Lucile Packard Children's Hospital AT STANFORD

Big Data in Biomedicine: Discovering new drugs and diagnostics from a trillion points of data

Atul Butte, MD, PhD abutte@stanford.edu

@atulbutte

Chief, Division of Systems Medicine, Department of Pediatrics, Genetics, and by courtesy, Medicine, Pathology, and Computer Science

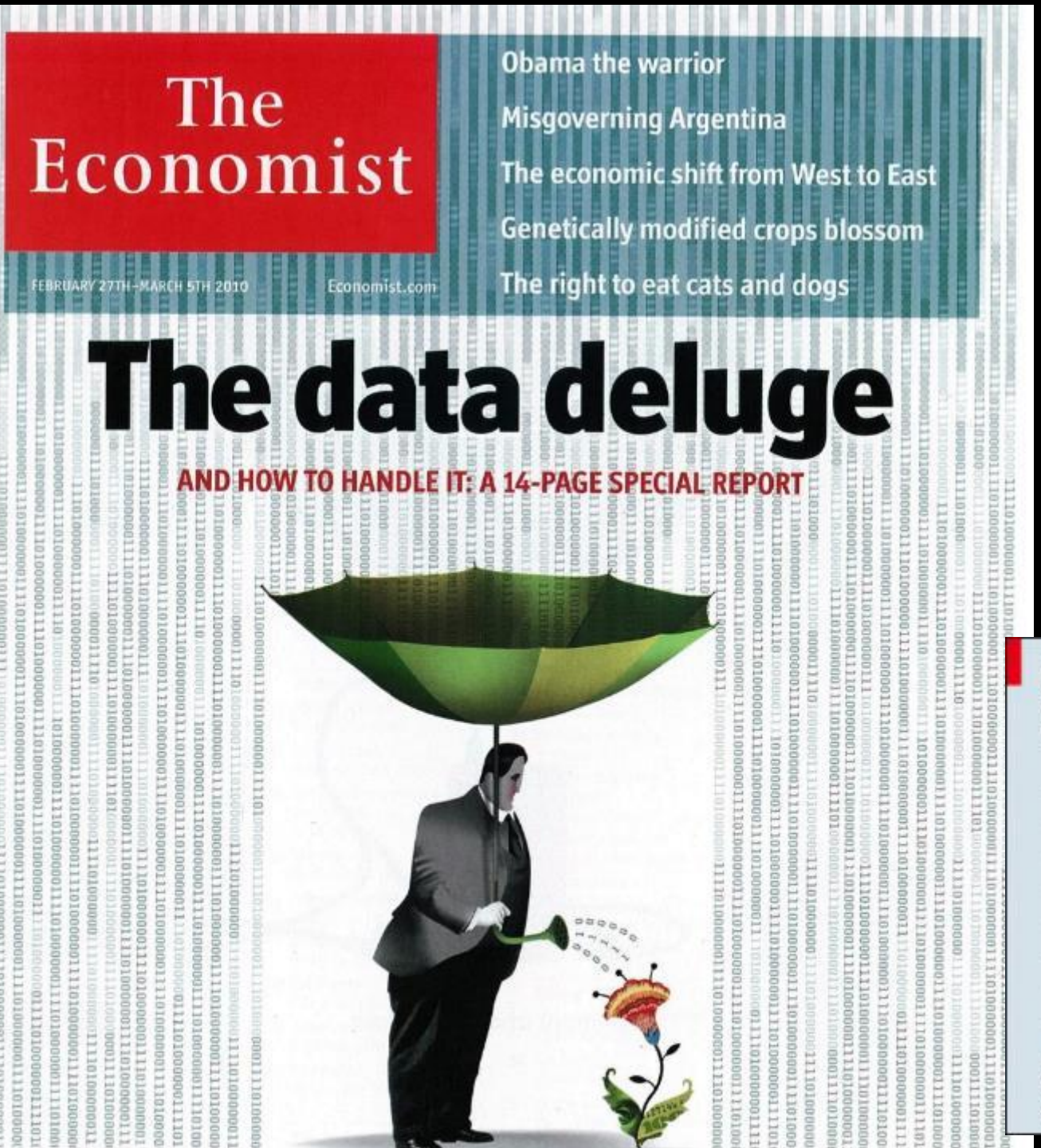
Center for Pediatric Bioinformatics, LPCH

Stanford University

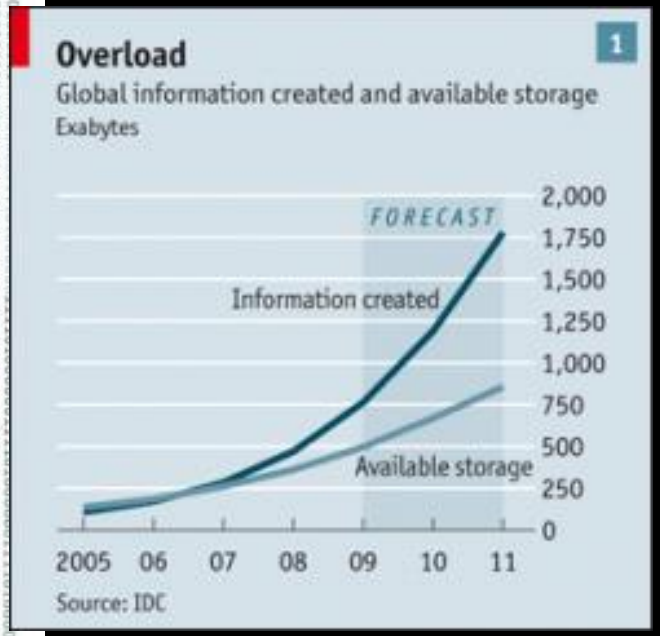
Butte Lab
Systems Medicine • Stanford Pediatrics • Packard Children's Hospital

Disclosures

- Scientific founder and advisory board membership
  - Genstruct
  - NuMedii
  - Personalis
  - Carmenta
- Past or present consultancy
  - Lilly
  - Johnson and Johnson
  - Roche
  - NuMedii
  - Genstruct
  - Tercica
  - Ecoeos
  - Ansh Labs
  - Prevendia
  - Samsung
  - Assay Depot
  - Regeneron
  - Verinata
  - Geisinger
- Honoraria
  - Lilly
  - Pfizer
  - Siemens
  - Bristol Myers Squibb
  - AstraZeneca
- Corporate Relationships
  - Aptalis
  - Thomson Reuters
- Speakers' bureau
  - None
- Companies started by students
  - Carmenta
  - Serendipity
  - NuMedii
  - Stimulomics
  - NunaHealth
  - Praedicat
  - MyTime
  - Flipora



Kilo
Mega
Giga
Tera
Peta
Exa
Zetta



Big Data in Biomedicine

The Next Scientific Revolution

The Four Paradigms of Science

THEORY
Beginning in ancient Greece and China, people tried to explain their observations through natural laws instead of supernatural causes.

EXPERIMENTATION
By the 17th century, scientists like Isaac Newton tried to make predictions for new phenomena and would verify hypotheses by conducting experiments.

COMPUTATION AND SIMULATION
The advent of high-performance computers in the latter half of the 20th century allowed scientists to explore regimes inaccessible to experiment and theory, such as climate modeling or galaxy formation, by numerically solving systems of equations on a large scale and in fine detail.

DATA MINING
Using more-powerful computers, scientists begin with the data and direct programs to mine enormous databases for relationships. In essence, they use computers to discover the rules by studying the data.

nature

www.nature.com/nature

Vol 461 | Issue no. 7261 | 10 September 2

Data's shameful neglect

Research canno

Brooks Hanson is Deputy Editor for physical sciences at

Making Data Maximally Available

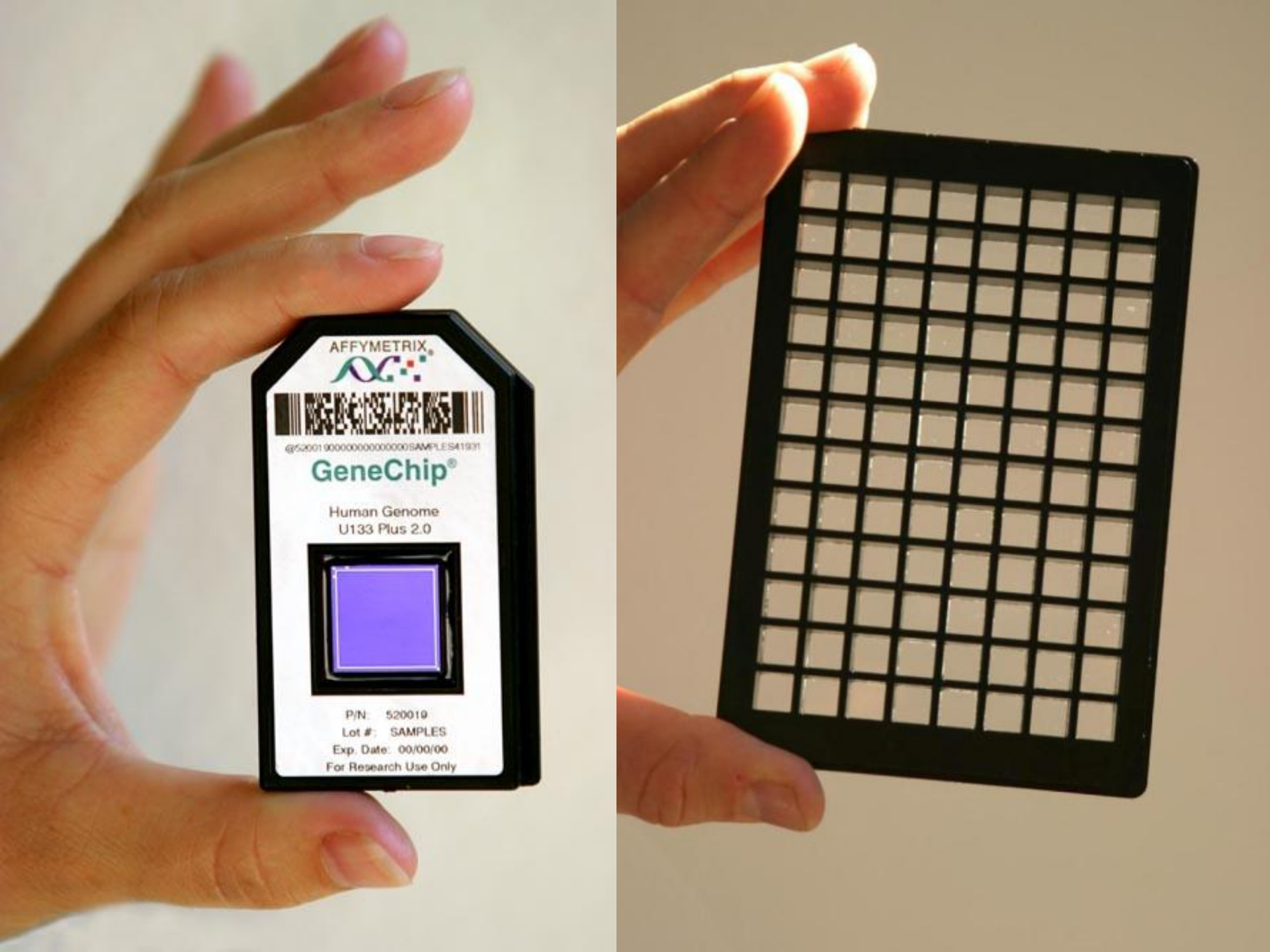
Sharing research data to improve public health

The purpose of medical research is to analyse and understand health and disease. A key and expensive element is the study of populations to explore how interactions between behaviour and environment, in the context of genetic diversity, determine causation and variation in

that every last ounce of knowledge will be wrung from the research.

Ensuring data are made widely available to the research community accelerates the pace of discovery and enhances the efficiency of the research enterprise.

Published Online
January 10, 2011
DOI:10.1016/S0140-6736(10)62234-9


AFFYMETRIX
GeneChip®
Human Genome
U133 Plus 2.0
P/N: 520019
Lot #: SAMPLES
Exp. Date: 00/00/00
For Research Use Only

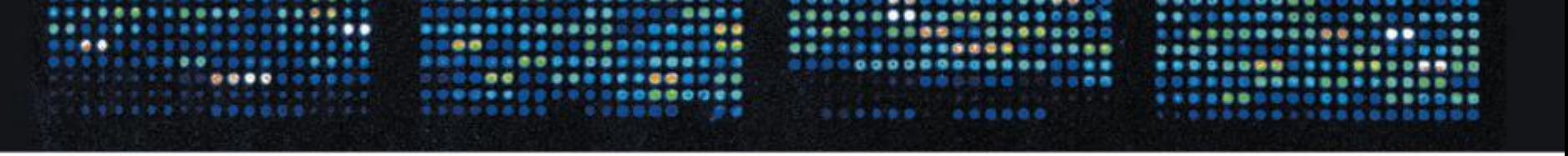
Show me the data!

The potential and power of gene expression analysis using DNA microarrays has lead to the widespread use of this technology. These expression or 'profiling' studies (as they are commonly known) are providing a new and unprecedented view of complex biological systems[1–9]. The com-

Charles M. Perou

Lineberger Comprehensive Cancer Center and Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. Correspondence should be addressed to C.M.P. (e-mail: cperou@med.unc.edu).

Perou CM. Nature Genetics 2001, 29:373.



DNA microarrays allow researchers to analyse the expression of a huge number of genes simultaneously.

GENOMICS

Gene data to hit milestone

With close to one million gene-
researchers can identify diseas

epositories,
y.

BY MONYA BAKER

Purvesh Khatri sits in front of an oversized computer screen, trawling for treasure in a sea of genetic data. Entering the search term 'breast cancer' into a public repository called the Gene Expression Omnibus (GEO), the postdoctoral researcher retrieves a list of 1,170 experiments, representing nearly 33,000 samples and a hoard of gene-expression data that could reveal previously unseen patterns.

That is exactly the kind of search that led Khatri's boss, Atul Butte, a bioinformatician at the Stanford School of Medicine in California, to identify a new drug target for diabetes. After downloading data from 130 gene-expression studies in mice, rats and humans, Butte looked for genes that were expressed at higher levels in disease samples than in controls. One gene was strikingly consistent: *CD44*, which encodes a protein found on the surface of white blood cells, was differentially expressed in 60% of the studies (K. Kodama *et al. Proc. Natl Acad. Sci. USA* **109**, 7049–7054; 2012). The CD44

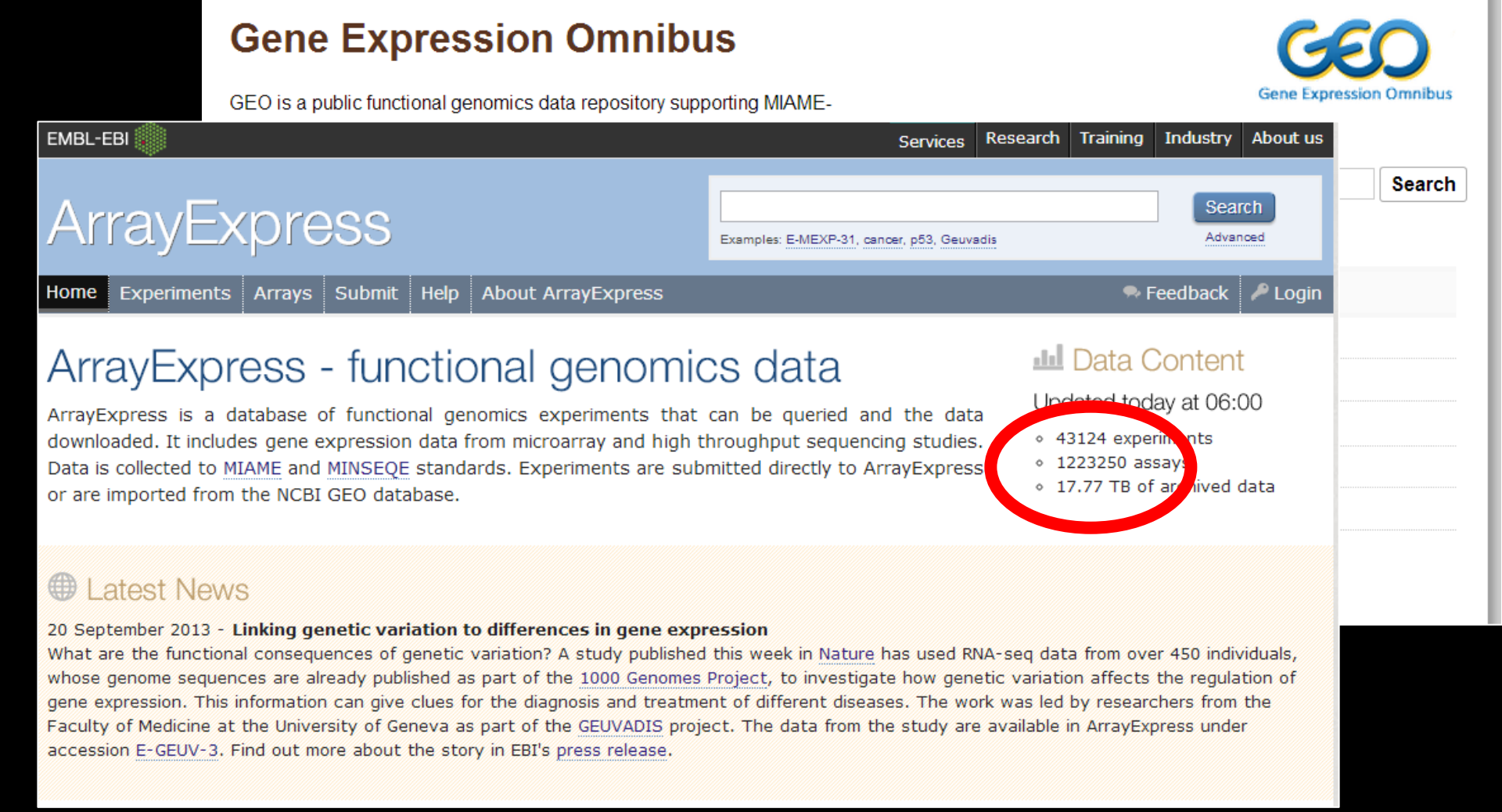
DATA DUMP

The number of gene-expression data sets in publicly available databases has climbed to nearly one million over the past decade.

