

From: [Richard Chen](#)
To: [John West](#); [Carol Tillis](#); [Christian Haudenschild](#); [Eliane Sousa](#)
Subject: Final Board Slides
Date: Tuesday, July 17, 2012 5:05:11 PM
Attachments: [board.07.18.12.v6.roc.ppt](#)

Includes change suggested by John. Ready for print.

-Rich

Research and Development

Personalis, Inc.

07.18.2012

Richard Chen, CSO

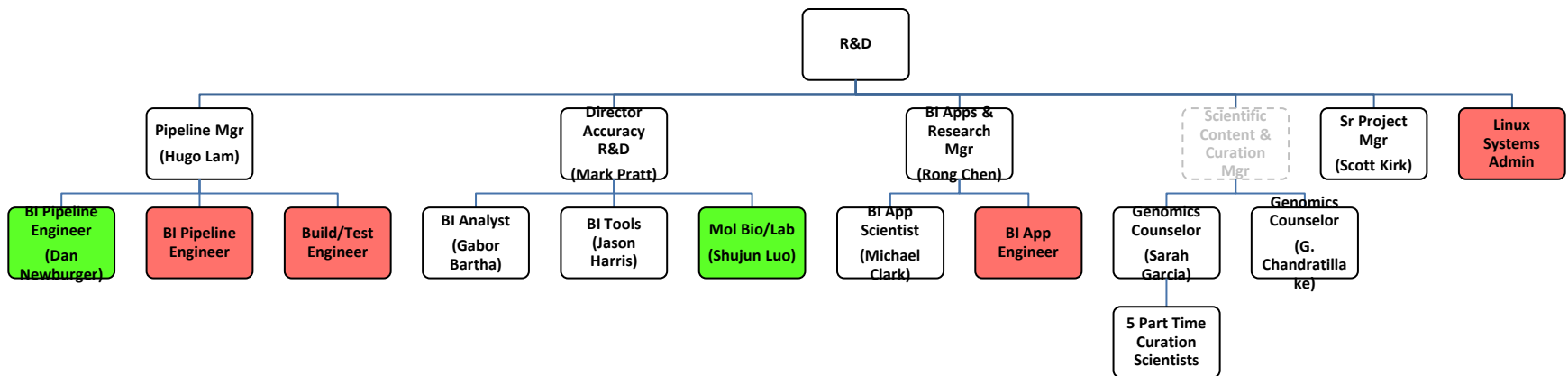
R&D Overview

- Hires
- Product Development
 - Genome Pipeline Service
 - Disease Variant Discovery Service
 - Genome Interpretation Service
- Content Development
- Research Collaborations
- Computing Infrastructure



R&D Hiring Update

- Slower hiring ramp than plan, looking for great hires
- New R&D Hires
 - Daniel Newburger (Pipeline R&D, started mid-June)
 - Shujun Luo (Accuracy R&D, started mid-June)
- Near Term Open Positions
 - Linux systems admin – hired new recruiter, looking for consulting help
 - Test/Build Engineer
 - Applications Engineer
 - BI Pipeline Engineer



-  Original Team
-  New hires (started)
-  New hires (not yet started)
-  Near Term Open Positions

