

A case study in cross-platform, cross-species data integration to yield novel discovery in the biology of transcription networks and cancer
Vancouver
September 8 2005

Win Hide, Ludwig Kerr International Fellow
SANBI, University of the Western Cape

Using gene expression integration to understand disease

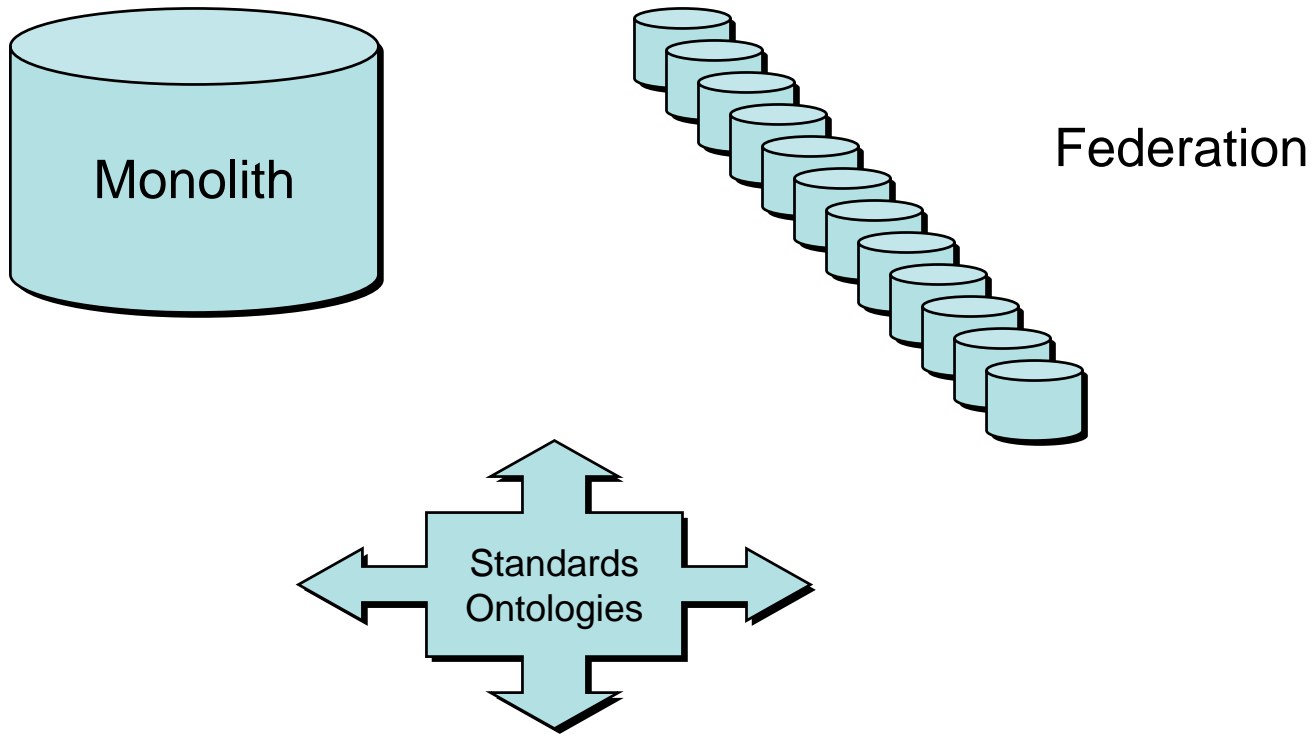


Applying genome biology

- Apprehend biological data
- Place it into association and context
- Define the relationships between objects in the form of networks
- Seek to understand relationships and perturbations of the networks
- Gain understanding of biology as a system



Integration Approaches



Using Ontologies: Definition

A (potentially incomplete) set of **terms**

Which cover some area of interest ("**domain**").

Each term is **defined**

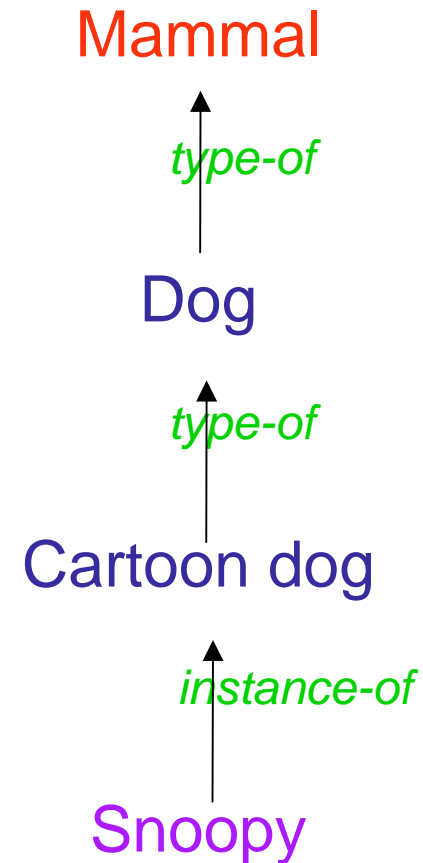
Each has a specified **relationship** with its parent and child terms

- All include
 - Vocabulary of terms
 - Definition of the meaning of each term



Essential aspects of an ontology

- Domains
- Concepts
- Relations
- Instances
- Axioms



Axioms: A dog is a fissiped mammal with nonretractile claws and typically long muzzle

An ontology for expression

- Gene expression data only have meaning in the context of a detailed description of the sample.
 - Tissue, age, sex, disease, time of day, nutritional status...
- Free text dump



Libraries that match to a Unigene entry

well-differentiated endometrial adenocarcinoma, 7 pooled tumors ;spleen ;epithelioid carcinoma ;breast ;moderately-differentiated adenocarcinoma ;hypernephroma ;glioblastoma with EGFR amplification ;anaplastic oligodendroglioma with 1p/19q loss ;Lens ;lung_normal ;prostate_normal ;sciatic nerve ;dorsal root ganglia ;Iris ;adrenal cortex carcinoma, cell line ;prostate_tumor ;colon_est ;colon_ins ;hypothalamus ;leiomyosarcoma ;nervous_normal ; melanocyte ;hypernephroma, cell line ;adenocarcinoma, cell line ;nervous_tumor ;moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors ; placenta_normal ;mammary adenocarcinoma, cell line ;placenta ;thymus, pooled ;normal epithelium ;adenocarcinoma ;Pooled human melanocyte, fetal heart, and pregnant uterus ;Burkitt lymphoma ;colon ;colon tumor, RER+ ;B cells from Burkitt lymphoma ;large cell carcinoma ;Fetal brain ;T cells from T cell leukemia ;lymphocyte ;multiple sclerosis lesions ;skin ;normal pigmented retinal epithelium ;melanotic melanoma, high MDR (cell line) ; melanotic melanoma, cell line ;leukopheresis ;rhabdomyosarcoma ;amelanotic melanoma, cell line ;neuroblastoma cells ;lung_tumor ;abdominal aortic adventitia from an aneurysm specimen ;pooled brain, lung, testis ;hippocampus ;epithelioid carcinoma cell line ;leiomyosarcoma ;osteosarcoma, cell line ;kidney_tumor ;cervical carcinoma cell line ;leukocyte ;medulla ;pooled lung and spleen ;duodenal adenocarcinoma, cell line ;lens ;human retina ;pooled pancreas and spleen ;fetal eyes ;insulinoma ;fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid ;cartilage ;pooled colon, kidney, stomach ;fetal eye ;cervix ;endometrium, adenocarcinoma cell line ;optic nerve ;RPE and Choroid ;Ascites ;Stomach ;Lung Focal Fibrosis ;heart ;Chondrosarcoma ;senescent fibroblast ;squamous cell carcinoma ;Primary Lung Cystic Fibrosis Epithelial Cells ;hepatocellular carcinoma, cell line ;Placenta ;epidermoid carcinoma, cell line ;Chondrosarcoma Grade II ;breast_normal ;Osteoarthritic Cartilage ; Fibrosarcoma ;neuroblastoma, cell line ;testis ;Chondrosarcoma Cell line ;Retina ;glioblastoma without EGFR amplification ;ductal carcinoma, cell line ;retina ;carcinoma, cell line ;teratocarcinoma, cell line ;embryonal carcinoma, cell line ;Cell lines ;bone marrow stroma ;neuroblastoma ; lymph ;lymphoma, follicular mixed small and large cell ;head_neck ;small cell carcinoma ;2 pooled tumors (clear cell type) ;moderately differentiated adenocarcinoma ;2 pooled high-grade transitional cell tumors ;five pooled sarcomas, including myxoid liposarcoma, solitary fibrous tumor, malignant fibrous histiocytoma, gastrointestinal stromal tumor, and mesothelioma ;adenocarcinoma cell line ;pnet ;Islets of Langerhans ;serous papillary carcinoma, high grade, 2 pooled tumors ;Liver and Spleen ;retinoblastoma ;renal cell adenocarcinoma ;pooled ;Lung ;ovary (pool of 3) ;anaplastic oligodendroglioma ;whole brain ;renal cell tumor ;lung ;poorly differentiated adenocarcinoma with signet ring cell features ;melanotic melanoma ;choriocarcinoma ;brain ;glioblastoma (pooled) ;Metastatic Chondrosarcoma ;Subchondral Bone ;B-cell, chronic lymphocytic leukemia ;pooled germ cell tumors ;uterus ;prostate ;Human Lung Epithelial cells ;germinal center B cell ;Cell Line ;Alveolar Macrophage ;Primary Lung Epithelial Cells ;Alveolar Macrophage ;kidney ;poorly-differentiated endometrial adenocarcinoma, 2 pooled tumors ;colonic mucosa from 3 patients with Crohn's disease ;colon tumor RER+ ;two pooled squamous cell carcinomas ;aorta ;ovarian tumor ;normal prostate ;normal prostatic epithelial cells ;tumor ;colon tumor ; dorsal root ganglion ;metastatic prostate bone lesion

