Personalized Medicine through Tumor Sequencing

Jeffrey Rosenfeld, Ph. D

My Background

- I am an Assistant Professor of Pathology and Laboratory Medicine at the Rutgers Robert Wood Johnson Medical School
- I supervise the computational aspects of the Precision Medicine Initiative at Rutgers University Cancer Institute of New Jersey (CINJ) in addition to working in the bioinformatics core
  - We are the only NCI Designated Comprehensive Cancer Center in NJ
- I have been working with Golden Helix software since 2009. First with SVS and now with VarSeq

Cancer

- Traditionally cancer has been treated with chemotherapy with all of its harsh effects
- Treatment was mainly focused on the site of a tumor
- Since cancer is caused by mutations driving tumor growth, targeting those mutations should control the cancer
- Starting in 2001, focused drugs such as imatinib/Gleevec, have been developed
  - Targets the bcr-abl fusion protein in chronic myelogenous leukemia (CML)

- BRAF V600E
  - Mutated in a significant percentage of cancers, especially melanoma

KNOW YOUR TUMOR℠

Powerful Knowledge. Personal Treatment.℠

CANCER IT'S PERSONAL

THE RIGHT PATIENT. THE RIGHT TREATMENT.

MY CANCER GENOME®

GENETICALLY INFORMED CANCER MEDICINE

- We are realizing that cancer is many diseases not one disease
- Individual patients have unique tumors and the treatment should be personalized

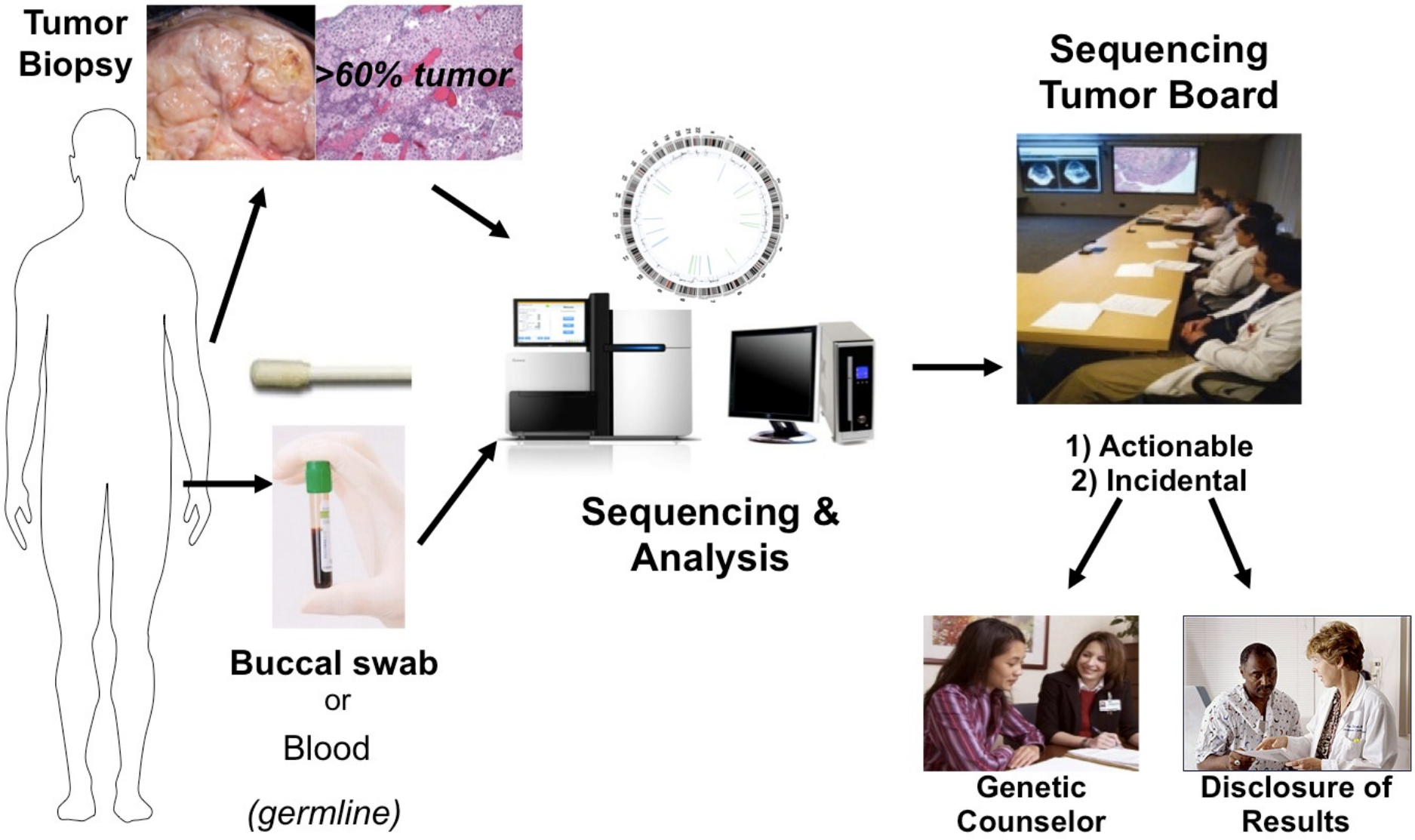
pancan.org, thegrouproom.tv,mycancergenome.org

