Informatics and Visualization Tools for Structural Genomics Research

Tom Ferrin, Ph.D.
Departments of Pharmaceutical Chemistry and Biopharmaceutical Sciences
Resource for Biocomputing, Visualization, and Informatics
University of California, San Francisco
Resource for Biocomputing, Visualization, and Informatics

The RBVI is a NIH/NCRR Biomedical Technology Research Center

We create innovative computational and visualization-based data analysis methods and algorithms, turns these into easy-to-use software tools, and apply these tools for solving a wide range of genomic and molecular recognition problems within the complex sequence → structure → function triad
Application areas

Gene characterization and interpretation
Drug design
Variation in drug response due to genetic factors
Protein engineering
Biomaterials design
Bioremediation
Prediction of protein function from sequence and structure
“It’s sink or swim as a tidal wave of data approaches”

- Petabyte (1,000 terabytes)
- Exabyte (1,000 petabytes)
- Zettabyte (1,000 exabytes)
- Yottabyte (1,000 zettabytes)

Tony Reichhardt
Nature 399:517-520 10 June 1999
“Many biologists are still in denial, never having faced the amount of information now pouring into databases such as Genbank and SwissProt… They haven’t really thought about how they’re going to use all this data…”

Ibid.
The Growing Gap in Functional Knowledge

- Rapid DNA sequencing invented
- EST sequencing begun

Graph showing the increasing number of publications and DNA sequences over time from 1975 to 1995.
Sample RVBI projects

• New methods for large-scale data collection, storage, analysis, and presentation for polymorphism (SNP) genotyping project

• Extensible visualization tools for comparative studies of protein sequence, structure, and function
ADVERSE REACTIONS

Genetic tests to prevent adverse drug reactions may save tens of thousands of lives a year, but for a troubled boy named Michael they came too late.

By David Stipp
Photography by Suzanne Captan

The death of nine-year-old Michael Adams-Corboy didn't seem at first like a signal event in medicine. It seemed like homicide.

Michael's short life was an uphill struggle from the start. Malnourished and anemic at birth, he was taken from an abusive mother and placed in a temporary foster home before his first birthday. By the time he was 6, his medical record bulged with bad news: Michael was cognitively delayed and exhibited mood and attention disorders, as well as with obsessive-compulsive disorder, tic-inducing Tourette's syndrome, and attention-deficit hyperactivity disorder.

Over the next few years he achieved a semblance of normality, thanks to the steady hands of the resourcefully effective couple who adopted him at age 2 and the daily doses of drugs to check his tics and obsessions. Small for his age, he took pride at finally being able to fight his way up onto the promenade rocks at his home on Manasquan Creek, N.J., a once-overgrown town two hours north of Philadelphia. He was learning to bowl in a league for handicapped kids and help his dad tend the garden.

October 30, 2000  F O R T U N E  •  171
Case Report #1: Michael Adams-Conroy

Young child born to abusive mother, adopted at age 3, with signs of fetal alcohol syndrome, obsessive-compulsive disorder, Tourette's syndrome, and attention-deficit hyperactivity disorder. Prescribed Prozac to help control emotional outbursts.

Child dies suddenly; toxicology tests show massive overdose of Prozac. Adoptive parents investigated for homicide and their other two children put into protective custody.
Sharp-eyed psychiatrist notices unusually high levels of other metabolites in toxicology report, indicating child may have had an enzyme deficiency inhibiting Prozac from being metabolized normally.

Subsequent genetic testing showed child had defect in 2D6 gene which resulted in abnormal liver enzyme that metabolizes antidepressants.

