Lessons Learned from Sequencing Clinical Specimens

NCI Workshop: Next-Generation DNA Sequencing as a Tool for Clinical Decision-making in Cancer Patient Management

Maureen Cronin, Ph.D.

May 3, 2012

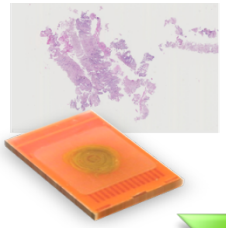
Cancer Genomics – Benefits and Challenges in Comprehensive NGS Testing for the Clinic

- One technology captures the full distribution of somatic genetic aberrations
  - Many important mutations occur at low frequency
  - 75% or more of variation missed by "hot-spot" analysis
  - Full genomic profiling supports pathway views of cancer
- Many individual mutation tests, with each test likely to have a negative result is unrealistic
  - There is not enough tumor sample, time or money for so many tests, and the logistics are impractical
- Mutation profiles from NGS are complex to validate
- Mutations profiles currently are difficult to translate into patient care

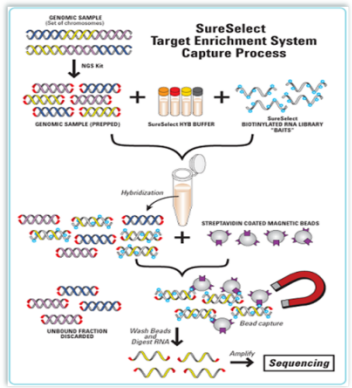
Successful NGS Testing Integrates Multiple Optimized Components

- Pre-analytic assessment of the specimen pathology with defined acceptance criteria
- Validated/standardized DNA extraction method with QC metrics for DNA yield and quality
- Validated/standardized sequencing library preparation method with QC metrics for successful library prep
- For targeted sequencing, performance-validated pool of capture probes with QC metrics for capture performance
- Validated/standardized sequencing protocol with QC metrics for acceptable sequence run results
- Post-analytic validated/standardized processes for data analysis pipeline and mutation profile assignment with defined sensitivity and specificity performance

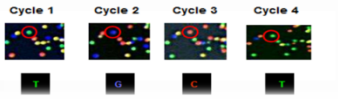
Targeted Cancer Genome Profiling Workflow



Fixed sample slides



DNA extraction, Library construction, Hybrid capture



Sequencing (Illumina HiSeq™ 2000)

