

From: [Richard Chen](#)
To: [John West](#); [Carol Tillis](#)
Subject: updated board slides
Date: Monday, June 11, 2012 4:03:47 PM
Attachments: [board.06.05.12.v3.roc.ppt](#)

Latest version

Research and Development

Personalis, Inc.

06.05.2012

Richard Chen, CSO

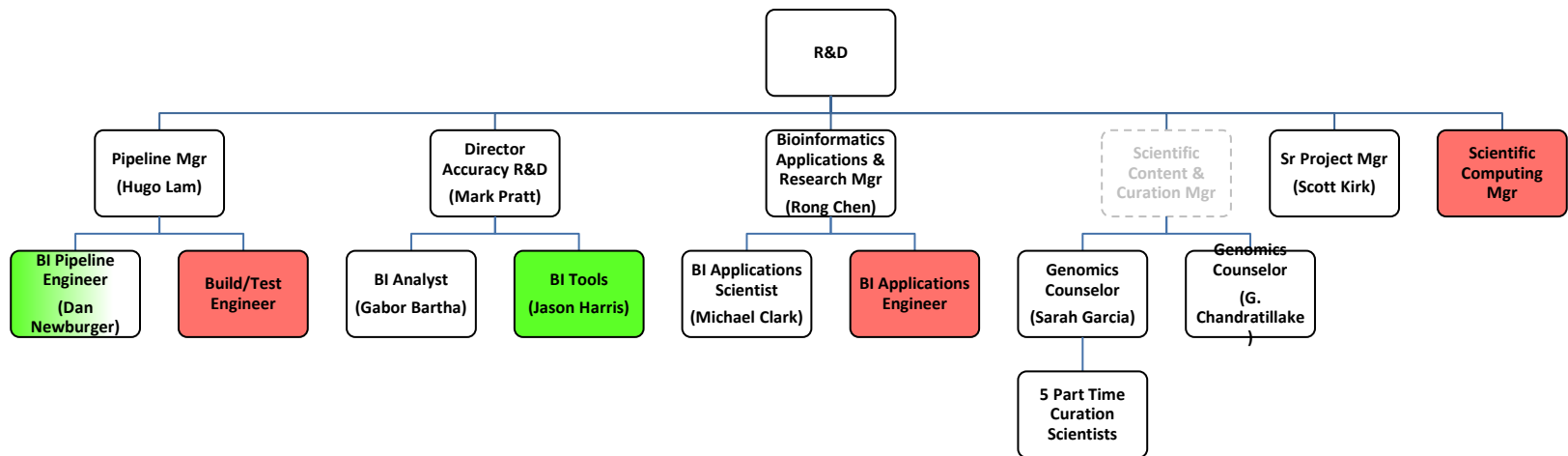
R&D Overview

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



R&D Hiring Update

- New R&D Hires
 - Jason Harris (Accuracy R&D, started late May)
 - Daniel Newburger (Pipeline R&D, starting mid-June)
- Near Term Open Positions
 - Scientific Computing Mgr – increased computing needs, size of R&D team making this more critical
 - Test/Build Engineer
 - Applications Engineer



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

Product Dev: Summary of Resource Allocation

	Accuracy R&D Mark Pratt Gabor Barth Jason Harris	Pipeline R&D Hugo Lam Daniel Newburger (Start 7/15)	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				