Adoptive parents exonerated.
Case Report #2

Patient: 3-year old boy

Diagnosis: Acute Lymphoblastic Leukemia (ALL)

Standard therapy: 6-mercaptopurine (6-MP)

Result: Adverse Drug Reaction leading to acute bone marrow suppression
Normal Mechanism of Action

6-MP

TPMT

6-METHYLMERCAPTOPURINE

S

CH₃

S

6-THIOGUANINE NTs
THIOPURINE METHYLTRANSFERASE (TPMT) GENES ARE DEFECTIVE IN 1:300 PEOPLE
This leads to elevated levels of Thioguanine Nucleotides
PEOPLE DIFFER IN THEIR RESPONSE TO DRUGS

- NO RESPONSE
- THERAPEUTIC RESPONSE
- ADVERSE DRUG REACTION (ADR)
TESTING FOR TPMT GENES IS NOW AVAILABLE
CHILDREN WITH DEFECTIVE TPMT GENES SHOULD RECEIVE A LOWER DOSE OF 6-MP
## Adverse Drug Reactions

**ADRs may kill 30,000 - 40,000 Americans each year and cause 2,200,000 serious nonfatal reactions.**  
*JAMA 1998 June 3;279(21):1684*

### Drugs with known genetically-linked potential for fatal adverse reactions (partial list):

<table>
<thead>
<tr>
<th>Drug (Brand Name)</th>
<th>Prescribed For...</th>
<th>Adverse Reaction</th>
<th>Gene at Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imipramine (Tofranil)</td>
<td>Depression, ATD</td>
<td>Heartbeat irregularity</td>
<td>CYP2D6</td>
</tr>
<tr>
<td>Isoniazid (Laniazid)</td>
<td>Tuberculosis</td>
<td>Liver toxicity</td>
<td>NAT2</td>
</tr>
<tr>
<td>Warfarin (Coumadin)</td>
<td>Prevention of blood clots</td>
<td>Internal bleeding</td>
<td>CYP2C9</td>
</tr>
<tr>
<td>5-fluorouracil (Adrucil)</td>
<td>Cancer</td>
<td>Severe immune suppression</td>
<td>DPD</td>
</tr>
<tr>
<td>Clarithromycin (Biaxin)</td>
<td>Antibiotic</td>
<td>Heartbeat irregularity</td>
<td>KCNE2</td>
</tr>
<tr>
<td>Azathioprine (Imuran)</td>
<td>Rheumatoid arthritis</td>
<td>Severe immune suppression</td>
<td>TPMT</td>
</tr>
</tbody>
</table>

ADRs may kill 30,000 - 40,000 Americans each year and cause 2,200,000 serious nonfatal reactions. **JAMA 1998 June 3;279(21):1684**

### Adverse Drug Reactions

ADRs may kill 30,000 - 40,000 Americans each year and cause 2,200,000 serious nonfatal reactions. **JAMA 1998 June 3;279(21):1684**

### Drugs with known genetically-linked potential for fatal adverse reactions (partial list):

<table>
<thead>
<tr>
<th>Drug (Brand Name)</th>
<th>Prescribed For...</th>
<th>Adverse Reaction</th>
<th>Gene at Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imipramine (Tofranil)</td>
<td>Depression, ATD</td>
<td>Heartbeat irregularity</td>
<td>CYP2D6</td>
</tr>
<tr>
<td>Isoniazid (Laniazid)</td>
<td>Tuberculosis</td>
<td>Liver toxicity</td>
<td>NAT2</td>
</tr>
<tr>
<td>Warfarin (Coumadin)</td>
<td>Prevention of blood clots</td>
<td>Internal bleeding</td>
<td>CYP2C9</td>
</tr>
<tr>
<td>5-fluorouracil (Adrucil)</td>
<td>Cancer</td>
<td>Severe immune suppression</td>
<td>DPD</td>
</tr>
<tr>
<td>Clarithromycin (Biaxin)</td>
<td>Antibiotic</td>
<td>Heartbeat irregularity</td>
<td>KCNE2</td>
</tr>
<tr>
<td>Azathioprine (Imuran)</td>
<td>Rheumatoid arthritis</td>
<td>Severe immune suppression</td>
<td>TPMT</td>
</tr>
</tbody>
</table>
Pharmacogenetics of Membrane Transporters

$12$-million, 4-year NIH grant
• Kathleen Giacomini and Ira Herskowitz, co-PIs, plus ~20 other UCSF researchers

Major Project Goal:
• Understand the genetic basis for variation in response to drugs which interact with membrane transporters. This class of proteins is of great pharmacological importance, as it provides the target for about 30% of the most commonly used prescription drugs and is a major determinant of the absorption, distribution and elimination of many others.
PMT project goals - continued

- Determine the amount of genetic variation (single-nucleotide polymorphisms) in at least 40 transporter genes by examining the DNA from an ethnically diverse sample of 250 people.

