# Focused and Customized IT Support of Bioinformatics



# Research Information Lechnology Systems

THE SIDNEY KIMMEL COMPREHENSIVE CANCER CENTER AT JOHNS HOPKINS

Michael Fox, Alla Guseynova, L. Allan Grimm https://www.rig.onc.jhmi.edu

## Introduction:

- o web-based storage for experiment results
- o statistical analysis of microarray data
- o querying of commercial gene chips
- o correlating microarray data with specimen bank data for clinical trials

#### Constraints:

comparing two conditions and you have two replicates in each condition. This tool will

ormalization procedure or by the Agilent error model) The source data needs to be

You can try any file with the same structure from your hard drive.

The output will be an html file with a list of genes, each hyperlinked to NCBI annotation

© Microarray Data Analysis | Design by (313)

The column(s) to combine/compare are: Output is produced at after print columns nev

Loading required package: tools Welcome to Bioconductor Vignettes contain introductor

material. To view, simply type: openVignette() For details on reading vignettes, see the

tab delimited file (you can look at the 🗪

You have entered the files:

Started reading file at Fri Apr 28 14:21:38 2006

Finished reading file at Fri Apr 28 14:21:41 2006

Finished loading libraries at Fri Apr 28 14:22:15 2006

Finished first half at Fri Apr 28 14:22:11 2006

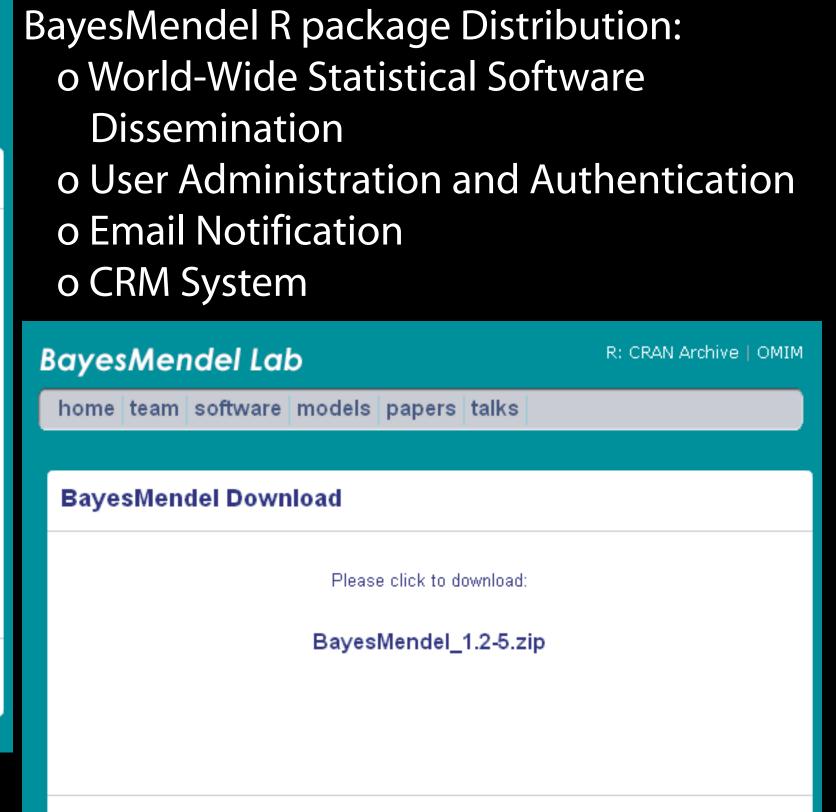
Finished program at Fri Apr 28 14:22:17 2006

File names are:test2.txt

openVignette help page.

- o staff of three
- o customized user requirements
- o users' tight budgets
- o compressed production schedules

BayesMendel Lab R: CRAN Archive   OMIM	Bayes				
home team software models papers talks	o Wo				
	Dis				
BayesMendel Download	o Us				
Download Information					
Please provide us some information about you.					
Name	Bayes				
Title	home				
Institution					
Email  To have the BayesMendel package downloaded, please first agree to the BayesMendel License.	Bayes				
I agree   I decline					
Home   Contact Us   © 2004 Giovanni Parmigiani and Wenyi Wang for the BayesMendel Lab. Last updated 03/05/2006					



