shown no difference in survival rates after BCT or mastectomy for stage I and II breast cancer (1–4). Studies comparing the psychological effects of BCT with mastectomy have shown that patients treated with

recurrence compared with mastectomy. A local recurrence rate of 10% in 10 years follow-up is generally considered as clinically acceptable for $T_{1-2}N_{0-1}$ breast cancers. However, local recurrence up to 30% have been reported in young patients (7, 8).

Several risk factors for local recurrence after BCT have been

gene of interest
FABP3



Last GDS update:	06/20/2007

Subset and Sample Info

Find gene in this DataSet.. fabp3 Go

4 assigned subsets

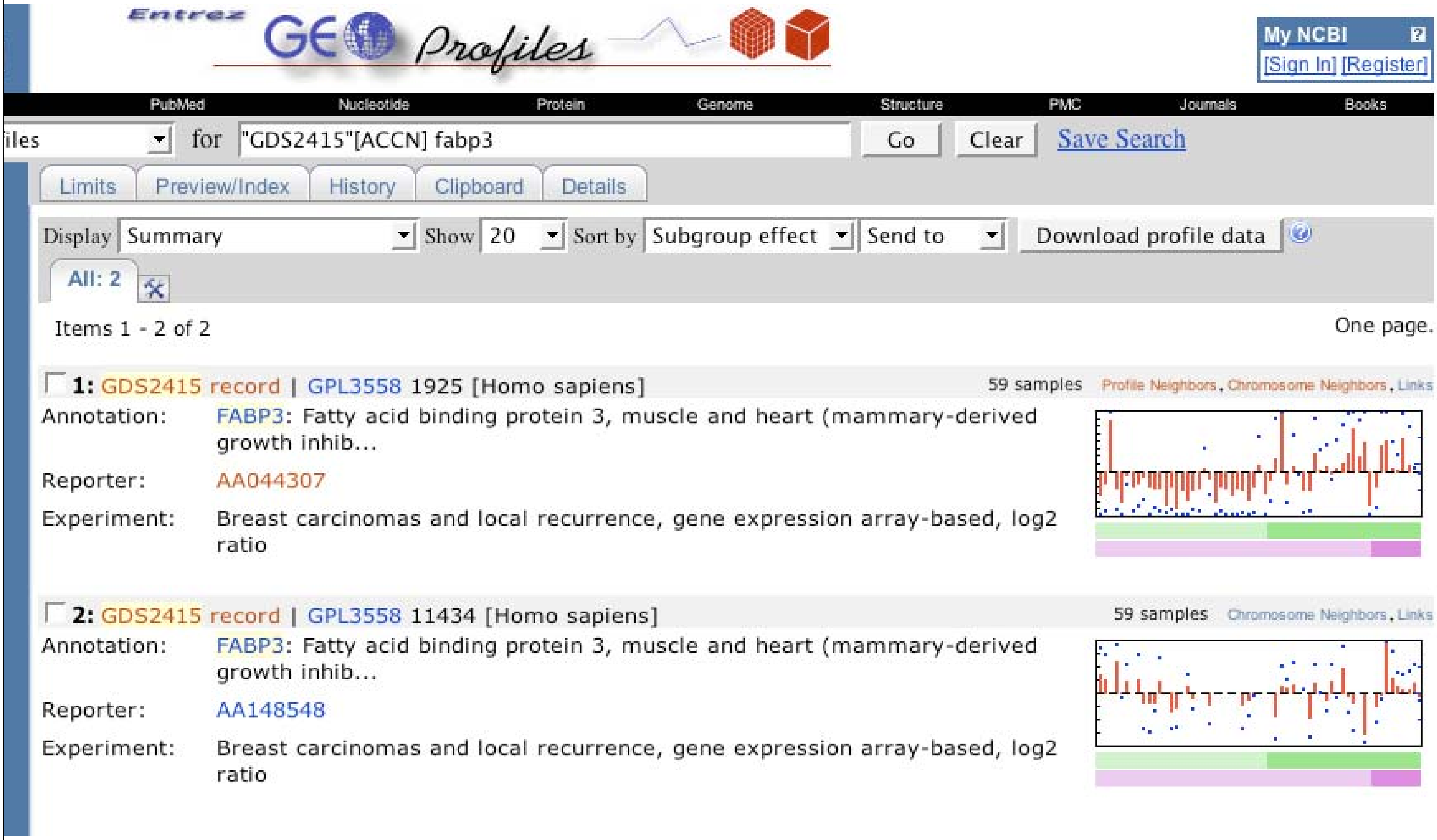
Find genes differentially expressed between groups.. Two-tailed t-test (A vs B)

Samples		Type	Description	A	0.050 significance level	B
☑ (31)	☐	disease state	no tumor recurrence	☑	↔	☐
☑ (28)		disease state	tumor recurrence	☐	↔	☑
☑ (50)	☐	specimen	primary tumor	☐	↔	☐
☑ (9)		specimen	recurrent tumor	☐	↔	☐

☑ GDS2415 only ☑ ranks ☑ values

59 samples, order: none

☑ GSM110395 : wsb 1428	☑ GSM110396 : wsb 1631	☑ GSM110397 : wsb 1642	☑ GSM110398 : wsb 1694
src 1: breast carcinoma	src 1: breast carcinoma	src 1: breast carcinoma	src 1: breast carcinoma
src 2: breast carcinoma	src 2: breast carcinoma	src 2: breast carcinoma	src 2: breast carcinoma



for "GDS2415"[ACCN] fabp3

Display Summary
Show 20
Sort by Subgroup effect
Send to

Items 1 - 2 of 2

One page.

☐ **1:** GDS2415 record | GPL3558 1925 [Homo sapiens] 59 samples

Annotation: FABP3: Fatty acid binding protein 3, muscle and heart (mammary-derived growth inhib...

Reporter: AA044307

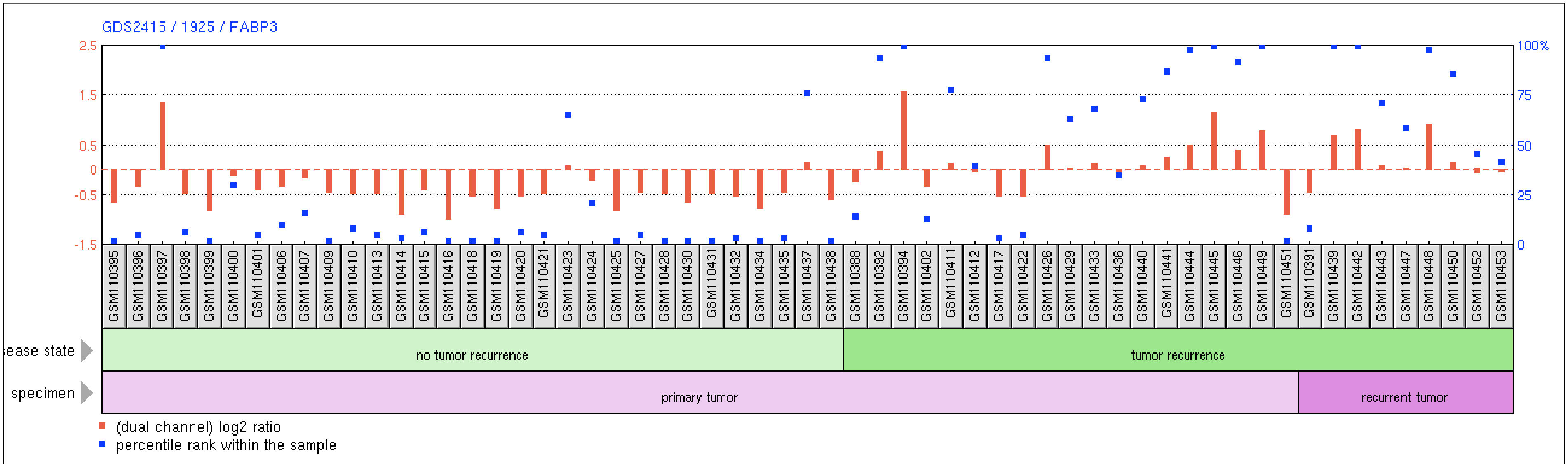
Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

