



Building Bridges to Discovery: Bioinformatics and Computational Biology at NIH

Karin A. Remington, Ph.D.

NCIBI 4th Annual Research Conference 2009
Ann Arbor, Michigan





NIGMS Center for Bioinformatics and Computational Biology (CBCB)

■ NIGMS:

- R01 portfolio in bioinformatics (interpreted very broadly), modeling, software development and maintenance...
- Systems Biology Centers
- MIDAS (Models of Infectious Disease Agent Study)
- Training grants in Bioinformatics, and Biostatistics

■ Trans-NIH:

- Roadmap: National Centers for Biomedical Computing
- BISTI: Biomedical Informatics Science and Technology Initiative

■ Inter-agency Coordination





Building Bridges to discovery....

- Develop new tools and computational approaches
- Enable new science
- Promote collaboration
- Encourage Reduce/Reuse/Recycle





Building Bridges to Discovery:

The CHALLENGES of interdisciplinary funding at the NIH





...WITHIN NIGMS





CBCB sponsored Research Program Grants (R01s):

Some current and emerging interests:

- Structural and functional prediction of proteins
- Biostatistics – modeling, study design, data analysis
- Numerical linear algebra and optimization algorithms
- Network analysis/graph theory
- Systems biology approaches
- Innovative genomic analysis techniques
- Interdisciplinary training emphasizing collaborative science between Math/CS and Biology
- Biomedical data infrastructures supporting aggregation, interoperability and preservation

***Fully integrating mathematics and computer science
in the context of biological studies***





Models of Infectious Disease Agent Study

Computational and mathematical investigations of:

- Dynamics of emergence and spread of pathogens and their products
- Identification and surveillance of infectious diseases
- Effectiveness and consequences of intervention strategies
- Host/pathogen interactions
- Ecological, climatic, and evolutionary dimensions of infectious diseases outbreaks.



NIGMS Predoctoral Research Training Program



The Vaccination Theory of education(?)

English is not history and history is not science and science is not art and art is not music, and art and music are minor subjects and English, history and science major subjects, and a subject is something you 'take' and when you have taken it, you have 'had' it, and if you have 'had' it, you are immune and need not take it again.

From TEACHING AS A SUBVERSIVE ACTIVITY
Neil Postman & Charles Weingartner





NIGMS Predoctoral Research Training Program

- Major mission of NIGMS
- Predoctoral-only training grants
- 11 pre- Ph.D. areas
- 1 pre - M.D.- Ph.D. area
- Support for 3,200 trainees annually





Training Goals

- Multidisciplinary and multi-departmental training
- Faculty from different academic units provide breadth of research opportunities
- Training to master a core scientific area in depth
- Acquire skills and knowledge of related fields





TRANS-NIH





The prevailing view of NIH...





Promoting the nation's health through research.

- Not a monolithic Agency - 27 Institutes and Centers
- Intramural research (NIH labs) - 10% of the budget, 6,000 scientists
- Extramural research (grants) - 80-90% of the budget

NCI \$4.7B	NIAID \$4.3B	NHLBI \$2.9B	NIDDK \$1.8B	NINDS \$1.5B	NIMH \$1.3
NICHD \$1.2B	NIA \$1.0B	NIDA \$0.9B	NEI \$0.6B	NIHHS \$0.6B	NIAMS \$0.5B
NIAAA \$0.4B	NIDCD \$0.3B	NIDCR \$0.3B	NCMHD \$0.1B	NCCAM \$0.1B	NINR \$0.1B
NIGMS \$1.9B	NCRR \$1.0B	NHGRI \$0.4B	NLM \$0.3B	NIBIB \$0.2B	
CC	CIT	CSR	FIC	OD	