eVOC – Expression VOCabulary

- A **hierarchical** controlled vocabulary
- Describes the **sample source** of cDNA and SAGE libraries, MPSS, Ditag, CAGE and target cDNAs for microarray experiments

Aim

- Link public expression data to get a broad, rich view of the transcriptome
 - EST
 - SAGE, CAGE
 - Microarray



eVOC Domains

- Human and mouse
- A growing number of orthogonal ontologies
 - **Anatomical system** *eg: digestive system, stomach*
 - **Cell type** *eg: granulocyte, T-cell*
 - **Developmental stage** *eg: embryo, adult*
 - **Pathology** *eg: leukemia, normal*
 - **Pooling** *eg: pooled donor, pooled tissue*
 - **Treatment**
 - **Array platform**
- Appropriately detailed set of terms
 - data-driven approach to determining the level of granularity required



Vocab Viewer

File

Anatomical Site

- Lymphoreticular
- Multisystem
- Developmental
- Alimentary
 - liver and biliary system
 - oral cavity
 - vermiform appendix
 - pharynx
 - large intestine
 - pancreas
 - stomach
 - small intestine
 - jejunum
 - duodenum
 - ileum
 - mesentery
 - omentum
 - oesophagus
- Musculoskeletal
- Genital
- Endocrine
- Urinary
- Dermal
- Respiratory
- Cardiovascular
- Haematological
- Nervous
 - auditory apparatus
 - spinal cord
 - ganglion
 - brain
 - meninges
 - diencephalon
 - tract
 - ventricular system
 - midbrain
 - cerebellum
 - medulla oblongata
 - cerebrospinal fluid
 - cerebrum

Development Stage

- Adult
- Toddler
- Fetus
 - 16 weeks
 - 13 weeks
 - 24 weeks
 - 23 weeks
 - 32 weeks
 - 18 weeks
 - 14 weeks
 - 22 weeks
 - 12 weeks
 - 9 weeks
 - 17 weeks
 - 28 weeks
 - 10 weeks
 - 20 weeks
 - 26 weeks
 - 19 weeks
- Infant
- Child
- Embryo
 - 4 cells
 - 8 weeks
 - 3 weeks
 - 6 weeks
 - 16 cells
 - 8 cells
 - 7 weeks
 - 4 weeks

Pathology

- Congenital anomalies
- Other disorders
- Normal
- Metabolic disorders
- Degenerative disorders
- Neoplasia
 - benign
 - malignant
 - astrocytoma
 - medulloblastoma
 - leukaemia
 - sarcoma
 - melanoma
 - chorioepithelioma
 - neuroblastoma
 - erythremia
 - seminoma
 - oligodendroglioma
 - lymphoma
 - hypernephroma
 - Wilms
 - myeloma
 - retinoblastoma
 - papilloma
 - carcinoma
 - teratocarcinoma
 - papillary serous carcinoma
 - choriocarcinoma
 - carcinoma in situ
 - squamous carcinoma
 - adenocarcinoma
 - medullary
 - glioblastoma
 - tumour
- Auto-immune disorders
- Inflammatory disorders

Cell Type

- goblet cell
- Müller cell
- hepatoblast
- pinealocyte
- A cell
- pronormoblast
- fibrocyte
- oligodendrocyte
- granulosa cell
- spermatocyte
- odontoblast
- telogial cell
- prospmatogonium
- pericyte
- plasma cell
- spermatid
- myelocyte
- peptic cell
- zymogenic cell
- mammotroph
- lemmal cell
- eosinophil
- oxyntic cell|parietal cell
- lymphocytic stem cell
- hyalocyte
- Merkel cell
- neuron
- haemocytoblast
- epithelium
- C cell
- Claudius cell
- monoblast
- normoblast
- ependymal cell
- lymphoid stem cell
- Purkinje cell
- histiocyte
- myeloid stem cell
- adipocyte

Untangled ontologies

- eVOC is an untangled ontology
 - Pure hierarchical tree structure
 - Easy to construct
 - Easy to visualise
 - Easy to query



Querying

- **Query returns:**

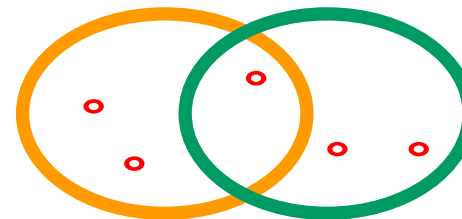
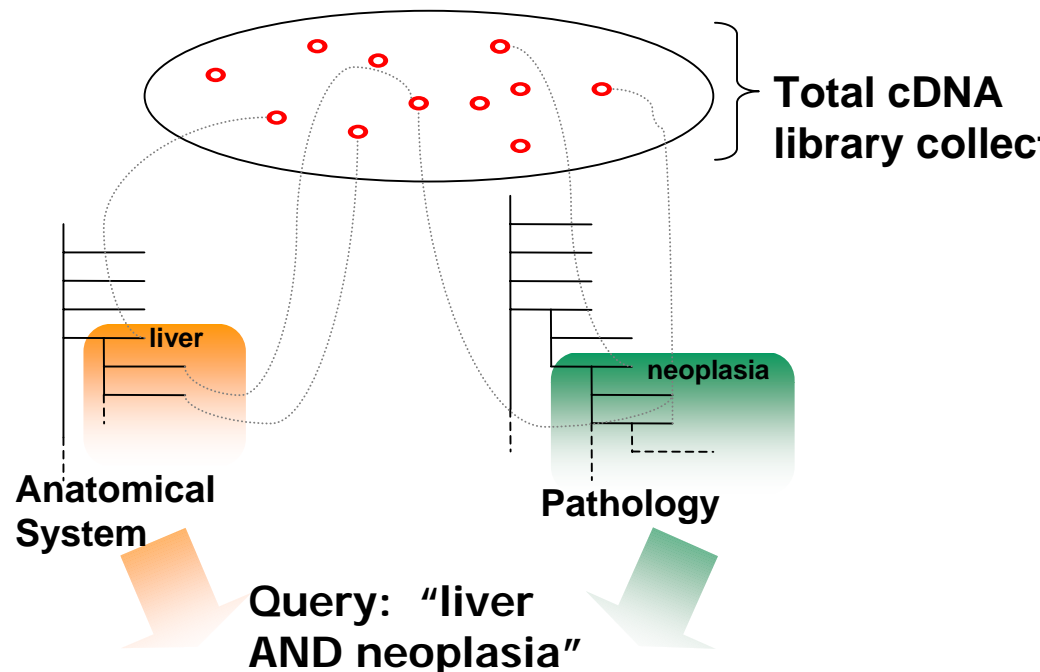
- The node with which that term is associated
- Nodes in the entire subtree rooted at that node.
- By following the mappings from the ontology nodes to public databases (eg: cDNA libraries) a query is translated to a set of **cDNA libraries**

