Databases to Support Disease-Focused Research

Type 1 Diabetes

Huntington's Disease

Nat Goodman
Institute for Systems Biology
January 2003

The Basic Idea

- Database (website) to support research of scientists working on diseases of interest
- ☞Key challenge: make it useful!
- Data must be
  - relevant to current research
  - rigorously accurate
  - timely
  - coordinated with other databases
- Steering committee provides scientific direction
- Also, easy-to-use, yadda, yadda, yadda

What Else is Like This?

- Other disease-focused websites
  - ✓ Alzheimer Research Forum (Alzforum) http://www.alzforum.org
  - ? ALS Therapy Development Foundation (ALS-TDF)
  - ✖ Technology × disease databases
    - Stanford breast cancer microarray website
  - Any others?
- Model organism databases
  - MGD, FlyBase, WormBase, TAIR, SGD, …
- Protein family databases
  - GPCRs, cytochrome P450s, …
- Locus-specific databases
  - HLA, CF, …
- Alliance for Cellular Signaling (AfCS)-Nature Gateway

Potential Data Scope

- Genomic regions
- Genes & proteins
  - functional summaries
  - curated sequences, genomic context, structures
  - orthologs, families, multiple alignments
- Microarray results
- Genotypes
- Protein-protein interactions
- Pathway models
- Empirical results on hot topics
- Reagents
