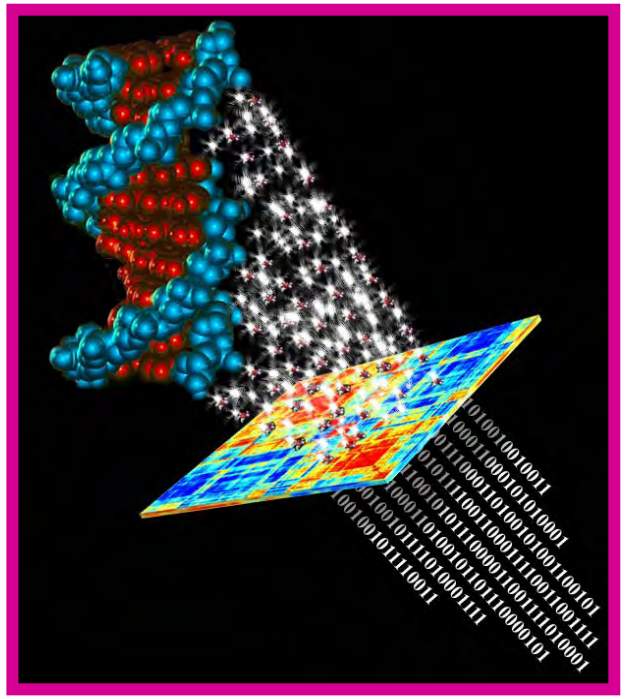
Integrative Pharm-Tox and the NCI-60: Genomics, Proteomics, and Bioinformatics



John N. Weinstein, M.D., Ph.D.
The Genomics & Bioinformatics Group
Laboratory of Molecular Pharmacology
Center for Cancer Research
National Cancer Institute

Center for Cancer Research

Integrative Pharm-Tox and the NCI-60: Genomics, Proteomics, and Bioinformatics

USA TODAY

Home | News | Travel | Money | Sports | Life | Tech | W

Tech » Science & Space ▪ April Holladay ▪ Dan Vergano ▪ Shop for Gadgets

Study: Worldwide carbon dioxide emissions soar

Updated 56m ago | Comments 101 | Recommend 15

E-mail | Save | Print |

By Dan Vergano, USA TODAY

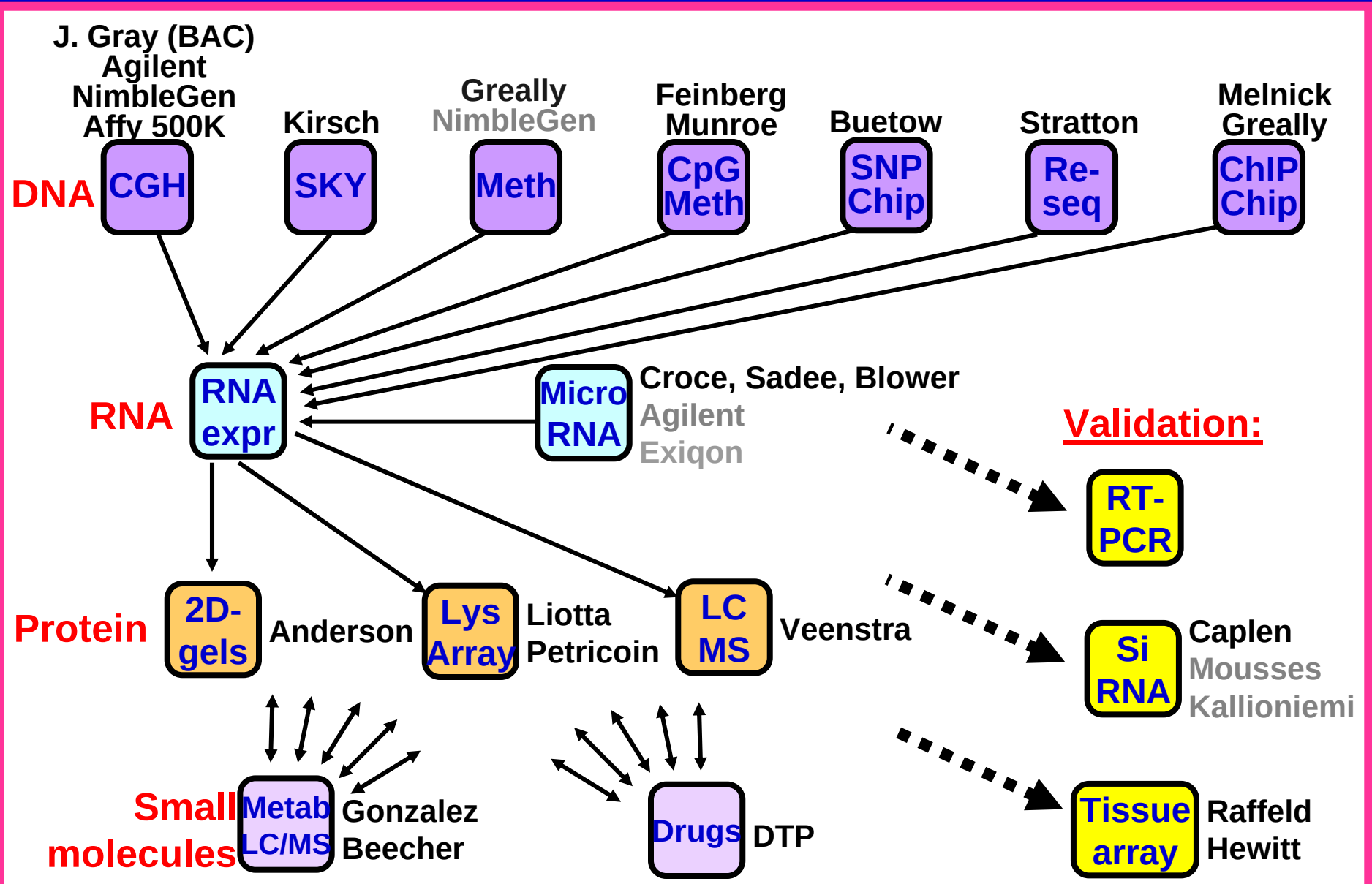
Warnings about global warming may not be dire enough, according to a climate study that describes a runaway-train acceleration of industrial carbon dioxide emissions.

Fueled by rapid growth in coal-reliant China, rates of carbon dioxide emission from industrial sources increased from 2000 to 2004 "at a rate that is over three times the rate during the 1990s," says a report released by the journal *Proceedingsof the National Academy of Sciences.*

Laboratory of Molecular Pharmacology
Center for Cancer Research
National Cancer Institute

Center for Cancer Research

Integrative Molecular Profiling : The NCI-60, etc. (the GBG and Collaborators)



GENOMICS AND BIOINFORMATICS GROUP
LMP, CCR, National Cancer Institute

Integrative Molecular Profiling : The NCI-60, etc. (the GBG and Collaborators)

J. Gray (BAC)
Agilent
NimbleGen
Affy 500K

Kirsch

Greally
NimbleGen

Feinberg
Munroe

Buetow

Stratton

Melnick
Greally

Integromics

(Weinstein & Pommier, Comptes Rendus Biol. 2003; 326:909)

ABC

Gottesman

The Cancer Genome Atlas (TCGA)

GeneLogic

GeneLogic

The Cancer Genetic Markers of Susceptibility Project (CGEMS)

Center for Cancer Research

Genomics And Bioinformatics Group
LMP, CCR, National Cancer Institute

CellMiner: An Integrative Database/Tool for Molecular Profiling Data

Molecular Profile Data (transcript and protein profiles for TP53 in the NCI-60):

CellMiner™ Build:4

Home | Cell Line Metadata | Dataset Metadata | Download Datasets | Query Datasets | Data Intersection | Mutation Data | Credits | Citing CellMiner

Chip/Normalization Method	Probe Id	Gene Symbol	Chromosome	BR:MCF7	BR:MDA_MB_231	BR:HS578T	BR:BT_549	BR:T47D	CNS:SF_268	CNS:SF_295	CNS
Protein: Lysate Array/log2_condensed	TP53_26_GBL00064	TP53	17	0.24	4.18	2.79	3.37	4.24	6.69	4.07	0.4
RNA: cDNA array/log2	236338	TP53	17	0.803183718	0.926968296	0.366112977	1.178144685	0.316370002	1.262515144	0.018824549	0.125

Go to top | Back

Cell Metadata (for the NCI-60 and drug-resistant cells):

CellMiner™ Build:4

Genomics and Bioinformatics Group
LMP, CCR, National Cancer Institute

Home | Cell Line Metadata | Dataset Metadata | Download Datasets | Query Datasets | Data Intersection | Mutation Data | Credits | Citing CellMiner

Cell Name	Panel Name	Age	Sex	Prior Treatment	Tissue of Origin	Histology	Source	Ploidy	p53	mdr	Doubling Time
CNS:SF_268	CNS	24	F	Rad	non-epi	"Glioblastoma, ud"	NA	"2n+, Hyperdiploid (47-57)"	MT	-38	33.1
CNS:SF_295	CNS	67	F	Rad	non-epi	"Glioblastoma, ud"	NA	"5n+/-, Near-pentaploid 115+/- (104-126)"	MT	91	29.5
CNS:SF_539	CNS	34	F	Rad/BCNU/5FU/HU/6MP	non-epi	Glial cell neoplasm	NA	"4n+/-, Near-tetraploid 92+/- (81-103)"	WT	-40	35.4
CNS:SNB_19	CNS	47	M	None	non-epi	"Glioblastoma, ud"	NA	"3n+/-, Near-triploid 69+/- (58-80)"	MT	-41	34.6
CNS:SNB_75	CNS	NA	F	None							
CNS:U251	CNS	75	M	NA							

(Shankavaram, et al., Mol. Cancer Ther., 2007)