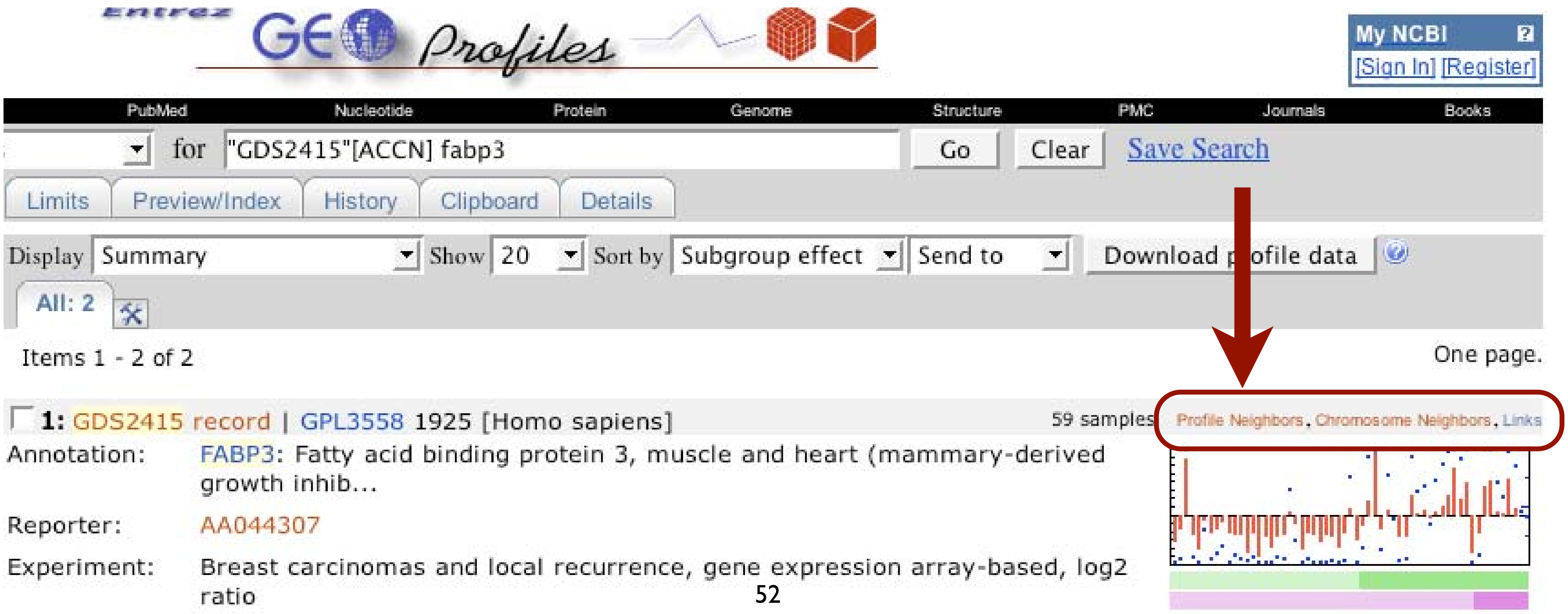
☐ **2:** GDS2415 record | GPL3558 11434 [Homo sapiens] 59 samples

Annotation: FABP3: Fatty acid binding protein 3, muscle and heart (mammary-derived growth inhib...

Reporter: AA148548

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio





"GDS2415"[ACCN] fabp3

Display Summary Show 20 Sort by Subgroup effect

Items 1 - 2 of 2

One page.

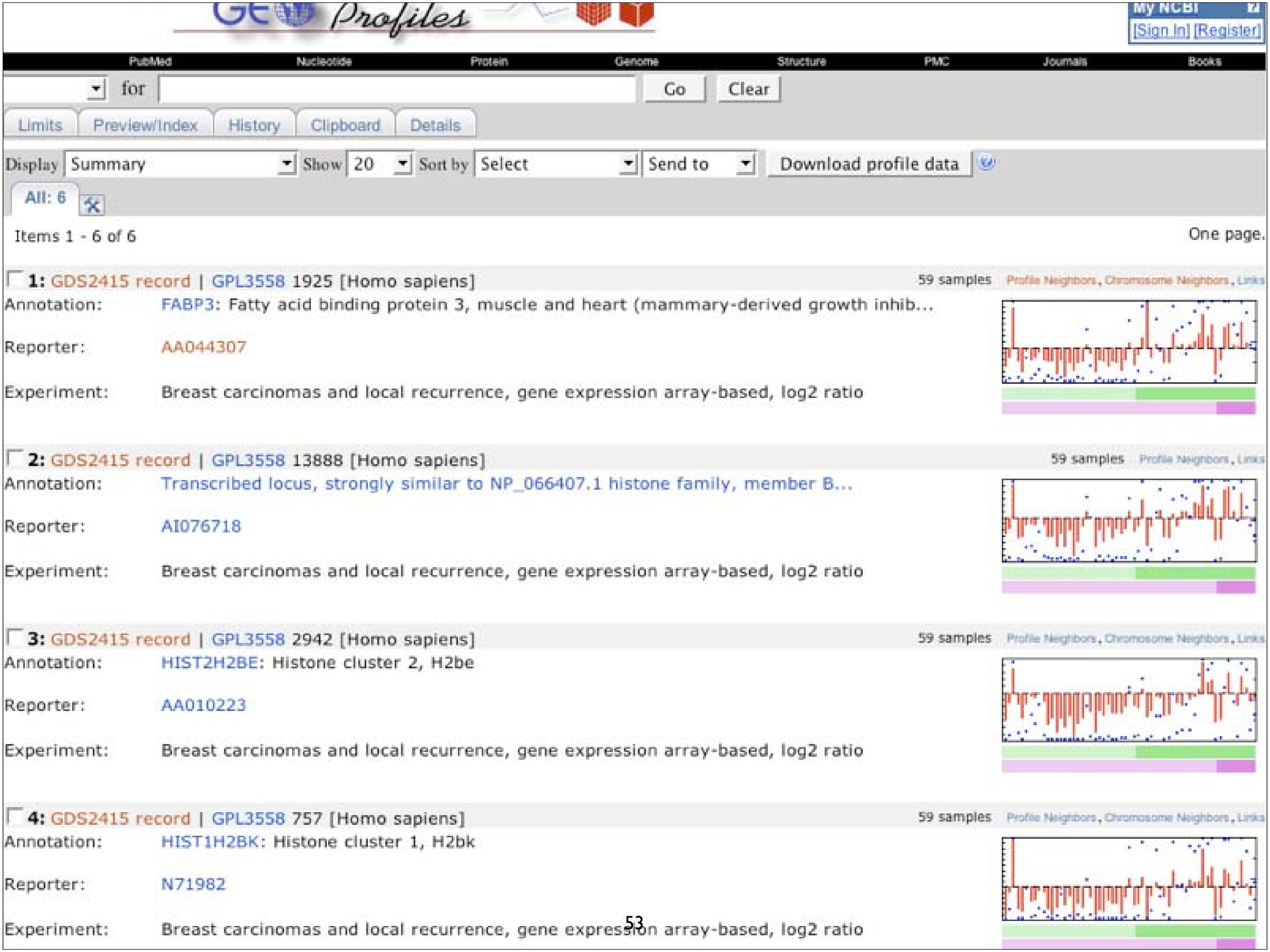
☐ **1:** GDS2415 record | GPL3558 1925 [Homo sapiens] 59 samples

Profile Neighbors, Chromosome Neighbors, Links

Annotation: FABP3: Fatty acid binding protein 3, muscle and heart (mammary-derived growth inhib...

Reporter: AA044307

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio



Items 1 - 6 of 6

One page.

☐ **1:** GDS2415 record | GPL3558 1925 [Homo sapiens] 59 samples Profile Neighbors, Chromosome Neighbors, Links

Annotation: FABP3: Fatty acid binding protein 3, muscle and heart (mammary-derived growth inhib...

Reporter: AA044307

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **2:** GDS2415 record | GPL3558 13888 [Homo sapiens] 59 samples Profile Neighbors, Links

Annotation: Transcribed locus, strongly similar to NP_066407.1 histone family, member B...

Reporter: AI076718

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **3:** GDS2415 record | GPL3558 2942 [Homo sapiens] 59 samples Profile Neighbors, Chromosome Neighbors, Links

Annotation: HIST2H2BE: Histone cluster 2, H2be

Reporter: AA010223

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

☐ **4:** GDS2415 record | GPL3558 757 [Homo sapiens] 59 samples Profile Neighbors, Chromosome Neighbors, Links

Annotation: HIST1H2BK: Histone cluster 1, H2bk

Reporter: N71982

Experiment: Breast carcinomas and local recurrence, gene expression array-based, log2 ratio

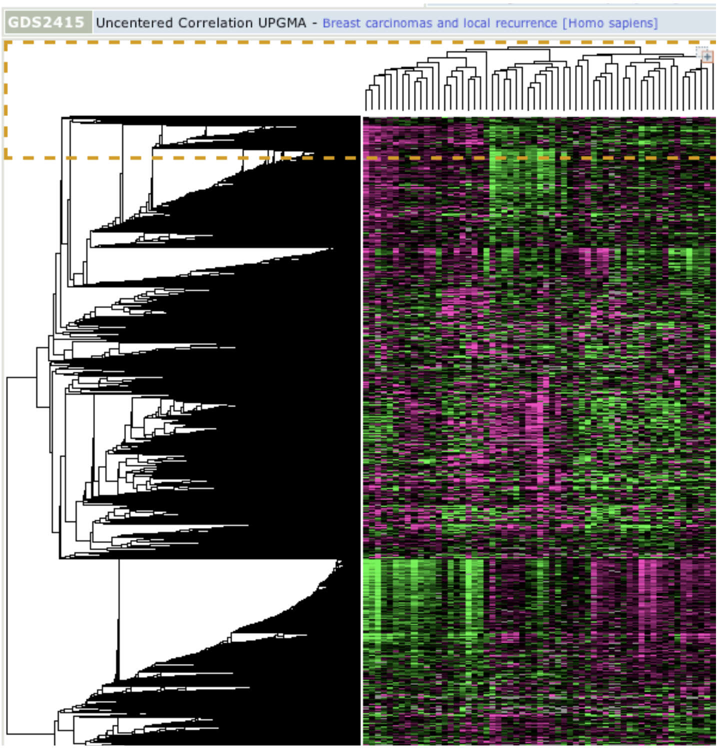
Profile Neighbors

Take home message: GEO data analysis tools are great for quick identification of interesting leads; you download the data to carry out more robust statistical analyses

- Connects groups of genes that have similar expression profiles within a DataSet
- pre-computed
- calculated by Pearson correlation coefficients

Other Features

- Cluster Heat Maps
  - precomputed sample and gene hierarchical cluster heat maps provided
  - different methods available; can select, expand, download
- GEO BLAST
  - retrieve gene expression profiles by sequence similarity