  - antibodies, mouse models, clones, constructs, …
- Therapeutic studies
  - drug, transplantation, gene transfer
  - molecular, cellular, lower organism, mouse, other mammals
  - clinical
- Patient information
  - clinical & pathologic features
- Biomarkers
- Literature scanning and alerting
- Reports of negative and “ho-hum” results
- Lay explanations

Practical Concerns

Too much data

prioritize!

- ☞ Steering committee to the rescue

Too much overlap

collaborate!

- Alzforum
- RefSeq
- GO
- Stanford HOPES!!!
- OMIM
- BIND
- ? MGD
- ✖ MEDLINE

Too much software

reuse!

- Alzforum
- other collaborating databases
- PubCrawler
- GBrowse
- BioPerl
- Generic Model Organism Database (GMOD)

Some Differences Between Projects

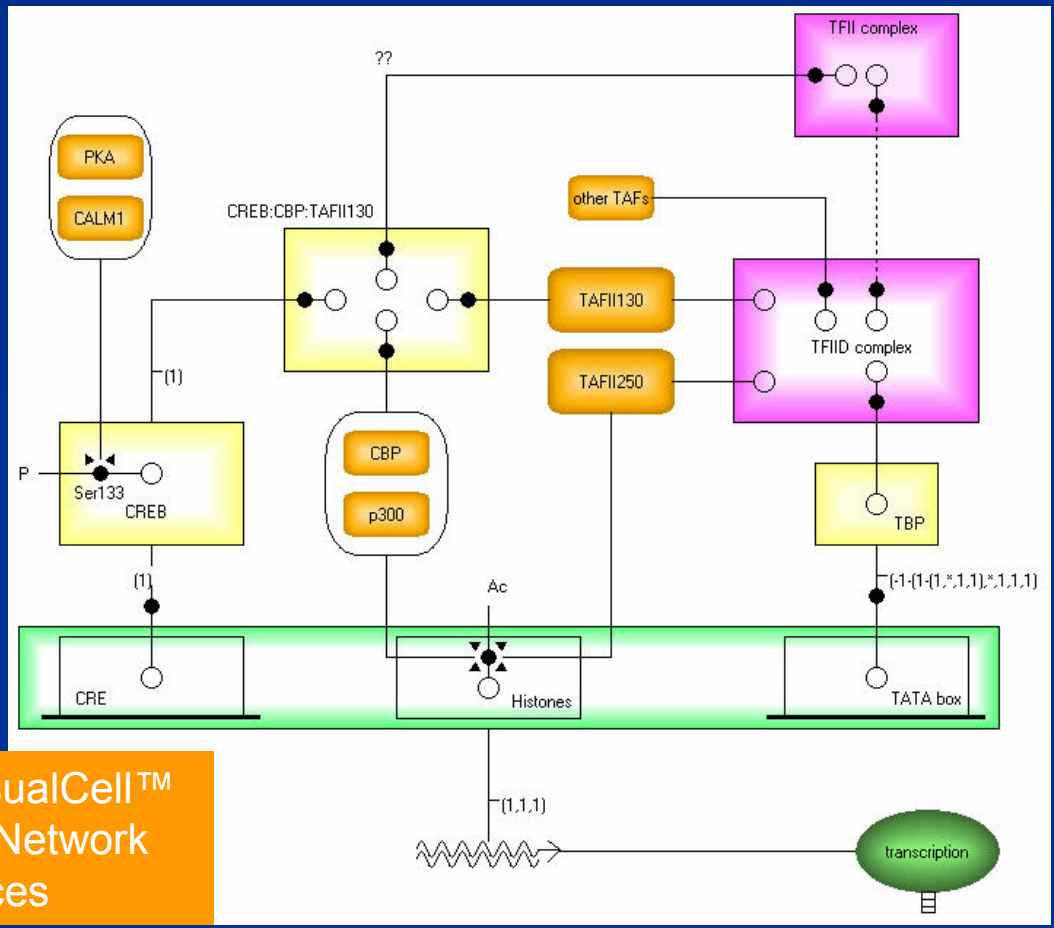
Data Type	Type 1 Diabetes	HD
Genomics	~17 susceptibility regions	Single gene disorder
Genes	Several hundred genes in susceptibility regions	~40 huntingtin (Htt) interactors ~100 genes of interest
Microarray	A few datasets available	Hereditary Disease Array Group led by Jim Olson Others ?
Genotyping	Consortium for fine-scale mapping	Two efforts to map age-of-onset modifiers
Therapies	Coordinated program for islet cell transplantation Gene & drug therapy Pharma, too!	Semi-coordinated program for drug screening Separate clinical studies Orphan disease