- Test the performance of these transporter variants in cell cultures and determine, through clinical phenotype studies, if people with those variants respond differently to drugs in a clinically significant way.

- Provide access to the data from these studies to the general scientific community through the World Wide Web to facilitate collaborative research and to speed development of new drug treatments.
The Corriel Cell Collection

African American (AA) - 100
Caucasian (CA) - 100
Asian American (AS) - 30
Mexican American (ME) - 10
Pacific Islander (PA) - 7

TOTAL - 247
UCSF Pharmacogenetics of Membrane Transporters

The UCSF Pharmacogenetics of Membrane Transporters (GMT) Project is sponsored by the National Institutes of Health's National Institute of General Medical Sciences (grant U01 GM61190). The Project is part of the Pharmacogenetics Research Network and Knowledgebase. Information about the entire Network can be found at PharmGKB.

Pharmacogenetics is the study of the genetic basis for variations from person to person in response to drugs. Membrane transporters play a major role in drug response in two ways. First, many drugs work by affecting function of transporters. Second, transporters determine the level of drugs within the body and thus determine whether drug levels are adequately high for therapeutic effect. The goal of the UCSF GMT Project is to understand the genetic basis for variation in drug response for drugs which interact with membrane transporters.

Model of organic cation transporter (OCT3)
PMT Intranet Website

Used by ~100 researchers at UCSF

Effective data analysis and display driven by iterative design/refinement cycle, successful because the bioinformatics team works closely with the molecular biologists

• Jill Mesirov, Whitehead: “Bioinformatics needs to be tightly integrated with the scientific research, not a service function”

Flexibility key!

• Multiple ways to display same data
• Simple download mechanism for scientists who want to load raw data into Excel spreadsheets
PMT Scientist-Users Are a Demanding Bunch...
OCT1 Transporter

PMT Project Investigator
Dr. Kathleen Giacomini

HGNC symbol
SLC22A1

Chromosome
6q26

Aliases
ORGANIC CATION TRANSPORTER; OCT1
SOLUTE CARRIER FAMILY 22, MEMBER 1; SLC22A1

Background information
NCBI data
Opmrn data
LocusLink data

Reference Entry
U77086.1
Homo sapiens organic cation transporter 1 (hOCT1) mRNA, complete cds.

Exons
A schematic representation of the gene showing its exons with the exons scaled in size relative to one another. Variant results are indicated by color(s) of the exon block.

Click on number to view desired individual exon results.

<table>
<thead>
<tr>
<th>Exon</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>red</td>
<td>blue</td>
<td>green</td>
<td>white</td>
<td>gray</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes</td>
<td>non-synonymous</td>
<td>indels (insertions &amp; deletions)</td>
<td>synonymous</td>
<td>intronic</td>
<td>no changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Amplicons
UCSC Blat search of all amplicons

Variants

- Variant identification summary (tab-delimited text version)
- Variant identification summary by ethnicity
- Variant identification per-sample data (list) (tab-delimited text version)
- Variant identification per-sample data (plot)
# OCT1, Exon 1