Launched in the November, 2006 Issue

AACR CENTENNIAL 1907 - 2007

American Association for Cancer Research

Molecular Cancer Therapeutics Introduces

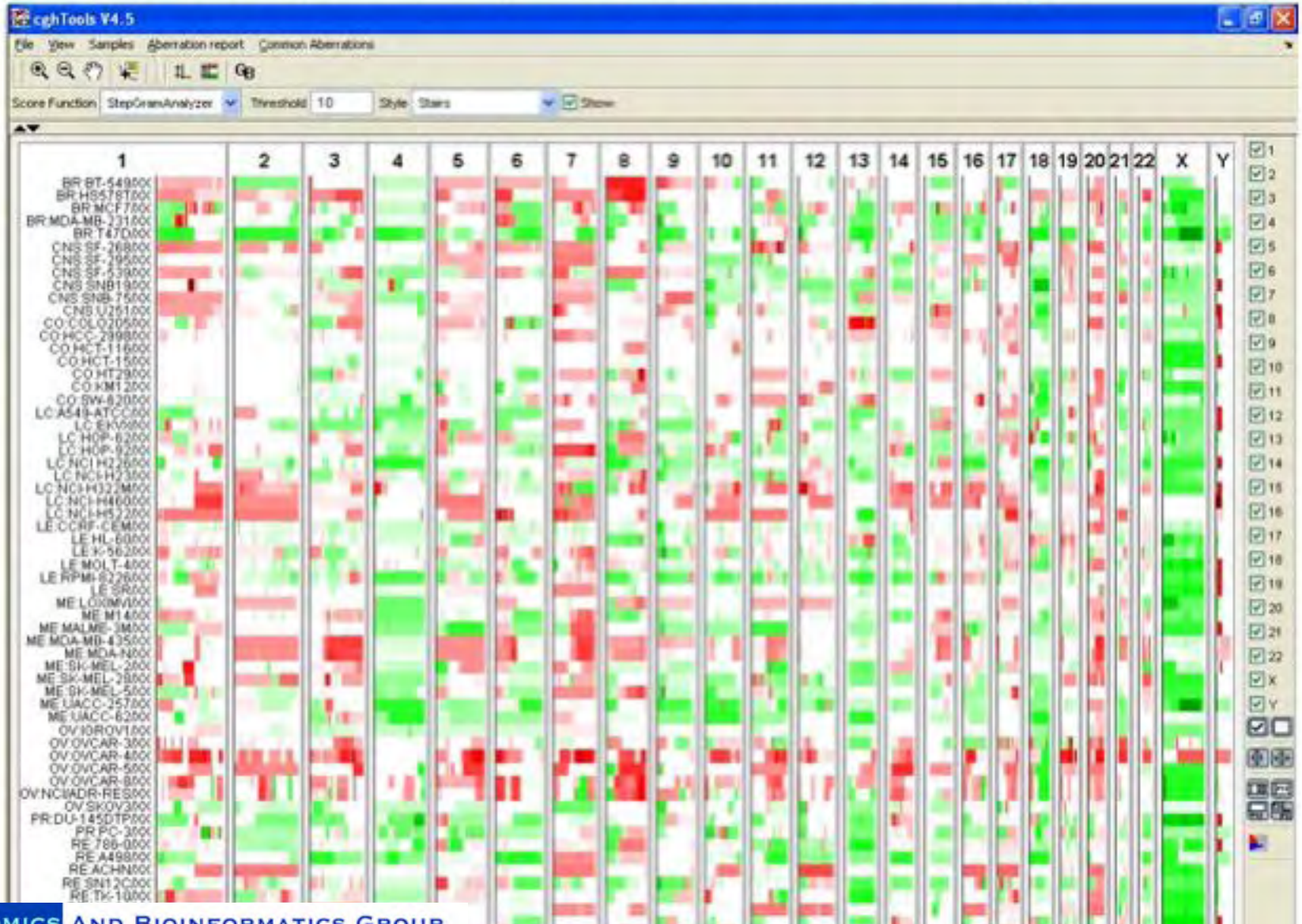
Spotlight on Molecular Profiling

A new series of pharmacogenomic and pharmacoproteomic articles focused on the NCI-60 cell line panel, highlighting the power and promise of "integromic" molecular profiling for therapeutics and for cancer patients in the post-genomic era.

Now available FREE online at: **http://mct.aacrjournals.org**

High-Resolution Array CGH of the NCI-60

William Reinhold, Kimberly Bussey, Uma Shankavaram, John Weinstein
(Genomics & Bioinformatics Group, NCI
Amir Ben-Don, Anya Tsalenko, Laurakay Bruhn, Zohar Yakhini, Doron Lipson, Michael Barrett (Agilent)



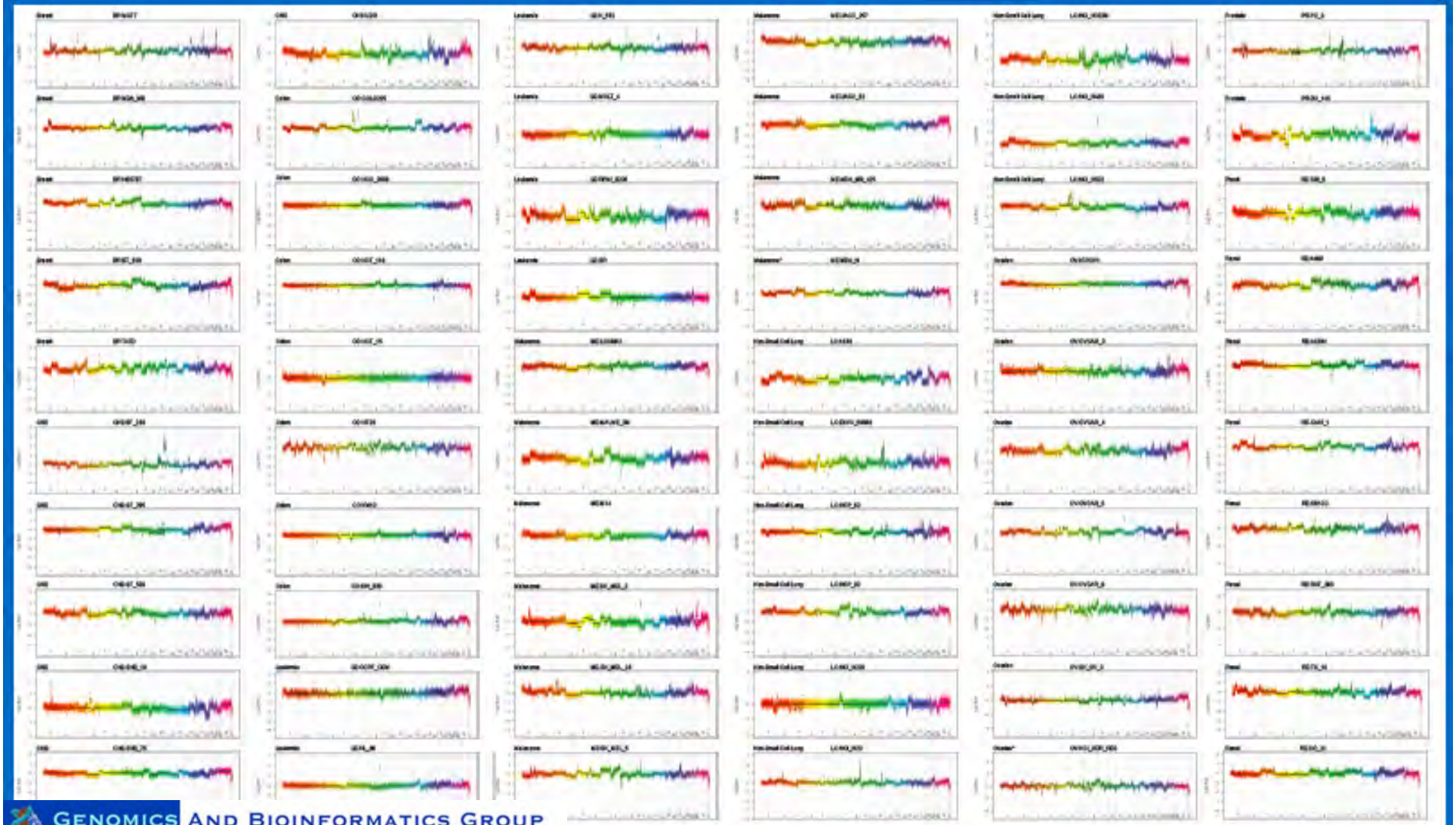
GENOMICS AND BIOINFORMATICS GROUP
LMP, CCR, National Cancer Institute

Agilent, Inc.

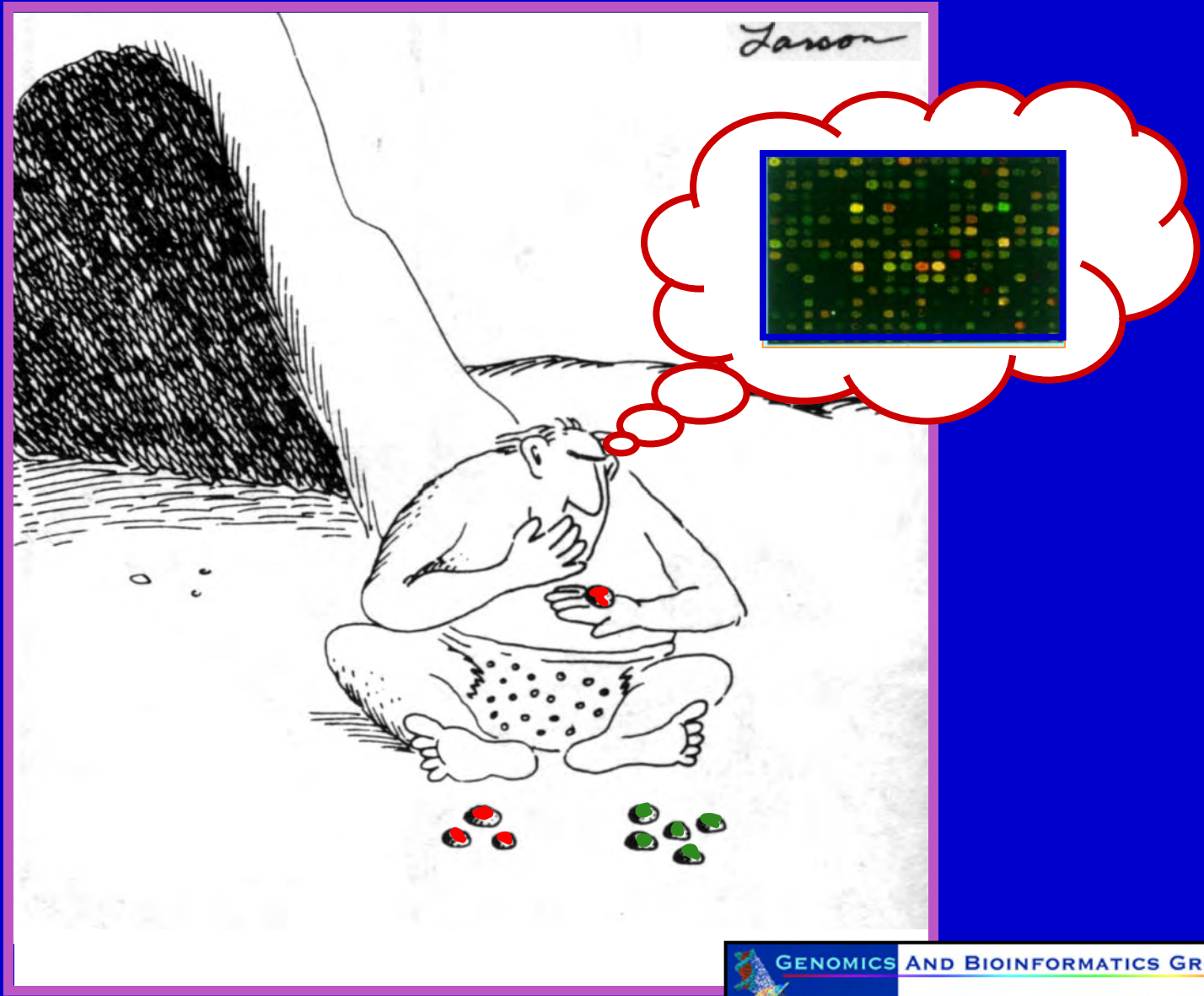
High-resolution oligonucleotide array CGH analysis of the NCI-60 cell

John N. Weinstein(1), William C. Reinhold(1), Kimberly J. Bussey(1), Mark Reimers(1), Uma Shankavaram(1),
Rebecca R. Selzer(2), Todd A. Richmond(2), Peggy S. Eis(2).

(1) Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD.
(2) NimbleGen Systems Inc., One Science Ct., Madison, WI.



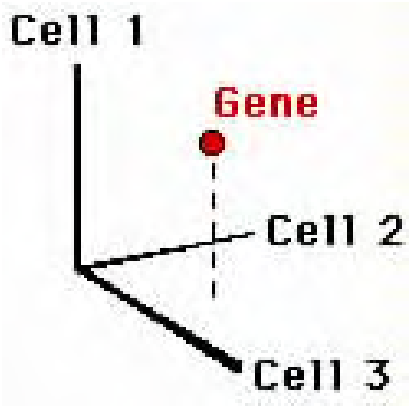
Analyzing and Integrating Omic Data:



GENOMICS AND BIOINFORMATICS GROUP

Statistics of High-Dimensional Omic Data

From Classical Statistics



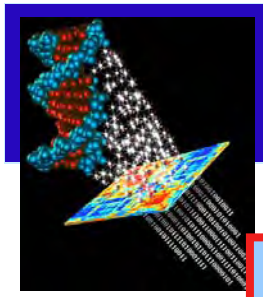
- Multivariate linear regression
- Principal component analysis
- Principal component regression (PCR)
- Multidimensional Scaling
- Partial least squares regression
- Stepwise regression
- ROC analysis
- SAM, PAM, Gene Shaving
- Correlation (Pearson, Spearman)
- Cluster analysis (e.g., hierarchical)
- Clustered image map (e.g., Weinstein, et al., Science 275:343,1997)

Computer-intensive statistical methods

- Bootstrap, Jackknife cross-validation
- Monte Carlo data randomizations, permutation tests
- Hidden Markov chain Monte Carlo
- Hierarchical error modeling with Gibbs sampling

From Machine Learning

- Neural networks (Weinstein, et al., Science 258:447,1992)
- Kohonen self-organizing maps (SOMs)
- Genetic function approximation
- Random Forest



Freely Available Bioinformatic Resources from the Genomics & Bioinformatics Group, LMP, CCR, NCI

(Head: John N. Weinstein, M.D., Ph.D.; jw4i@nih.gov)

(URL:discover.nci.nih.gov -- highlighted in Science (Netwatch)

MatchMiner: translates among gene identifier types for lists of hundreds or thousands of genes. Includes: GenBank accession numbers, IMAGE clone IDs, common gene names, HUGO names, gene symbols, UniGene clusters, FISH-mapped BAC clones, Affymetrix identifiers, and chromosome locations. (Bussey, et al., Genome Biology 2003; 4:R27)

GoMiner and High-throughput GoMiner (HTGM) leverage the Gene Ontology for large lists of genes, for example in microarry experiments. HTGM deals with multiple microarrays in a study, providing calculations and CIM graphics to integrate the data over the entire set of experiments (Zeeberg, et al., Genome Biology 2003; 4:R28 and BMC Bioinformatics 2005:6; 168).