First Data Scope for HD Website

Data Type	Details
Large scale datasets	Mouse & molecular drug screening Protein-protein interactions (Hughes, Myriad Proteomics) Protein abundance in cerebrospinal fluid (Watts, ISB)
Gene list	Human, mouse, rat orthologs Sequences Functional summaries
Empirical results	Example: Htt interaction with transcription factors - binding, transcriptional activity, cell death
Reagents	Antibodies Genetic constructs
Pathway models	Hypothesized disease mechanisms Example: Htt & CREB-mediated transcription

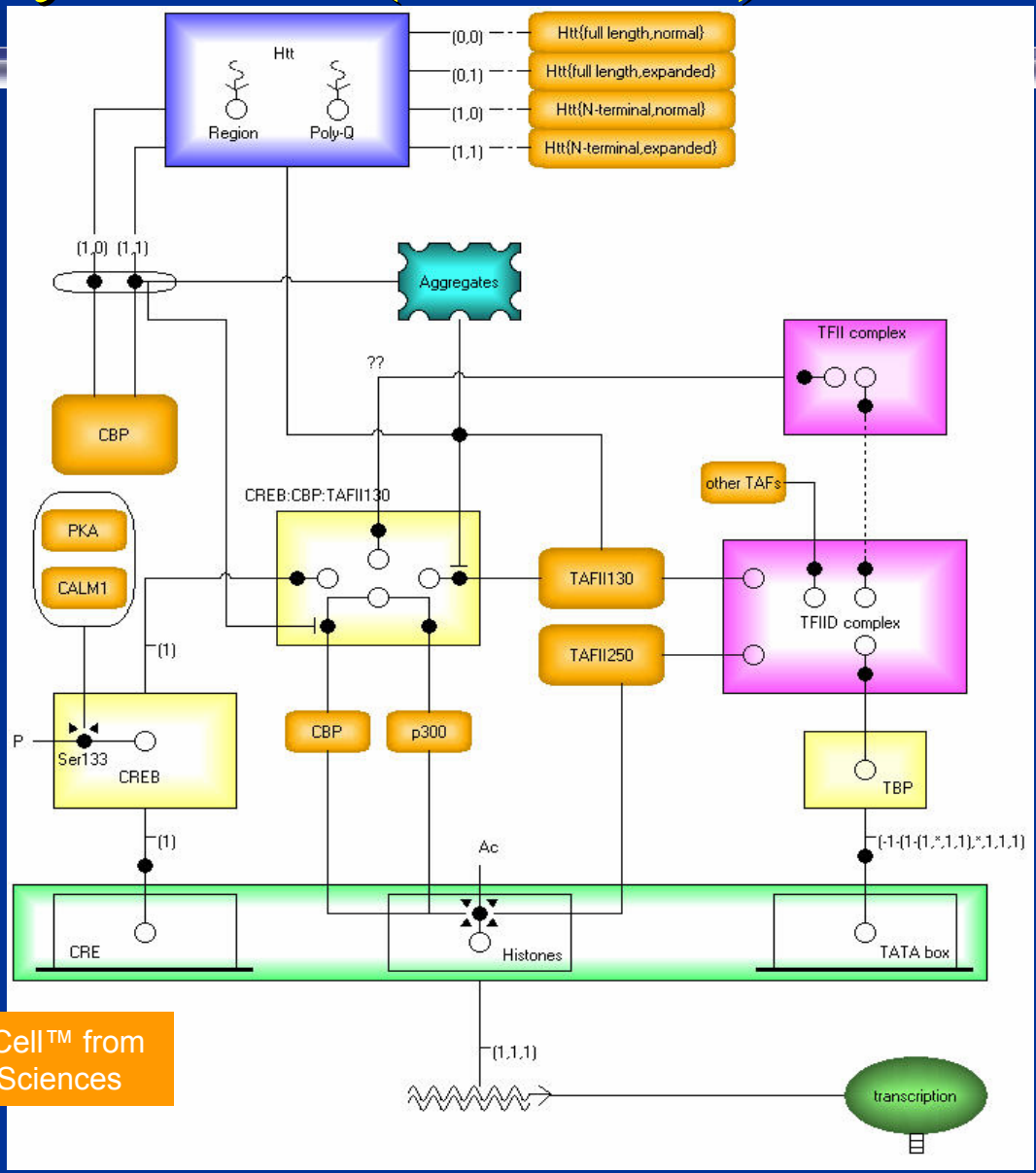
Pathway Model (Wild type)

Normal CREB-mediated transcription



Software: VisualCell™ from Gene Network Sciences

Pathway Model (Diseased)