SOURCES: NIH, EBI

ean Bioinformatics
h, UK. Some time in
umber of deposited
n (see 'Data dump').
dented resource that
sts and speed up pro-
ease. Gene-sequence
tensively, but expres-
plex and can reveal
t active in, say, liver
eased versus healthy
s often look at many
urpose the data sets,
n those posed by the

many data sets are
h harder is working
d. Heather Piwowar,
h the National Evolu-
om the University of
uver, Canada, found
sited in GEO in 2005
7 had been cited by
e rates are certainly



Over 1.2 million microarrays available

Doubles every 2-3 years

Butte AJ. Translational Bioinformatics: coming of age. *JAMIA*, 2008.

NCBI Resources How To Sign in to NCB

GEO DataSets

GEO DataSets breast cancer Search

Save search Advanced Help

Show additional filters

Entry type
DataSets (141)
Series (1582)
Samples (39835)
Platforms (35)

Organism
Select ...

Study type
Expression profiling by array
Methylation profiling by array
More ...

Author
Select ...

Attribute name
tissue
strain
More ...

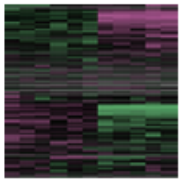
Publication

Display Settings: Summary, 20 per page, Sorted by Default order **Send to:**

<< First < Prev Page 1 of 2080 Next > Last >>

Results: 1 to 20 of 41593

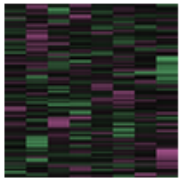
1. Leukemia inhibitory factor effect on Sin3a-silenced MCF7 **breast cancer** cell line



Analysis of SIN3 transcription regulator homolog A (Sin3a)-depleted MCF7 cells stimulated with LIF cytokine to activate signal transducer and activator of transcription 3 (STAT3). STAT3 transcription factor is a potent oncogene. Results provide insight into role of Sin3a in mediating STAT3 activity.

Organism: Homo sapiens
Type: Expression profiling by array, transformed count, 2 agent, 2 genotype/variation sets
Platform: GPL570 Series: GSE35696 11 Samples
Download data: GEO (CEL)
DataSet Accession: GDS4388 ID: 4388
PubMed Full text in PMC Similar studies GEO Profiles Analyze DataSet

2. Co-expression of tyrosine kinase receptors HER2 and HER3 in mammary epithelial cells MCF10A grown in three-dimensional cultures



Analysis of MCF10A mammary epithelial cells expressing HER2, HER3, or HER2/HER3 heterodimer. Co-expression of HER2 and HER3 induced migration and invasion of MCF10A cells. Results provide insight into the role of HER2 and HER3 in **breast cancer**.

Filters: Manage Filters

Top Organisms [Tree]
Homo sapiens *(38361)*
Mus musculus *(3059)*
Rattus norvegicus *(184)*
Canis lupus familiaris *(68)*
Human herpesvirus 8 *(6)*
More...

Find related data
Database: Select
Find items

Search details

```
"breast neoplasms"[MeSH Terms] OR breast cancer[All Fields]
```

Search See more...

Recent activity

NCBI Resources How To Sign in to NCB

GEO DataSets

GEO DataSets breast cancer Search

Save search Advanced Help

Show additional filters

Entry type
DataSets (141)
Series (1582)
Samples (39835)
Platforms (35)

Organism
Select ...

Study type
Expression profiling by array
Methylation profiling by array
More ...

Author
Select ...

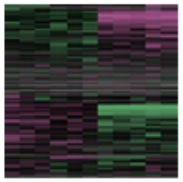
Attribute name
tissue
strain
More ...

Publication

Display Settings: Summary, 20 per page, Sorted by Default order **Send to:**

<< First < Prev Page 1 of 2080 Next > Last >>

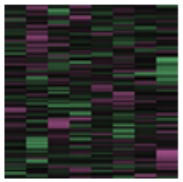
Results: 1 to 20 of 41593



1. Leukemia inhibitory factor effect on Sin3a-silenced MCF7 **breast cancer** cell line

Analysis of SIN3 transcription regulator homolog A (Sin3a)-depleted MCF7 cells stimulated with LIF cytokine to activate signal transducer and activator of transcription 3 (STAT3). STAT3 transcription factor is a potent oncogene. Results provide insight into role of Sin3a in mediating STAT3 activity.

Organism: Homo sapiens
Type: Expression profiling by array, transformed count, 2 agent, 2 genotype/variation sets
Platform: GPL570 Series: GSE35696 11 Samples
Download data: GEO (CEL)
DataSet Accession: GDS4388 ID: 4388
PubMed Full text in PMC Similar studies GEO Profiles Analyze DataSet



2. Co-expression of tyrosine kinase receptors HER2 and HER3 in mammary epithelial cells MCF10A grown in three-dimensional cultures

Analysis of MCF10A mammary epithelial cells expressing HER2, HER3, or HER2/HER3 heterodimer. Co-expression of HER2 and HER3 induced migration and invasion of MCF10A cells. Results provide insight into the role of HER2 and HER3 in **breast cancer**.

Filters: Manage Filters

Top Organisms [Tree]
Homo sapiens *(38361)*
Mus musculus *(3059)*
Rattus norvegicus *(184)*
Canis lupus familiaris *(68)*
Human herpesvirus 8 *(6)*
More...

Find related data
Database: Select
Find items

Search details
```
"breast neoplasms"[MeSH Terms] OR breast cancer[All Fields]
```
Search See more...

Recent activity

Public big data = retroactive crowd-sourcing

Teen develops algorithm to diagnose leukaemia

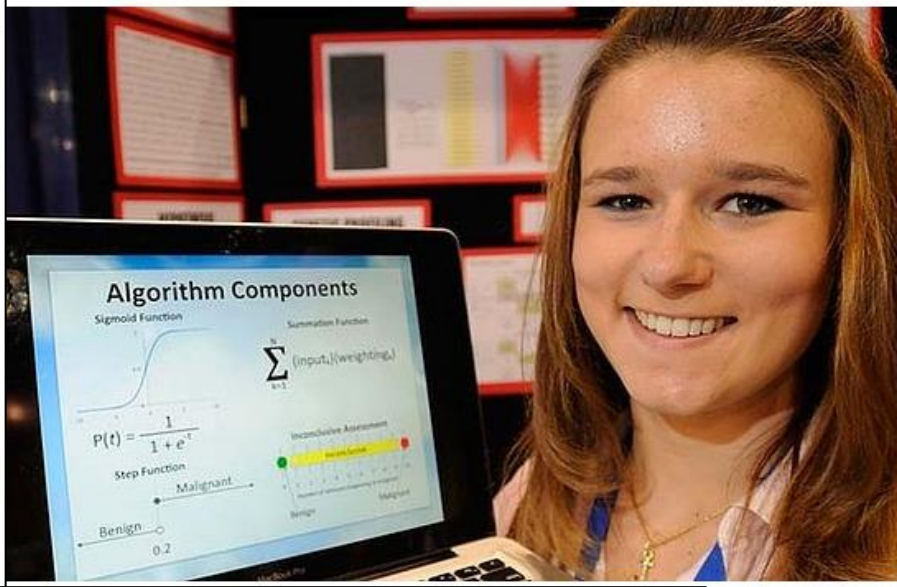
May 22, 2013 - 8:44AM

Read lat

achandran

Recommend 131 Share 0 Share Pin it submit

Print



17-year-old programs artificial 'brain' to diagnose breast cancer

Published July 25, 2012 / InnovationNewsDaily Staff

Print

Email

Share

Like 787

Tweet 303

Share 24

A high school junior has created a computer brain that c
cancer with 99 percent sensitivity.

Seventeen-year-old Brittany Wenger of Sarasota, Fla.,
diagnosing app based on an artificial neural network, ba
program whose structure is inspired by the way brain ce
another. She won grand prize at the Google Science Fa
ceremony held in Palo Alto, Calif. last night (July 23).

Like other artificial intelligence programs, artificial neura
to do by analyzing examples they're given and they perf
more examples. In addition, they're able to detect patte
complex for human brains or other types of programs to
June, Google researchers built a neural network that lea
on the Internet without any outside input.

RELATED STORIES

- MTV reality star Diem Brown chooses to delay cancer treatment to

enger. *Photo: Intel*

er how to diagnos

nswer

our

enger, 18

SCIENCE & NATURE

Jack Andraka, the Teen Prodigy of Pancreatic Cancer

A high school sophomore won the youth achievement Smithsonian American Ingenuity Award for inventing a new method to detect a lethal cancer

REDDIT DIGG STUMBLE EMAIL

By Abigail Tucker

Smithsonian magazine, December 2012, Subscribe

View More Photos »

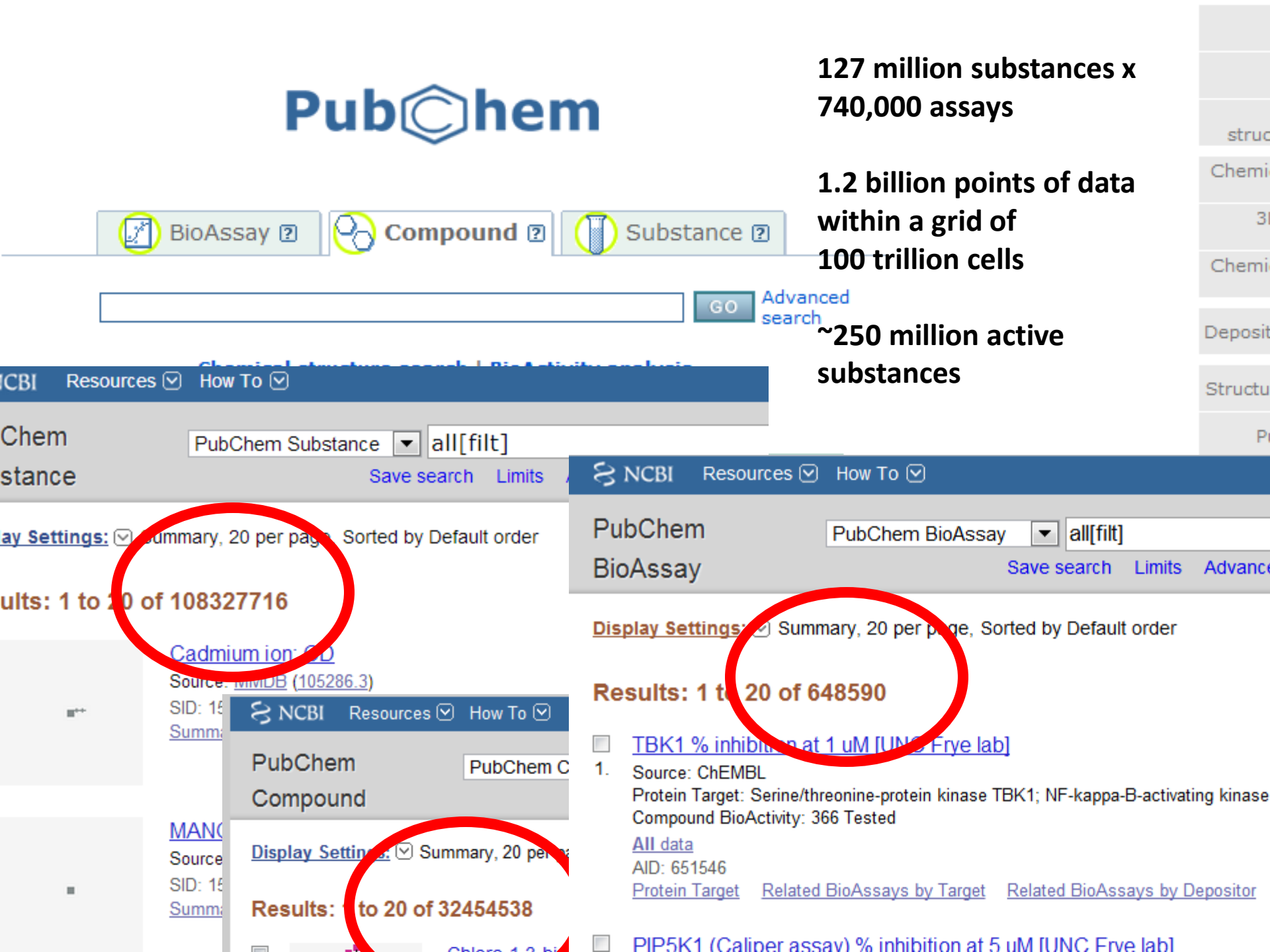
Available Cancer Types	# Cases Shipped by BCR	# Cases with Data	Date Last Updated (mm/dd/yy)
Acute Myeloid Leukemia [LAML]	200	200	6/24/2013
Adrenocortical carcinoma [ACC]	80	0	
Bladder Urothelial Carcinoma [BLCA]	201	184	7/5/2013
Brain Lower Grade Glioma [LGG]	296	271	7/3/2013
Breast invasive carcinoma [BRCA]	1007	961	7/5/2013
Cervical squamous cell carcinoma and endocervical adenocarcinoma [CESC]	163	163	7/5/2013
Colon adenocarcinoma [COAD]	439	425	6/28/2013
Esophageal carcinoma [ESCA]	63	63	7/5/2013
Glioblastoma multiforme [GBM]	514	510	6/28/2013
Head and Neck squamous cell carcinoma [HNSC]	427	376	7/3/2013
Kidney Chromophobe [KICH]	66	66	7/5/2013
Kidney renal clear cell carcinoma [KIRC]	512	512	7/3/2013
Kidney renal papillary cell carcinoma [KIRP]	158	144	6/28/2013
Liver hepatocellular carcinoma [LIHC]	152	128	7/3/2013
Lung adenocarcinoma [LUAD]	500	499	7/3/2013
Lung squamous cell carcinoma [LUSC]	500	494	7/5/2013
Lymphoid Neoplasm Diffuse Large B-cell Lymphoma[DLBC]	18	18	7/3/2013
Mesothelioma [MESO]			
Ovarian serous cystadenocarcinoma [OV]			
Pancreatic adenocarcinoma [PAAD]			
Pheochromocytoma and Paraganglioma [PCPG]			
Prostate adenocarcinoma [PRAD]			
Rectum adenocarcinoma [READ]	169	168	6/28/2013
Sarcoma [SARC]	111	75	7/5/2013
Skin Cutaneous Melanoma [SKCM]	357	336	7/5/2013
Stomach adenocarcinoma [STAD]	343	325	7/3/2013
Testicular Germ Cell Tumors [TGCT]	0	0	