CIMminer: generates color-coded Clustered Image Maps (CIMs) ("heat maps") to represent "high-dimensional" data sets such as gene expression profiles. We introduced CIMs in the early 1990's. Clustering of the axes brings like together with like to create patterns of color. (Weinstein, et al., Science 1997; 275:343-349).

MedMiner: searches and organizes the biomedical literature on genes, gene-gene relationships, and gene-drug relationships. MedMiner speeds up 5-fold the capture and organization of literature from PubMed searches. (Tanabe, et al., BioTechniques 1999; 27: 1210).

AbMiner: relational database of commercially available antibodies and quality-control information on them.

LeadScope/ LeadMiner™: links molecular markers and the drug discovery process. It links gene expression profiles for the NCI-60 (or other screening cell panel) to a set of 27,000 chemical substructure descriptors of compounds tested against the cells. (Blower, et al., Nature's Pharmacogenomics J. 2002; 2:259)

SmudgeMiner: introduces new algorithms and visualizations for detecting regional artefacts on microarrays and assessing their influence. (Reimers and Weinstein, BMC Bioinformatics 2005;6: 166)

MIMminer: electronic Molecular Interaction Maps (MIMs) (Kohn, et al., Mol.Biol. Cell 2006;17: 1)

CellMiner: DNA, RNA,and protein molecular profiles on the NCI-60 and drug-resistant cells in a queryable, user-friendly web database/tool (Shankavaram, et al., Molec. Ca. Ther., in press).

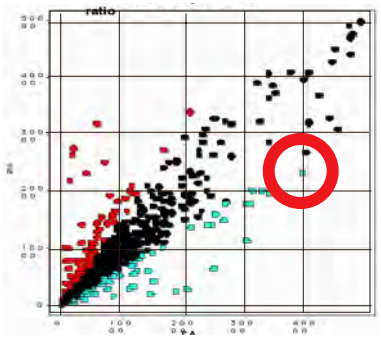
SpliceMiner: Organizes information on splice variants of genes (Kahn, et al., BMC Bioinf. 2007;8:75).

http://discover.nci.nih.gov (lead software engineer: D. Kane, SRA)

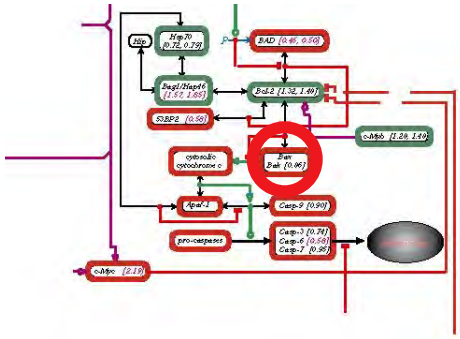
The Miner Suite: Tools for Integromic Research

Input: DNA, RNA, Protein, Functional, Structural, Chemical, Pharmacological, Clinical Data

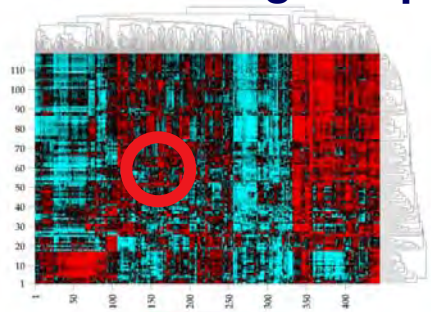
Microarray Data



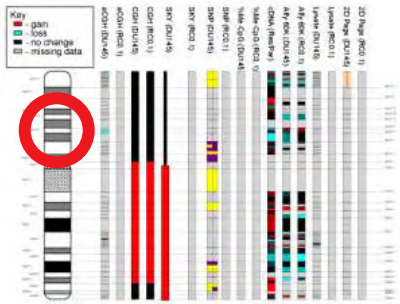
Molec. Interaction Maps



Clustered Image Maps



Chromosome Icons



GBG Software:

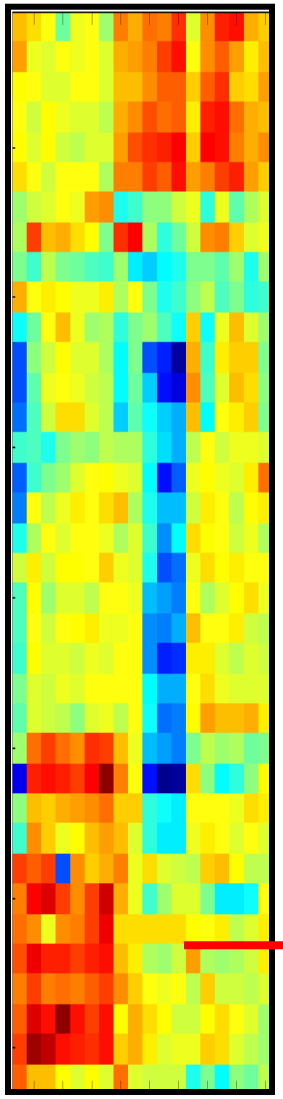
- MIMminer
- CIMminer
- MatchMiner
- GoMiner
- MedMiner
- AbMiner
- LeadMiner
- MethMiner
- SmudgeMiner
- CellMiner
- SpliceMiner

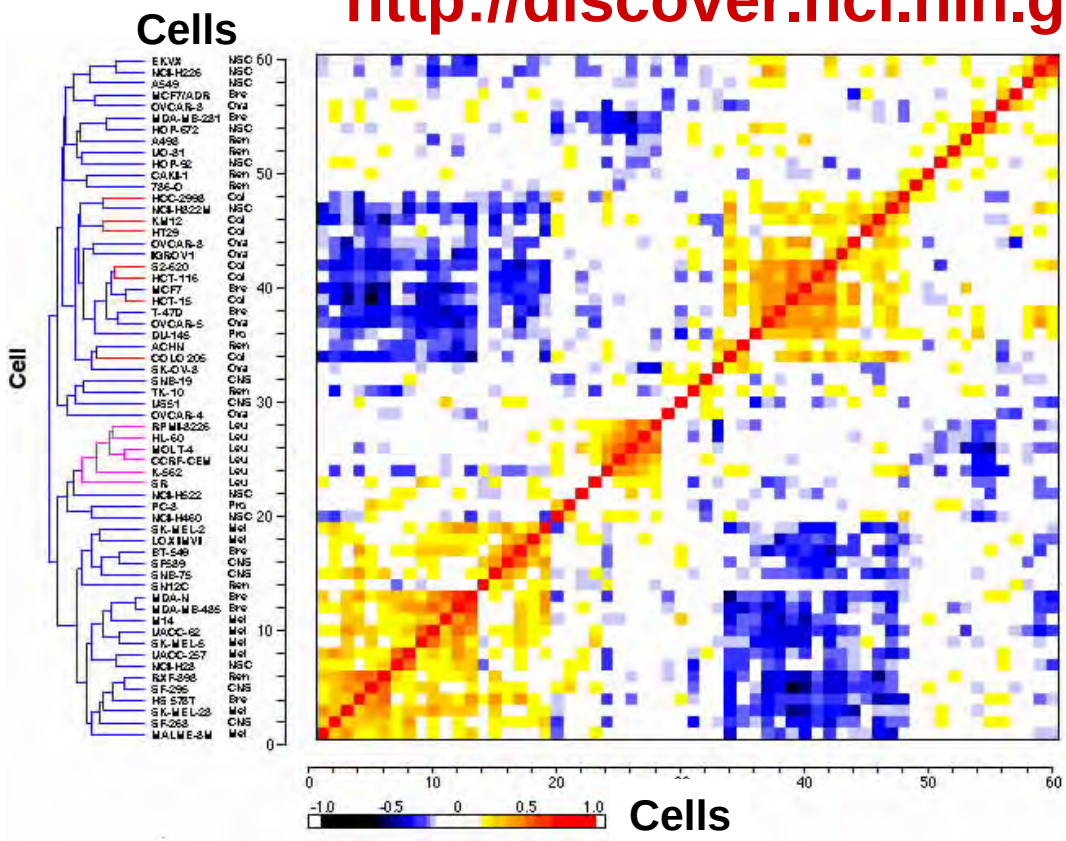
External Resources:

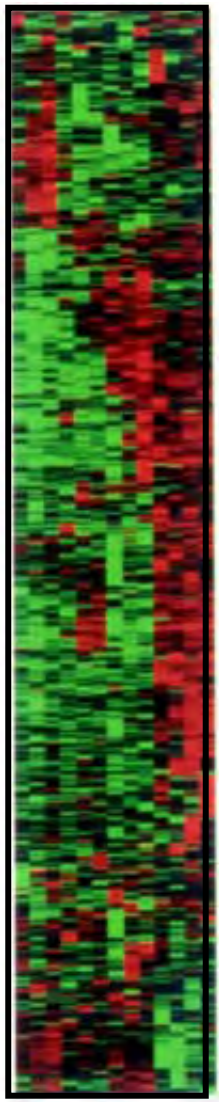
- Spotfire
- GenBank
- GeneCards
- CGAP
- MAExplorer
- UniGene
- UCSC
- LocusLink
- OMIM

