Welcome to Workshop

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015

Please set computers, cell phones, and blackberries on silent mode, and answer all calls in the hallway

Webcast at https://collaboration.fda.gov/ivdtw111215/

Food and beverages will be available for purchase by workshop participants at the Sodexo kiosk in the registration lobby during breaks / lunch

Wi-Fi can be accessed in the Great Room area using *guestaccess*

Links to the meeting transcript and the archived webcast will be posted to the workshop registration webpage approximately 6-8 weeks after the meeting

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015

Welcome

Paula Caposino, PhD

Office of In Vitro Diagnostics and Radiological Health
Center for Devices and Radiological Health
U.S. Food and Drug Administration

Robert Califf, MD

Deputy Commissioner for Medical Products and Tobacco

U.S. Food and Drug Administration

Disclaimer

NGS is raising new policy / regulatory issues; thoughts presented here are preliminary and do not represent proposed or finalized FDA policy – these are just preliminary ideas or approaches we are considering

Precision Medicine Initiative: Modernizing FDA Regulation of Genomic Tests

November 12, 2015

E. David Litwack, Ph.D.
Personalized Medicine Staff
Office of In Vitro Diagnostics and Radiological Health (OIR)
Food and Drug Administration (FDA)

Precision Medicine – Tailoring Health Care to Each Patient

- The right treatment
- The right patient
- The right time

The President's Precision Medicine Initiative (PMI)

To enable a new era of medicine through research and technology that empowers patients, researchers, and providers to work together toward development of individualized treatments.

Success of Precision Medicine Requires:

- ***Safe and accurate diagnostic tests*** that reliably identify individual variation
- ***Learning health systems*** that enable researchers and clinicians to learn from and inform the patient experience
- ***Development of targeted therapies*** that are more efficacious or have less deleterious side effects for specific individuals
- ***Updated research and regulatory policies*** that catalyze the development of new treatments while protecting patients

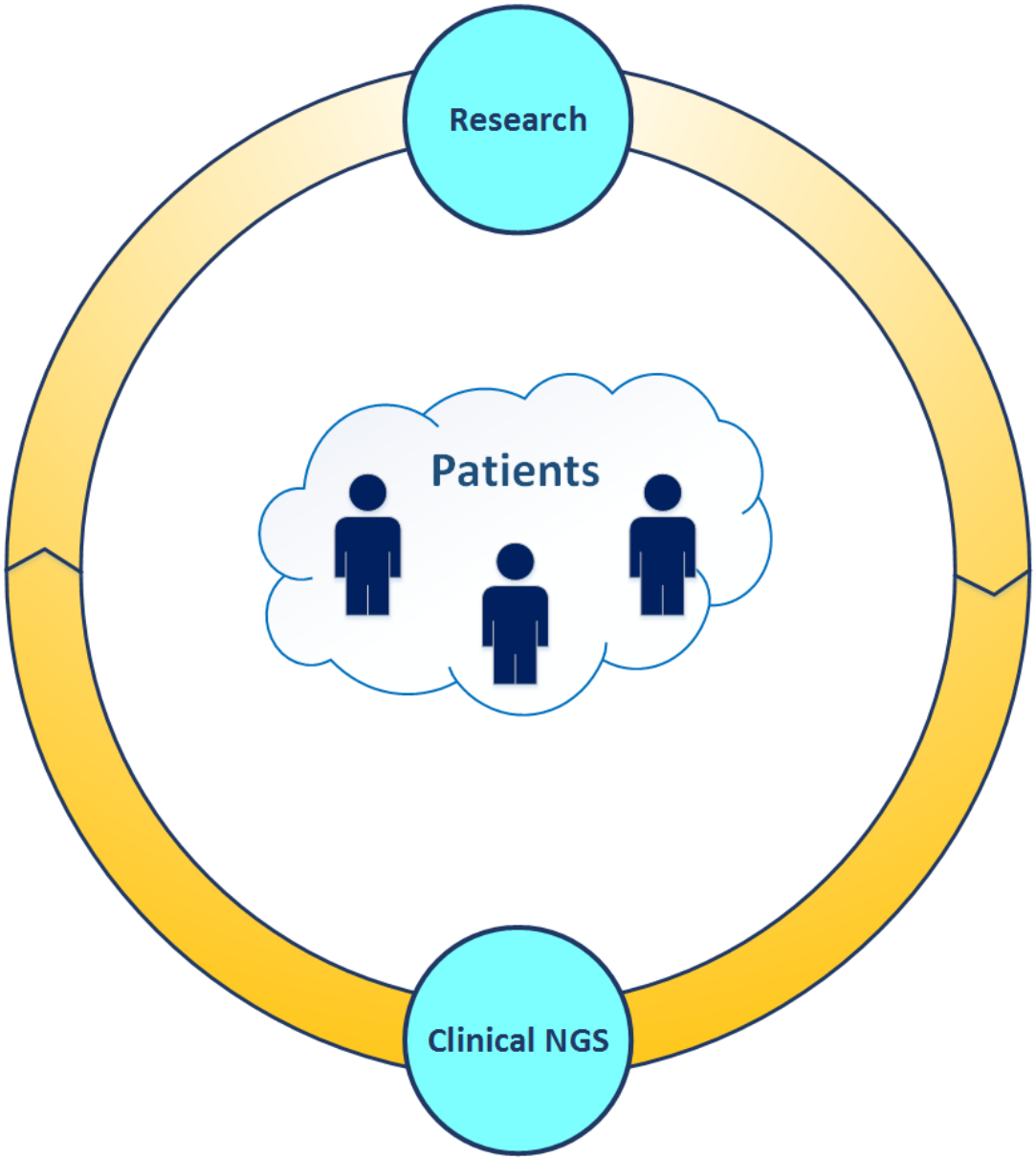
Precision Medicine Initiative: Modernizing FDA Regulation of Genomic Tests

Traditional testing

Next generation

NGS tests often lack a specific intended use

- Can’t predefine the results that will be obtained
- Often don’t know the disease that will be diagnosed until the test is performed


Research
Patients
Clinical NGS

Precision Medicine Initiative - FDA

Vision: Implement new regulatory policies to promote research and accelerate the translation of precision medicine technologies into treatments that *benefit patients*.

- Near Term: Implement standards and shared resources that will enable the development of knowledge for research and patient decision making
- Longer Term: Implement standards-based regulation of diagnostic tests that will ensure that the tests patients receive provide accurate, reproducible, and meaningful results

Modernizing FDA Regulation of Genomics

- Develop and implement ***standards*** to assure quality
- Develop ***open-source tools*** to help test developers meet standards (*precisionFDA)*
- Support the development of a ***data commons*** for evidence on the clinical relevance of genetic variation

Develop and implement an adaptive standards-based regulatory approach

Public Engagement

- FDA workshop on Feb 20, 2015 discussed the overall vision
  - http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm427296.htm
- Follow-on workshops to discuss technical details
  - Analytical performance - Nov 12, 2015
    - http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm459449.htm
  - Clinical interpretation – Nov 13, 2015
    - http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm459450.htm
  - Oncopanels - Feb 2016
  - Patient/provider perspectives – March 2016

Partners

- Stakeholders involved for each area
  - PMI partners
  - NIST / GIAB Consortium
  - UCSF-Stanford CERSI
  - Professional Societies
  - CDC
  - Industry
  - Clinical laboratories

Standards Based Approaches to NGS Analytical Validity

Zivana Tezak, PhD

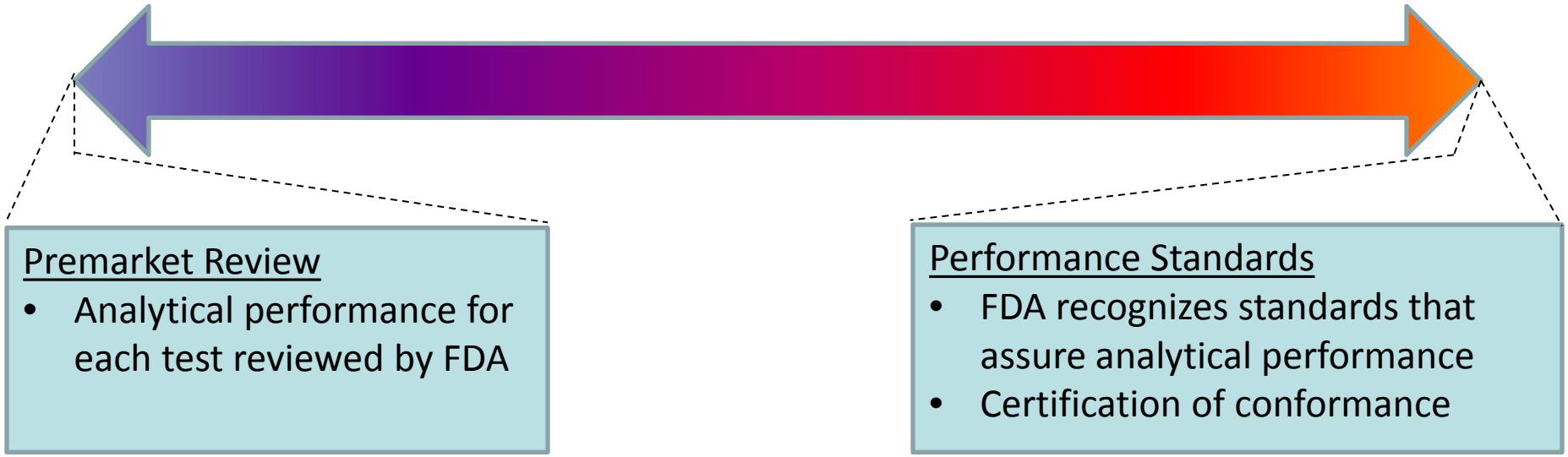
Personalized Medicine Staff
Office of In Vitro Diagnostics and Radiological Health
U.S. Food and Drug Administration

Silver Spring, MD, November 12, 2015

Analytical performance

- Analytical validity
  - Does the test correctly and reliably detect what it says it does?
  - Accuracy, precision, interference, cross-reactivity, limits of detection, other information/data relevant to test performance
- Challenges for NGS tests
  - Amount of data / bioinformatics needs
  - Almost unlimited number of variants could be detected
  - Rapidly evolving technology
  - Intended use often not easily defined
  - Different clinical applications

February 2015 - Analytical performance standards as a regulatory tool for NGS?



- Are performance standards feasible? Sufficient?
- How should they be implemented?
- **Who** should develop them?
- How should conformance to standards be verified?

FDA - discussion paper, public meeting on February 20, 2015
http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm427296.htm

November 2015 – A spectrum of approaches from performance standards to design concept standards

Performance Standards

- Finalized test evaluation
- Establishing specific metrics and acceptance criteria that the test would have to satisfy

Design Concept Standards

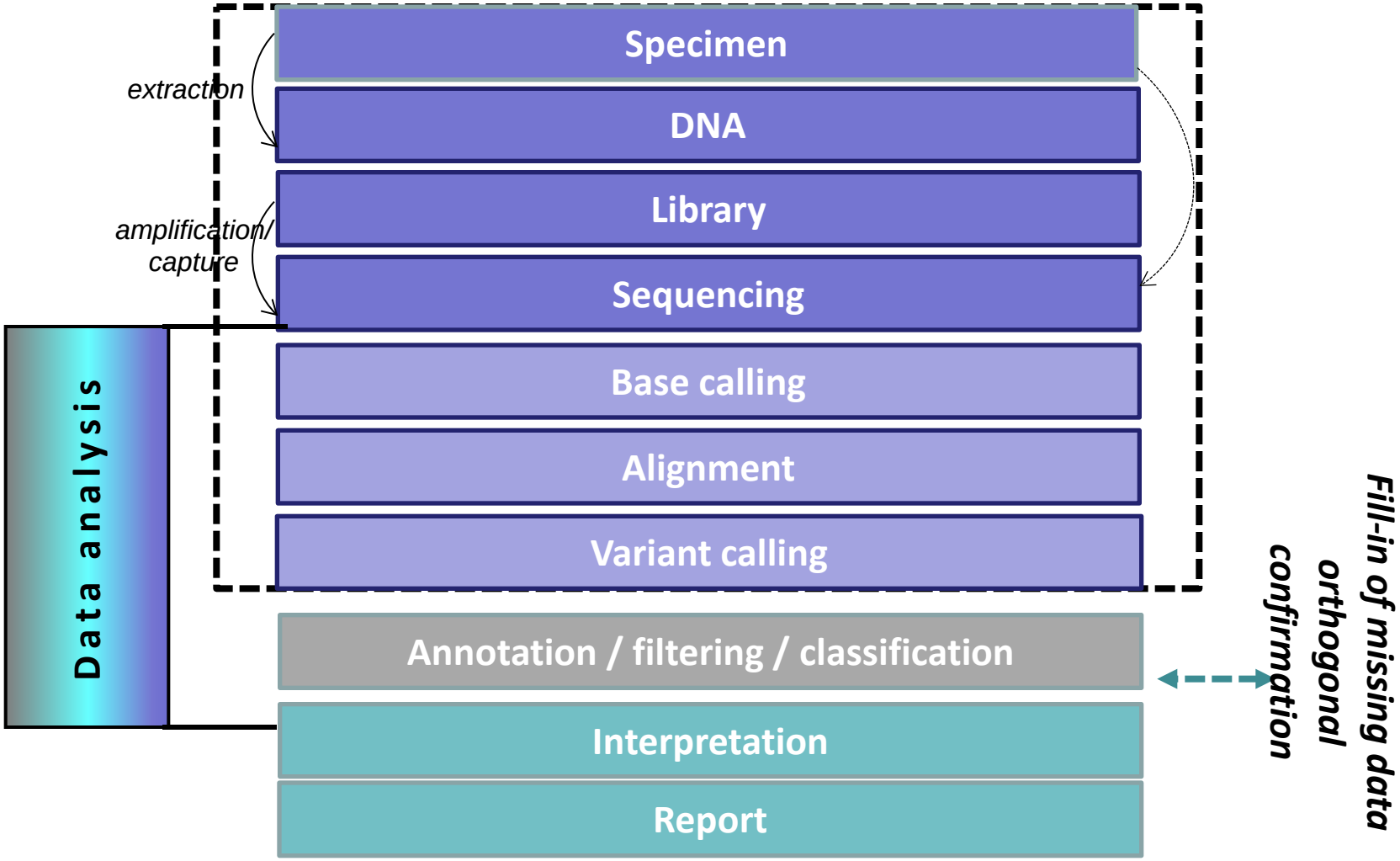
- Test design and development
- Ensure that the test developer can achieve the proper design and validation of an NGS test

- Which approach?
- Hybrid approach? If combination, which elements from each approach?

Transparency?

FDA - discussion paper(s) - public meeting on November 12, 2015
http://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM468521.pdf
http://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM468520.pdf

NGS test steps



Performance standard approach

- Intended use of the test (e.g., carrier screening, diagnosis of inherited disease, cancer risk prediction)
  - ***What*** to measure - metrics (e.g., accuracy, precision, sensitivity/specificity, quality scores)
  - ***How*** to measure - protocols and studies to perform (e.g., specimen types, sample numbers, comparator methods)
  - **Defined** acceptance criteria, thresholds for measurements

  (Performance data / result presentation)

Design concept approach

What is the purpose of testing = **Intended Use**

Verification

- Who will be tested, disorders, target panel composition / coverage of clinically relevant targets, variant types
- What are the test inputs? (specimen type, volume, turnaround time, workflow, limitations)
- What performance is needed? (level of accuracy, precision, what and how to evaluate)
- What strategy will be used? (targeted, exome, amplification, capture)
- What components are needed? (reagents, instruments, software)
- How should test components be put together? (protocols, processes)

Consider each critical factor in the test development process and their **impact on the overall design** of the test

Validation = performance evaluation

Standards and guidelines

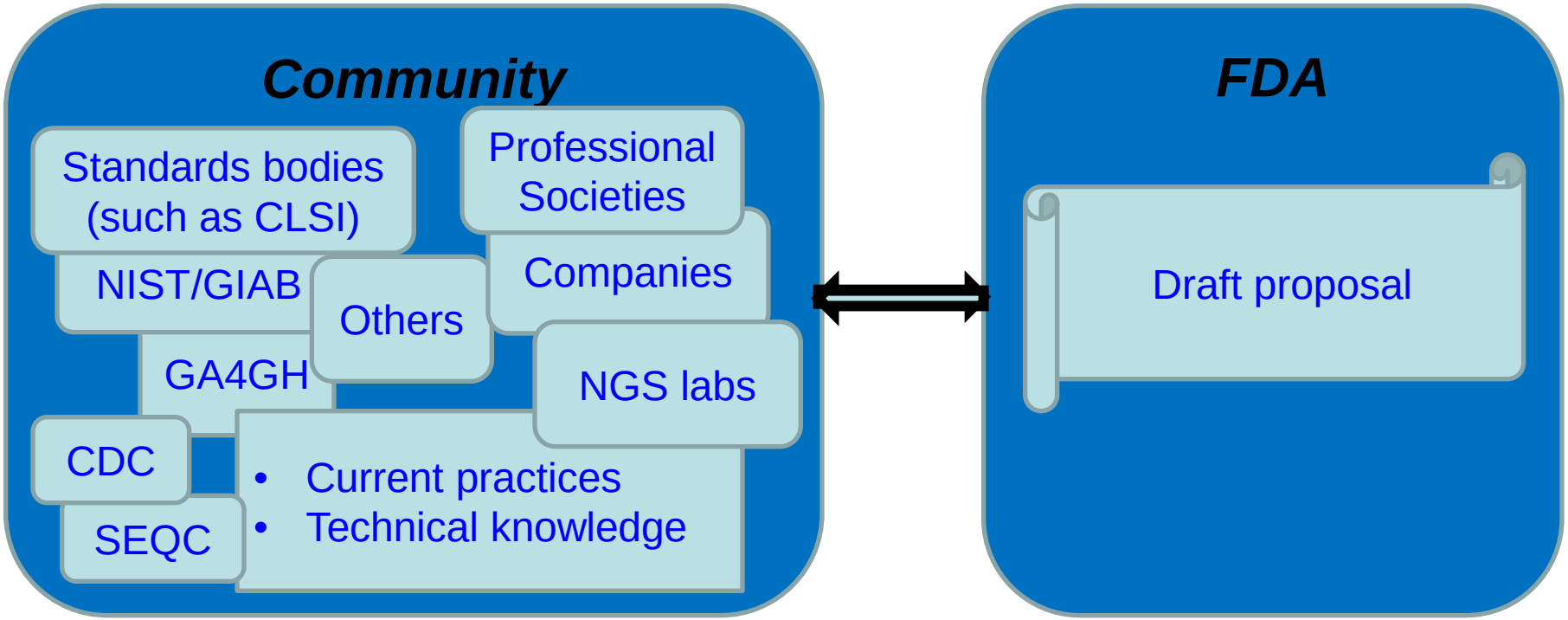
- Examples of currently existing or ongoing standard development, how they correspond to FDA needs:
  - Material standards – e.g., NIST, sample panels
  - Methodological and performance standards – e.g., NYSDOH, CAP, MolDX, GA4GH, Nex-StoCT, ACMG, HGVS, CPIC, CLSI
- What can be used, for what purpose
- Identify gaps
- Questions:
  - Who should fill identified gaps
  - What is acceptable timeline?
- Need built-in flexibility / adaptability

Areas in need of standards?

Initial feedback and assessments; where to start?

- **Documentary standards**
  - by intended use
  - by methodology, test building blocks
  - by classes of variants, genomic regions
- **Material standards**
  - reference samples, disease-specific sample sets, controls
- **Computational solutions**
  - run quality, process metrics
- **Terminology/nomenclature** (language agreement and definitions)
- Appropriate comparators and "gold standards"
- **Data formats** and interchangeability / interoperability

Development and use of analytical standards



Today's workshop

- Panel 1 discussion:

  Standards-based approaches to analytical validation – A Spectrum from Design Concept Standards to Performance Standards

- Panel 2 discussion:

  Developing analytical standards for NGS-based assays

- Panel 3 discussion:

  Developing bioinformatics strategies and tools for evaluating NGS tests

- PrecisionFDA:

  A cloud-based community platform for NGS assay evaluation and exploration

- Public comment

Please Comment!

- We also ask for broader responses to our questions
- Submit written comments at:

http://www.regulations.gov/#!docketDetail

- Docket is open until November 25th

Panel Discussion of Topic 1
Standards-based approaches to analytical validation – A Spectrum from Design Concept Standards to Performance Standards

Moderator: Zivana Tezak, PhD

Speakers and Panelists

- **Birgit Funke, PhD (Harvard Partners Healthcare)**
- **John Pfeifer, MD, PhD (Washington University School of Medicine)**
- **Jared Maguire, PhD (Counsyl)**
- **Geoff Otto, PhD (Foundation Medicine)**
- **Arend Sidow, PhD (Stanford University)**
- **Erasmus Schneider, PhD (New York State Dept of Health)**

Birgit Funke, PhD

Partners Laboratory for Molecular Medicine

Standards-based approaches to analytical validation

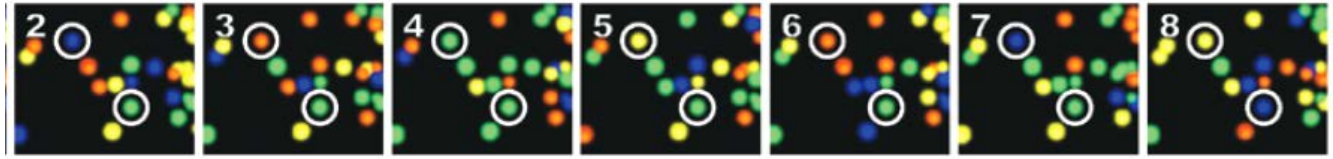
A Spectrum from Design Concept Standards to Performance Standards

FDA Workshop November 12, 2015

Birgit Funke, PhD, FACMG

Associate Professor of Pathology, Harvard Medical School

Director Clinical R&D: Laboratory for Molecular Medicine



http://www.nature.com/nrg/journal/v11/n1/images/nrg2626-f2.jpg

DISCLOSURES

Consultant:

- 23andMe

Academic affiliation:

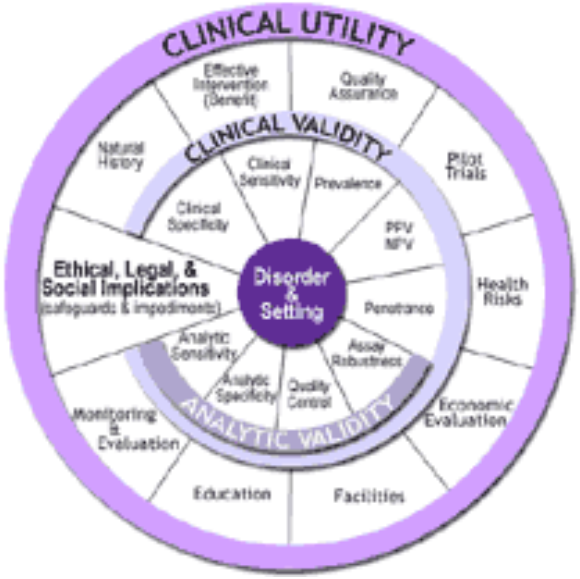
- Harvard Medical School, MGH

DISCLAIMER

Although based on many discussions with and documents of national expert committees, this presentation reflects my own opinion.