Browse dbGaP

db GaP GENOTYPE and PHENOTYPE

By Studies | By Diseases | Advanced Search

Study	Embargo Release	Details	Participants	Type of St
CIDR: Genome Wide Association Study in Familial Parkinson Disease (PD)	Feb 13, 2009	V D A	1991	Case-cont
Framingham SHARe	Version 1: Oct 19, 2008 Version 2: Feb 01, 2009 Version 3: Jul 08, 2009	V D A	14277	Longitudir
GAIN: Collaborative Association Study of Psoriasis	Aug 13, 2008	V D A	2875	Case-cont
GAIN: Genotyping the 270 HapMap samples for GAIN by Broad		V D A	-	Parent-offsprir
GAIN: Genotyping the 270 HapMap samples for GAIN by Perlegen		V D A	-	Parent-offsprir
GAIN: International Multi-Center ADHD Genetics Project	Mar 26, 2008	V D A	2835	Parent-offsprir
GAIN: Linking Genome-Wide Association Study of Schizophrenia	Version 1: Nov 07, 2008 Version 2: Dec 03, 2008	V D A	5066	Case-cont
GAIN: Major Depression: Stage 1 Genomewide Association in Population-Based Samples	Jul 09, 2008	V D A	3741	Case-cont
GAIN: Search for Susceptibility Genes for Diabetic Nephropathy in Type 1 Diabetes	Jul 09, 2008	V D A	1825	Case-cont
GAIN: Whole Genome Association Study of Bipolar Disorder	Version 1: Nov 25, 2008 Version 2: Dec 01, 2008	V D A	3261	Case-cont
GAW16 Framingham and Simulated Data	Oct 19, 2008	V D A	7130	Longitudir population-b
Genome-wide Association Studies in the Hutterites		V D A	632	Population-b
Genome-wide Association Study of Neuroblastoma		V D A	1032	Case-cont
Genome-wide Study in Amyotrophic Lateral Sclerosis and Controls: First Stage Analysis	Jun 26, 2008	V D A	544	Case-cont
Ischemic Stroke Genetics Study (ISGS)	Jun 26, 2008	V D A	485	Case-cont
Mayo-Perlegen LEAPS (Linked Efforts to Accelerate Parkinson's Solutions) Collaboration	Mar 03, 2008	V D A	1550	Case-cont
NEI Age-Related Eye Disease Study (AREDS)	Jun 11, 2007	V D A	600	Case-cont
NINDS Parkinson's Disease	Oct 12, 2007	V D A	535	Case-cont
NINDS Parkinsonism Study	Oct 12, 2007	V D A	1283	Case-se
NINDS Repository Cerebrovascular Disease/Stroke Study	Jun 26, 2008	V D A	870	Case-se
NINDS Repository Motor Neuron Disease/ALS Study	Jun 26, 2008	V D A	1790	Case-se
NINDS Repository Neurologically Normal Control Collection	Oct 12, 2007	V D A	2723	Control-s
POPRES: Population Reference Sample		V D A	5919	Population sa Control-s
SEARCH GWA Study of Statin-Induced Myopathy		V D A	175	Case-cont
Study of Irish Amyotrophic Lateral Sclerosis (SIALS)		V D A	432	Case-cont
The Finland-United States Investigation of NIDDM Genetics (FUSION) study		V D A	2335	Case-cont
Whole Genome Association Study of Systemic Lupus Erythematosus		V D A	4651	Case-cont



127 million substances x 740,000 assays

1.2 billion points of data within a grid of 100 trillion cells

~250 million active substances

ScienceInsider
Breaking news and analysis from the world of science policy

White House Unveils Long-Awaited Public Access Policy

by Jocelyn Kaiser on 22 February 2013, 5:40 PM | 1 Comment

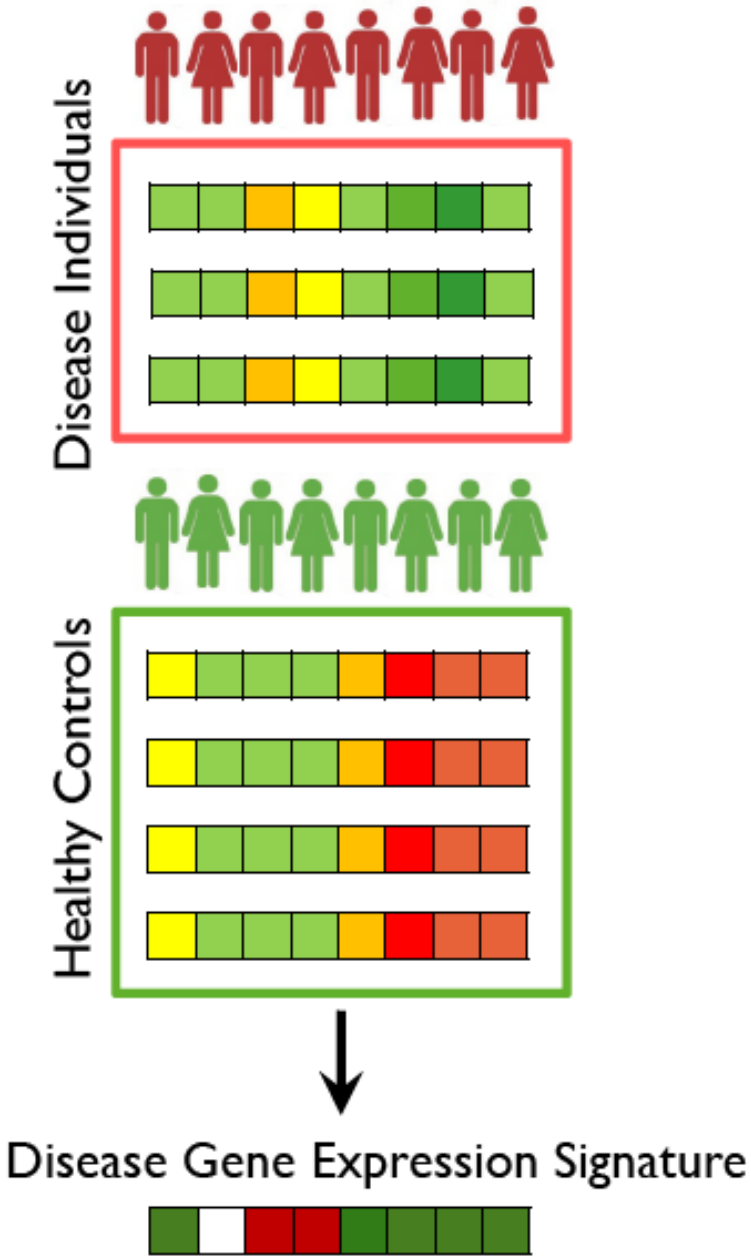
Email Print | +1 6 More PREVIOUS ARTICLE NEXT ARTICLE

In a victory for open access advocates, the White House science office today released a long-awaited policy aimed at sharing the results of federally funded research with the public. The policy will require that science agencies make papers that they fund freely available online within 12 months after the results appear in a journal.

The policy follows several years of consultations and a petition to the White House from open access advocates last year. It appears to have found a middle ground between the two sides in a decadelong debate over so-called "open access"—the issue of whether and when scientific papers funded with taxpayer dollars should be available, for free, to the public. Traditionally, publicly funded scientists journals that charge fees for access to the papers. That system has b the advent of digital technologies and new research funding models. M resisted complete and immediate open access, arguing that it will des survive.

The new federal directive is a "landmark" and a "watershed moment," Scholarly Publishing and Academic Resources Coalition, an open acc American Publishers, which has called some public access mandates of journals, said the directive "outlines a reasonable, balanced resoluti research funded by federal agencies."

John Holdren, Director of the Office of Science and Technology Policy, "has directed Federal agencies with more than $100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and requiring researchers to better account for and manage the digital data resulting from federally funded scientific research."


Disease Individuals
Healthy Controls
Disease Gene Expression Signature

Questioning standardization in science

Richard Paylor

Some scientists suggest that environmental standardization may lead to spurious findings. The implication from this hypothesis will likely be controversial.

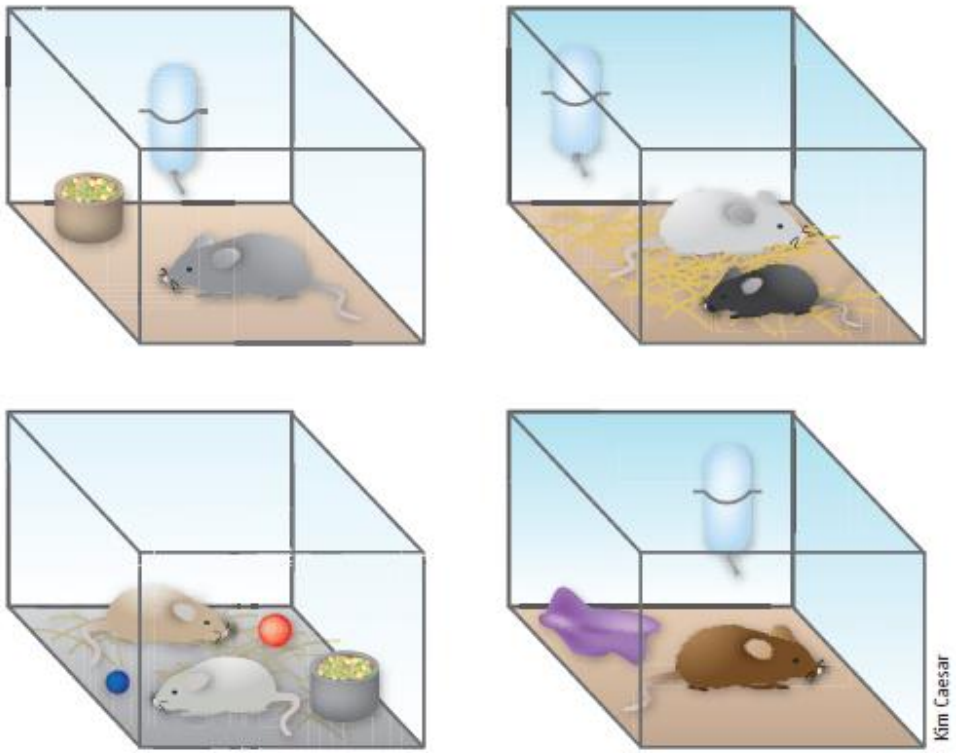


Figure 1 | Heterogenous conditions lead to more reproducible behavioral results.

Richard Paylor is in the Department of Molecular and Human Genetics, and Neuroscience, Baylor College of Medicine, Houston, Texas, USA.
e-mail: paylor@bcm.edu


Disease Individuals
Healthy Controls
Disease Gene Expression Signature

browse by disease

— A —

— I —

— R —

browse by disease

A

- Anal Cancer
- Anemia
- Asthma

B

- Bladder Cancer
- Brain Cancer
- Breast Cancer

C

- Carcinoid
- Cervical Cancer
- Chronic Obstructive Pulmonary Disease

I

- Idiopathic Pulmonary Fibrosis

K

- Kidney Cancer

L

- Leukemia
- Liver Cancer
- Lung Cancer

M

- Melanoma
- Monoclonal Gammopathy

R

- Rheumatoid Arthritis

S

- Sarcoidosis
- Scleroderma
- Systemic Lupus Erythematosus

T

- Testicular Cancer

U

- Uterine Cancer

Search Results

You've Selected:

Disease: Leukemia (X)

Clear All Selections

Category

Products (21)

Tissue

Bone Marrow (9)
Peripheral Blood (12)

Cell Type

B Cells CD19 (2)
B Cells Negative Selection (2)
Buffy Coat (1)
CD45 (2)
Fresh (2)
Mononuclear Cells (2)
Plasma (1)
Serum (1)
Special Processing (2)
T Cells CD3 (2)
T Cells Negative Selection (2)
Viable Plated Cells (2)

Units

0.3mL (1)
0.5 million cells (10)
0.5mL (2)
1 unit (2)
5.0 million cells (2)

Price

$0.00 - $1,000.00 (17)
$1,000.00 - $2,000.00 (2)

Leukemia

21 Items Previous | 1 | 2 | Next View as:

15 Items Per Page | Sort By...

bma

Bone Marrow | B Cells, Negative Selection | Leukemia
SKU: BMA-BCE-LE
$500.00

bma

Bone Marrow | B Cells, CD19 | Leukemia
SKU: BMA-CD19-LE
$500.00

bma

Bone Marrow | T Cells, CD3 | Leukemia
SKU: BMA-CD3-LE
$500.00

bma

Bone Marrow | CD45 | Leukemia
SKU: BMA-CD45-LE
$500.00

bma

Bone Marrow | Fresh | Leukemia
SKU: BMA-FRE-LE
$2,500.00

bma

Bone Marrow | Mononuclear Cells | Leukemia
SKU: BMA-MON-LE
$750.00

bma

Bone Marrow | Special Processing | Leukemia
SKU: BMA-SPE-LE
$500.00

bma

Bone Marrow | T Cells, Negative Selection | Leukemia
SKU: BMA-TCE-LE
$500.00

Search Results

You've Selected:

Disease: Leukemia (X)

Clear All Selections

Category

Products (21)

Tissue

Bone Marrow (9)
Peripheral Blood (12)

Cell Type

B Cells CD19 (2)
B Cells Negative Selection (2)
Buffy Coat (1)
CD45 (2)
Fresh (2)
Mononuclear Cells (2)
Plasma (1)
Serum (1)
Special Processing (2)
T Cells CD3 (2)
T Cells Negative Selection (2)
Viable Plated Cells (2)

Units

0.3mL (1)
0.5 million cells (10)
0.5mL (2)
1 unit (2)
5.0 million cells (2)

Price

$0.00 - $1,000.00 (17)

Leukemia

21 Items Previous 1 2 Next View as:

15 Items Per Page | Sort By...

pbl

Peripheral Blood | Mononuclear Cells | Leukemia
SKU: PBL-MON-LE
$500.00

pbl

Peripheral Blood | Plasma | Leukemia
SKU: PBL-PLA-LE

pbl

Peripheral Blood | Serum | Leukemia
SKU: PBL-SER-LE
$55.00

pbl

Peripheral Blood | Special Processing | Leukemia
SKU: PBL-SPE-LE
$500.00

pbl

Peripheral Blood | T Cells, Negative Selection | Leukemia
SKU: PBL-TCE-LE
$600.00

pbl

Peripheral Blood | Viable Plated Cells | Leukemia
SKU: PBL-VPC-LE
$1,000.00

Previous 1 2 Next

pbl

Peripheral Blood | Plasma | Leukemia

SKU: PBL-PLA-LE

$55.00

Be the first to review this product

Quick Overview: 0.5mL Plasma specimen collected in K2EDTA tube and stored in 1.0mL cryovial. Sample stored at -80C and shipped on dry ice.

# of Samples per Patient

-- Please Select --
-- Please Select --
1
2 +$55.00
3 +$110.00
4 +$165.00
5 +$220.00
6 +$275.00
7 +$330.00

Disease subtype

-- Please Select --

Units

-- Please Select --

add to cart

Add to Saved Items | Email A Colleague

Overview › | Product Reviews ›

Additional Information

SKU	PBL-PLA-LE
Treatment Status	Any Treatment, Pre Treatment, Post Treatment, Active Treatment, Recurrent/Refractory Disease, Remission

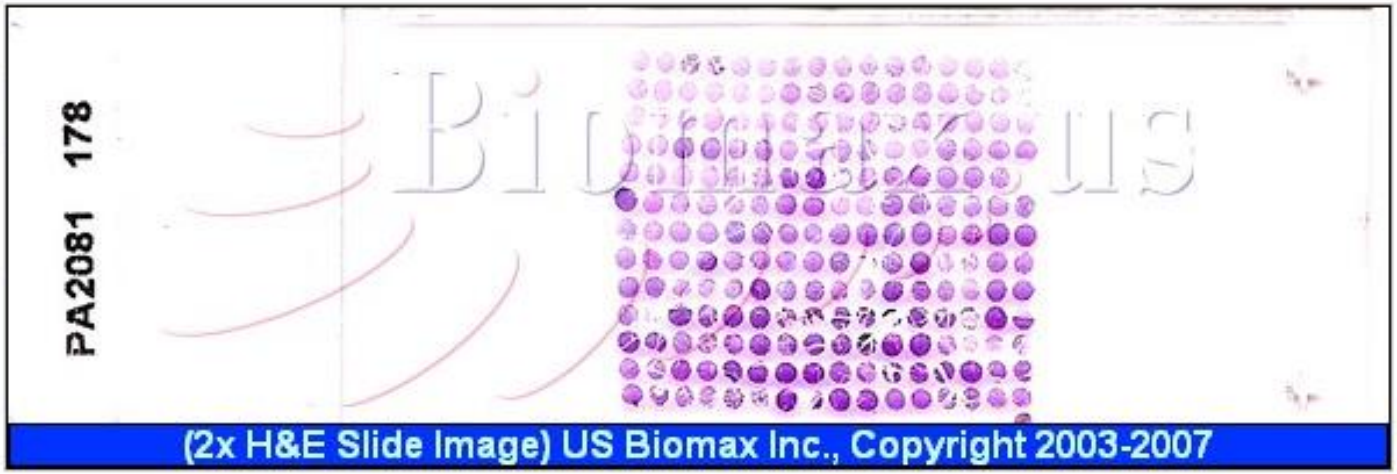
PA2081

Pancreas disease spectrum (pancreatic cancer progression) tissue array, 101 cases/208 cores

