

Creating a database and developing a website:

ReproMine

Summary

- I. ReProMine's project
- II. ReProMine's database
- III. ReProMine's website

I. ReproMine's project

- Gather raw data on reproductive diseases and re-normalize for comparison
- Analyze these data to identify individual mRNAs related conditions or stages of disease development
- Create a list of molecular pathways, constituting each set of disease and development associated with mRNA transcript identified in the previous step
- Create a database to facilitate data access to users

Why generate this database?

- Microarray technologies allow tens of thousands of genes to be analysed simultaneously
- Microarrays have been applied extensively to developmental and reproductive disease
- There is a realisation that more biological meaning can be obtained from microarray data when it is analysed at the level of pathways and gene sets
- There is also a realisation that concordant data from several separate microarray studies is more reliable than data from a single study

An example of bringing microarray data from several different studies together: OncoPrint

Welcome to OncoPrint 4.3 Research Edition

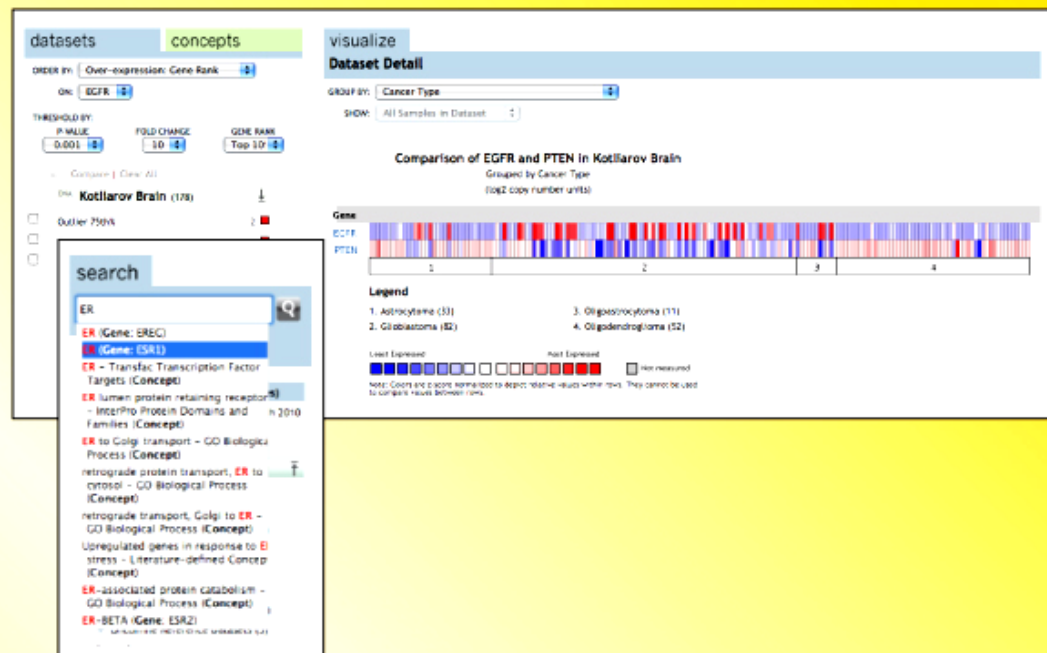
Upgrade to the *OncoPrint Research Premium Edition*: advanced features for cutting-edge cancer research