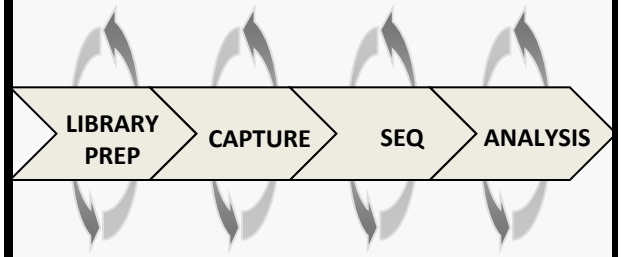
DEVELOPMENT AND VALIDATION OF AN (NGS) TEST

DESIGN



http://www.cdc.gov/genomics/gtesting/ACCE/

OPTIMIZATION



- Test each step
- String together
- Optimize
- Lock down protocol

VALIDATION

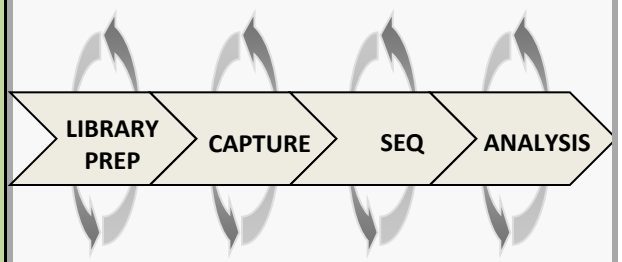
Test analytical performance using locked down protocol

NGS TEST DESIGN

DESIGN

- Select test contents
- Characterize contents
- Clinical needs
- Technical limitations
- Supplemental assays?
- Fill-in and variant confirmation ?
- Establish development + validation plan

OPTIMIZATION



- Test each step
- String together
- Optimize
- Lock down protocol

VALIDATION

Test analytical performance using locked down protocol

DESIGN - A THREE DIMENSIONAL EXERCISE

Content

Gene selection
- Clinical validity

Gene characteristics
- Homology, High GC/AT?
- Pathogenic variant types
- Mosaicism?

X

Technology

Sequencer
- capacity/ln
- Read length

Bioinformatics pipeline
- SNV
- Indel
- CNV
- SV

=

Techn. Reality Check

Operational
- Target size
- # samples/lane
- **_test cost, TAT**

Diagnostic
- % gene unsequenceable
- Difficult variants prevalent
- Ability to detect low AF

Clinical Reality Check

Consult domain experts
- Difficult gene critical?
- Alternate assays offered elsewhere
- Required analytical sensitivity*?
- Required completeness

**** TP/TP+FN***

Development + validation plan

Supplemental assay/s when:
- NGS suboptimal but gene or variant type essential

Validation (reference) samples
- Specimen types
- Variant Types
- Reference material sources

Fill-in +confirmation strategy

HOW TO STANDARDIZE…THERE IS NO "ONE SIZE FITS ALL"

Two difficult to sequence genes – two solutions* **(Laboratory for Molecular Medicine)**

Both present similar challenge

	Renal Panel (PKD1)	Hearing Loss Panel (STRC)
Essential gene (detection rate)	YES (85%)	YES (10%)
Sequenceable via NGS	Highly homologous (partial)	Highly homologous (entire gene)
Pathogenic variant spectrum	SNVs, indels >>CNVs	CNVs >>SNVs, indels

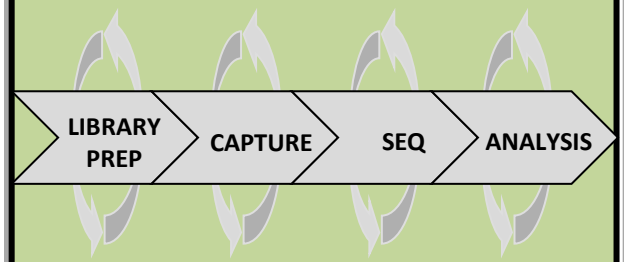
Two very different solutions

TEST DEVELOPMENT/OPTIMIZATION

DESIGN

- Select test contents
- Characterize contents
- Clinical needs
- Technical limitations
- Supplemental assays?
- Fill-in and variant confirmation ?
- Establish development + validation plan

OPTIMIZATION



- Test each step
- String together
- Optimize
- Lock down protocol

VALIDATION

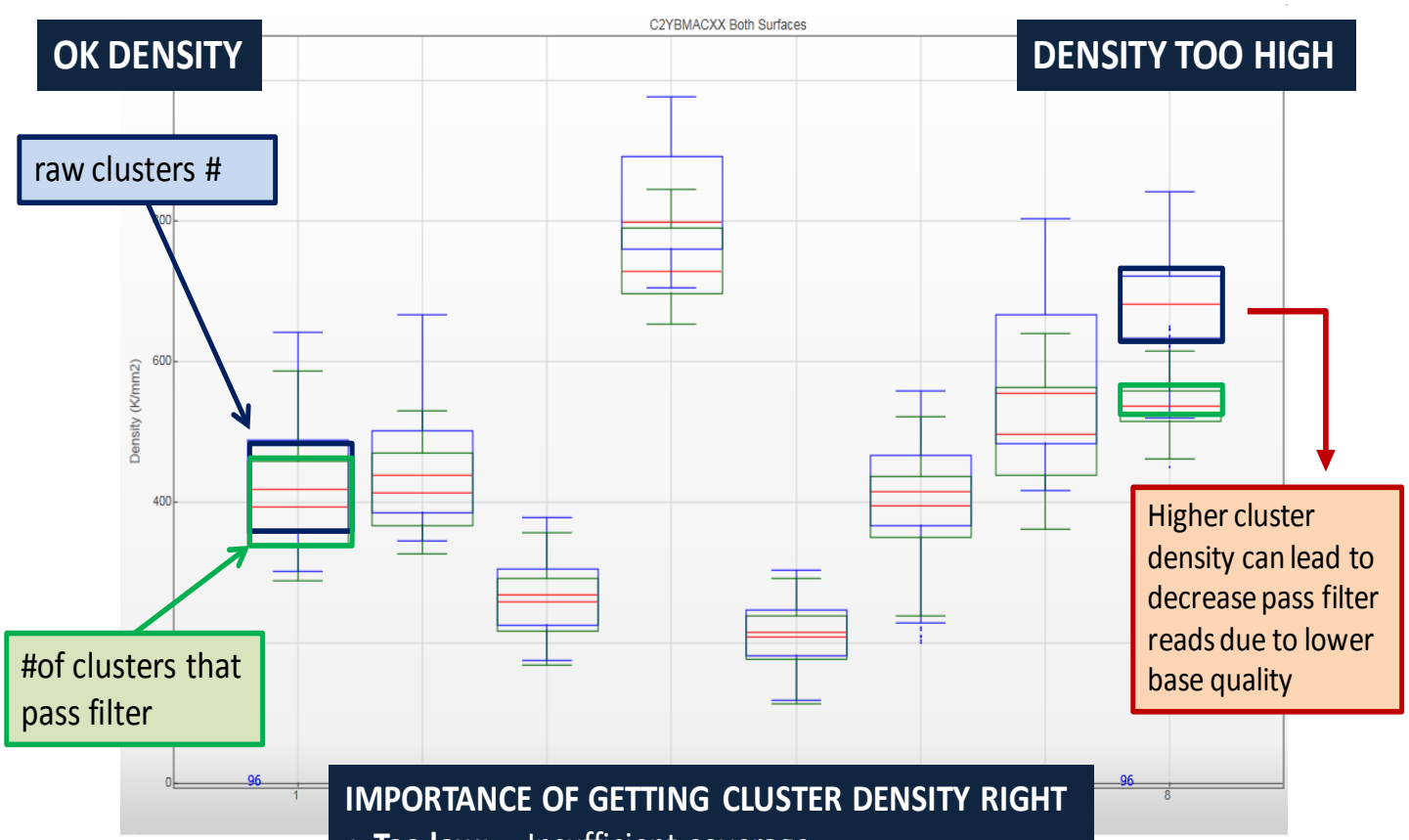
Test analytical performance using locked down protocol

QUALITY CONTROL STEPS

DNA Extraction	• OD 260/280
Library Generation	• Fragment size (each step)
Sequencing	• Cluster Density, Error rate • Total reads, % reads pass filter
Alignment, Variant Calling	• % reads on target, avg coverage • Min cov/base, library complexity, • MQ, Strand Bias, GC/AT bias, etc
Filtration, Prioritization	
Variant confirmation	• Concordance w/ NGS result
Variant + Gene assessment	• Double review
Interpretation+ Report	• Double review of variants and interpretation

- **Q/C Approach?**
  - Individual steps?
  - End-to-end?
  - → Both is best!
- **End-to-end critical measures**
  - Completeness
  - Accuracy

EXAMPLE Q/C STEPS - CLUSTER DENSITY (HiSeq)



IMPORTANCE OF GETTING CLUSTER DENSITY RIGHT

- **Too low:** Insufficient coverage
- **Too high:** increased FP

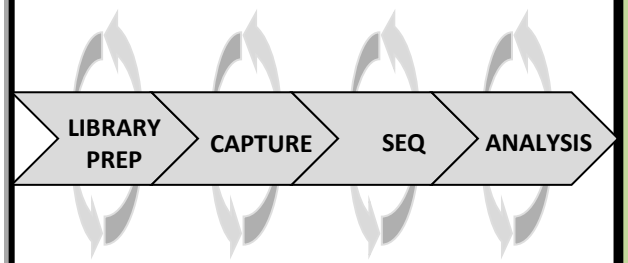
Difficult to prescribe standard – reduced quality at this step can be acceptable depending on the overall test design (fill in, confirmation)

TEST DEVELOPMENT

DESIGN

- Select test contents
- Characterize contents
- Clinical needs
- Technical limitations
- Supplemental assays?
- Fill-in and variant confirmation ?
- Establish development + validation plan

DEVELOPMENT



- Test each step
- String together
- Optimize
- Lock down protocol

VALIDATION

Test analytical performance using locked down protocol

VALIDATION LAYERS
EXAMPLE: ANALYTICAL SENSITIVITY

TIER 1: BASE LINE (**METHODS BASED**, CUMULATIVE) **CAP**

NA12878	**Test 1** (100 genes)	**Test 2** (70 genes)	**Test 3** (30 genes)	**Test 4** (200genes)	**NIST**
Variants	600	300	100	800	→ Analytical sensitivity **(+CI)** *(=TP/TP+FN)*

TIER 2: **DISEASE/GENE SPECIFIC** CONSIDERATIONS

- Well character. gene + n samples (e.g. frequent pathogenic variants)
- In/del rich gene + n samples
- CNV rich gene + n samples
- Mosaicism + n samples

CDC

Need for more reference materials!

VALIDATION FAQs

- **Should there be a minimum # of SAMPLES?**
  - **No** for baseline sensitivity/specificity. For sensitivity:
    - # variants tested is more important than # samples (stat. conf.)
    - # variants depends on target size (the larger the target the more)
  - **Probably** for precision (inter/intra run variability)
  - **Possibly** for disease/gene specific validation add-ons

- **What is an acceptable analytical sensitivity?**
  - It depends!!
  - While generally sensitivity should be very high there are scenarios where reduced sensitivity is OK
  - Needs disease/gene expert input

- **What is an acceptable analytical specificity?**
  - Reduced specificity is usually not acceptable for clinically significant variants (germline)
  - Optimal specificity can be achieved by way of orthogonal confirmation of variants (most molecular dx tests are a patchwork!)

MANY AVAILABLE RESOURCES

Content

Expert created
- GeneReviews
- Clinical validity assessments (eg **ClinGen**)
- Must test variant lists
- (**ACMG, AMP, etc**)

X

Technology

Vendor and expert created
- Instrument performance
- Analysis – ability to detect different variant types
- Other resources (homologous regions, etc)

=

Techn. Reality Check

Done by testing lab

Clinical Reality Check

Collaborative exercise
(testing lab – physician)

Development + validation plan

- Published prefessional guidelines
  - **ACMG, AMP, CAP, CLSI, CDC, etc**
- Reference materials
  - **Coriell, CDC, EQMN, Etc**

- Knowledge, resources and guidelines exists!
- Growing professional society + other community efforts
- **But….. De-centralized!**
- Requires knowledge to find information
- May be a problem for "newcomers"

WHAT DOES THE COMMUNITY NEED TO DO??

HARMONIZE EXISTING RESOURCES (e.g. by disease or specialty focused)

- Lab director answers structured questionnaire
- Answers trigger automated creation of validation plan
  - Tool outputs reference materials needed + where to get them
  - Provides in silico reference data sets
  - Lists problematic regions need for supplemental tests
- Lab director can
  - Upload fastq + order custom mutagenesis to create controls for most important genes/variants
  - Upload validation data + get automated performance metrics for standard reference materials (e.g. NA12878)
  - Access reporting and interpretation guidelines

SUMMARY

Design concept approach vs. performance standards

- Molecular genetic testing community is familiar with **BOTH**
- Current practice is a mix of **BOTH**
- Performance standards are difficult to generalize

What is needed

- Centralization of resources + guidelines
- **$$** to create additional resources

ACATAATCTACGAACTATCAATGTTTATGATGG

ATTTGAATCTGATAATGCGAAGAGTTGCTAATA

ATCTTGGATTCTATCGATAACAGCCGAGGTGCC

TACTTTGGATACTTTGACAGGTGGACACTCAAA

ATGGCAAACGTATTCCTGAGACTGCCAGAGCTG

GGCTTTATTGAAGTACCATCTTACATTTTAAAC

TCACGTTACGAAAGATAACAT**THANKS!**AGTCT

ATATCAAAAGTGATCATAATTCTGAAAATCCTT

AATGGATATTTAACCTTGGCTCCTAATTTCGGT

AGGTGGTTTTGTAACTATTTGCAGACATCCATC

GTATAATAAAAAGATCAGAAGGGTTTACTATTA

TTTAATGCTGATTTTGATGGAGATGAGATGACA

CAATCTCGAACAAGCTTTGATTATGAACTCACG

GCAATCCAATGTTCGGCTTGGTCCAAGATCAAA

John Pfeifer, MD, PhD

Washington University
School of Medicine

Workshop: Standards based approaches to analytical validity

John D. Pfeifer, MD, PhD

Vice Chair for Clinical Affairs, Department of Pathology

Washington University School of Medicine

Washington University in St. Louis
School of Medicine

GENOMICS AND PATHOLOGY SERVICES
GPS@WUSTL
CLINICAL TESTING • RESEARCH • EDUCATION

Disclosures

Consultant:

- Novartis
- Agilent
- Illumina
- Strand Analytical Labs

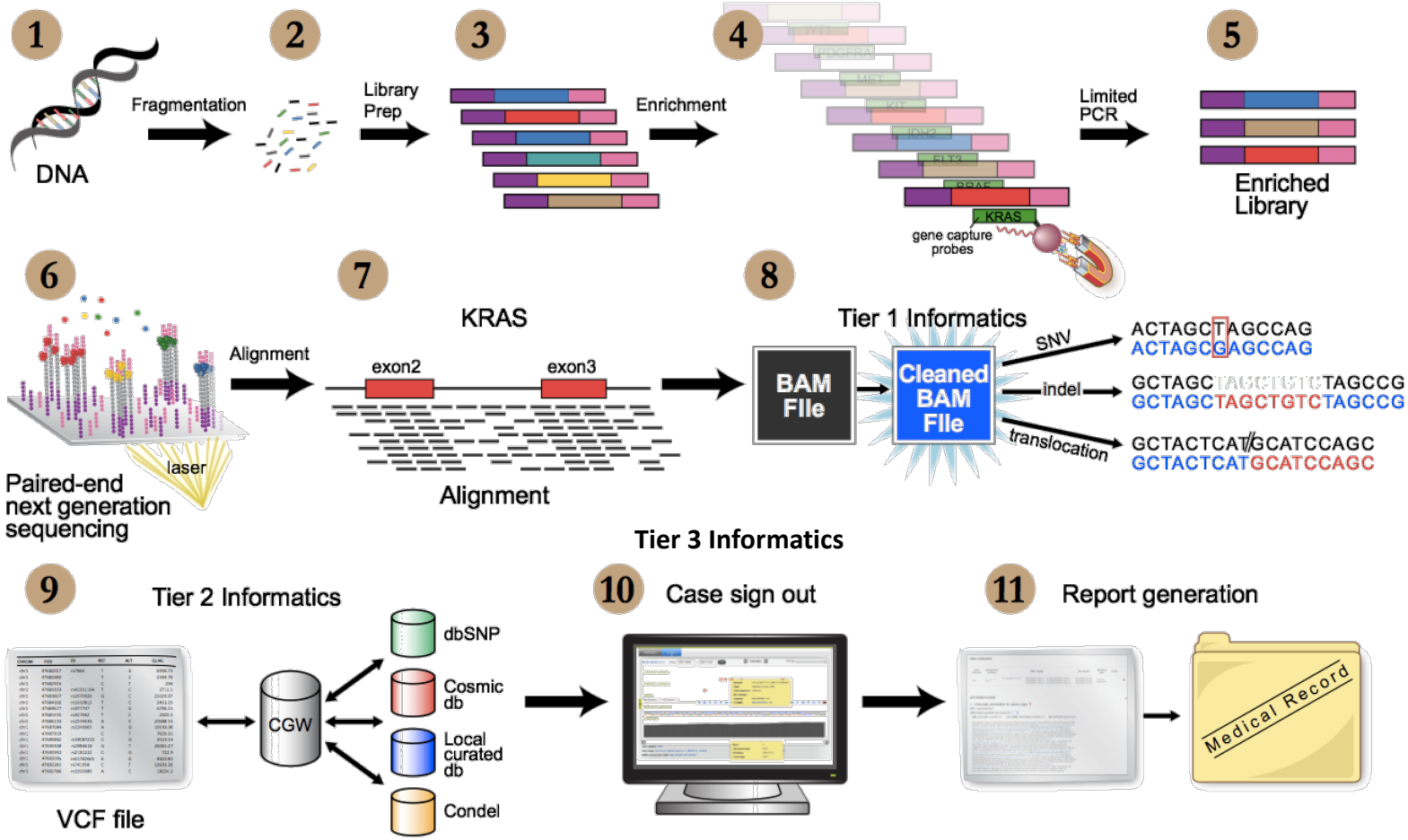
Co-founder:

- PierianDx
- P&V Licensing, LLC

Academic affiliation:

- Washington University School of Medicine is an academic, tertiary care, urban medical center
- GPS is a reference lab, within the Department of Pathology, that performs a range of molecular testing, including NGS

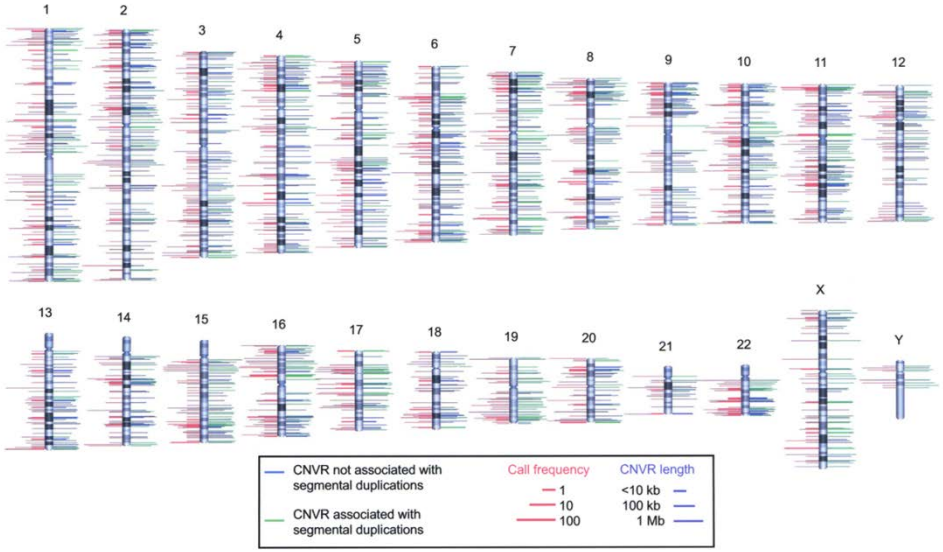
Clinical NGS testing for acquired variants is similar to testing for inherited variants…