GEO, the gene expression omnibus

- public repository of expression data from many different experimental platforms
- Main uses
  - ✓ search for experiments of interest
  - ✓ search for expression information about gene of interest
- submit, search, analyses tools available
- data standards required MIAME, MiNiML

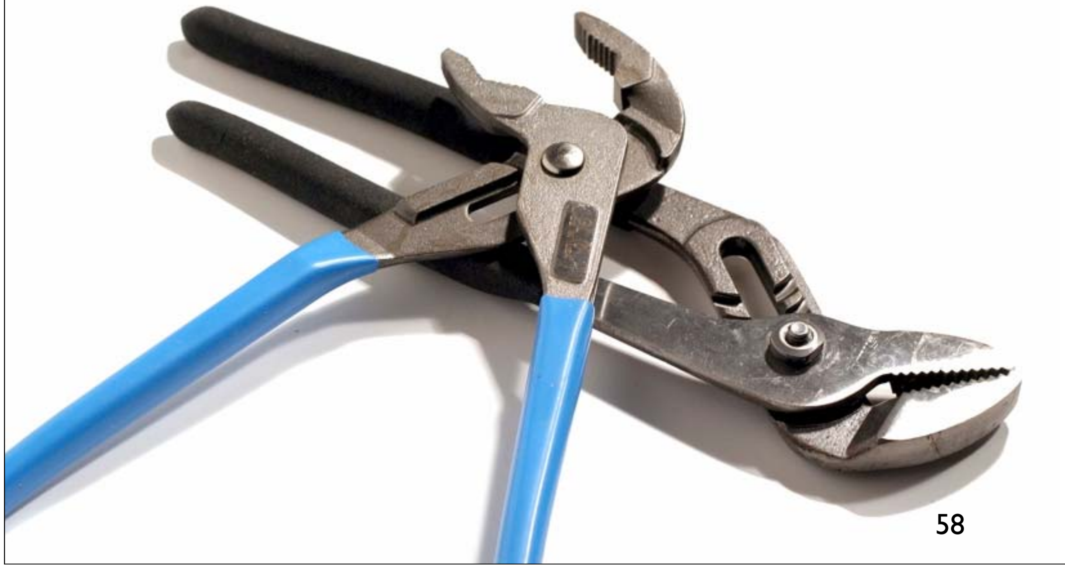
Credits & References

- NCBI GEO: mining tens of millions of expression profiles–database and tools update. Barrett T, et al. Nucleic Acids Res. 35 (2007) D760-5. [PMID: 17099226]
- GEO: the Gene Expression Omnibus

  http://www.ncbi.nlm.nih.gov/projects/geo/info/GEOHandoutFinal.pdf
- Dr. Paul Pavlidis, UBC Bioinformatics Centre

Bioinformatics

Session 3.2 - Pathway Resources for Systems Biology



Proteomics

- How large is the human proteome, anyway?

Class	Size	Description
Non Redundant Proteins	20,000-25,000	representative protein from every gene locus
Variants	50,000-500,000	different proteins obtained by splicing or proteolysis
Combinatorial Variants	>10,000,000	different proteins generated by somatic DNA rearrangements
Protein Species	>100,000	proteins that differ in chemical composition due to PTM
Protein Alleles	75,000-150,000	proteins that differ by genetic variation (coding SNPs)

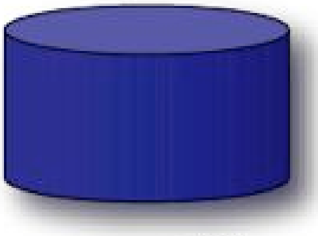
source: Uhlen M et al, Molecular & Cellular Proteomics (2005)

Cellular Pathways

- A striking similarity between intracellular signaling pathways and the Tokyo subway system

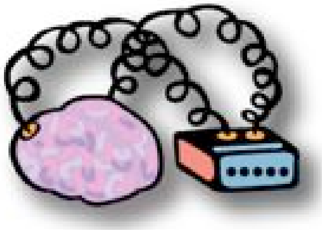
Pathway Information

- Databases
  - Fully electronic
  - Easily computer readable

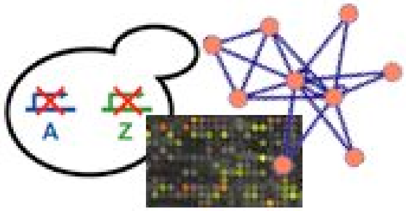


- Literature
  - Increasingly electronic
  - Human readable

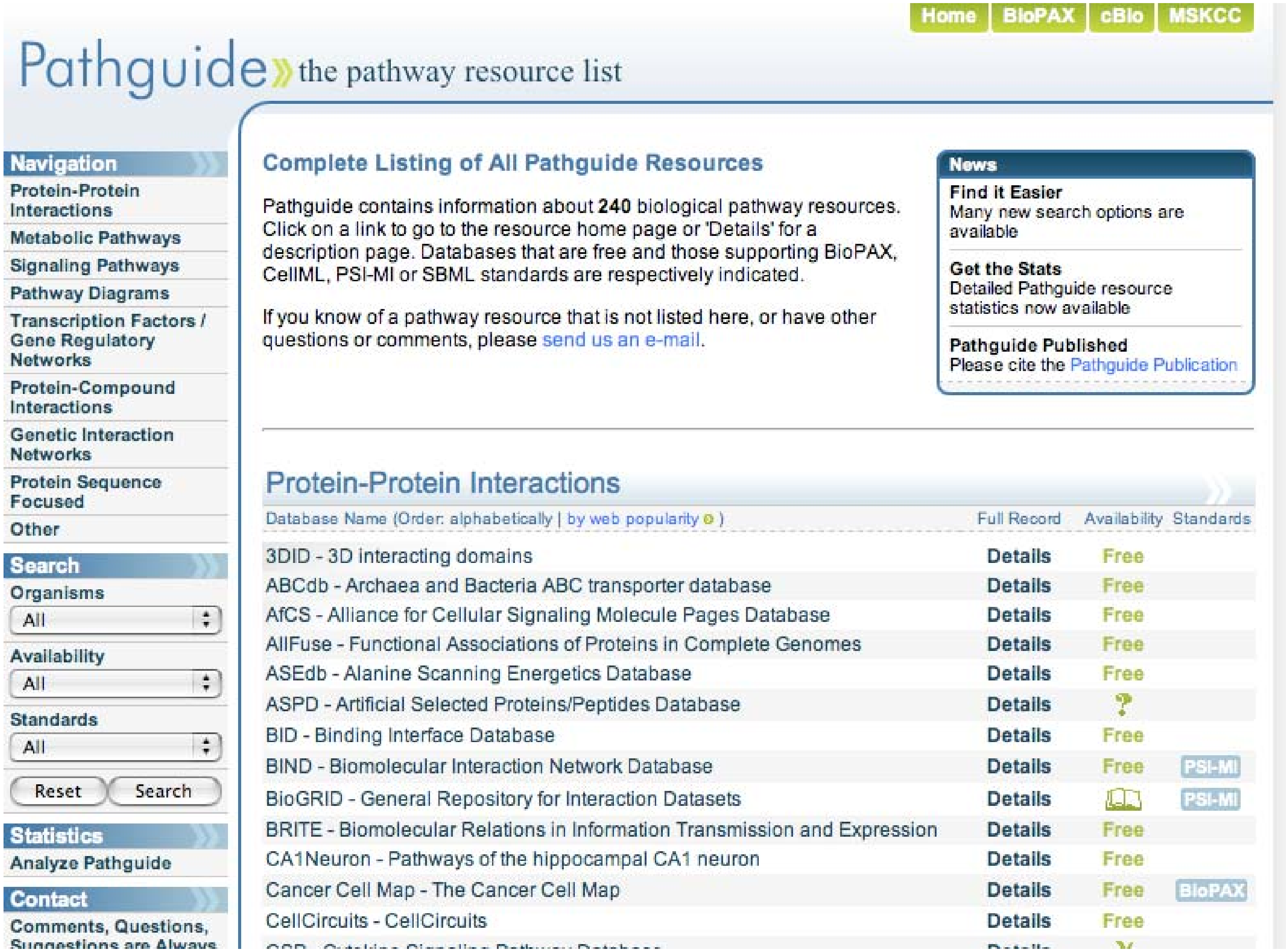
- Biologist's brains
  - Richest data source
  - Limited bandwidth access



- Experiments
  - Basis for models



http://www.pathguide.org/



Pathguide » the pathway resource list

Navigation

- Protein-Protein Interactions
- Metabolic Pathways
- Signaling Pathways
- Pathway Diagrams
- Transcription Factors / Gene Regulatory Networks
- Protein-Compound Interactions
- Genetic Interaction Networks
- Protein Sequence Focused
- Other

Search

Organisms: All

Availability: All

Standards: All

Statistics

Analyze Pathguide

Contact

Comments, Questions, Suggestions are Always

Complete Listing of All Pathguide Resources

Pathguide contains information about **240** biological pathway resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

