A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma

STEPHEN H. FRIEND*, RENE BERNARDS*, SNEZNA ROGELJ*, ROBERT A. WEINBERG†, JOYCE M. RAPAPORT‡, DANIEL M. ALBERT§ & THADDEUS P. DRYJA§

*Whitney Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA
†Division of Hematology-Oncology, The Children’s Hospital, Dana-Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA
‡Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
Article

Functional Discovery via a Compendium of Expression Profiles

Timothy R Hughes¹, Matthew J Marton¹, Allan R Jones¹, Christopher J Roberts¹, Roland Stoughton¹, Christopher D Armour¹, Holly A Bennett¹, Ernest Coffey¹, Hongyue Dai¹, Yudong D He¹, Matthew J Kidd¹, Amy M King¹, Michael R Meyer¹, David Slade¹, Pek Y Lum¹, Sergey B Stepaniants¹, Daniel D Shoemaker¹, Daniel Gachotte², Kalpana Chakraburty³, Julian Simon⁴, Martin Bard⁵ and Stephen H Friend¹, II, *
time of transition

symptoms to molecular components

singular syndromes to variable presentations
The way we like to think:

The way it is:
We focus on a world where **biomedical research** is about to fundamentally change. We think it will be often conducted in an **open, collaborative** way where **teams of teams** far beyond the **current guilds of experts** will contribute to making better, faster, relevant discoveries.
1

incentives and rewards

sharing among
pre-defined researchers

Synapse- within defined Communities
Synapse enabling large-scale collaborative science

a tool to improve transparency and reproducibility of data intensive science
Synapse enabling large-scale collaborative science

enabling researchers to contribute to large-scale collaborative science
Two approaches to building common scientific knowledge

Text summary of the completed project
Assembled after the fact

Every code change versioned
Every issue tracked
Every project the starting point for new work
All evolving and accessible in real time
Social Coding
Goal is for Synapse is to function as a github for Biomedical Data

- Data and code versioned
- Analysis history captured in real time
- Work anywhere, and share the results with anyone
- Social/Interactive Science
Disease Models for Neuropsychiatric Disease
NIH AMP Program in Alzheimer’s
All data shared quarterly and teams participating on bi-weekly calls
NIH-led effort launches Big Data portal for Alzheimer’s drug discovery

Innovative collaboration, rapid data-sharing opens research to wider community

A National Institutes of Health-led public-private partnership to transform and accelerate drug development achieved a significant milestone today with the launch of a new Alzheimer’s Big Data portal — including delivery of the first wave of data — for use by the research community. The new data sharing and analysis resource is part of the Accelerating Medicines Partnership (AMP), an unprecedented venture bringing together NIH, the U.S. Food and Drug Administration, industry and academic scientists from a variety of disciplines to translate knowledge faster and more successfully into new therapies.

The opening of the AMP-AD Knowledge Portal and release of the first wave of data will enable sharing and analyses of large and complex biomedical datasets. Researchers believe this approach will ramp up the development of predictive models of Alzheimer’s disease and enable the selection of novel targets that drive the changes in molecular networks leading to the clinical signs and symptoms of the disease.

“We are determined to reduce the cost and time it takes to discover viable therapeutic targets and bring new diagnostics and effective therapies to people with Alzheimer’s. That demands a new way of doing business,” said NIH Director Francis S. Collins, M.D., Ph.D. “The AD initiative of AMP is one way we can revolutionize Alzheimer’s research and drug development by applying the principles of open science to the use and analysis of large and complex human data sets.”

“We are determined to reduce the cost and time it takes to discover viable therapeutic targets and bring new diagnostics and effective therapies to people with Alzheimer’s.”