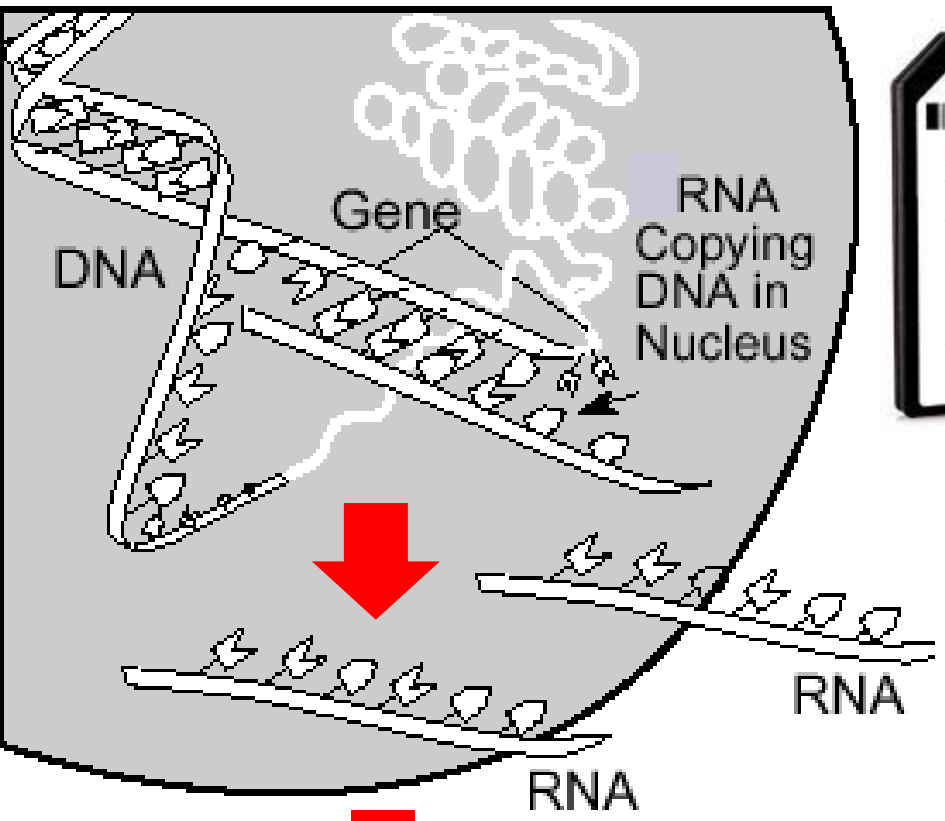
Unstained: $395.00
Trial (PA2081t*): $105.00
H&E (PA2081s*): $395.00

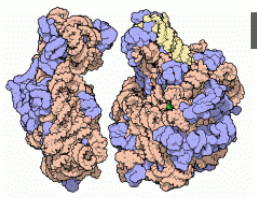
Overview | Specification Sheet | Core Image | Overlapping Cases | Print | Export for Excel

Microarray Panel:	Pancreatic carcinoma tissue microarray, containing 42 cases of ductal adenocarcinoma, 3 adenosquamous carcinoma, 1 islet cell carcinoma, 6 metastatic carcinoma, 10 islet cell tumor, 2 hyperplasia, 10 inflammation, 20 adjacent normal tissue and 10 normal tissue from autopsy, duplicated cores per case
Cores:	208
Cases:	101
Layout:	16 cols × 13 rows
Core Diameter:	1 mm
Thickness:	5 µm
Quality Control:	Anti-Actin confirmed
Applications:	Routine histology procedures including ImmunoHistochemistry (IHC) and In Situ Hybridization (ISH), protocols which can be found on our support page.
Notes:	Slides are recommended to be baked at 60°C for 2 hours and to undergo antigen retrieval procedures before regular IHC or ISH procedures. Some slides are covered by a thin layer of paraffin to prevent oxidization and decay.

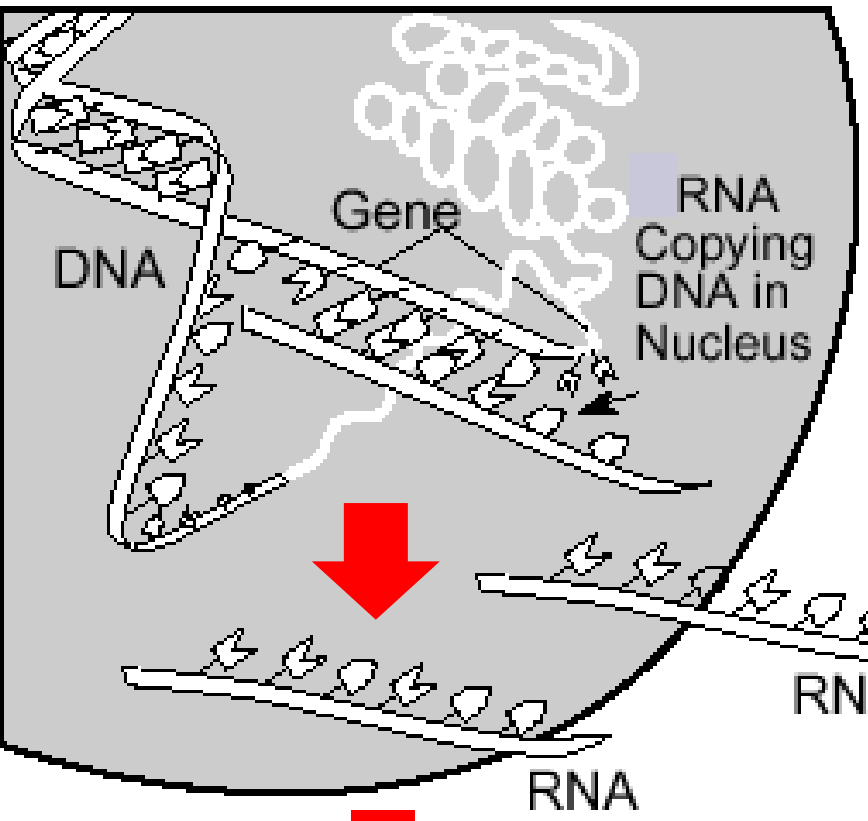


Cancer markers



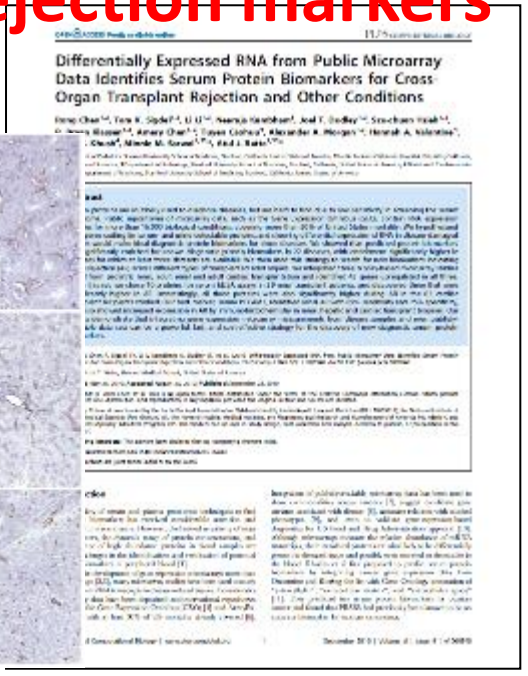


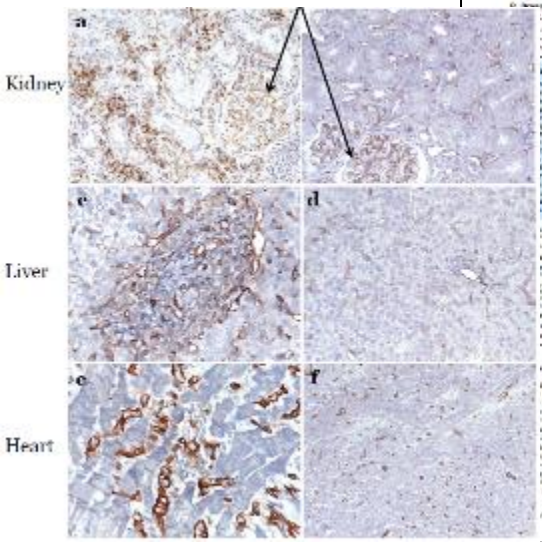
Protein

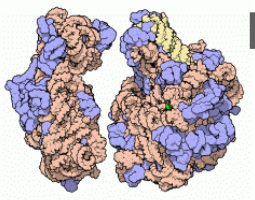


Cancer markers

Transplant Rejection markers







Protein

Preeclampsia: large cause of maternal and fetal death

- Incidence
  - 5-8% of all pregnancies in the U.S. and worldwide
  - 4.1 million births in the U.S. in 2009
  - Up to 300K cases of preeclampsia annually in the U.S.
- Mortality
  - Responsible for 18% of all maternal deaths in the U.S.
  - Maternal death in 56 out of every 100,000 live births in US
  - Neonatal death in 71 out of every 100,000 live births in US
- Cost
  - $20 billion in direct costs in the U.S annually
  - Average hospital stay of 3.5 days

Linda Liu
Matt Cooper
Bruce Ling

tion, sample and factor annotations [clear]

maturity OR preeclampsia OR eclampsia

s data only

viewer login

Advanced query syntax

ArrayExpress Browser Help

Filter on [reset]

All species | All arrays | All assays by molecule | by | All technologies

Displa 50 De

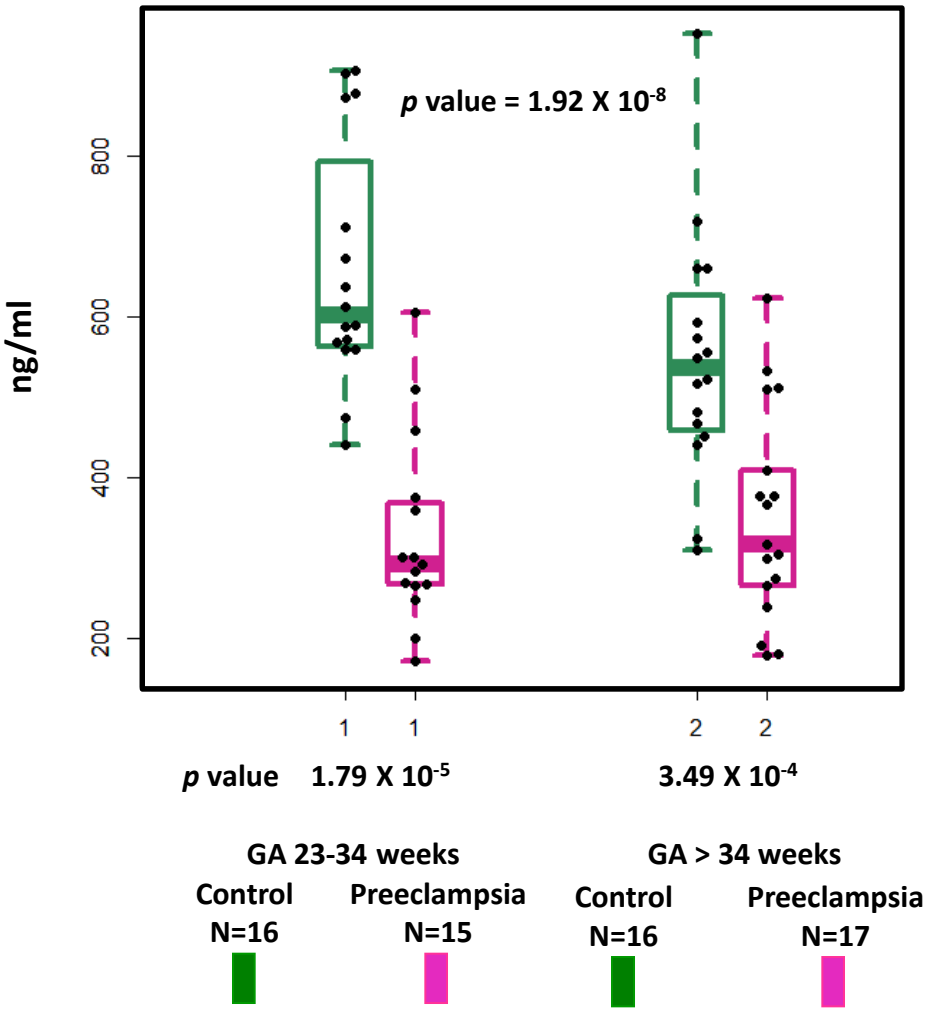
	Title	Assays	Species
89	Differential expression of microRNAs in the placentae of Chinese patients with severe pre-eclampsia	8	Homo sapie
87	Latexin is downregulated in human upper digestive track tumors and exhibits tumor suppressor potential	9	Homo sapie
98	Effects of estrogen and progesterone on glycosyltransferase expression in mice	42	Mus muscul
26	Gene expression profiling of placenta from women with early- and late-onset preeclampsia reveals differences in angiogenesis	15	Homo sapie
83	Gene expression profile analysis in the mid-gestation human intestine discloses greater immaturity of the colon as compared to ileum	12	Homo sapie
06	Transcriptional Profiling of Human Placentas from Pregnancies Complicated by Preeclampsia Reveals Disregulation of Sialic Acid Acetylesterase and Immune Sign...	60	Homo sapie
10	Human Trophoblast Responses to P. gingivalis Infection	8	Homo sapie
17	The miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor	6	Mus muscul
29	Gene expression profiling for placentas from pre-eclamptic, unexplained FGR and normal pregnancies.	24	Homo sapie
88	Placental gene expression in severe preeclampsia.	43	Homo sapie
90	A Dysregulated Proinflammatory Status in Preterm School-aged Children with PVL-induced Spastic Cerebral Palsy	10	Homo sapie
16	Temporal development of necrotizing enterocolitis in the preterm pig	26	Sus scrofa
15	Amniotic fluid and necrotizing enterocolitis in preterm pigs	17	Sus scrofa
98	Gene expression microarray analysis of early oxygen-induced retinopathy (OIR)	10	Rattus norv
09	Differentially expressed mRNA transcripts in the placenta delivered by preterm versus term spontaneous labour	10	Homo sapie
99	Comparative analysis of the human and mouse placental transcriptome and proteome	2	Homo sapie
55	Transcription profiling of mouse placental labyrinth and human villus tree	4	Homo sapie
	Transcription profiling of human decidua basalis to identify pre-eclampsia susceptibility genes	104	Homo sapie
22	Severe Preeclampsia-Related Changes in Gene Expression at the Maternal-Fetal Interface Include Siglec-6 and Pappalysin-2	46	Homo sapie
67	Transcription profiling of human chorionic villus samples (CVS) from preeclampsia patients	12	Homo sapie
75	Transcription profiling of human choriocarcinoma cells overexpressing STOX1 mimicks transcriptional alterations observed in preeclamptic placentas	2	Homo sapie
9	Transcription profiling of human 27 non-labored normal preterm and 9 non-labored normal term basal plate samples to investigate gene expression at the human...	72	Homo sapie
9	Transcription profiling of human myometrium from samples obtained at term (n=6) or preterm (n=6) with and without labor	12	Homo sapie
	MicroRNA profiling of human chorioamniotic membranes from women at term in labor, women at term not in labor and women who deliverd preterm	30	Homo sapie
16	Gene expression profiling indicates inflammatory pathways involved in IUGR due to placental insufficiency	16	Homo sapie
4	Transcription profiling of human placentae of different gestational ages: an opru network and uw scor study	12	Homo sapie
9	Transcription profiling of mouse uterus from _gravidl_d18 wild type vs. Cox-1 knockout animal	8	Mus muscul
3	Transcription profiling of human adipose tissue, decidua and placenta from patients with preeclampsia vs. unaffected indviduals reveals dysregulation of the circul...	6	Homo sapie
0	Transcription profiling of human trophoblast cell cultures	2	Homo sapie
7	Transcription profiling of placenta from women who underwent a normal pregnancy and from those suffering from early or late onset pre-eclampsia	14	Homo sapie
92	Transcription profiling of mice	12	Mus muscul
0	Transcription profiling of decidua basalis from women with normal pregnancies or preeclampsia and/or fetal growth restriction	35	Homo sapie