News

Find it Easier
Many new search options are available

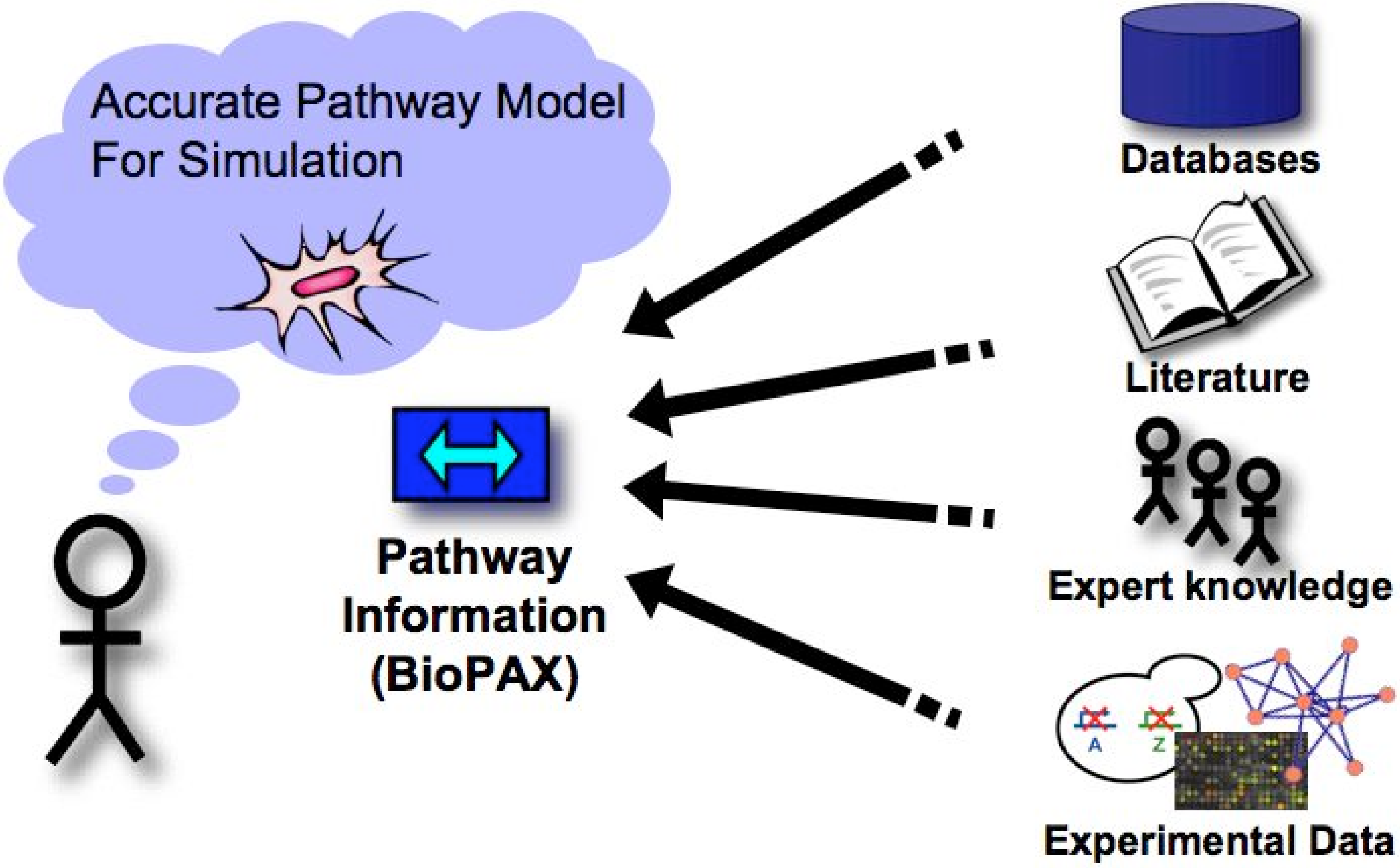
Get the Stats
Detailed Pathguide resource statistics now available

Pathguide Published
Please cite the Pathguide Publication

Protein-Protein Interactions

Database Name (Order: alphabetically | by web popularity)	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	Free	
ABCdb - Archaea and Bacteria ABC transporter database	Details	Free	
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Details	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	Details	Free	
ASEdb - Alanine Scanning Energetics Database	Details	Free	
ASPD - Artificial Selected Proteins/Peptides Database	Details	?	
BID - Binding Interface Database	Details	Free	
BIND - Biomolecular Interaction Network Database	Details	Free	PSI-MI
BioGRID - General Repository for Interaction Datasets	Details	📖	PSI-MI
BRITE - Biomolecular Relations in Information Transmission and Expression	Details	Free	
CA1Neuron - Pathways of the hippocampal CA1 neuron	Details	Free	
Cancer Cell Map - The Cancer Cell Map	Details	Free	BioPAX
CellCircuits - CellCircuits	Details	Free	

Using Pathway Information



Accurate Pathway Model For Simulation

Pathway Information (BioPAX)