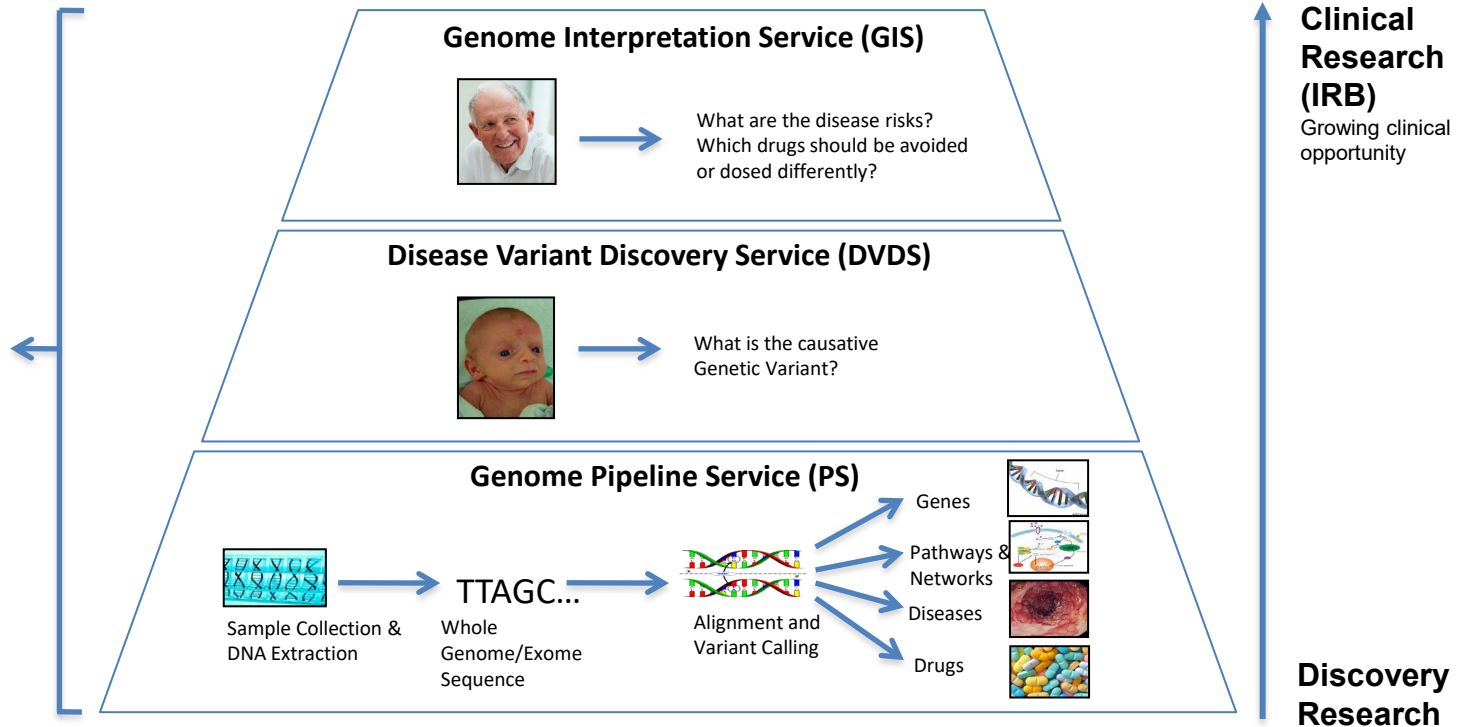
Founder/SAB Working Groups

- Formed small dev teams engaging with a specific founder/SAB member on an key part of our product roadmap
- Small group weekly meetings (2-4 hrs)
- All groups convene every 4-5 weeks for review and updates