ANALYSIS PIPELINE

Point Mutations
Bayesian algorithm

Short Insertions/deletions
Local assembly

Copy Number Alterations
Comparison with process-matched normal control

Rearrangements
Analysis of chimeric pairs

Annotation & Interpretation

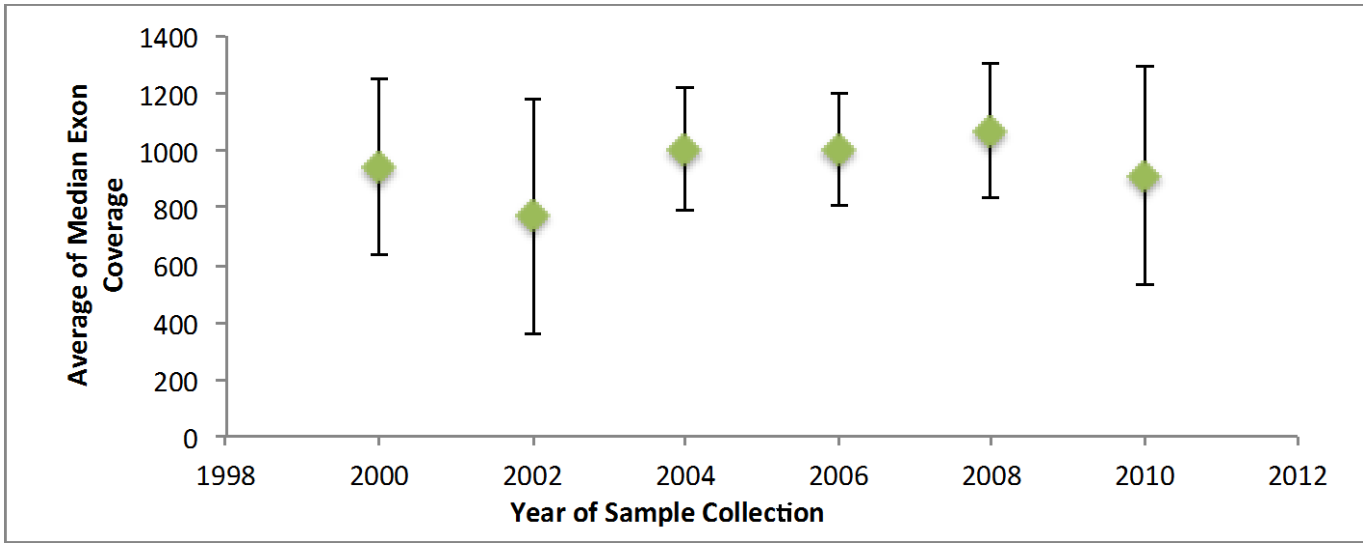
dbSNP
COSMIC
Med. Literature

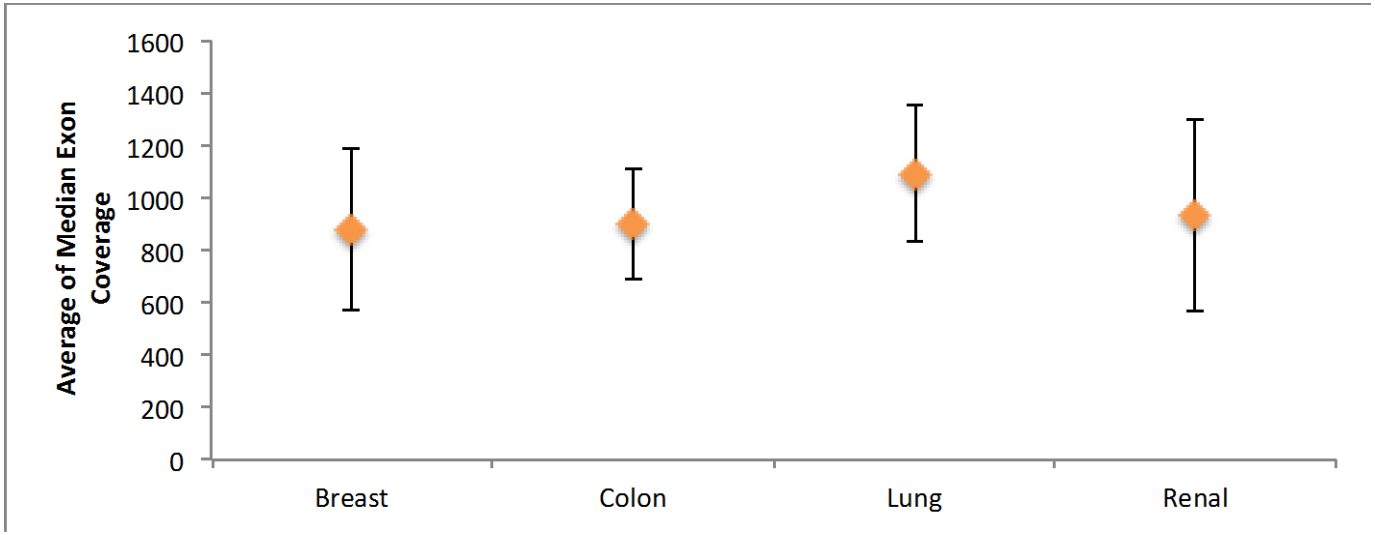
Report

Pre-analytic Specimen Assessment
Solid Tumor Samples and Extracted DNA

- Provide guideline for sample types, amount of specimen and sample purity to submit (e.g., 40 μm of FFPE cut from NSCLC, CRC, BrCa, PrCa, resected or core biopsy specimens with ≥ 1mm^2 surface area) and have every submitted sample path reviewed (H&E slide)
- Standardized DNA extraction protocol with a minimum yield requirement validated to perform in the standard sequencing process (e.g., 150ng of dsDNA by picogreen® fluorescence assay)
- Internal control sample (e.g., sections cut from a block that has been successfully extracted and DNA sequenced); validation data for extraction success in all common tumor types
- Lessons Learned:
  - DNA yield and purity varies with tumor and sample types
  - DNA yield from a standard amount of tissue is unpredictable
  - It is important in tumor specimens to note approximate tumor purity (quartiles or better)
  - An extraction control does not always predict sample performance
  - Yield of high purity, dsDNA is important to successful library preparation