Reference: Cottrell CE, et al. *J Mol Diagn* 2014;16:89-105

… but there are some important differences

- Often more variability in test design (hot spot versus whole gene; amplification-based versus hybrid capture based; tumor only versus paired tumor-normal)
- Often more variability in the number of genes tested (dozens, to hundreds, to exome, to whole genome)
- Often more complexity in the mutational profile (coexisting SNVs, small and large indels, CNVs, and SVs)
- Wider variation in variant allele frequency (VAF)
- Overlap with germline variation
  - SNPs, indels, and CNVs



Reference: Redon R, et al. *Nature* 2006;444:444-454

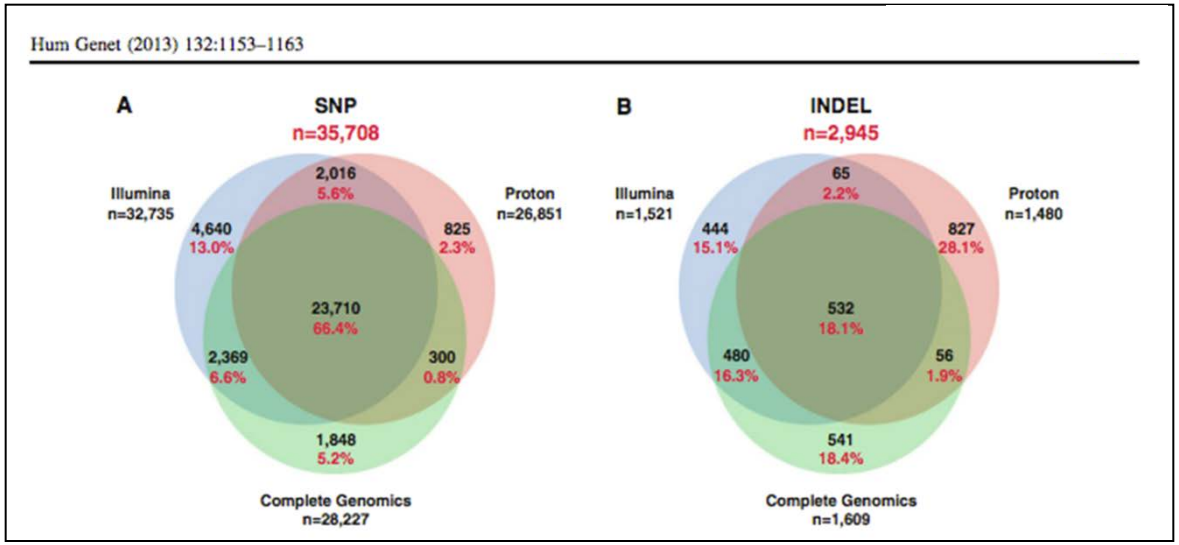
Clinical "Cancer NGS" is not one thing

Analytically:

- Different platforms for "NGS"
- Different assay designs for "NGS"
- Different bioinformatic pipelines
- Different annotation and interpretation schemes for identified variants
- Different assay validation schemes

Conceptually:

- NGS as a clinical test
- NGS as a method



Zivana's introduction: Analytical performance

Birgit's presentation: THERE IS NO "ONE SIZE FITS ALL"

Challenges for NGS tests

- Amount of data / bioinformatics needs
- Almost unlimited number of variants could be detected
- Rapidly evolving technology
- Intended use often not easily defined
- Different clinical applications

THERE IS NO "ONE SIZE FITS ALL"

Two difficult to sequence genes – two solutions (Laboratory for Molecular Medicine)

Both present similar challenge	Hearing Loss Panel (STRC)	Renal Panel (PKD1)
Essential gene (detection rate)	YES (10%)	YES (85%)
Sequenceable via NGS	Highly homologous (entire gene)	Highly homologous (partial)
Pathogenic variant spectrum	CNVs >>SNVs, indels	SNVs, indels >>CNVs
Well sequenced gene	NO	YES
Clinical specificity	LOW	HIGH
Keep on panel	YES	NO - alternate test for ADPKD
Supplemental assay	LR-PCR x 2	N/A
Operational factors	High failure LR-PCR #1 – DROP	N/A
Impact on performance metric	Reduced analytical sensitivity - OK	N/A

Two very different solutions

Implications

Clinical NGS has three independent components (test design, platform, and bioinformatic pipeline) that create novel issues around test validation and lab accreditation

In addition to the customary "process based" approaches, there is a need for "results based" or "standards based" approaches that can provide objective measures of:

- Test metrics (sensitivity, specificity, and so on) that permit comparison of tests across design, platform, and bioinformatics subsumed by the phrase "clinical NGS"
- Standards that account for the databases used for annotation and interpretation as well as the analytic test components

To accomplish this, there will need to be (at least):

- Agreement on what constitutes the "gold standard"
- Understanding that "clinical NGS" is rapidly evolving, and so whatever the initial guidelines/metrics/requirements, they are just a starting point

“Process based” versus “Standards based”

"Standards based" approaches will require novel test validation *and* laboratory accreditation models, including…

Patient samples

- limited supply
- not comprehensive

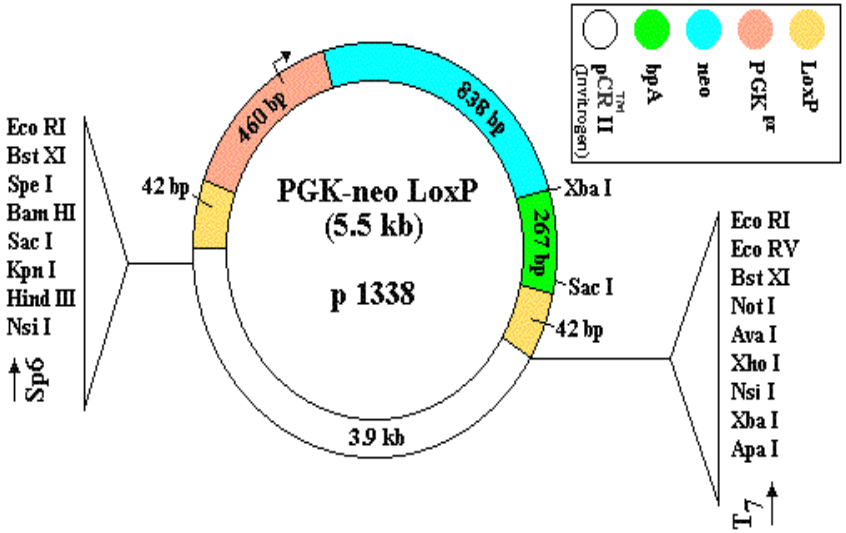
Cell lines

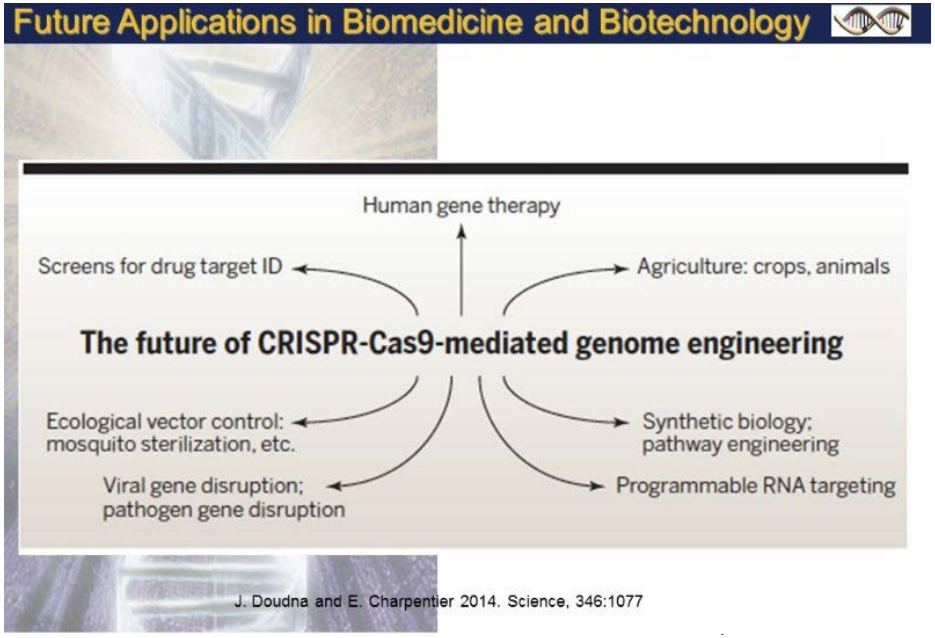
- complex mixtures possible
- wide variety of variants, but not comprehensive
- inexhaustible
- many commercial and NFP sources

… as well as, all of which can challenge the full NGS assay from soup to nuts

Engineered constructs and cell lines

- complex mixtures possible
- wide variety of variants, but still not comprehensive
- Inexhaustible after production
- many commercial sources
- expensive to produce
- artificial
- hard to model different VAFs

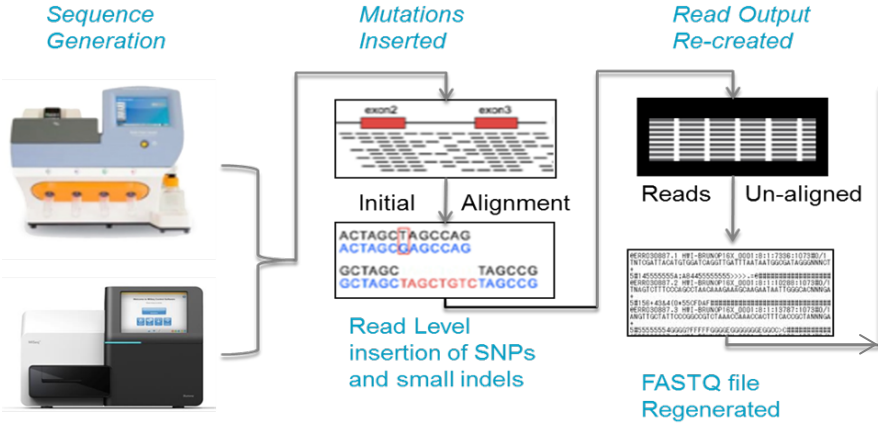




Another "standards based" approach for assessing analytical validity is via "in silico datasets"

Sequence files that have undergone in silico mutagenesis (i.e., sequence files from NGS of a well characterized specimen that have been manipulated by computerized algorithms to introduce relevant sequence variants into the reference sequence files) have many advantages:

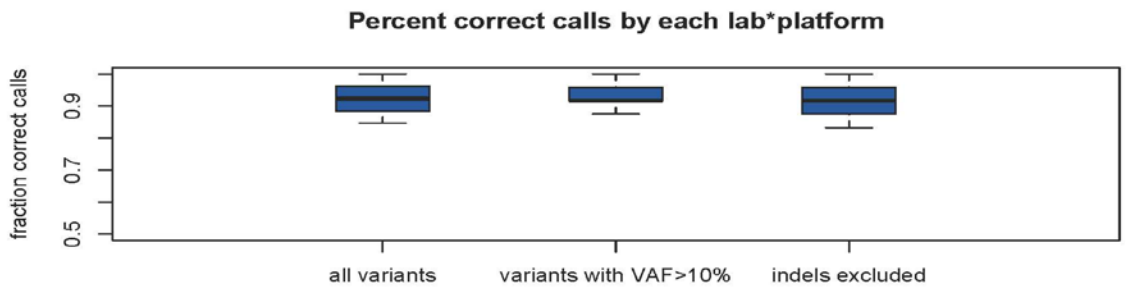
- mixtures of variants and VAFs characteristic of inherited diseases and cancer
- inexpensive, comprehensive, current
- challenge an NGS test's bioinformatic pipeline from alignment through variant detection, annotation, and interpretation (but therefore supplement but do not replace traditional methods)
- broad applicability
  - constitutional and/or somatic variants, nuclear and/or mitochondrial variants
  - can be extended to analysis of cfDNA, RNA, methylation, infectious organisms, the microbiome, etc.
  - can be used with vendor/institutional/regulatory/pharma files for "ISVT"

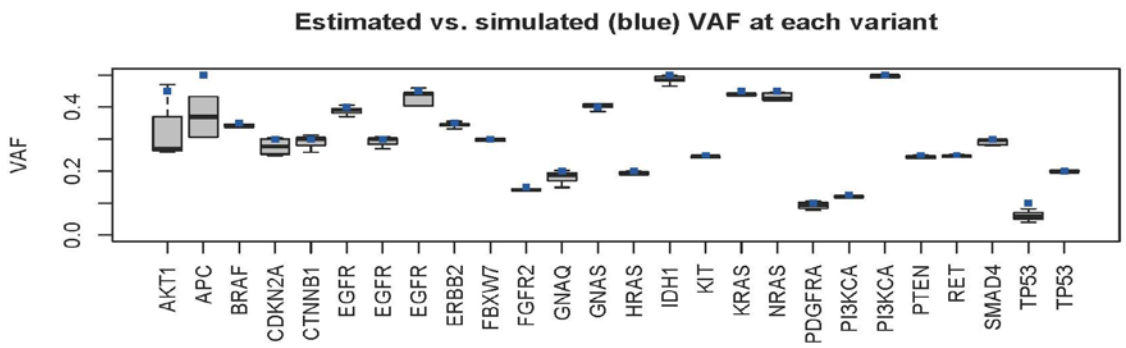


References: Frampton M et al. *PLoS One* 2012;7: e49110.
Schrijver I et al. *J Mol Diagn* 2014;16:283-7.
Kalman LV et al. *Arch Pathol Lab Med* 2013;137:983-8.

In silico methods can model variants as well as VAFs

- CAP sponsored demonstration project showed the utility of in silico proficiency testing employing platforms and amplification-based NGS kits in widespread clinical use
- Variants introduced included SNPs and small indels, modeled from actual somatic mutations reported in the COSMIC database





Reference: Duncavage et al. *J Mol Diagn* 2015;17:797

Summary

- The unique aspects of massively parallel sequencing create novel challenges for test validation *and* laboratory accreditation
- The range of test designs, sequencing platforms, and bioinformatics adds to the challenge
- It's time to consider standards based approaches in addition to process based approaches
- Novel reagents, including engineered cell lines and in silico datasets, can be used to create mixtures of variants that mimic the complexity of clinical samples to support standards based approaches to analytical validity

Panel Discussion of Topic 1
Standards-based approaches to analytical validation – A Spectrum from Design Concept Standards to Performance Standards

Moderator: Zivana Tezak, PhD

Speakers and Panelists

- **Birgit Funke, PhD (Harvard Partners Healthcare)**
- **John Pfeifer, MD, PhD (Washington University School of Medicine)**
- **Jared Maguire, PhD (Counsyl)**
- **Geoff Otto, PhD (Foundation Medicine)**
- **Arend Sidow, PhD (Stanford University)**
- **Erasmus Schneider, PhD (New York State Dept of Health)**

BREAK

10:45 – 11:00

Welcome to Workshop

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015

Panel Discussion of Topic 2
Developing analytical standards for NGS-based assays

Moderator: Adam Berger, PhD

Speakers and Panelists

- **Deanna Church, PhD (Personalis)**
- **Girish Putcha, MD, PhD (Palmetto/MolDX)**
- **Gil Alterovitz, PhD (MIT/Global Alliance for Genomics and Health)**
- **Lisa Kalman, PhD (CDC)**
- **Catherine Rehder, PhD (Duke/ACMG)**
- **Karl Voelkerding, MD (ARUP, College of American Pathologists)**

Girish Putcha, MD, PhD

Palmetto / MolDX

MolDX®

PALMETTO GBA®

A CELERIAN GROUP COMPANY

Public Workshop - Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

12 November 2015

Girish Putcha, MD, PhD
Director of Laboratory Science

Please note that the opinions expressed herein are my own.

Why Analytical Standards?

- Explosion of NGS-based tests for many applications
  - Inherited diseases, oncology, infectious disease, etc.
  - Germline and somatic disorders
  - Tissue-based and cell free DNA (cfDNA) testing
  - Tumor only versus matched tumor-normal testing
  - Single cell analysis (e.g., CTCs, CTSCs, etc.)
  - “Hotspot”, whole exome, and whole genome panels
- Potential for direct, high impact on patient care
  - Diagnostic, prognostic, and predictive
  - Treatment selection (e.g., pharmacological, surgical, etc.)
  - Patient monitoring/surveillance protocols

Why Analytical Standards?

- Analytical performance characteristics (e.g., accuracy, precision, and limits of detection) may differ for:
  - Different classes of variants (SNVs, insertions, deletions, CNVs, and SVs)
  - Different variant allele frequencies (VAFs) or tumor content
  - Different "sizes" (e.g., insertions ≤10 bp vs >10 bp; CNV ≤8 vs >8, etc.)
  - Different sequence contexts (e.g., GS-rich regions)
  - Different specimen types (e.g., FFPE, whole blood, etc.)
  - Different technology platforms and associated sequencing chemistries
- Such differences can impact patient care.

Why Analytical Standards?

- To my knowledge, existing standard development efforts focus on <u>process</u>, not on <u>product</u>
  - Former is necessary but not sufficient for the latter (i.e., a "good" test)
  - Does the "product" includes clinical interpretation?
- Absence of clinically driven analytical requirements?
  - If clinical requirements are unknown, then analytical criteria based on what is <u>reasonably</u> achievable may be a start.
  - For example,
    - For EGFR SNVs and indels, is there a minimum VAF below which drug response is not seen? Does response vary with VAF?
    - For HER2 CNVs, is there a minimum amplification level below which drug response is not seen? Does response vary with amp level?

Overview of NGS-based Oncology Testing Landscape: Past, Present, and Future?

Develop a framework that contemplates not just where we are today but also where we are probably going (e.g., ctDNA, single cell CTC and/or CTSC analysis, etc.)

Variant Type	Germline-only Reporting	Tumor-only Testing, Somatic-only Reporting	Matched Tumor-Normal Testing, Somatic-only Reporting	Matched Tumor-Normal Testing, Somatic and Germline Reporting
Single nucleotide variants	Inherited Disease Panels	"Hotspot" Panels	?	?
Insertions			?	?
Deletions			?	?
Copy number alterations – Amplifications	?	Comprehensive Genomic Panels	?	?
Copy number alterations – Homozygous Deletions	?		?	?
Translocations	?		?	?

MolDX Analytical Standards for Comprehensive Genomic Panels (CGP) for Metastatic NSCLC

- Hybrid of process and performance standards
- Process standards
  - Address the pre-analytical, analytical and post-analytical phases of testing, including clinical interpretation
  - Attempt to leverage existing standards
    - General requirements include CAP accreditation or final test approval from NYDOH
    - Both have NGS-specific procedural requirements (e.g., for validation, bioinformatics, etc.)
    - Not “design concept” standards but process standards
    - Only standards with NGS-specific requirements AND independent, periodic 3rd party review and/or inspection to ensure compliance
  - Participation in external PT program for NGS-based testing that addresses all reported variant classes, if available.

MolDX Analytical Standards for Comprehensive Genomic Panels (CGP) for Metastatic NSCLC

- Analytical requirements
  - Using SOPs, equipment, reagents, and personnel involved in routine clinical testing (i.e., not R&D)
  - Accuracy and precision of sequencing on reference cell line mixes (with minimum specifications)
  - Accuracy of sequencing on clinical samples
    - No minimum specifications, only reported
  - Other applicable validation elements reported with no minimum specifications. For example:
    - Analytical specificity (e.g., interfering substances)
    - Sample stability (e.g., shipping, freeze thaw)
    - Reagent stability (e.g., closed/shelf-life, open/in-use)
    - Indication(s) for use and intended use population(s)

MolDX Analytical Standards for Comprehensive Genomic Panels (CGP) for Metastatic NSCLC

- Under consideration for future revisions
  - Addition of “dry lab samples” (e.g., BAM or other “raw” data files)
    - “Stress test” bioinformatic pipelines producing variant calls (and potentially their interpretation)
    - Perhaps decrease need (and cost) for “representative” reference materials addressing the aforementioned sources of variability in analytical performance (e.g., VAF, indel size, sequence context, specimen types, etc)
  - Inclusion of commercially available reference materials to assess effects of aforementioned variables
  - More specific attention to the impact of sequence context
- Comments and feedback always welcome

MolDX Analytical Standards for Comprehensive Genomic Panels (CGP) for Metastatic NSCLC

Row	Variant Type	Detail	Lower 95% Confidence Interval[a]		
			PPA[b]	PPV[b]	Intermediate Precision[c]
1	Single nucleotide variants	Expected VAF > 10%	≥ 99.0%	≥ 99.0%	≥ 95.0%
2	Single nucleotide variants	Expected VAF = 5-10%	≥ 95.0%	≥ 95.0%	≥ 90.0%
3	Insertions	≤ 10 bp Expected VAF > 20%	≥ 85.0%	≥ 95.0%	≥ 80.0%
4	Insertions	≤ 10 bp Expected VAF = 10-20%	≥ 75.0%	≥ 85.0%	≥ 75.0%
5	Insertions	11-70 bp Expected VAF > 20%	≥ 85.0%	≥ 95.0%	≥ 80.0%
6	Insertions	11-70 bp Expected VAF = 10-20%	≥ 75.0%	≥ 85.0%	≥ 75.0%
7	Deletions	≤ 10 bp Expected VAF > 20%	≥ 85.0%	≥ 95.0%	≥ 80.0%
8	Deletions	≤ 10 bp Expected VAF = 10-20%	≥ 75.0%	≥ 85.0%	≥ 75.0%
9	Deletions	11-70 bp Expected VAF > 20%	≥ 85.0%	≥ 95.0%	≥ 80.0%
10	Deletions	11-70 bp Expected VAF = 10-20%	≥ 75.0%	≥ 85.0%	≥ 75.0%

MolDX Analytical Standards for Comprehensive Genomic Panels (CGP) for Metastatic NSCLC

Row	Variant Type	Detail	Lower 95% Confidence Interval[a]		
			PPA[b]	PPV[b]	Intermediate Precision[c]
12	Copy number alterations – Amplifications	Ploidy < 4 Expected CN ≥ 8 20-30% tumor nuclei	≥ 60.0%	≥ 80.0%	≥ 75.0%
13	Copy number alterations – Homozygous Deletions	Ploidy < 4 Expected CN = 0 > 30% tumor nuclei	≥ 80.0%	≥ 85.0%	≥ 60.0%
14	Copy number alterations – Homozygous Deletions	Ploidy < 4 Expected CN = 0 20-30% tumor nuclei	≥ 50.0%	≥ 75.0%	≥ 50.0%
15	Translocations	≥ 20% tumor nuclei	≥ 85.0%	≥ 85.0%	≥ 90.0%

[a] For calculating 95% CIs, use Score or Clopper-Pearson method as described in CLSI EP12-A.