LocusLink ID

**MicroArrays Database** 

of Agilent chips

o query of a maintained database

o search by any or combination of

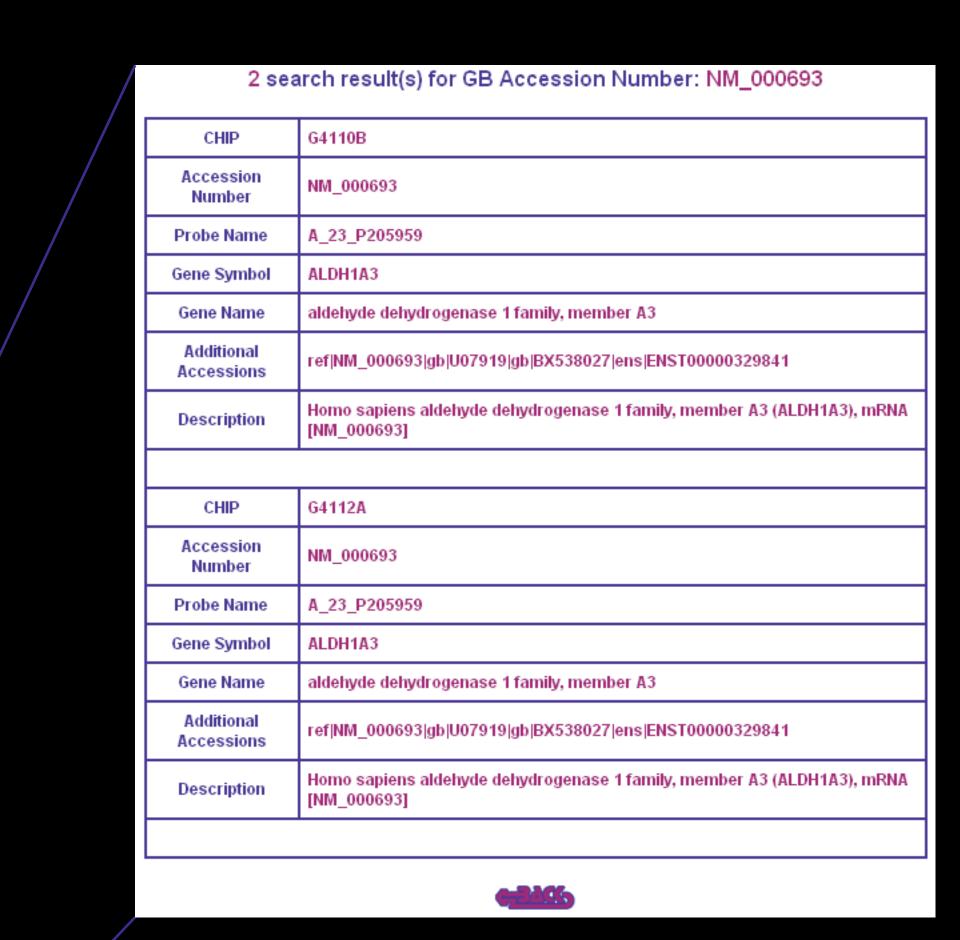
symbol, gene name, Unigene

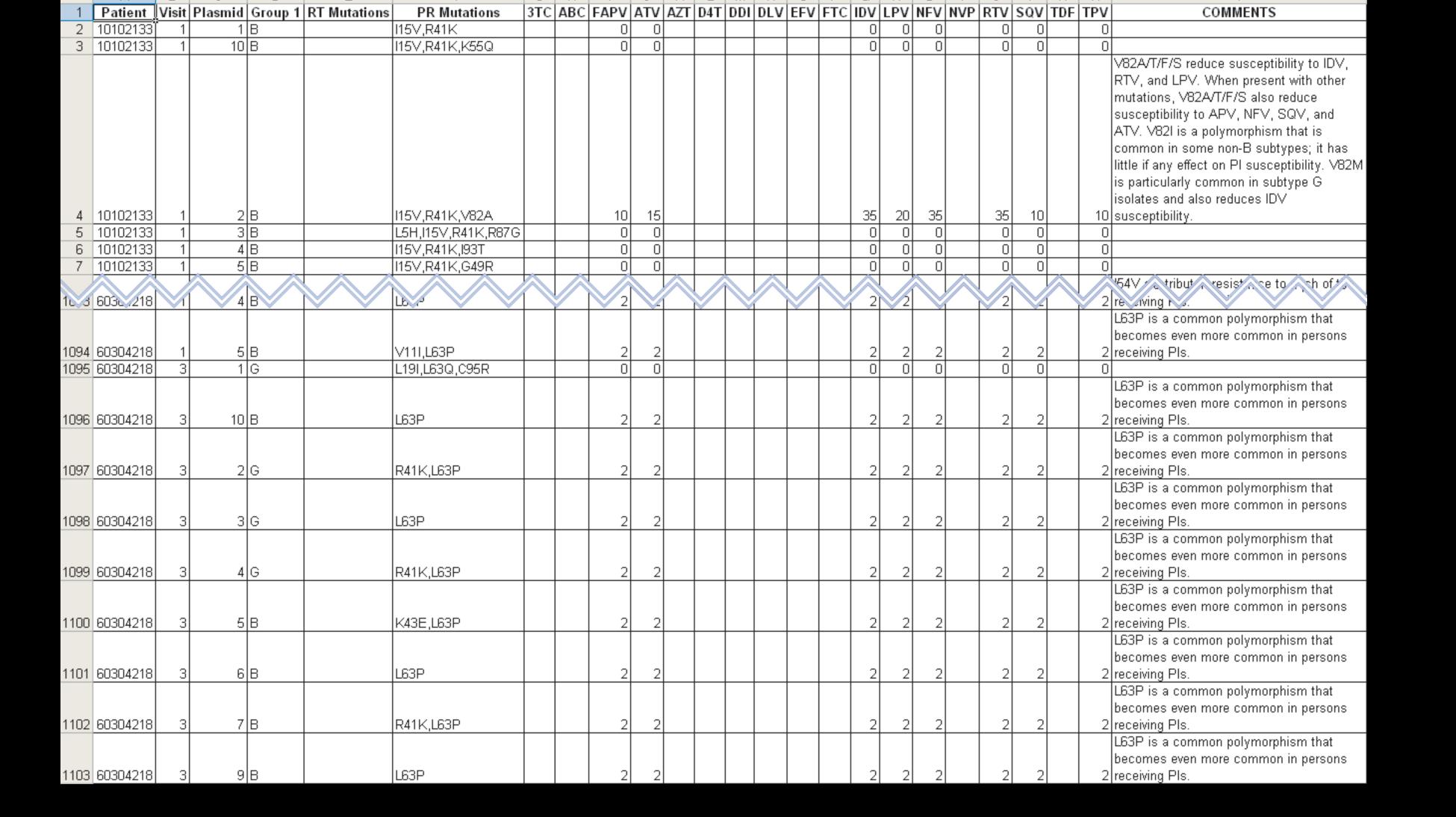
o enables investigators to select

GB accession number, gene

name, or LocusLink ID

the optimal chip for an





frastructure, be sure to check out this new offering

HIV Treatment Websites

The Stanford HIVDB web service (Sierra) allows users to obtain the latest

of resistance to 17 FDA-approved protease and RT inhibitors. However the

We developed the following algorithm to automate the process:

5) perform an XSL transformation on each XML document

7) send users output in dynamically created Excel format

6) insert the parsed results into our Oracle database

Using Cold Fusion program:

provided Perl program requires modification to be used for each sequence.