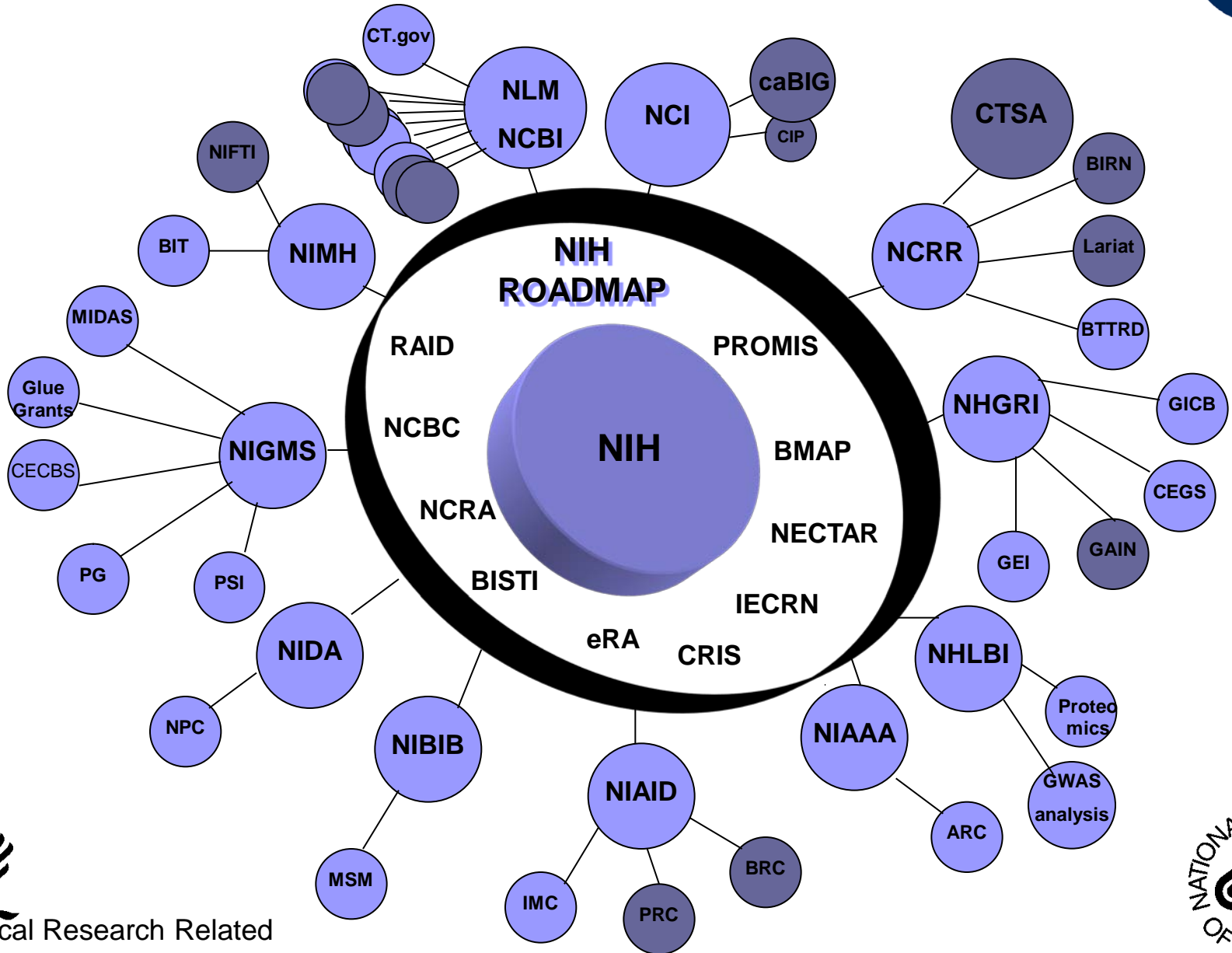
“BISTI”

***The Biomedical Information Science and
Technology Initiative***





BISTI related NIH activities...





BISTI Website: <http://www.bisti.nih.gov>

The screenshot shows the BISTI website in a Windows Internet Explorer browser window. The address bar displays <http://www.bisti.nih.gov/>. The page header includes the BISTI logo and the text "Biomedical Information Science and Technology Initiative" and "Your source for biomedical informatics at NIH". A navigation menu contains links for HOME, FUNDING, LIBRARY, RELATED INITIATIVES, INTRAMURAL RESEARCH, and CONTACT US. The main content area is divided into several sections:

- Introduction:** "The Biomedical Information Science and Technology Initiative is a consortium of representatives from each of the NIH institutes and centers. BISTI was established in May 2000 to serve as the focus of biomedical computing issues at the NIH." "The mission of BISTI is to make optimal use of computer science and technology to address problems in biology and medicine by fostering new basic understandings, collaborations, and transdisciplinary initiatives between the computational and biomedical sciences." "In support of this mission, the BISTI coordinates research grants, training opportunities, and scientific symposia associated with biomedical computing. Regular monthly meetings are conducted to discuss program status, future needs and directions, and topics of interest to the bioinformatics community."
- Funding:** "Funding for biomedical informatics is spread broadly across NIH Institutes and Centers. The NIH Guide is available for customized searches of current announcements. BISTI coordinates several cross-cutting opportunities. [Read more about funding >](#)"
- Library:** "BISTI-related reports, white papers, and other documents are posted in our online library for reference. [Read more in the BISTI library >](#)"
- Related Initiatives:** "Links to an alphabet soup of programs related to BISTI are consolidated here in"
- Intramural Research:** "NIH is home to a rich intramural research community. [Read more about](#)"

The footer of the page shows the URL <http://www.bisti.nih.gov/index.asp> and the status bar indicates "Internet | Protected Mode: On" and "100%".





BISTI Website: <http://www.bisti.nih.gov>

BISTI Biomedical Information Science and Technology Initiative
Your source for biomedical informatics at NIH

HOME FUNDING LIBRARY RELATED INITIATIVES INTRAMURAL RESEARCH CONTACT US

Return to: [Home](#) > [Funding Contacts](#) > NIH R01 Contacts

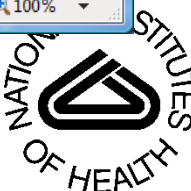
BISTI R01 Contact List

IC	Contact	Email	Phone No.
National Institute of Allergy and Infectious Diseases (NIAID)	Dr. Cheryl Kraft	CKraft@niaid.nih.gov	301-496-7551
National Institute of Allergy and Infectious Diseases (NIAID)	Ms. Valentina Di Francesco	VDiFrancesco@niaid.nih.gov	301-496-1884
National Institute of Neurological Disorders and Stroke (NINDS)	Dr. Yuan Liu	liuyuan@ninds.nih.gov	301-496-0012
National Institutes of Diabetes and Digestive and Kidney Diseases (NIDDK)	Dr. Arthur Castle	CastleA@mail.nih.gov	301-541-7719
National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)	Dr. Gayle Lester	gl83g@nih.gov	301-594-5055

Funding Categories

- [Funding Announcements](#)
- [Funding Process](#)
- **Funding Contacts**
- [Funded Projects](#)

<http://www.bisti.nih.gov/index.asp>



DIVERSE DISCIPLINES, ONE COMMUNITY

BiomedicalComputation REVIEW



BRINGING THE
Fruits of Computation
TO BEAR ON
Human Health:

It's a Tough job
But the **NIH**
Has to Do It

By Katharine Miller



DIVERSE DISCIPLINES, ONE COMMUNITY

Biomedical Computation REVIEW



“As functional as [BISTI] has been over the years, it has really been unable to look across institutes in a real data-driven way, to analyze across NIH where our investments are going,” says Remington.