multi-gene search

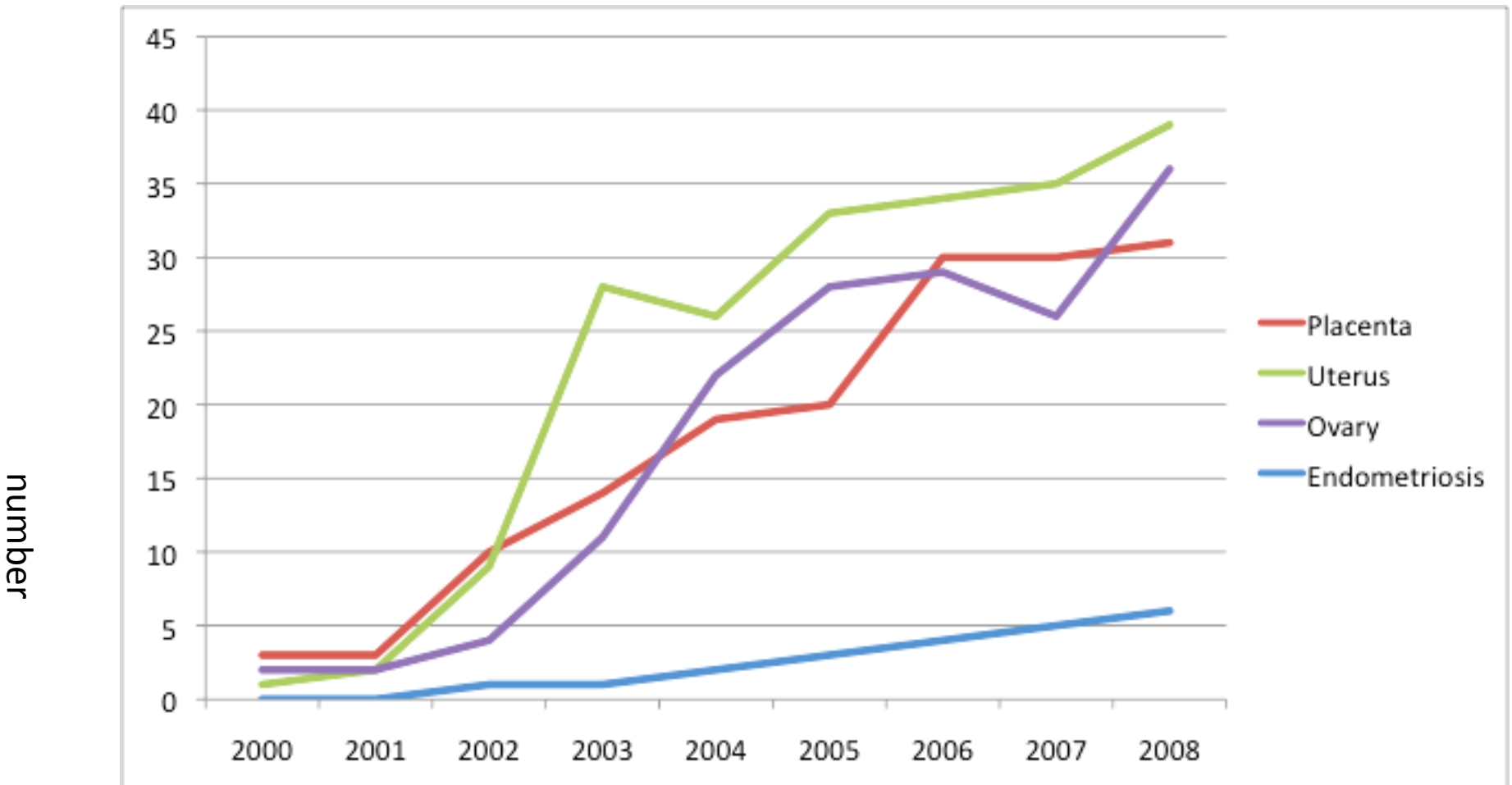
Searching has never been easier.

With OncoPrint Research Premium Edition any number of genes, a concept, and countless filters can be added to your search in order to compare the expression of multiple genes at once and understand the relationships between multiple genes in a sample.