Stanford HIV Drug Resistance Algorithm scores, comments, and inferred levels

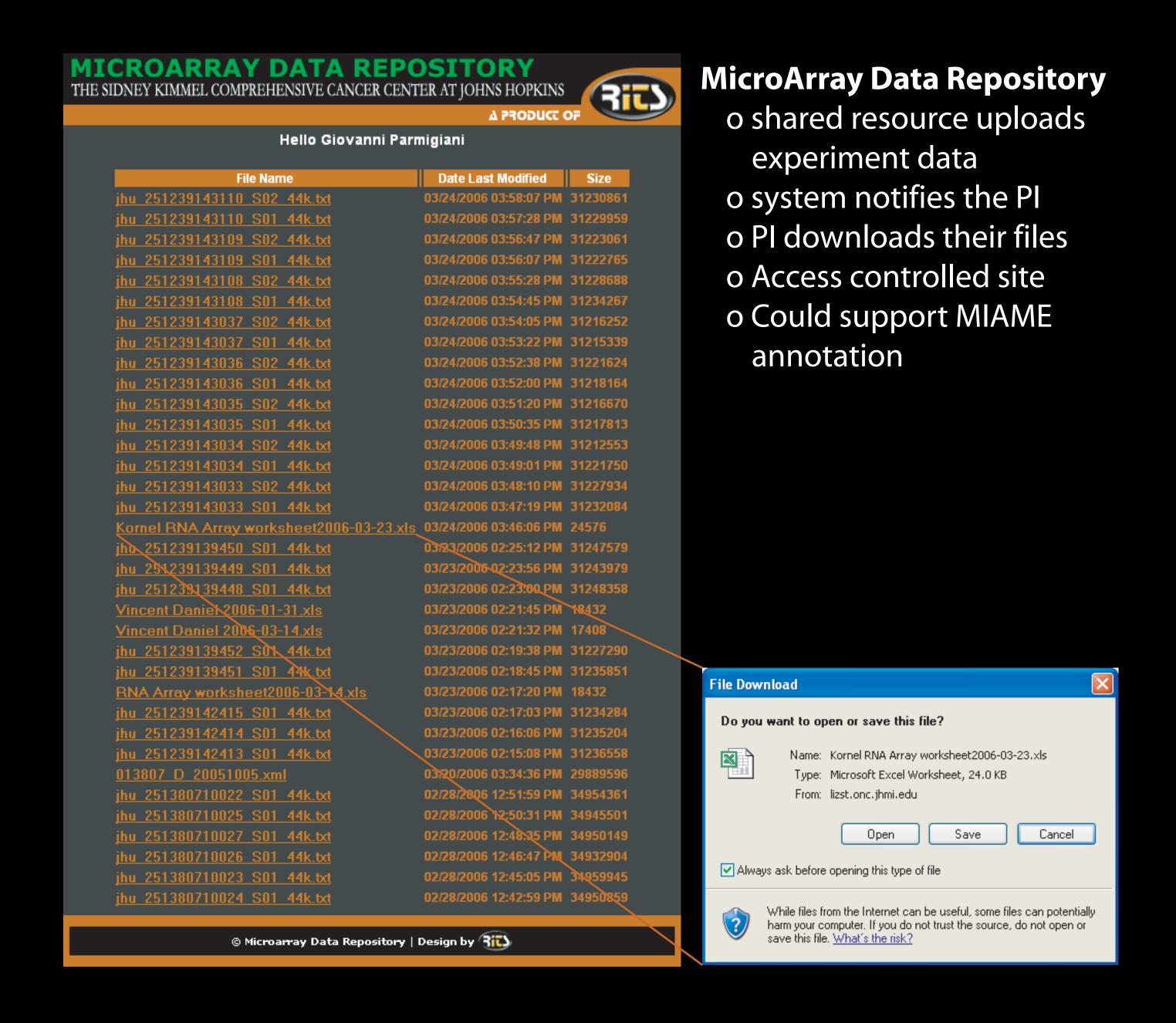
1) read 1100 sequences from files with all provided characteristic such as