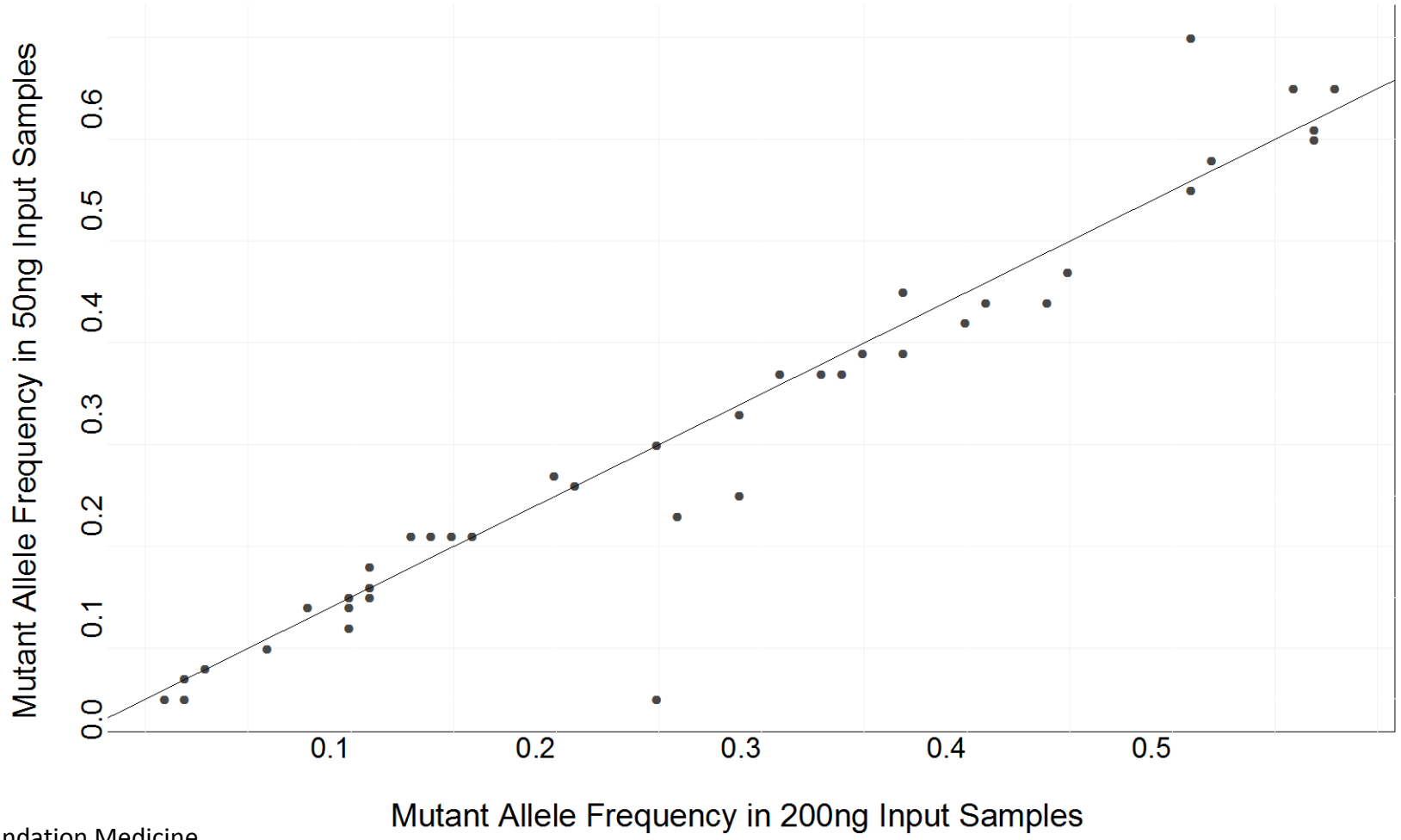
Sequencing coverage independent of tissue type and sample age