Software: VisualCell™ from Gene Network Sciences

Steering Committee Response

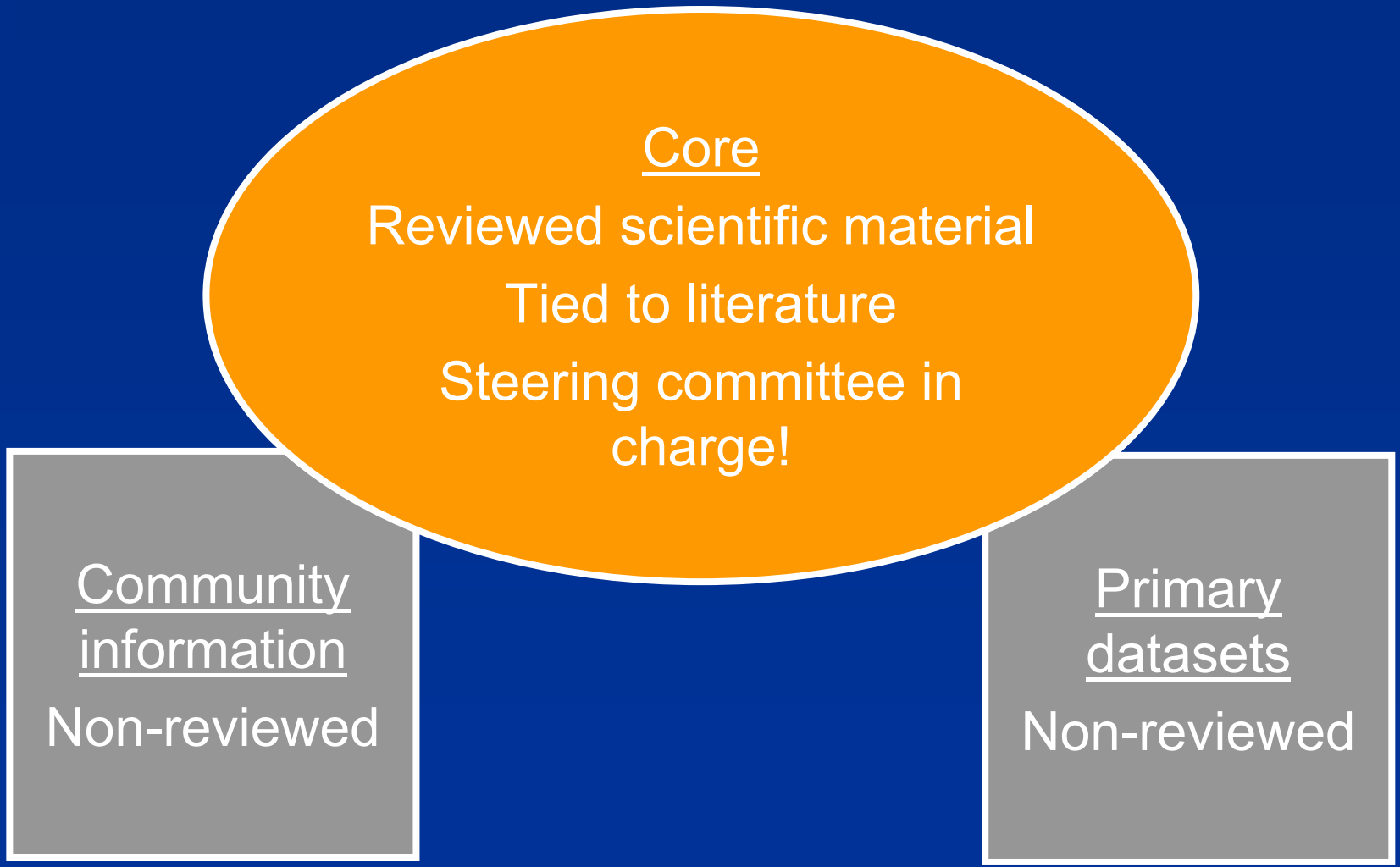
Steering Committee Guidelines

- Peer-review!
- Connect everything to literature
- Rigorously scrutinized, but diverse, science
- Data – “just the facts, Ma’am” – not conjecture
- Hypotheses presented as such – not as fact

My Response

Hmm… this is kinda narrow for a community website

Compromise


Core
Reviewed scientific material
Tied to literature
Steering committee in charge!
Community information
Non-reviewed
Primary datasets
Non-reviewed

Current Core Data Scope

Data Type	Details
Comprehensive bibliography Milestone papers	Annotation by curators & committee User comments
Published drug screens in mouse	Bibliography & dataset
Mouse models	Bibliography & dataset
Antibodies	Bibliography & dataset
Published microarray studies	Bibliography, lists of changed genes, links to full datasets
Gene list	Bibliography Human, mouse, rat orthologs Sequences Htt interactions Short functional descriptions

Current Core Services

- Genome / gene browser
  - View genes in human, mouse, rat syntenic regions
  - Accesses UC Santa Cruz DAS server plus local databases
  - All standard Santa Cruz information visible here, too
  - Based on GBrowse – collaboration with L. Stein
- Literature alerting
  - Specify MEDLINE queries
  - Can include our bibliographies
  - System runs periodically to get new hits
  - Based on PubCrawler– collaboration with K. Wolfe, K. Hokamp