CIMminer: Clustered Image Maps (CIMs)

http://discover.nci.nih.gov







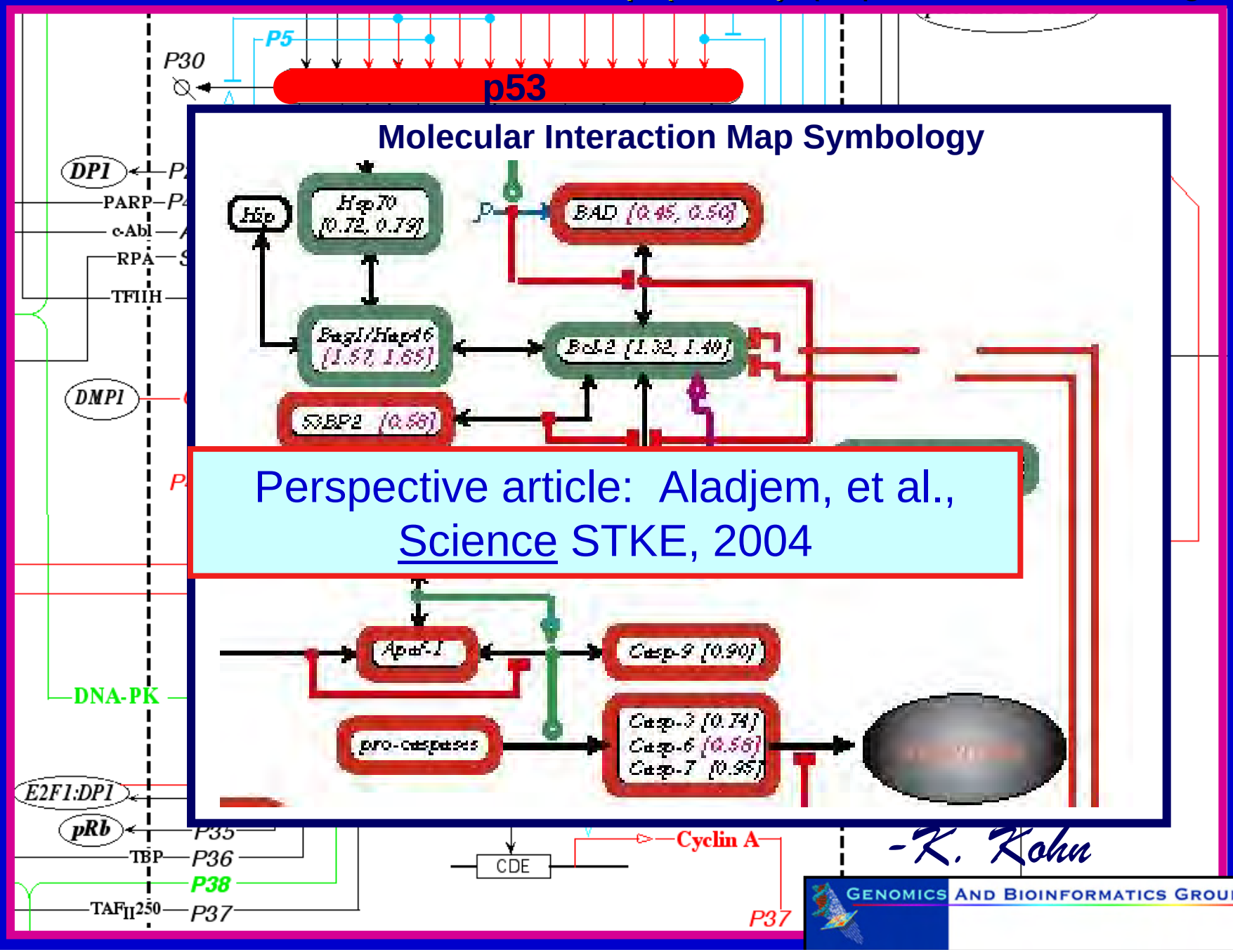
Weinstein et al, Stem Cells, 1994

Weinstein et al, Science, 1997

Myers…Weinstein, Electrophor.,1997

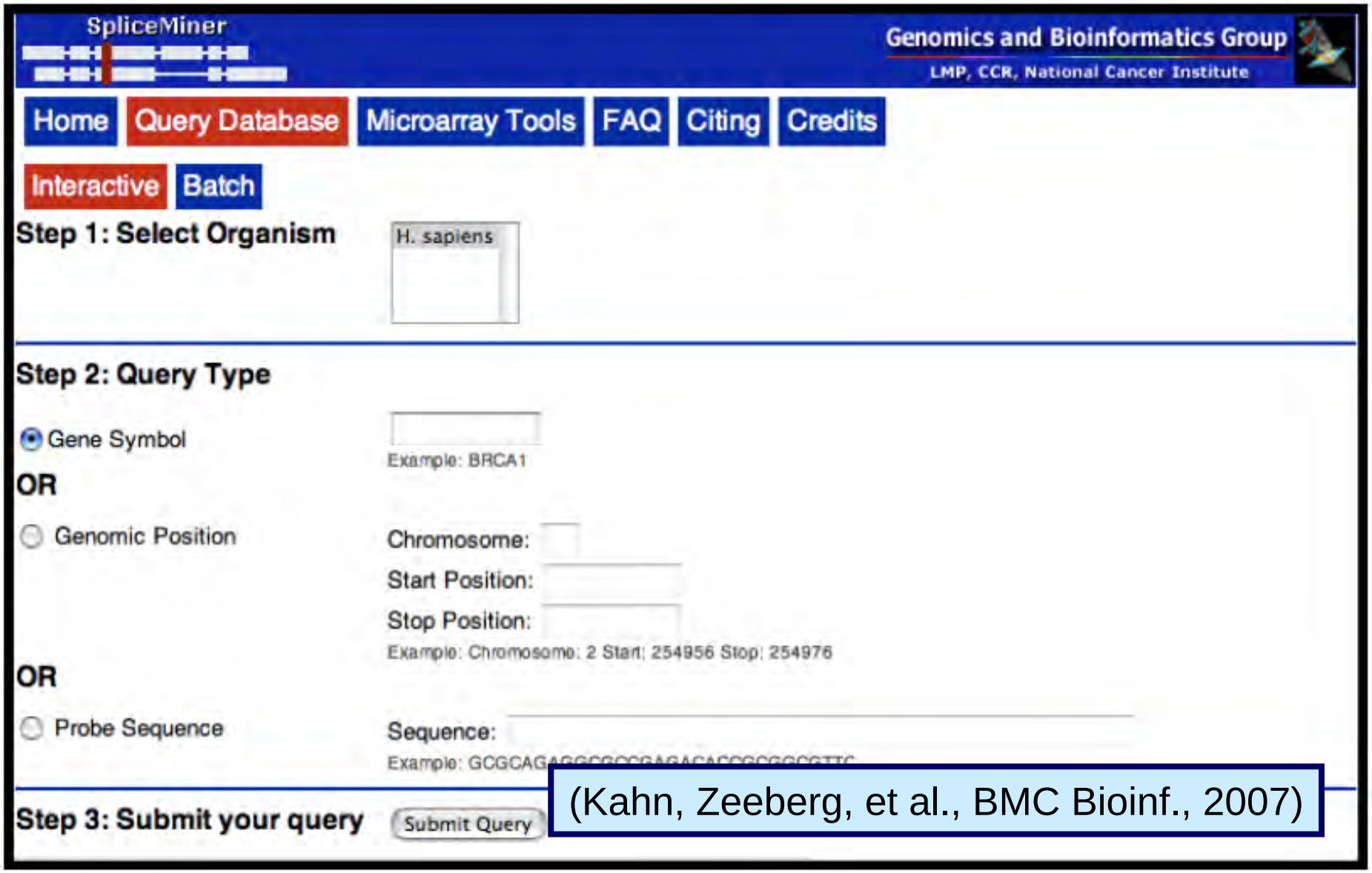
Eisen et al, PNAS, 1998

Electronic Molecular Interaction Map (eMIM) (http://discover.nci.nih.gov)



Perspective article: Aladjem, et al., Science STKE, 2004

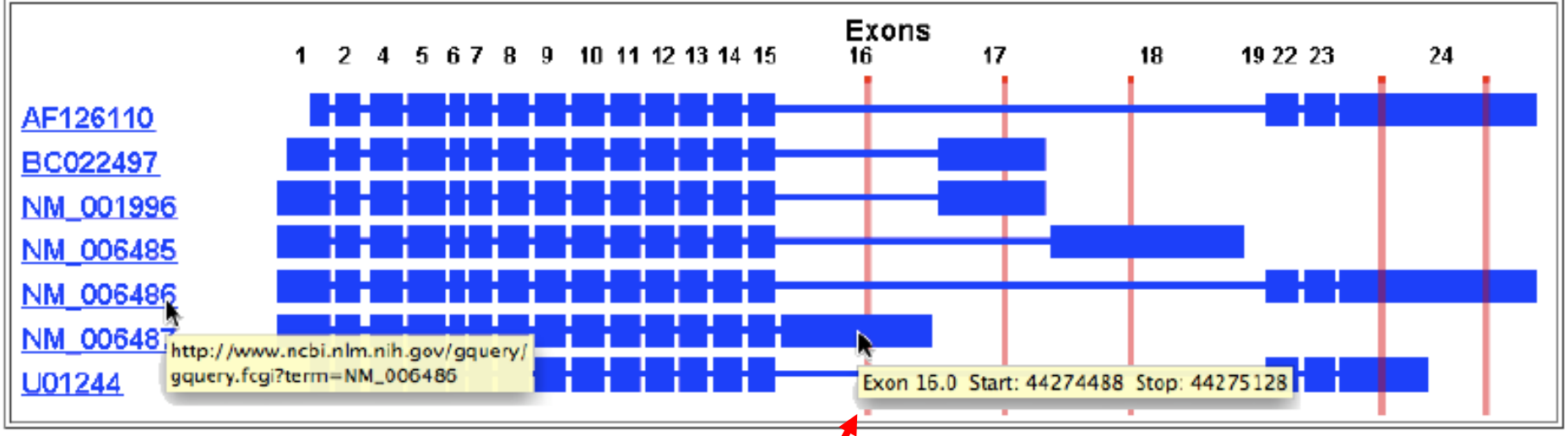
SpliceMiner Input



(Kahn, Zeeberg, et al., BMC Bioinf., 2007)

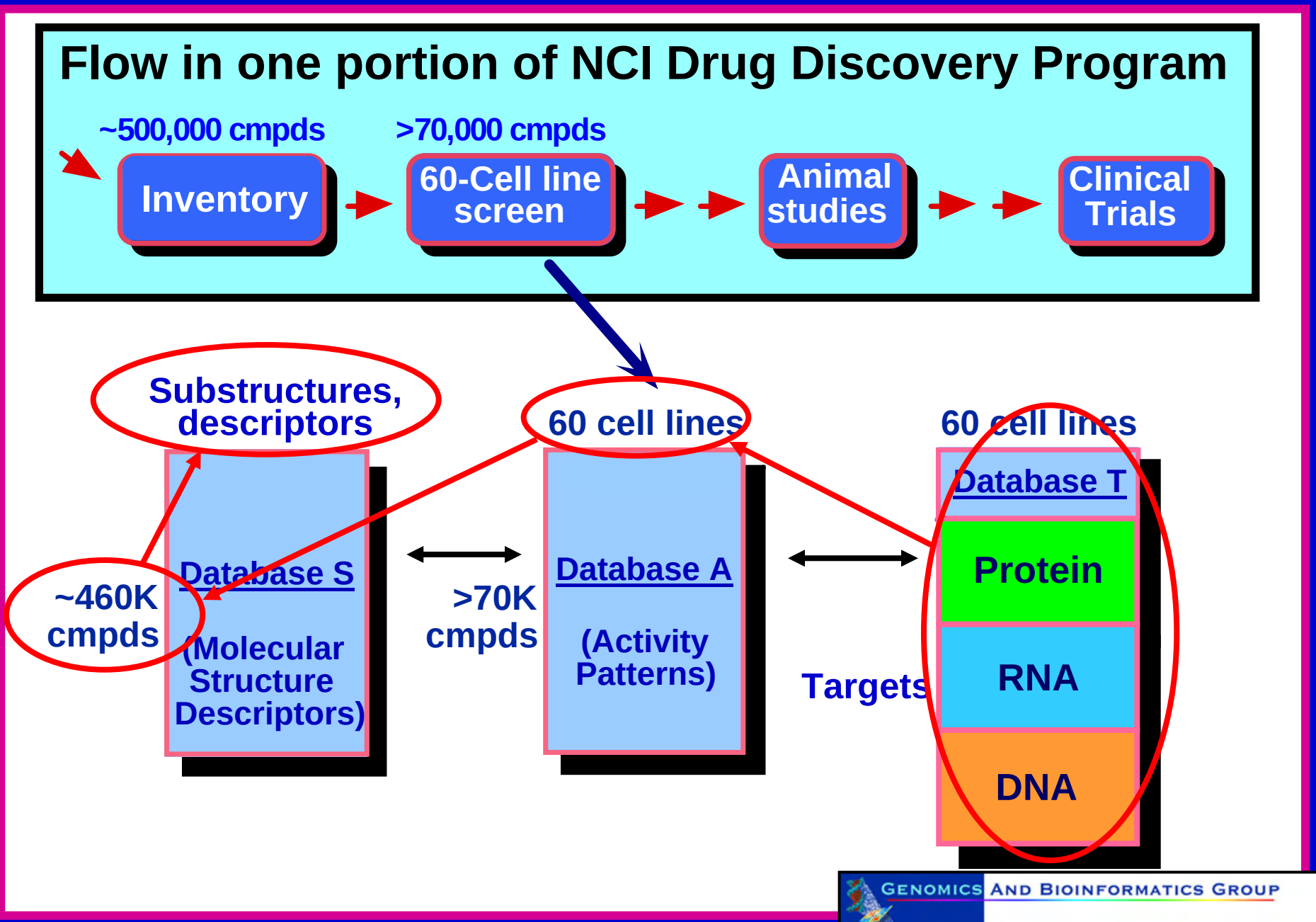
SpliceMiner Output

Gene Symbol	Alternate Symbols	Chromosome	Strand	Start	Stop	Total Exons
FBLN1	FBLN	22	+	44219256	44317551	24