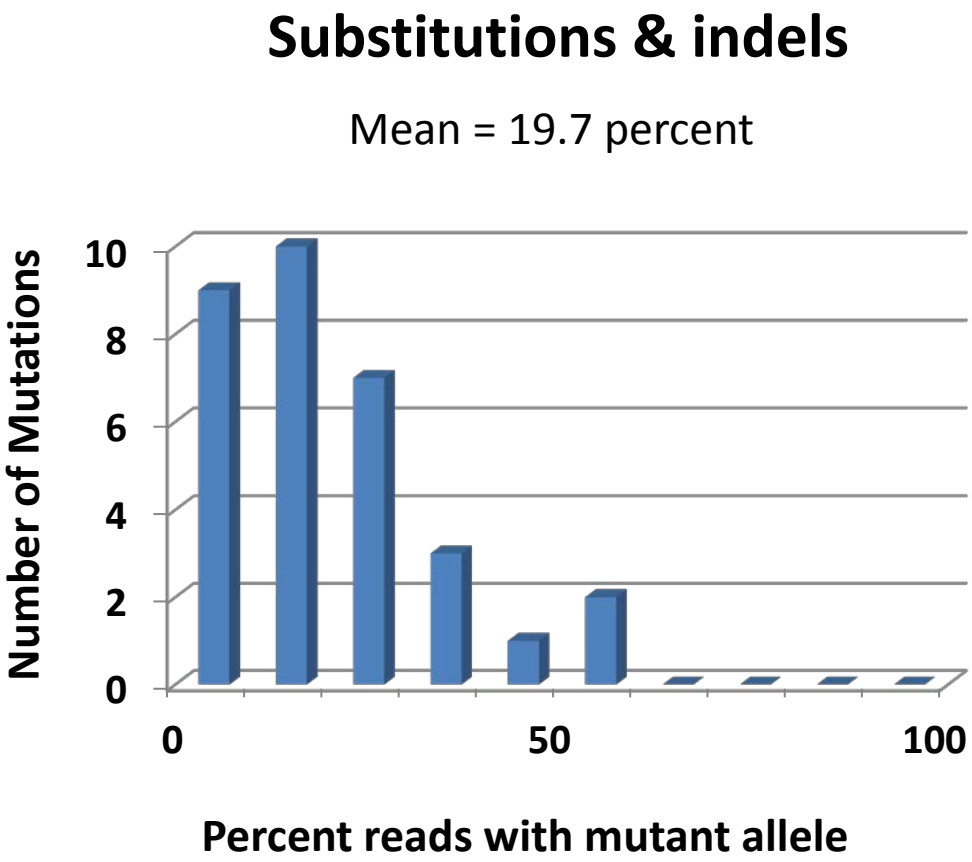
Foundation Medicine, 2011 AMP Meeting

Known somatic alterations replicate between 200ng and 50ng FFPE tumor specimen libraries