Which have associated **ESTs**

Which can be linked to **genes**

Which have **GO-annotated functions**

Which are mapped to **pathways**



Result: Intersection of libraries mapped to "liver" and to "neoplasia"

Ensembl EnsMart Genome Browser (MartView) - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Mail

Address http://www.ensembl.org/Multi/martview/2R6ee14a01.mart

Transcripts per gene:

 Single

 Multiple

Entries with a 5' UTR

 Only

 Excluded

Entries with a 3' UTR

 Only

 Excluded

GENE ONTOLOGY:

Evidence code for mapping: ANY

Molecular Function
 E.g. GO:0008083 or growth factor activity

Biological Process
 E.g. GO:0008219 or cell death

Cellular Component
 E.g. GO:0005623 or cell

EXPRESSION:

Entries expressed in:

 eGenetics/SANBI data

 GNF data (AFFY-HG-U95)

Anatomical Site: brain

Development Stage: embryo

Pathology:

Cell Type:

MULTI SPECIES COMPARISONS:

Homologous genes

Entries associated with homologous mouse genes

 Only

 Excluded

Entries associated with homologous

 Only

anatomical_site - Microsoft Internet Explorer

- Anatomical Site
 - Alimentary
 - Cardiovascular
 - Dermal
 - Developmental
 - Endocrine
 - Genital
 - Haematological
 - Lymphoreticular
 - Multisystem
 - Musculoskeletal
 - Nervous
 - cerebellum
 - cerebrospinal fluid
 - cerebrum
 - diencephalon
 - medulla oblongata
 - meninges
 - midbrain
 - pons
 - tract
 - ventricular system
- auditory apparatus
- brain
- ganglion
- olfactory apparatus
- peripheral nerve
- spinal cord
- sympathetic chain
- visual apparatus

Prioritising disease gene candidates

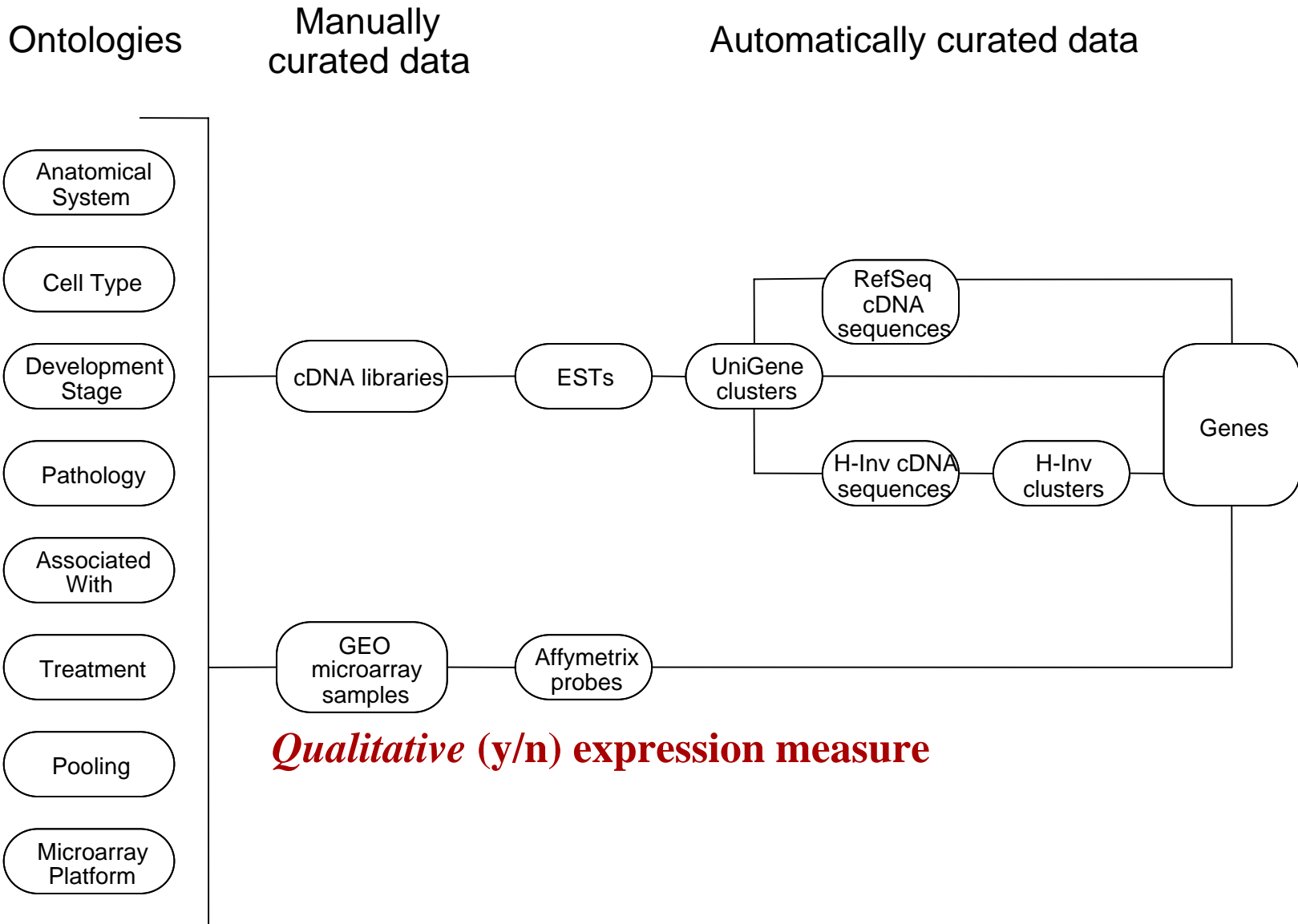
Within a region identified by classical genetic techniques:

- Examine the expression of genes in the region as a pointer to potential candidates.