— Francis S. Collins, M.D., Ph.D.
Director, National Institute of Health
incentives and rewards
sharing beyond pre-defined researchers
Crowdsourcing using “Challenges” empowered by Synapse
Structure of a Challenge

Crowdsourcing → Predictions

Data

Unbiased Evaluation → Acceleration of Research

Collaboration

Measurements

Ground Truth

Engaging New experts
The Sage Bionetworks/DREAM Breast Cancer Prognosis Challenge

**Goal:** use crowdsourcing to forge a computational model that accurately predicts breast cancer survival

**Training data set:** genomic and clinical data from 2000 women diagnosed with breast cancer (Metabric data set)

**Data access and analysis tools:** Synapse

**Compute resources:** each participant provided with a standardized virtual machine donated by Google

**Model scoring:** models submitted to Synapse for scoring on a real-time leaderboard
Incentivizing Continuous Participation

• Monthly leaderboard winners
  – Winner is highlighted within the Challenge community
  – Winner posts a blog on winning model to Synapse

• Communities that link to the Leaderboard
  – Stackoverflow: Q&A site with 1,000,000 users
  – Science Translational Medicine community
Predict cancer-associated mutations from whole-genome sequencing data.

Develop predictive models to infer genes that are essential to cancer cell viability using gene expression and/or gene copy number features.
DREAM 10 Challenge Planning Already Underway:

DREAM ALS Stratification Prize4Life Challenge

- Data set is ready: fully harmonized clinical data set of >9,000 ALS patients
- Challenge objective: algorithms that can assist with stratification of ALS patient populations to help improve individual patient-level prognosis and to improve the success rate of clinical trials

Project Datasphere Prostate Cancer DREAM Challenge

Challenge Focus: predict survival for prostate cancer patients based on patients’ clinical variables using Comparative Arm data from Industry Trials
Epic Sciences Inks Collaboration with Prostate Cancer Clinical Trials Consortium
3

searching for targets

looking beyond those that are sick
The Resilience Project: A Search for Unexpected Heroes
power to diagnose
yet lack power to fully treat
Unexpected Hero candidate disease annotation

Sample collection

Initial screening

Search Your Genome Tool

Panel applied

Unexpected Hero core allele panel

High-throughput filtering

Feedback: Allele removal addition

Low-throughput filtering

Health record review

Sanger confirmation

~50 selected candidates

9 Unexpected Hero candidates

>10,000 initial candidates

596,610 study subjects
Table 4: Unexpected Hero Candidates. All strong candidates identified so far harbor mutations found in the UHP core allele panel, with two compound heterozygotes.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Gene</th>
<th>cDNA mutation</th>
<th>Protein mutation</th>
<th>Panel source</th>
<th># of Candidates</th>
<th>Zygosity</th>
<th>Studies</th>
<th>References (PMID)</th>
<th>Validation status *</th>
<th>Further evaluation status</th>
<th>Sample Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystic fibrosis</td>
<td>CFTR</td>
<td>c.1558G&gt;T</td>
<td>p.V520F</td>
<td>core allele panel</td>
<td>2</td>
<td>bocc</td>
<td>23andMe</td>
<td>23974870</td>
<td>C1, C2, C3, G1, G2</td>
<td>Pending for Sanger confirmation</td>
<td>2 adults, one declared no manifestation</td>
</tr>
<tr>
<td>Smith-Lemli-Opitz Syndrome</td>
<td>DHCR7</td>
<td>c.944-1G&gt;C</td>
<td></td>
<td>core allele panel</td>
<td>2</td>
<td>bocc</td>
<td>UK1K</td>
<td>10677299; 23042628; 11001807</td>
<td>C1, C2, G1, G2</td>
<td>Pending for individual medical history</td>
<td>Not obtained</td>
</tr>
<tr>
<td>Familial Dysautonomia</td>
<td>IKBAP</td>
<td>c.2204+6T&gt;C</td>
<td>p.R74C</td>
<td>core allele panel</td>
<td>1</td>
<td>bocc</td>
<td>23andMe</td>
<td>18197058; 18197057</td>
<td>C1, C2, G1, G2</td>
<td>Pending for Sanger confirmation</td>
<td>No disease reported by individual</td>
</tr>
<tr>
<td>MPS IIIA</td>
<td>SGSH</td>
<td>c.220C&gt;T</td>
<td>p.R74C</td>
<td>core allele panel</td>
<td>1</td>
<td>bocc</td>
<td>23andMe</td>
<td>9285796; 1179348</td>
<td>C1, C2, G1, G2</td>
<td>Pending for Sanger confirmation</td>
<td>No disease reported by individual</td>
</tr>
<tr>
<td>Epidermolysis Bullosa Simplex</td>
<td>KRT14</td>
<td>c.373C&gt;T</td>
<td>p.R125C</td>
<td>core allele panel</td>
<td>1</td>
<td>het</td>
<td>BGI</td>
<td>1717157; 16786515</td>
<td>C1, C2, G1, G2</td>
<td>Pending for Sanger confirmation</td>
<td>Not obtained</td>
</tr>
<tr>
<td>Pfeiffer syndrome</td>
<td>FGFR1</td>
<td>c.755C&gt;T</td>
<td>P252R</td>
<td>core allele panel</td>
<td>1</td>
<td>het</td>
<td>SWE-SCZ (MSSM)</td>
<td>7874169; 16418739</td>
<td>C1, C2, G3, G1, G2</td>
<td>Pending for individual medical history</td>
<td>No abnormal morphology</td>
</tr>
<tr>
<td>APECED</td>
<td>AIRE</td>
<td>c.769C&gt;T</td>
<td>p.R257X</td>
<td>core allele panel</td>
<td>1</td>
<td>bocc</td>
<td>23andMe</td>
<td>20718774; 20407228</td>
<td>C1, C2, G1, G2</td>
<td>Pending for Sanger confirmation</td>
<td>No disease reported by individual</td>
</tr>
</tbody>
</table>