Foundation Medicine,
2011 AMP Meeting

Mutant Allele Percentages in 22 NSCLCs

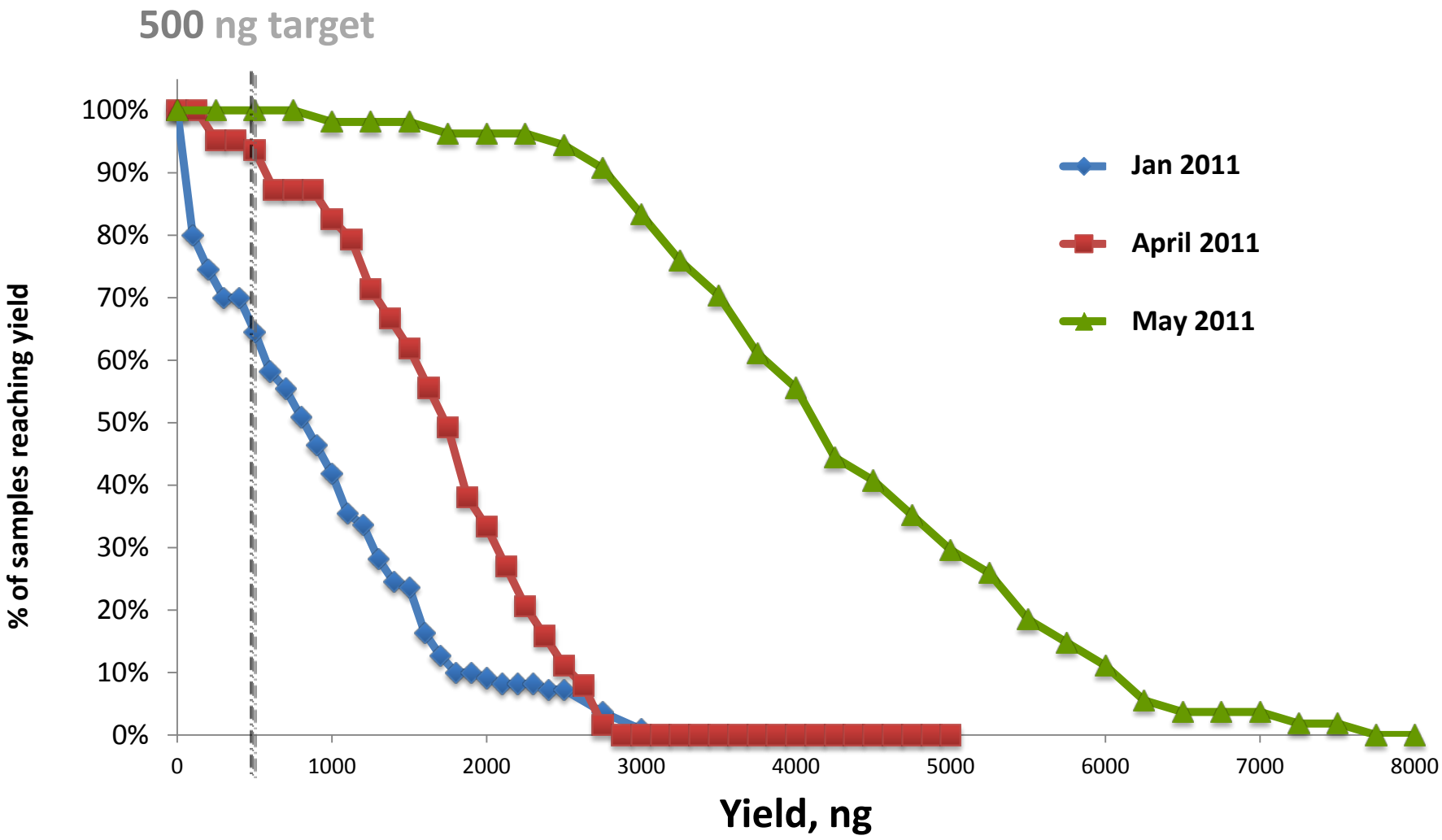


Foundation Medicine, 2011
Agilent Users Meeting, Boston, MA

Ligation-Based Library Construction

- Represents unbiased whole genome to accurately quantify genomic alterations
  - Base substitution mutations
  - Small deletions/insertions
  - Genomic rearrangements
  - Copy number alterations
- Compatible with FFPE input (short dsDNA)
- Process validation defines
  - Minimum input DNA requirement
  - Specifies parameters of all process steps
  - Defines QC criteria for successful library output (complexity and mass)
- Process controls may be characterized cell lines, normal tissue to tumor tissue but must include a control to reflect patient tissues (FFPE)
- Lessons Learned
  - PCR parameters are important and must be characterized and limited
  - Low quality DNA samples give low quality libraries
  - Low quality DNA may be “rescued” by pooling multiple independent libraries and increasing the mass of DNA used in library construction
  - Library quality is the primary determinant of quality of hybridization capture for targeted sequencing