NIH Roadmap Strategy

Building Blocks Pathways
Molecular Libraries
Structural Biology
Nanomedicine
Bioinformatics and
Computational Biology

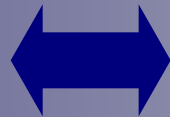
Translational
Research
Initiatives

Clinical
Research
Informatics

Clinical and Translational
Science Awards

Clinical Research Training
National Clinical Research Associates

Bench



Bedside



Practice

Interdisciplinary Research
Pioneer Award
Nanomedicine

Public Private
Partnerships

Integrated Research Networks
Clinical Outcomes





NCBC Centers



www.ncbcs.org

National Centers for Biomedical Computing





NCBC Goals

- ***Development*** of cutting edge computer science
- ***Translation*** of this computer science into biomedical computation, i.e., effective algorithms and environments for solving real biological problems
 - Enable the analysis, modeling, understanding, and prediction of dynamic and complex biomedical systems across time and distance scales
 - Allow the integration of biomedical and behavioral data and knowledge at all levels of organization





NCBC Organization by Core

- **Core 1: Algorithm Development**
 - (computer science)
- **Core 2: Software Engineering**
 - (biomedical computation)
- **Core 3: Driving Biological Problems**





Core 3: *Driving Biological Problems*

- Collaborative science
- Investigator(s) with challenging biomedical problems which focus the NCBC's computational research.
- Problems selected for their broad biomedical significance and compatibility with the core computational expertise of the specific NCBC.





Cores 4 - 6

- (4) providing *infrastructure* to serve the needs of the broad community of biomedical and behavioral researchers;
- (5) enhancing the *training* for a new generation of biomedical researchers in appropriate computational tools and techniques;
- (6) *disseminating* newly developed tools and techniques to the broader biomedical research community;





...ACROSS AGENCIES





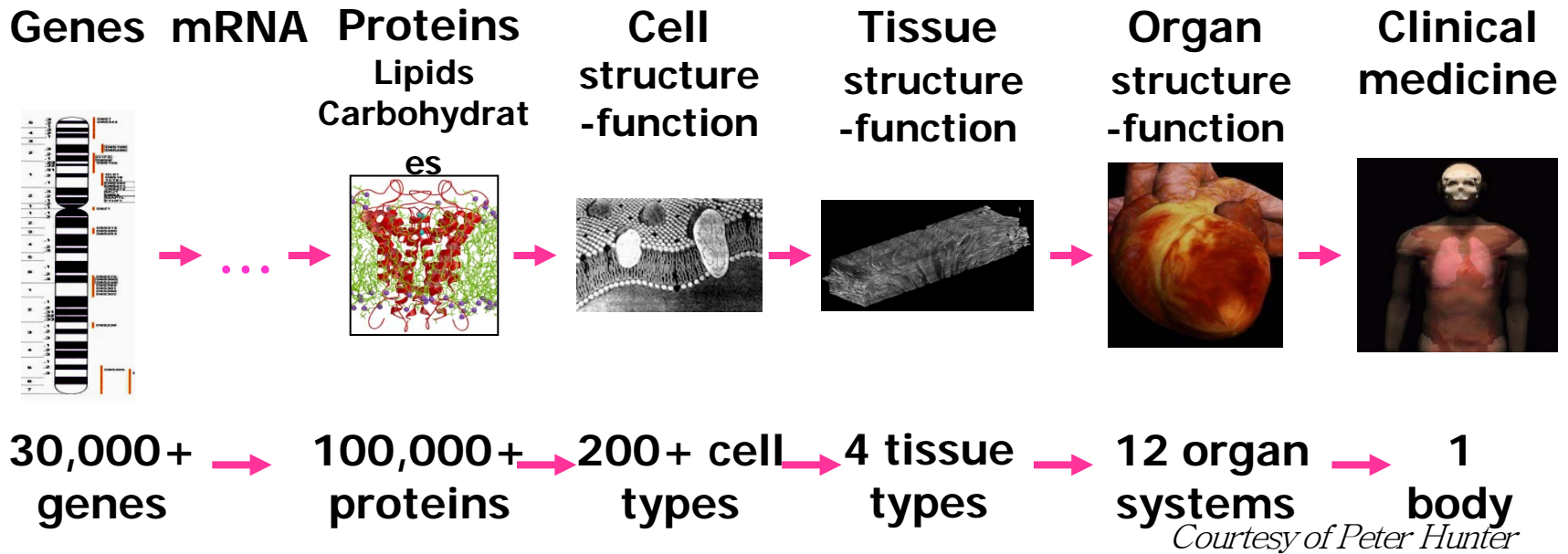
also....Many Related Inter-Agency Activities:

- Networking and Information Technology Research and Development (NITRD)
- Interagency Modeling and Analysis Group (IMAG) NSF 04 607
- Collaborative Research in Computational Neurosciences (CRCNS) NSF 04 514
- Dynamic Data Driven Applications Systems NSF 05 570
- Multi-Agency Tissue Engineering Science (MATES) Working Group
- The Biomaterials and Medical Implant Science (BMIS) Coordinating Committee
- Roundtable on Biomedical Engineering Materials and Applications (BEMA)
- Interagency Image Guided Interventions (IGI)
- Joint NSF-NIH Initiative to Support Research in the Areas of Mathematical Biology NSF 04 572, NSF 06 607



Multiscale Modeling

Spanning scales from molecular to population, designed to illuminate biological processes and further the predictive capability in biological, biomedical and environmental systems. Multi-scale modeling encompasses concepts of space, time and state space.