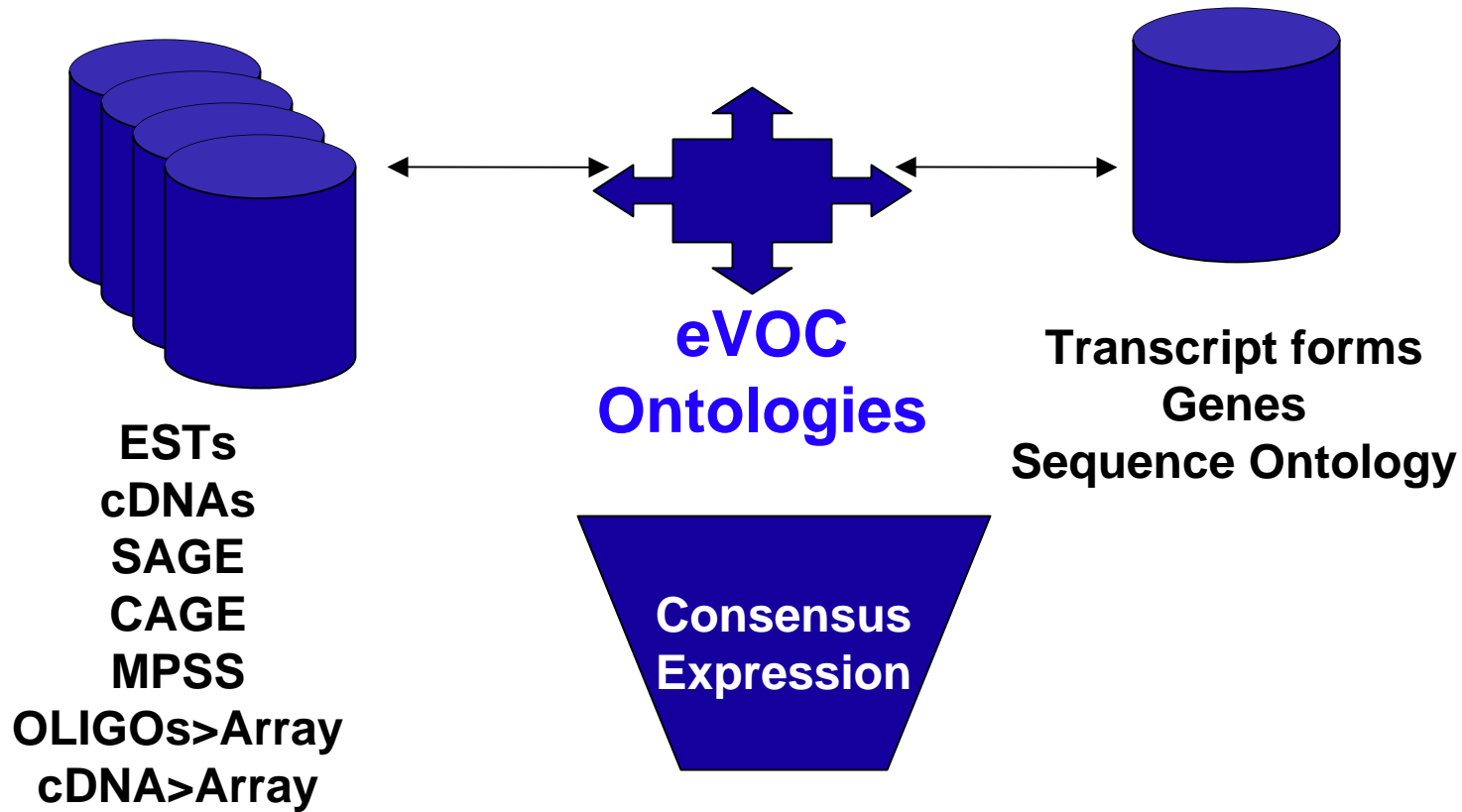
Tested on the region implicated in Retinitis pigmentosa before the identification of the causative gene....

Total known genes	21 787
Genes between 8q11.23 and 8q12.1	38
Genes with expression in retina	(RP1) 7

Cross-platform mining and validation



Evolving Expression Integration



eVOC domains

- Human, Mouse, Rat and model organisms
- Mouse human sequence ontology
- Fantom3
- Hinv and Hinv disease edition
- ENSEMBL/ENSmart
- Alternate Splice Database (EBI)
- Ludwig Institute for Cancer Research
- UniProt (SwissProt)
- Pathogens



Give me all genes that...

- Are common to mouse and human **AND**
- Expressed in (ONLY IN) heart during the first heart beat **AND**
- Are also found in neoplasia **OR**
- Are elevated in expression during stress **AND**
- Have an alternate exon expressed more often in neoplasia **AND**
- Are co-localised on the genome **OR**
- Are found associated in the text with heart disease **AND**
- Are protein kinases or involved in signal transduction **AND**
- Have orthologs in worm and also fly

Monthly release of eVOC mappings

Latest data release: eVoke data v2.3

Ontologies:

Electric Genetics - Oct 3, 2004

cDNA libraries:

GenBank release 143 - Aug 15, 2004

GenBank daily updates - Oct 2, 2004

GEO microarray samples:

Gene Expression Omnibus - Nov 20, 2003

EST sequences:

GenBank release 143 - Aug 15, 2004

GenBank daily updates - Oct 2, 2004

Affymetrix probes

Gene Expression Omnibus - Nov 20, 2003

Kelso J, Visagie J, Theiler G, Christoffels A, Bardien S, Smedley D, Otgaar D, Greyling G, Jongeneel CV, McCarthy MI, Hide T, Hide W.

eVOC: A Controlled Vocabulary for Unifying Gene Expression Data. Genome Research. 2003 Jun;13(6):1222-30

RefSeq cDNA sequences:

UniGene Build # 175 - Sep 30, 2004

H-Inv cDNA sequences:

H-InvDB (Version_1.7) - Jul 1, 2004

UniGene clusters:

UniGene Build # 175 - Sep 30, 2004

H-Inv clusters:

H-InvDB (Version_1.7) - Jul 1, 2004

Genes:

LocusLink - Oct 3, 2004





eVOC Ontology . Org



- [Home](#)
- [Browse eVOC](#)
- [Download eVOC Ontologies](#)
- [Download eVOC Data](#)
- [Publications](#)
- [Contributors](#)
- [Citing Evoc](#)
- [Software?](#)
- [Contact Us](#)
- [Links](#)

BrowseEvoc

Select Ontology: Select Data Type:

- Anatomical System
- ▶ developmental
 - ▶ anatomical site
 - ▶ cardiovascular
 - ▶ respiratory
 - ▶ haematological
 - ▶ lymphoid tissue
 - ▼ alimentary
 - ▶ oral
 - ▶ pharynx
 - esophagus
 - stomach
 - ▶ gut
 - mesentery
 - omentum
 - peritoneum
 - liver and biliary system
 - **pancreas**
 - ▶ genitourinary
 - ▶ endocrine
 - ▶ musculoskeletal
 - ▶ dermal
 - ▶ nervous
 - unclassifiable

[Cancel Data Download](#)

Current Term: pancreas

- NIH_MGC_120
- Barstead pancreas HPLRB1
- Human Fetal Pancreas 1A
- Human Fetal Pancreas 1B
- Human Pancreas
- Human pancreas expression library
- Human pancreatic cancer (CWallrapp)
- Human pancreatic cancer cell line PANC1
- Human pancreatic cancer cell line Patu 8988t
- Human pancreatic tumor
- NCI_CGAP_Pan1
- NCI_CGAP_Pan3
- NIH_MGC_39
- NIH_MGC_42
- NIH_MGC_70
- NIH_MGC_78
- Pancreas II
- Pancreas I
- Pancreas tumor III
- Pancreas tumor II
- Pancreas tumor I
- Pancreas tumor, subtracted (abundant clones)
- Strataqene pancreas (#937208)

Search Website

Recent Site Updates

- [Evoc Related Links](#) (October 21)
- [Evoc Contact Details](#) (October 21)
- [Evoc Contributors](#) (October 21)
- [Evoc Publications](#) (October 21)
- [Citing Evoc](#) (October 21)
- [Evoc Ontology Dot Org](#) (October 21)
- [Short Recent Changes](#) (October 21)

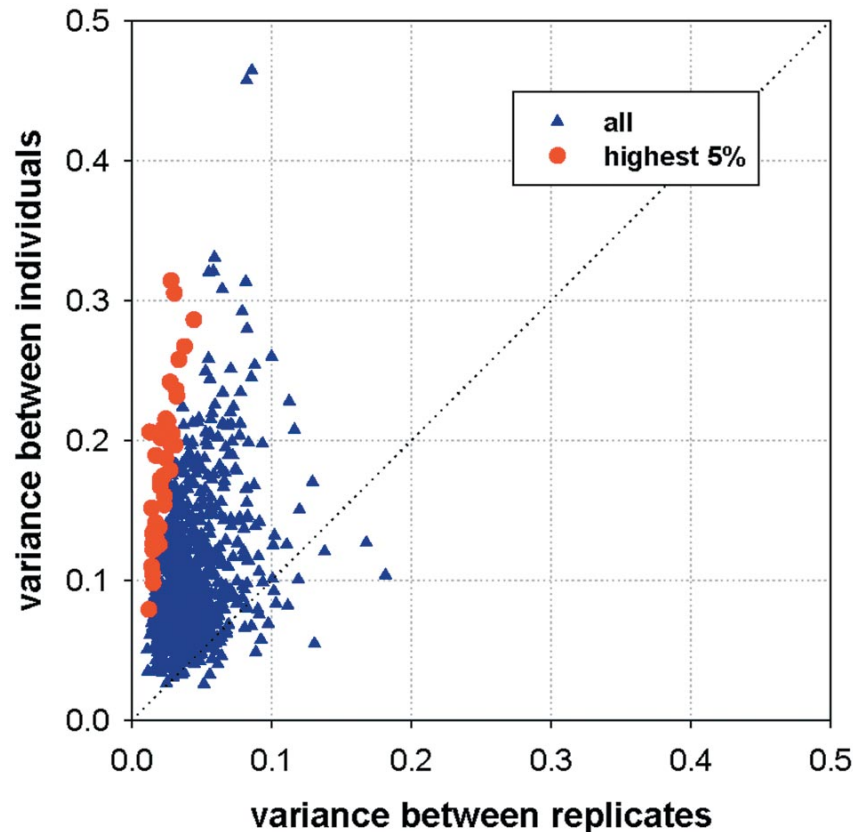
Sampling and Data

- ARRAY : False negative:
 - under represent gene expression
- EST libraries: True positives
 - Very broad representation of expression experiments
 - Absolute measure
- TEXT: False positives
 - Lots of results, which has true meaning?