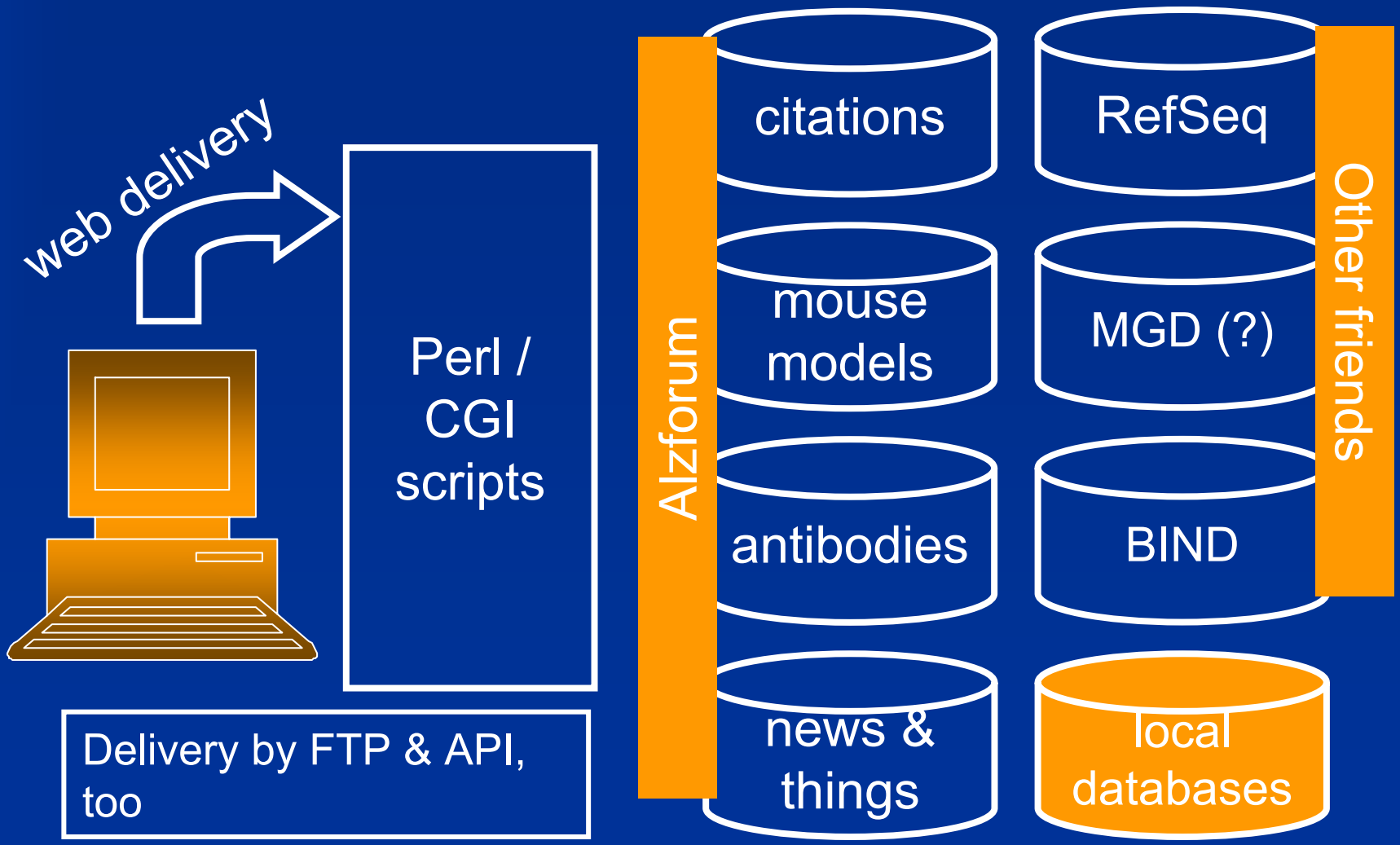
Current Satellite Data Scope

Data Type	Details
News	Like news in *Science* and *Nature*
Forum	Interviews with leading scientists Live discussions on hot topics with subsequent transcripts Web delivery of presentations Mini-reviews derived from above
Calendar of events	Conferences, etc.
Contact info for HD researchers	With permission!
Lay explanations	For major sections, at least
Primary datasets	Protein-protein interactions (Hughes) Protein abundance in CSF (Watts)