**Polymorphisms and Allele Frequencies**

<table>
<thead>
<tr>
<th>Exon</th>
<th>SNP #</th>
<th>CDS Pos</th>
<th>Exon Pos</th>
<th>Nucleotide Change</th>
<th>Amino Acid Position</th>
<th>Amino Acid Change</th>
<th><strong>Total Freq</strong></th>
<th>AA Freq</th>
<th>CA Freq</th>
<th>AS Freq</th>
<th>ME Freq</th>
<th>PA Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>41</td>
<td>41</td>
<td>C -&gt; T</td>
<td>14</td>
<td>Ser -&gt; Phe</td>
<td>0.013 (0.000)</td>
<td>0.031 (0.007)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>67</td>
<td>67</td>
<td>C -&gt; G</td>
<td>23</td>
<td>Leu -&gt; Val</td>
<td>0.002 (0.974)</td>
<td>0.005 (0.956)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>113</td>
<td>113</td>
<td>G -&gt; A</td>
<td>38</td>
<td>Gly -&gt; Asp</td>
<td>0.002 (0.974)</td>
<td>0.000 (n/a)</td>
<td>0.005 (0.959)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>150</td>
<td>156</td>
<td>T -&gt; C</td>
<td>52</td>
<td>syn</td>
<td>0.262 (0.437)</td>
<td>0.263 (0.260)</td>
<td>0.206 (0.939)</td>
<td>0.433 (0.385)</td>
<td>0.278 (0.249)</td>
<td>0.286 (0.427)</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>181</td>
<td>181</td>
<td>C -&gt; T</td>
<td>61</td>
<td>Arg -&gt; Cys</td>
<td>0.031 (0.619)</td>
<td>0.000 (n/a)</td>
<td>0.072 (0.444)</td>
<td>0.000 (n/a)</td>
<td>0.056 (0.860)</td>
<td>0.000 (n/a)</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>253</td>
<td>253</td>
<td>C -&gt; T</td>
<td>85</td>
<td>Leu -&gt; Phe</td>
<td>0.004 (0.949)</td>
<td>0.010 (0.919)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
<td>0.000 (n/a)</td>
</tr>
</tbody>
</table>

Values in red have a frequency of 0.010 or higher.

**Amplicon**

Color Scheme: Primer Intron Exon SNP
Sequencing Interrogation: Primer Intron Exon SNP

PCR primers:

```
TGGAGGAAGCAATGACCTCTGACCACTGACACTGACAGATGCTGACACGCCTGTGAGAGCTGACCCCTC
CTGGTCTGACGTCGCTTGACACGTGGCCTGGGGTGGTGGTGTTGGTGGCTGAC
```

SNP information: mouse over SNP for information
**Variants**


- Variant identification summary (tab-delimited text version)
- Variant identification summary by ethnicity
- Variant identification per-sample data (list) (tab-delimited text version)
- Variant identification per-sample data (plot)
- Population genetics statistics (tab-delimited text version)
- Consensus sequences with mammalian species

**Transmembrane prediction** for OCT1 protein. Non-synonymous amino acid changes shown in **red**, indels (insertions and deletions) in **blue**, and synonymous changes in **green**.

---

**Cellular Phenotyping Results**

**Clinical Studies**

- Influence of hOCT1 Genotypes on 123I-MIBG Distribution to the Liver
- Effect of Genetic Variation in the Liver Transporter, OCT1, on Response to Metformin in Healthy Subjects
You've now seen DNA and AA sequences. What about structure?
Why is Structure Important?

Sequence → Structure → Function

- Current research areas:
  - Prediction of structure from sequence
  - Prediction of function from sequence and structure
  - Understanding evolutionary changes
  - Engineering proteins for specialized function

- Applications in pharmacogenomics ...
  - Improvements in drug discovery and development process
  - Prediction of drug response
  - Avoidance of toxic side effects
Growth in Protein Structures

Years: 1972 to 2004

- Deposited structures for the year
- Total available structures

Last updated: 01-Mar-2004
The Structural Genomics Initiatives

“The next step beyond the human genome project”

$150 million in NIH grants to establish 9 U.S. centers

• Goals:
  - Speed the determination of three-dimensional atomic-scale maps of proteins
  - 35,000 structures by 2005
  - Identify all proteins expressed in an organism – “proteomics”