[b]

		Orthogonal Reference Results		
		Positive	Negative	Total
NGS Results	Positive	A	B	A+B
	Negative	C	D	C+D
	Total	A+C	B+D	A+B+C+D

"Positive" = Variant called/identified when compared to current build of human genome (e.g., hg19 human genome assembly, version 37)

Positive percent agreement (PPA) = A/(A+C); Positive predictive value (PPV) = A/(A+B)

[c] Based on at least 3 different operators, runs/batches/days, manufacturing reagent lots (for critical reagents), and (if applicable) sequencers and sites. Agreement is defined as concordance at the variant level for all called (i.e., not necessarily reported but excluding no calls) variants. In other words, did each called variant agree across different operators, runs/batches/days, manufacturing reagent lots, and (if applicable) sequencers and sites?

Lessons Learned & Open Questions

- Hybrid of design concept and performance standards recommended
- For performance standards, the Agency should . . .
  - Specify
    - What metrics (e.g., PPA, PPV, etc.)
    - How these metrics are calculated
    - Any associated statistics required
    - Any other details relevant to understanding and comparing analytical performance
  - If minimum performance criteria are specified,
    - What “real world” data to include?
    - Should these be agnostic to technology platform and associated sequencing chemistries?

Final Thoughts

- At a minimum, regulation should . . .
  - Facilitate transparency of actual performance characteristics
  - Allow “apples to apples” comparisons for “consumers” (e.g., patients, providers, payers, etc.)

- Ideally, regulation would also . . .
  - Clarify indication(s) for use and intended use population(s) – i.e., who should be tested, when and why?

Thank you.

Questions and comments are welcome.

girish.putcha@palmettogba.com

MolDX®

Deanna Church, PhD

Personalis

Analytical performance standards

Deanna M. Church, PhD
Senior Director of Genomics and Content
Personalis, Inc

Public Workshop - Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

Nov 12, 2015

Pioneering Genome-Guided Medicine

© 2014 Personalis, Inc. All rights reserved.

Personalis®

Conflict of Interest Disclosure

In relation to this presentation, I declare the following real or perceived conflict of interest:

Deanna Church is an employee of Personalis, Inc., and as such receives salary from, and is a shareholder in, Personalis, Inc.

Personalis, Inc. is a for-profit provider of genomic sequencing and interpretation services for researchers and clinicians. Personalis' services include an enhanced clinical exome sequencing test: **The ACE Clinical Exome™ Test.**

Personalis

Acknowledgements

- Sarah Garcia
- Jennifer Yen
- Jason Harris
- Stephen Chervitz
- Aldrin Montana
- Brian Linebaugh
- Shujun Luo
- Gabor Bartha
- Elena Helman
- Sean Boyle
- Ravi Alla
- Michael Clark
- Martina Lefterova
- Massimo Morra
- Gemma Chandratillake
- John West
- Richard Chen

Personalis

Cancer clinical validation study

© 2014 Personalis, Inc. All rights reserved.

Personalis®

DNA Validation– Sensitivity and PPV

28 cancer cell lines with known cancer variants

- Melanoma
- Epidermoid
- Renal
- Lung
- Breast Ductal
- Colorectal
- Colorectal
- Colon
- Leukemia
- Prostate
- Breast
- AML
- NSC Lung
- NSC Lung
- Ovarian
- Prostate
- Melanoma
- Melanoma
- Breast Ductal
- Glioblastoma
- Melanoma

Sequence all cell lines to establish base-line sensitivity and positive predictive value.

These cell lines represent a diverse set of tumor types.

Together, they contain:

- 817 known cancer SNVs
- 62 known cancer indels
- 21 known deletions
- 23 known amplifications

Personalis

DNA Validation– Sensitivity and PPV at LOD

21 cancer cell lines with known cancer variants

- Melanoma
- Epidermoid
- Renal
- Lung
- Breast Ductal
- Colorectal
- Colorectal
- Colon
- Leukemia
- Prostate
- Breast
- AML
- NSC Lung
- NSC Lung
- Ovarian
- Prostate
- Melanoma
- Melanoma
- Breast Ductal
- Glioblastoma
- Melanoma

Mix cell lines in equal ratios to achieve particular allelic representations.

≤5% allelic representation
All 20 cell lines mixed at equal ratios

5-10% allelic representation
10 cell lines mixed at equal ratios (two pools)

5-95% allelic representation
2 cell lines mixed at unequal ratios (five pools)

10-20% allelic representation
5 cell lines mixed at equal ratios (four pools)

25-50% allelic representation
2 cell lines mixed at equal ratios (three pools)

Sequence all cell lines to establish base-line sensitivity and positive predictive value.

Sequence all pools to determine sensitivity and positive predictive value at particular allele frequencies.

Personalis

RNA Validation

30 cancer cell lines with known cancer fusions

- *EML4-ALK*
- *RUNX1-RUNX1T1*
- *MLL-MLLT3*
- *PAX3-FOXO1*
- *EWSR1-FLI1*
- *SLC34A2-ROS1*
- *TPM3-NTRK1*
- *CCDC6-RET*
- *BCR-ABL1*
- Various
- Various
- *EGFR-*
- *PPARGC1A*
- Various
- Various
- Various
- Various
- Various
- Various
- Various
- NPM1-ALK
- PICALM-MLLT10
- SL34A2-ROS1
- Various
- Various
- Various
- Various
- Various
- Various
- Various
- Various

Sequence all cell lines to establish base-line sensitivity and positive predictive value for fusions.

These cell lines contain a number of known cancer-relevant gene fusions.

These include:

- *ALK* fusions
- *BRC-ABL1*
- *ROS1* fusions
- *NTRK1* fusion

ACE Cancer Panel CLIA Validation Results

ACE Cancer Panel Performance Specifications		
Sensitivity	Base Substitutions	>99% MAF ≥ 5%
	Indels	>99% MAF ≥ 10%
	CNAs	97% tumor content ≥ 30%
	Gene Fusions	>99%
Specificity	>99%*	
Typical Median Depth	>500X	
Sample Types	Fresh Frozen or FFPE Tumor Samples ≥ 30% Tumor	

* Based on Base substitutions and Indels, others pending

Performance metrics not directly comparable to other tests due to independent validation sets

Personalis

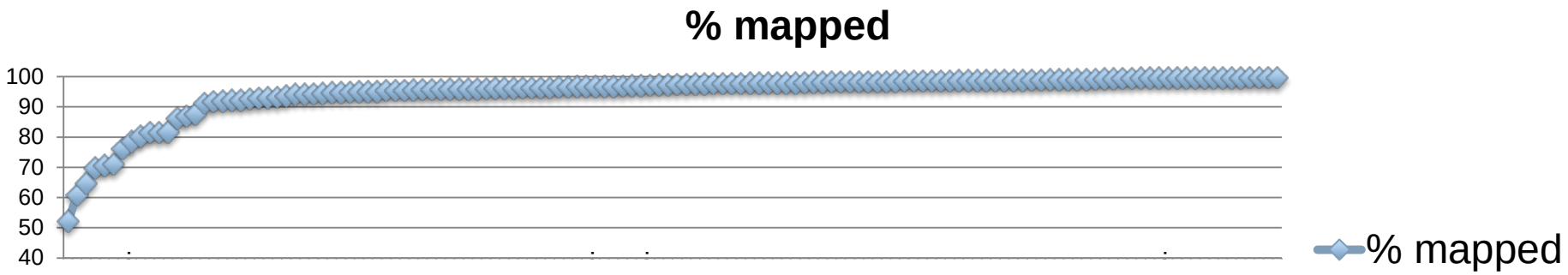
Real life samples

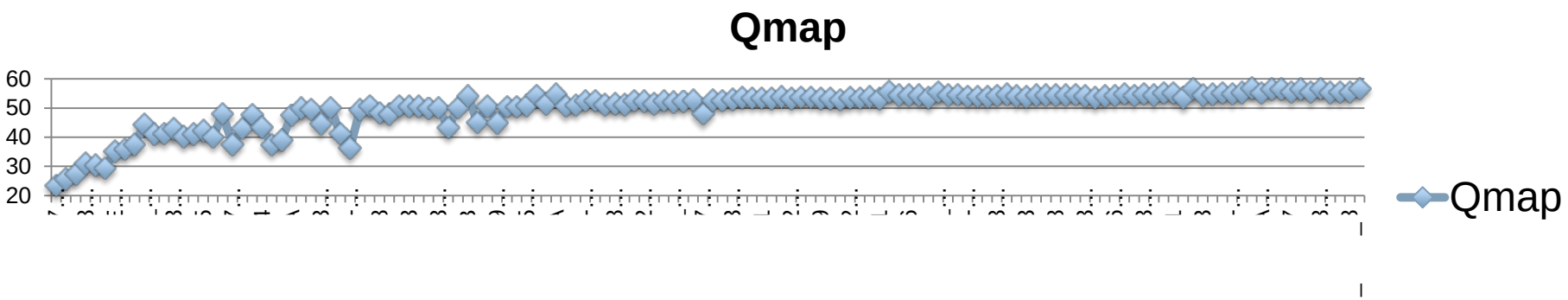
© 2014 Personalis, Inc. All rights reserved.

Personalis®

Real FFPE sample variability

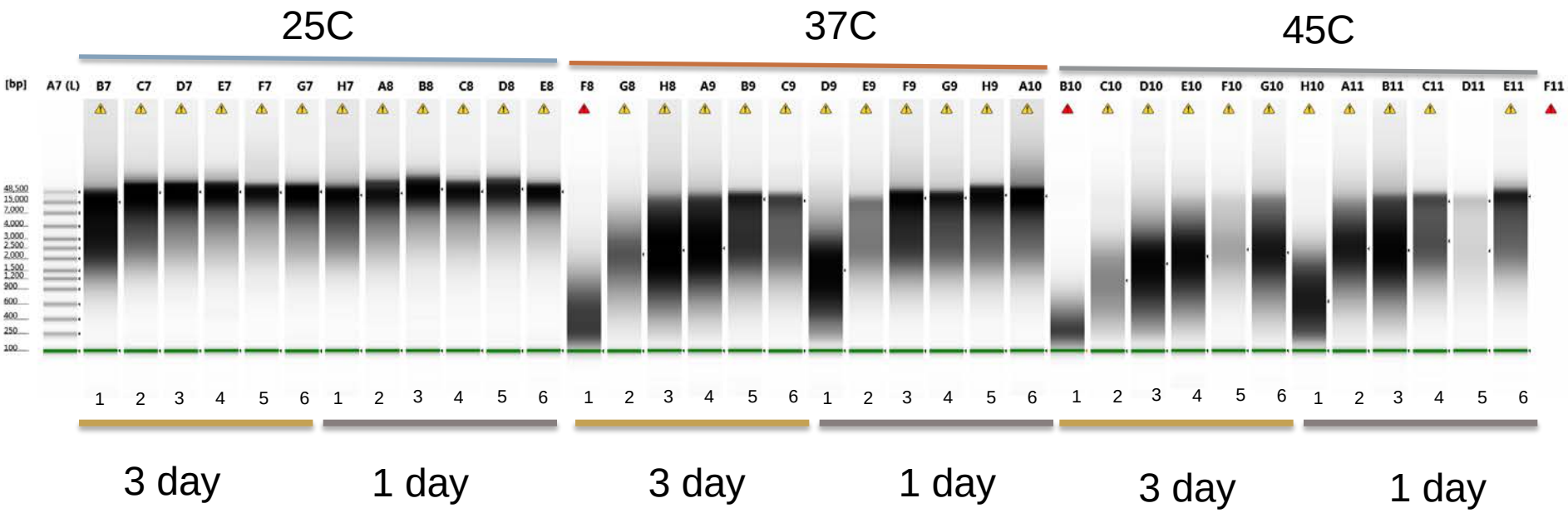
Large range of performance for FFPE samples





Clinical Samples: quality may vary

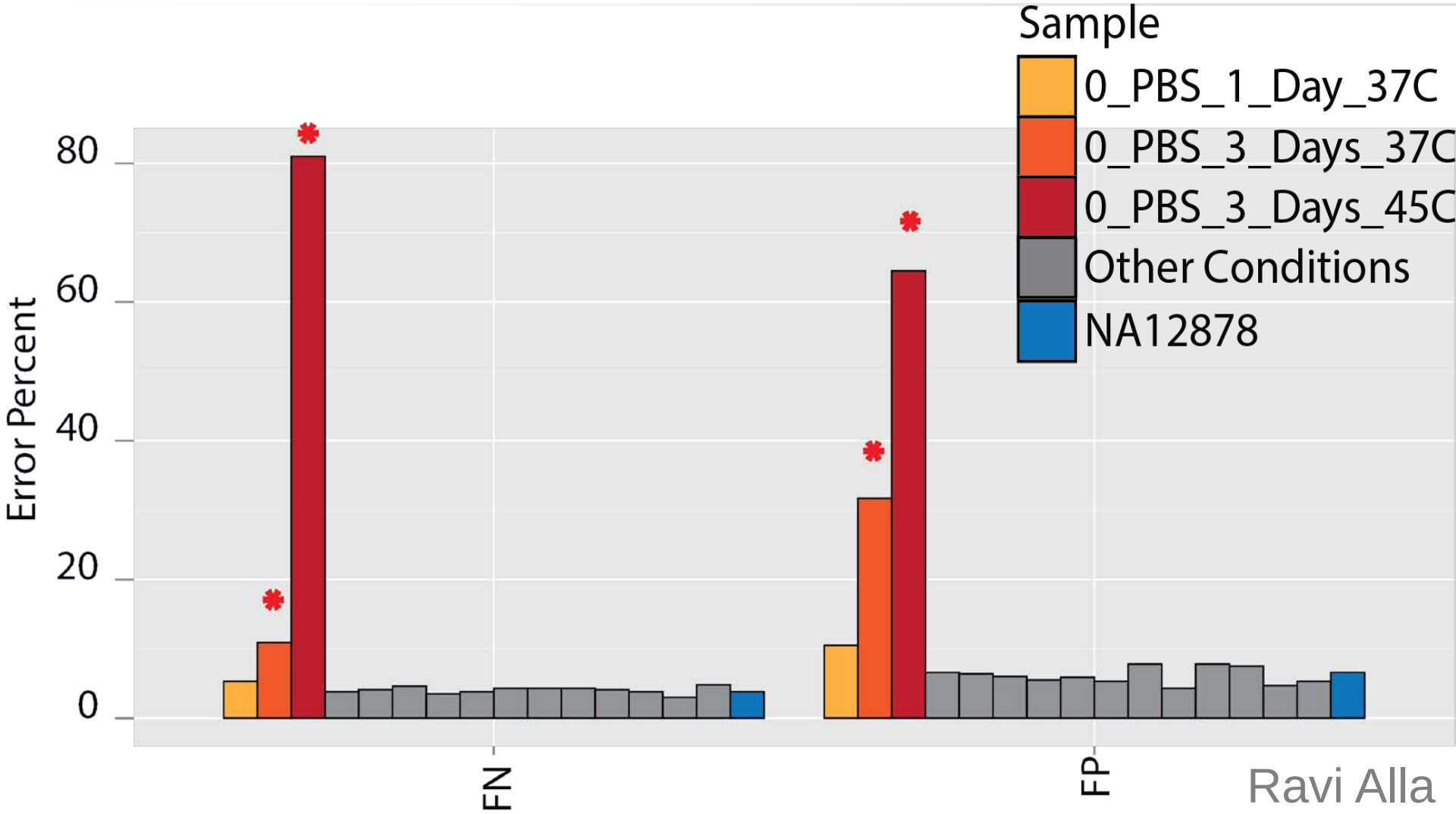
DNA fragmentation assay



1. 0% PBS
2. 20% PBS
3. 40% PBS
4. 60% PBS
5. 80% PBS
6. 100% PBS

Personalis

Clinical Samples: quality may vary



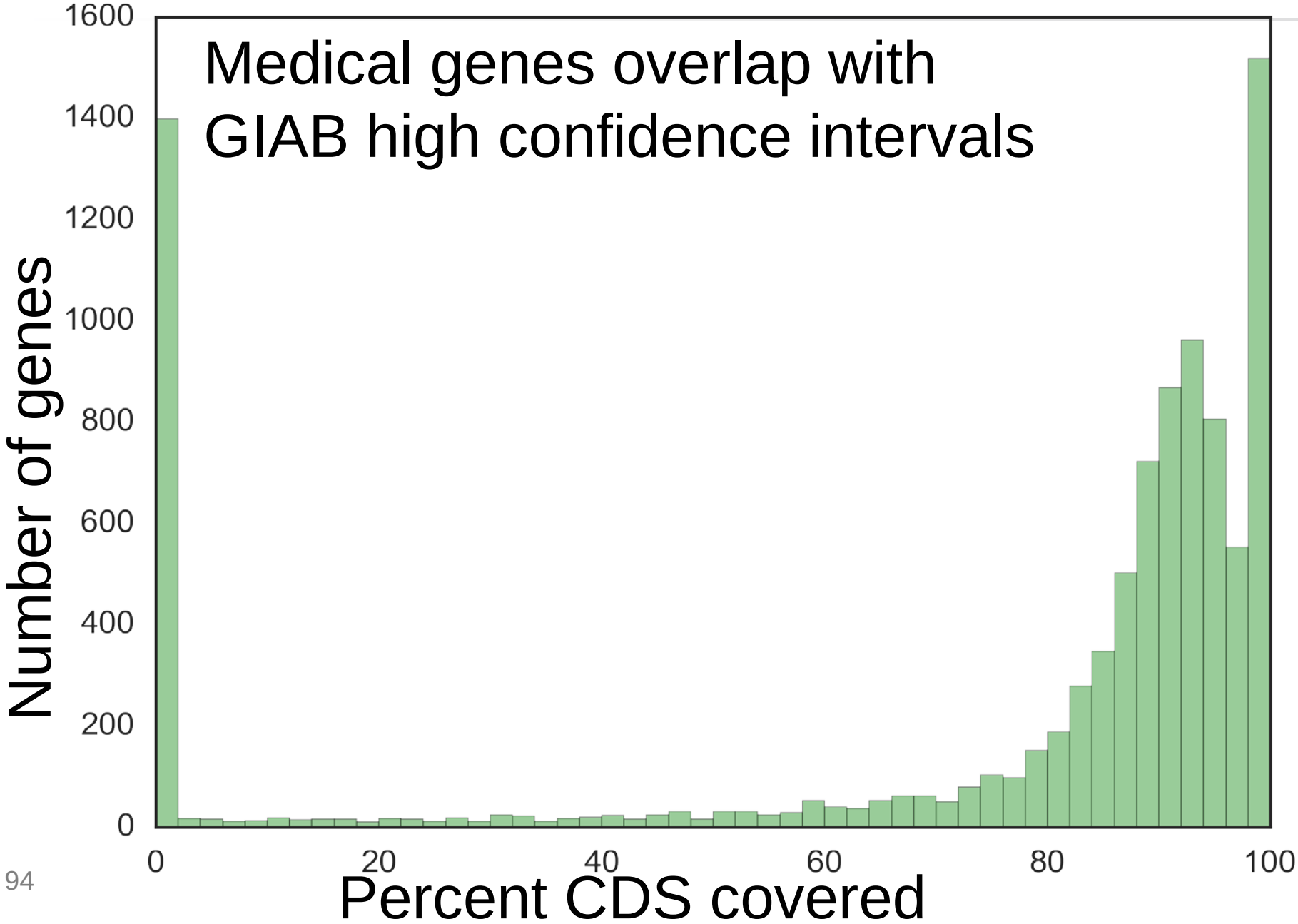
Ravi Alla

Benchmarking sets

© 2014 Personalis, Inc. All rights reserved.