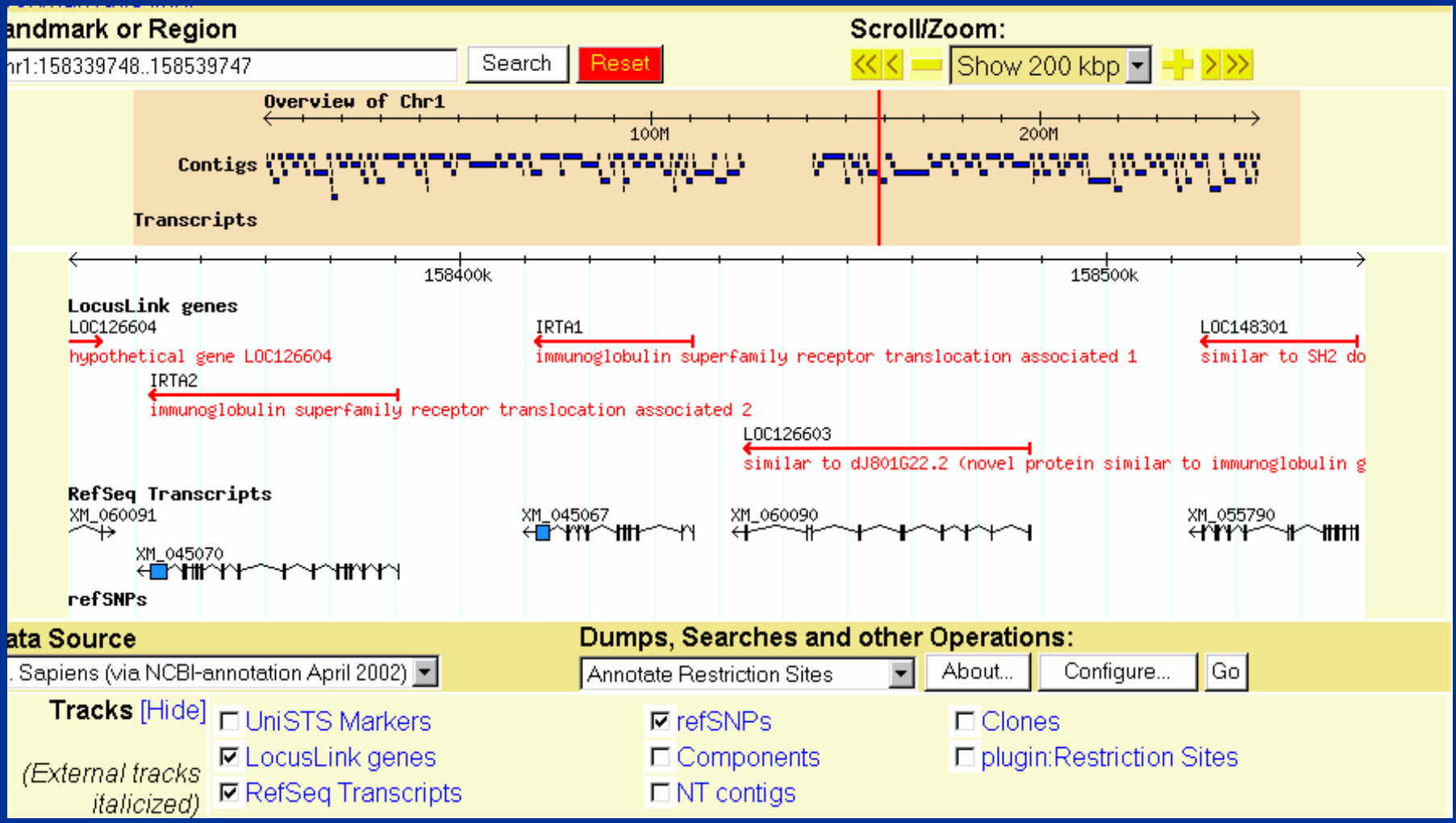
Help From Our Friends

Data Type	Who	What
All bibliographies	Alzforum	citation database
Comprehensive bibliography	Alzforum	scanning & librarian
Mouse models	Alzforum	database
Antibodies	Alzforum	database & curator
Published microarray studies	HDAG	data & review
Gene list	MGD	orthologs (we hope)
	RefSeq	sequences, descriptions
	GO	annotations
	BIND	Htt interactions
News, forum, calendar, contacts	Alzforum	
Lay explanations	HOPES	
Primary datasets	Myriad, ISB	data

Software Architecture



Genome Browser Screenshot



ALZHEIMER RESEARCH FORUM

NETWORKING FOR A CURE

Search | Get Newsletter

advanced search

WHAT'S NEW

Home | Contact Us | Become a Member | Your Profile | Login

PAPERS OF THE WEEK

- Current Papers
- ARF Recommends
- Milestone Papers
- Search All Papers

NEWS

- Research News
- Drug News

RESEARCH

- AD Hypotheses
  - Current Hypotheses
  - Hypothesis Factory
- Forums
  - Live Discussions
  - Virtual Conferences
  - Interviews
- Compendia
  - Antibodies
  - Mutations
  - Genes
  - Transgenic Mice
  - Patents
- Resources
  - Jobs
  - Conference Calendar
  - Library
  - Grants
  - Research Tools

DISEASE MANAGEMENT

- About Alzheimer's
  - FAQs
- Diagnosis
  - Clinical Guidelines
  - Tests
  - Brain Banks
- Treatment
  - Drugs and Therapies
  - Drugs in Clinical Trials
- Caregiving
  - Patient Care
  - Support Directory
  - Eldercare Chats

RESEARCH NEWS

- MRI Keeps an Eye on Those Transplanted Stem Cells In Vivo
  19 November 2002. Embryonic stem (ES) cells have the inherent potential to replace or repair almost every tissue in the body. One key to this prowess is the ability to migrate...