Druggable Mutations

- Currently, there are 70 drugs FDA approved for targeted cancer treatment
  - Some of these drugs are highly similar and simply competing products
- In many cases, the drug will be approved for one type of cancer, but doctors will prescribe it off-label for other types of cancer

- A large number of drugs are currently making their way through clinical trials
  - Genetic testing is helping to be targeted in trials

Procedure



Determining Mutations in a Tumor

- When sequencing is performed, it can use either a tumor-only or a tumor-normal approach

Tumor Mutations	-	Germline Mutations	=	Mutations only in the Tumor

  - Can be drivers or passengers

- The number of sequenced tumors is rapidly growing and it may limit the need for sequencing a matched normal
  - TCGA - 11,000 samples
  - ICGC - 17,867 samples

- Sequencing a tumor and a matched normal sample is twice as expensive as just doing the tumor

Three ways for a hospital to implement tumor profiling

Commercial Lab

1. Tumor biopsy samples are sent to an outside lab for processing
   - The lab does all of the work including billing
   - Easy to implement, but no possibility of customization
   - Hard to get access to the raw BAM files for research

Commercial Providers

FOUNDATION MEDICINE®

CANCER GENETICS INCORPORATED

Personalis®
Pioneering Genome-Guided Medicine

GeneDx
DNA DIAGNOSTIC EXPERTS

Paradigm

Integrated ONCOLOGY
LabCorp Specialty Testing Group

PGD
Personal Genome Diagnostics

Genomic Alterations Detected	FDA Approved Therapies (in patient's tumor type)	FDA Approved Therapies (in another tumor type)	Potential Clinical Trials
EGFR N771_P772>KFP	Erlotinib Gefitinib	Cetuximab Panitumumab	Yes, see clinical trials section
CCND1 amplification	None	None	Yes, see clinical trials section
ARID1A Q633*	None	None	None

In-house

2. Develop everything in-house
   - Example: Memorial Sloan Kettering MSK-IMPACT
   - Test is a custom designed panel of genes
   - Analysis software is maintained by a team of local developers
   - Database of mutations is curated by local scientists
   - Allows for complete control over the data

   - Requires a large investment of time and a substantial amount of individuals with a wide range of skills

Hybrid Approach

3. Run the test locally, using different commercial parts

- Companies have developed specialized products including target panels, analysis software and annotation databases
- Lesser requirement of resources than developing everything locally
- Maintain control of the data
- Not relying simply on a “black-box” for the testing

The Rutgers Clinical Genomics Laboratory Sequencing Workflow

Our Pipeline in a Nutshell

- RainDance Thunderbolts Cancer Panel
- MiSeq
- BaseSpace
- VarSeq
- N-Of-One

- Designed to be a low-cost primary test
- Patients who do not receive a test result beneficial to their treatment can be relaxed to a larger panel

RainDance

- Enriches the genome for mutational hotspots in 50 top cancer genes
- Droplet PCR allows for even amplification