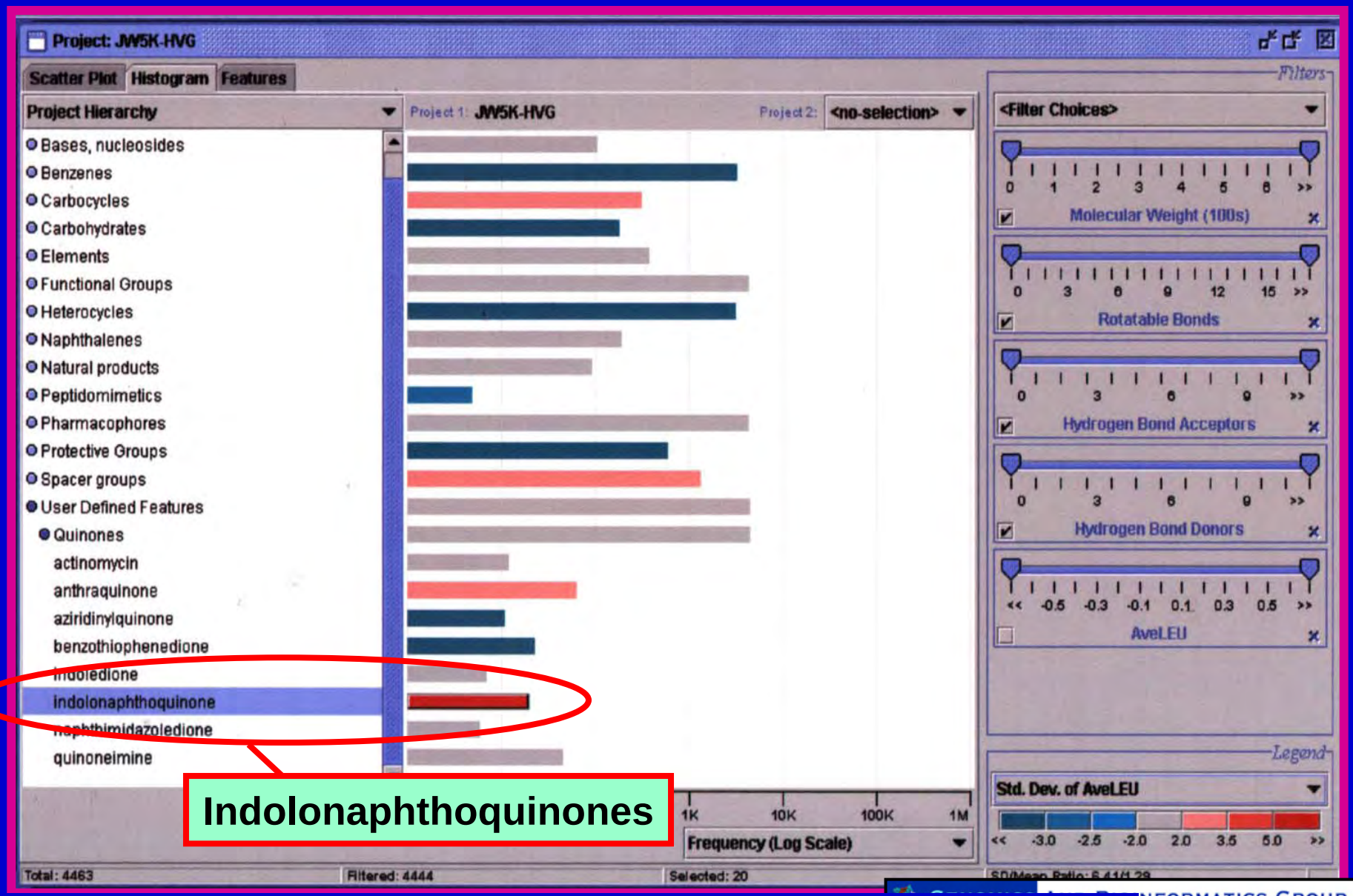


Probe

Leadscope/LeadMiner: From Genes to Chemical Substructures

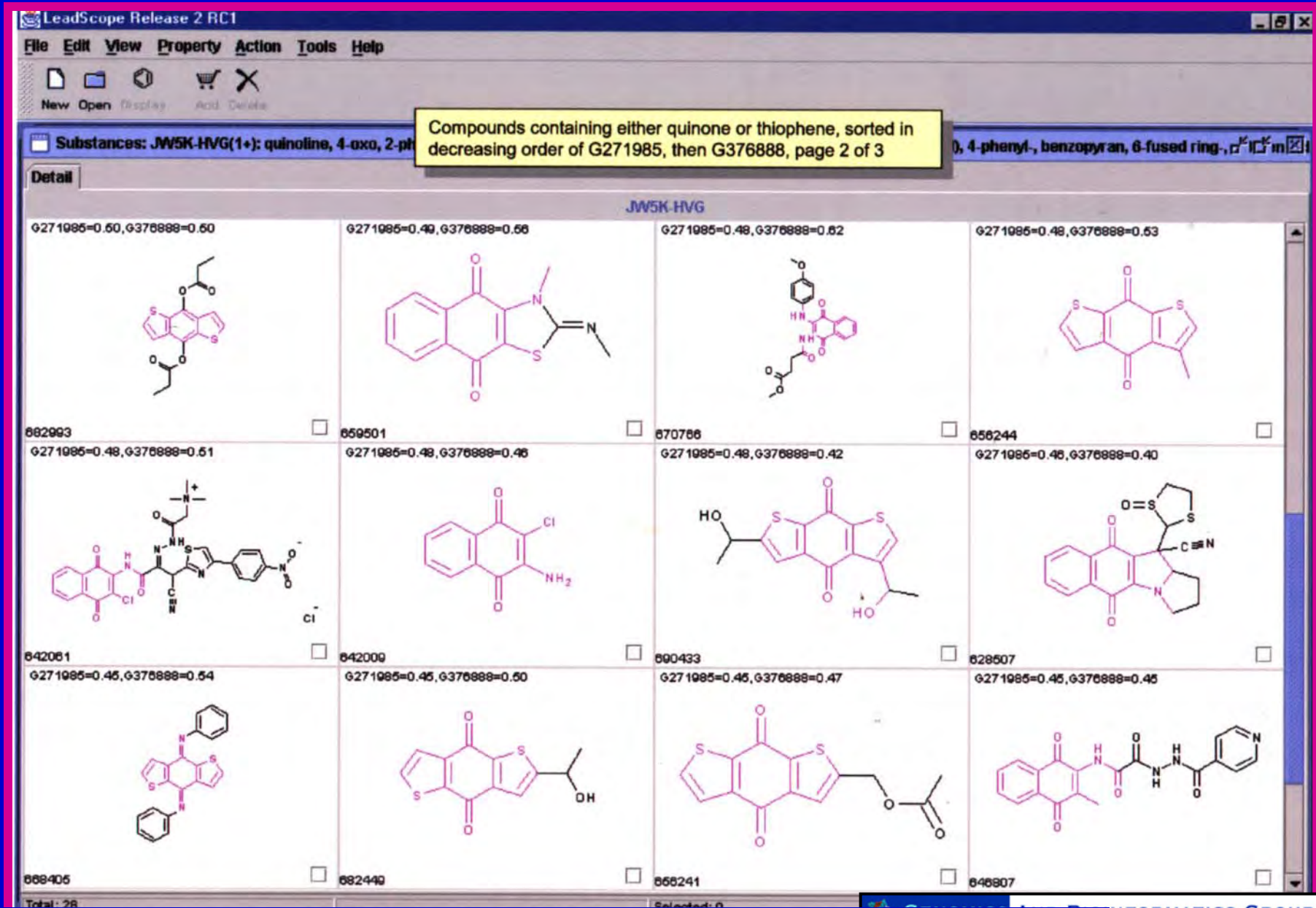


Leadscope/LeadMiner: From Genes to Chemical Substructures



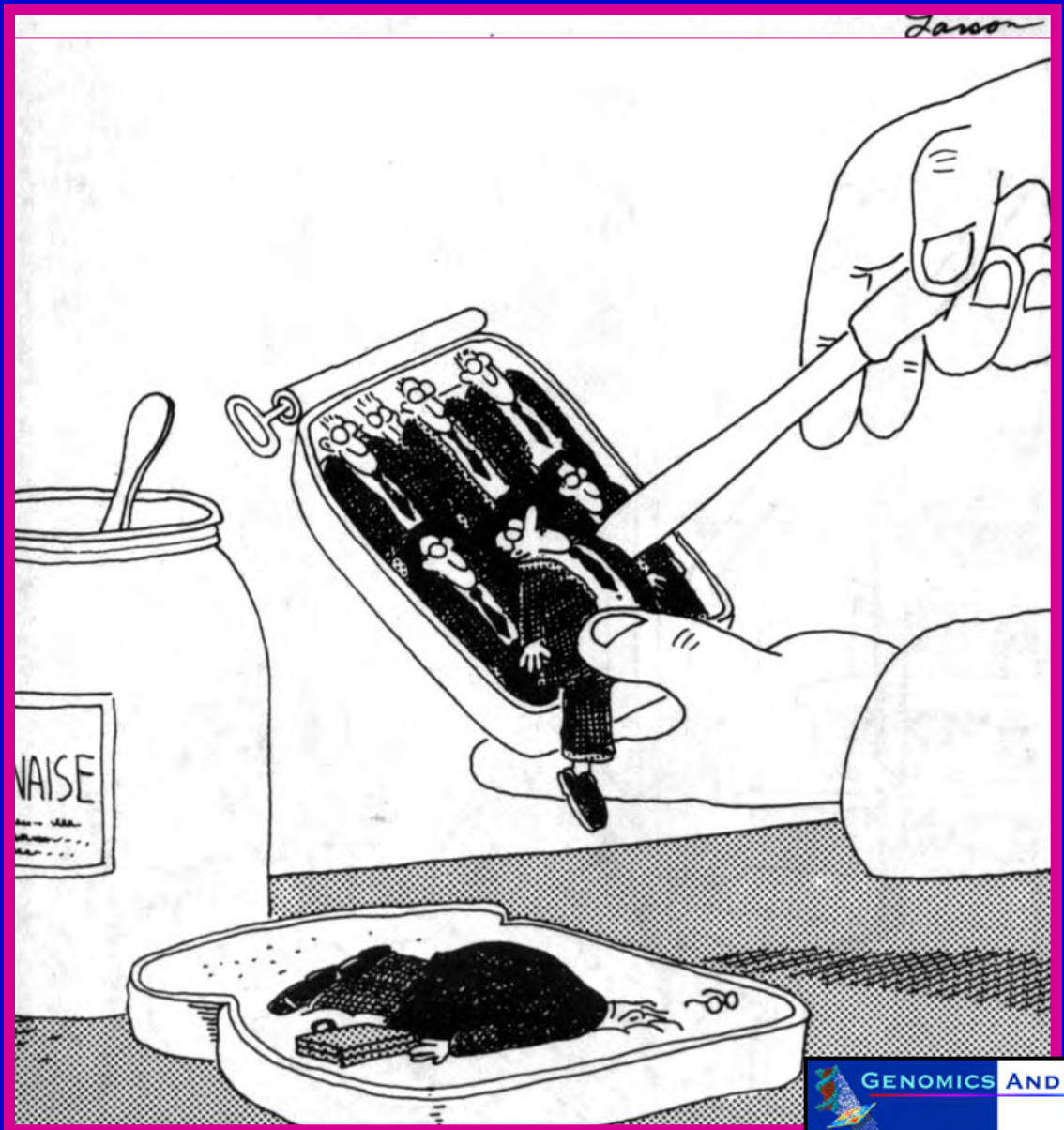
(Blower, et al., Nature's Pharmacogenomics Journal 2:259, 2002)