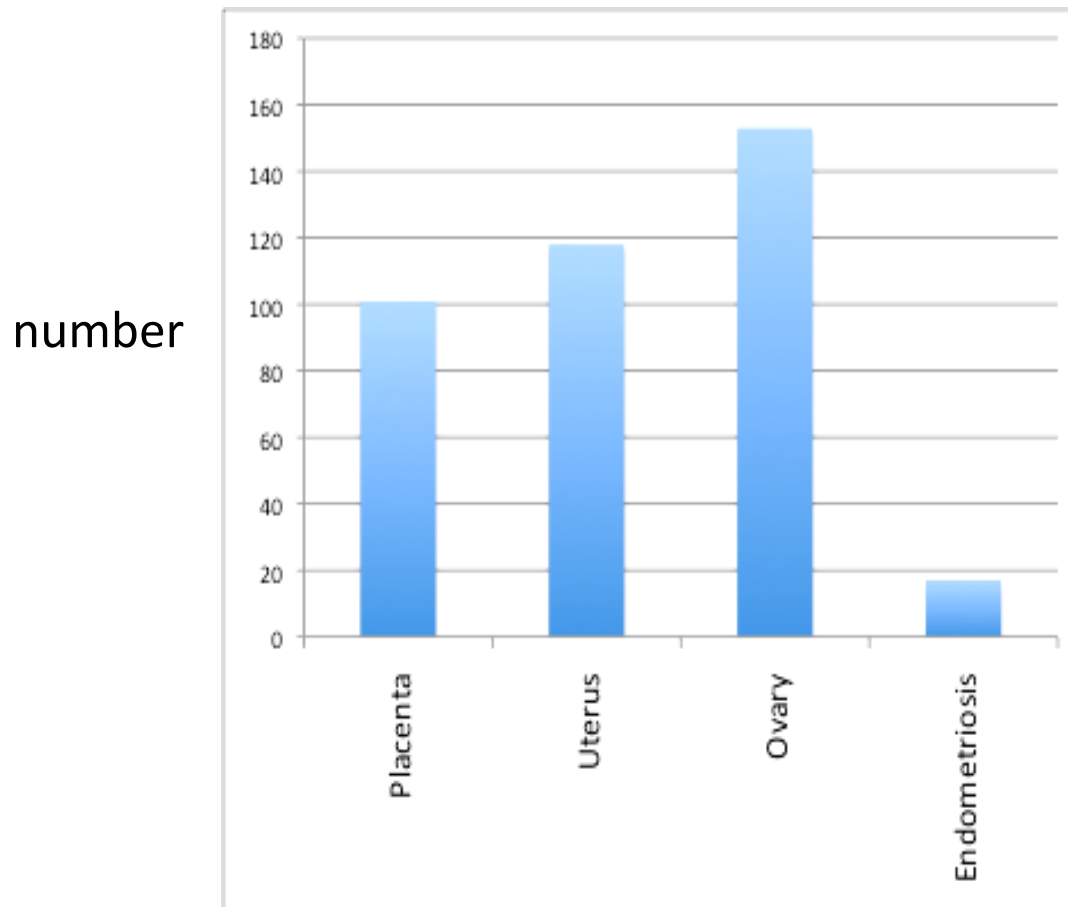
With the smart search box, when you type two characters, a list of options appears with terms from genes to tissue types.



What data is there to put into this database? Many microarray studies of reproduction and development have been performed over the last ten years