Variation of gene expression for 813 genes in 35 individuals

Cheung et al Feb 2003, Nature Genetics



**Varied by a factor of 2.4 or greater
With highest by a factor of 17
No account for oligo location on gene
or isoform exon cassette**

Scatter plot of variance in expression level between individuals and between replicates for 813 genes. The genes with the highest variance ratio (top 5%) are highlighted in red. The dotted line indicates a variance ratio of 1.0.

CT-Antigens

- Show testis upregulation signature
- Large number of isoforms and recent family members
- Half are X-Linked



GNF

Gene Expression Atlas: Text Query - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://expression.gnf.org/cgi-bin/index.cgi#Q> Search Print

Home Bookmarks The Mozilla Organiza... Latest Builds

13 r

2 hits to U95A

Hs_285115
IL13RA1
interleukin 13 receptor, alpha 1

359_at

Probe	Average Difference Value
tin	58.5
jun	41.5
ih	45
es	64
us	83
la	10.5
us	112.5
un	127.6
nd	83.5
od	488.2
is	156.7
ss	153
ka	83.5
nd	131
ia	14.5
er	72.4
ts	127.6
ry	214
8E	348
9S	243
us	57.5
6G	67
nd	158
sa	130
ng	351.5
us	206.5
eh	45
nd	155
ey	216
er	139
er	204.6
rk	63.5
ec	206.6
vt	79
ly	46.5
62	90.5
4	44
di	37
li	55.5
3b	152
5B	175
h2	128.5
10	142.5
60	142.5
22	75
es	89
5U	77.5

Med#3

Median

1
359_at Original Affymetrix Annotation
Y10659 cds HSIL13RA H.sapiens IL-13Ra mRNA | GNF

Current Annotation Sun Nov 2 05:19:26 PST 2000

Locuslink id 3597
name interleukin 13 receptor, alpha 1
symbol IL13RA1
product interleukin 13 receptor, alpha 1 precursor
sumfunc none
summary Summary: The protein encoded by this gene is the alpha subunit of the interleukin 13 receptor. This subunit forms a receptor complex with IL4 receptor alpha, a subunit of the IL4 receptor. This subunit serves as a ligand-binding subunit of the IL13 receptor, and is also a component of IL4 receptors. This protein has been shown to bind tyrosine kinase TYK2, and thus may mediate signaling processes that lead to the activation of STAT3 and STAT6 induced by IL13 and IL4.
go receptor activity;cell surface receptor ligand activity;transduction;interleukin-13 receptor complex;membrane;plasma membrane;hematopoietin/interleukin-3 receptor activity;cytokine receptor activity;

extannot none
phenotype none
map none
maplink none
nm NM_001560 Notes
IL13 receptor alpha-1 chain; CD213a1 antigen
db MIM:601551

Done



an
academic web site
of the
Weizmann Institute of Science



Search GeneNote by

Choose data type: **Currently, all results are displayed using GeneNote's gene-centric view.**

[GeneNote Home](#)

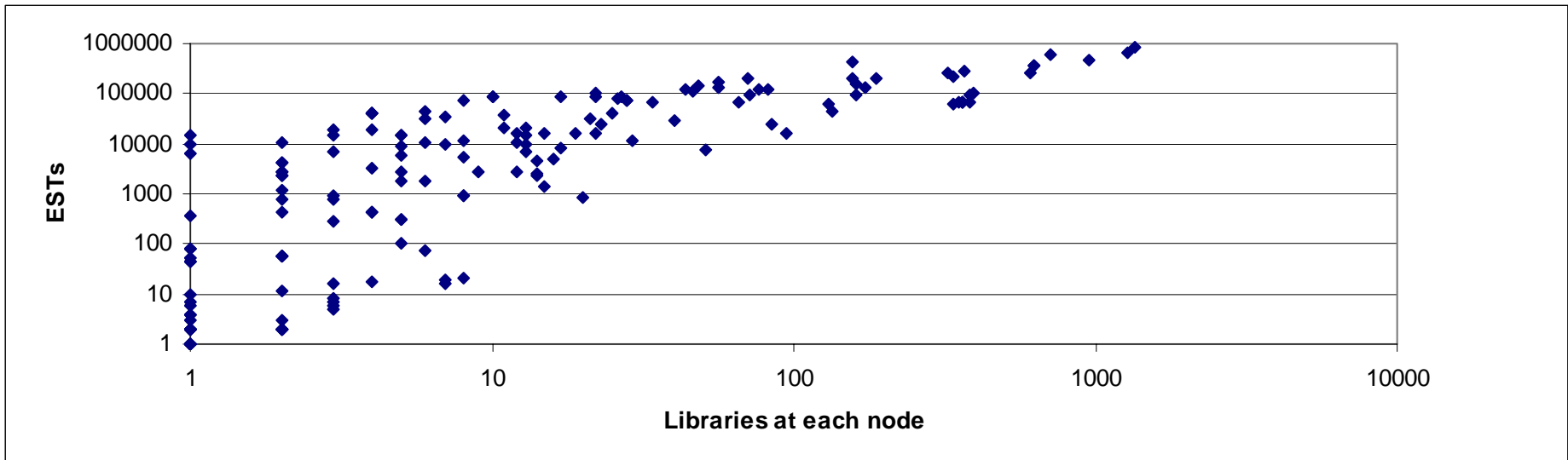
Ensembl ID	ENSG00000131724
LocusLink ID	3597
Chromosomal Location	Chromosome: X, Orientation: +, Start: 114847163, End: 114914104
Aggregate Expression	<p>GeneNote - expression arrays 1 probe-set</p> <p>Normalized intensity</p> <p>White rectangles above bars show min-max range for duplicates</p>

Tissue distribution considered

Expression Category	Tissue Distribution Definition
1	Testis Restricted.
2	≤ 2 non – gametogenic tissues.
3	≤ 6 non – gametogenic tissues; Differentially Expressed
4	≥ 6 non – gametogenic tissues; Ubiquitously Expressed.

Table 1 showing the expression categories devised by LICR⁴ based on tissue distribution of the CT genes.