See http://www.nigms.nih.gov/funding/psi.html for additional information

<table>
<thead>
<tr>
<th>Center</th>
<th>Lead Institution</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southeast Collab. for Struct. Gen.</td>
<td>Univ. of Georgia</td>
<td>Bacteria/roundworm/human</td>
</tr>
<tr>
<td>Berkeley Struct. Genomic Center</td>
<td>Lawrence Berkeley Lab.</td>
<td>Bacteria</td>
</tr>
<tr>
<td>Joint Ctr. for Struc. Genomics</td>
<td>Scripps Research Inst.</td>
<td>Roundworm/human</td>
</tr>
<tr>
<td>TB Struct. Genomics Consortium</td>
<td>Los Alamos Nat. Lab.</td>
<td>Tuberculosis</td>
</tr>
<tr>
<td>Midwest Ctr. for Struct. Genomics</td>
<td>Argonne National Lab.</td>
<td>Archaea/bacteria/eukarya</td>
</tr>
<tr>
<td>Ctr. for Eukaryotic Struct. Genomics</td>
<td>Univ. of Wisconsin</td>
<td>Arabidopsis thaliana</td>
</tr>
<tr>
<td>Struct. Gen. of Pathogenic Protozoa</td>
<td>Univ. of Washington</td>
<td>Protozoans</td>
</tr>
</tbody>
</table>
Stereo pairs?
Chimera is an extensible interactive 3-D modeling system designed to allow developers to quickly incorporate novel algorithms and analysis tools

- ~30 extensions written to date
- Extensions are written in the Python programming language
  - Easy to learn, even for novice programmers
  - Offers object-oriented language features
- Extensions can control standard user interface features (e.g. camera, help, menus, toolbar) as well as their own custom interfaces
Sample Chimera Extension

Multalign Viewer
- simultaneous display of protein sequence and structure
Chimera Demo
Tools for Comparative Protein Studies

**MinRMS** - exhaustive search for all plausible structural alignments of two proteins

**AlignPlot** - interactive exploration of structural alignments

**MultAlign Viewer** - integrates sequence and structure space

**Chimera** - extensible 3-D molecular modeling system
Example Study

Structural comparison of glutamine synthetase (GS) and creatine kinase (CK)

- GS: 468 residues, PDB entry 2gls
- CK: 380 residues, PDB entry 1crk
- No significant sequence similarity, both have multimeric forms, proposed similar tertiary structures, and catalyze similar reactions
GS and CK catalysis

Glutamate + ATP → Glutamine

Creatine + Mg ATP → Phosphocreatine

Glutamine + NH₃ → Glutamate

Phosphocreatine + MgADP + H⁺ → Creatine + Pi
Glutamine synthetase and creatine kinase
After MinRMS alignment

Glutamine synthetase

Creatine kinase
AlignPlot GUI

| # Matches | 120 |
| RMSD      | 1.99 |
| Longest Distance | 3.37 |
| -log(probability) | 8.25 |
Resulting structure-based sequence alignment

1crk.pdb  TVHEKRKLFP PSADYPDLRK HNNCMAECLT PAIYAKLDRK LTPNGYSLDQ CIQGTVDNPG HPFIKTVGMV AGDEESYEV
2gls.pdb          ................................. ........................................

1crk.pdb  DAS...... .................................................... ....... .......... .......... ......... ..
2gls.pdb  SAEH VLTMLNEHEV KFVDLRFDT KGR..EIQHT IPALQVNAEF FEEGKMFDGS

1crk.pdb  KI...T . H GCF  .......... ..DERYLS.
2gls.pdb  SIGGWKGINE SDMVLMPDAS TAVIDPFFAD STLIIRCDIL EPGTLQGYDR DP.RSIAKRA .E.DYLIRATG IADT ... V

1crk.pdb  PPACSR ........................................... VV...V E.RRE  VENYVTAL.
2gls.pdb  YFPVRPV....S  AODIRSE.MC

1crk.pdb  AGV...K.G.D.L SGKYSLLTNM SERDQQQQLID DHFLFDKPV.S PLTCAGMAR DWPDAEGI.W HNNDKTFLV.. WINEED...
2gls.pdb  L.VGEO.MGI ........................................... V V ..... E.A HHH. EVATA

1crk.pdb  GT...V..VA TR FN ....MT..K ADEIQIYKYV VHVNAHFEGK T.....T FM .......... P.KPMFGDNG SCMCHS.. L
2gls.pdb  R.K...LISDKPRFPK I.....L.E NIRL

1crk.pdb  V.EVLTCPQNL GT . S.LGAGEV.
2gls.pdb  TQ...V.AIYACL .S PKARRI.EV ..RF ....PD PAAN .PYLC .PAAILMAGL GT ....K

1crk.pdb  L.P ...... ........................................... Q ....FGR ............................
2gls.pdb  AKNGTNLFSG DRYAGLSEQ . ........ ALYYIGGVI KHA.KAINAL ANPTNTSYKR LGPGYEAPVM LAYSARNBSA