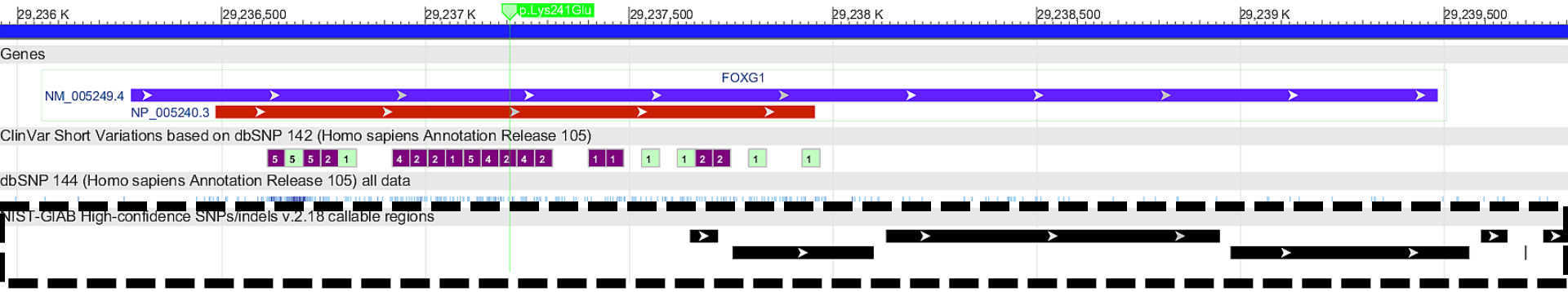
Personalis®

Medical genes CDS coverage by GIAB



FOXG1

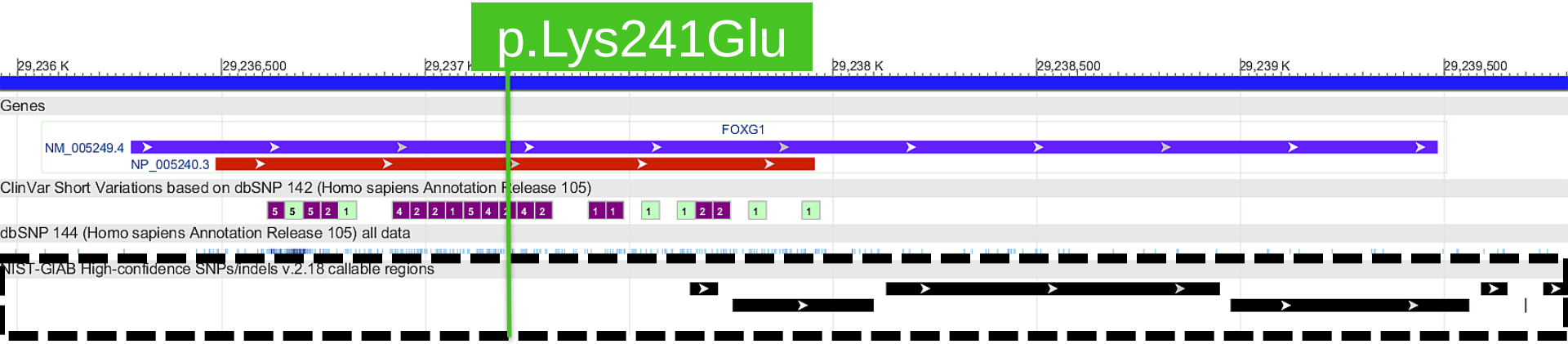
Rett Syndrome



NIST High confidence regions (v2.18)

FOXG1

Rett Syndrome

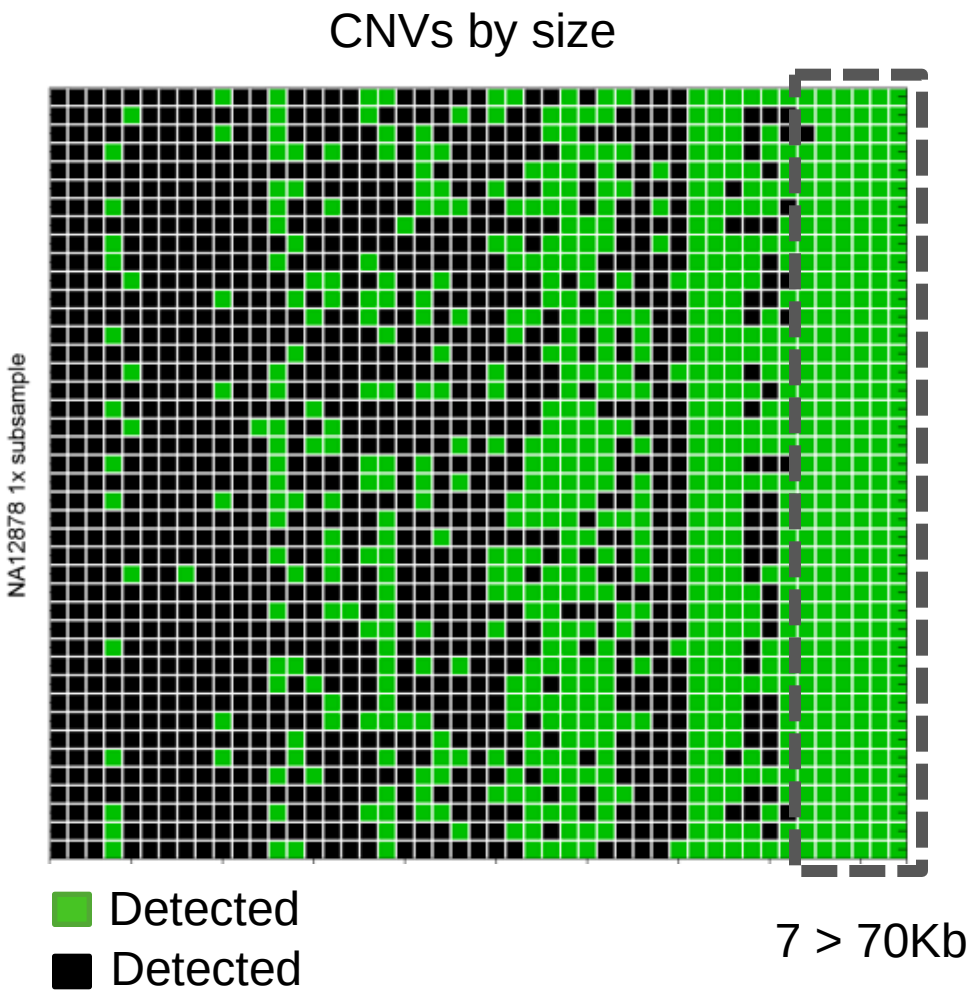


NIST High confidence regions (v2.18)

Variant reported in our lab

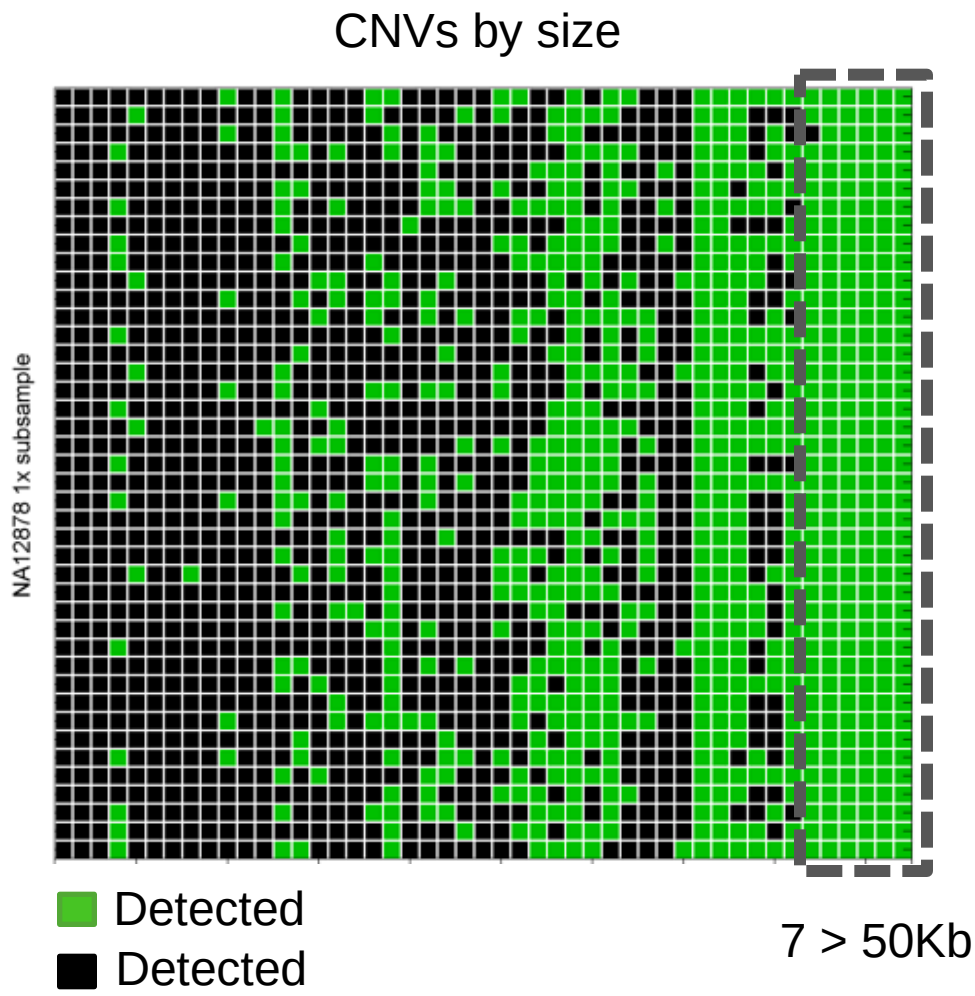
CNV detection needs

NA12878: 47 high quality CNVs >10Kb



CNV detection

NA12878: 47 high quality CNVs >10Kb



Sequence another 28 Samples to get another 48 CNV >50Kb

Knowing the unknowns

- **Common benchmarks**
  - Improved transparency
  - Easier comparison across labs

Personalis, Inc. | Confidential and Proprietary

Knowing the unknowns

- **Common benchmarks**
  - Improved transparency
  - Easier comparison across labs
- **Reference samples reflective of real life**

Personalis, Inc. | Confidential and Proprietary

Knowing the unknowns

- **Common benchmarks**
  - Improved transparency
  - Easier comparison across labs
- **Reference samples reflective of real life**
- **Improved ‘truth’ sets**
  - Cover more of the genome
  - Cover hard parts of genomes
  - Cover more variant types

Personalis, Inc. | Confidential and Proprietary

Panel Discussion of Topic 2
Developing analytical standards for NGS-based assays

Moderator: Adam Berger, PhD

Speakers and Panelists

- **Deanna Church, PhD (Personalis)**
- **Girish Putcha, MD, PhD (Palmetto/MolDX)**
- **Gil Alterovitz, PhD (MIT/Global Alliance for Genomics and Health)**
- **Lisa Kalman, PhD (CDC)**
- **Catherine Rehder, PhD (Duke/ACMG)**
- **Karl Voelkerding, MD (ARUP, College of American Pathologists)**

LUNCH BREAK

12:30 – 1:30

Welcome to Workshop

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015

Panel Discussion of Topic 3
Developing bioinformatics strategies and tools for evaluating NGS tests

Moderator: Sharon Liang, PhD

Speakers and Panelists

- **Kevin Jacobs, PhD (23andMe/Global Alliance for Genomics and Health)**
- **Narayanan Veeraraghavan, PhD (Baylor College of Medicine)**
- **Sean Davis, PhD (NCI)**
- **Tina Hambuch, PhD (Illumina)**
- **Sean Hofherr, PhD (Children's National Medical Center)**
- **Niall Lennon, PhD (Broad Institute)**

Kevin Jacobs, PhD

23andMe

Global Alliance for Genomics and Health

GIAB

Bioinformatic strategies for evaluating NGS tests

Kevin Jacobs

11/12/2015

Presenting for the Genome in a Bottle Consortium (GIAB) &

Global Alliance for Genomics and Health (GA4GH) Benchmarking Group

Goals when evaluating an NGS test

- Validate reagents and equipment
- Validate components
  - Includes software, processes, analysis
  - Each component should have a validation plan
- Validation studies to evaluate the test
  - Ensure all components, reagents and equipment function properly when integrated together
  - Measure sensitivity, specificity, stability, limits of detection, interfering substances, etc.
  - Analytical claims may differ by variant type and genomic context

NGS software components

- NGS analysis software
  - Primary: sequence generation and basecalling analysis
  - Secondary: Assembly/alignment of reads and variant calling
  - Tertiary: Annotation and interpretation of variant calls
- LIMS: Laboratory information management system
- Laboratory instrument automation software
- Laboratory data analysis software
  - Spreadsheets for sample and reagent dilution
  - qPCR analysis programs for library quantification
  - etc.

Characteristics of most bioinformatic software

- Written by smart people, but most often
  - not developing software within a regulated environment
  - not familiar with the requirements for developing software within a regulated environment
  - not able to commit to long-term maintenance and support of their software
  - more focused on submitting a paper or graduating
- Software implement algorithms and models that are extremely complex
  - often have dozens of parameters or hard-coded values with apropriate values that are depend on the sample type, laboratory process, equipment used, assay parameters, etc.
  - apply heuristics that may only be valid under certain conditions
  - are intended to be used only by experts and then with careful validation
- Even larger and more mature software efforts require substantial effort to validate
  - even when using off-the-shelf or "best practices" pipelines
- For guidance, see "General Principles of Software Validation; Final Guidance for Industry and FDA Staff" , issued January 11, 2002

Concepts for effective software validation

- Determinisim
  - Give the same input, does the software produce the same output every time it is run?
  - A suprising amount of commonly used software is not deterministic, but can be corrected
- Versioning
  - Can you identify when any software component or data file has been updated?
- Deep testing
  - When a component has changed, do you have an automated testing mechanism to validate those changes locally and the effect of those changes on your entire process?
- Continuous integration
  - Can you make software testing automatic and incremental, so that new failures can be detected quickly and associated with the precise change that caused them?
- Third party qualification and validation
  - Can you also implement these kinds of testing procedures for third-party software?

Example NGS testing methodology

- Ideally, a clinical NGS lab would like to be able to re-run the data analysis for every sample every processed to ensure that no results un-intentionally changed due to a software or data change.
  - In theory, this sort of testing would detect any differences.
  - In practice, NGS data are large, analysis is time consuming and most differences are “noise”.
  - As a result, this kind of testing is almost never done.

- Alternatively, save the sequencing reads in the region of all reportable findings and re-analyze those after every software or data change.
  - Much more computationally tractible.
  - Can detect every instance where a reportable finding changes.

Analytical benchmarking of NGS tests

- Work done within
  - Genome In a Bottle Consortium
  - Global Alliance for Genomics and Health Data Working Group

- Being implemented in precisionFDA prototype

- In order to measure the analytical accuracy of an NGS assay
  - Compare sequencing results
  - Evaluate the equivalence of variant calls
  - Assign true/false positives/negatives to each call
  - Aggregate concordance statistics across variant types and genomic contexts
  - Produce meaningful summary statistics and visualizations

A seemingly trivial problem

- Do two descriptions of a variant represent the same nucleic acid change?
- Answering this question correctly is at the heart of other important questions
  - How does my VCF file differ from a reference material's VCF file?
  - Is a specific variant present in an individual's genome?
  - Is a given HGVS variant present in an individual's VCF file?
  - Does this individual's VCF file contain any pathogenic or likely-pathogenic variants in ClinVar?

HGVS nomenclature ambiguity

- Question: Are two HGVS variants equivalent?

- Transcript mapping
  - NM_001166478.1:c.35_36insT — NM_001166478.1:c.35dupT
  - NM_001166478.1:c.35_36dup — NM_001166478.1:c.36_37dupCT
  - NM_001166478.1:c.59delG — NM_001166478.1:c.60delG
  - NM_001166478.1:c.2_7delinsTTTAGA — NM_001166478.1:c.3_4delGAinsTT

  These equivalences can usually be obtained via normalization rules

- Transcript → genome, transcript → transcript mapping
  - NM_001166478.1:c.59delG — NC_000006.11:g.49917099delC
  - NM_001166478.1:c.59delG — NC_000006.11:g.49917098delC
  - NM_001166478.1:c.59delG — NM_000014775.4:c.201delG

  These equivalences require standardized alignment information

HGVS nomenclature ambiguity

- Consider c.2_4delCTAinsGC
  - REF: ACTAC
  - Allele: =G-C=

- This variant can also be spelled
  - c.[2C>G; 3del; 4A>C]
  - c.[2C>G; 3T>C; 4del]
  - c.[2del; 3T>G; 4A>C]
  - …

Implications of imprecise variant comparison

- When benchmarking
  - Excess false positive and negative call rates
- When comparing observed variants to a reference database
  - Excess false negative calls, resulting in known variants being treated as novel
- False call rates will be inflated most for complex variants, those most likely to be functionally relevant

Community efforts

- GIAB, GA4GH and many commercial NGS lab and vendors are working together on solutions
  - HGVS variant normalization and translation
    - uta & hgvs projects led by Reece Hart and supported by Invitae
  - NGS benchmarking software
    - vgraph: https://github.com/bioinformed/vgraph
    - hap.py developed by Illumina: https://github.com/Illumina/hap.py
    - vcf-eval developed by Real Time Genomics: https://github.com/RealTimeGenomics/rtg-tools
  - Test suites for variant comparison
    - Community curated consensus on complex cases
  - Refinements to NGS standards
    - Proposals to tighten semantics for VCF files

Narayanan Veeraraghavan, PhD

Human Genome Sequencing Center

Baylor College of Medicine

HGSC
HUMAN GENOME SEQUENCING CENTER

Baylor Miraca
Genetics Laboratories

Baylor College of Medicine
GIVING LIFE TO POSSIBLE

FDA Public Workshop

Panel 3: Developing bioinformatics strategies and tools for evaluating NGS tests

Narayanan Veeraraghavan, PhD.
Human Genome Sequencing Center
Baylor College of Medicine

Disclaimer

- All views expressed are mine only and does not reflect the views and opinion of my employers.