Information density map of eVOC ontology anatomy term node levels

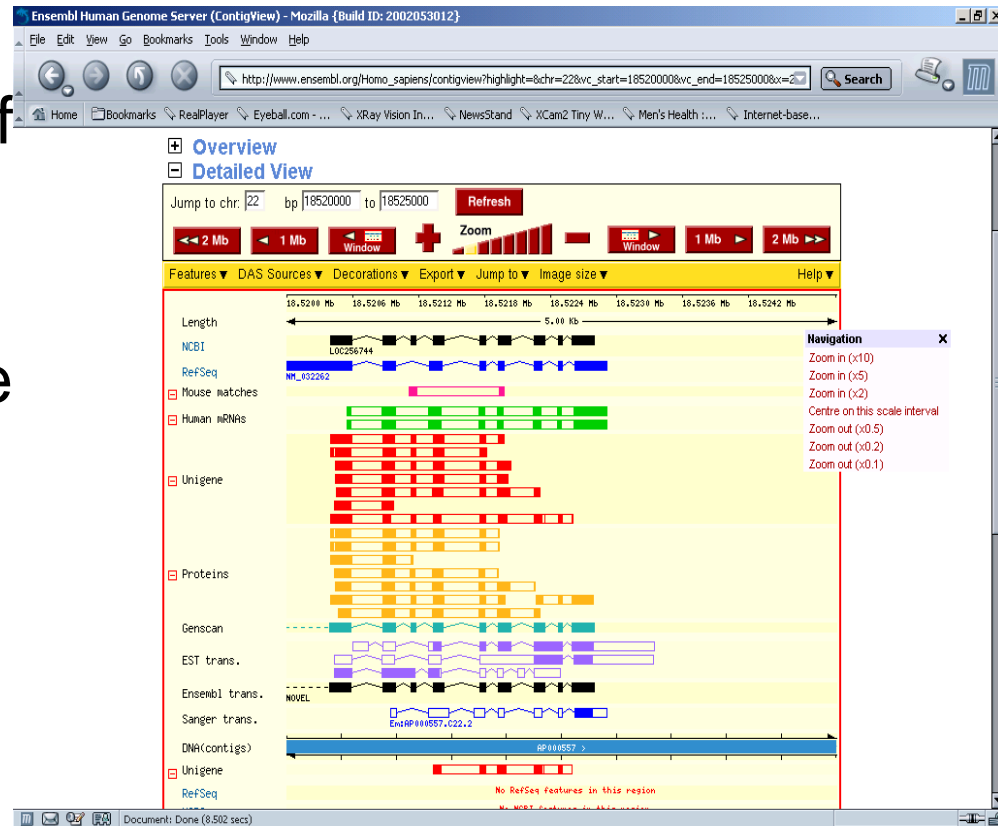


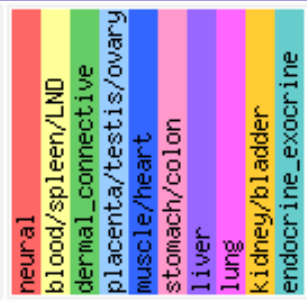
72 term nodes have > 10 libraries



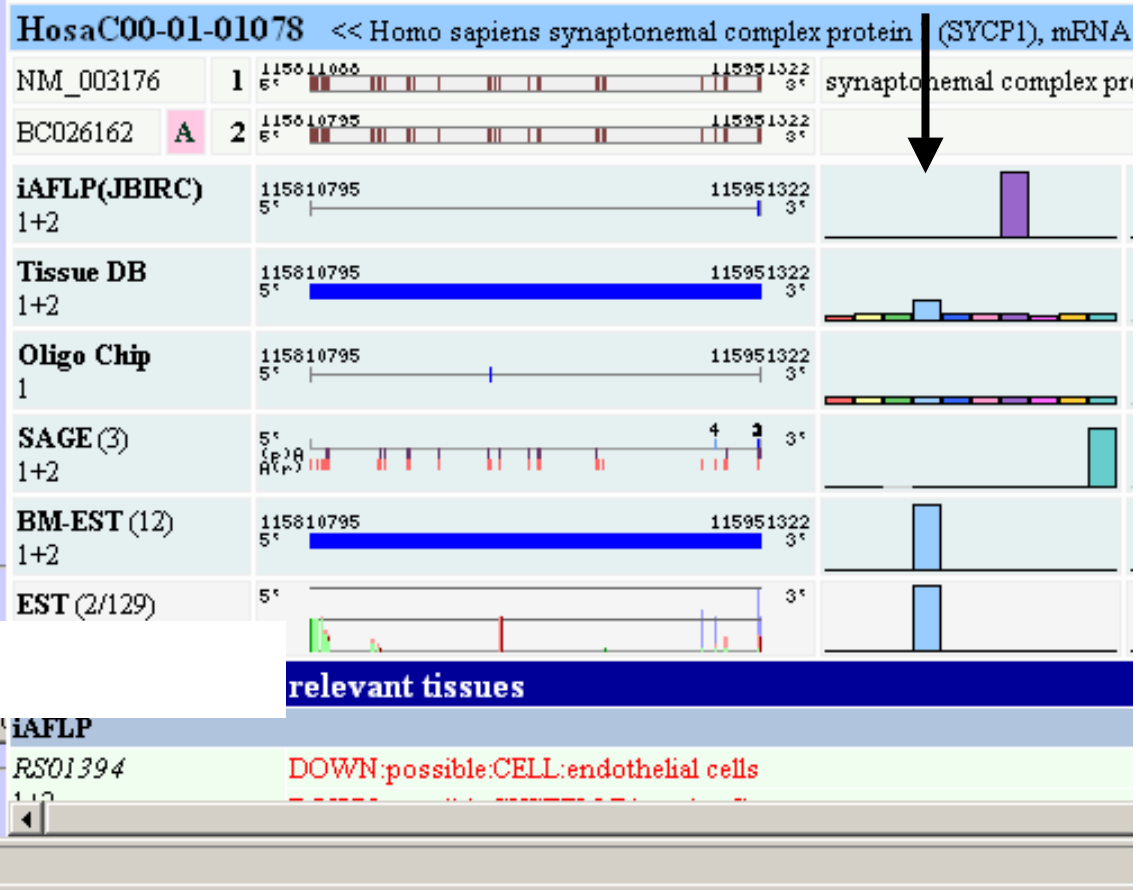
Can we use an evidence based approach to expression characterisation?

- Genome Annotation requires several types of evidence to support a 'solved' gene model
- Exon boundaries require support from transcript, protein, prediction, cross-species comparison





SYCP-1



One of many examples of CT-Antigens that show inconsistency between expression distribution between SAGE, Chip, PCR-RT and EST library

Direct => **H-Inv cluster ID:**

Search

Pattern Search

▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
0	0	0	100	0	0	0	0	0	0
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Select Plat Form

<input checked="" type="checkbox"/> iAFLP(JBIRC)	<input checked="" type="checkbox"/> iAFLP(Osaka)
<input checked="" type="checkbox"/> HuGEIndex	<input checked="" type="checkbox"/> Aburatani
<input checked="" type="checkbox"/> Oligo Chip	<input checked="" type="checkbox"/> SAGE
<input checked="" type="checkbox"/> CNRS	<input checked="" type="checkbox"/> MPSS
<input checked="" type="checkbox"/> BM-EST	

Correlation cos pearson

9291 items

NO.	H-Inv cluster ID	ACCESSION NO	expression	Correlation	experiment
1	01-0000012	AF015592,AB003698,AF005209		0.95	BM-EST
2	01-0000274	AF316543,AB061675,A,Y009398		0.78	Aburatani
3				1.00	BM-EST
4	01-0000304	AF064102,AF064103,AF000367,BC038979,AF122013		0.75	Aburatani
5				0.97	Aburatani
6	01-0000404	A,Y044866,BC051817,AF060509		0.86	BM-EST
7	01-0000445	AF076464,AF076463,M33478		0.73	HuGEIndex
8	01-0000453	AF116571,AF083105,AF098915		0.81	BM-EST
9	01-0000490	AF087651,AF091501,AF119569		1.00	BM-EST
10	01-0000516	AF105202,AK074957		0.91	BM-EST
11	01-0000520	AF120274,AF109401		1.00	BM-EST
12	01-0000557	AF127138,AF186380,AF236117		0.75	BM-EST
13	01-0000614	AF161426,AF379631		0.75	BM-EST
14	01-0000631	AF299350,AF331033,AF176813,AF271058		0.83	Aburatani
15				1.00	BM-EST
16	01-0000698	AF273043,D67035		0.93	Aburatani
17				0.98	Aburatani
18				1.00	BM-EST

FANTOM consortium

- Large scale data generation mouse and human transcriptome project
- Network analysis
 - promotor and alternate transcription start site
 - polyadenylation, and exon isoform cassette usage
- Transcriptional regulation at fine resolution and the resulting promotor association
- Cap Analysis of Gene Expression (CAGE)
- Developmental libraries of the mouse,
- Insight into cancer biology.