Gene List

ABL1	EGFR	GNAQ	KRAS	PTPN11
AKT1	ERBB2	GNAS	MET	RB1
ALK	ERBB4	HNF1A	MLH1	RET
APC	EZH2	HRAS	MPL	SMAD4
ATM	FBXW7	IDH1	NOTCH1	SMARCB1
BRAF	FGFR1	IDH2	NPM1	SMO
CDH1	FGFR2	JAK2	NRAS	SRC
CDKN2A	FGFR3	JAK3	PDGFRA	STK11
CSF1R	FLT3	KDR	PIK3CA	TP53
CTNNB1	GNA11	KIT	PTEN	VHL

- Total of ~200 amplicons
- All amplicons together only equal <40kb of DNA
- Does NOT cover full genes or even all exons, only hotspots

Sequencing and primary analysis

- Samples can be multiplexed to put 8 on a single Mi-seq run
  - Keeps the costs down as compared to the need for a HiSeq lane for other panels
- Alignment and variant calling are performed using BaseSpace using standard tools
- Only SNPs and indels can be called
  - No fusions, amplifications or deletions can be detected
- Calls the most highly studied tumor mutations

Data transfer

- Extensive infrastructure is needed to maintain CLIA/CAP and HIPPA compliance
- The sequencing lab and the analysis machines are located on different networks with intervening firewalls
  - BAM and VCF files need to be encrypted to be copied
  - Transfer is automatic based upon a daemon recognizing files in a designated directory
- Data files are automatically replicated archived for mandated long-term storage

Analysis Software Tradeoffs

- Commercial vs. Open source
  - Robust support from professionals
  - Software will exist beyond the timeline of a graduate student or postdoc
  - Need to pay for the software
- Graphical vs. Command Line
  - Usability for a clinician or technician without extensive computational skills
  - Potential for a bioinformatician to automate jobs and build pipelines
- Pay-per-sample vs. License model
  - Set cost that can be built into a medical bill
  - Fixed annual expenditure
  - No need to pay for re-running of samples or for non-billable/research samples

VarSeq

- Produced by Golden Helix
- A graphical interface to filter variants and to classify them using public and private data
- Roughly equivalent to what can be done using ANNOVAR or VAAST, but with a graphical interface (GUI)
- A command-line version is available where the analysis parameters can be set up in the GUI, but run automatically from a script
- Software is highly configurable by the user
- Can process our samples in < 1 minute each
- A full genome or exome will take a few minutes or longer to process
- Reduces the number of variants we find in a tumor from hundreds to <10

- Software is purchased as an annual license rather than paying per-sample
  - At Rutgers we were very attracted to the annual license model
  - Allows us to run research samples without worrying about the cost

Variant Filtering

- The raw VCF has >100 variants called for each tumor
- Most of these variants are not important in the treatment of the cancer
  - Could be germ-line variants
  - Could be passenger mutations from highly mutating tumors