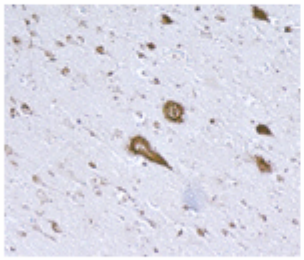
- Spinal Taps for Tau Tests
  18 November 2002. Assays of Aβ or tau protein in cerebrospinal fluid can identify Alzheimer's disease in patients with mild cognitive impairment, say the authors of a study in today's Archives of Neurology...

- Early Results Hint That Insulin-Sensitizing Drug Improves Cognition
  18 November 2002. Insulin resistance is emerging as a possible risk factor for Alzheimer's, and data from a pilot trial presented at the 2002 Neuroscience meeting indicate that an insulin-sensitizing drug may represent a novel treatment approach...

GO TO ALL RESEARCH NEWS

IN THE SPOTLIGHT

- Virtual Conference: Subcellular Aβ Accumulation and Origin of Plaques in Alzheimer's Disease



Presentation (with audio) from the 8th International Conference on Alzheimer's Disease and Related Disorders, held in Stockholm, Sweden, 21-25 July 2002. Presentation by Gunnar Gouras.

Abstract: We have been studying the generation and subcellular localization of Aβ neurons by biochemical and immunocytochemical assays, and more recently have turned to immuno-gold electron microscopy (EM) to determine the more precise localization of Aβ peptides and amyloid precursor protein (APP) within neurons of normal brain and to study the earliest site of Aβ accumulation and neuropathology with aging in AD mouse models of β-amyloidosis...

Forum Calendar

Upcoming Live Discussions

26 November 2002 at 12 noon-1pm EST: "Intraneuronal Aβ Accumulation-More Evidence, Less Controversy?" Live discussion with Gunnar Gouras.

Desperately Seeking

- Antibodies
- Collaborators

Polls & Opinions

Might neurogenesis play a role in AD?

- Loss of neurogenesis could contribute to AD.
- It doesn't occur at a high enough rate to be significant.
- A possible role for APP and PS1, someone please

PAPERS OF THE WEEK

Words in Title, and Abstract: AND
Author: Select Author (e.g. Doe, J) AND
Comments By: Select Commentator AND
Journal: Select Journal AND
Current Papers: AND
Publication Date: From
To AND
Limited To: Annotated ARF Recommends Milestone
Sort By:
Reset ARF Search

ARF RECOMMENDS

1 to 20 of 246 results BACK | NEXT

ARF RECOMMENDS

Boon K, Osorio EC, Greenhut SF, Schaefer CF, Shoemaker J, Polyak K, Morin PJ, Buetow KH, Strausberg RL, De Souza SJ, Riggins GJ. **An anatomy of normal and malignant gene expression.** *Proc Natl Acad Sci U S A.* 2002 Aug 20;99(17):11287-92. Pubmed Abstract

VOTE/COMMENT VIEW ANNOTATION

ARF RECOMMENDS

Panov AV, Gutekunst CA, Leavitt BR, Hayden MR, Burke JR, Strittmatter WJ, Greenamyre JT. **Early mitochondrial calcium defects in Huntington's disease are a direct effect of polyglutamines.** *Nat Neurosci.* 2002 Aug ;5(8):731-6. Pubmed Abstract

VOTE/COMMENT VIEW ANNOTATION

ARF RECOMMENDS

Town T, Zolton J, Shaffner R, Schnell B, Crescentini R, Wu Y, Zeng J, DelleDonne A, Obregon D, Tan J, Mullan M. **p35/Cdk5 pathway mediates soluble amyloid-beta peptide-induced tau phosphorylation in vitro.** *J Neurosci Res.* 2002 Aug 1;69(3):362-372. Pubmed Abstract

VOTE/COMMENT VIEW ANNOTATION

ARF RECOMMENDS

Hartig JS, Najafi-Shoushtari SH, Grüne I, Yan A, Ellington AD, Famulok M. **Protein-dependent ribozymes report molecular interactions in real time.** *Nat*

Related Links

- Science's SAGE

- Coordination Chemistry Reviews Special Issue on Recent Topics in Aluminium Chemistry

Journal of Alzheimer's Disease

Free access to ARF members until the end of 2002!