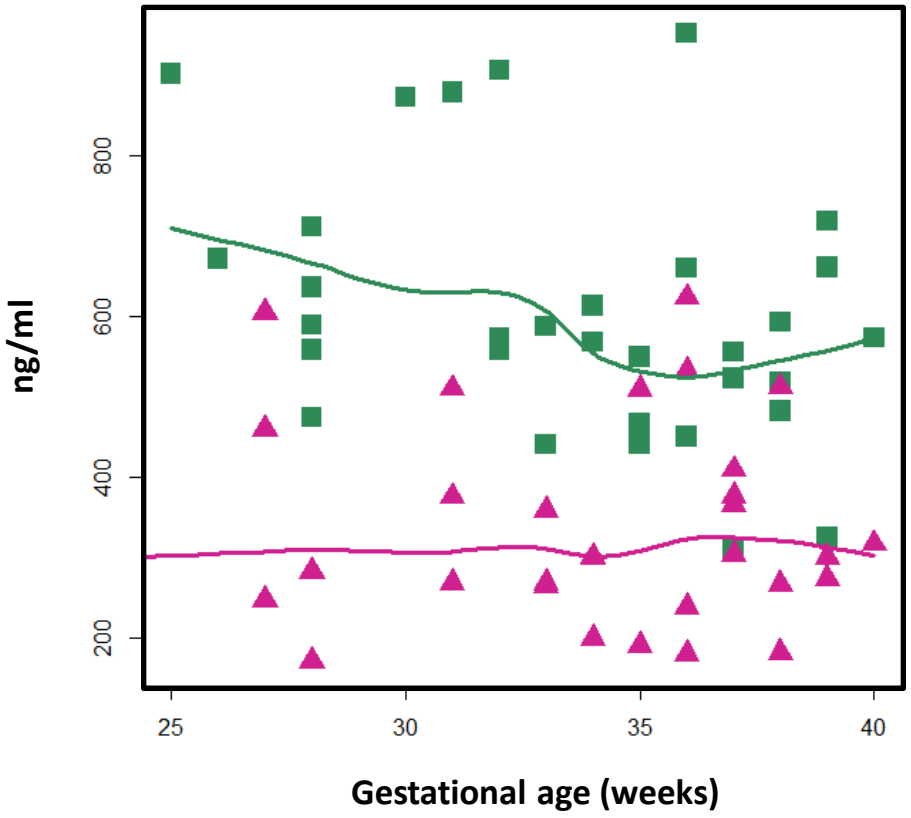
42 experiments, 1221 assays.

march of dimes®
prematurity research center
at STANFORD University School of Medicine

Linda Liu
Bruce Ling

New markers for preeclampsia





Carmenta BIOSCIENCE

Need a diagnostic for preeclampsia

Public big data available

March of Dimes Center for Prematurity Research

Data analyzed, diagnostic designed

SPARK grant ($50k)

Life Science Angels, other seed investors ($2 million)

STOCK WATCH Express, Wet Seal, Avago Jump

Carmenta Bioscience Secures Over $2 Million in Oversubscribed Seed Financing

Camille Samuels Accepts Seat on Carmenta Board of Directors

BusinessWire **Press Release:** Carmenta Bioscience, Inc. – Wed, Aug 7, 2013 9:05 AM EDT

RELATED CONTENT

PALO ALTO, Calif.--(BUSINESS WIRE)--

Carmenta Bioscience, Inc., a privately held medical technology

BUSINESS | 2/10/2012 @ 7:41AM | 173,575 views

The Truly Staggering Cost Of Inventing New Drugs

83 comments, 66 called-out + Comment Now + Follow Comments

During the Super Bowl, a representative of the pharmaceutical company Eli Lilly posted the on the company's corporate blog that the average cost of bringing a new drug to market is $1.3 billion, a price that would buy 371 Super Bowl ads, 16 million official NFL footballs, two pro football stadiums, pay of almost all NFL football players, and every seat in every NFL stadium for six weeks in a row. This is, of course, ludicrous.

Image by AFP/Getty Images via @daylife

The average drug developed by a major pharmaceutical company costs at least $4 billion, and it can be as much as $11 billion.

Research Spending Per New Drug

Company	Ticker	Number of drugs approved	R&D Spending Per Drug ($Mil)	Total R&D Spending 1997-2011 ($Mil)
AstraZeneca	AZN	5	11,790.93	58,955
GlaxoSmithKline	GSK	10	8,170.81	81,708
Sanofi	SNY	8	7,909.26	63,274
Roche Holding AG	RHHBY	11	7,803.77	85,841
Pfizer Inc.	PFE	14	7,727.03	108,178
Johnson & Johnson	JNJ	15	5,885.65	88,285
Eli Lilly & Co.	LLY	11	4,577.04	50,347
Abbott Laboratories	ABT	8	4,496.21	35,970
Merck & Co Inc	MRK	16	4,209.99	67,360
Bristol-Myers Squibb Co.	BMY	11	4,152.26	45,675
Novartis AG	NVS	21	3,983.13	83,646
Amgen Inc.	AMGN	9	3,692.14	33,229

Sources: InnoThink Center For Research In Biomedical Innovation; Thomson Reuters Fundamentals via FactSet Research Systems

Pharma Summits Patent Cliff in 2012; $290B in Sales at Risk Through 2018

By Nuala Moran

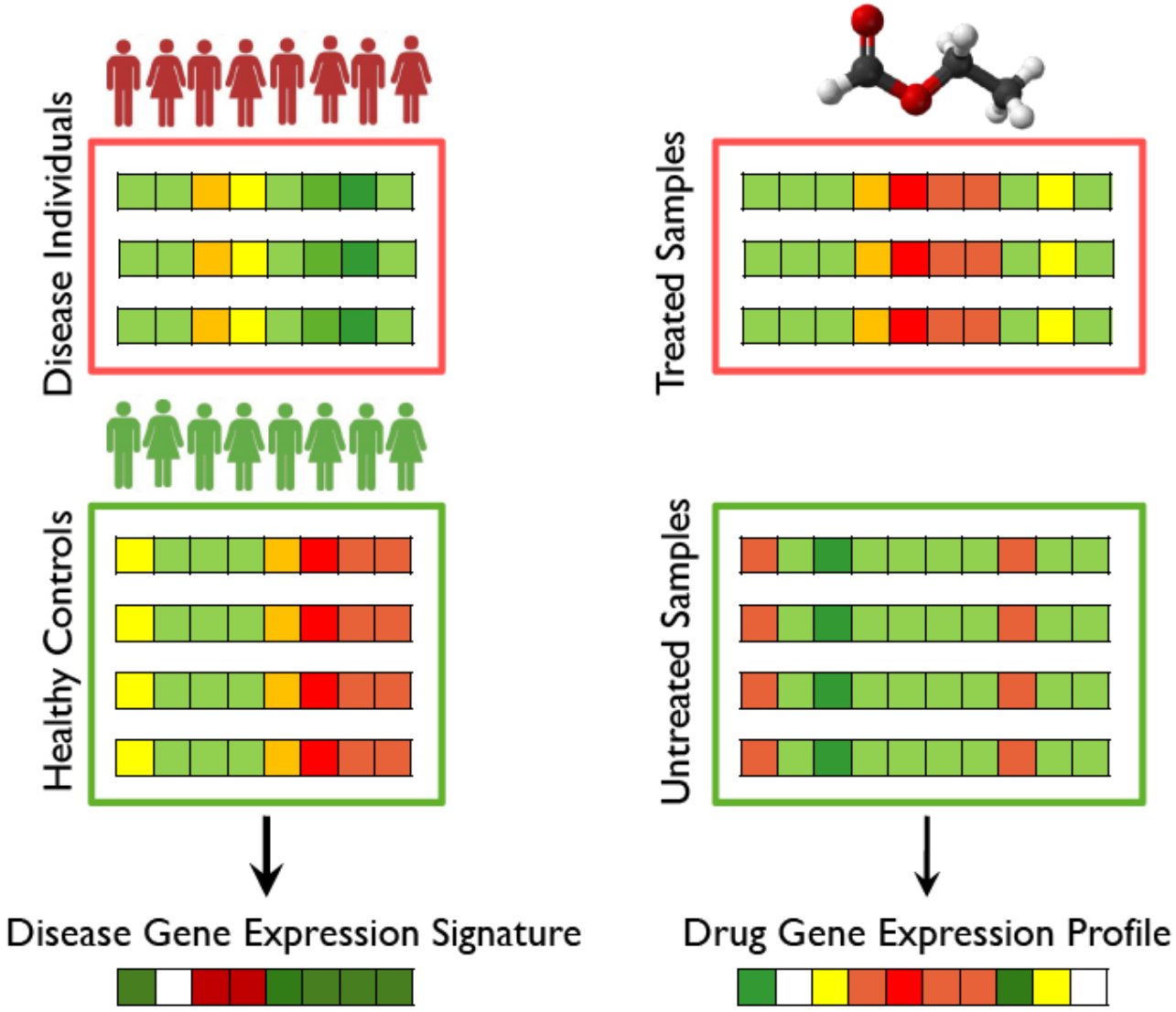
Staff Writer

BOSTON – 2012 may be the year that the patent cliff reaches its height – with $33 billion of sales at risk – but the impact of loss of exclusivity will continue to reverberate across the decade, with more than $290 billion of prescription drugs sales due to be exposed to generic competition between now and 2018.

"This is the worst year, but it will also be bad in succeeding years," said Jonathan de Pass, founder and CEO of EvaluatePharma, the consulting firm that compiled the data. The somewhat depressing conclusions of the report, World Preview 2018, were discussed at BIO 2012 on Tuesday, as the largest partnering-fest of the year got into its swing.

In the past 10 years, a huge amount of money has been thrown at acquisitions and the restructuring of R&D, in an attempt to replace the revenues that are under threat from patent expiries. Over the same time, EvaluatePharma estimated that $1.1 trillion has been invested in R&D in a bid to revitalize pipelines.

Pharma industry executives "have blown an awful lot of cash" in a bid to swerve around the patent cliff. Unfortunately, it looks as if "the response is not sufficient," de Pass said. The forecast returns from new drugs in 2018 "are not that great."



Lamb J, ..., Golub TR. *Science*, 2006.
Sirota M, Dudley JT, ..., Sweet-Cordero A, Sage J, Butte AJ. *Science Translational Medicine*, 2011.

assay depot

the marketplace for pharmaceutical research services

All the Services you need

Researchers around the world use Assay Depot to find the latest research services

Take the Tour

Search... Search or click on a button below to browse research services

biology chemistry dmpk pharmacology toxicology

assay depot
the marketplace for pharmaceutical research services

biology | chemistry | dmpk | pharmacology | toxicology

Bone Models

Bone Metastases
Osteoarthritis
Osteoporosis

Cardiovascular Models

Atrial Arrhythmias
Coronary Artery Disease
Hypertension
Ischemia
Myocardial Infarction
Restenosis
Ventricular Tachycardia

Dermatology Models

Acne
Atopic Dermatitis
Hair Growth
Lupus
Psoriasis
Rosacea
Skin Graft
Wound Healing

Diabetes Models

BB/W Rats
Food Intake
Goto-Kakizaki Rats
Non Obese Diabetic Mice
Obese Mice
Primate Diabetes
Streptozotocin Mice
Streptozotocin Rats
More...

Genitourinary Models

Chronic Kidney Disease
Cystometry
Endometriosis
IGA Glomerulonephritis
Interstitial Cystitis
Spinalized Rats

In Vitro Models

In Vitro Bone Models
In Vitro CVD Models
In Vitro Diabetes Models
In Vitro Eye Models
In Vitro Oncology Models
In Vitro Skin Models

In Vivo Technologies

Cognition
EEG
Electrophysiology
Imaging
Microdialysis

Infectious Disease

Bacterial Infection
Dengue Virus
Hepatitis C Virus
Influenza
LCMV Mouse
Malaria

Inflammation Models

Arthritis
Delayed Type Hypersens
Edema
Hemophilia
Irritable Bowel Disease
Irritant
LPS Acute Response
Mucositis
More...

Neurological Models

Alzheimer's Disease
Anxiety
Behavioral Tests
Cerebral Palsy
Circadian Profiling
Depression
Epilepsy
Olfactory Testing
More...

Oncology Models

Angiogenesis
Cachexia

Ophthalmic Models

Cataract
Corneal Dystrophy

Otology Models

Hearing Loss
Meniere's Disease

Pain Models

General Pain
Inflammatory Pain

Respiratory Models

Ascaris Lung Allergy
Cough

ob/ob Diabetes Model - 16 Mice

Service Description

Provider: Links Biosciences is a US company with laboratories in Hangzhou, China. The laboratory has been offering exploratory (non-GLP) pharmacology services to US and Chinese biopharma since 2004.

Background: The obese mutant mouse model was first reported by Ingalls A *et al* from the Jackson Laboratory in 1951 (Obese, a New Mutation in the House Mouse [164 KB]). The obese mouse resulted from a spontaneous mutation in a gene that was named *ob* in the V stock. Mice homozygous for the obese spontaneous mutation, (Lep^ob^; commonly referred to as *ob* or *ob/ob*), are first recognizable at about 4 weeks of age. Homozygous mutant mice gain weight rapidly and may reach three times the weight of wild-type controls. In addition to obesity, mutant mice exhibit hyperphagia, a diabetes-like syndrome of hyperglycemia, glucose intolerance, elevated plasma insulin, subfertility, impaired wound healing, and an increase in hormone production from both pituitary and adrenal glands. Friedman J *et al* reported leptin in 1994, and demonstrated that leptin, the product of the *ob* gene, was produced in white adipose tissue and served as the peripheral signal to the central nervous system of nutritional status.

Service Details: This service offers a 28 day db/db mouse model of T2DM and obesity. Customer has various options that are conveyed to Links Biosciences using a Service Order Form. Customer assigns up to 16 mice to

assay depot

biology | chemistry | dmpk | pharmacology | toxicology

Home » Pharmacology » Diabetes and Obesity » Obese Mice

ob/ob Diabetes Model - 16 Mice

$9,000.00 USD
per service

9 week
turn around time

Provided By
Links Biosciences

Request Info

Add to Cart

SHARE

Service Description

Provider: Links Biosciences is a US company with laboratories in Hangzhou, China. The laboratory has been offering exploratory (non-GLP) pharmacology services to US and Chinese biopharma since 2004.

Background: The obese mutant mouse model was first reported by Ingalls A *et al* from the Jackson Laboratory in 1951 (Obese, a New Mutation in the House Mouse [164 KB]). The obese mouse resulted from a spontaneous mutation in a gene that was named *ob* in the V stock. Mice homozygous for the obese spontaneous mutation, (Lep^ob^; commonly referred to as *ob* or *ob/ob*), are first recognizable at about 4 weeks of age. Homozygous mutant mice gain weight rapidly and may reach three times the weight of wild-type controls. In addition to obesity, mutant mice exhibit hyperphagia, a diabetes-like syndrome of hyperglycemia, glucose intolerance, elevated plasma insulin, subfertility, impaired wound healing, and an increase in hormone production from both pituitary and adrenal glands. Friedman J *et al* reported leptin in 1994, and demonstrated that leptin, the product of the *ob* gene, was produced in white adipose tissue and served as the peripheral signal to the central nervous system of nutritional status.

Service Details: This service offers a 28 day db/db mouse model of T2DM and obesity. Customer has various options that are conveyed to Links Biosciences using a Service Order Form. Customer assigns up to 16 mice to

Validation methods are increasingly commoditized

Scroll down to browse a list of available research models for **Type I and Type II diabetes, hyperglycemia, insulin resistance, diet-induced obesity and related diseases**. Use the filters on the left to refine the list and then click on any listing to view technical information or to ask a question.

Click on the Vendors tab to view a complete list of CROs that offer diabetes and obesity pharmacology models.

VIEW SERVICES | VIEW VENDORS | **133 results** | get help

Search Filters

Diabetes and Obesity

- BB/W Rats
- Food Intake
- Goto-Kakizaki Rats
- Non Obese Diabetic Mice
- Obese Mice
- Obese Primates
- Primate Diabetes
- Streptozotocin Mice
- Streptozotocin Rats
- db/db Diabetic Mice
- fa/fa Zucker Diabetic Rats

Certifications help

- GLP (48)
- AAALAC (28)
- GMP (20)
- ISO 9001 (7)
- GCP (7)
- FDA (5)
- USDA (4)

more

Locations

United States (64)

Univ. of Maryland School of Medicine Obesity and Diabetes Research Center

University of Maryland School of Medicine Obesity and Diabetes Research Center focuses on research of obesity, diabetes, and aging in nonhuman primates.

vendor info | Add

Transgenic Rabbit Models

Transgenic Rabbit Models offers transgenic rabbit models for the study of atherosclerosis, ophtalmology, hypertrophic myopathies, diabetes, obesity, hemostasis, respiratory diseases, AIDS, and cancer.

vendor info | Add

Ophthy-DS

Ophthy-DS offers ophthalmic model services for macular degeneration, diabetes, uveitis, and dry eye.

vendor info | Add

PharmaNess

PharmaNess offers pharmacokinetics, pharmacodynamics, formulations, behavioral assay, in vivo screening, ex vivo screening, microscopy, stereology and histology staining services.

vendor info | Add

Wisconsin National Primate Research Center

Wisconsin National Primate Research Center focuses on research of regenerative medicine, reproduction, immunology, virology, aging, and metabolic diseases.

vendor info | Add

Selected Vendors

Scroll down to browse a list of available research models for **Type I and Type II diabetes, hyperglycemia, insulin resistance, diet-induced obesity and related diseases**. Use the filters on the left to refine the list and then click on any listing to view technical information or to ask a question.

Click on the Vendors tab to view a complete list of CROs that offer diabetes and obesity pharmacology models.