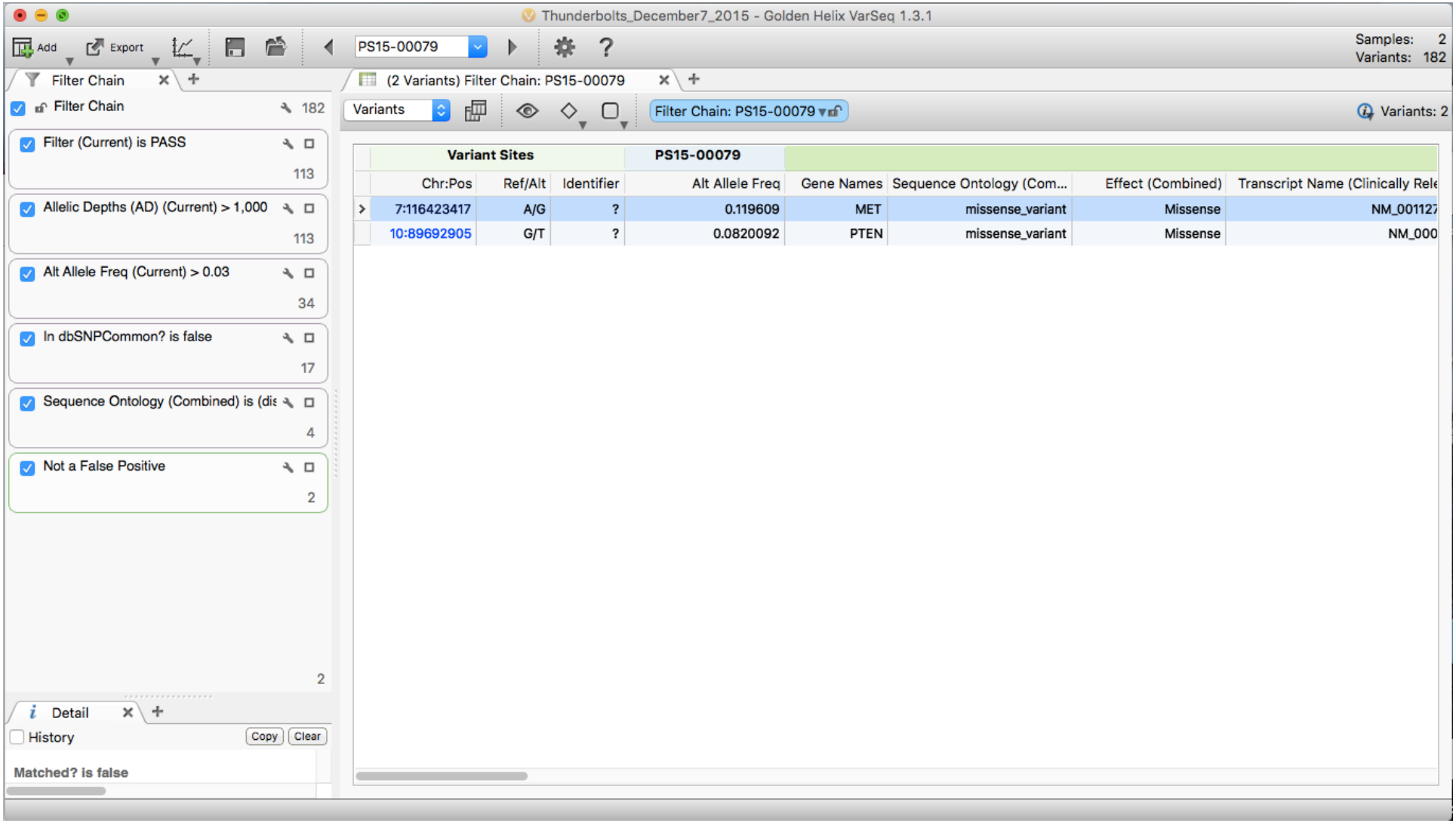
- Variants could have low read depth
  - Could reflect sequencer errors

- Minor allele frequency could be very low
  - Variant could be noise
  - A mutation only found in a very small proportion of the tumor sequencing reads is not an important clinical target