Biological Scales



Interagency Modeling and Analysis Group

IMAG Multiscale Modeling (MSM) Consortium Working Groups

- Filament Dynamics and Simulation (FDS)
- Cardiac and Skeletal Muscle Physiology
- Macro-To-Micro Scale Transport in Human Systems
- Cell Level Modeling
- High Performance Computing, Computational Issues and Algorithms
- Tissue Mechanics
- Multiscale Imaging
- Theoretical Methods
- Nano-modeling
- Model Sharing



Contact: Grace Peng, NIBIB, email: penggr@mail.nih.gov





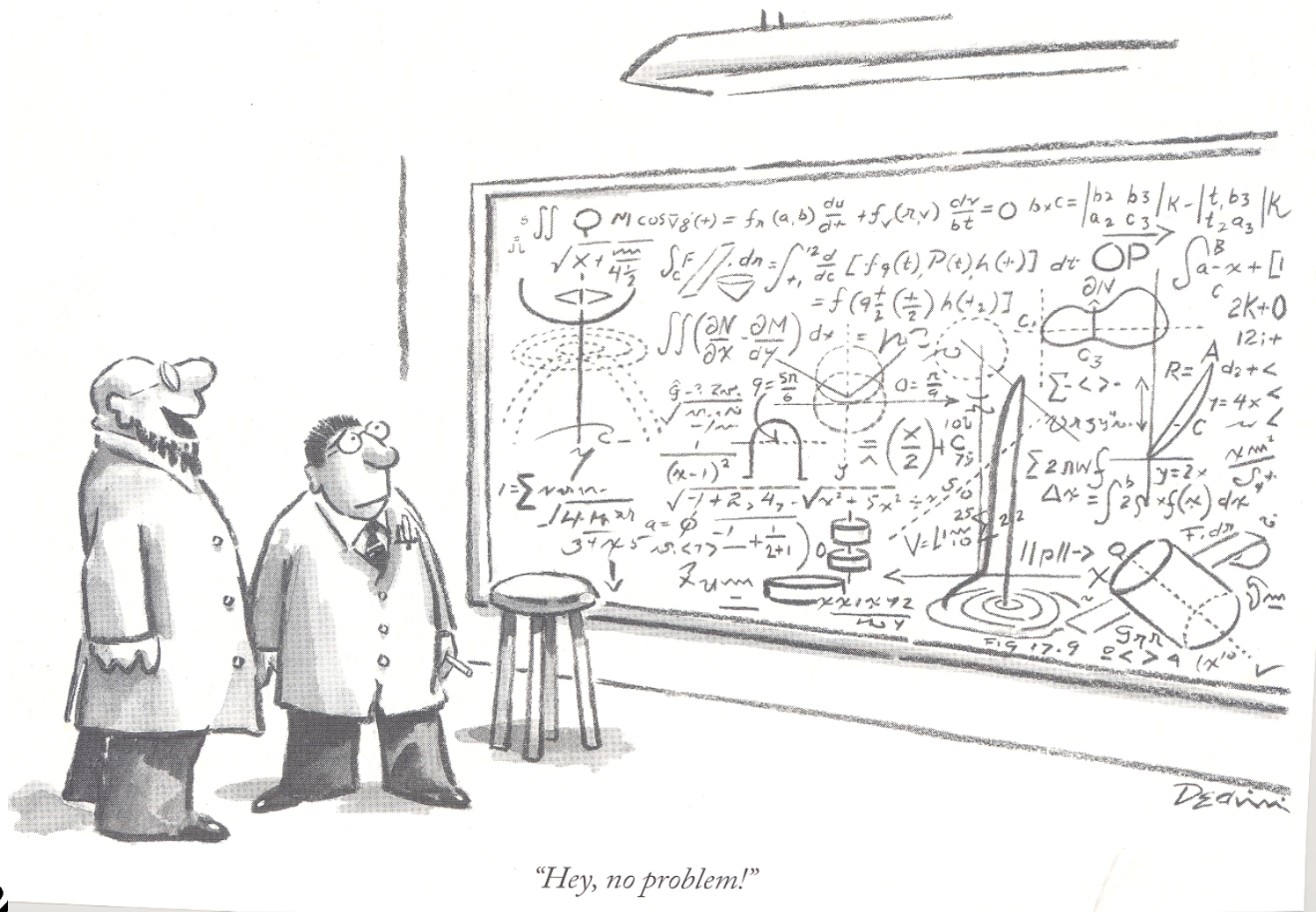
Building Bridges to Discovery:

The CHALLENGES of interdisciplinary funding at the NIH





Language barriers:



"Hey, no problem!"

NIGMS Predoctoral Research Training Program



The Vaccination Theory of education(?)

English is not history and history is not science and science is not art and art is not music, and art and music are minor subjects and

English, history and science major subjects, and a subject is something you 'take' and when you have taken it, you have 'had' it, and if you have 'had' it, you are immune and need not take it again.