VIEW SERVICES | VIEW VENDORS | 133 results | get help

Search Filters

Diabetes and Obesity

- BB/W Rats
- Food Intake
- Goto-Kakizaki Rats
- Non Obese Diabetic Mice
- Obese Mice
- Obese Primates
- Primate Diabetes
- Streptozotocin Mice
- Streptozotocin Rats
- db/db Diabetic Mice
- fa/fa Zucker Diabetic Rats

Certifications help

- GLP (48)
- AAALAC (28)
- GMP (20)
- ISO 9001 (7)
- GCP (7)
- FDA (5)
- USDA (4)

more

Locations

United States (64)

Univ. of Maryland School of Medicine Obesity and Diabetes Research Center

University of Maryland School of Medicine Obesity and Diabetes Research Center focuses on research of obesity, diabetes, and aging in nonhuman primates.

vendor info | Add

Transgenic Rabbit Models

Transgenic Rabbit Models offers transgenic rabbit models for the study of atherosclerosis, ophtalmology, hypertrophic myopathies, diabetes, obesity, hemostasis, respiratory diseases, AIDS, and cancer.

vendor info | Add

Ophthy-DS

Ophthy-DS offers ophthalmic model services for macular degeneration, diabetes, uveitis, and dry eye.

vendor info | Add

PharmaNess

PharmaNess offers pharmacokinetics, pharmacodynamics, formulations, behavioral assay, in vivo screening, ex vivo screening, microscopy, stereology and histology staining services.

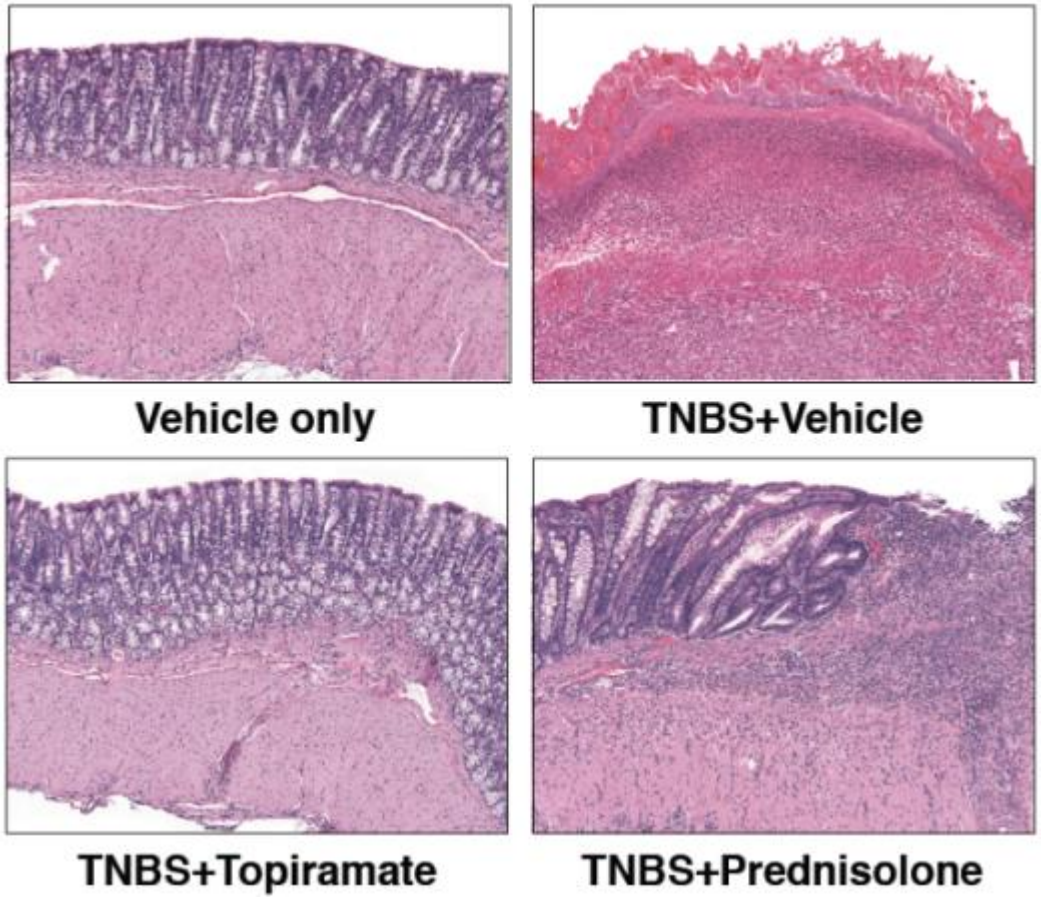
vendor info | Add

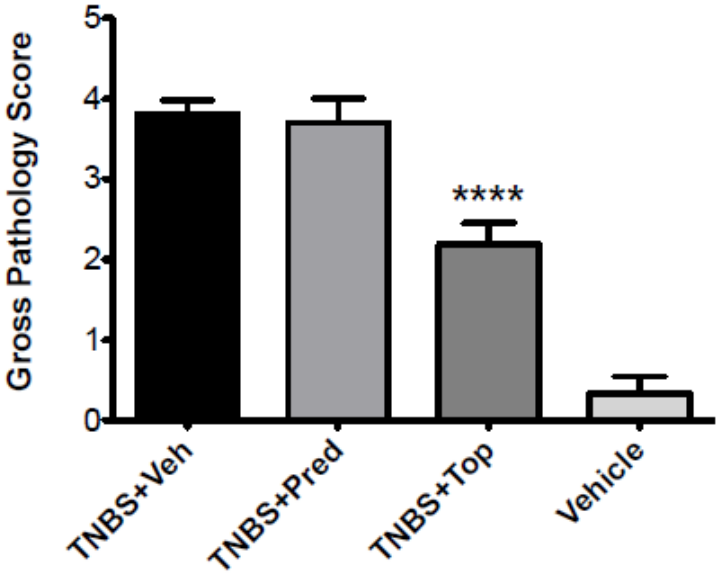
Wisconsin National Primate Research Center

Wisconsin National Primate Research Center focuses on research of regenerative medicine, reproduction, immunology, virology, aging, and metabolic diseases.

vendor info | Add

Anti-seizure drug works against a rat model of inflammatory bowel disease

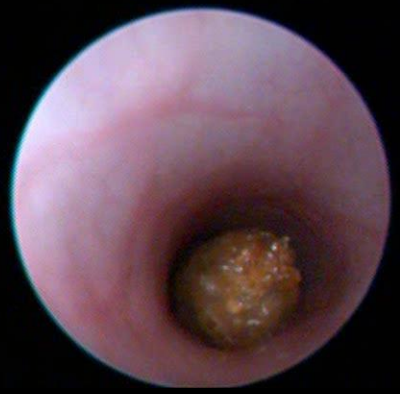




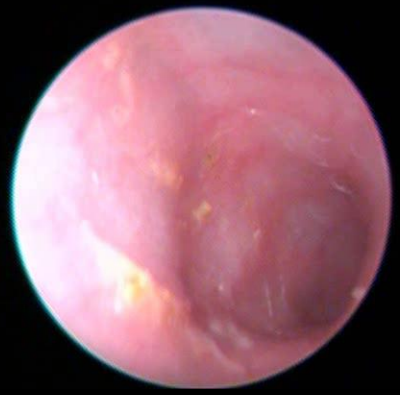
Marina Sirota
Joel Dudley
Mohan M Shenoy
Jay Pasricha

Dudley JT, Sirota M, ..., Pasricha J, Butte AJ. *Science Translational Medicine*, 2011.

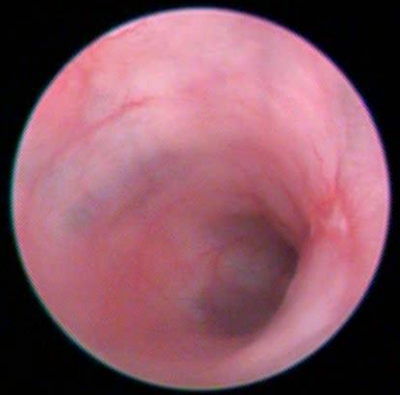
Anti-seizure drug works against a rat model of inflammatory bowel disease



Rat colonoscopy



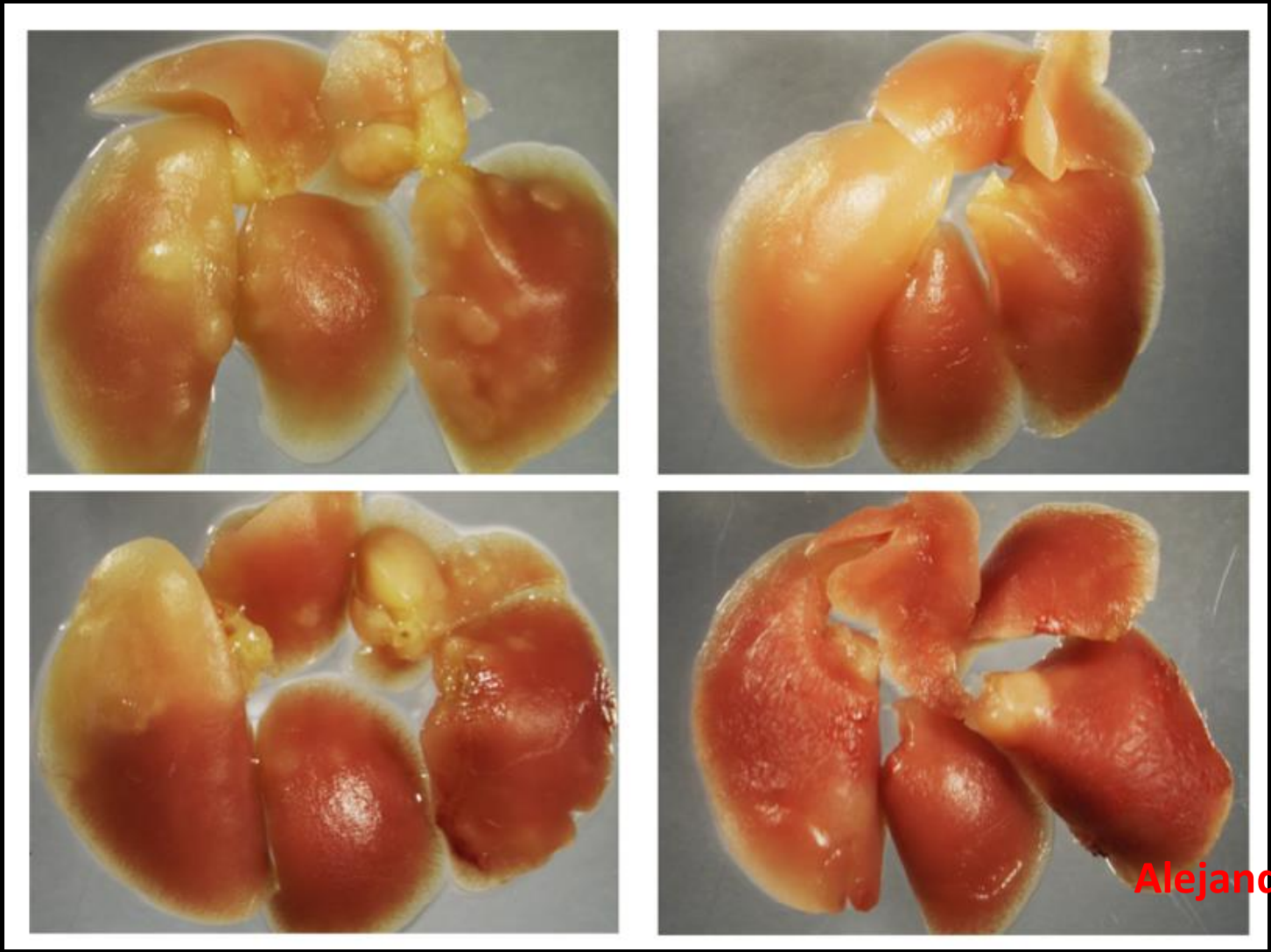
Rat with Inflammatory Bowel Disease



Inflammatory Bowel Disease After Anti-seizure Drug

Dudley JT, Sirota M, ..., Pasricha J, Butte AJ. *Science Translational Medicine*, 2011.

Anti-depressant Imipramine Shows Significant Activity Against Small Cell Lung Cancer



Vehicle control

Imipramine

p53/Rb/p130 triple knockout model of SCLC

Mice dosed after tumor formation

Joel Dudley
Nadine Jahchan
Julien Sage
Alejandro Sweet-Cordero
Joel Neal
NuMedii

Need more drugs for more diseases

Public big data available

NIH funding

Data analyzed, method designed

Company launched, ARRA, Stanford license, first deal

Claremont Creek, Lightspeed ($3.5 million)

NuMedii
Translating Big Data into new medicines

BIOTECHsf

Jun 26, 2013, 5:30am PDT

Venture capital

'Digital drug development' company NuMedii snags $3.5 million

Ron Leuty
Reporter-
San Francisco Business Times
Email | Twitter | Google+ | Twitter

NuMedii Inc., the Palo Alto startup looking to convert pages of drug safety data into faster drug-development times, lined up

Enlarge

NuMedii CEO Gini Deshpande: Tapping old data for new drugs.

FierceBiotechIT NEWS TOPICS ANALYSIS FEAT

Topics: Big Data

UPDATED: 'Big Data' startup wins deal with Aptalis Pharma

October 3, 2012 | By Ryan McBride

SHARE

Email

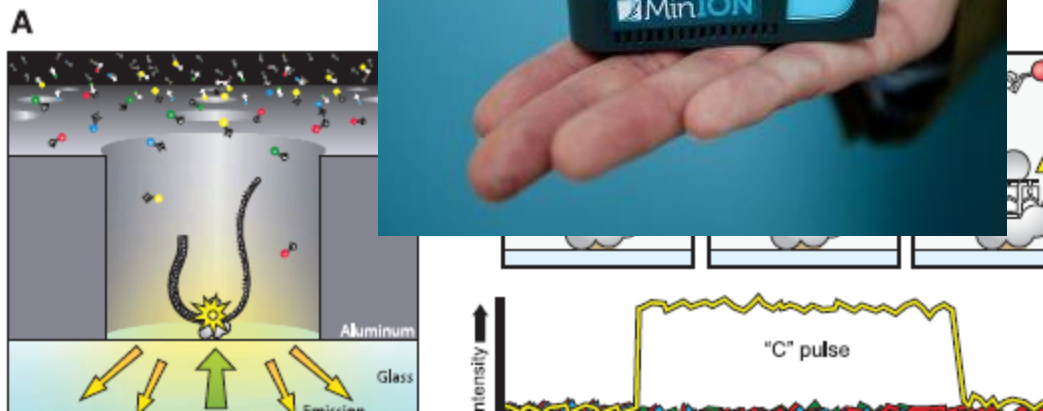
31

Tweet

NuMedii has landed a deal with Aptalis Pharma in which the Stanford University spinoff will apply its predictive "Big Data" technology. The companies aim to hun down and advance drugs to combat gastrointestinal ailments and cystic fibrosis which are two areas of focus at Aptalis. The deal boosts the commercial credentials of NuMedii, building on the startup's role in a pair of papers last year that showed how its computational method could quickly pair approved and gene drugs with new potential uses against diseases.

Sequencing Excitement

- 454/Roche, Life Technologies
- Helicos: $30k genome
- Pacific Biosystems: sequence human genome in 15 minutes
- Run times in minutes at a cost of hundreds of dollars
- Complete Genomics: 80 genomes/day
- Ion Torrent and Illumina: ~$1500 per genome
- Oxford: USB stick



Clinical assessment incorporating a personal genome

Euan A Ashley, Atul J Butte, Matthew T Wheeler, Rong Chen, Teri E Klein, Frederick E Dewey, Joel T Dudley, Kelly E Ormond, Aleksandra Pavlovic, Alexander A Morgan, Dmitry Pushkarev, Norma F Neff, Louanne Hudgins, Li Gong, Laura M Hodges, Dorit S Berlin, Caroline F Thorn, Katrin Sangkuhl, Joan M Hebert, Mark Woon, Hersh Sagreiya, Ryan Whaley, Joshua W Knowles, Michael F Chou, Joseph V Thakuria, Abraham M Rosenbaum, Alexander Wait Zaranek, George M Church, Henry T Greely, Stephen R Quake, Russ B Altman

Summary

Background The cost of genomic information has fallen steeply, but the clinical translation of genetic risk estimates remains unclear. We aimed to undertake an integrated analysis of a complete human genome in a clinical context.

Methods We assessed a patient with a family history of vascular disease and early sudden death. Clinical assessment included analysis of this patient's full genome sequence, risk prediction for coronary artery disease, screening for causes of sudden cardiac death, and genetic counselling. Genetic analysis included the development of novel methods for the integration of whole genome and clinical risk. Disease and risk analysis focused on prediction of genetic risk of variants associated with mendelian disease, recognised drug responses, and pathogenicity for novel variants. We queried disease-specific mutation databases and pharmacogenomics databases to identify genes and mutations with known associations with disease and drug response. We estimated post-test probabilities of disease by applying likelihood ratios derived from integration of multiple common variants to age-appropriate and sex-appropriate pre-test probabilities. We also accounted for gene-environment interactions and conditionally dependent risks.