Filtration

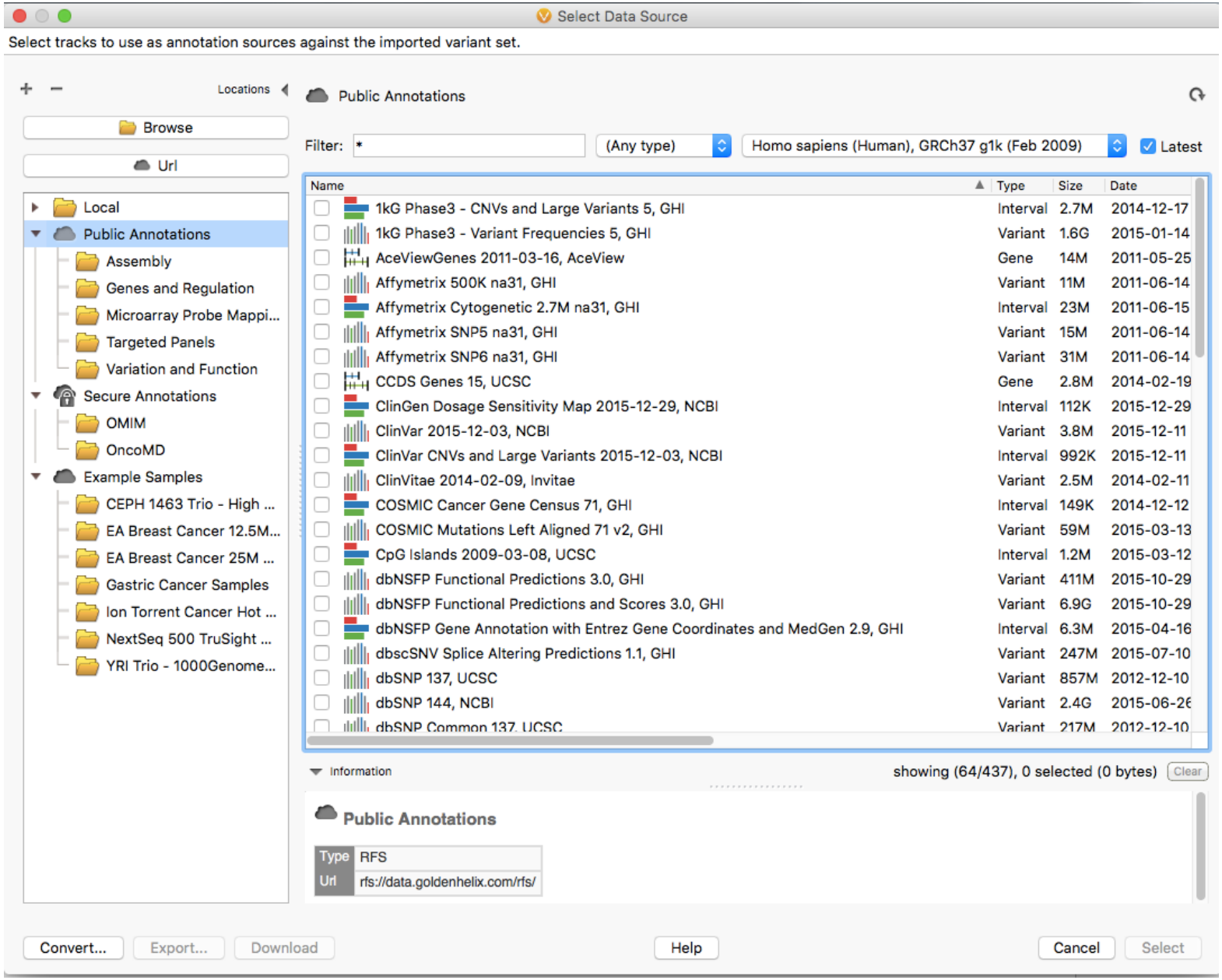
- Passes the standard BaseSpace filters
- Read Depth is > 1000X
- Minor allele frequency is > 0.05
- Not found in the >=1% in any 1000 Genomes population
- Non-synonymous