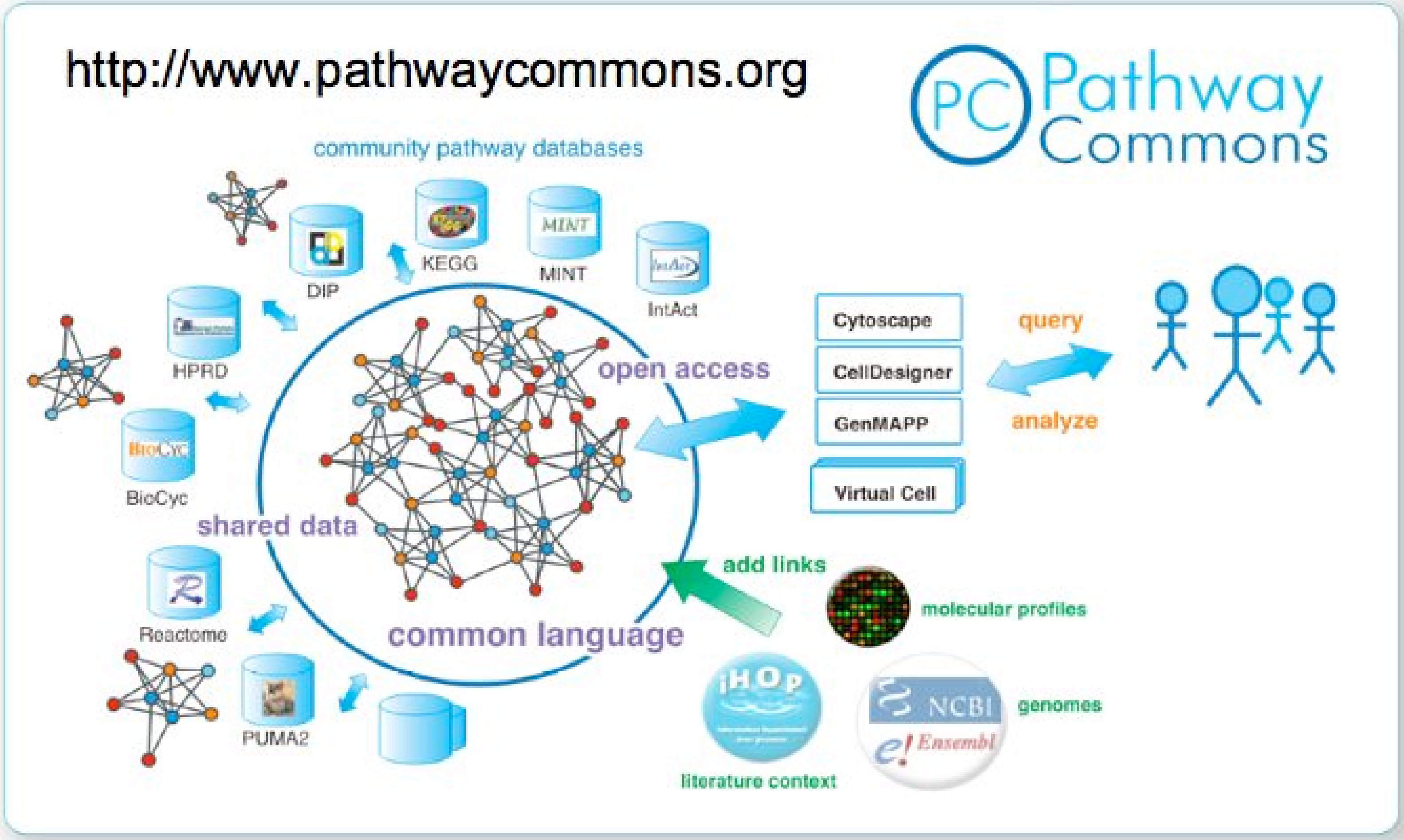
Databases

Literature

Expert knowledge

Experimental Data

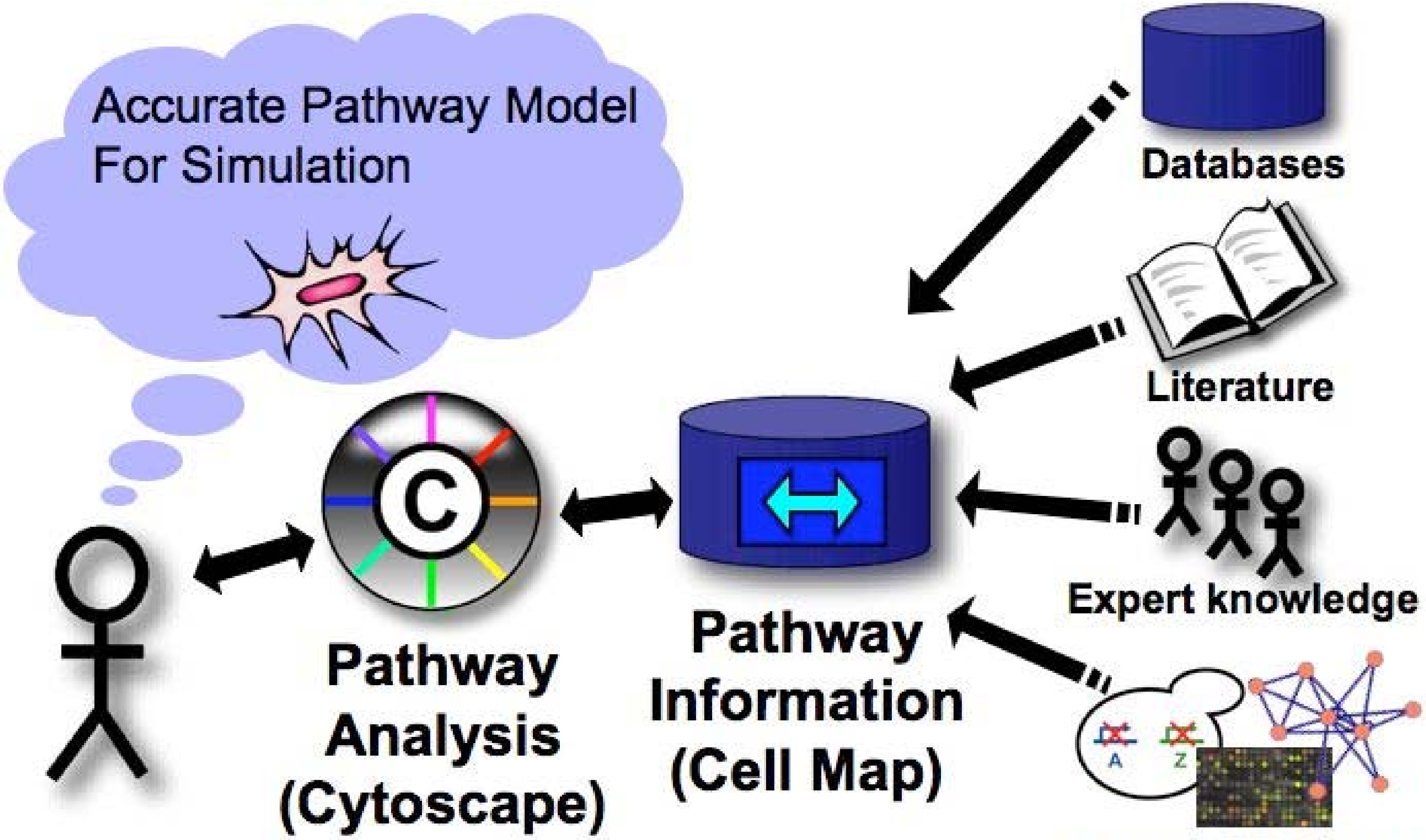
Aim: Convenient Access to Pathway Information



Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge to integrated cell map

Using Pathway Information



Accurate Pathway Model For Simulation

Databases

Literature

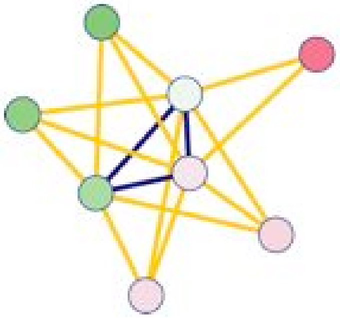
Expert knowledge

Experimental Data

Pathway Analysis (Cytoscape)

Pathway Information (Cell Map)

Cytoscape - Network Visualization and Analysis

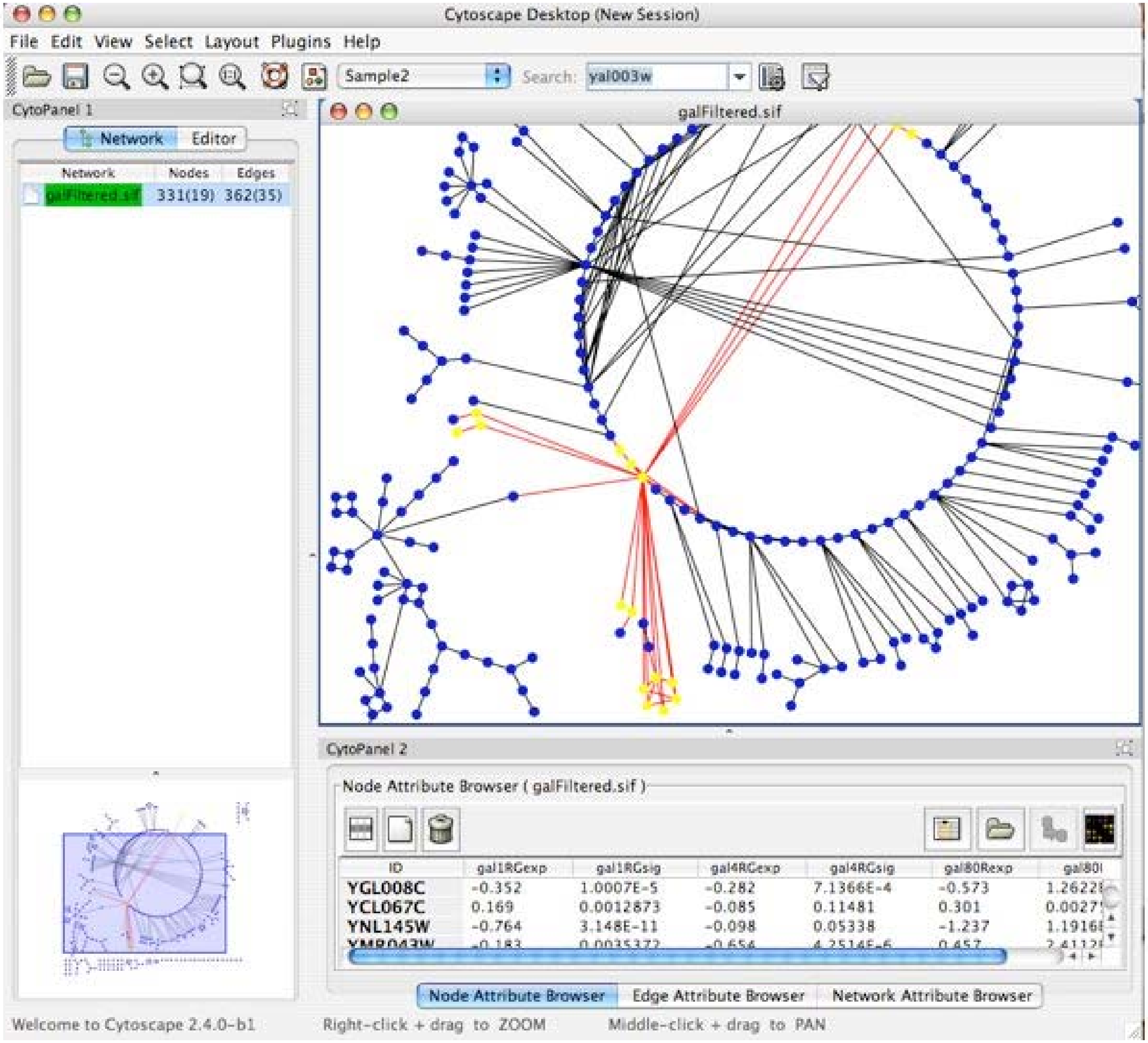


http://cytoscape.org

- Freely-available (open-source, java) software
- Visualizing biological networks (e.g. molecular interaction networks)
- Analyzing networks with gene expression profiles and other cell state data

UCSD, ISB, Agilent, MSKCC, Pasteur, UCSF, UToronto

Other software: Osprey, BioLayout, VisANT, Navigator, PIMWalker, ProViz

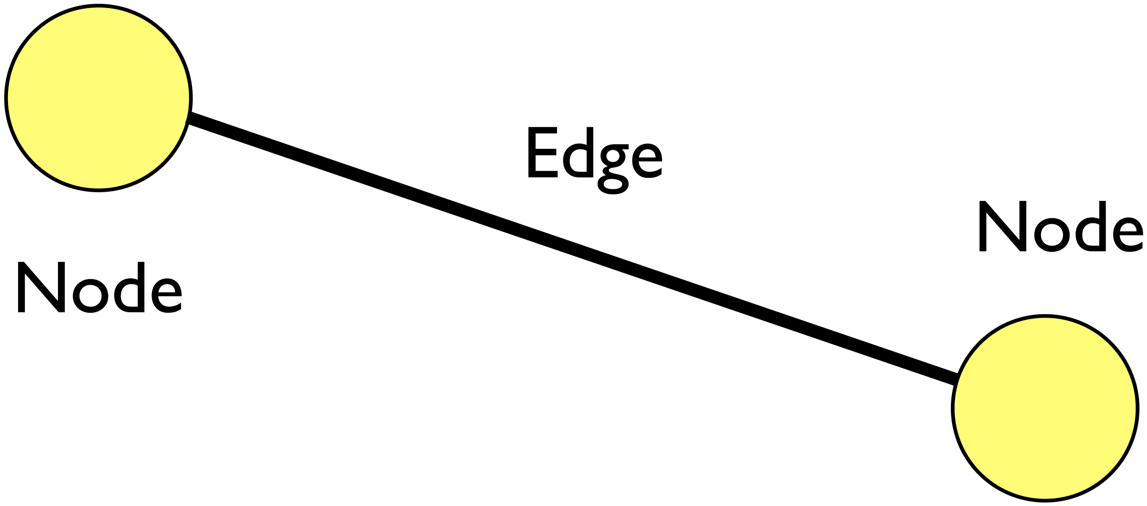


Network	Nodes	Edges
galFiltered.sif	331(19)	362(35)