Leadscope/LeadMiner: From Genes to Chemical Substructures



(Blower, et al., Nature's Pharmacogenomics Journal 2:259, 2002)

Integromics: Translational Vignettes



GENOMICS AND BIOINFORMATICS GROUP

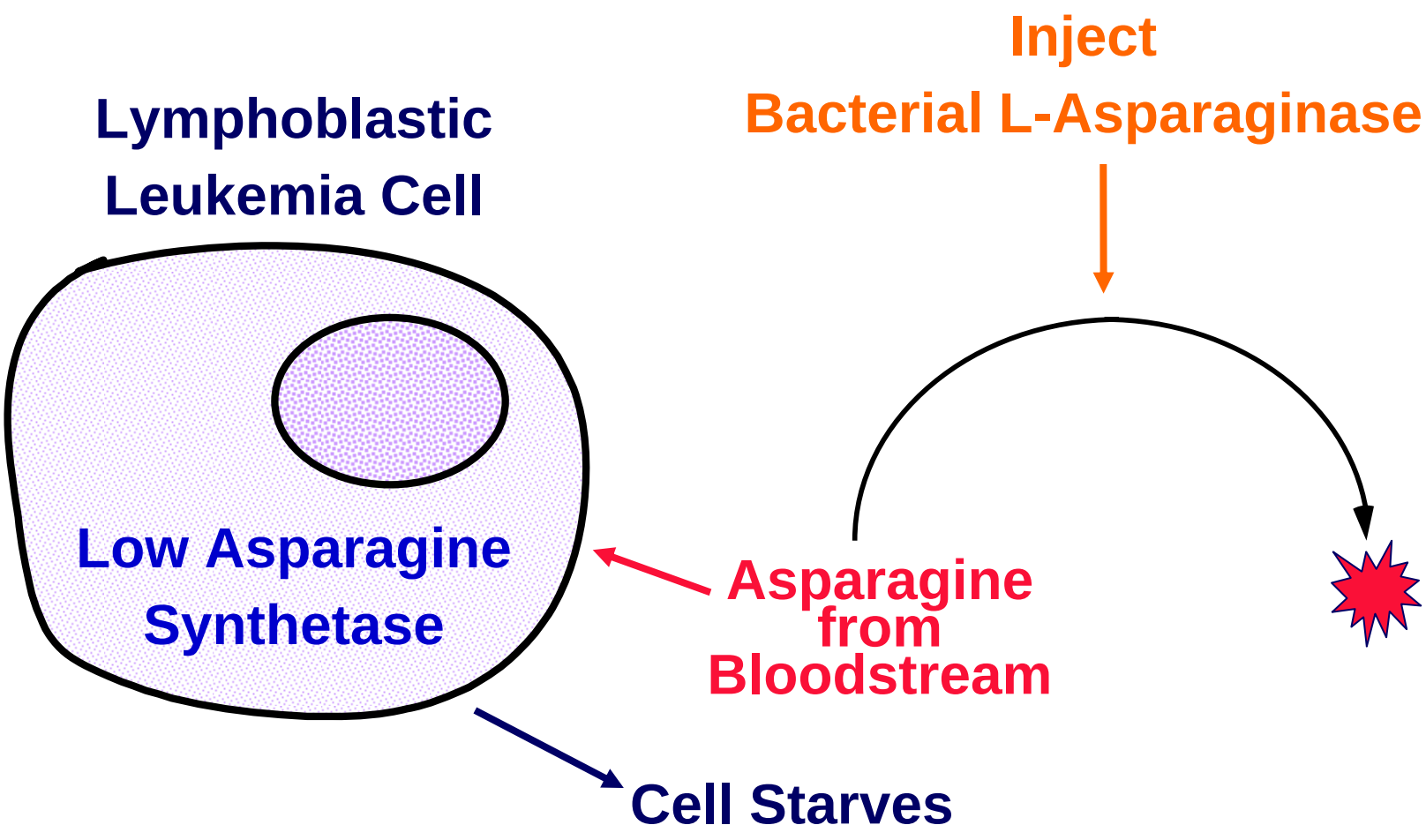
Integromic Vignette 1: Pharmacogenomics

Asparagine Synthetase Expression

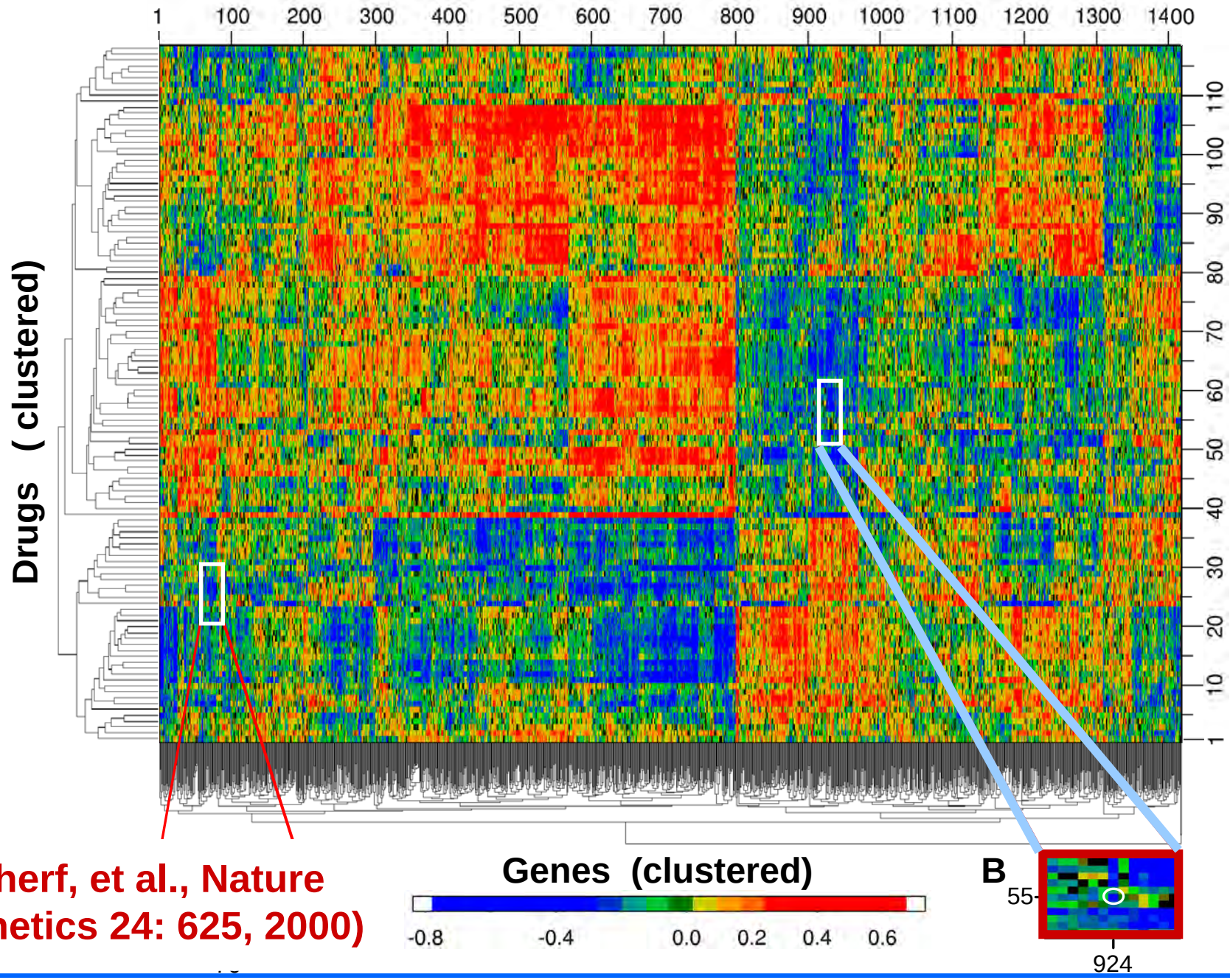
and

L-Asparaginase Activity

Rationale for Therapy with L-Asparaginase

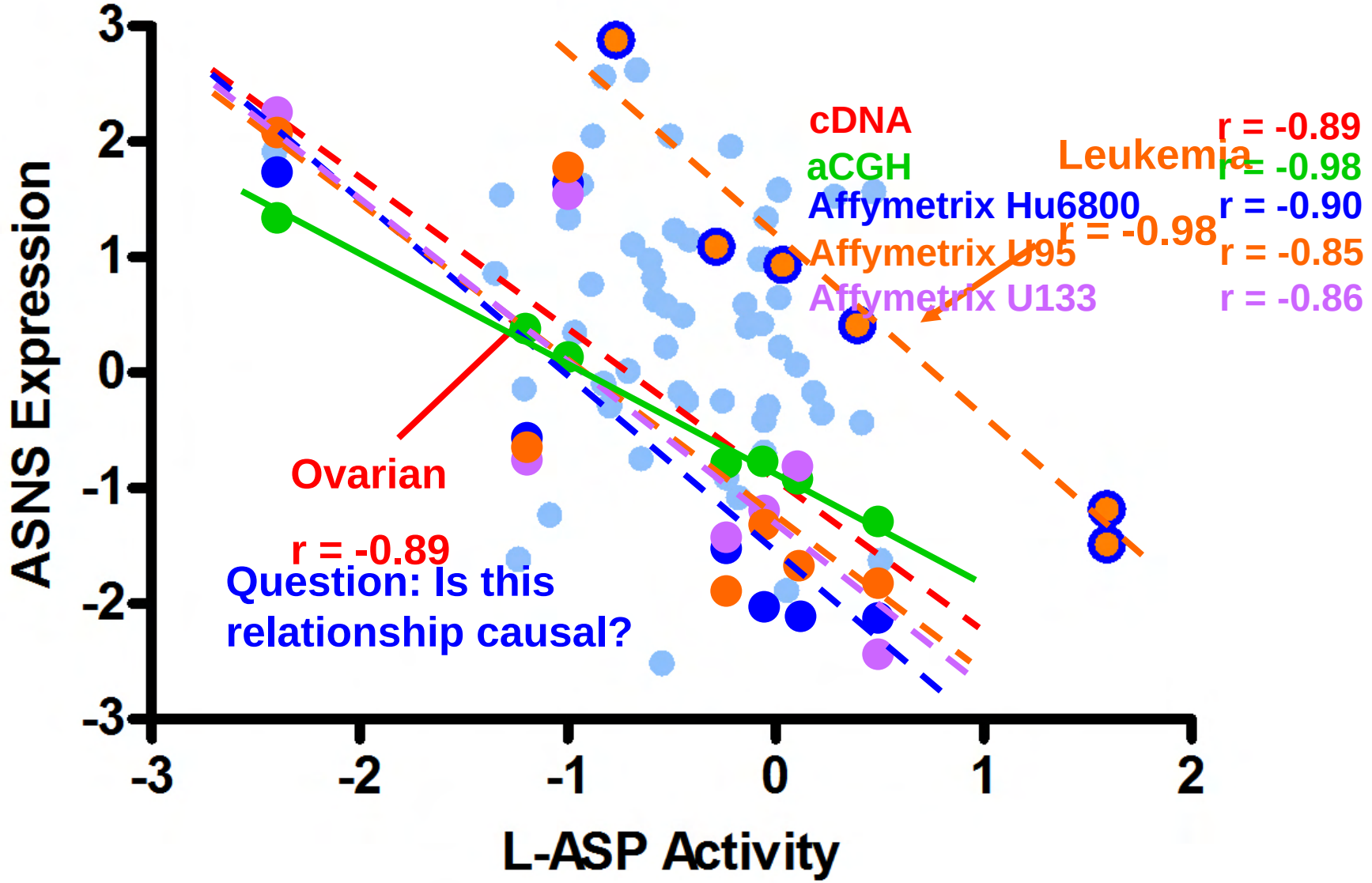


Correlating Gene Expression Pattern with Drug Activity (CIM)