- Not a known false positive from the assay

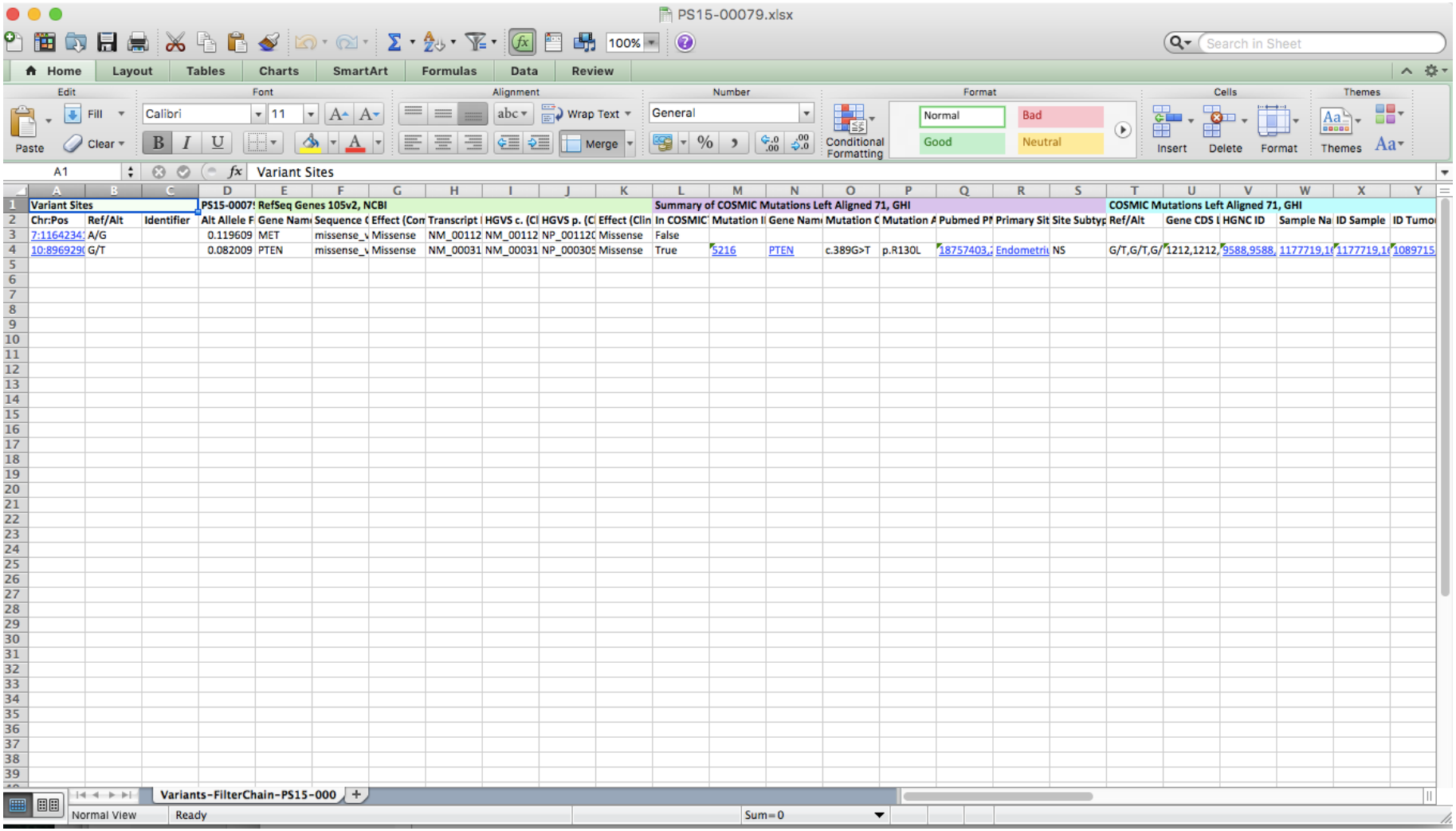


- Reduced list of variants from 182 => 2

Many Available Data Sources



Results



PS15-00079.xlsx

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Variant Sites			PS15-0007	RefSeq Genes 105v2, NCBI							Summary of COSMIC Mutations Left Aligned 71, GHI								COSMIC Mutations Left Aligned 71, GHI					
2	Chr:Pos	Ref/Alt	Identifier	Alt Allele F	Gene Nam	Sequence (	Effect (Con	Transcript	HGVS c. (Cl	HGVS p. (C	Effect (Clin	In COSMIC	Mutation I	Gene Nam	Mutation C	Mutation A	Pubmed PI	Primary Sit	Site Subtyp	Ref/Alt	Gene CDS I	HGNC ID	Sample Na	ID Sample	ID Tumo
3	7:1164234	A/G		0.119609	MET	missense_v	Missense	NM_00112	NM_00112	NP_00112(	Missense	False													
4	10:896929(	G/T		0.082009	PTEN	missense_v	Missense	NM_00031	NM_00031	NP_00030	Missense	True	5216	PTEN	c.389G>T	p.R130L	18757403,2	Endometri	NS	G/T,G/T,G/	1212,1212,	9588,9588,	1177719,1	1177719,1	1089715

Variants-FilterChain-PS15-000

Normal View | Ready | Sum=0

Clinical evaluation of variants

- After the variant list is narrowed down, there is a need to determine whether the variants are clinically actionable
- Literature on variants is continually changing
  - Hundreds of papers are published each week

- Clinical Trials are constantly opening and closing
  - Trials can have very strict criteria for accepting patients
  - Patients have different geographical constraints on where they can travel for treatment

Ingenuity

Expert Findings

Expert Findings are experimentally demonstrated Findings that are manually curated and reviewed from the full-text of articles in top journals for accuracy and contextual details. Ph.D. scientists, who follow a strict content extraction protocol to ensure inclusion of contextual details, model these Findings.

