Open Access Integrative bionetworks: Why will disease biology and causal disease models enter the pre-competitive space?

Create an open access, integrative bionetwork evolved by contributor scientists working to eliminate human disease

Stephen Friend MD PhD
Sage Bionetworks
Fred Hutchinson Cancer Research Center

IOM National Policy Forum
Extending the Spectrum of Pre-Competitive Oncology Biomedical Research
February 9th, 2010
What drives current behavior?
Extending the Spectrum of Pre-Competitive Oncology Biomedical Research

Sharing Consultants
Trading Scientists
Sharing Revenues
Sharing Analysis Tools
Sharing Samples and Non-Compound Data
Sharing Disease Models
Sharing Clinical Compound Data
Sharing Toxicities
Sharing Compounds
Extending the Spectrum of Pre-Competitive Oncology Biomedical Research

Remember that if it is a company and often also if a University they are obligated to justify the underlying business case.

Can I get there without others cheaply? Will others beat me to it if I do not join? Can I build a first mover advantage? Can I sustain my advantage and sustain my return?
What are forces driving desire to consider precompetitive options?
Thanks to Bill Proud

"Purpose of visit?"
Computational methods for integrating massive molecular and clinical datasets obtained across sizable populations into predictive disease models can recapitulate complex biological systems.

Data should feed and refine a set of models that inform our understanding of disease causality as well as generate new mechanisms, targets, diagnostics and knowledge.
Profiling signatures to ID responder populations
The Merck/Moffitt Strategy: Direct patient selection from database matching molecular signatures to clinical trials

Profiles stored at hospital

“BRCA ness”

Select Target LOF
DNA Damage  PI3K  Ras

Disease recurrence and trial eligibility
• Validation of molecular hypotheses
• Patient selection on available profiling data

Biomarker Driven Branched Subpopulation based Adaptive Trials
HEART
VASCULATURE
KIDNEY
IMMUNE SYSTEM

transcriptional network
protein network
GI TRACT
metabolite network
Non-coding RNA network

BRAIN

ENVIRONMENT

ENVIRONMENT

ENVIRONMENT

ENVIRONMENT

ENVIRONMENT

ENVIRONMENT
The truth is we have little idea on the underlying causes of common human diseases.

We need to more fully embrace the complexity to develop a better understanding.
The “Rosetta Integrative Genomics Experiment”: Generation, assembly, and integration of data to build models that predict clinical outcome

Merck Inc. Co.
5 Year Program
Based at Rosetta
Total Resources
>$150M

- Generate data need to build networks
- Assemble other available data
- Integrate and build models
- Test predictions
- Develop treatments
- Design Predictive Markers
Genetics allows Causality to be established

• DNA Variation in populations represents complex set of perturbations that with the environment drive disease

• Map DNA variation to intermediate phenotypes (RNA expression) and clinical phenotypes (outcomes)

• For shared DNA variation that drives both intermediate traits and clinical traits can use statistical modeling to derive likely order

Network Modeling

- **Probability-based networks**: Bayesian network
  
  Bayes’ theorem
  
  \[ p(B|A) = \frac{p(A|B) \times p(B)}{p(A)} \]

- **Correlation-based networks**: 
  Weighted Co-expression network
  Tissue-tissue network
Decipher Biological Systems via Multiple Networks

Genotype Data

Gene Expression

Clinical Traits

Other Data

Bayesian Network
Zhu et al., 2004

Causality Network
Schadt et al., 2005

Co-Expression Network
Zhang & Horvath, 2005

Integration
Identification of Causal Regulators of Metabolic & CV Disease

Network Causal for Diabetes, Obesity & Cardiovascular Traits

Test all sub-networks to identify those causal for disease
Identify one sub-network that contains many causal regulators:
- 75% of genes supported as causal for atherosclerotic lesions
  - (117 of 157 genes; p=5.6e-103)
- 50% genes supported as causal for obesity
  - (366 of 735; p=3.8e-235)
- 45% genes supported as causal for diabetes
  - (263 of 571; p=1.5e-133)

Angiogenesis-enriched Module
Vascularization-enriched Module
Cell Cycle Module

Variations in DNA elucidate molecular networks that cause disease
### Extensive Publication Validates Scientific Approach

• >60 Publications from Rosetta Genetics Group (~30 scientists) over 5 years including high profile papers in Nature and Nature Genetics

|                  | "Validation of candidate causal genes for obesity that affect..." Nat Genet. (2009)  
|                  | ...... Plus 10 additional papers in Genome Research, PLoS Genetics, PLoS Comp.Biology, etc  

|                 | ...... Plus 5 additional papers in Genome Res., Genomics, Mamm.Genome  

| Bone            | "Integrating genotypic and expression data …for bone traits…” Nat Genet. (2005)  
|                 | “...approach to identify candidate genes regulating BMD…” J Bone Miner Res. (2009)  


---

Metabolic Disease  
CVD  
Bone  
Methods
Unbiased Reconstruction of a Mammalian Transcriptional Network Mediating Pathogen Responses

Gene Expression Signatures Diagnose Influenza and Other Symptomatic Respiratory Viral Infections in Humans