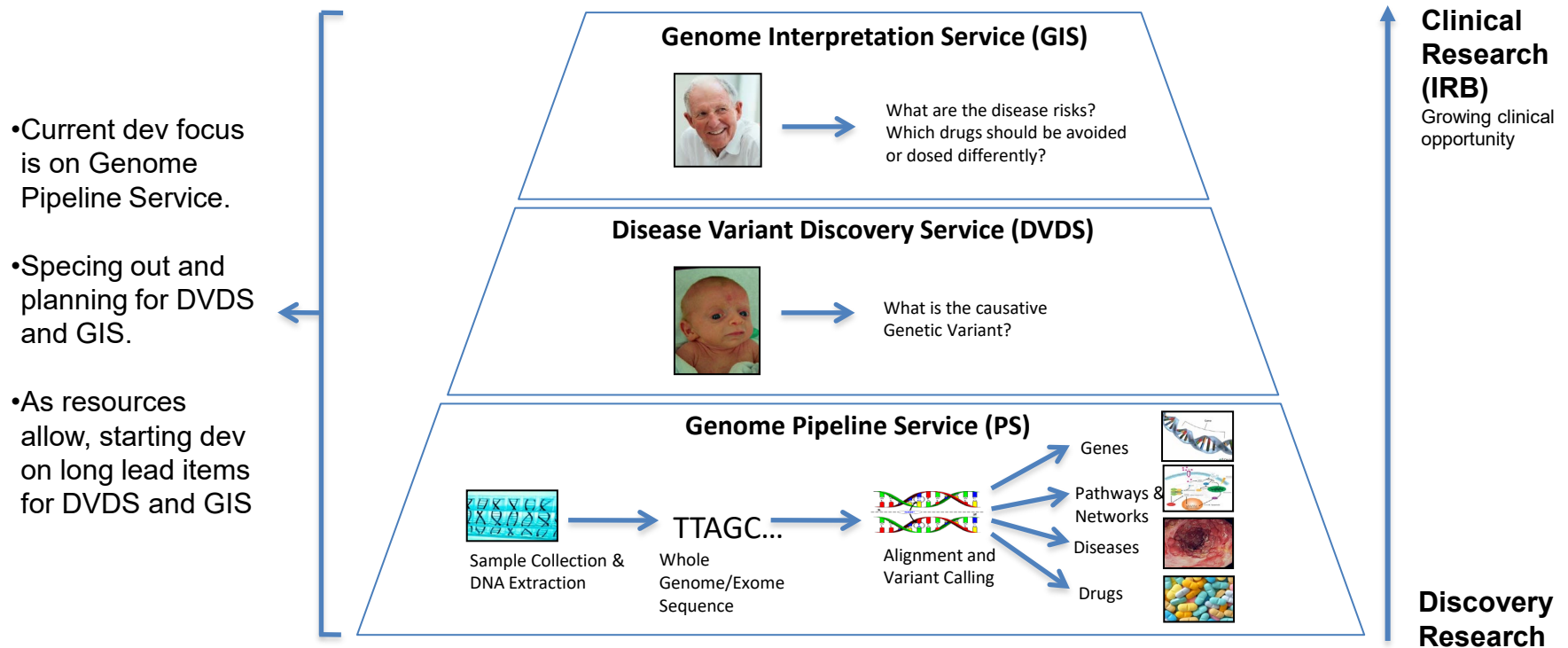
Personalis EX2070

Summary of Resource Allocation

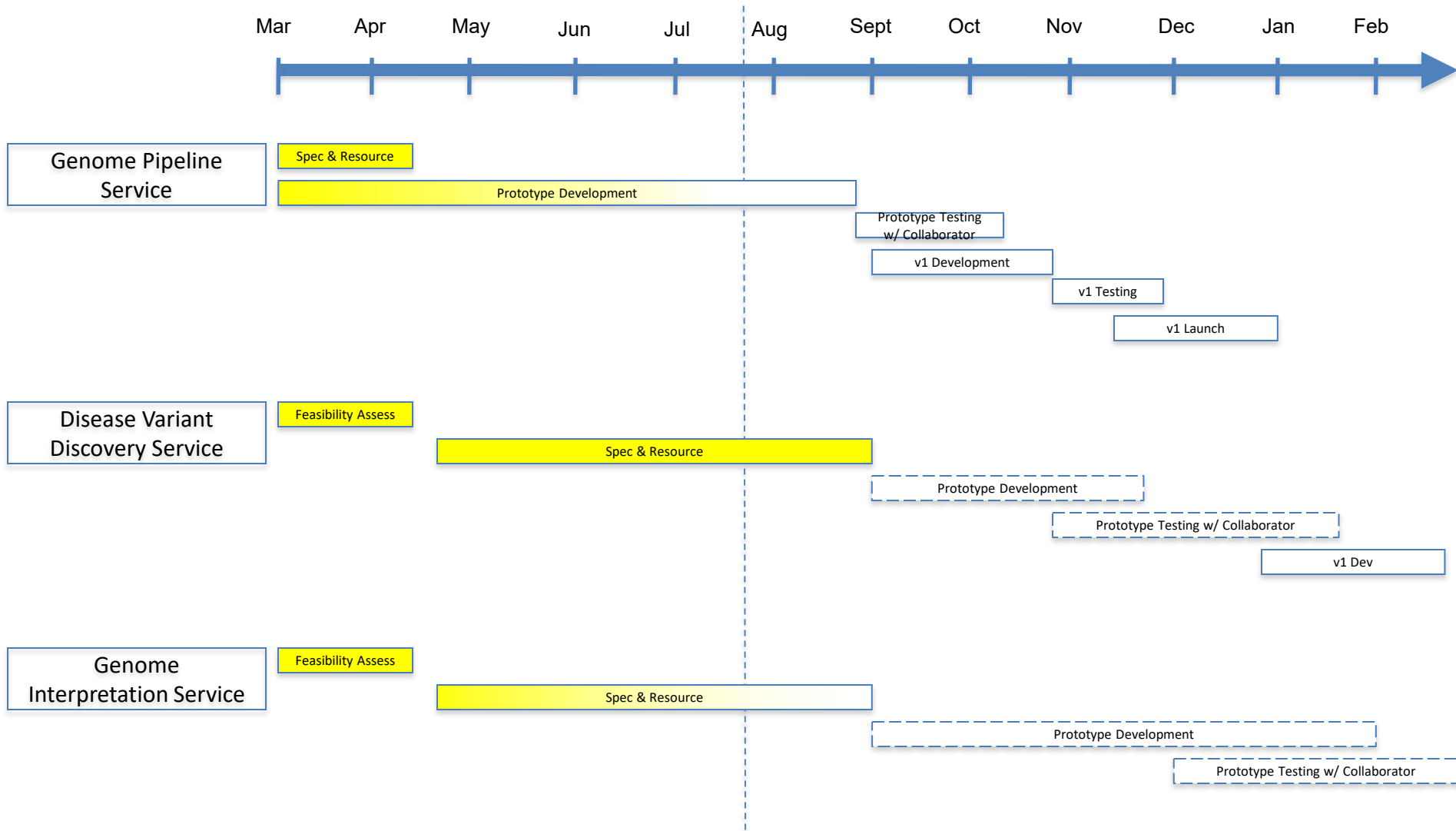
	Accuracy R&D Mark Pratt Gabor Barth Jason Harris Shujun Luo	Pipeline R&D Hugo Lam Daniel Newburger	Apps R&D Rong Chen Michael Clark	Content R&D Sarah Garcia Gemma Chandratillake PT Curators
Pipeline Service				
Disease Variant Discovery Service				
Genome Interpretation Service				
Research Collaborations				
Ethics Board				

Personalis EX2070

R&D Roadmap



Summary of Estimated Timelines



Personalis EX2070

Update: Genome Pipeline Service

- R&D team focus on development of the prototype Pipeline Service
- Successful move to new location with minimal impact to R&D
- Key development milestones
 - HugeSeq pipeline component updates
 - Ethnicity specific genomes used for alignment
 - Variant annotation engine completed
 - Case/control analysis methodology for key use cases developed
 - First prototype output reports being generated
- Remain on track for testing prototype with collaborators in Aug 2012
- Risks/Challenges
 - Hiring – working aggressively with recruiter for scientific computing and BI hires, further delay in hires could affect schedule
 - Computing infrastructure – rackspace constraints relieved with move to new building, looking to install scalable storage solutions for R&D, additional computing capacity installed
 - Licensing of 3rd party tools and content
 - Identifying friendly research collaborators to work with in August