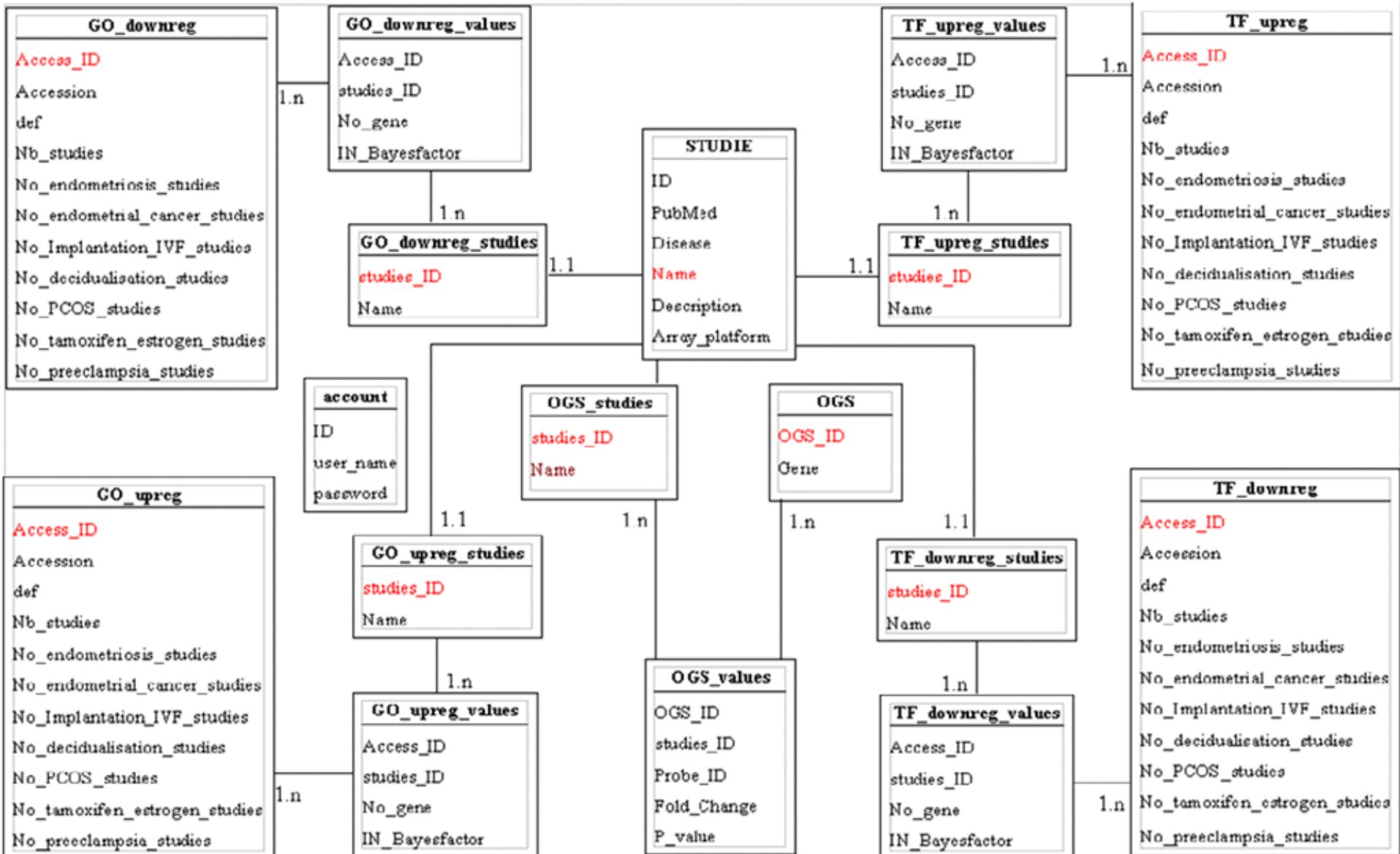
Based on these studies, there are now significant amounts of microarray data available for several reproductive diseases



II. ReProMine's database

- **Demand Analysis:** definition of user needs, simple database, easily accessible, which they can easily find the data that the user would need, to contain all relevant information
- **Solution Design:** Conceptual Data Model (CDM)
- **Physical design:** realization of the database
- Database Client server

II. ReproMine's database



III. ReproMine's website

Tools:

WampServer (Windows Apache MySQL PHP) platform for web development Windows

Php: interpreted language, mainly used for producing dynamic web pages, it allows to enter data in the database, create the website, and the interaction between the website and database

HTML: markup language, for writing the hypertext, to format the pages to include multimedia resources such as images, entry forms, and programmable elements

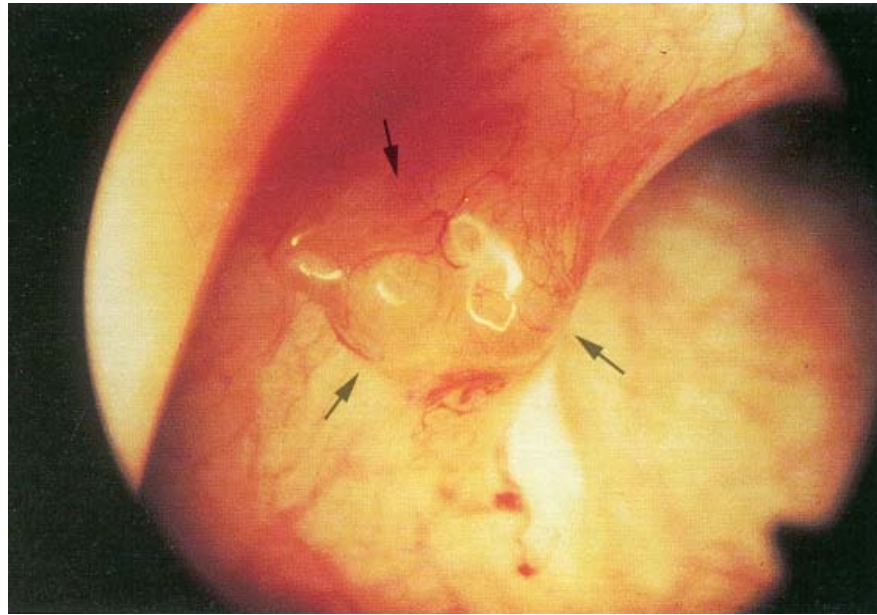
Mysql management system database, which allows sorting, filtering and data processing