***Lancet*, 375:1525, May 1, 2010.**

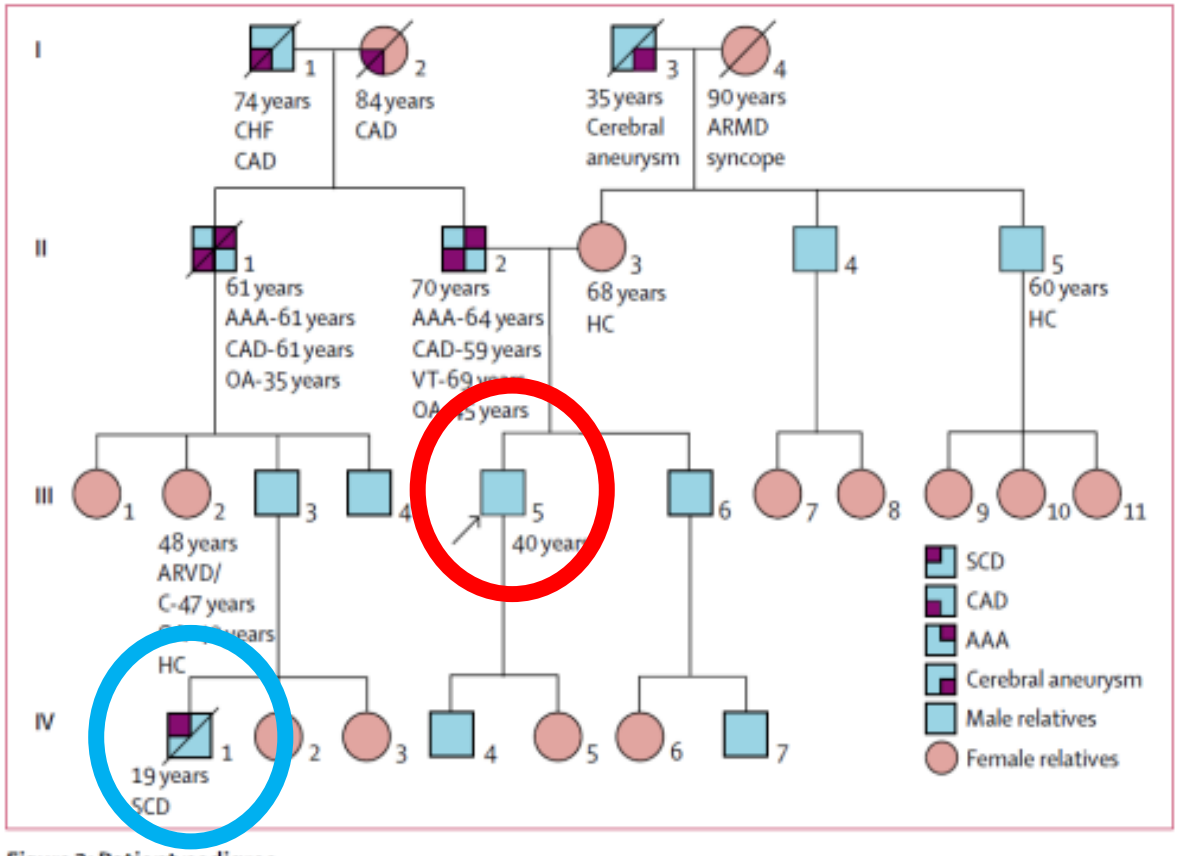


***Figure 2:* Patient pedigree**

The arrow shows the patient. Diagonal lines show relatives who are deceased. Years are age at death or diagnosis. AAA=abdominal aortic aneurysm. ARMD=age-related macular degeneration. ARVD/C=arrhythmogenic right-ventricular dysplasia or cardiomyopathy. CAD=coronary artery disease. CHF=congestive heart failure. HC=hypercholesterolaemia. HTN=hypertension. OA=osteoarthritis. SCD=sudden cardiac death (presumed). VT=paroxysmal ventricular tachycardia.

Credit: Euan Ashley, Russ Altman, Steve Quake, Lancet

ORIGINAL ARTICLE

Association of *IL23R*, *TNFRSF1A*, and HLA-DRB1*0103 Allele Variants with Inflammatory Bowel Disease Phenotypes in the Finnish Population

Maarit Lappalainen, MSc,[*†] Leena Halme, MD, PhD,[‡] Ulla Turunen, MD,[§] Päivi Saavalainen, PhD,*[|] Elisabet Einarsdottir, PhD,*[|] Martti Färkkilä, MD, PhD,[§] Kimmo Kontula, MD, PhD,*[†] and Paulina Paavola-Sakki, MD, PhD[†,§]*

Background: Crohn's disease (CD) and ulcerative colitis (UC), 2 major forms of inflammatory bowel disease (IBD), are complex disorders with significant genetic predisposition. The first CD-associated gene, *CARD15/NOD2*, was recently identified and since then several reports on novel IBD candidate genes have emerged. We investigated disease phenotype association to genetic variations in *IL23R, ATG16L1, DLG5, ABCB1/MDR1, TLR4, TNFRSF1A*, chromosome 5 risk haplotype including *SLC22A4* and *SLC22A5*, and HLA-DRB1*0103 allele among Finnish IBD patients.

Methods: A total of 699 IBD patients were genotyped for disease-associated variants by polymerase chain reaction (PCR) and restriction enzyme digestion or Sequenom iPLEX method.

Results: Five markers spanning the *IL23R* gene were associated with CD. The SNP (single nucleotide polymorphism) rs2201841 gave the strongest association ($P = 0.002$). The rare HLA-DRB1*0103 allele was found to associate with UC ($P = 0.008$), and the *TNFRSF1A* A36G variant was associated with familial UC ($P = 0.007$). Upon phenotypic analysis we detected association between familial UC and rare *TNFRSF1A* alleles 36G and IVS6+10G ($P = 0.001$ and $P = 0.042$, respectively). In addition, *IL23R* markers were associated with stricturing CD ($P = 0.010–0.017$), and ileocolonic CD was more prevalent in the carriers of the same 2 *TNFRSF1A* variants ($P = 0.021$ and $P = 0.028$, respectively). Less significant genotype–phenotype associations were observed for the *TLR4* and HLA variants.

Conclusions: We were able to replicate the association of the *IL23R* variants with CD as well as HLA-DRB1*0103 with UC; confirmation of *TNFRSF1A* association with UC needs additional studies. Our findings also suggest that polymorphisms at *IL23R* and *TNFRSF1A*, and possibly HLA and *TLR4*, loci may account for phenotypic variation in IBD.

(*Inflamm Bowel Dis 2008;14:1118–1124*)

Key Words: Finnish, inflammatory bowel disease, HLA-DRB1*0103, IL23R, TNFRSF1A

Since the initial discovery of the association of *CARD15/NOD2* gene variants with Crohn's disease (CD),[1–3] several new susceptibility genes for inflammatory bowel disease (IBD) have been reported. In 2004 the positional cloning approach led to the identification of the associated variants in solute carrier family 22 (*SLC22A* members 4 and 5)[4] and the discs large homolog 5 (*DLG5*)[5] genes that are implicated in fatty acid oxidation and in maintaining epithelial integrity, respectively. It has not, however, been unequivocally proved that the *SLC22A* genes represent the actual disease genes.[6–13] Most of the studies have confirmed the association of CD with the *SLC22A* gene variants or with the chromosome 5 risk haplotype; however, a study of more than 981 Belgian IBD patients could not replicate the association with IBD, CD, or ulcerative colitis (UC).[14] A recent study by Silverberg et al[15] using a large cohort of IBD trios excluded the *SLC22A5* gene variant as the potential causal variant. The association of genetic variations in the *DLG5* gene with IBD and CD was initially described in 2 large European study samples.[5] The haplotype A, tagged by SNP DLG5_e26 ins/delA, was significantly undertransmitted in IBD and CD, whereas haplotype D, tagged by the SNP G113A (R30Q), was significantly overtransmitted in both IBD and CD. Several groups have not been able to replicate the association since the original report.[13,14,16] However, in 1 case gender-specific analysis revealed an association.[17]

The association of IBD with genetic variation in the Toll-like receptor 4 (*TLR4*) gene has been investigated by many groups but the results have been controversial, which

Received for publication January 23, 2008; Accepted January 26, 2008.

From the *Research Program for Molecular Medicine, Biomedicum Helsinki, Finland, †Department of Medicine, University of Helsinki, Helsinki, Finland, ‡Department of Transplantation and Liver Surgery, Helsinki University Hospital, Helsinki, Finland, §Department of Gastroenterology, Helsinki University Hospital, Helsinki, Finland, |Department of Medical Genetics, Biomedicum Helsinki, Finland.

Supported by a grant from the Special State Funds of the Helsinki University Central Hospital (EVO), University of Helsinki, the Finnish Academy, and the EU commission (MEXT-CT-2005-025270).

Reprints: Kimmo Kontula, Department of Medicine, University of Helsinki, Haartmaninkatu 4, FIN-00290 Helsinki, Finland (e-mail: kimmo.kontula@hus.fi).

DOI 10.1002/ibd.20423
Published online 13 March 2008 in Wiley InterScience (www.interscience.wiley.com)

Rong Chen
Optra Systems

- Study published in 2008 in *Inflammatory Bowel Disease*
- Crohn's Disease and Ulcerative Colitis
- Investigated 9 loci in 700 Finnish IBD patients
- We record 100+ items
  - GWAS, non-GWAS papers
  - Disease, Phenotype
  - Population, Gender
  - Alleles and Genotypes
  - p-value (and confidence)
  - Odds ratio (and confidence)
  - Technology, Study design
  - Genetic model
- Mapped to UMLS concepts

ORIGINAL ARTICLE

Association of *IL23R*, *TNFRSF1A*, and HLA-DRB1*0103 Allele Variants with Inflammatory Bowel Disease Phenotypes in the Finnish Population

Maarit Lappalainen, MSc,[*†] *Leena Halme, MD, PhD,*[‡] *Ulla Turunen, MD,*[§] *Päivi Saavalainen, PhD,*[*|] *Elisabet Einarsdottir, PhD,*[*|] *Martti Färkkilä, MD, PhD,*[§] *Kimmo Kontula, MD, PhD,*[*†] *and Paulina Paavola-Sakki, MD, PhD*[†,§]

Background: Crohn's disease (CD) and ulcerative colitis (UC), 2 major forms of inflammatory bowel disease (IBD), are complex disorders with significant genetic predisposition. The first CD-associated gene, *CARD15/NOD2*, was recently identified and since then several reports on novel IBD candidate genes have emerged. We investigated disease phenotype association to genetic variations in *IL23R, ATG16L1, DLG5, ABCB1/MDR1, TLR4, TNFRSF1A*, chromosome 5 risk haplotype including *SLC22A4* and *SLC22A5*, and HLA-DRB1*0103 allele among Finnish IBD patients.

Methods: A total of 699 IBD patients were genotyped for disease-associated variants by polymerase chain reaction (PCR) and restriction enzyme digestion or Sequenom iPLEX method.

Conclusions: We were able to replicate the association of the *IL23R* variants with CD as well as HLA-DRB1*0103 with UC; confirmation of *TNFRSF1A* association with UC needs additional studies. Our findings also suggest that polymorphisms at *IL23R* and *TNFRSF1A*, and possibly HLA and *TLR4*, loci may account for phenotypic variation in IBD.

(*Inflamm Bowel Dis 2008;14:1118–1124*)

Key Words: Finnish, inflammatory bowel disease, HLA-DRB1*0103, IL23R, TNFRSF1A

Since the initial discovery of the association of *CARD15/NOD2* gene variants with Crohn's disease (CD),[1–3] sev-

TABLE 1. Case-control Analysis of the *IL23R* Gene Including 8 SNPs

dbSNP ID	Allele	Location	Controls $n = 292$	IBD $n = 697$	*P* value	CD $n = 238$	*P* value	UC $n = 459$	*P* value
rs1004819	C	Intron 5	0.751	0.704	**0.037**	0.671	**0.005**	0.721	0.215
	T		0.249	0.296		0.329		0.279	

sinki, Haartmaninkatu 4, FIN-00290 Helsinki, Finland (e-mail: kimmo.kontula@hus.fi).

DOI 10.1002/ibd.20431

Published online13 March 2008 in Wiley InterScience (www.interscience.wiley.com).

port.[13,14,16] However, in 1 case gender-specific analysis revealed an association.[17]

The association of IBD with genetic variation in the Toll-like receptor 4 (*TLR4*) gene has been investigated by many groups but the results have been controversial, which

- Study published in 2008 in *Inflammatory Bowel Disease*
- Crohn's Disease and Ulcerative Colitis
- Investigated 9 loci in 700 Finnish IBD patients
- We record 100+ items
  - GWAS, non-GWAS papers
  - Genetic model
- Mapped to UMLS concepts

VARIMED: Variants Informing Medicine

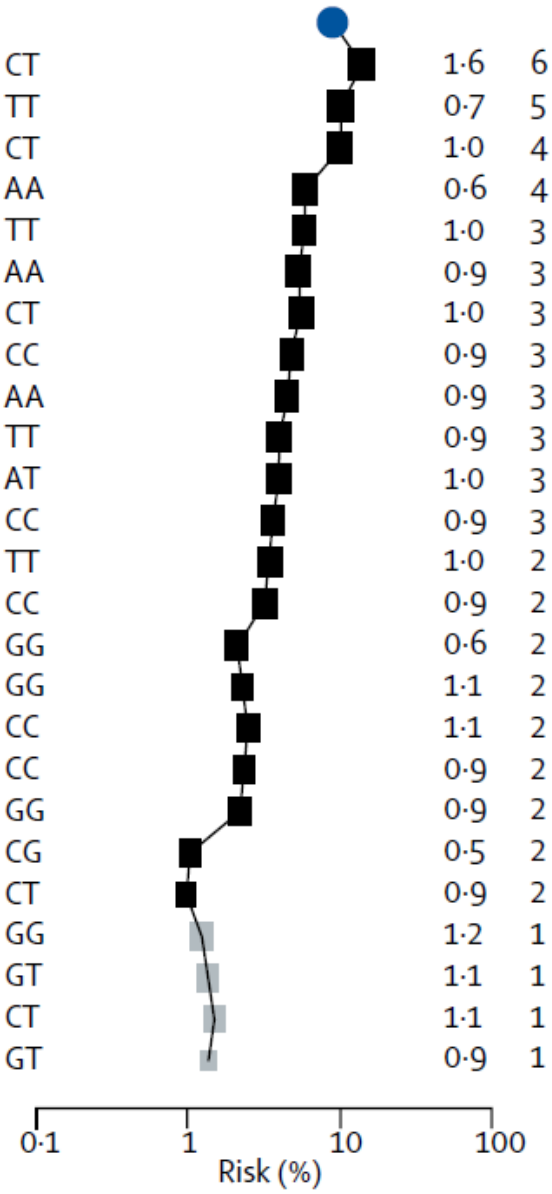
Number of papers curated	Number of records	Distinct SNPs	Diseases and phenotypes
~19,000	~1.6 million	~473,000	~7,400

Chen R, Davydov EV, Sirota M, Butte AJ.
PLoS One.
2010 October: 5(10): e13574.