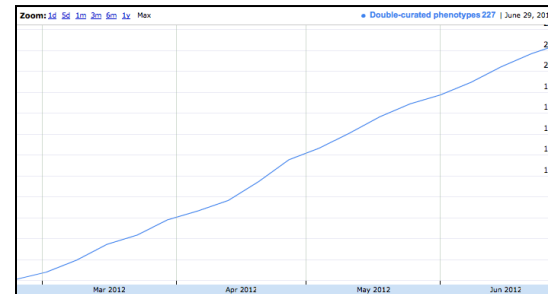
Update: DVDS and GIS

- Success of Pipeline and Accuracy efforts are critical for these products
- Resources focused on pipeline development, but working concurrently with SAB/Founder working groups on defining and spec'ing components of DVDS and GIS
 - Key software modules for DVDS already being developed through SAB consulting arrangement
 - GIS Complex disease report, Mendelian disease report, and PGx specs being developed by working groups

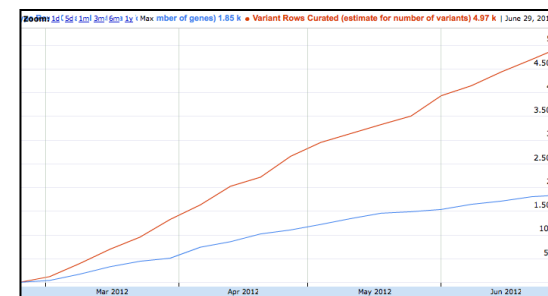
Working Group	Founder/SAB Lead	Personalis Lead
Pipeline	Mike, Carlos	Hugo, Scott
Accuracy	Mike, Carlos	John, Mark
GIS: Complex Dz	Atul	Rong/Sarah/Gemma
GIS: Mendelian Dz	Euan/Rick	Sarah/Gemma
GIS: Pharmacogenomics	Russ	Rich
DVDS	Euan/Rick	Michael

MendelDB Dashboard

- High quality curated DB of Mendelian phenotypes and their variants
 - Managed by Sarah Garcia and Gemma Chandratillake, both Genetics Ph.D. & Genetic Counselors
- Critical content to support Disease Variant Discovery Service and Genome Interpretation Service
- On track to finish all (~560) GeneReviews by Aug 2012, corresponding to >200 common Mendelian disease phenotypes
- Will then continue phenotype curation and variant curation from other sources
- Flexible workforce for other curation tasks in the future



phenotypes



genes (blue)
variants (red)



productivity
(rows/hr)



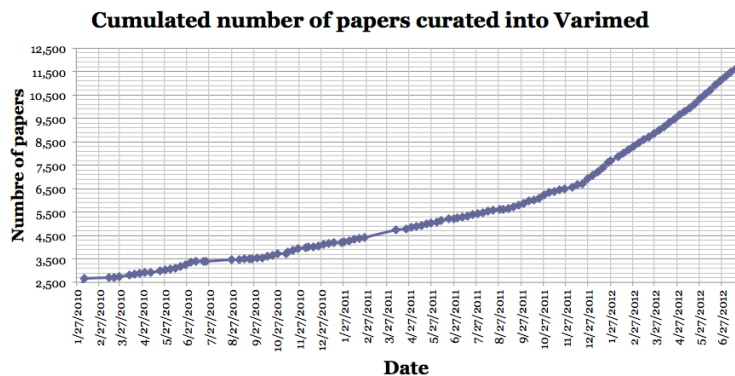
Raw error rate
(# errors/row)
(prior to QC)