HGSC
HUMAN GENOME SEQUENCING CENTER

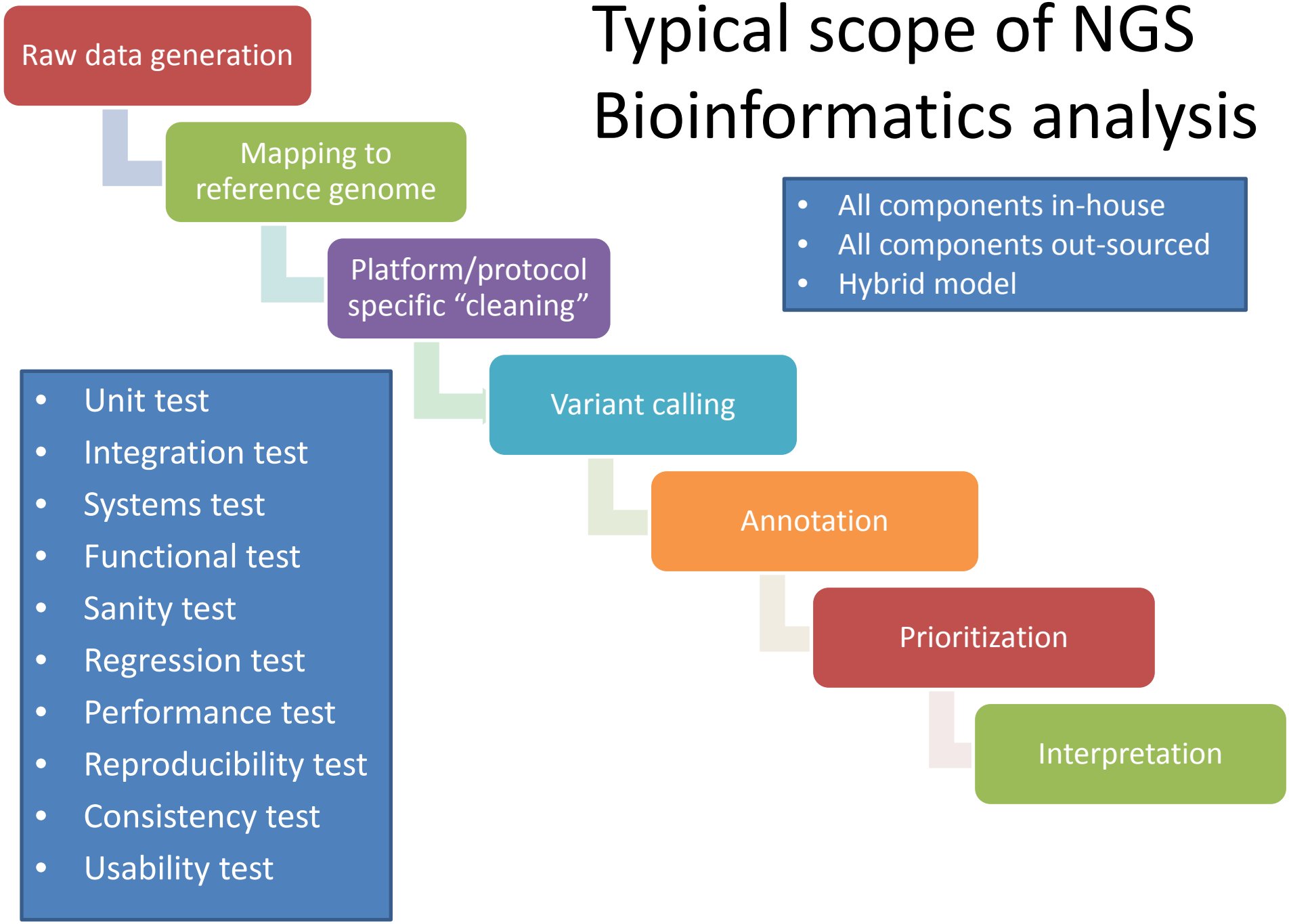
Baylor Miraca
Genetics Laboratories

Baylor
College of
Medicine
GIVING LIFE TO POSSIBLE

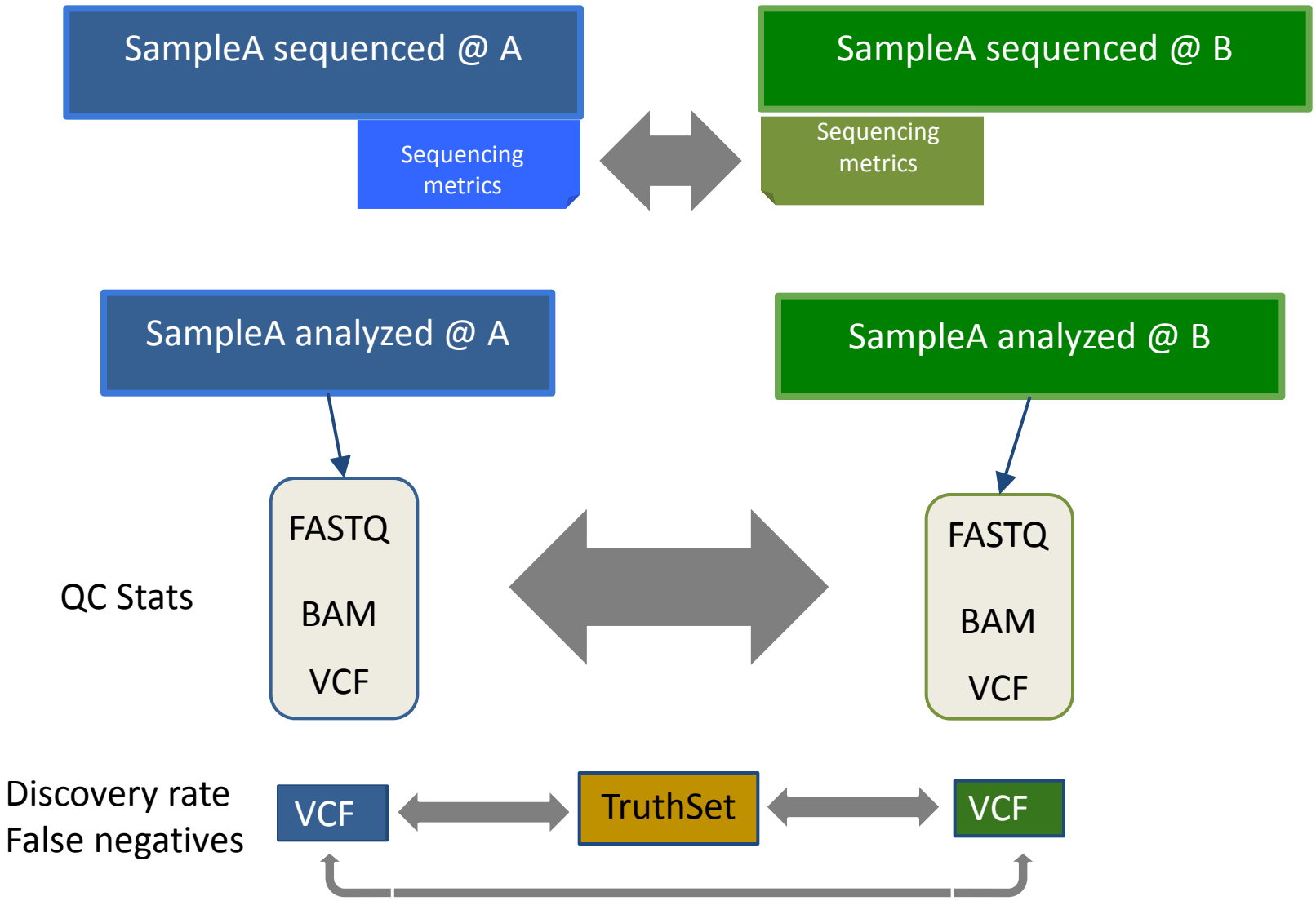
Summary

- Specific challenges:
  - What is the analyte? FASTQ? Variant calls? Interpretation?
  - How to specify, calculate, evaluate and represent Accuracy and Reliability?
    - All components in-house
    - All components out-sourced
    - Hybrid model
  - How to define acceptance criteria?
    - What is the application/circumstance?
    - How is confirmation performed?
    - Who is the end consumer?
- Unit optimization = system optimization?

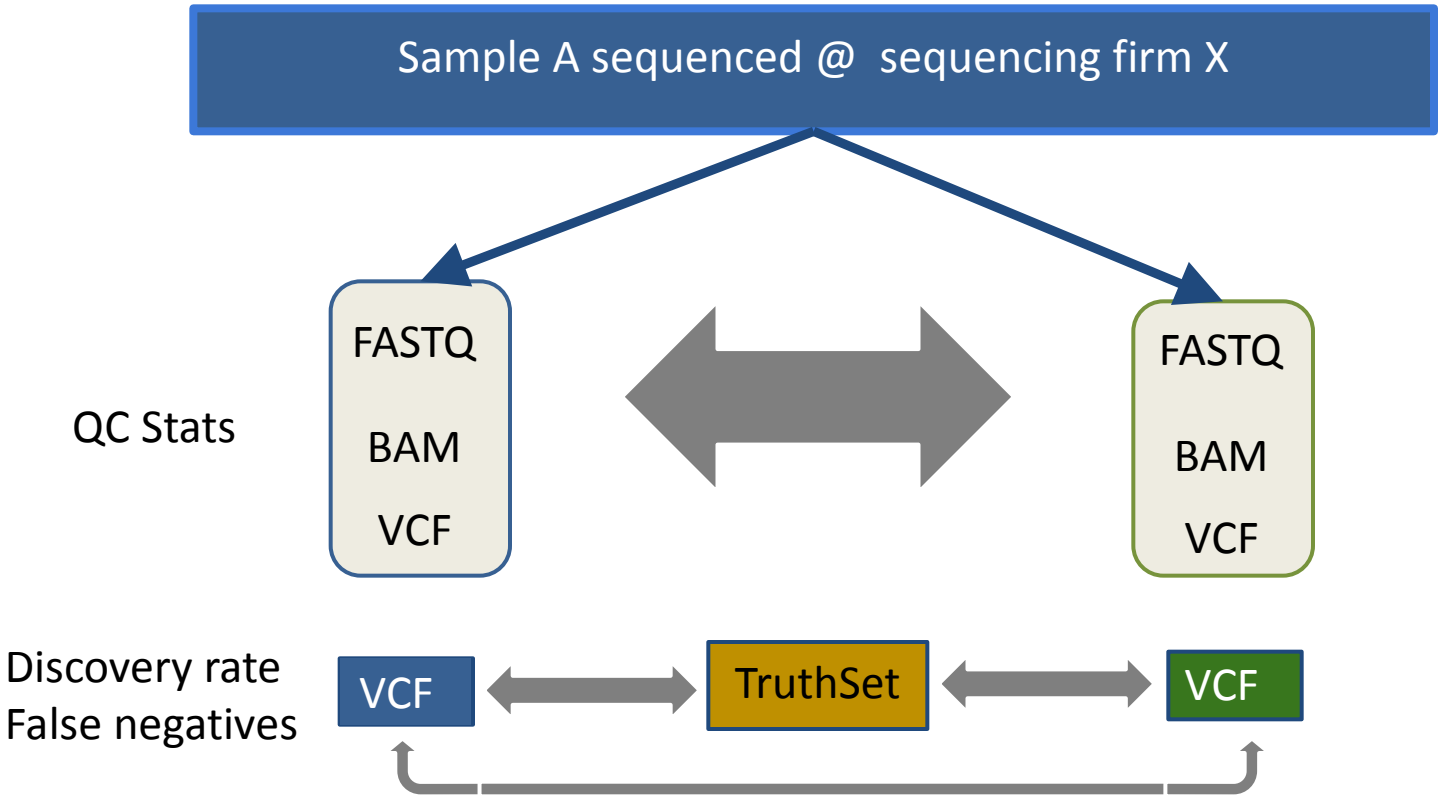
Typical scope of NGS Bioinformatics analysis



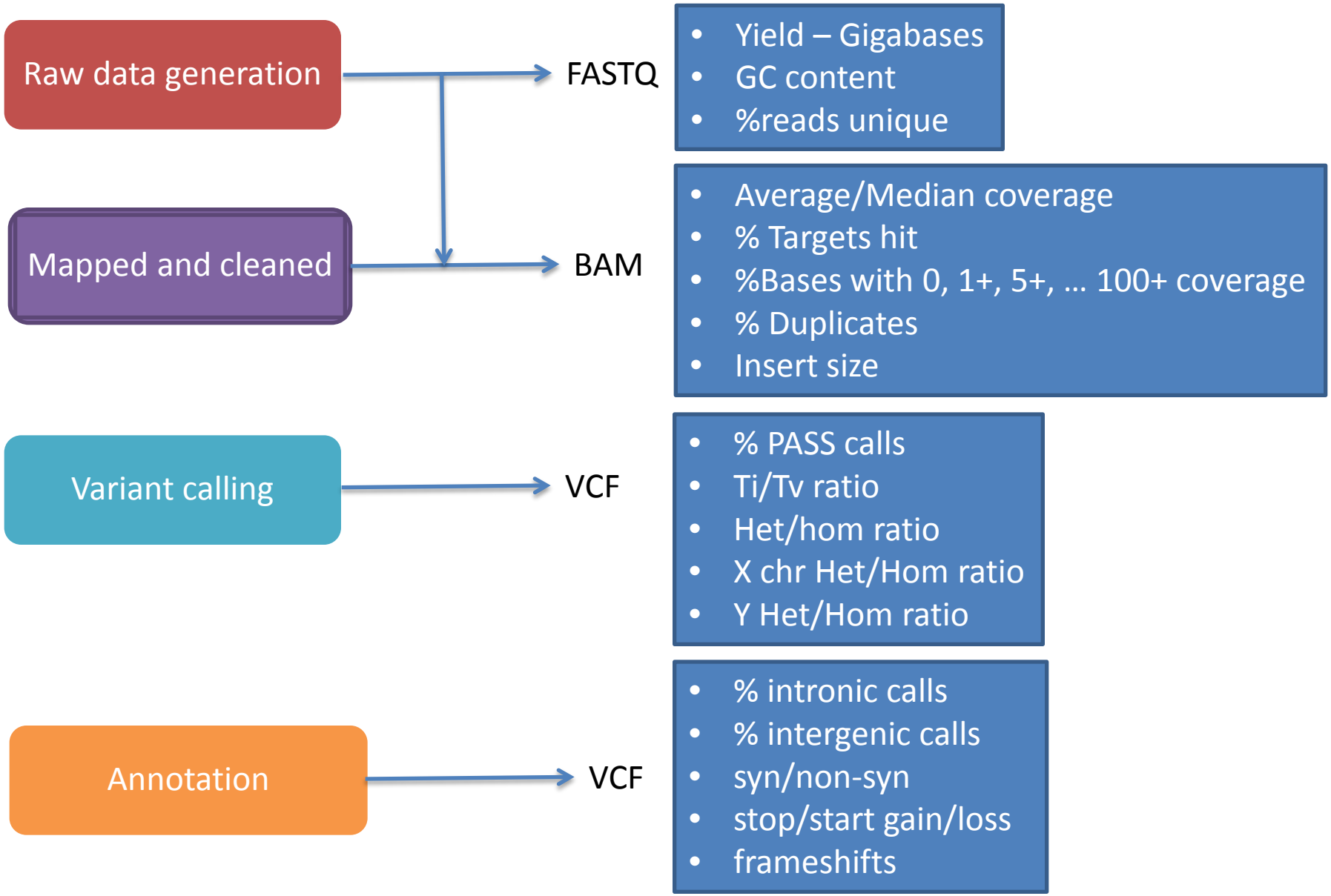
Comparing end-to-end pipeline



Comparing bioinformatic pipeline



How to specify, calculate, ….



Example benchmarks – NA12878 CEPH

Category	Pipeline A	Pipeline B
Synonymous_SNV	11476	12633
Non-Synonymous_SNV	10525	12792
StopGain SNV	145	190
StopLoss SNV	34	14
Het/Hom ratio	1.69	2.08
Ti/Tv ratio	2.84	2.75
X het count	369	462
X hom count	199	204
Y het count	3	13
Y hom count	0	4

Example benchmarks – NA12878 CEPH

Category	Pipeline A	Pipeline B
StopGain SNV	145	190
StopLoss SNV	34	14

Stop**Gain** SNVs

	A & B	Only A (# of variants)	Only B (# of variants)
Total	80	65	110
1KG MAF < 1.0%	8	1	1
1KG MAF < 0.1%	5	0	1
ESP6500 Eu < 1.0%	9	7	5
ESP6500 Eu < 0.1%	4	4	3
ARIC Eu < 1.0%	9	9	4
ARIC Eu < 0.1%	4	5	3

1KG MAF – 1000 Genome, Mean allele Frequency
ESP6500 Eu – ESP6500, European allele frequency
ARIC Eu – Atherosclerosis Risk In Communities, European allele frequency

Example benchmarks – NA12878 CEPH

Category	Pipeline A	Pipeline B
StopGain SNV	145	190
StopLoss SNV	34	14

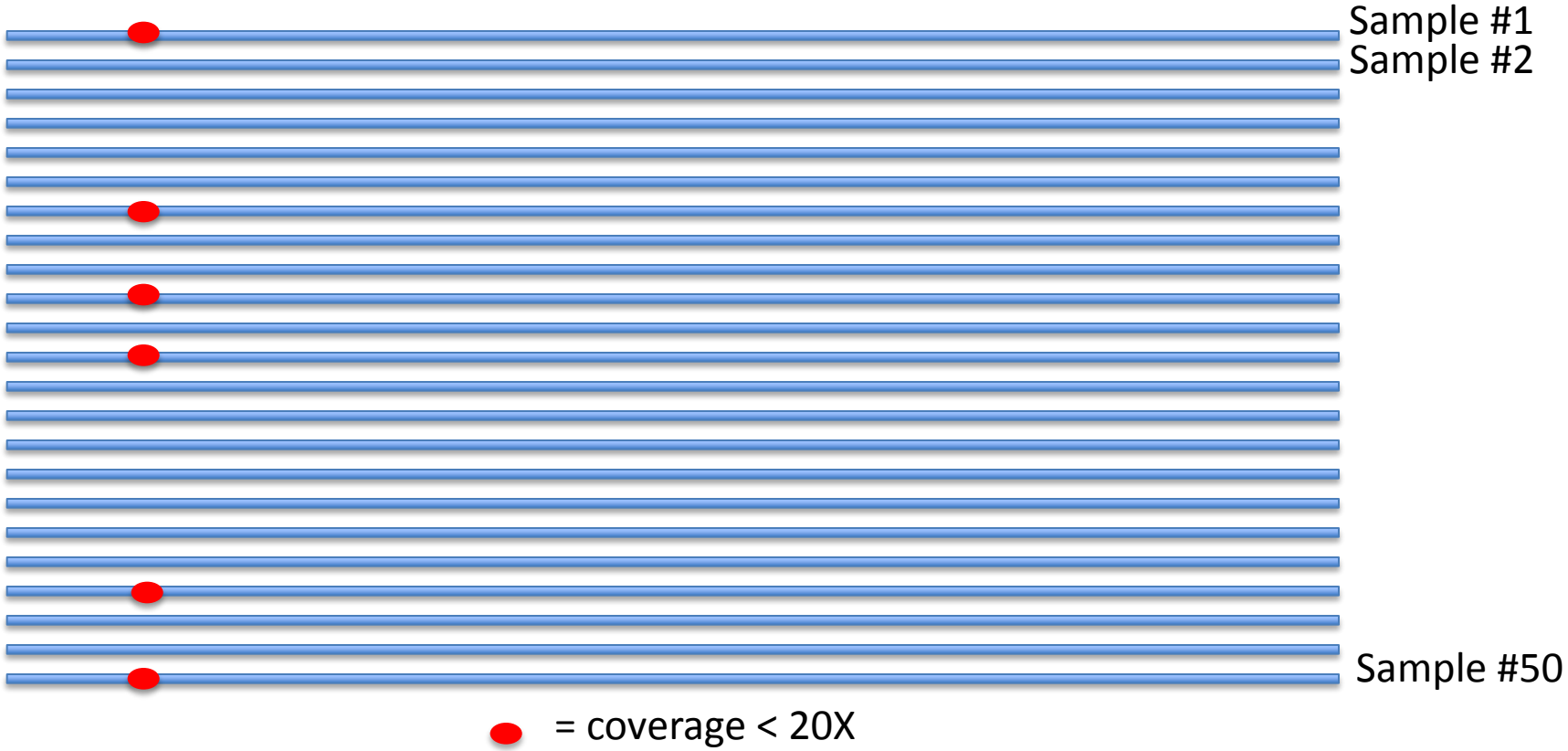
Stop**Loss** SNVs

	A & B	A (# of variants)	B (# of variants)
Total	9	25	5
1KG MAF < 1.0%	1	0	0
1KG MAF < 0.1%	0	0	0
ESP6500 Eu < 1.0%	2	1	1
ESP6500 Eu < 0.1%	0	1	1
ARIC Eu < 1.0%	1	3	0
ARIC Eu < 0.1%	0	0	0

1KG MAF – 1000 Genome, Mean allele Frequency
ESP6500 Eu – ESP6500, European allele frequency
ARIC Eu – Atherosclerosis Risk In Communities, European allele frequency

How to specify, calculate, ….

- **Question**: Will the assay adequately cover the genes/regions of interest? [Lets dig a bit deeper]



If at a particular base across N samples, if p% of samples show coverage <20X, I'll classify that site as poorly covered by the assay.

How to specify, calculate, ….

- **Question**: Will the assay adequately cover the genes/regions of interest?

Metric:
Base coverage
% of gene/region covered at >= 20X
May want to use certain base quality (say, phred score >= 20) and mapping quality (>= 1).

Method:
Determine # of reads with mapping quality > =1 at a particular position and only count those with base quality >=20 at that position.

Concerns:

- How would you treat 19 reads with extremely high base quality and 1 read with a "just-below-the-mark" base quality? Do we need a better method to capture such nuances?
- What about a high base quality site flanked by low base quality sites?
- What read quality cut-offs are used by your variant caller?
- Does using trios allow for tolerating a lower performance?

Methods used to calculate metrics need to capture the complexity of factors affecting the question.

i.e. simplicity is at a cost

How to define acceptance criteria?

- What is the circumstance?
  - NICU setting
  - Carrier Testing
  - Germline
  - Somatic

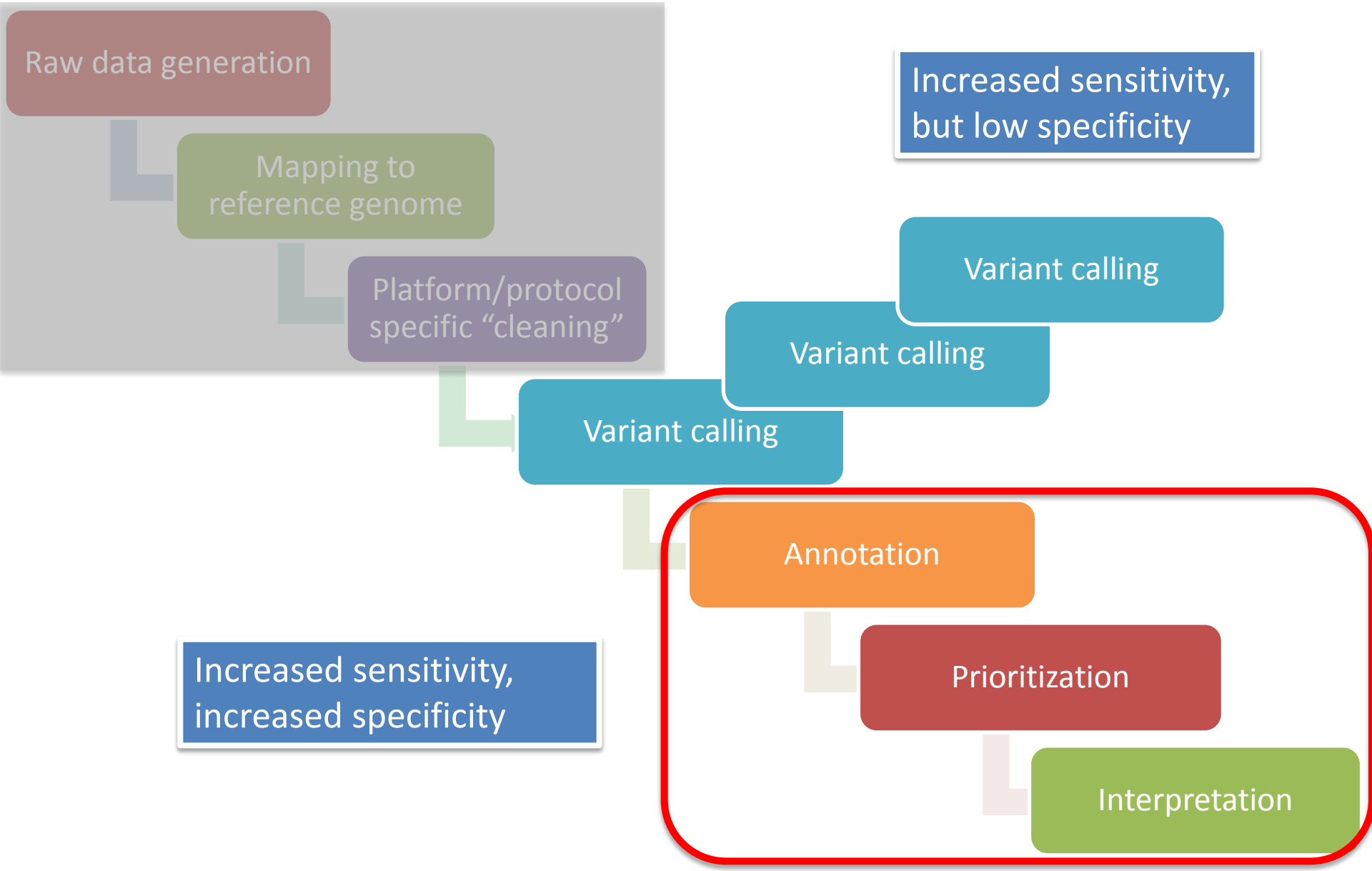
- How is confirmation performed?
  - Orthogonal platforms: Sanger, Long-range PCR, MLPA
  - Consensus tools

- Who is the end consumer?
  - CLIA lab
  - Bioinfo team
  - Physician
  - Patient/customer

Metrics are irrelevant without accompanying questions they serve to measure.