Aimee K. Zaas,1,3,9 Minhua Chen,2,9 Jay Varkey, Timothy Veldman,4 Alfred O. Hero III, Joseph Lucas,3 Yongsheng Huang, Ronald Turner,6 Anthony Gilbert,6 Robert Lambkin-Williams,6 N. Christine Øien,3 Bradly Nicholson,7 Stephen Kingsmore,8 Lawrence Carin,2 Christopher W. Woods,1,3,7 and Geoffrey S. Ginsburg,3,*
The transcriptional network for mesenchymal transformation of brain tumours

Maria Stella Carro, Wei Keat Lim, Mariano Javier Alvarez, Robert J. Bollo, Xudong Zhao, Evan Y. Snyder, Erik P. Sulman, Sandrine L. Anne, Fiona Doetsch, Howard Colman, Anna Lasorella, Ken Aldape, Andrea Califano & Antonio Iavarone

NATURE 463:318, 21 JANUARY 2010
Why Build Sage Bionetworks and a Commons?
Sage: Transitioning from Clinicians as Archivists to a Contributor Community by jointly building evolving Models of Health and Disease

Current Approaches

**Nature Science**

**PLoS**

Integrative genomic approaches
Using causal bionetwork approaches
To build evolving disease models

GWAS

TCGA

Requires: Reliable Data- Quality Controls
Well annotated- engage standards
Curation of platforms
Coordination of Computational Models
Tools for Interactive Work place

(lineverage existing public, private efforts EBI, NCBI etc.)
How to take this forward: Jettison from a single entity

Recognition that the benefits of bionetwork based molecular models of diseases are great but that they require significant resources

Appreciation that it will require decades of evolving representations

Excited to enable patients and their advocacy groups to build models for who responds and does not respond to marketed drugs

Merck saw that by donating data, tools, know-how, it might seed a “commons” allowing a potential long term gain to the whole community provided by evolving models of disease built via a contributor network
Sage Bionetworks

Non-profit to create a “commons” where integrative bionetworks is evolved by contributor scientists to accelerate the elimination of human diseases.

Data, Models, and Tools, with interconnectivity, data standards, governance rules in a public/private interface supported by Foundations, Universities, Govt, Patients
Current Status- Progress Report?
Sage Bionetworks Strategic Priorities

- Integrative genomics and network biology research
- Repository and tools to establish the Commons platform
- Interdisciplinary scientist training to enable widespread participation
Active Partners

Clinicians- Cooperative Trial Groups
- Cambridge UK- Carlos Cordon/ UBC- Sam Aparicio
- CML –Jerry Radich
- NSABP- Soon Paik
- TCGA- Lynda Chin

Government Sources
- NCI Grant as a CCSB (awaiting formal designation)
  LSDF Washington State- (finalist awaiting word in February)
- RO1- Novel Statistical methods – Sage Harvard (submitted)

Foundations
- Cure Huntington’s Disease Initiative

Pharma Biotech CRO Payer Partners (Pfizer)
- Building out specific Disease Models
  Oncology, Diabetes, Cardiovascular Disease
- Quintiles

Information Technology Partners
- Amazon, U Miami, U of Arizona

Software Tool Providers
- GenomeSpace-Broad

Patient Advocacy Groups
- Clearity Foundation-Ovarian Cancer
- Multiple Myeloma Research Foundation
- Live Strong
Migrating Coherent Datasets to Populate the Repository

Merck > 30 datasets
TCGA datasets
NSABP
Rest of the World

In progress
Agreed
Agreed
?

Sage Repository & Commons
Sage Non-Responders Project

Patient Oriented Cohort Study to ID Non-Responders to Approved CA Drugs

Co-Chairs- Rich Schilsky and Stephen Friend

Multiple Myeloma- Ken Anderson/ Kathy Guiste
AML at First Relapse- Fred Applebaum/
Non-Small Cell Lung Cancer- Roy Herbst/
Ovarian Cancer- Beth Karlan/ Laura Shawver

Molecular Profiling: Saywers, Golub, Levine, Polit, Levine

Patient Outreach: Anne Wojcicki 23andMe, Live Strong
Relationship between Sage Bionetworks and Sage Commons
Four Requirements to enable a Platform for Clinicians Scientists and Patients

- Data Repository (Commons)
- Platform Architecture System
- Probabilistic Causal Network Models Of Disease
- Rules And Governance
Building a Commons/ Platform; Role of a Contributor Social Network

Regional Meetings: UK, Boston, San Francisco, Beijing

First Inaugural Sage Congress April 23-24 San Francisco, Hosted by UCSF

Participants and Speakers:
- Modelers: Vamsi Mootha, Tery Ideker, Andrea Califano, Eric Schadt
- Contributor Networks: Jamie Heywood, Jeff Hammerbacher, Bob Cook-Degan
- IT Partners: Google, MSFT, Oracle, Amazon, SONY..
- Government/Payers: NIH, NHLBI, NCI, EBI, MEDCO, FDA..
- Pharma/Biotech: Pfizer, Lilly, Genentech, GSK..

Goals and Outcomes
- Standards and Ontologies for Integration Analysis
- Citations of Probabilistic Causal Network Models
- Ratifying Commons Governance Rules for Data and Model Sharing
Extending the Spectrum of Pre-Competitive Oncology Biomedical Research

Remember that if it is a company and often also if a University they are obligated to justify
The underlying business case

Sharing  Disease Models  Unsharing

Can I get there without others cheaply? UNLIKELY
Will others beat me to it and I do not join?  LIKELY
Can I build a first mover advantage?  UNLIKELY
Can I sustain my advantage and sustain my return?  UNLIKELY

Visible