From TEACHING AS A SUBVERSIVE ACTIVITY
Neil Postman & Charles Weingartner





Support of NIH Study Sections, when it comes to interdisciplinary teams and emerging areas:





To encourage fair review of the desired collaborations:

- Special Emphasis Panels for targeted FOAs
- Multiple PI applications
- Encouragement of “New” Investigators
 - IC-level policies to boost likelihood of funding
- “Enhancing Peer-Review” effort underway





Support of research and promotion, when it comes to interdisciplinary teams and emerging areas:





The genome sequence of *Drosophila melanogaster*.

Science. 2000 Mar 24;287(5461):2185-95

Adams MD, Cline TM, Akeno N, Ohta T, Taylor T, Eichler E, Lander E, Manly B, McKnight K, Mortrud M, Nierman M, Schumacher S, Whittaker P, Adams B, Agbayani A, An H, Ankeny R, Aronson R, Axelrod D, Ayres M, Baker B, Bakker S, Balakrishnan A, Baliga S, Balogh Z, Balslev D, Banerjee A, Baranowski T, Barbero J, Barbero P, Barin J, Barlow J, Barrow D, Barrett C, Barrett T, Barzilai R, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hosten D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalili M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McCluskey TC, McLeod M, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nishizawa T, Nishizawa Y, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Ringden O, Romanov S, Rosetti F, Scheeler F, Shen H, Shue BC, Sidén-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner D, Wang E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Whittaker P, Whittaker J, Whittaker K, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith RA, Myers EW, Rubin GM, Venter JC.



Another challenge: Appropriate funding mechanisms

- Interdisciplinary science – interdisciplinary teams
- Multiple PIs - strong leadership and shared vision
- Sustainable funding plans





Cooperative Agreements in a Competitive Science World





Models of Infectious Disease Agent Study

Computational and mathematical investigations of:

- Dynamics of emergence and spread of pathogens and their products
- Identification and surveillance of infectious diseases
- Effectiveness and consequences of intervention strategies
- Host/pathogen interactions
- Ecological, climatic, and evolutionary dimensions of infectious diseases outbreaks.





NIH/NIGMS Center for Bioinformatics and Computational Biology

Karin A. Remington, Ph.D.