- Findings from ~300 top journals are curated from the full text, including tables and figures
- Supports computation and answering in-depth biological questions in the relevant context
- Weekly updates keep information fresh and up-to-date

ExpertAssist Findings

ExpertAssist Findings are manually reviewed, automatically extracted Findings from the abstracts of a broad range of recently published biomedical journals. The extraction protocol is modeled on the same process used by our Expert Findings. Additionally, the information is manually reviewed before import into the Knowledge Base. This means that unlike other automatic approaches, QIAGEN's Ingenuity ExpertAssist Findings maintain a very high level of quality, have proper synonym resolution, capture both contextual details and broad functional relationships, and are computationally accessible.

- Weekly updates capture information published from the prior week's publications
- Includes Findings from ~3600 journals
- Supports computation and answering in-depth biological questions in the relevant context

N-Of-One

- Instead of needing to maintain a database of publications and trials, we outsource to N-Of-One
- Provided the analysis for Foundation Medicine several years
- Company maintains a thorough database of literature and clinical trials relating to genome variants

N-Of-One

- Data is exported from VarSeq as a text or Excel file and then converted to XML
- We transfer our results to them using XML and they identify all known clinical information about a variant
- Report is returned to us as XML and used to generate a signed clinical report for physicians

- Results are returned within a few hours for known mutations, but may require a few days for a novel mutation

Reporting

- The annotations from N-Of-One are reviewed by the Molecular Pathologist and a final signed report is generated
- This report becomes part of the patient's electronic medical record (EMR) and is used to help determine their treatment

Report

Solid Tumor Targeted Mutation Panel by Next Generation Sequencing

Patient name: **Report Date:** May 03, 2016

Patient Information		Specimen Information	
RUCDR Record #:		**Tumor Type:**	Small bowel adenocarcinoma
MRN:		**Specimen Type:**	FFPE Slides
Date of Birth:		**Specimen Site:**	Right Ovary
Gender:		**Date of collection:**	
Referring MD:	Kennedy, Timothy/Gibbon, Darlene	**Date received:**	
		Pathology Specimen ID:	

1. Results

1.1 Positive Biomarkers

Mutations/Hotspots Identified	Biological Association	Allele Freq	FDA Approved Treatments	Off Label Potential Treatments	Clinical Trials
GNAS R201H	G protein	0.36	No	Yes	Yes
KRAS G12D	KRAS	0.57	No	Yes	Yes

1.2 Variants of uncertain clinical significance

Mutations/Hotspots Identified	Biological Association	Allele Freq
ATM A1942V	DNA repair	0.61

VUS: The functional or therapeutic consequences of VARIANTS OF UNCERTAIN CLINICAL SIGNIFICANCE are unknown.

Molecular Tumor Board

- At CINJ, we have a weekly molecular tumor board where oncologists, pathologists, and bioinformaticians review the findings from the molecular testing
- With all of the professionals working together, and taking into account the patient's condition, we determine the proper course of treatment

Clinical Data Warehouse

- The molecular test results along with diagnostic and treatment data from the EMR and histology images are collected into a database
- This data can be queried to ask innovative research questions on the population of patients
- How well did breast cancer patients with a particular mutation do on a specific drug?
- Can we better determine which mutations should be targeted by specific drugs by using a set of mutations rather than just a single mutation?

Summary

- Clinical Tumor Sequencing is dramatically changing cancer treatment
- There are a large number of computational steps that are needed for doing clinical tumor sequencing
- Through the use of these panels, we have greatly enhanced the outcomes for our patients.