Total 9

* see Supplementary Table 7 for validation status code definitions
4

engaging the public as partners

expanding the dimensions of diseases
Incorporating open data, patient wisdom and public involvement into biomedical research
A Research Study with Feedback Loops

Anecdotes into Signals

Partners

Study Participant

Discussion

Answers

Individual Data Trackers

Questions

Results

Bridge

Population Data

Study Researcher

Partners

Anecdotes into Signals
"I am aiming to screen for Parkinson's disease using voice recordings alone."

- Max Litt, Parkinson's Voice Initiative

Whether you are healthy or living with Parkinson's, help provide the voice information needed to build a system to screen for and monitor the symptoms of this debilitating disease. All you need to do: make a low-cost, anonymous, three-minute phone call.
Tanner
Kieburtz
Kruger
Bloem
Current measures of disease are often relatively insensitive, episodic, provider-centered, clinic-based …

Parkinson disease outcome measures – MDS-UPDRS

3.10 GAIT

Instructions to examiner: Testing gait is best performed by having the patient walking away from and towards the examiner so that both right and left sides of the body can be easily observed simultaneously. The patient should walk at least 10 meters (30 feet), then turn around and return to the examiner. This item measures multiple behaviors: stride amplitude, stride speed, height of foot lift, heel strike during walking, turning, and arm swing, but not freezing. Assess also for “freezing of gait” (next item 3.11) while patient is walking. Observe posture for item 3.13

0: Normal: No problems.

1: Slight: Independent walking with minor gait impairment.

2: Mild: Independent walking but with substantial gait impairment.

3: Moderate: Requires an assistance device for safe walking (walking stick, walker) but not a person.

4: Severe: Cannot walk at all or only with another person’s assistance.

Source: Movement Disorder Society – Unified Parkinson’s Disease Rating Scale
Smartphones and wearable devices can greatly improve current measures of Parkinson disease outcome measures.

<table>
<thead>
<tr>
<th>Current measures</th>
<th>Smartphone measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insensitive</td>
<td>Sensitive</td>
</tr>
<tr>
<td>Subjective</td>
<td>Objective</td>
</tr>
<tr>
<td>Episodic</td>
<td>Continuous</td>
</tr>
<tr>
<td>Provider-centered</td>
<td>Individual-centered</td>
</tr>
<tr>
<td>In clinic</td>
<td>Remote</td>
</tr>
<tr>
<td>Uni-dimensional</td>
<td>Multi-dimensional</td>
</tr>
<tr>
<td>Limited or absent feedback</td>
<td>Real-time feedback</td>
</tr>
</tbody>
</table>
TIMED TAPPING TASK
Tests for Bradykinesia

For this task, please lay your phone on a flat surface to produce the most accurate results.
Tests for Bradykinesia

Once you tap Get Started, you will have 5 seconds before the first interval set appears.
Tests for Bradykinesia

Next, use 2 fingers on the same hand to alternately tap the buttons for 20 seconds. Time your taps to be as consistent as possible.
Tests for Bradykinesia

After the intervals are finished, your results will be visible on the next screen.
Completing more activities increases the effectiveness of the study.