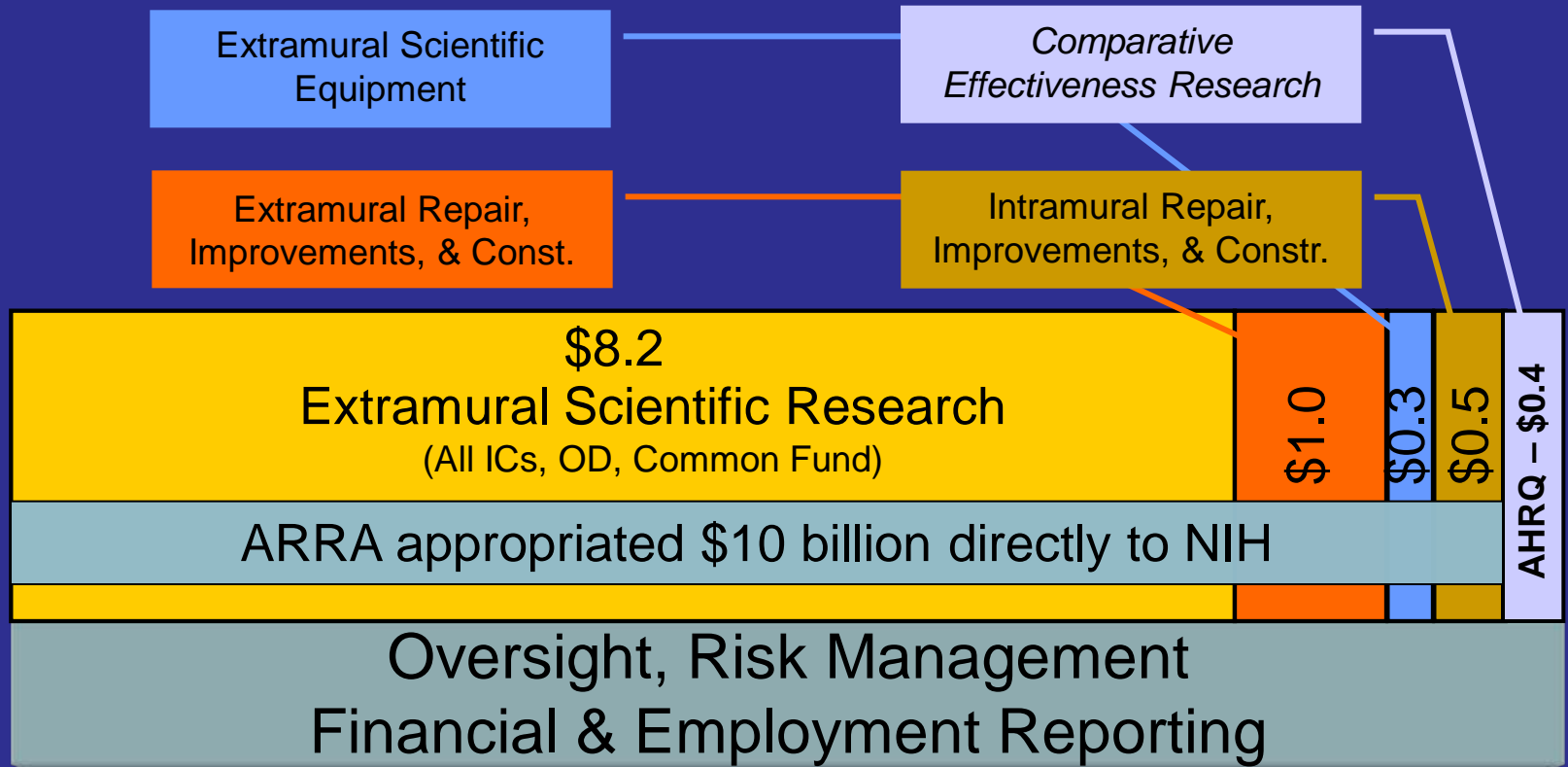
remingka@nigms.nih.gov



NIH Implementation of ARRA



NIH Allocation of ARRA Funds Dollars In Billions



Funding Impact

- Stimulate the economy
- Create and preserve jobs
- Advance biomedical research



http://www.nih.gov/recovery



The screenshot shows the NIH website's 'Recovery' section. At the top, it features the U.S. Department of Health & Human Services logo and the NIH logo with the tagline 'The Nation's Medical Research Agency'. A navigation menu includes links for HOME, HEALTH, GRANTS, NEWS, RESEARCH, INSTITUTES, and ABOUT NIH. A search bar is located in the top right corner. The main content area is titled 'NIH and the ARRA' and 'American Recovery & Reinvestment Act'. It includes a 'Quick Links' sidebar with options like Mission, Leadership, Research Planning, Budget, Organization, History, Photo Galleries, NIH Almanac, Outreach & Education, Public Involvement, Small Business Opportunities, Jobs, FAQ, Visitor Information, Contact NIH, and Find NIH Staff. The main text provides an overview of the Recovery Act, signed in February 2009, and lists links for an overview and implementing programs. A 'RECOVERY.GOV' section offers more information about grant programs. An 'Announcements' section highlights the availability of \$1.5 billion in Recovery Act funds and the NIH Challenge Grants initiative.

U.S. Department of Health & Human Services

www.hhs.gov

Employee Info | Staff Directory | En Español

NATIONAL INSTITUTES OF HEALTH

National Institutes of Health
The Nation's Medical Research Agency

search

>> Advanced Search

HOME HEALTH GRANTS NEWS RESEARCH INSTITUTES ABOUT NIH

Home Email this page

NIH and the ARRA

Quick Links

- Mission
- Leadership
- Research Planning
- Budget
- Organization
- History
- Photo Galleries
- NIH Almanac
- Outreach & Education
- Public Involvement
- Small Business Opportunities
- Jobs
- FAQ
- Visitor Information
- Contact NIH
- Find NIH Staff

American Recovery & Reinvestment Act

✉ Sign up to receive NIH and the American Recovery & Reinvestment Act e-mail updates.

Overview of the American Recovery and Reinvestment Act of 2009 (Recovery Act). The American Recovery and Reinvestment Act of 2009 (Recovery Act) was signed into law by President Obama on February 17th, 2009. It is an unprecedented effort to jumpstart our economy, create or save millions of jobs, and put a down payment on addressing long-neglected challenges so our country can thrive in the 21st century. The Act is an extraordinary response to a crisis unlike any since the Great Depression, and includes measures to modernize our nation's infrastructure, enhance energy independence, expand educational opportunities, pres...

those in greatest need.

- » Overview of the Recovery Act
<http://www.hhs.gov/recovery/overview/index.html>

RECOVERY.GOV

- » Implementing the Recovery Act
<http://www.hhs.gov/recovery/programs/index.html>
- » Learn more about programs that issue grants under the Recovery Act
<http://grants.nih.gov/recovery/>

Announcements

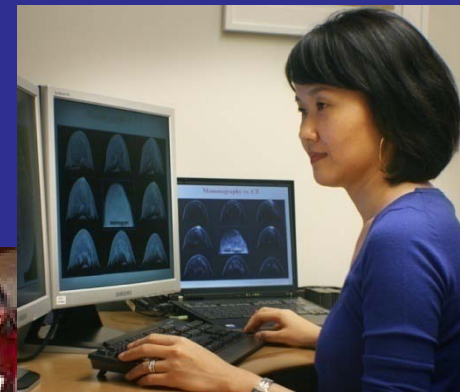
- » Applications for \$1.5 Billion in Recovery Act Funds Now Available
<http://www.nih.gov/news/health/mar2009/od-10.htm>
- » The NIH has designated at least \$200 million in FYs 2009—2010 for a new initiative called the NIH Challenge Grants in Health and Science Research. This new program will support research on topic areas that address specific scientific and health research challenges in biomedical and behavioral research that would benefit from significant 2-year