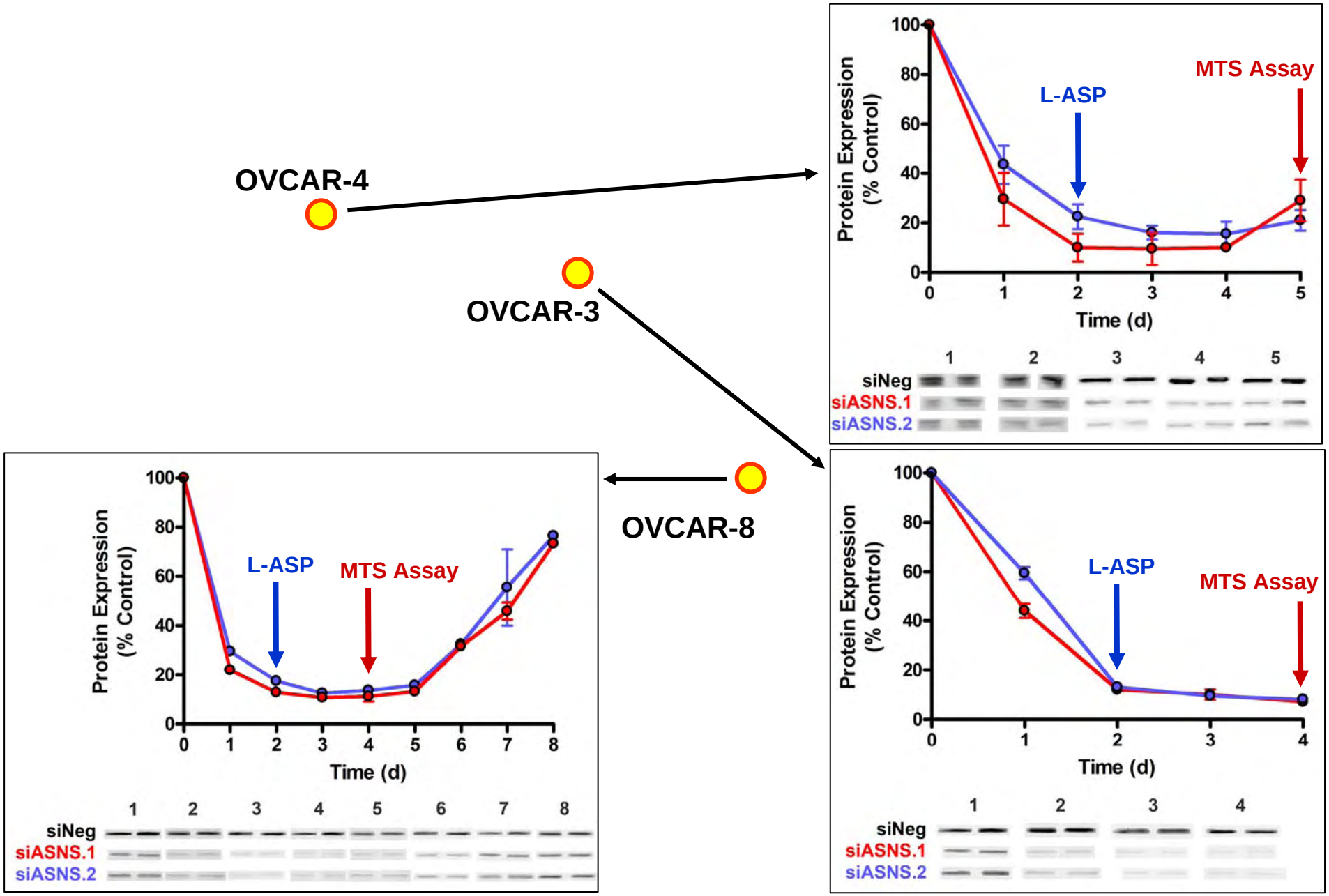
(Scherf, et al., Nature Genetics 24: 625, 2000)

ASNS Expression vs. L-ASP Activity in the NCI-60 Cell Lines



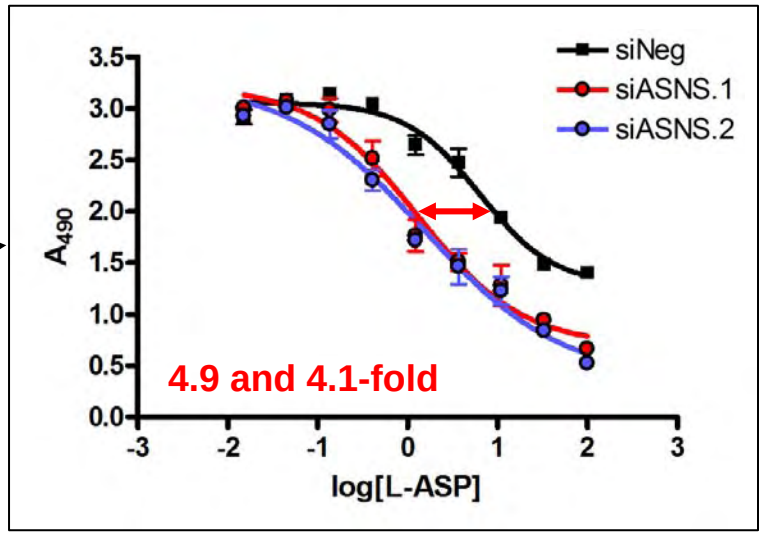
Scherf et al., *Nat Genet* (2000); Bussey et al., *Mol Ca. Ther* (2006); Lorenzi et al., *Mol Cancer Ther*, submitted