An example of using this database:
the disease endometriosis

What is endometriosis?

- A disease where endometrium (the lining of the uterus) gets into the wrong place
- It causes severe pain and also infertility
- It is thought to require several processes including cell proliferation and an immune response

An example of an endometriotic lesion seen in a women's abdomen through an endoscope



When you compare endometriosis to normal endometrium lining the uterus, what genes appear to be regulated in common in 7 studies?

Query

Download

- Endometriosis meta-analysis endometriosis vs endometrium
- Hever endometriosis vs endometrium
- Hever Reanalysis 500 endometriosis vs endometrium
- Hever Reanalysis 1057 endometriosis vs endometrium
- Hull endometriosis vs endometrium
- Hull reanalysis 500 endometriosis vs endometrium
- Hull reanalysis 682 endometriosis vs endometrium
- Borghese endometriosis vs endometrium
- Eyster endometriosis vs endometrium
- Eyster reanalysis 500 endometriosis vs endometrium
- Eyster reanalysis 751 endometriosis vs endometrium
- Zafrales endometriosis vs endometrium
- Burney endometriosis vs control endometrium



Study	Gene	Pvalue	Fold change
Hever endometriosis vs endometrium		1356209_at	2.04
Hever endometriosis vs endometrium		1580399_at	2.34
Hever endometriosis vs endometrium		1580755_at	6.85
Hull endometriosis vs endometrium		200605_at	-2.21597
Hull endometriosis vs endometrium		200974_at	2.0333493
Hull endometriosis vs endometrium		201020_s_at	3.17875724
Hever endometriosis vs endometrium		201110_s_at	3.56456074
Hull endometriosis vs endometrium		201111_s_at	13.03
Hull endometriosis vs endometrium		201131_s_at	-4.8466797
Hull endometriosis vs endometrium		201141_s_at	2.76284684
Hull endometriosis vs endometrium		201302_at	-2.5360409
Hull endometriosis vs endometrium		201326_at	-3.4884398
Hever endometriosis vs endometrium		201389_at	2.04
Hull endometriosis vs endometrium		201428_at	-6.8395079
Hever endometriosis vs endometrium		201450_s_at	6.14
Hull endometriosis vs endometrium		201451_s_at	2.931861981
Hever endometriosis vs endometrium		201468_s_at	6.84
Hull endometriosis vs endometrium		201497_s_at	8.73204195
Hever endometriosis vs endometrium		201521_s_at	6.31
Hull endometriosis vs endometrium		201563_at	-3.6929152
Hull endometriosis vs endometrium		201612_s_at	2.3278692
Hever endometriosis vs endometrium		201621_s_at	2.79
Hull endometriosis vs endometrium		201650_at	-2.0796791
Hever endometriosis vs endometrium		201666_at	6.04
Hull endometriosis vs endometrium		201689_s_at	-3.852376
Hever endometriosis vs endometrium		201741_s_at	3.13
Hull endometriosis vs endometrium		201785_at	2.88382023
Hull endometriosis vs endometrium		201792_s_at	2.78847744
Hull endometriosis vs endometrium		201811_s_at	2.01834268
Hull endometriosis vs endometrium		201829_s_at	-4.4248991
Hull endometriosis vs endometrium		201843_s_at	2.03873003
Hull endometriosis vs endometrium		201926_s_at	-3.4232276
Hull endometriosis vs endometrium		201957_at	2.1282204
Hull endometriosis vs endometrium		201971_s_at	-2.8976095
Hull endometriosis vs endometrium		202003_s_at	3.32251255
Hull endometriosis vs endometrium		202073_s_at	2.17442142
Hull endometriosis vs endometrium		202177_at	2.5640124
Hull endometriosis vs endometrium		202191_s_at	2.62268361



Only 8 genes in common between all studies!

endometrium lining the uterus, what biologically coherent gene sets (from the Gene Ontology database) appear to be regulated in common in 7 studies?