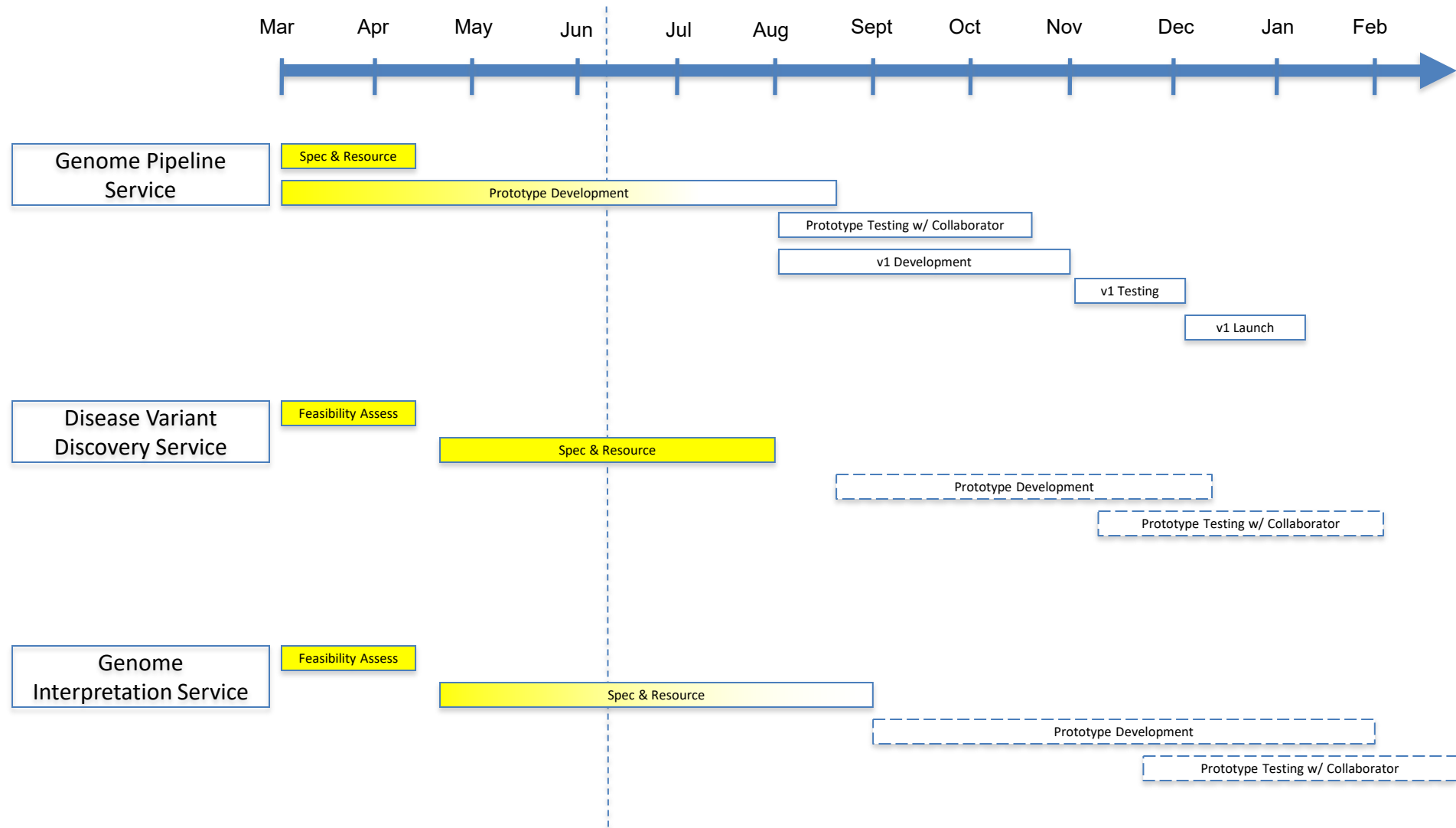
Working Group	Founder/SAB Lead	Personalis Lead
Pipeline	Mike, Carlos	Hugo, Scott
Accuracy	Mike, Carlos	John, Mark
GIS: Complex Dz	Atul	Rong/Sarah/Gemma
GIS: Mendelian Dz	Euan/Rick	Sarah/Gemma
GIS: Pharmacogenomics	Russ	Rich
DVDS	Euan/Rick	Michael

R&D Update

- Moving forward on all three services.
- Current dev focus is on Genome Pipeline Service.
- Specing out and planning for DVDS and GIS.
- As resources allow, starting dev on long lead items for DVDS and GIS



Product Dev: Summary of Estimated Timelines



Personalis EX2062

Update: Genome Pipeline Service

- Team focused on development of the prototype Pipeline Service
- On track for testing prototype with “friendly” collaborators in Aug 2012
- Agile development/Scrum (2-3 week cycles) Dev Process
- Development environment in place:
 - JIRA (for project management, collaboration)
 - GIT (version control system)
 - Confluence (JIRA plug-in for internal WIKI)
- Key Risks/Challenges
 - Product differentiation through accuracy and content – actively managing and iterating on features/product design to maximize differentiation/value
 - Hiring – working aggressively with recruiter for scientific computing and BI hires, delay in hires could affect schedule
 - Computing infrastructure - currently constrained by rackspace at Regus, with move, now working to address computing CPUs, storage, network speed needs
 - Move to new office –move plan minimizes disruption, should impact the schedule < 3-5 days
 - Licensing of 3rd party tools and content
 - Identifying friendly research collaborators to work with in August – need time to put IRB and collaboration agreements in place