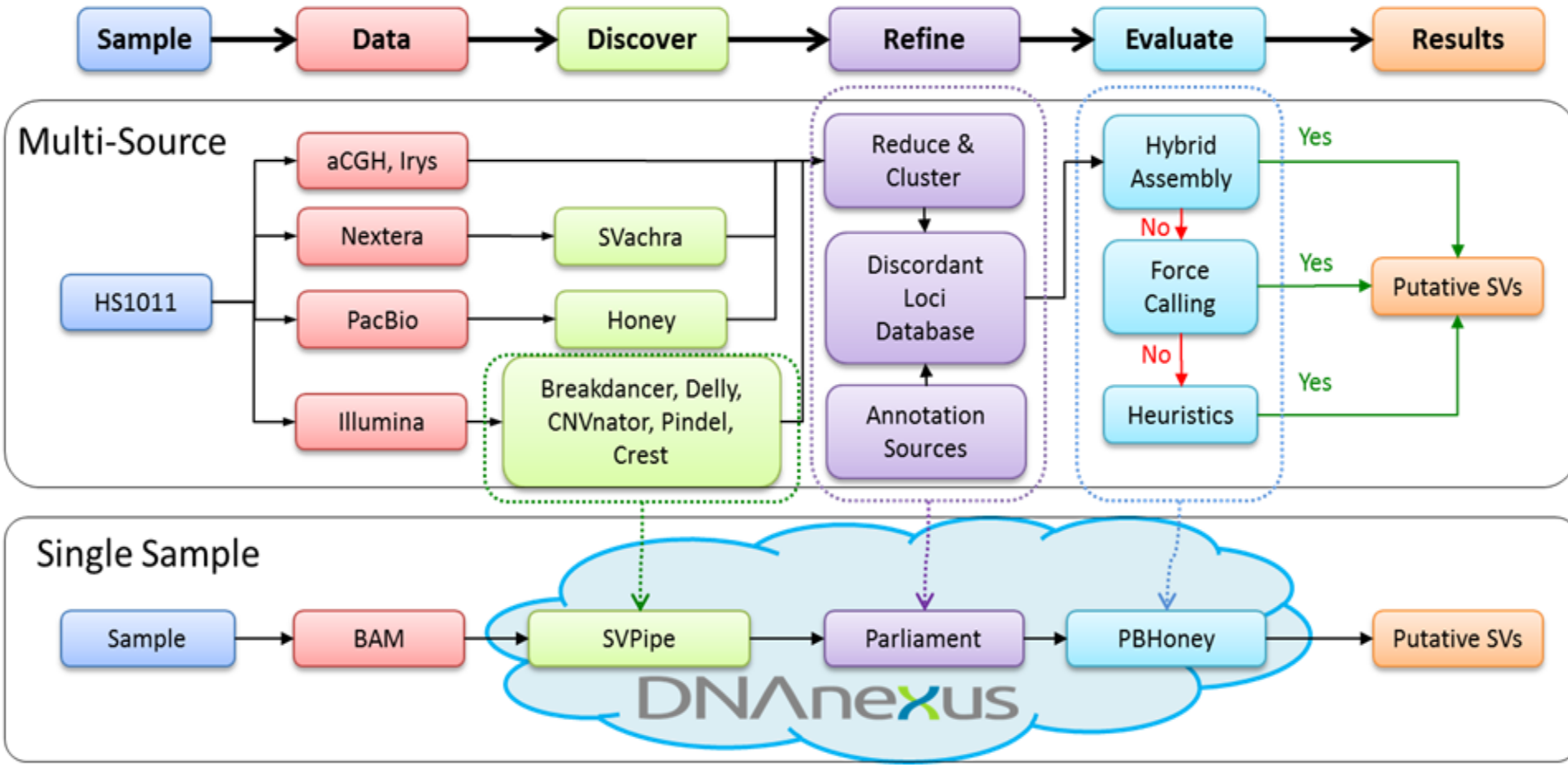
i.e. context is important

Unit optimization = system optimization?



Unit optimization = system optimization?

Structural Variation Parliament



Speed optimization around the longest running process

http://sourceforge.net/projects/parliamentsv/

http://sourceforge.net/projects/pb-jelly/

English et al., 2015, *BMC Genomics*, PMID: 25886820

A “systems” view might dominate “unit” view. So, performance metrics must be evaluated in that light.

i.e. work with the end goal in sight

Thanks!

Metrics and Standards are critical and must be carefully developed, however there is no "monolithic solution"

HGSC
HUMAN GENOME SEQUENCING CENTER

Baylor Miraca
Genetics Laboratories

Baylor
College of
Medicine
GIVING LIFE TO POSSIBLE

Panel Discussion of Topic 3
Developing bioinformatics strategies and tools for evaluating NGS tests

Moderator: Sharon Liang, PhD

Speakers and Panelists

- **Kevin Jacobs, PhD (23andMe/Global Alliance for Genomics and Health)**
- **Narayanan Veeraraghavan, PhD (Baylor College of Medicine)**
- **Sean Davis, PhD (NCI)**
- **Tina Hambuch, PhD (Illumina)**
- **Sean Hofherr, PhD (Children's National Medical Center)**
- **Niall Lennon, PhD (Broad Institute)**

BREAK

Welcome to Workshop

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015

Public Comment Speaker #1

Tara Burke, PhD

Association for Molecular Pathology

Public Comment Speaker #2

David Flannery, MD

American College of Medical Genetics and Genomics

Public Comment Speaker #3

Gillian Hooker, PhD, ScM, CGC

NextGxDx

We catalog and track almost 60,000 testing products.

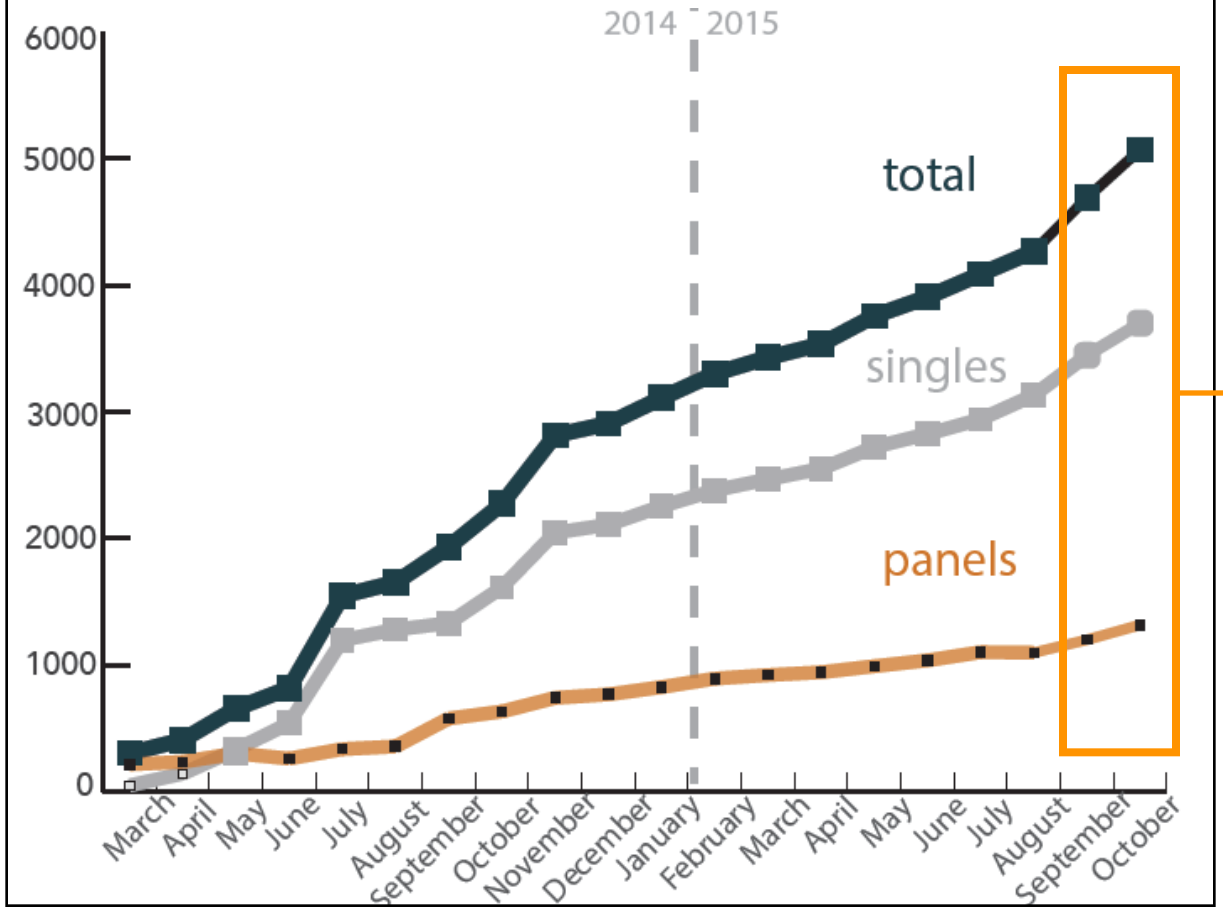
Recommendations

#1. Standards must be scalable and flexible to prevent stifling of innovation.

The Genetic Testing Market is Growing Rapidly

On average, 10 new tests enter the market every day.

Market Growth 2014–2015



~ 40% of new tests use NGS.

Recommendations

#1. Standards must be scalable and flexible to prevent stifling of innovation.

#2. Common terminology must be adopted prior to development of analytic standards.

Labs use many different names for the same entity

AKAP Lbc
ARHGEF13
BRX
c lbc
HA 3
Ht31
PROTO LB

→ AKAP13:d

Common terminology is a necessary precursor to analytical standards.

Recommendations

#1. Standards must be scalable and flexible to prevent stifling of innovation.
#2. Common terminology must be adopted prior to development of analytic standards.
#3. Standards should be built to bring transparency to clinicians and better facilitate clinical decisions around molecular testing.

Public Comment Speaker #4

Natalie LaFranzo, PhD

Horizon Discovery, Ltd.

HORIZON DISCOVERY

Engineered, Cell-Line Based Reference Standards Offer a Renewable Resource for Analytical Validation of NGS Assays

FDA Workshop - November 12, 2015

Natalie LaFranzo, PhD
Product Manager – NGS Products

horizon

Disclaimer

This Presentation does not constitute or form any part of an offer to sell, or invitation to purchase or apply for or enter into any contract or make any other commitment whatsoever in relation to, securities. Although reasonable care has been taken to ensure that the facts stated in this Presentation are accurate and that the opinions expressed are fair and reasonable, the contents of this Presentation have not been formally verified by Horizon Discovery plc (the "Company") or any other person. Accordingly, no representation or warranty, expressed or implied, is made as to the fairness, accuracy, completeness or correctness of the information and opinions contained in this Presentation and no reliance should be placed on such information or opinions. Further, the information in this Presentation is not complete and is subject to updating, revision, further verification and amendment. Neither the Company, nor any of its subsidiaries, nor any of its respective members, directors, officers or employees nor any other person accepts any liability whatsoever for any loss howsoever arising from any use of such information or opinions or otherwise arising in connection with this Presentation.

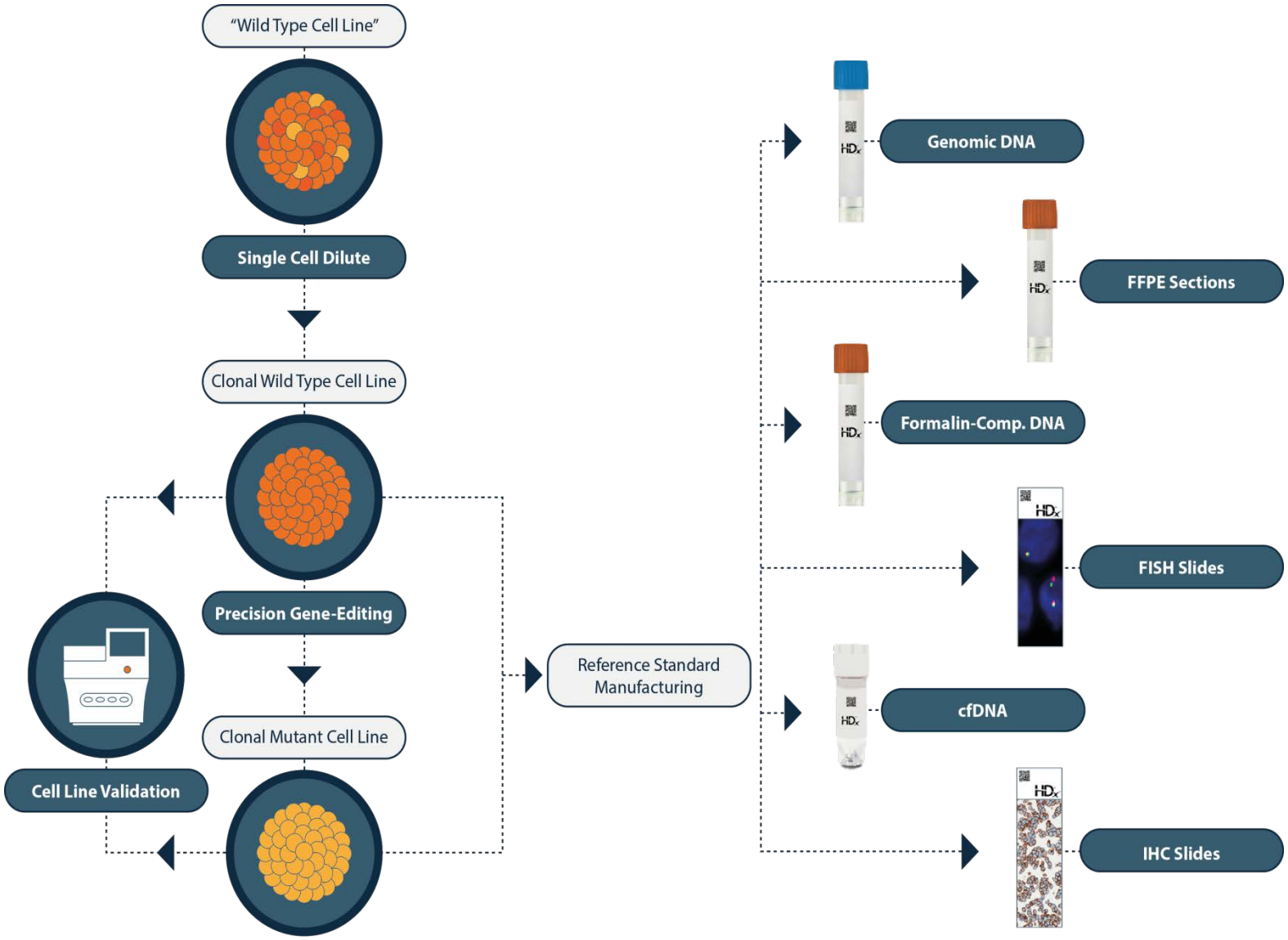
Accordingly, information contained in the Presentation is being supplied to you solely for your information and may not be copied, reproduced or further distributed to any person or published in whole or in part, for any purpose. In particular, the distribution of this Presentation in certain jurisdictions may be restricted by law, and persons into whose possession this Presentation comes should inform themselves about, and observe, any such restrictions. Any failure to comply with these restrictions may constitute a violation of laws of any such jurisdiction.

This Presentation includes certain forward-looking statements, estimates and projections with respect to the anticipated future performance of Horizon Discovery plc, its products and the markets in which it operates. Forward-looking statements involve risks and uncertainties. Actual events could differ materially from those projected herein and such statements, estimates and projections reflect the various assumptions made by the Company which assumptions may or may not prove to be correct. These forward-looking statements speak only as at the date of this Presentation. The Company expressly disclaims any obligation or undertaking to disseminate any updates or revisions to any forward-looking statements contained in the Presentation to reflect any change in the Company's expectations with regard thereto or any change in events, conditions or circumstances on which any such statements are based.

No part of this Presentation, or the fact of its distribution, should form the basis of or be relied upon in connection with any contract or commitment or investment decision whatsoever. This Presentation does not constitute a recommendation regarding the securities of the Company.

By participating in and/or accepting delivery of this Presentation you agree to be bound by the foregoing restrictions and the other terms of this disclaimer.

Engineering Cell Lines into Reference Standards



n.lafranzo@horizondiscovery.com

Advantages of Engineering Cell Lines into Reference Materials

Mimics individual patient genetics

Variants presented in relevant genomic context

Range of allelic frequencies

Quality-controlled and validated

Prepared under a certified quality management system

Renewable and affordable

n.lafranzo@horizondiscovery.com

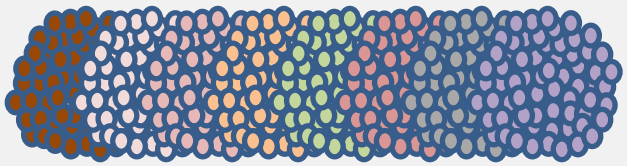
Questions for the Community

- What disease-relevant mutations are most needed as reference standards?
  - *Uncommon in clinical archives, hard to source, and/or difficult to detect*
- What formats are required to best evaluate analytical performance of NGS assays?
  - *Formalin treatment affects both the pre-analytical and analytical workflow*
- How can we develop reference standards which address and harmonize all of the regulatory guidelines [CAP, CLIA, NY State, future FDA]?
- How should reference standards be accounted for in reimbursement/billing of diagnostic tests?

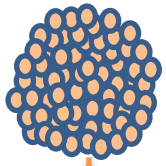
n.lafranzo@horizondiscovery.com

Developing a Universal Reference Standard

Genetically Defined Mutant Cell Lines



EML4-ALK FFPE curl	
FISH	✓
IHC	✓
gDNA	✓
RNA	✓

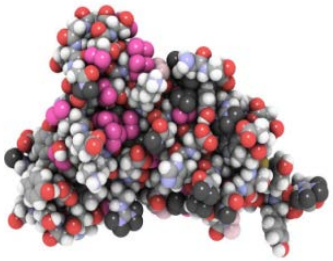


EML4-ALK

FISH

gDNA

RNA



IHC

n.lafranzo@horizondiscovery.com

Your Horizon Contact:

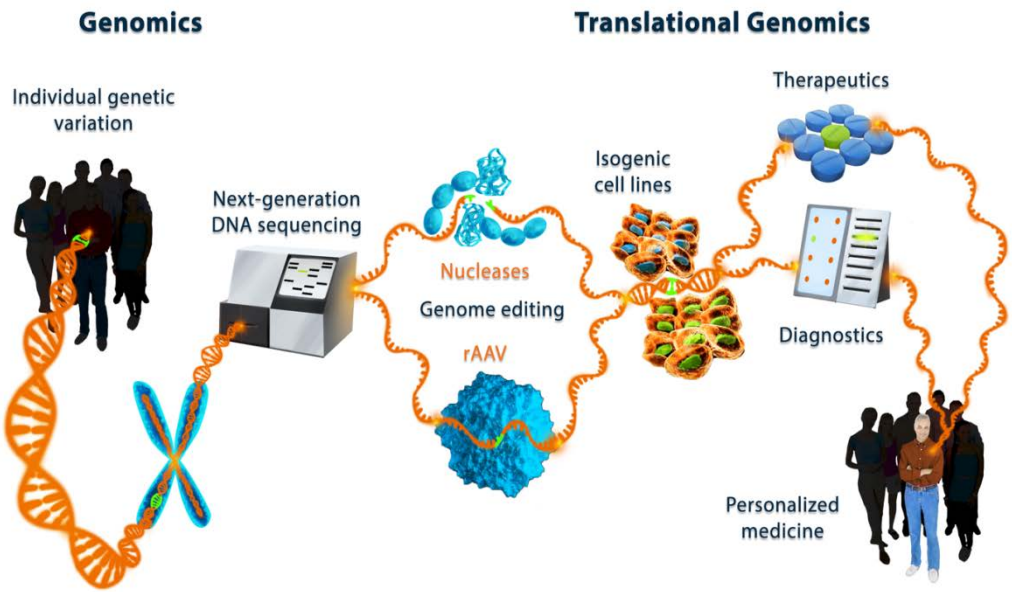
Natalie LaFranzo, PhD
Product Manager, Next Gen Sequencing Products

n.lafranzo@horizondiscovery.com

Direct: (314) 400-6636

About Horizon Discovery

Powering Genomic Research and Translational Medicine, from Sequence to Treatment



Horizon's mission is *to be a fully integrated life science company that provides enabling products, services and research programs to clients engaged at every stage of the healthcare continuum from sequence to treatment*

n.lafranzo@horizondiscovery.com

Abstract Provided

Horizon Discovery utilizes gene-engineering strategies to introduce disease-relevant mutations into single-cell diluted human cell lines. These engineered lines are presented as FFPE cell lines, high molecular weight DNA, or fragmented DNA prepared under strict manufacturing and quality control guidelines. These cell-line derived materials offer significant advantages as potential standards for next generation sequencing (NGS) assays. Horizon is keenly interested in understanding the needs of the IVD and companion diagnostics community, defining and adhering to regulatory guidance, and providing input on feasibility and scientific best practices.

horizon diagnostics

n.lafranzo@horizondiscovery.com

How are our standards prepared?