ID	gal1RGexp	gal1RGsig	gal4RGexp	gal4RGsig	gal80Rexp	gal80I
YGL008C	-0.352	1.0007E-5	-0.282	7.1366E-4	-0.573	1.2622
YCL067C	0.169	0.0012873	-0.085	0.11481	0.301	0.0027
YNL145W	-0.764	3.148E-11	-0.098	0.05338	-1.237	1.1916
YMR043W	-0.183	0.0035372	-0.654	4.2514E-6	0.457	2.4112

Welcome to Cytoscape 2.4.0-b1   Right-click + drag to ZOOM   Middle-click + drag to PAN

Pathway Graphs

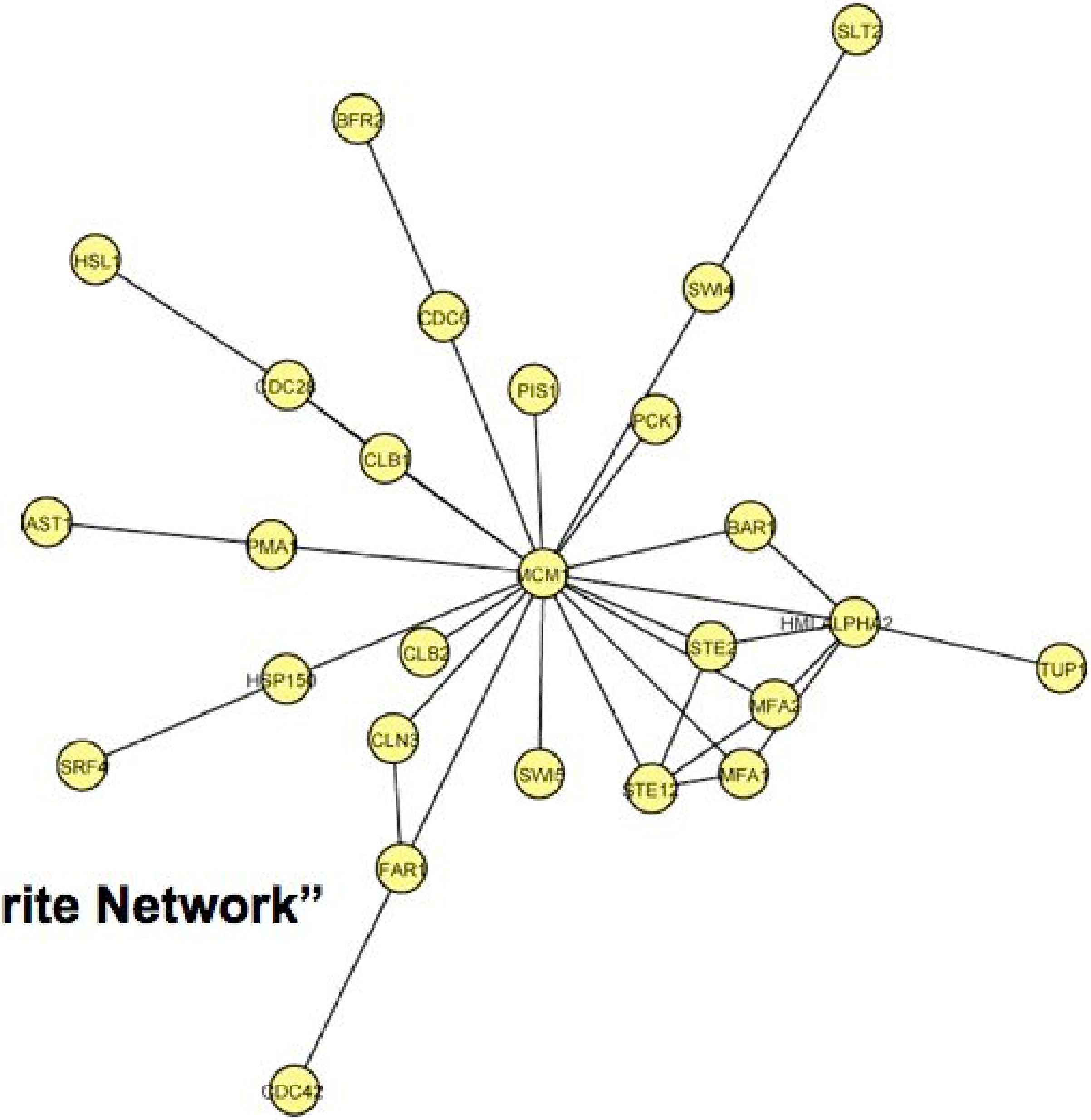


- In addition to describing the network topology, nodes and edges can each have their own attributes

Visual Style

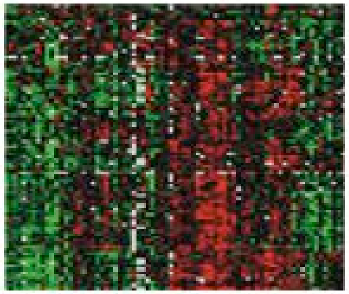
- Customized views of experimental data in a network context
- Network has node and edge attributes
    - E.g. expression data, interaction type, GO function
- Mapped to visual attributes
    - E.g. node/edge size, shape, colour…
- E.g. Visualize gene expression data as node colour gradient on the network

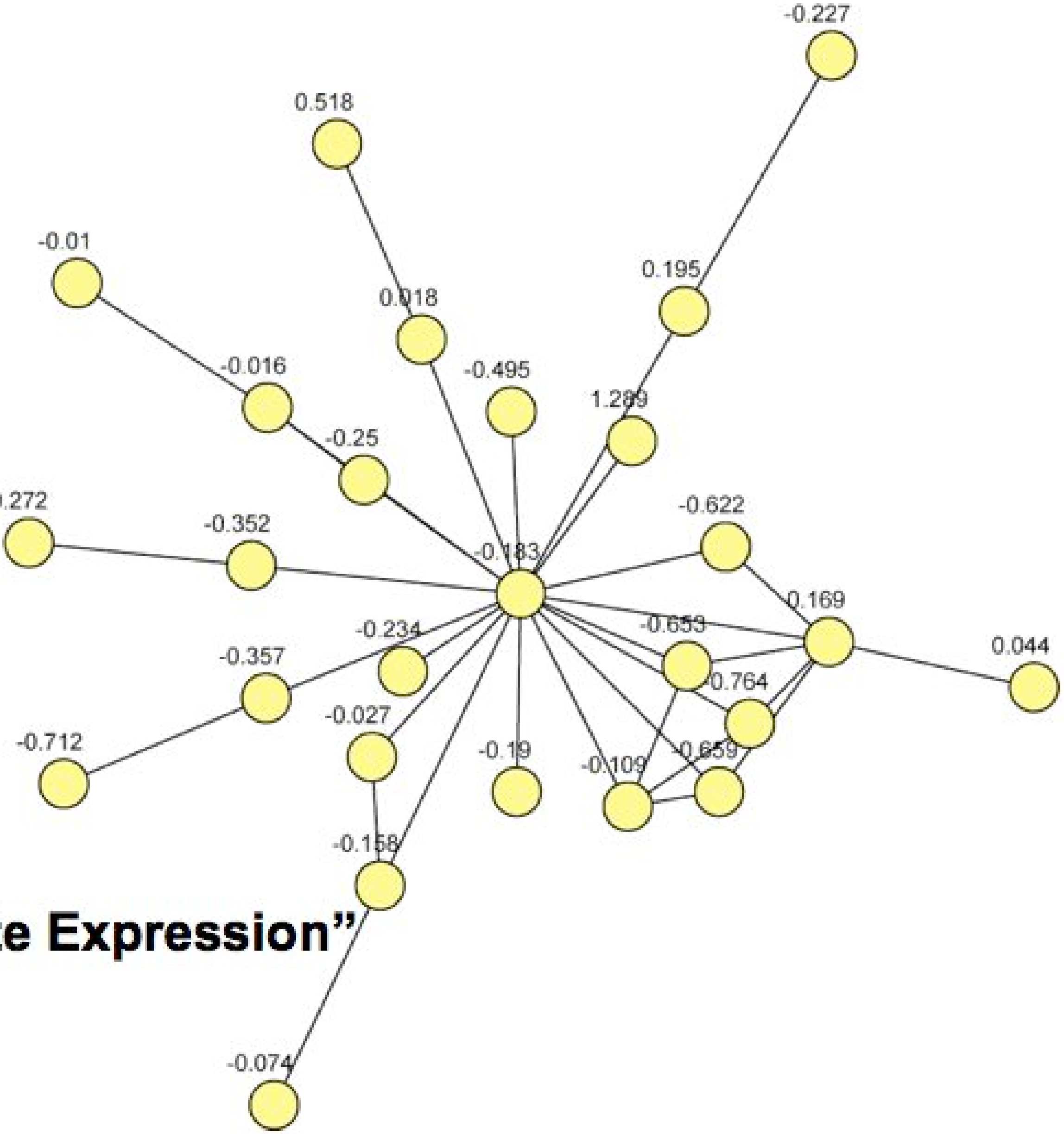
Visual Style



Load “Your Favorite Network”