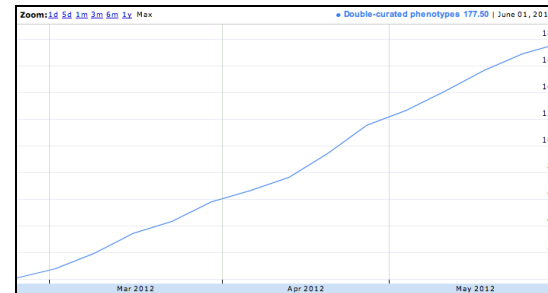
Personalis EX2062

Update: DVDS and GIS

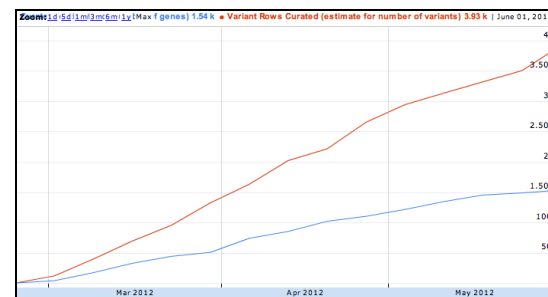
- Success of Pipeline and Accuracy efforts are critical for these products
- Resources focused on pipeline development, but working concurrently on defining and spec'ing the DVDS and GIS
- Completed first draft of functional spec and sample report for DVDS, now in project planning and resourcing phase

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)

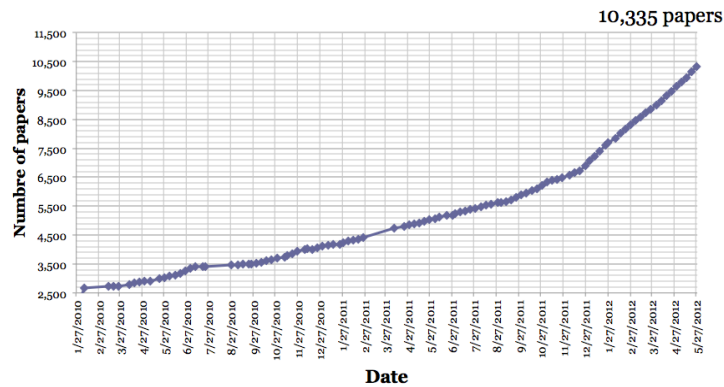


Raw error rate
(# errors/row)
(prior to QC)