Query

Download

- Endometriosis meta-analysis endometriosis vs endometrium
- Hever endometriosis vs endometrium
- Hever Reanalysis 500 endometriosis vs endometrium
- Hever Reanalysis 1057 endometriosis vs endometrium
- Hull endometriosis vs endometrium
- Hull reanalysis 500 endometriosis vs endometrium
- Hull reanalysis 682 endometriosis vs endometrium
- Borghese endometriosis vs endometrium
- Eyster endometriosis vs endometrium
- Eyster reanalysis 500 endometriosis vs endometrium
- Eyster reanalysis 751 endometriosis vs endometrium
- Zafrakas endometriosis vs endometrium
- Burney endometriosis vs control endometrium



GO	Accession	GO term	N	Biological
GO:0002642	GO:0002642	immune response	31	46.25
GO:0002642	GO:0002642	response to toxic substance	27	42.41
GO:0002642	GO:0002642	defense response	26	40.62
GO:0002642	GO:0002642	response to heat shock or hypoxia	23	35.38
GO:0002642	GO:0002642	response to oxidative stress stimulus	22	34.09
GO:0002642	GO:0002642	response to hypoxia	21	32.80
GO:0002642	GO:0002642	response to oxidative stress stimulus	20	31.51
GO:0002642	GO:0002642	response to hypoxia	19	30.22
GO:0002642	GO:0002642	response to oxidative stress stimulus	18	28.93
GO:0002642	GO:0002642	response to oxidative stress stimulus	17	27.64
GO:0002642	GO:0002642	response to oxidative stress stimulus	16	26.35
GO:0002642	GO:0002642	response to oxidative stress stimulus	15	25.06
GO:0002642	GO:0002642	response to oxidative stress stimulus	14	23.77
GO:0002642	GO:0002642	response to oxidative stress stimulus	13	22.48
GO:0002642	GO:0002642	response to oxidative stress stimulus	12	21.19
GO:0002642	GO:0002642	response to oxidative stress stimulus	11	20.00
GO:0002642	GO:0002642	response to oxidative stress stimulus	10	18.81
GO:0002642	GO:0002642	response to oxidative stress stimulus	9	17.62
GO:0002642	GO:0002642	response to oxidative stress stimulus	8	16.43
GO:0002642	GO:0002642	response to oxidative stress stimulus	7	15.24
GO:0002642	GO:0002642	response to oxidative stress stimulus	6	14.05
GO:0002642	GO:0002642	response to oxidative stress stimulus	5	12.86
GO:0002642	GO:0002642	response to oxidative stress stimulus	4	11.67
GO:0002642	GO:0002642	response to oxidative stress stimulus	3	10.48
GO:0002642	GO:0002642	response to oxidative stress stimulus	2	9.29
GO:0002642	GO:0002642	response to oxidative stress stimulus	1	8.10



17 biologically coherent gene sets, especially related to inflammation and immune response

Interestingly, inflammation and immune response are known to play important roles in this disease, and many drugs used to treat this disease act by blocking inflammation and immune response

Key results from the endometriosis search

.Very few genes (8) are regulated in all studies, this is not unusual and reflects different experimental protocols

.Despite this, the genes regulated in all studies share common themes - they are significantly enriched for 17 biologically coherent gene sets (e.g. immune response, which is known to be important for endometriosis)

.The genes regulated in the various studies are also significantly enriched for 21 common transcription factor motifs in their promoters, which appear to be related to what is known about this disease (e.g. cell division, response to growth factors, immune response)

The future

- Repromine will be maintained and expanded
- It will be able to incorporate RNA-seq data as well as data of other types such as genetic data and miRNA data
- It will be continued as a collaboration France-NZ!!!

acknowledgements

I would like to thank for their help, advice and their generosity Cristin PRINT, Enid LAM, Louis RANJARD, Peter TSAI and all people of the Bioinformatics Intitute