3) use the WSDL specification to determine the number and structure of

4) interface with the Stanford algorithm for each sequence to produce an

patient ID, patient visit etc and insert them into the Oracle table

2) extract the user requested number of codons from our database



### MICROARRAY DATA ANALYSIS RICTO Array Data Analysis

o using algorithms from bioinformaticists

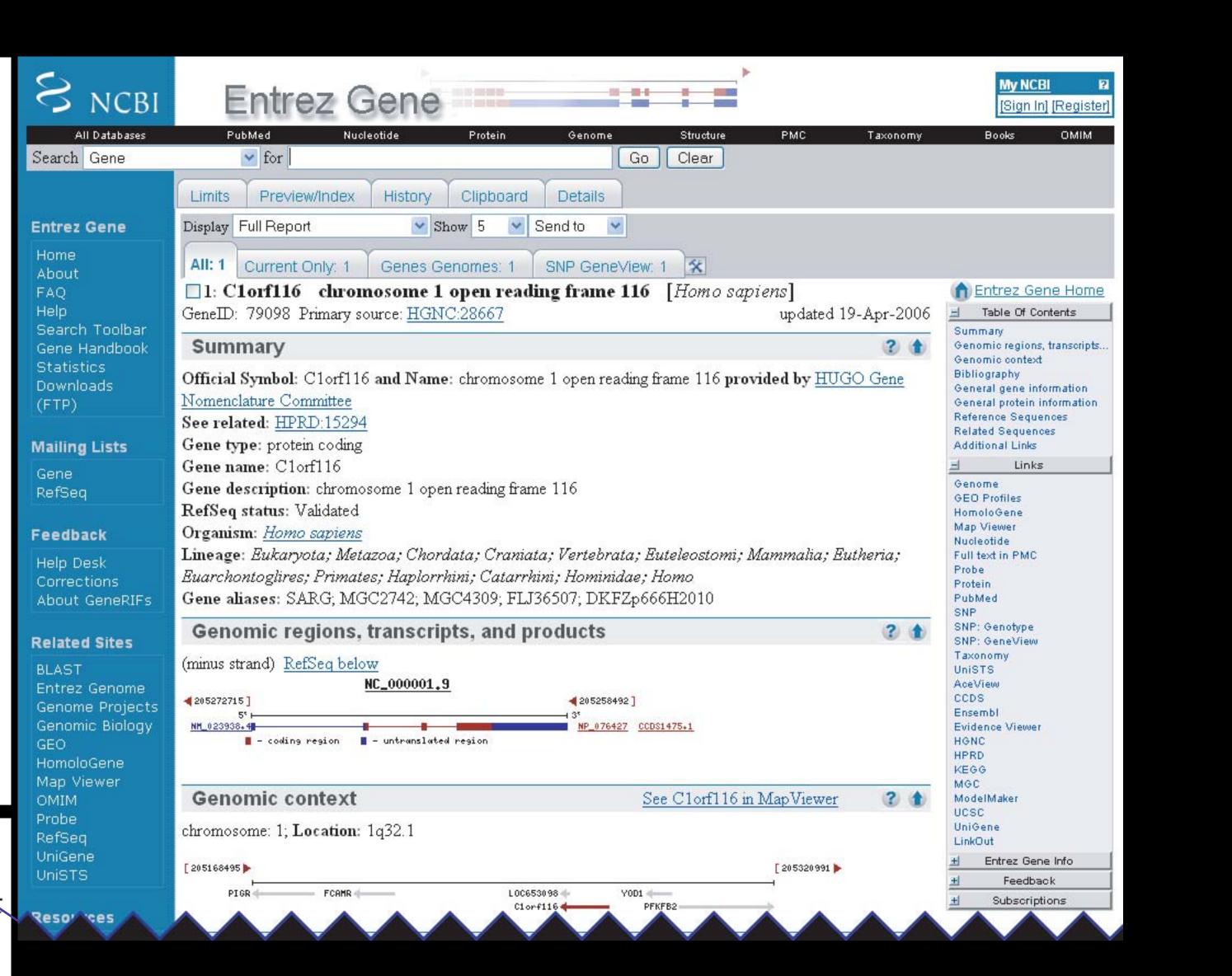
o users can upload experiment data results o yields annotated results

Differentially Expressed Genes

LocusID	Affy ID	Gene symbol	Chromosomal location	Probability	SAM Statistic	Difference in log2(Expression)Tumor-Norma
79098	219476_at	SARG	1q32.1	1	3.03	6.75
7021	214451_at	TFAP2B	бр21-р12	1	2.8	7.51
<u>934</u>	208650_s_at	CD24	6q21	0.999	2.46	6.26
<u>4629</u>	201496_x_at	MYH11	16p13.13-p13.12	0.999	2.4	6.01
<u>3855</u>	209016_s_at	KR <b>Y</b> 7	12q12-q13	0.999	2.38	6.48
8842	204304_s_at	PROM1	4p15.32	0.999	2.3	5.75
79949	219857_at	C10orf81	18q25.3	0.998	2.2	5.15
3872	212236_x_at	KRT17	17q12-q21	0.997	2.23	6.87
<u>55691</u>	208475_at	FRMD4A	10p13	0.995	2.02	4.85
<u>4316</u>	204259_at	MMP7	11q21-q22	0.994	1.98	4.49
10521	213998_s_at	DDX17	22q13.1	0.994	1.96	4.42
<u>12</u>	202376_at	SERPINA3	14q32.1	0.993	1.93	4.4
10205	203780_at	EVA1	11q24	0.992	1.95	4.73
3866	204734_at	KRT15	17q21.2	0.992	2.11	8.32
6337	203453_at	SCNN1A	12p13	0.991	1.9	4.34
<u>4070</u>	202286_s_at	TACSTD2	1p32-p31	0.991	1.91	4.68
<u>934</u>	266_s_at	CD24	6q21	0.989	1.84	4.08
<u>3880</u>	201650_at	KRT19	17q21.2	0.988	1.96	6.28
<u>3854</u>	213680_at	KRT6B	12q12-q13	0.988	1.86	4.49
25984	218963_s_at	KRT23	17q21.2	0.986	1.83	4.46