ARRA Grant Funding Opps
and Info at
Grants.nih.gov/Recovery

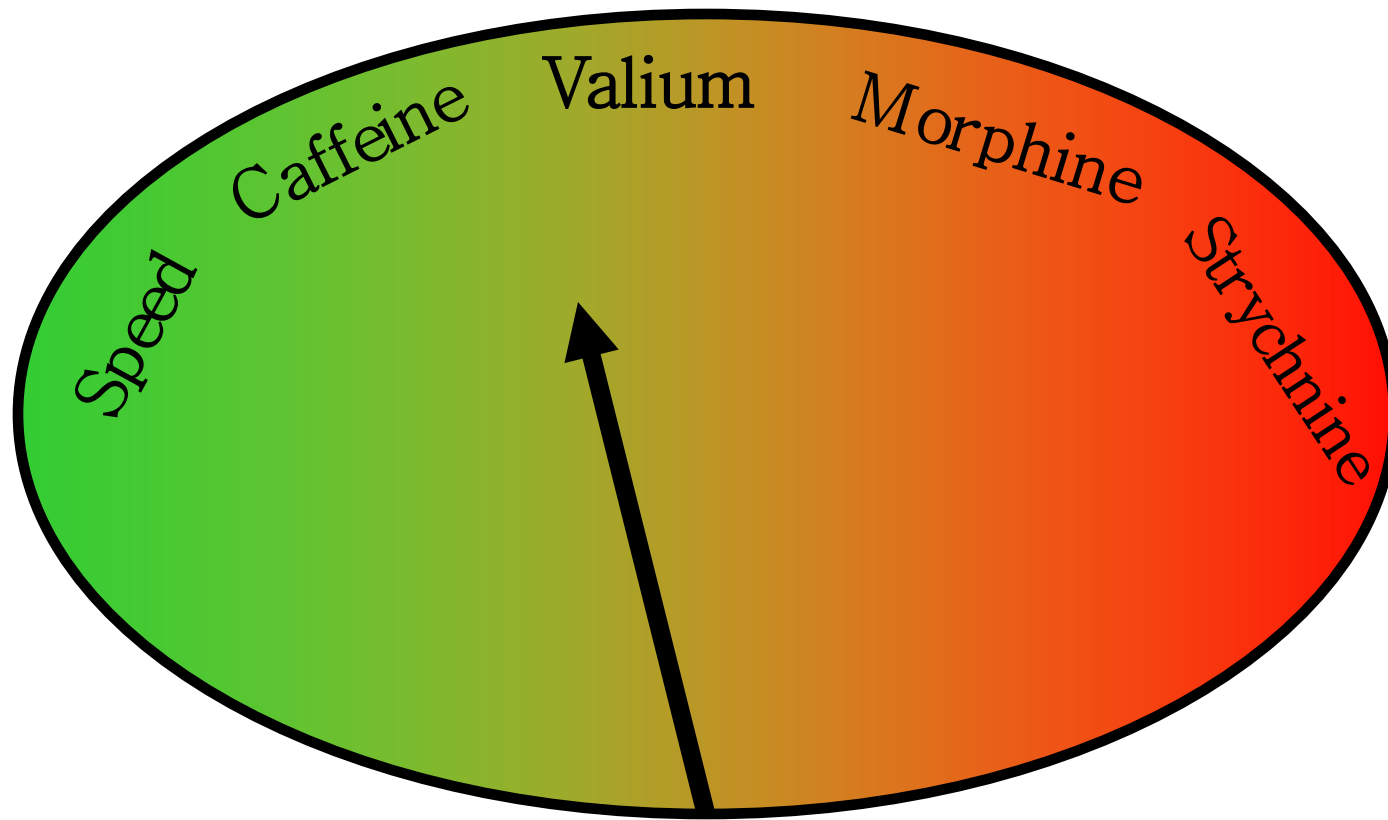


Scientific Research Approach

- Stimulate and accelerate biomedical research with existing mechanisms
- Expand science with new programs



NIGMS alone has \$500M over two years (...really 18 months)



ARRA Meter





Support of research and promotion, when it comes to interdisciplinary teams and emerging areas:



First Genome Sequenced 1995
Haemophilus influenzae

1.8 million basepairs





The genome sequence of *Drosophila melanogaster*.

Science. 2000 Mar 24;287(5461):2185-95

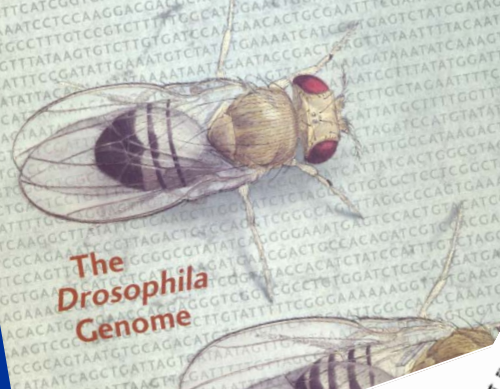
Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Linsley TR, Rose RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yant H, White MK, Anderson KL, Rowe LB, Gnanapavan R, McEwen GS, McBryden OJ, Chen Y, Chen Z, Chen L, Chen B, Brodeur GM, Clark RA, Doolittle R, Faloutsos M, Gabor GC, Ghosh S, Golub M, Gunter T, Harris M, Harte V, Heisterkamp R, Johnson D, Lachure J, Lee A, Leighton J, Liebman M, Lintz N, Luo C, Lusk E, Malek S, Mansoori S, McPherson D, McWilliam M, Miller N, Miller L, Mitchell B, Mortrud M, Murray K, Nelson D, Nelson J, Nelson P, Nelson S, Nelson T, Nizenz L, Patenaude B, Pevsner I, Platt R, Pockwinse S, Reagin J, Renwick S, Sandhu J, Scheer A, Schenck A, Shen W, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner D, Unger T, Wang AH, Wang JX, Wang ZY, Wassarman DA, Weinstock GM, Weissbach J, Weissman SM, Wedgate T, Wendler J, Westbrook JC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang J, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith DR, Adams RA, Myers EW, Rubin GM, Venter JC.

Remington K.

Science

24 March 2000

Vol. 287, No. 5461
Pages 2105-2364 \$8



The Drosophila Genome

A Comparison of Mouse Chromosomes

Richard J. Mural,¹
Ron Wilder,¹
Richard A. H.
Wentworth,¹
Ashur

¹The Institute for Genomic Research, Rockville, MD 20850, USA, ²The Center for Advancement of Genomics, Rockville, MD 20850, USA.

³These authors contributed equally to this work.