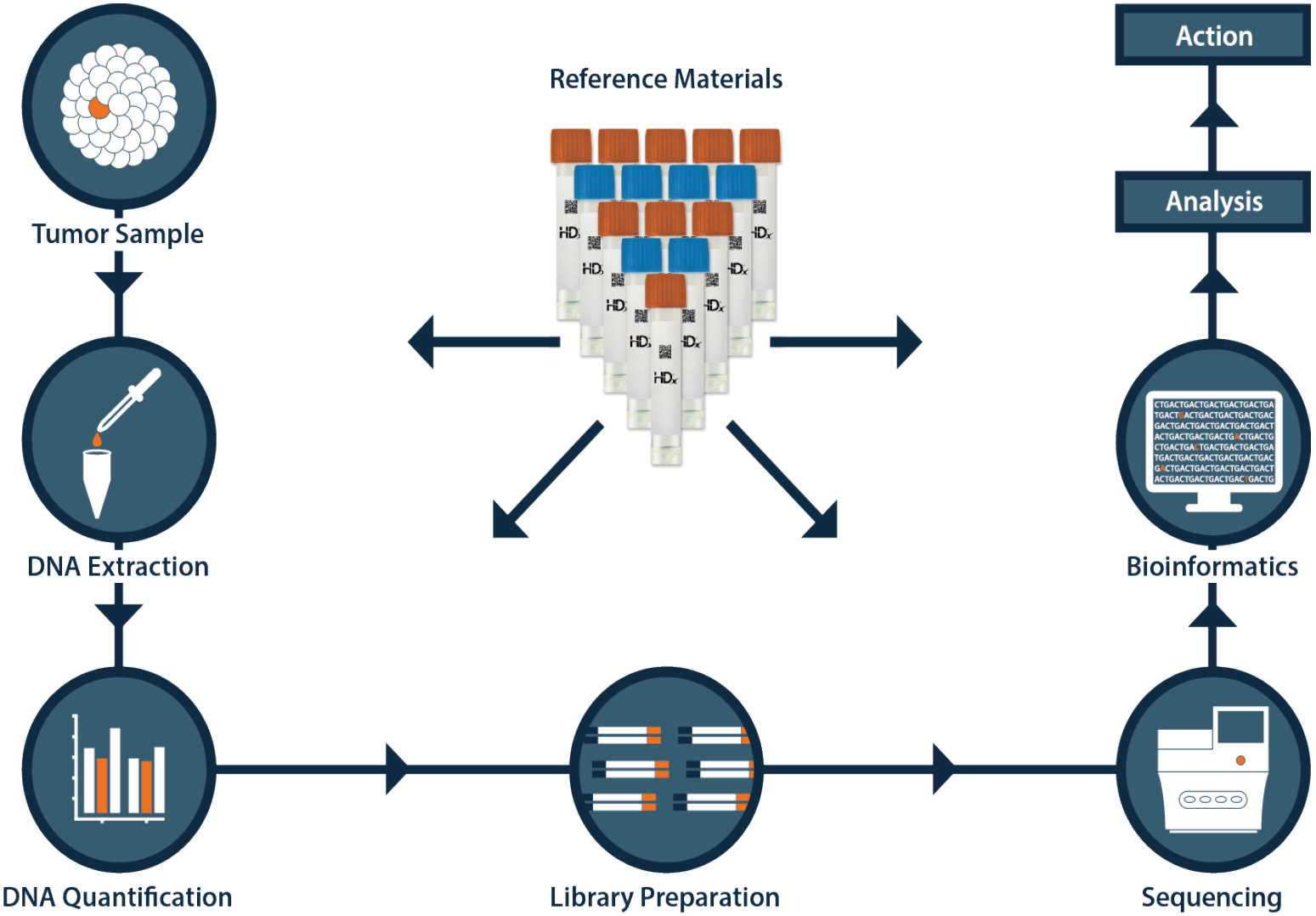
Genomic DNA
Stoichiometric Dilutions

Engineered variants (usually) start at 50% allelic frequency (heterozygous mutation) and may be diluted down to lower allelic frequencies using the isogenic, matched wild type.

Alternatively, different mutations may be mixed together to generate a “multiplexed” reference standard.

n.lafranzo@horizondiscovery.com

Sources of Variability in the Next Generation Sequencing Workflow



n.lafranzo@horizondiscovery.com

Public Comment Speaker #5

Steven Lincoln

Invitae, Inc.

Public Comment Speaker #6

Matthew Marton, PhD

Merck Research Labs

Public Comment Speaker #7

Mitch Nelles, PhD

Chief Operating Officer

CareDx Inc

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing in Vitro Diagnostic Tests

Mitch Nelles, PhD

Chief Operating Officer

CareDx Inc

Brisbane, CA

November 12, 2015

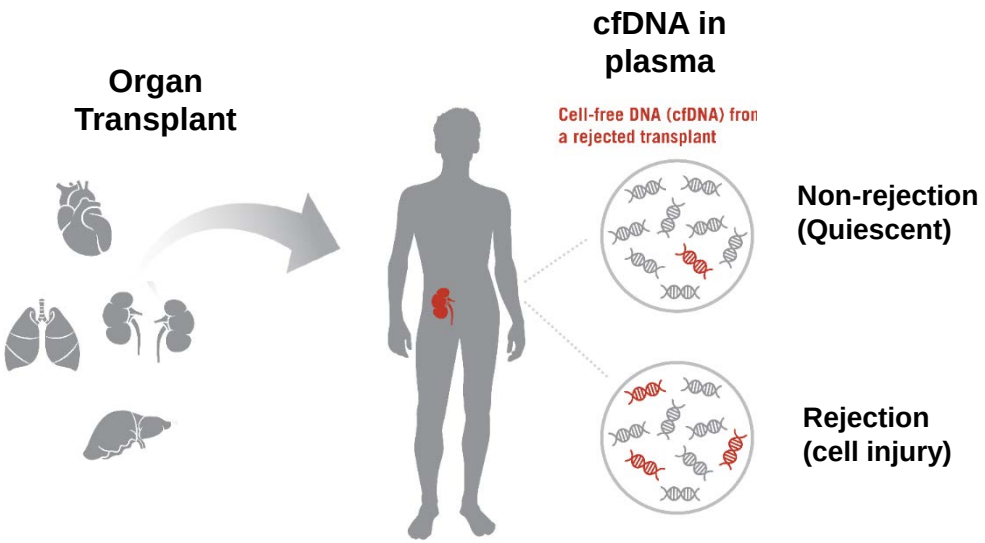
AlloMap

- Assesses the gene expression profile of RNA isolated from PBMC
  - Intended to aid in the identification of heart transplant recipients with stable allograft function who have a low probability of moderate/severe ACR at the time of testing in conjunction with standard clinical assessment.
  - Indicated for heart transplant recipients:
    - 15 years of age or older
    - At least 2 months (≥55 days) post-transplant
- Test measures the expression of 20 genes
  - 9 control and normalization genes
  - 11 informative genes
  - Score is a composite of the expression all 11 reporting genes representing multiple diverse molecular pathways (Score range: 0-40)

Pathway and Gene	Gene Expression Level
T cell priming	
ITGA4 **Integrin alpha-4** α subunit of VLA-4; involved in T cell trafficking and adhesion	↑
PDCD1 **Programmed cell death** T cell costimulatory molecular (inhibitory); CD28 family	↑
Proliferation and mobilization of erythrocytes	
MARCH8 **Cellular mediator of immune response (MIR)** E3 ubiquitin ligase	↑
WDR40A **WD repeat domain 40A** Uncharacterized protein of the WD-repeat protein family	↑
Platelet activation	
PF4 **Platelet factor 4** Chemokine-like molecule expressed in platelets	↓
C6orf25 **G6b inhibitory receptor** Putative inhibitory receptor of the Ig superfamily expressed in platelets	↓
Steroid response	
IL1R2 **Interleukin-1 receptor type II** IL-1 decoy receptor inhibits cytokine signaling; steroid-dependent expression	↓
ITGAM **Integrin alpha-M** α subunit of MAC-1; involved in cell trafficking	↓
FLT3 **FMS-like tyrosine kinase** Signaling molecule expressed in monocytes	↓
Unknown role	
SEMA7A **Semaphorin 7A** Expressed by T cells, B cells, and immature granulocytes	↑
RHOU **Ras homolog gene family, member U** Member of the Rho GTPase family involved in the modulation of cytoskeleton organization	↑

Donor-derived cell-free DNA (dd-cfDNA)

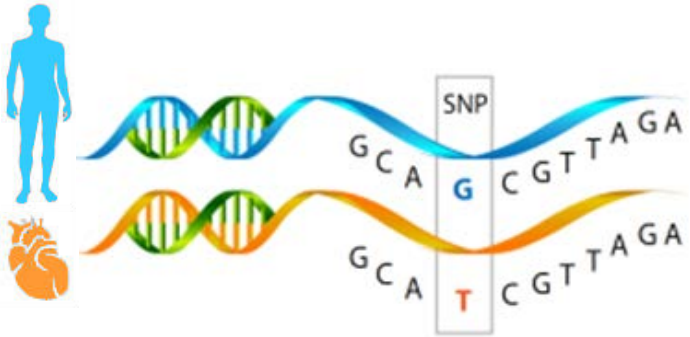
New Biomarker for Surveillance ofTransplanted Organ



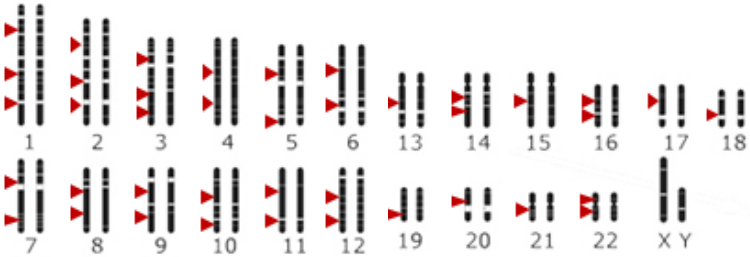
% of dd-cfDNA can be measured in plasma and increases as cells from the transplanted organ are injured (apoptosis, necrosis)

Measuring Donor-Derived Cell-Free DNA in Transplant Patients

- dd-cfDNA is measured by determining the fraction of donor-specific nucleotides at single nucleotide polymorphism (SNP) loci in a transplant patient
- We developed a method to measure the small amount of dd-cfDNA from plasma:
  - Extract cfDNA
  - Amplify 266 SNP loci selected to be sufficient to accurately measure dd-cfDNA. Amplify by concurrent multiplex PCR (Access Array technology, Fluidigm)
  - Sequence amplified SNP loci (MiSeq next-generation sequencing, Illumina)
  - Calculate the fraction of donor-specific nucleotides (CareDx)



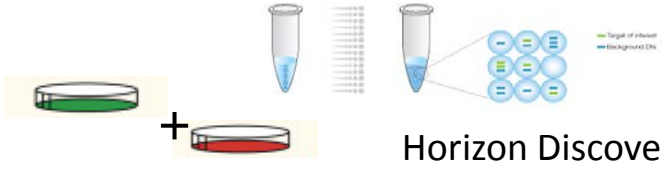
266 SNPs selected from across the human genome

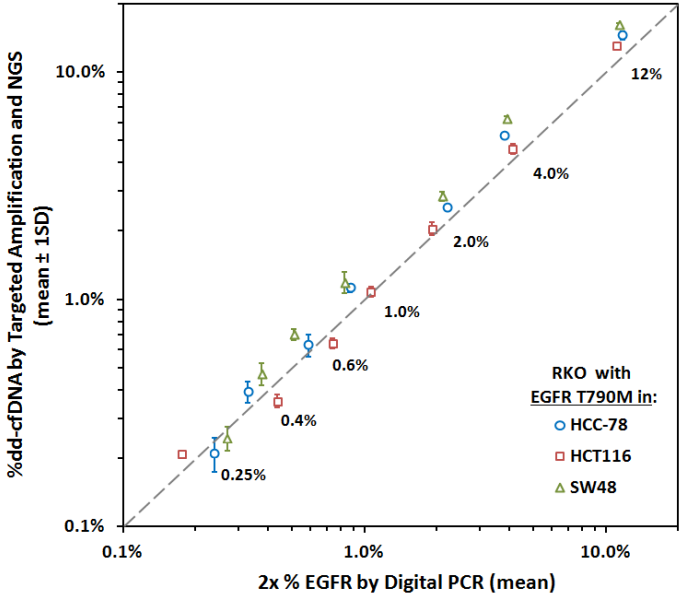


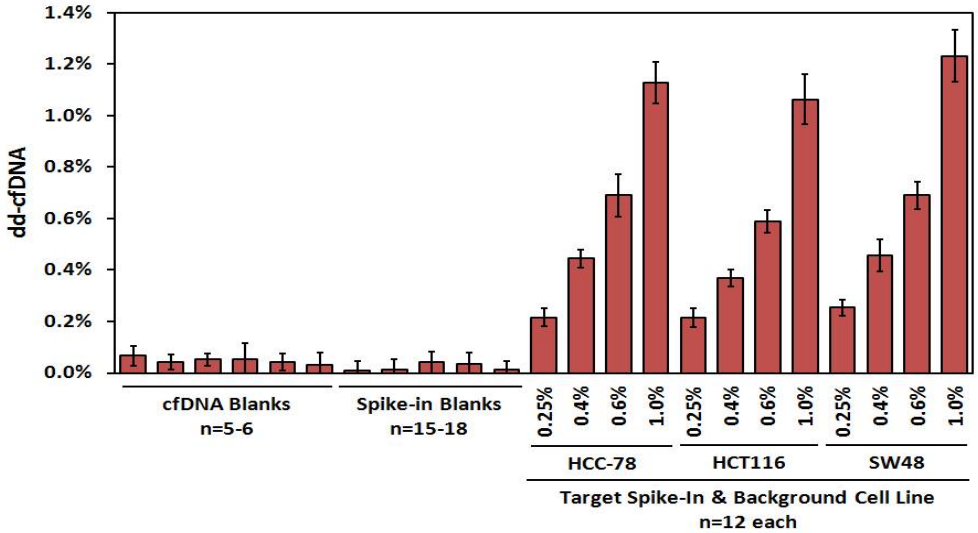
Fit-for-Purpose Design Concept and Individual Performance Standards

- NGS-based tests produce data for numerous individual variants and not always feasible to develop and review performance of all variants in a test
  - Design concept standards
    - Select representative variants with boundary properties from different regions to reflect reasonable demonstration of device's overall performance
    - Performance metrics for some variants may not be inferred with high confidence
  - Individual performance standards
    - Development of individual performance metrics for each variant is burdensome
    - Due to technological challenges of some sequencing variants, insurance of inferred performance may not be suitable

Reference Standards are Critical







- Reference standards should be sustainable and available for community-wide use
- Optimally quantitative to permit analytical performance review
- Use orthologous quantitation procedure(s) to increase confidence in result

Methods-based Proficiency is Important

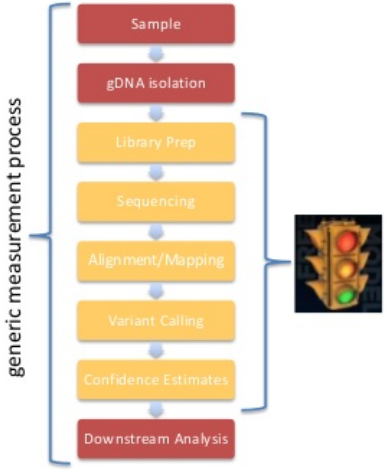
NIST Human Genome Reference Materials (RMs)

- NIST RM 8398 is available!
  - tinyurl.com/giabpilot
  - DNA isolated from large growth cell cultures
  - Stable, homogeneous
  - Best for regulated uses
  - DNA from same cell line at Coriell (NA12878)

Genome in a Bottle Consortium genomeinabottle.org

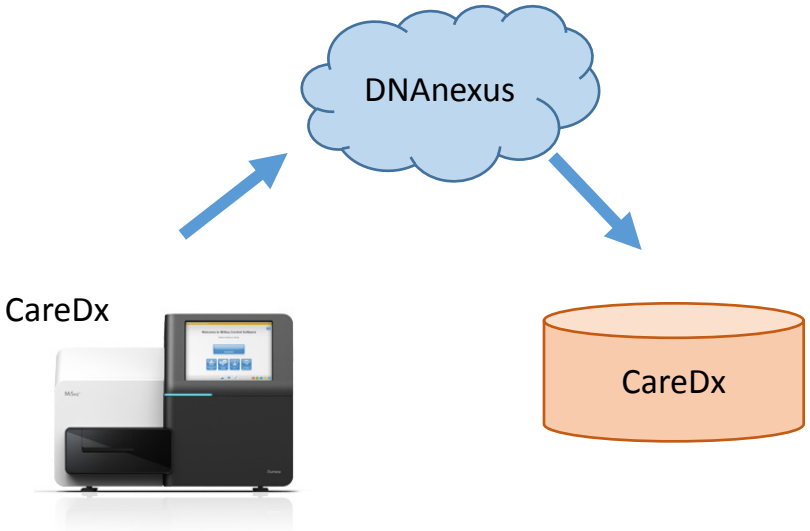
Measurement Process

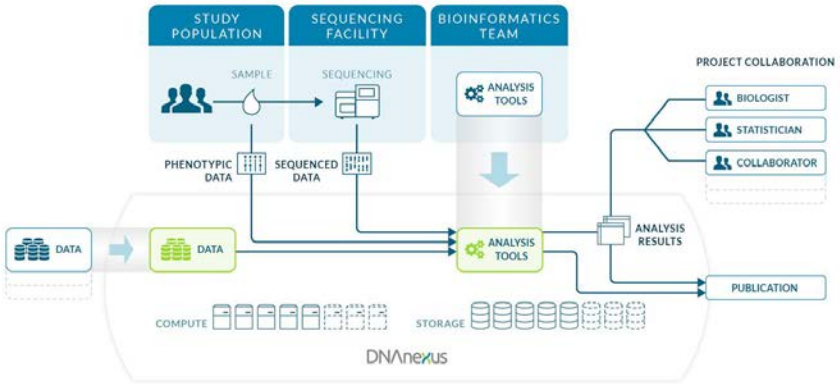
- gDNA reference materials will be developed to characterize performance of a part of process
  - materials will be certified for their variants against a reference sequence, with confidence estimates



- Alternative assessment procedures complement analyte-specific proficiency
- Delineates a step or process within the entire workflow for review

Standardization of Computational Analysis





- Explore and encourage cloud based computing to permit independent review of custom pipeline analysis
- In silico constructed standards can play an important role in computational validation

Summary

For clinical-grade NGS testing, my colleagues at CareDx and I encourage the FDA and our industry to

- Apply a combination of Design Concept Standards and predefined individual variant performance standards
- Develop and utilize well characterized, sustainable reference materials
- Take full advantage of pre-existing recommendations for methods-based proficiency testing in conjunction with the availability of new genome reference resources from Genome in a Bottle and NIST
- Implement rigorous and standardized strategies for computational pipeline analysis

Public Comment Speaker #8

Girish Putcha, MD, PhD

Palmetto GBA/MolDX

MolDX®

PALMETTO GBA®

A CELERIAN GROUP COMPANY

Public Workshop - Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

12 November 2015

Girish Putcha, MD, PhD
Director of Laboratory Science

Please note that the opinions expressed herein are my own.

Inconvenient Truths

- As an industry, we are committing fratricide, yet our fate is largely shared. Rather than focus on what divides us, we might do better to focus on what unites us . . .

Inconvenient Truths

- A broken regulatory "system" with 2 parallel universes:
  - System 1: Generally high, fairly uniform and transparent, and externally assessed evidentiary standards, but considered slow and expensive
  - System 2: Highly variable, generally opaque and self-determined evidentiary standards that allows almost anyone to make almost any claim without even the minimum transparency required to judge the veracity of those claims, but considered facile and inexpensive.
- A broken reimbursement system that's highly fragmented; idiosyncratically variable in its coverage decisions despite presumably asking similar if not identical questions; and archaic insofar as (unlike most other areas of healthcare) it bases pricing on cost instead of market, let alone on value.

Please note that the opinions expressed herein are my own.

Inconvenient Truths

- Unfortunately, NGS is just the latest sexy new technology to suffer from all the same old problems.
- But it also presents us with an opportunity to try to fix at least some of what ails us:
  - Creating at least analytical standards that can be applied to all NGS-based tests, or in a more nuanced manner to tests based on their intended use
  - Improving the transparency around the performance characteristics of these tests in such a way that “apples to apples” comparisons are at least possible in the marketplace by end users, such as patients, providers, payers, etc.

Please note that the opinions expressed herein are my own.

Inconvenient Truths

- All regulation is inherently social engineering that seeks to balance risks and benefits so is there a problem here that needs a solution? And is the solution appropriately titrated to the problem?
  - Do we know? The absence of evidence is NOT evidence of absence. So who bears the burden of proof?
  - For example, do we know whether response rates to targeted drugs prescribed on the basis of NGS-based test results are at least non-inferior to that based on the original companion diagnostic? Is the answer different for different technology platforms and sequencing chemistries? (And to be absolutely clear, the same question can be asked of any technology platform, not just NGS.)

Please note that the opinions expressed herein are my own.

Inconvenient Truths

- If the answer is “we don’t know”, then we as a community can decide that it doesn’t matter, or that the benefits outweigh the risks based on what we do know. But let’s at least be intellectually rigorous enough to acknowledge this in the clear light of day. And I submit that the same is true for most, if not all, of the questions about the clinical applications of NGS.
- I sincerely commend the FDA for its efforts to be thoughtful, prudent and parsimonious in its approach today. They may not get it right, but they certainly seem to be trying. And we in the industry will have no one to blame but ourselves if we don’t at least work with them to try to do this “right”.

Please note that the opinions expressed herein are my own.

Thank you.

Questions and comments are welcome.

girish.putcha@palmettogba.com

Please note that the opinions expressed herein are my own.

Public Comment Speaker #9

Judi Smith

Precision for Medicine

Public Comment Speaker #10

Daniel Weaver

Genformatic

Thank you for your feedback!

Docket FDA-2015-N-2881-0001

Open through November 25, 2015

http://www.regulations.gov/#!docketDetail;D=FDA-2015-N-2881

Wrap-up

Elizabeth Mansfield, PhD

Office of In Vitro Diagnostics and Radiological Health
Center for Devices and Radiological Health
U.S. Food and Drug Administration

Thank You!

Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests

#PMIFDA
#precisionmedicine

Silver Spring, MD, November 12, 2015