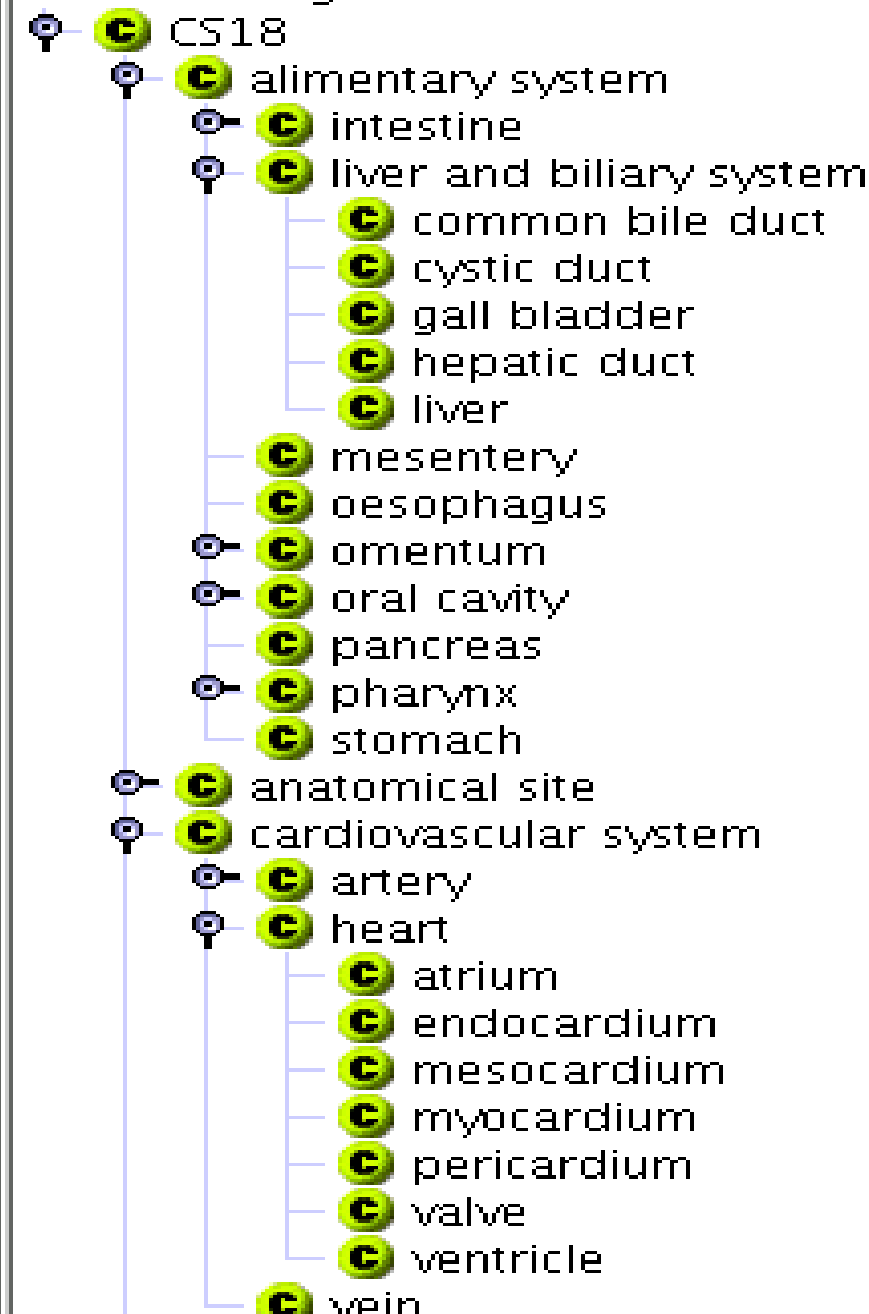
Mouse eVOC objectives

- develop comparable mouse (and human) developmental anatomical ontologies
- populate with gene expression data
- cross-species gene expression query to allow comparison:
 - human adult and mouse adult
 - mouse developmental stages
 - human developmental stages
 - mouse and human developmental stages