1crk.pdb  LG...V..VA TA  \A.AADV.VY ....DI..SN LD.RGMR....EVEL...V .QVIVDGVN7 .LDVCEKKELG KDKLTVPPP
2gls.pdb  ST...IPV . VA .....S PKARRI.EV ..RF ....PD PAAN .PYLC .PAAILMAGL GT ....K

1crk.pdb  IP ........................................... Q ....FGR ............................
2gls.pdb  KIHFGEPM DKNLVDLIPF EAKSETIPQVAG SLEEA..LNA LLDLDREPLKA GCYFTIDEAID AYIALRREED DRVRMTPHPV

1crk.pdb       
2gls.pdb  EFELTTSV
All known protein structures
All known protein sequences
Lists of small ligands
All known protein interactions

Center for Computational Proteomics Research

For as many proteins as possible:
Binding sites for ligands and proteins
Binding affinities
Geometries of complexes

INPUT

OUTPUT

Database
Graphical User Interface
Web server
Cross-links
Collaborations
Service
Education

Structural genomics
Functional annotation
Macromolecular assemblies
Cellular networks
Comparative proteomics
Pharmacogenetics
Drug discovery

INTERFACE

IMPACT
Genome-Wide Mapping of Protein Interactions

Ligand-Protein Docking Pipeline

1. Comparative modeling (Sali)
2. Refine protein models (Jacobson)
3. Identify ligand binding sites on models (Babbitt, Kortemme, Sali)
4. Virtual ligand libraries (Shoichet)
5. Build ligand-protein complexes (Dill, Jacobson, Babbitt)
6. Rescore ligand-protein complexes
7. Central database (Ferrin, Sali)
8. Graphical User Interface (Ferrin)

Protein-Protein Docking Pipeline

9. Build protein-protein complexes (Baker, Sali)
10. Specificity modeling of protein interactions (Kortemme)

Software and Information Technology

- Software backplane (Ferrin)
- Cluster hardware and software environments (IBM, Intel, Ferrin, Sali)
- Global optimization (Dill, Rosen)
- Information navigation and search (Hearst)
- Testing (Sali, Shoichet)
- Testing and Applications (Babbitt, Baker, Dill, Shoichet)

Computational applications

- D.1
- D.2
- D.3
- D.4
- D.5
- D.6
- D.7
- D.8
- D.9
- D.10
- D.11
- D.12
- D.13
- D.14
- D.15
- D.16
- D.17
- D.18-20

Genome-Wide Mapping of Protein Interactions

All known protein structures
All known protein sequences
Comparative modeling (Sali)
Refine protein models (Jacobson)
Identify ligand binding sites on models (Babbitt, Kortemme, Sali)
Annotated protein structure models
Virtual ligand libraries (Shoichet)
Build ligand-protein complexes (Dill, Jacobson, Babbitt)
Rescore ligand-protein complexes
Central database (Ferrin, Sali)
Graphical User Interface (Ferrin)
Summary

We are in the midst of a profound and exciting new era in bioinformatics and computational biology

The data made available by the various genome and structural genomics projects will occupy researchers for decades to come

High performance computing and the internet play a critical role in the navigation, analysis, and dissemination of this data and the resulting scientific knowledge

The tremendous volume of data makes for a critical need for tools and techniques that make information navigation easy

The potential impact on drug development and treatment of human disease is enormous
Acknowledgements

Collaborators & Staff
• Dr. Conrad Huang, Dr. Elaine Meng, Prof. Patricia Babbitt, Prof. Kathy Giacomini, Greg Couch, Eric Pettersen, Al Conde, Tom Goddard, Susan Johns, Doug Stryke, Michiko Kawamoto

NIH National Center for Research Resources
• P41-RR01081

National Institute of General Medical Sciences
• GM61390
Additional information

RBVI:
www.rbvi.ucsf.edu

PMT project:
www.pharmacogenetics.ucsf.edu

Chimera:
www.cgl.ucsf.edu/chimera

CCPR:
www.computationalproteomics.org