ALL IN ONE DAY PRECISION MEDICINE

Ketan Paranjape
General Manager Life Sciences

www.intel.com/healthcare
http://github.com/intel-hls
@IntelHealth, @portlandketan

THE STEPS TOWARDS PRECISION HEALTH ANALYTICS

1 ***Patient visit***

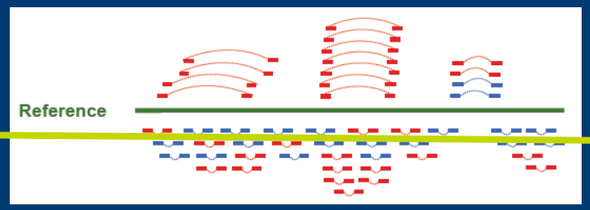
2 ***Genes causing disease and key pathways identified***

3 ***Gene-targeted drugs identified, treatment begins in earnest***

PRIMARY ANALYSIS

Individual

SECONDARY ANALYSIS, DNA/RNA PIPELINE + MORE



HC
HC
HC
HC
HC
HC
HC
HC
HC
HC

Multiple sample compute starts here

Joint genotyping

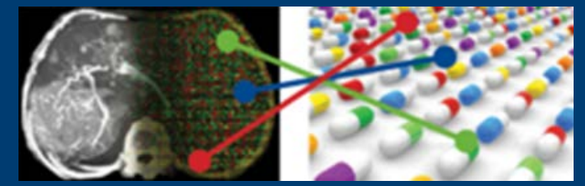
Variant store

Pop/dis study

Multiple individuals

PRECISION MEDICINE

Predicted actionable variants, data-driven association



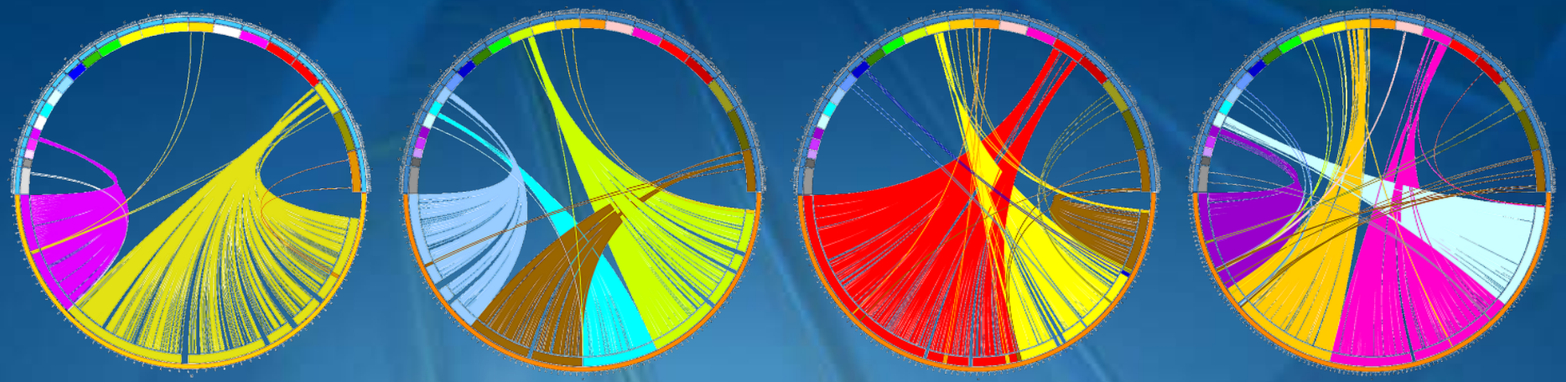
Clinically actionable variants

Clinical trial groups →
Data curation →

Knowledge database

1-4 DAYS | **WEEKS** | **MONTHS**

We Must Go Faster! And not via Phones, Faxes, & Fed-exing.



ALL IN ONE DAY BY 2020

MAIN COMPUTING CHALLENGES TO ACHIEVE 'ALL IN ONE DAY'

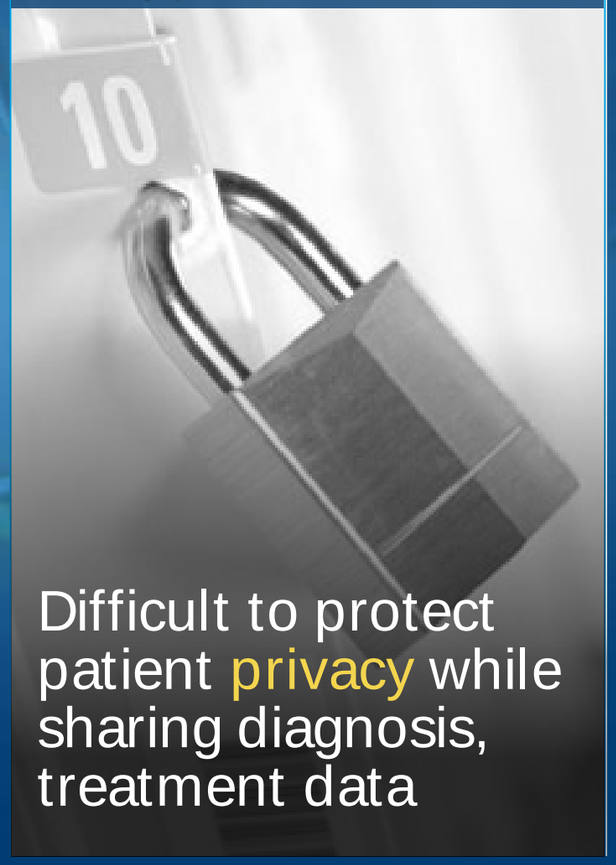
Size
Data management

Raw sequencing data sets are large, impractical to share

Speed
Workflow efficiency

Critical diagnosis, treatment pipelines can take weeks

Secure sharing
Privacy protection



Difficult to protect patient privacy while sharing diagnosis, treatment data

Scalability
Increasing requirements

Genomic sequencing outpacing capabilities of current architectures


4% of data currently in research
96% of data stuck in institutions
• Too large to move
• Protected
(privacy laws, business model IP)