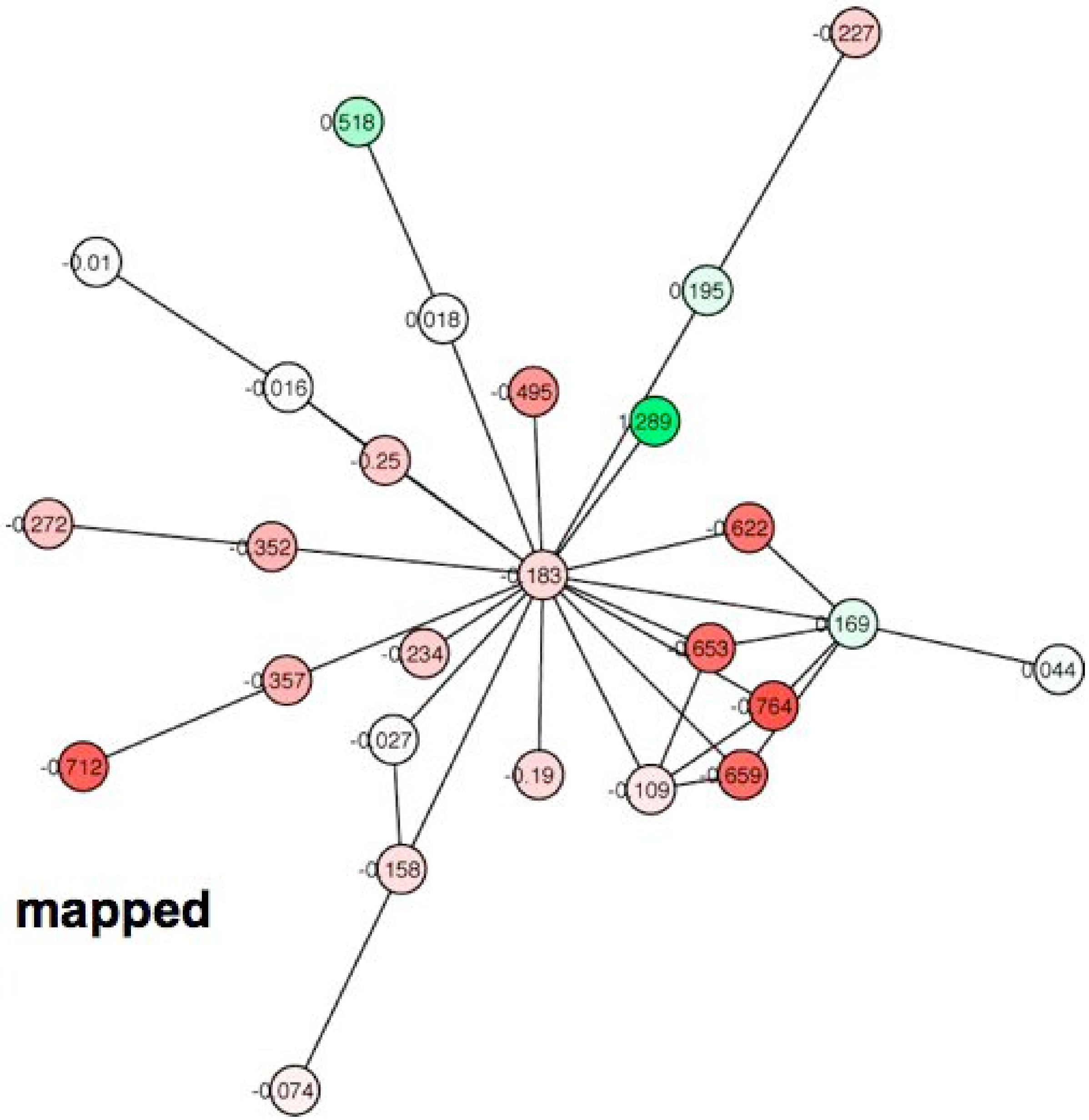
Visual Style





Load "Your Favorite Expression" Dataset

Visual Style



Expression data mapped to node colours

Systems Biology

- Goals:
  - ✓ integrating diverse data types, pathways
  - ✓ cellular simulations
- Community approaches:
  - ✓ pathguide, pathway commons, cytoscape
- Open data exchange key to success

Credits & References

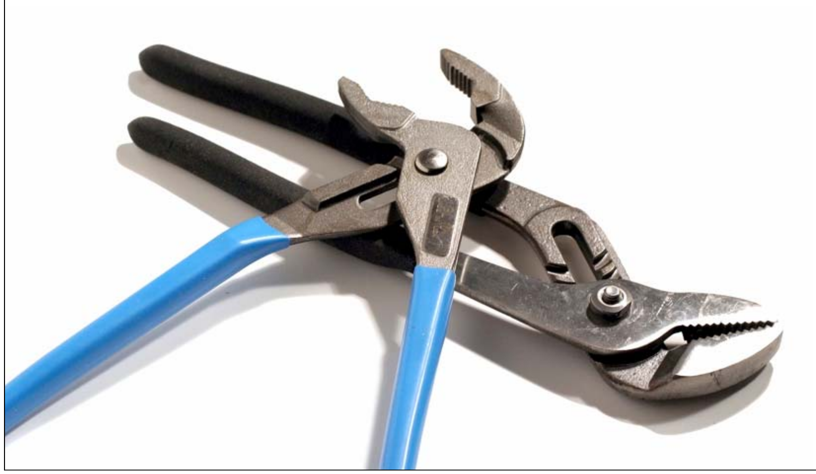
- Dr. Gary Bader, DCCBR, UofT

  slides/images used with permission

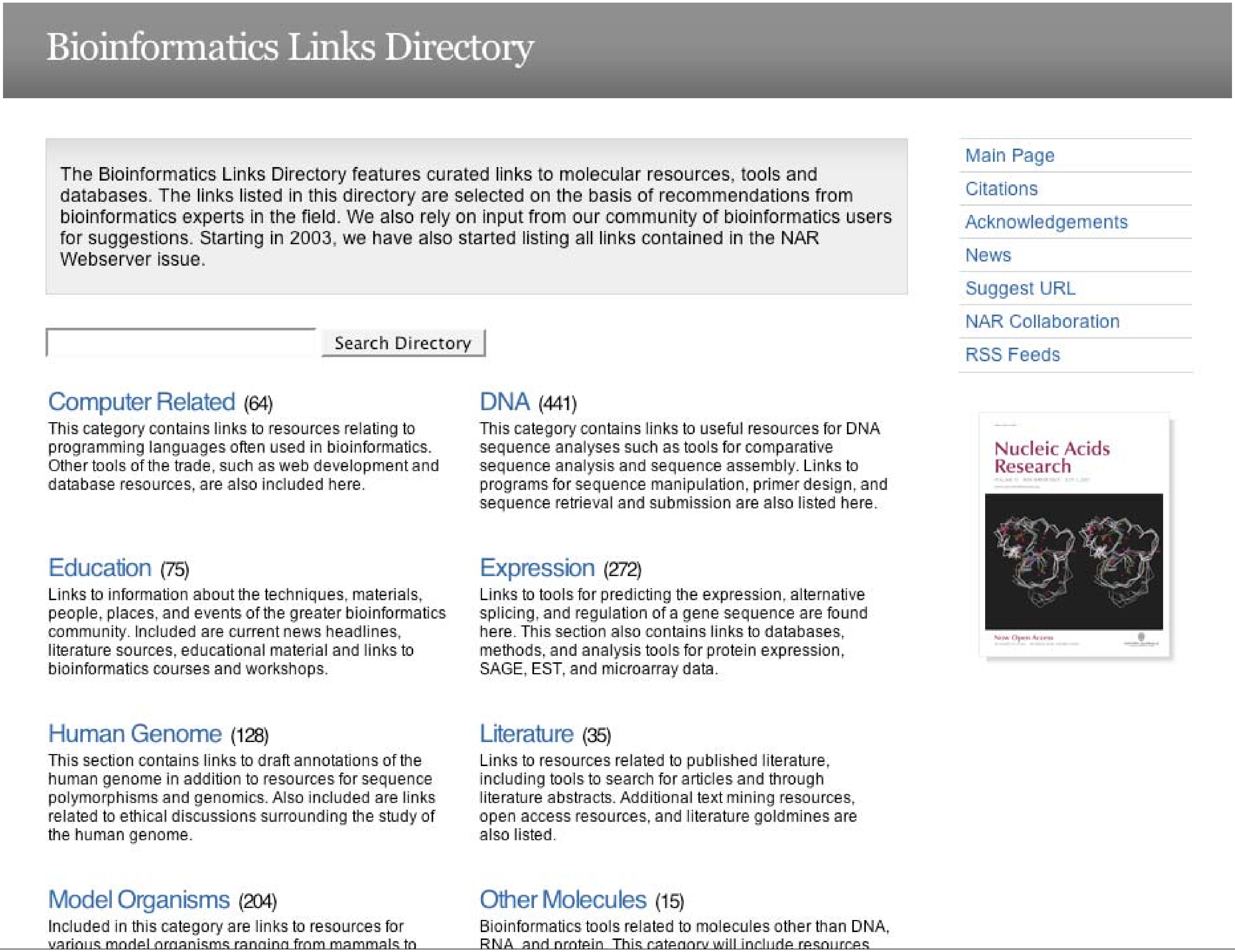
- Cary MP, Bader GD, Sander C "Pathway Information for Systems Biology", FEBS Letters (2005)

Bioinformatics Links Directory

Finding online tools & resources for Life Sciences research



http://bioinformatics.ca/links_directory/



Bioinformatics Links Directory

The Bioinformatics Links Directory features curated links to molecular resources, tools and databases. The links listed in this directory are selected on the basis of recommendations from bioinformatics experts in the field. We also rely on input from our community of bioinformatics users for suggestions. Starting in 2003, we have also started listing all links contained in the NAR Webserver issue.