DECEMBER 2003 VOL 301

SCIENCE www.sciencemag.org

The high degree of similarity between the mouse and human genomes, demonstrated through analysis of the sequence of mouse chromosome assembly (Mmu16), which was obtained as part of a whole-genome shotgun assembly of the mouse genome, is at least 10% greater than the structure and protein-coding potential of Mmu16 with that of the human genome segments and predicted genes on human chromosomes 1-22. Genes that are highly conserved between Mmu16 and Mmu16, and 22, are likely orthologs to the corresponding portion of the human genome. At least 14 genes for which we could find no human counterparts.

The laboratory review is an invaluable model for helping us understand human biology and disease. The mouse genome sequence, in combination with the recently reported human genome sequence (1, 2), offers the opportunity to rapidly improve our understanding of the relative importance of mouse models of human

SCIENCE VOL 296 31 MAY 2002

The Dog Genome: Survey Sequencing and Comparative Analysis

Ewen F. Kirkness,¹ Vineet Bafna,^{2*} Aaron L. Halpern,^{2*} Samuel Levy,^{2*} Karin Remington,^{2*} Douglas B. Rusch,^{2*} Arthur L. Delcher,¹ Mihai Pop,¹ Wei Wang,¹ Claire M. Fraser,¹ J. Craig Venter²

A survey of the dog genome sequence (6.22 million sequence reads; 1.5% coverage) demonstrates the power of sample sequencing for comparative analysis of mammalian genomes and the generation of species-specific resources. More than 650 million base pairs (>25%) of dog sequence align uniquely to the human genome, including fragments of putative orthologs for 18,473 of 24,567 annotated human genes. Mutation rates, conserved synteny, repeat content, and phylogeny can be compared among human, mouse, and dog. A variety of polymorphic elements are identified that will be valuable for mapping the genetic basis of diseases and traits in the dog.

Our understanding of how the human genome functions in health and disease will benefit from comparison of its structure with the genomes of other species (1, 2). The domestic dog is a particularly good example, where an unusual population structure offers unique opportunities for understanding the genetic basis of morphology, behaviors, and disease susceptibility (3, 4). The physical and behavioral characteristics of ~300 dog "breeds" are maintained by restricting gene flow between breeds. Many modern breeds are derived from few founders and have been bred for desired characteristics. This has led to a species with enormous phenotypic diversity, but with significant homogenization of

the gene pool within breeds. Many of the ~360 known genetic disorders in dogs resemble human conditions, and their causes may be more tractable in large dog pedigrees than in small, outbred human families (4, 5). The combination of genetic homogeneity and phenotypic diversity also provides an opportunity to understand the genetic basis of many complex developmental processes in mammals (6).

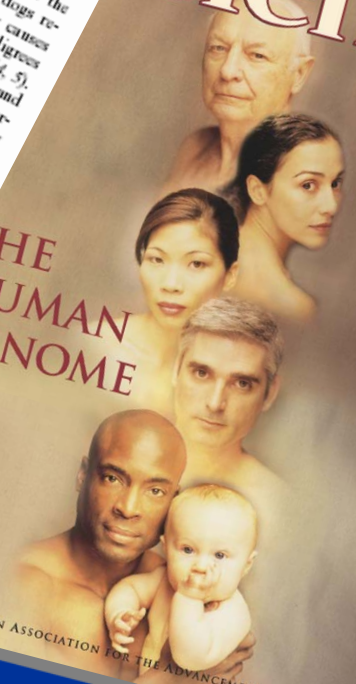
Because of the costs of sequencing mammalian genomes to completion, these projects have been restricted to a few species that are considered to be of greatest value to biomedical research. The decision as to whether future projects should aim for complete sequence coverage of a few more genomes, or whether the existing "reference genomes" can be exploited to characterize a wider variety of genomes that are sequenced to a lower level of coverage, must be made. Here,

Science

16 February 2001

Vol. 291 No. 5507
Pages 1145-1434 \$9

THE HUMAN GENOME



AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

THE *DROSOPHILA* GENOME

REVIEW

A Whole-Genome Assembly of *Drosophila*

Eugene W. Myers,^{1*} Granger G. Sutton,¹ Art L. Delcher,¹ Ian M. Dew,¹ Dan P. Fasulo,¹ Michael J. Flanigan,¹ Saul A. Kravitz,¹ Clark M. Mobarry,¹ Knut H. J. Reinert,¹ Karin A. Remington,¹ Eric L. Anson,¹ Randall A. Bolanos,¹ Hui-Hsien Chou,¹ Catherine M. Jordan,¹ Aaron L. Halpern,¹ Stefano Lonardi,¹ Ellen M. Beasley,¹ Rhonda C. Brandon,¹ Lin Chen,¹ Patrick J. Dunn,¹ Zhongwu Lai,¹ Yong Liang,¹ Deborah R. Nusskern,¹ Ming Zhan,¹ Qing Zhang,¹ Xiangqun Zheng,¹ Gerald M. Rubin,² Mark D. Adams,¹ J. Craig Venter¹

We report on the quality of a whole-genome assembly of *Drosophila melanogaster* and the nature of the computer algorithms that accomplished it. Three independent external data sources essentially agree with and support the assembly's sequence and ordering of contigs across the

end to sequence next in an interactive walk across the genome. Weber and Myers then proposed the whole-genome shotgun sequencing of the human genome in 1997 (8,





NIH Center for Bioinformatics and Computational Biology

Karin A. Remington, Ph.D.

remingka@nigms.nih.gov