Personalis EX2070

VarimedDB Dashboard

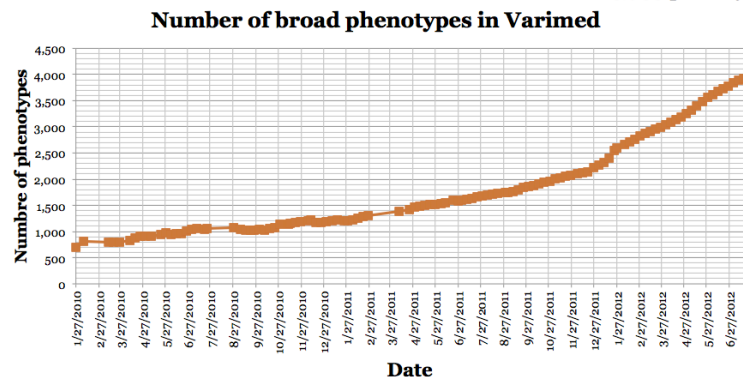
Number of papers

11,628 papers



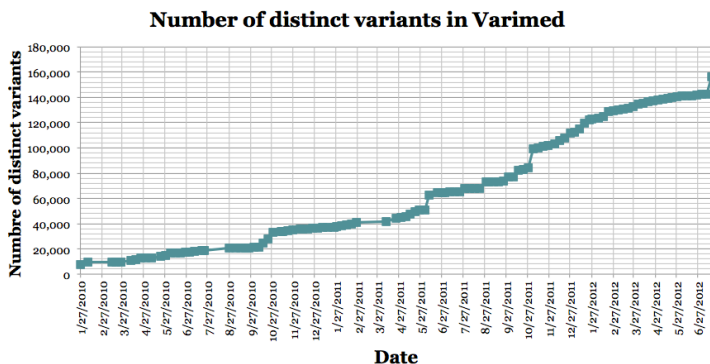
Number of broad phenotypes

3,933 phenotypes



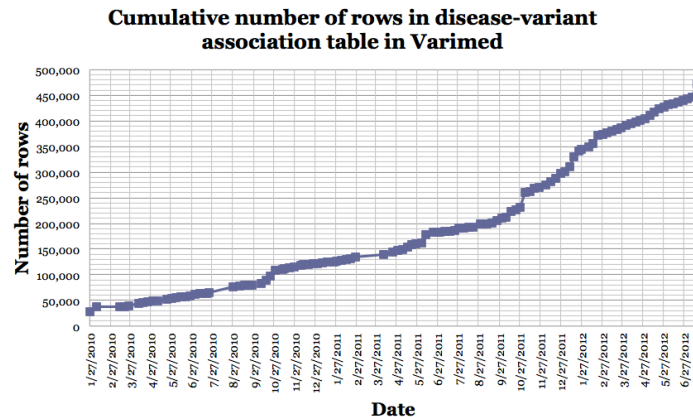
Number of distinct variants

156,038 variants



Number of variant-phenotype relationships (rows)

471,819 relationships



Personalis EX2070

Scientific Computing Infrastructure

- Increased our genome processing capability to 240 cores
- Looking to add ~120TB usable storage for R&D
 - Fielded NAS options from NetApp, Hitachi, Isilon, Dell
 - Most competitive offers from Hitachi and NetApp
- Met with Amazon, Clikr regarding cloud options (referred from MDV)
- Increasing speed of internal network to 10G
- Server backup planning initiated
- Continuing search for Linux/Unix sys admin

Overall R&D Risks and Mitigation

Risk	Mitigation
• R&D hiring delay of BI engineers and Unix sys admin. May impact schedule	• New IT recruiters now involved for Unix sys admin hire • Redoubled efforts on BI hires (new process, conferences, networking)
• Long lead time content curation and content quality items for GIS and DVDS	• Started dev on those items with long lead times to develop concurrently with Pipeline (MendelDB, etc)
• Licensing of applications and content for analysis	• Systematically reviewing and applying for licenses for key content and algorithms
• Ethical use of Genomes for R&D	• Ethics board approval for all Genomes used for R&D • Part of orientation training
• Product Differentiation	• Continuing to push differentiation in three key areas: Accuracy/Quality, Content, Features
• Identifying high quality collaborators for testing our services. Potentially long lead time.	• Begin to identify collaborators for each of the services to get IRB/ethical legal framework in place

Next Steps & Priorities

- Continuing hiring for open positions
- Continue focus on Pipeline Service development with Aug 2012 prototype target
- Identify friendly collaborators for testing Pipeline Service Prototype
- Continue to define specs and plan around Disease Variant Discovery Service and Genome Interpretation Service
- Expand computing storage capabilities
- Continue to execute on content priorities, develop longer term content strategy
- Continue to execute on Accuracy objectives and integration with laboratory capabilities