Special Issue: Challenging Views of AD

Please use the following access codes:
Username: IOSUser1
Password: ios0508

Just Published!

Research and Practice in Alzheimer's Disease 2002.
Vellas B. et al.
SERDI Publisher.

Mutations	Transgene/ Promoter and Regulatory Elements	Genetic Background	Behavioral Phenotype	Neurological Characteristics	Patents/ Availability	Primary Citation
Tau P301L (JNPL3)	Longest human Tau isoform with 4 repeats containing exon 10 and lacking exons 2 and 3 with P301L/ mouse prion promoter (MoPrP)	C57BL/DBA2/SW	Severe motor and behavioral disturbances observed early	Fibrillary gliosis in the anterior horns, axonal degeneration, neuronal lesions similar to FTDP-17	Unpatented Taconic Farm Under Development	Lewis J et al
Tau P301L	Human tau40 isoform with 4 repeats, exons 2 and 3 with P301L/ neuron-specific mouse Thy1.2 promoter	B6D2F1 x B6D2F1, founder animals were intercrossed with C57BL/6 mice to establish lines	Signs of Wallerian degeneration, neurogenic muscle atrophy, muscle weakness.	Numerous abnormal, tau-reactive nerve cell bodies and dendrites; large numbers of pathologically enlarged axons containing neurofilament and tau-reactive spheroids	Unpatented Contact Jurgen Gotz goetz@bli.unizh.ch	Gotz J et al
Tau G272V	Human tau40 with G272V mutation/ murine prion protein promoter	B6D2F1 x B6D2F1 founder mice inter-crossed with C57BL/6 mice	No neurological deficits readily noticeable	Filaments in murine oligodendrocytes, associated with tau phosphorylation at AT8 epitope 202/205 in vivo. In the spinal cord, fibrillary inclusions identified by thioflavin-S in oligodendrocytes and motor neurons	Unpatented Contact Jurgen Gotz goetz@bli.unizh.ch	Gotz J et al
Tau V337M	Human longest tau cDNA with V337M mutation/PDGF-b promoter, neuron-specific mouse Thy1.2 promoter	B6SJL background	Considerably smaller in size. Higher overall spontaneous locomotion . No significant difference in a Morris water	Neurons of irregular shape in hippocampus were immunoreactive for paired helical filament-associated tau, and showed signs of atrophic cell death	Unpatented Contact Contact Akihiko Takashima kenneth@brain.riken.go.jp	Tanemura et al.

A Few Words About IP

- Open source
- Open data
- Strong privacy

Four Rules for a Successful Website

1. Too much data
   - Prioritize!
   - What will be most useful?
   - Rely on scientific experts
2. Too much software
   - Reuse!
   - Lots of great software available
   - Developers willing to help
3. Too much overlap
   - Collaborate!
   - Many databases welcome this
   - Less work – better product -- more fun!
4. Obsess on quality
   - Bad data wastes everyone's time

Acknowledgements

ISB Project Team

George Lake
Michelle Whiting
Paul Edlefsen
Robert Hubley

HDF

Carl Johnson
Minka van Beuzekom

Alzforum

June Kinoshita

RefSeq

Kim Pruitt

GO Consortium

Evelyn Camon

HOPES

Bill Durham

HDAG

Jim Olson

Myriad Proteomics

Bob Hughes

ISB

Julian Watts

Steering Committee

Carl Johnson and Minka van Beuzekom, HDF

Dan Goldowitz
University of Tennessee

Emma Hockly
Guy's Hospital

Bruce Kristal
Cornell University

Marcy MacDonald
Massachusetts General Hospital

Ray Truant
McMaster University