Rong Chen
Anil Patwardhan
Michael Clark
Optra Systems
Personalis

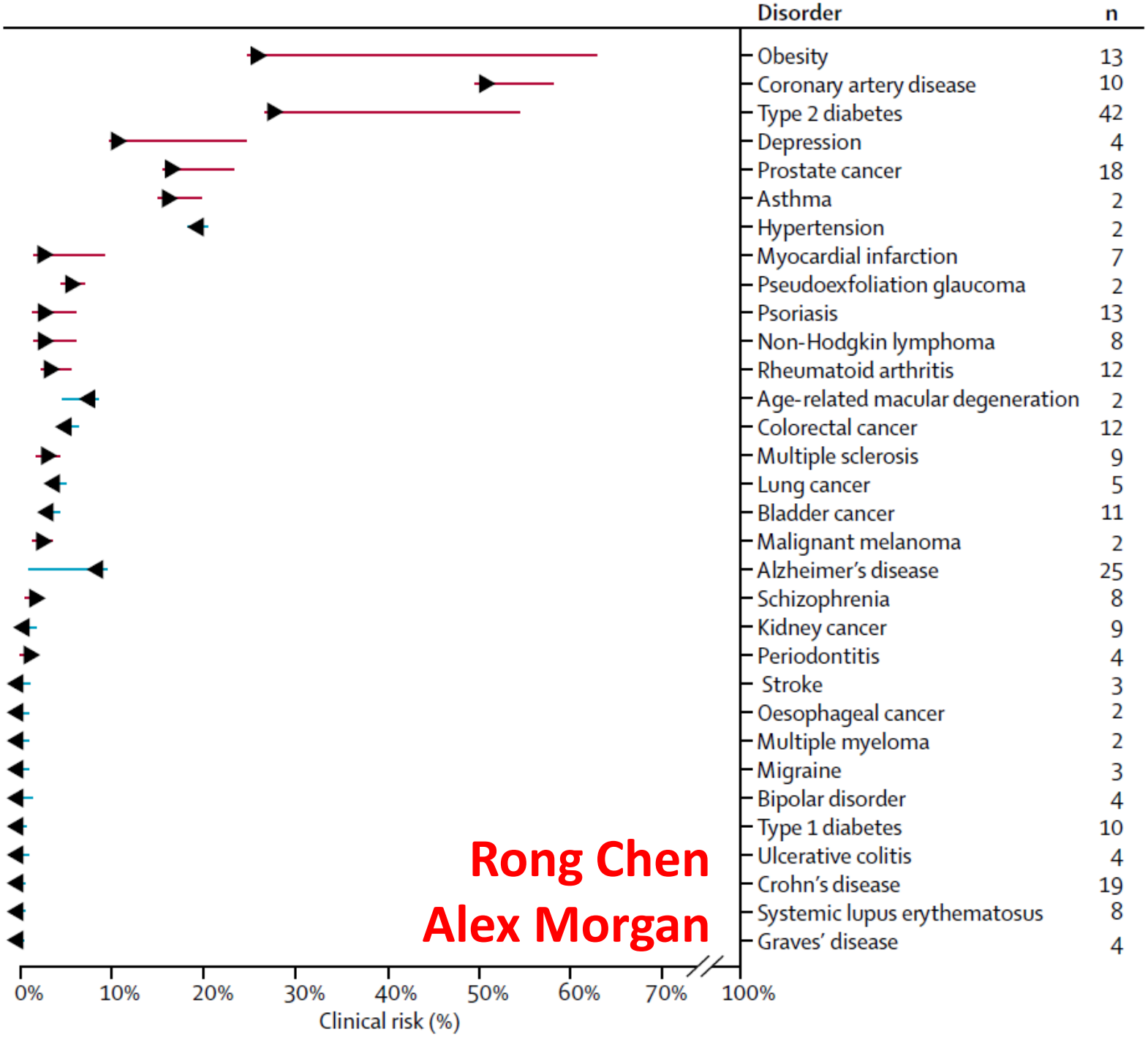
D Alzheimer's disease

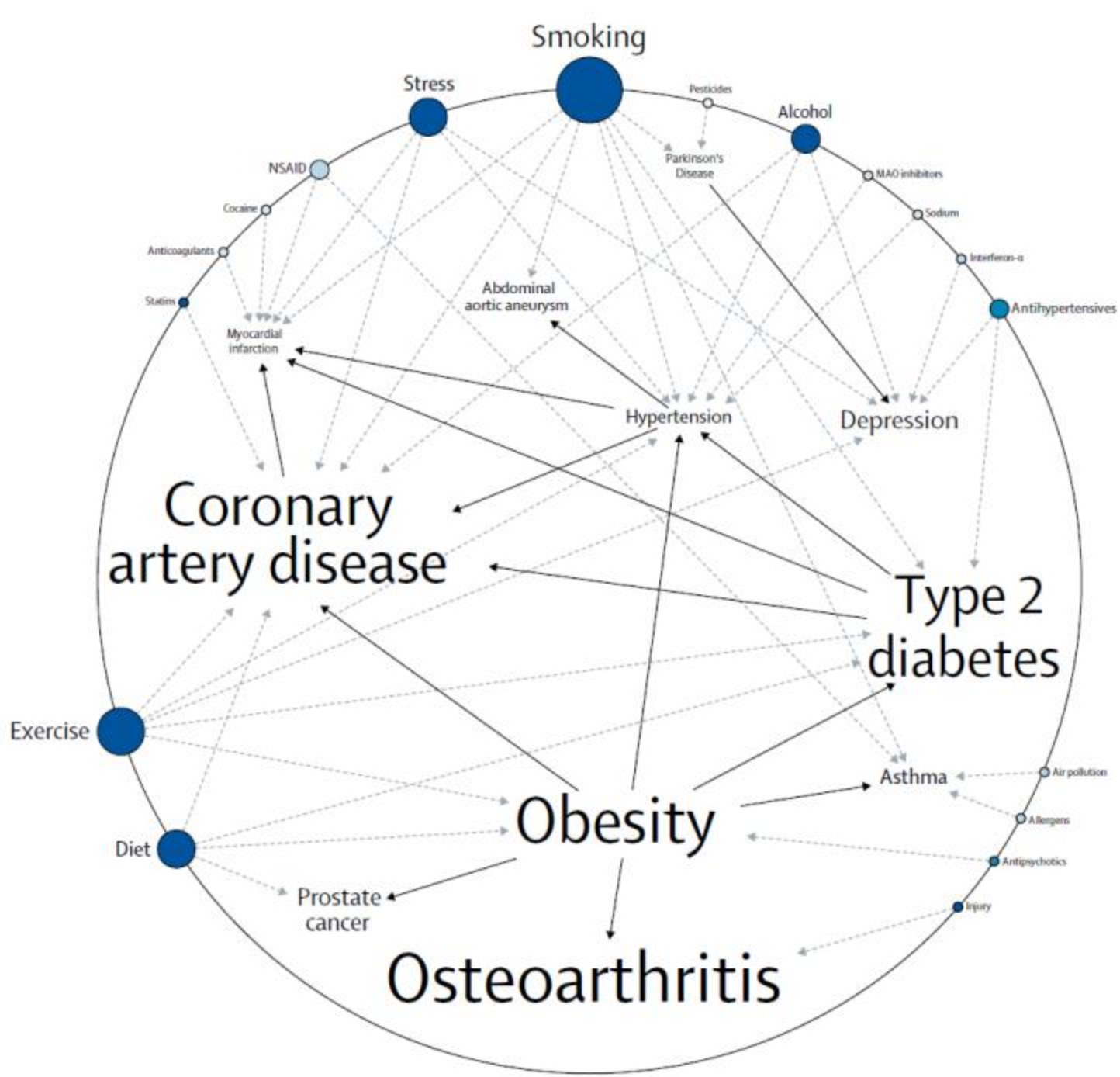
Gene*	SNP location	Patient genotype	LR	Studies†	Samples‡	Post-test probability (%)
						9·0%
TOMM40	rs157581	CT	1·6	6	7740	13·90%
DAPK1	rs4878104	TT	0·7	5	10397	10·19%
TRAK2	rs13022344	CT	1·0	4	6512	10·12%
DAPK1	rs4877365	AA	0·6	4	4841	5·89%
E8F3	rs11016976	TT	1·0	3	5736	5·87%
TNK1	rs1554948	AA	0·9	3	5736	5·32%
MYH13	rs2074877	CT	1·0	3	5366	5·55%
GALP	rs3745833	CC	0·9	3	5366	4·82%
PCK1	rs8192708	AA	0·9	3	5366	4·47%
	rs1859849	TT	0·9	3	5304	4·02%
	rs11622883	AT	1·0	3	5248	3·97%
WWC1	rs17070145	CC	0·9	3	2545	3·65%
LMNA	rs505058	TT	1·0	2	4646	3·49%
ACAN	rs2882676	CC	0·9	2	4590	3·22%
PGBD1	rs3800324	GG	0·6	2	4590	2·11%
GOLM1	rs10868366	GG	1·1	2	2156	2·30%
GOLM1	rs7019241	CC	1·1	2	2156	2·49%
	rs9886784	CC	0·9	2	2156	2·36%
	rs10519262	GG	0·9	2	2156	2·22%
	rs463946	CG	0·5	2	1922	1·04%
PLAU	rs2227564	CT	0·9	2	956	0·98%
ADAM12	rs1278279	GG	1·2	1	2320	1·23%
SORL1	rs2070045	GT	1·1	1	2031	1·36%
ABCA1	rs2230806	CT	1·1	1	1691	1·50%
PSEN1	rs165932	GT	0·9	1	170	1·37%



Rong Chen
Alex Morgan

Ashley EA*, Butte AJ*, Wheeler MT, Chen R, Klein TE, Dewey FE, Dudley JT, Ormond KE, Pavlovic A, Hudgins L, Gong L, Hodges LM, Berlin DS, Thorn CF, Sangkuhl K, Hebert JM, Woon M, Sagreiya H, Whaley R, Morgan AA, Pushkarev D, Neff NF, Knowles W, Chou M, Thakuria J, Rosenbaum A, Zaranek AW, Church G, Greely HT*, Quake SR*, Altman RB*. Clinical evaluation incorporating a personal genome. *Lancet*, 2010.


Disorder
n
Obesity 13
Coronary artery disease 10
Type 2 diabetes 42
Depression 4
Prostate cancer 18
Asthma 2
Hypertension 2
Myocardial infarction 7
Pseudoexfoliation glaucoma 2
Psoriasis 13
Non-Hodgkin lymphoma 8
Rheumatoid arthritis 12
Age-related macular degeneration 2
Colorectal cancer 12
Multiple sclerosis 9
Lung cancer 5
Bladder cancer 11
Malignant melanoma 2
Alzheimer's disease 25
Schizophrenia 8
Kidney cancer 9
Periodontitis 4
Stroke 3
Oesophageal cancer 2
Multiple myeloma 2
Migraine 3
Bipolar disorder 4
Type 1 diabetes 10
Ulcerative colitis 4
Crohn's disease 19
Systemic lupus erythematosus 8
Graves' disease 4
Rong Chen
Alex Morgan
0%
10%
20%
30%
40%
50%
60%
70%
100%
Clinical risk (%)



Rong Chen
Alex Morgan
Joel Dudley

Need to use genomes to predict disease

Publications available for curation

CHI startup funding

Science curated, methods designed

Company launched, Stanford license

MDV, Lightspeed, Abingworth ($20 million)

Same 3 plus Wellington Shields ($22 million)

STOCK WATCH Express, Wet Seal, Avago Jump

Personalis Awarded Contract From VA Million Veteran Program – Whole Genome Sequencing and Data Analysis for Over 1,000 Individuals

BusinessWire Press Rele

Recommend Twe

MENLO PARK, Calif.--(BUSI

The US Department of Vetera
sequencing and data analysis
samples from several VA sou
secure computing facility and
against an advanced human re
genetic analyses to help confi
laboratory genetic analysis, i

Personalis® PIONEERING GENOME GUIDED MEDICINE

Services Technology Applications Publications News & Events

News & Events Press Releases

Back to

October 07, 2013

Personalis Closes $22M Series B Financing

Menlo Park, CA – Personalis, Inc., the leading provider of advanced medical exome and genome sequencing an
interpretation services, announced today the closing of a $22 million Series B financing. "This financing brings th
investment in Personalis to over $42M, the relative scale of which gives us a significant advantage over most ne
now entering this space " said Personalis CEO John West. "We will use this to further build on our technologica
expand our product offerings, scale our operations, and expand our commercial team. Since our first customer
order 13 months ago, we have received more than $14M in customer orders, for the sequencing and analysis of

immport.niaid.nih.gov

ImmPort
Bioinformatics For The Future of Immunology

Log in ▼ | Sign up FREE

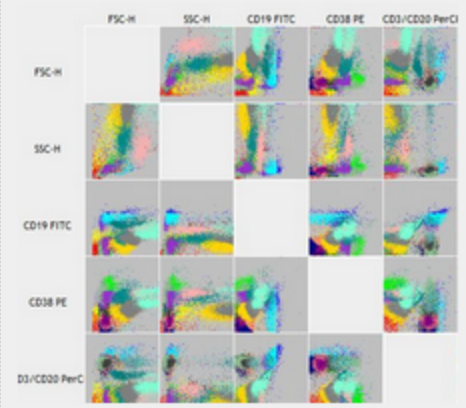
About ImmPort | Access Data | Tools | Resources | News & Events

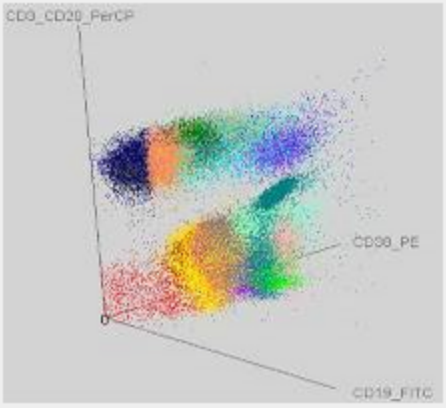
FLOCK

FLOCK (FLOw Clustering without K means) uses a density-based clustering approach to algorithmically identify biologically relevant cell populations. Provides statistical analysis of populations.

Two dimensional and three dimensional visualization of cell populations.

Compare population statistics from multiple samples.





Jeff Wiser
Patrick Dunn
Sanchita Bhattacharya

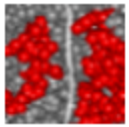
Flow Cytometry Analysis (FLOCK)

Flow cytometry analysis component includes:

- Automated cell population identification
- Result visualization in 2D and 3D
- Statistical analysis of population characteristics
- Automated mapping of populations across multiple samples

flock
Flow cytometry clustering without K

MHC Validation and Analysis



MHC Sequence Feature Variant Type (SFVT) Analysis enables genetic association analysis of classical HLA protein sub-regions defined with structural (e.g. helix) and functional (e.g. binding site) information.

MHC Alleles

Complete DNA and protein sequences, sequence features, and population frequencies of MHC, MIC and TAP alleles. Align MHC sequences horizontally to visualize extent of polymorphisms across all alleles in a locus.

New Data Release

August 16, 2013 - The National Institute of Allergy and Infectious Diseases (NIAID) released to the ImmPort user community new data from 6 clinical studies or trials and updates to 7 additional studies available here. Research areas include predictive influenza biomarkers, antibody responses to pH1N1 and oral immunotherapy for childhood allergies. This release brings the total number of shared studies to 60.

Data Summary

Studies	60
Subjects	13859
Experiments	569
Total Results	226998

◉ Research Program category ○ Research Focus category

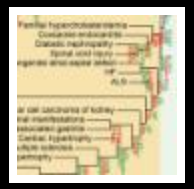
We are used to kids starting computer, mobile, and internet companies in garages and dorm rooms...

We are used to kids starting computer, mobile, and internet companies in garages and dorm rooms...

Maybe kids today need to start "garage biotechs"?

Take Home Points

- The patients, samples, molecular, clinical, and epidemiological data and tools are already publicly available to make an impact across medicine.

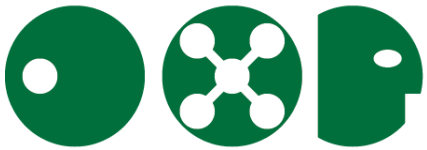


- Waiting for the perfect tools, perfect infrastructure, perfect data, and perfect annotations is waiting too long. Need for perfection is hiding data today.

- We need investigators who can imagine basic questions to ask of these repositories of clinical and genomic measurements.

Collaborators

- Jeff Wiser, Patrick Dunn, Mike Atassi / Northrop Grumman
- Ashley Xia and Quan Chen / NIAID
- Takashi Kadowaki, Momoko Horikoshi, Kazuo Hara, Hiroshi Ohtsu / U Tokyo
- Kyoko Toda, Satoru Yamada, Junichiro Irie / Kitasato Univ and Hospital
- Shiro Maeda / RIKEN
- Alejandro Sweet-Cordero, Julien Sage / Pediatric Oncology
- Mark Davis, C. Garrison Fathman / Immunology
- Russ Altman, Steve Quake / Bioengineering
- Euan Ashley, Joseph Wu, Tom Quertermous / Cardiology
- Mike Snyder, Carlos Bustamante, Anne Brunet / Genetics
- Jay Pasricha / Gastroenterology
- Rob Tibshirani, Brad Efron / Statistics
- Hannah Valantine, Kiran Khush/ Cardiology
- Ken Weinberg / Pediatric Stem Cell Therapeutics
- Mark Musen, Nigam Shah / National Center for Biomedical Ontology
- Minnie Sarwal / Nephrology
- David Miklos / Oncology



Butte Lab

Systems Medicine • Stanford Pediatrics • Packard Children's Hospital

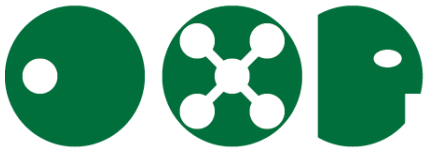
Support

- Lucile Packard Foundation for Children's Health
- NIH: NIAID, NLM, NIGMS, NCI; NIDDK, NHGRI, NIA, NHLBI, NCATS
- March of Dimes
- Hewlett Packard
- Howard Hughes Medical Institute
- California Institute for Regenerative Medicine
- Luke Evnin and Deann Wright (Scleroderma Research Foundation)
- Clayville Research Fund
- PhRMA Foundation
- Stanford Cancer Center, Bio-X, SPARK

Admin and Tech Staff

- Susan Aptekar
- Jen Cory
- Alex Skrenchuk

- Tarangini Deshpande
- Kimayani Butte



Butte Lab

Systems Medicine • Stanford Pediatrics • Packard Children's Hospital