To achieve better research and clinical outcomes, we need

Collaborative analytics

across institutions

- Sequencing data
- Treatments
- Outcomes

DESIGN CHALLENGES WE'VE BEEN FOCUSING ON

Get new insights in a secure, shared way without giving up control of your data

Move the **analytics**, not the data

Do it in an **open, secure** trust model

Do it at **scale**, for any researcher

Bring your own **tools** and **infrastructure**

Collaborative Cancer Cloud

Welcome DFCI, OICR

- Jointly run genomic pipelines
- Drive new analytic paradigms against data sets using integrated platform
- Advance big data analytics in a federated, secure, and scalable environment

CCC is an open platform, enabling either on-premise or public clouds along with best practice precision medicine analytics

Genomics tools

intel BROAD INSTITUTE

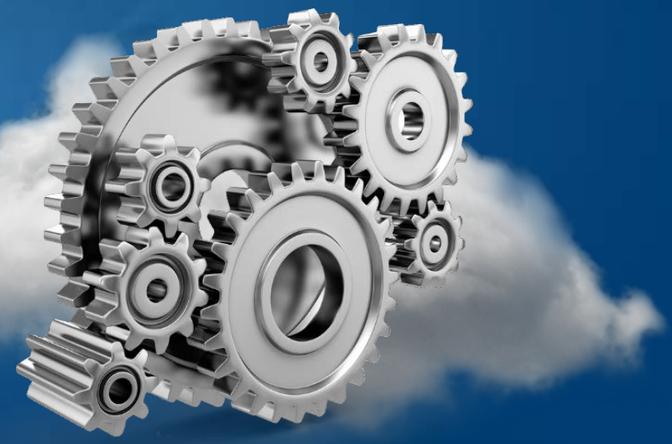
GATK
(The Genome Analysis Toolkit)

A wide variety of tools with a primary focus on variant discovery and genotyping

GenomicsDB
(Optimized for Genomics Data)

Store vast amounts of patient variant data and perform fast processing, at unprecedented scale

Cromwell
(Integrated workflow engine)



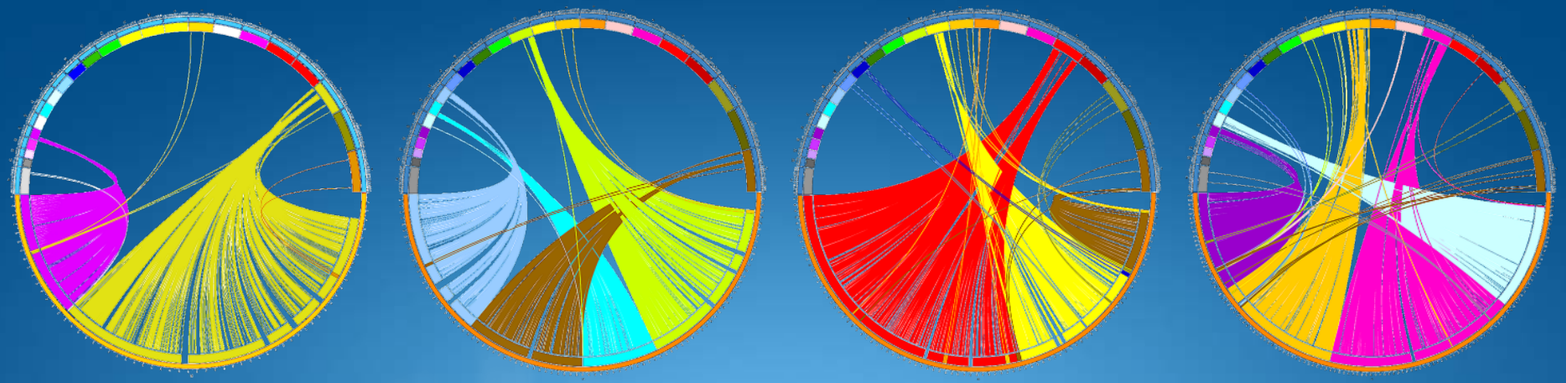
Launch genomic pipelines on private or public clouds in a portable and reproducible manner

AI in one day by 2020

- Need for federated, secure and scalable environments to enable precision medicine
- Tools –
- *http://github.com/intel-hls/GenomicsDB*
- *http://github.com/intel-hls/TileDB*
- *http://github.com/intel-hls/Cromwell*

www.intel.com/healthcare

@IntelHealth



ALL IN ONE DAY BY 2020

ALL-IN-ONE-DAY GOAL – A CHALLENGE FOR MEDICAL AND TECH INDUSTRIES

Patient

Get precision recommendations

Get DNA sequenced analyzed

Talk with doctor about precision medicine options within 24 hours.

Doctor

Deliver precision medicine options

Order DNA sequencing

Analyze results based on best available data across institutions

Share pros and cons of options with patient

IT/Technology

Enable collaboration

Deploy scalable infrastructure to drive down total cost of ownership

Allow for data sharing across institutions

Maintain security and control of local data

All within 24 hours in 2020