Today’s Activities

My Bradykinesia
13 Taps / second – Excellent

Performance Comparison

Learn More
Lorem ipsum dolor sit amet, etos et ya consectetur adip isicing elit, sed dol
FINDING

WINdows

FOR

INTERVENTION

Pre-enroll
Clinical trials

ANECdotes into signals

Validated anchors with serious friction

Structured Tests
Moderate validity
Friction

Poorly validated
Frictionless
Passive measures

Gait
Tremors
Mood
Cognition
Fatigue
Speech
Sleep

STRESS  SLEEP  MEDICATION  EXERCISE  FOODS  MEDITATION
OPPORTUNITIES FOR OPTIMIZING CLINICAL TRIALS

Parkinson’s Disease

Alzheimer’s Disease

MS

Diabetes

Rheumatoid Arthritis

(Breast Cancer Survivors)

Etc....

(mood social behavior, cognition, movement)
Apple Launches "ResearchKit" for Medical Studies
Smartphones set for large-scale health studies

Tech giant Apple introduces mobile platform for biomedical research.

Helen Shen

10 March 2015

Rights & Permissions
• http://www.apple.com/live/
combining the pieces

altering roles and responsibilities

a look at the process end to end of health care and drug discovery
Reagents & Resources
Created, discovered and shared by the SGC without restriction on use: innovation for everyone.

Director

Aled Edwards, PhD
Director, SGC

Dr. Aled Edwards is founding and current CEO of the Structural Genomics Consortium (SGC), Professor at the University of Toronto and Visiting Professor at the University of Oxford. Trained as a protein biochemist at McGill University (with Peter Braun), and at Stanford University (with Roger Kaback), his research interests include structural biology, host-virus interactions, functional proteomics and drug discovery.

In 1997, he and Dr. Cheryl Arosmiith, contemporaneously with others around the world, launched a pilot project in structural genomics. The publication of their pilot project represented the first report of a large-scale effort in structural biology. Since 2000, Al and his colleagues have contributed to over 2,000 unique protein structures into the Protein Data Bank.
The Precompetitive Space: Time to Move the Yardsticks

Thea Norman, Aled Edwards, Chas Bountra, Stephen Friend

Industry, government, patient advocacy groups, public funders, and academic thought leaders met in Toronto, Canada, to set into motion an initiative that addresses some of the scientific and organizational challenges of modern therapeutics discovery. What emerged from the meeting was a public-private partnership that seeks to establish proof of clinical mechanism (POCM) for selected “pioneer” disease targets using lead compounds—all accomplished in the precompetitive space. The group will reconvene in April 2011 to create a business plan that specifies the generation of two positive POCM results per year.
Sustainable pipeline for better care and treatments from and back to engaged public/patients driven by nurturing open collaborations exploring multi-dimensional clinical space
2015 Paris Assembly: Connecting Open Research, Open Education and Open Social Impact

For more information email: assembly2015@sagebase.org

April 16-18, 2015
Duclaux Lecture Theater, Institut Pasteur, Paris

Honorary Co-Chairs
Christian Brechot, President, Institut Pasteur and Director, Drugs for Neglected Diseases Initiative
Leland Hartwell, Arizona State University and Professor Emeritus, Fred Hutchinson Cancer Research Center

Contributing Entrepreneur
William Drayton, Founder and Chair, ASHOKA, Innovators for the Public

Organizers
Stephen Friend, Sage Bionetworks
Francois Taddei, Centre de Recherche Interdisciplinaire
John Wilbanks, Sage Bionetworks
Geoff Mulgan, NESTA

Planning Committee
Samir Brahmchari, Council of Scientific and Industrial Research, India
Charles Hugh-Jones, Sanofi North America
Accelerating open research through provenance and transparency
Changing the roles of whom we learn what from when

RDA
MARCH 10, 2015