Computer Related (64)
This category contains links to resources relating to programming languages often used in bioinformatics. Other tools of the trade, such as web development and database resources, are also included here.

DNA (441)
This category contains links to useful resources for DNA sequence analyses such as tools for comparative sequence analysis and sequence assembly. Links to programs for sequence manipulation, primer design, and sequence retrieval and submission are also listed here.

Education (75)
Links to information about the techniques, materials, people, places, and events of the greater bioinformatics community. Included are current news headlines, literature sources, educational material and links to bioinformatics courses and workshops.

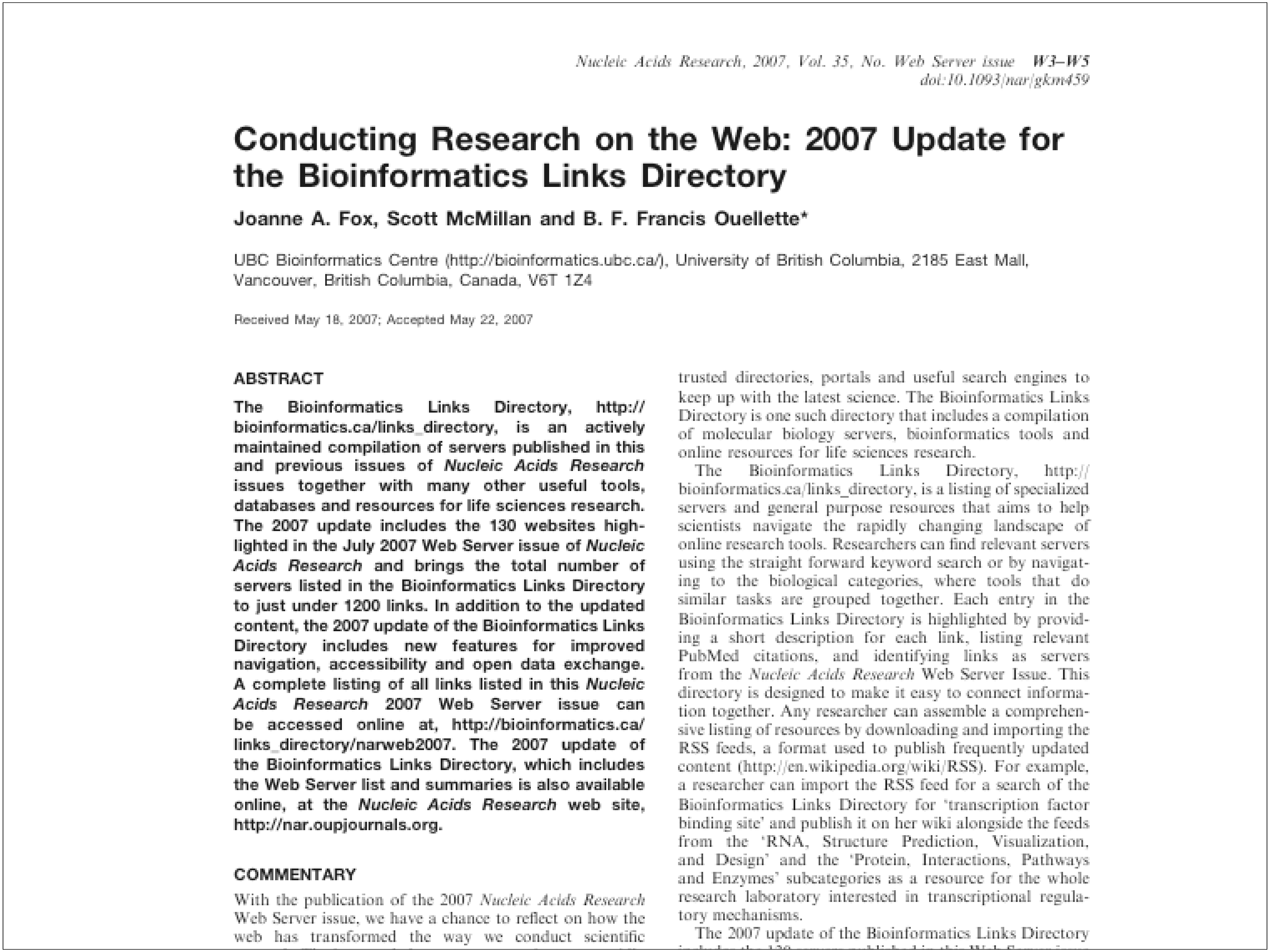
Expression (272)
Links to tools for predicting the expression, alternative splicing, and regulation of a gene sequence are found here. This section also contains links to databases, methods, and analysis tools for protein expression, SAGE, EST, and microarray data.

Human Genome (128)
This section contains links to draft annotations of the human genome in addition to resources for sequence polymorphisms and genomics. Also included are links related to ethical discussions surrounding the study of the human genome.

Literature (35)
Links to resources related to published literature, including tools to search for articles and through literature abstracts. Additional text mining resources, open access resources, and literature goldmines are also listed.

Model Organisms (204)
Included in this category are links to resources for various model organisms ranging from mammals to

Other Molecules (15)
Bioinformatics tools related to molecules other than DNA, RNA, and protein. This category will include resources



Nucleic Acids Research, 2007, Vol. 35, No. Web Server issue W3–W5
doi:10.1093/nar/gkm459

Conducting Research on the Web: 2007 Update for the Bioinformatics Links Directory

Joanne A. Fox, Scott McMillan and B. F. Francis Ouellette*

UBC Bioinformatics Centre (http://bioinformatics.ubc.ca/), University of British Columbia, 2185 East Mall, Vancouver, British Columbia, Canada, V6T 1Z4

Received May 18, 2007; Accepted May 22, 2007

ABSTRACT

The Bioinformatics Links Directory, http://bioinformatics.ca/links_directory, is an actively maintained compilation of servers published in this and previous issues of *Nucleic Acids Research* issues together with many other useful tools, databases and resources for life sciences research. The 2007 update includes the 130 websites highlighted in the July 2007 Web Server issue of *Nucleic Acids Research* and brings the total number of servers listed in the Bioinformatics Links Directory to just under 1200 links. In addition to the updated content, the 2007 update of the Bioinformatics Links Directory includes new features for improved navigation, accessibility and open data exchange. A complete listing of all links listed in this *Nucleic Acids Research* 2007 Web Server issue can be accessed online at, http://bioinformatics.ca/links_directory/narweb2007. The 2007 update of the Bioinformatics Links Directory, which includes the Web Server list and summaries is also available online, at the *Nucleic Acids Research* web site, http://nar.oupjournals.org.

COMMENTARY

With the publication of the 2007 *Nucleic Acids Research* Web Server issue, we have a chance to reflect on how the web has transformed the way we conduct scientific

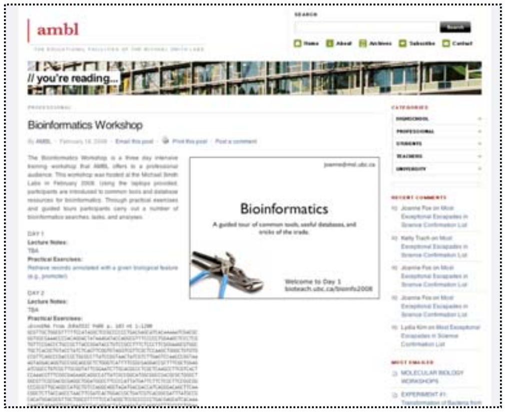
trusted directories, portals and useful search engines to keep up with the latest science. The Bioinformatics Links Directory is one such directory that includes a compilation of molecular biology servers, bioinformatics tools and online resources for life sciences research.

The Bioinformatics Links Directory, http://bioinformatics.ca/links_directory, is a listing of specialized servers and general purpose resources that aims to help scientists navigate the rapidly changing landscape of online research tools. Researchers can find relevant servers using the straight forward keyword search or by navigating to the biological categories, where tools that do similar tasks are grouped together. Each entry in the Bioinformatics Links Directory is highlighted by providing a short description for each link, listing relevant PubMed citations, and identifying links as servers from the *Nucleic Acids Research* Web Server Issue. This directory is designed to make it easy to connect information together. Any researcher can assemble a comprehensive listing of resources by downloading and importing the RSS feeds, a format used to publish frequently updated content (http://en.wikipedia.org/wiki/RSS). For example, a researcher can import the RSS feed for a search of the Bioinformatics Links Directory for 'transcription factor binding site' and publish it on her wiki alongside the feeds from the 'RNA, Structure Prediction, Visualization, and Design' and the 'Protein, Interactions, Pathways and Enzymes' subcategories as a resource for the whole research laboratory interested in transcriptional regulatory mechanisms.

The 2007 update of the Bioinformatics Links Directory

Your Feedback is Important!

navigate to:
bioteach.ubc.ca/bioinfo2008



Bioinformatics Workshop

DAY 1
Lecture Notes:
TBA
Practical Exercises:

DAY 2
Lecture Notes:
TBA
Practical Exercises:

Bioinformatics
A guided tour of common tools, useful databases, and tricks of the trade.
Welcome to Day 1
bioteach.ubc.ca/bioinfo2008

MOLECULAR BIOLOGY WORKSHOPS
EXPERIMENT #1:


Workshop Evaluation:
Click Here to take survey

Thanks for attending the AMBL Bioinformatics Workshop

- Questions? Please contact:

Dr. Joanne Fox
Michael Smith Laboratories
joanne@msl.ubc.ca