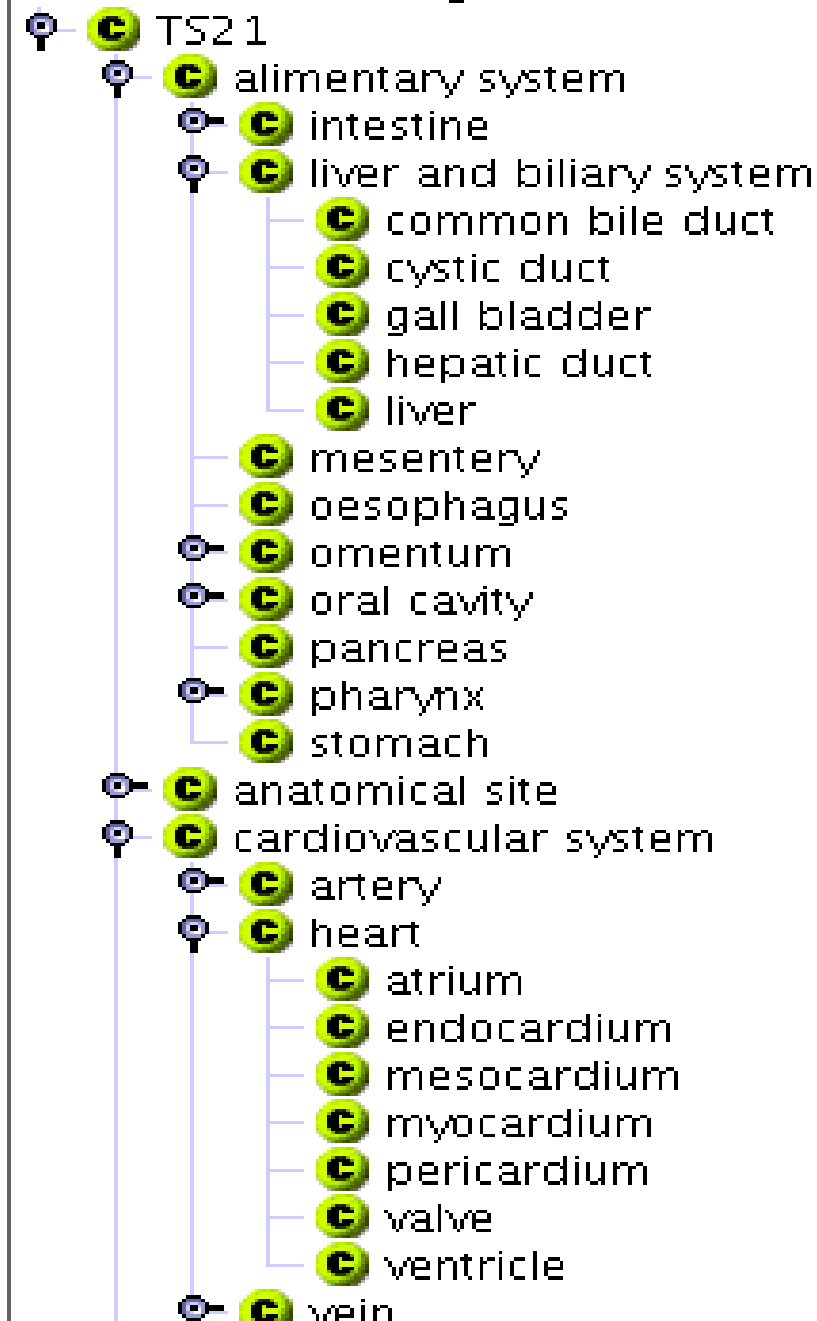
eVOC Ontologies

eVOC Ontologies

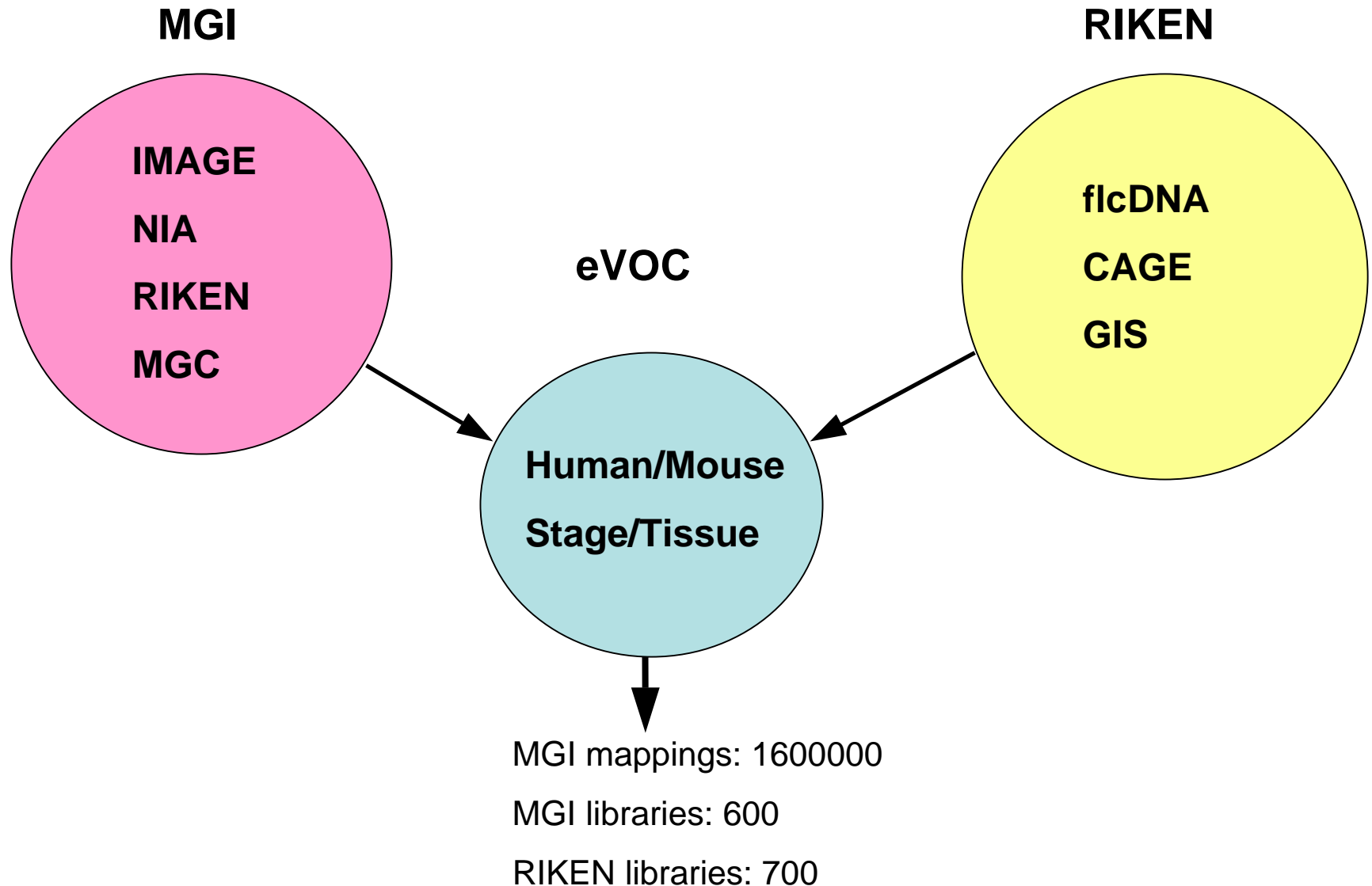


eVOC Mouse Ontologies

eVOC Mouse Ontologies



Mouse/Human Integration via eVOC



Binning expression

- All human data all expression grouped
 - into embryonic (Theiler stages 1-4)
 - fetal (Theiler stages 5 to 26)
 - adult (newborn onwards)
 - merge all stem- and germ cell expression into a fourth class



Alternative Splicing

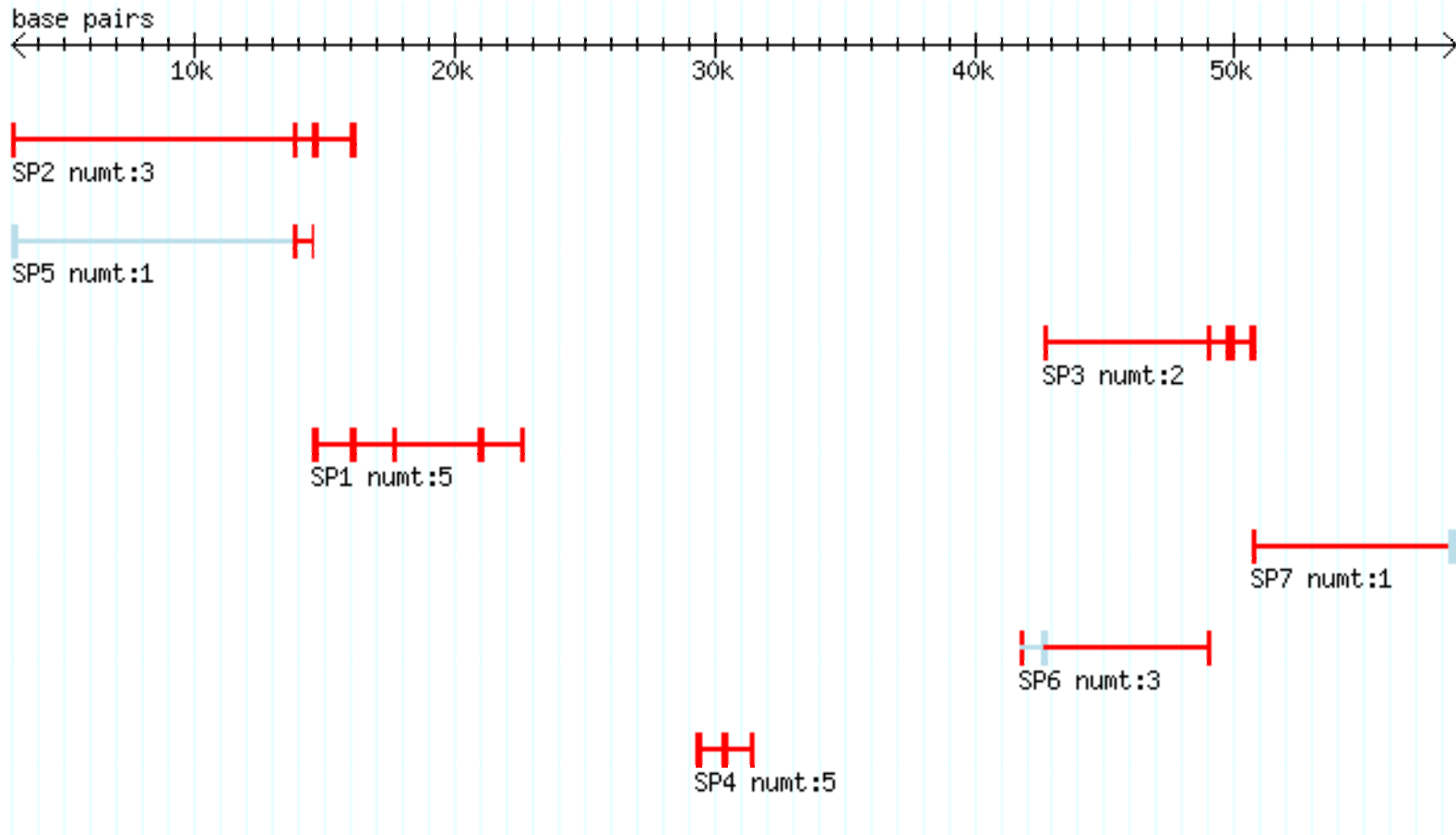
BRDT-NY
Fetal testis

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

BRDT
Adult



Alternative Splicing of BRDT



Splice Graph Queries

- Dependencies on expression state?
- Splice rate and events compared:
 - Rapidly changing tissues (gut, 1-year old brain)
 - Development
 - Cancer
- Different domains / epitopes?



Ontologies provide an underlying tool that can be broadly applied in disparate systems

- Cross platform and cross ontology mapping yields rapid results
- Approach requires data driven mapping for each instance
- Broader adoption of standardised ontologies will promote disparate data integration



Win Hide
Oliver Hoffman
Adele Kruger
Janet Kelso
Simon Cross
Allan Powell
Vladimir Bajic
Hong Pan



**Institute for
Infocomm Research**

Liza Groenewald



electric genetics
corporation



welcometrust



**NATIONAL
BIOINFORMATICS
NETWORK**



Cape Town, South Africa