ASNS Protein Expression after siRNA Transfection

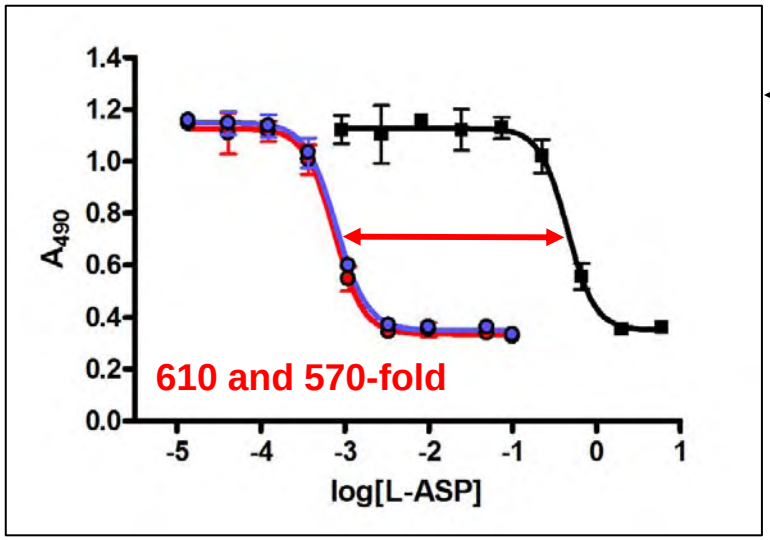


Response to L-ASP After siRNA Knock-down

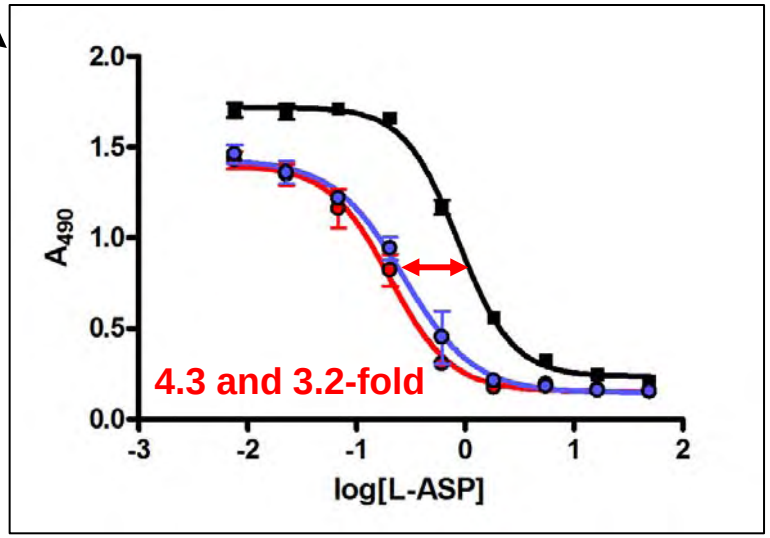
OVCAR-4



OVCAR-3

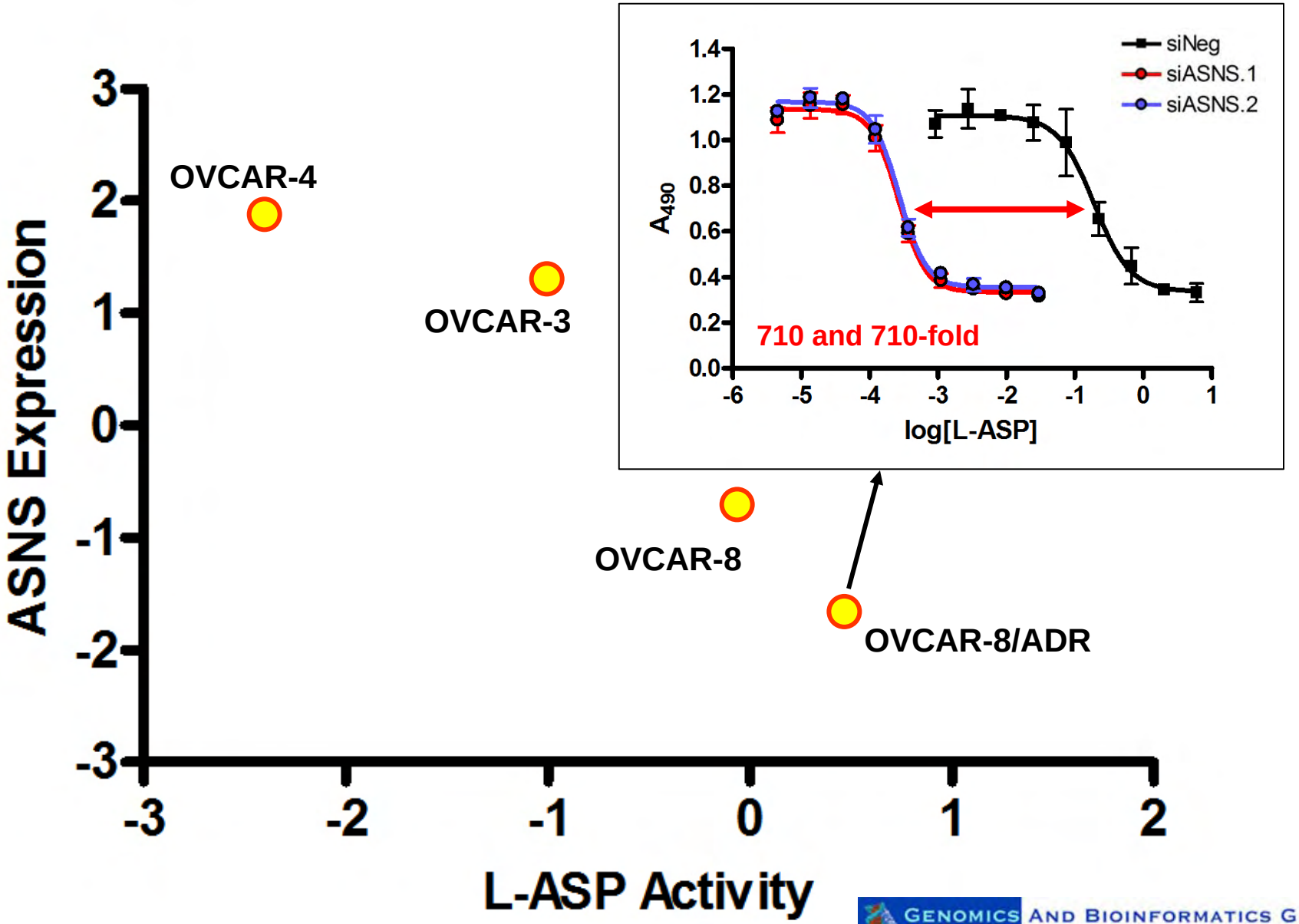


OVCAR-8



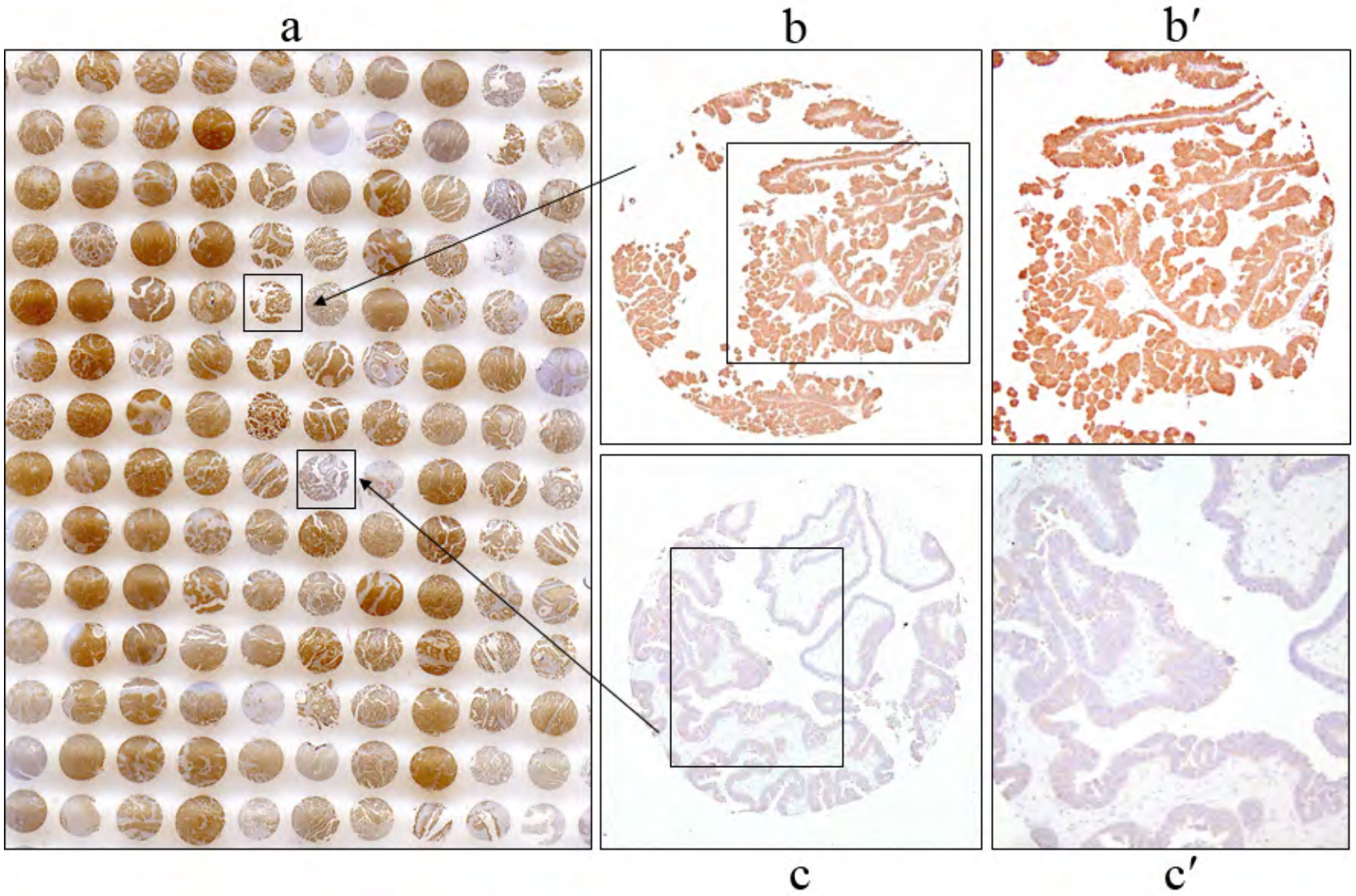
GENOMICS AND BIOINFORMATICS GROUP
LMP, CCR, National Cancer Institute

Multidrug Resistance and the ASNS/L-ASP Relationship



GENOMICS AND BIOINFORMATICS GROUP
LMP, CCR, National Cancer Institute

Ovarian Cancer Tissue Array



Conclusions: L-Asparaginase

- There is a pharmacogenomic / pharmacoproteomic relationship between L-ASP activity and the expression of ASNS in ovarian cancer cell lines.
- That relationship is causal (siRNA expts).
- The relationship is not affected by classical multidrug resistance.
- The findings provide rationale for clinical testing of L-ASP against low-ASNS subsets of ovarian cancers with ASNS as a causally linked biomarker.

Omics - Beyond Genomics and Proteomics

- Genc
- Prote
- Tran
- Kino
- CHO
- Immu
- Toxi
- Meta
- Meta
- Tran
- Clinc

Synergy between omic and hypothesis-driven research

Synergy between experimental science and bioinformatics

Synergy among scientific cultures

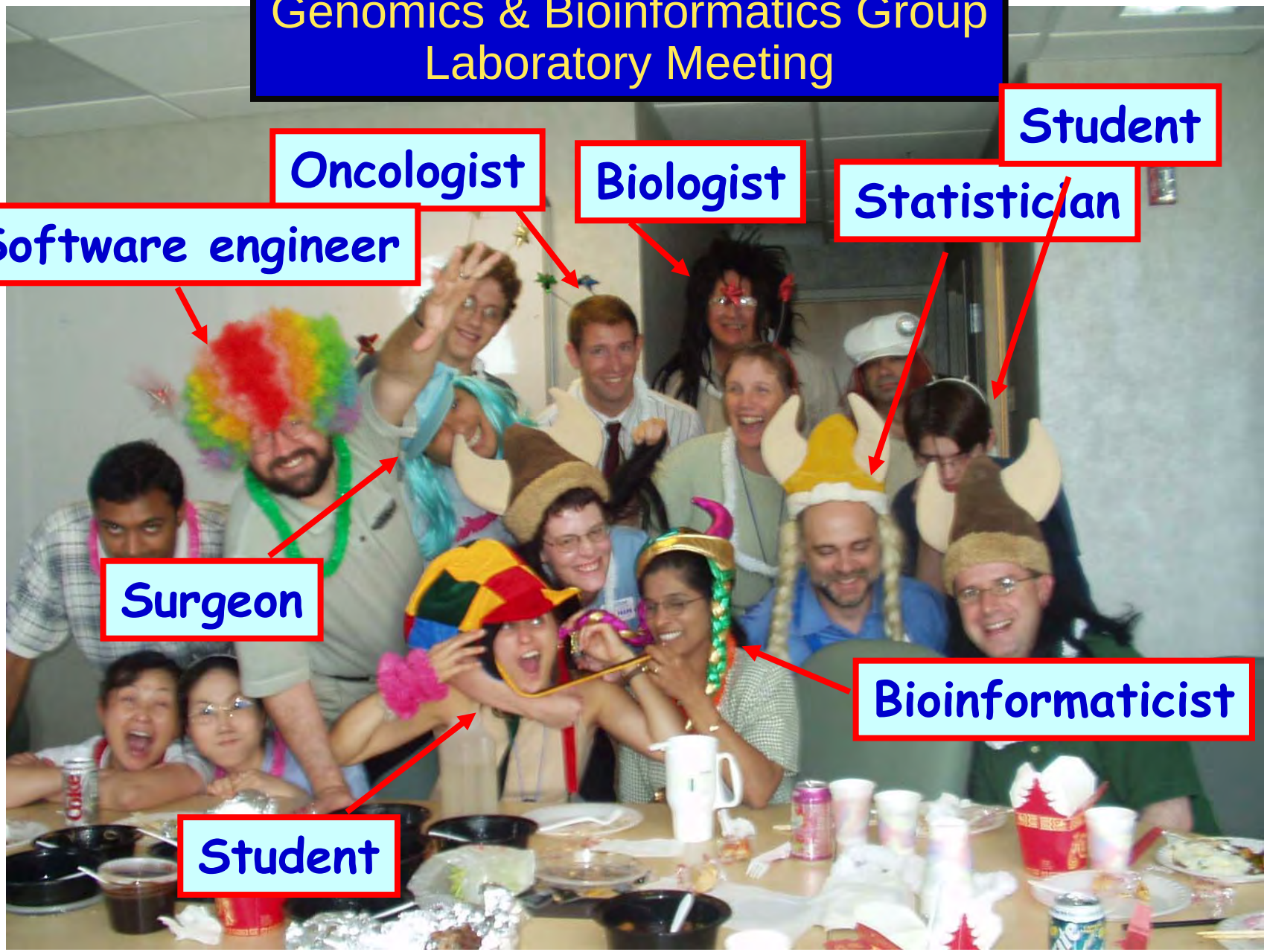
Synergy between domain and analytical expertise

Synergy among types of molecular data

Integromics™

nics
ics
omics
408 (2000)
• Pharmacogenomics
nics
• Pharmacomics
ics
s
nary
7, 1998)

Genomics & Bioinformatics Group Laboratory Meeting



Principal NIH Collaborators

Genomics & Bioinformatics Group

- John Weinstein
- William Reinhold
- Gabriel Eichler
- Philip Lorenzi
- Ogechi Ikediobi
- Barry Zeeberg

Past Members

- Uma Shankavaram
- Satoshi Nishizuka
- Mark Waltham
- Samir Lababidi
- Kim Bussey
- Jes Alexander
- Hosein Kouros-Mehr
- Alika Maunakea
- Jae Lee
- Fuad Gwadry
- Sohyoung Kim
- Sylvia major
- Uwe Scherf
- Ari Kahn
- Sudar Narasimhan (GBG, SRA)
- Ike Nnabuggwu (GBG, SRA)
- Ajay

GBG Computer Science Team

- David Kane (GBG, SRA)
- Margot Sunshine (GBG, SRA)
- Sohana Choudhry(GBG, SRA)
- Hong Cao (GBG, SRA)

LMP, NCI

- Yves Pommier
- Kurt Kohn
- Mirit Aladjem

DTP, NCI

- Ken Paull
- Dominic Scudiero
- Anne Monks
- David Covell
- Susan Holbeck
- Dan Zaharevitz
- Mark Kunkel
- Michael Grever

- Michael Boyd
- Bruce Chabner

NIH

- N Caplen
- P Goldsmith
- E Kohn
- M Raffeld
- M Gottesman
- S Ambs
- C Harris
- T Fojo
- S Bates
- S Hewitt
- G Szakacs
- J-P Annereau
- R Klausner
- JC Barrett
- J Riss
- I Kirsch
- A Roschke
- K Buetow
- D Nelson
- L Rubinstein
- T Veenstra
- F Gonzalez
- G Alvord
- J Roayeaei
- P Munson

Principal Collaborators Outside of NIH

Stanford

- D Ross, C. Perou, M Eisen, P Brown, D Botstein

UCSF

- J Gray, K Chen

Broad

- E Lander, T Golub, J Staunton, J Mesirov

Ohio State

- P Blower, W Sadee, Y Huang, C Croce, CG Liu

NCI, GMU

- L Liotta, E Petricoin, C Jamison

Albert Einstein

- J Greally, A Melnick

Georgetown U.

- A Fornace, Hongfang Liu

VTT Center, Finland

- O Kallioniemi

TGen

- D von Hoff, S Mousses, M Barrett

Large Scale Biology

- N L. Anderson

GeneLogic

- E Kaldjian, J. Cossman, D Dolginow

Agilent

- Z Yakhini, D Lipson, A Ben-Dor,

NimbleGen

- R Seltzer, P Eis, R Green

ExonHit

- R Einstein

Enzon

- I Horak

Metabolic Analyses

- C. Beecher