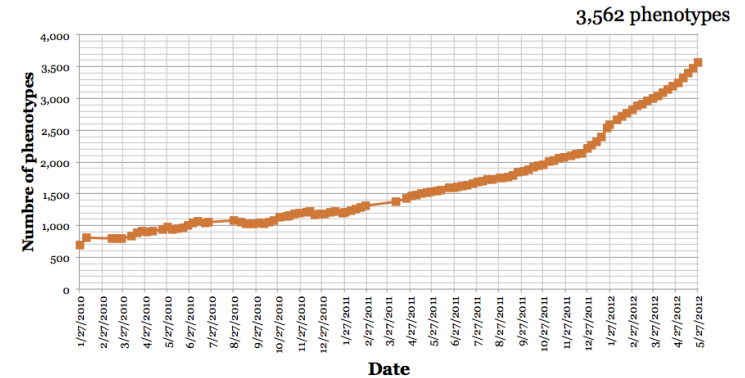
Personalis EX2062

VarimedDB Dashboard

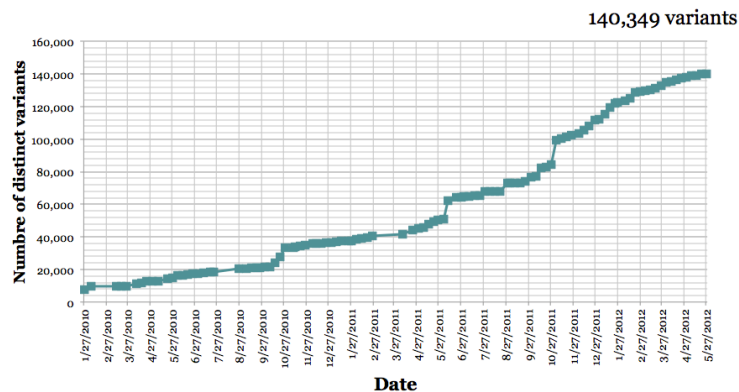
Number of papers



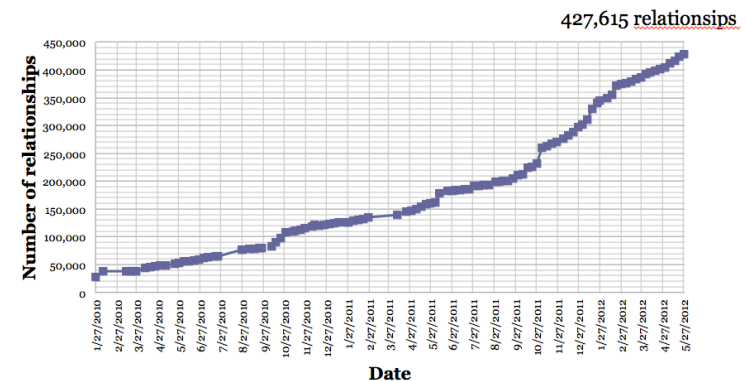
Number of phenotypes



Number of distinct variants



Number of variant-phenotype relationships



Personalis EX2062

Supercentenarian Collaboration

- Completed IRB and Stanford Collaboration Agreement with Dr. Stuart Kim and Stanford
- Moving forward with research collaboration
- Goal: To uncover genetic factors that permit extreme longevity
 - Do supercentenarians have fewer overall risk alleles for common diseases that help them avoid disease?
 - Do supercentenarians have protective alleles that allow them to avoid aging/disease?
 - Do supercentenarians have CNVs that provide protection against or promote aging/disease?
- Details:
 - ≥ 110 yrs old, 71 such people in the world, most female
 - ~ 17 supercentenarians (≥ 110 years old) have been sequenced (CG)
 - Analyzing variants using our content/algorithms
- Benefits:
 - Potentially high profile, impact publication
 - Will showcase the power of our content from Varimed
 - Research experience with genomes from extremely healthy, long lived people
 - Test case for our Pipeline product

Scientific Computing Infrastructure

- Increased needs from:
 - Accuracy and R&D effort
 - processing and reprocessing of numerous genomes has increased computing requirements substantially
- Additional New hardware geared at:
 - Increasing our genome processing capability
 - Dell R910 servers x 4 (for total of 6)
 - Dell R710 servers x 5
 - Increasing our overall storage to enable development work, download of numerous genomes for development and testing work requires more storage
 - NAS (100 to 150TB of storage needed)
 - Increasing speed of network
 - 10G switches
- Other items
 - Server backup planning
 - Move to new location

Overall R&D Risks and Mitigation

Risk	Mitigation
Hiring great people takes time. If not hired on schedule, may impact schedule	- Recruiters now involved, behind on scientific computing - Web site careers posting
Numerous development tasks are still research in nature making timeline more difficult to estimate	Continue rapid prototyping to resolve, get visibility on some key technical risks to plan appropriately, especially with Accuracy effort, GIS product and report generation
Long lead time content curation and content quality items for GIS and DVDS	Started on those items with long lead times to develop concurrently with Pipeline (MendelDB, etc)
Product dev process	Basic Project Mgmt infrastructure deployed including JIRA, GIT, JIRA, etc to enhance efficiency
Licensing of applications and content for analysis	Systematically reviewing and applying for licenses.
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
Scalability of products and services	Prioritize resources towards demonstrating value with our services, then scaling once that is demonstrated
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
- Begin to identify friendly collaborators for testing Pipeline Service Prototype
- Further refine specs and project plan around Disease Variant Discovery Service and Genome Interpretation Service
- Continue to execute on content priorities, develop longer term content strategy
- Continue to execute on Accuracy objectives and integration with laboratory capabilities