# Differentially Expressed Genes

LocusID	Affy ID	Gene symbol	Chromosomal location	Probability	SAM Statistic	Difference in log2(Expression)Tumor-Norma
<u>8788</u>	209560_s_at	DLK1	14q32	1	-3.44	-7.88
3326	214359_s_at	HSPCB	6p12	1	-2.85	-6.42
7718	206683_at	ZNF165	6p21.3	1	-2.78	-6.67
<u>4054</u>	219922_s_at	LTBP3	11q12	1	-2.61	-6.17
2192	201787_at	FBLN1	22q13.31	1	-2.5	-5.83
<u>3495</u>	213674_x_at	IGHD	14q32.33	1	-2.64	-7.32
2277	206742_at	FIGF	Xp22.31	1	-2.47	-5.72
<u>5966</u>	206036_s_at	REL	2p13-p12	0.999	-2.36	-5.52
<u>2354</u>	202768_at	FOSB	19q13.32	0.999	-2.38	-6.15
<u>8854</u>	207016_s_at	ALDH1A2	15q21.3	0.999	-2.28	-5.36
4239	212713_at	MFAP4	17p11.2	0.999	-2.27	-5.32
<u>3479</u>	209541_at	IGF1	12q22-q23	0.998	-2.31	-6.12
2192	202995_s_at	FBLN1	22q13.31	0.998	-2.16	-4.99
<u>3337</u>	200664_s_at	DNAJB1	19p13.2	0.998	-2.12	-4.67
<u>25978</u>	202538_s_at	DKFZP5640123	3p12.1	0.998	-2.13	-4.78
<u>1634</u>	211896_s_at	DCN	12q13.2	0.998	-2.18	-5.43
<u>5743</u>	204748_at	PTGS2	1q25.2-q25.3	0.996	-2.19	-6.44
<u> 26577</u>	219295_s_at	PCOLCE2	3q21-q24	0.996	-2.09	-5.36
2326	205666_at	FMO1	1q23-q25	0.996	-2.08	-5.26
1910	206701_x_at	EDNRB	13q22	0.995	-2	-4.47



### Summary of Conclusions:

o every application developed is web-based, supported by a robust database engine o Use of CDEs and ontologies from SPOREs and CaBIG o Rapid Application Development

51 search result(s) for Gene Name: aldehyde dehydrogenase Gene Name | aldehyde dehydrogenase 3 family, member A2 | Please type GB accession number: Homo sapiens aldehyde dehydrogenase 3 family, member A2 (ALDH3A2), mRNA If you don't know GB accession Number you can select Gene Symbol and/or type Gene Name, UniGene ID, LocusLink ID: dehyde dehydrogen Probe Name | A 23 P396086 Gene Name | 2-aminoadipic 6-semialdehyde dehydrogenase Homo sapiens 2-aminoadipic 6-semialdehyde dehydrogenase, mRNA (cDNA clone IMAGE:3684334), complete cds. [BC015096] - © Microarrays Database | Design by 🙃

	3 search result(s) for Gene Symbol: ALDH1A3
CHIP	G4110B
Accession Number	NM_000693
Probe Name	A_23_P205959
Gene Symbol	ALDH1A3
Gene Name	aldehyde dehydrogenase 1 family, member A3
Additional Accessions	ref NM_000693 gb U07919 gb BX538027 ens ENST00000329841
Description	Homo sapiens aldehyde dehydrogenase 1 family, member A3 (ALDH1A3), mRNA [NM_000693]
CHIP	G4112A
Accession Number	NM_000693
Probe Name	A_23_P205959
Gene Symbol	ALDH1A3
Gene Name	aldehyde dehydrogenase 1 family, member A3
Additional Accessions	ref NM_000693 gb U07919 gb BX538027 ens ENST00000329841
Description	Homo sapiens aldehyde dehydrogenase 1 family, member A3 (ALDH1A3), mRNA [NM_000693]
CHIP	G4112A
Accession	BC009245

aldehyde dehydrogenase 1 family, member

IMAGE:2960240), complete cds. [BC009245]

Probe Name | A\_24\_P917280

Gene Symbol ALDH1A3