Library Construction can be Optimized

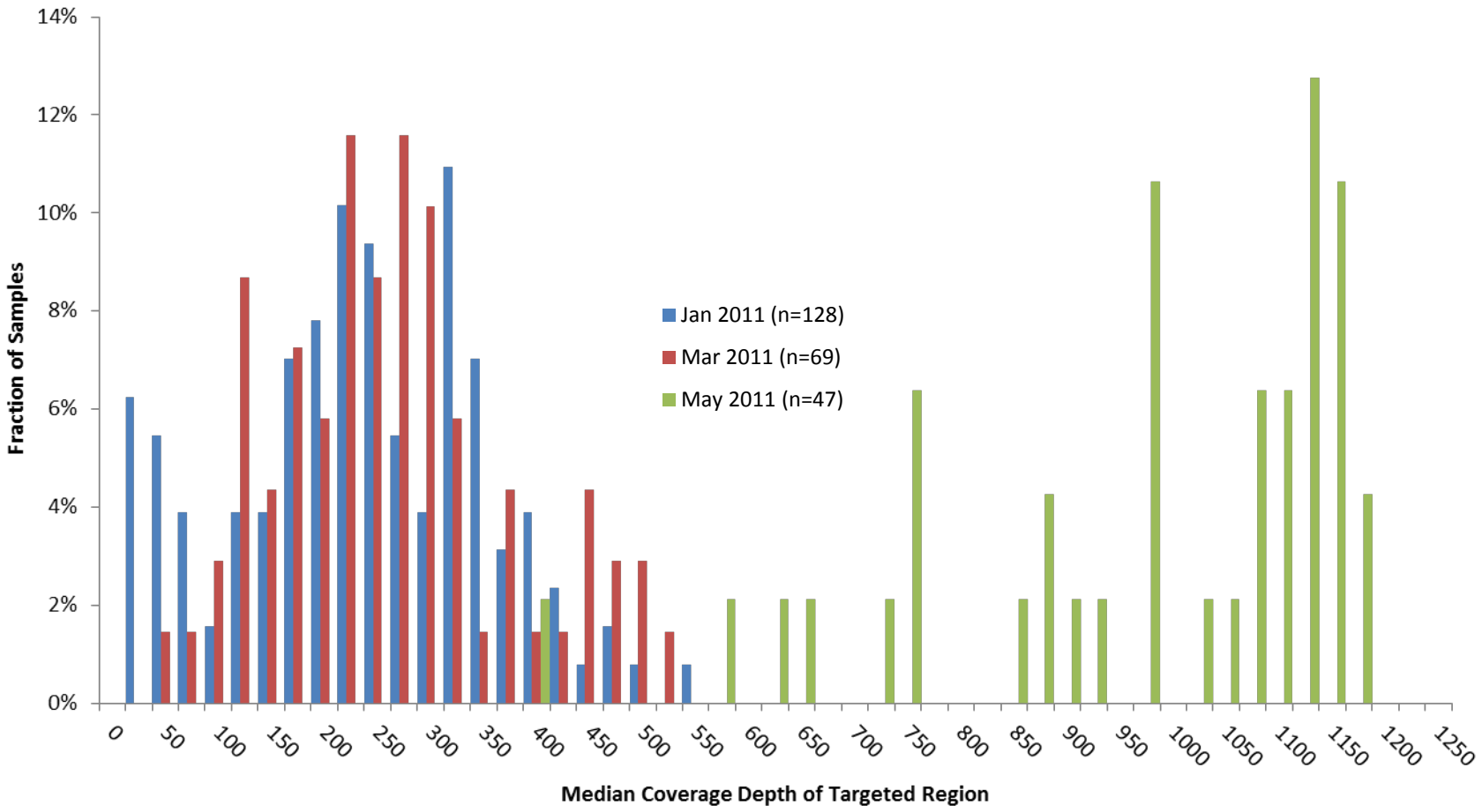


Foundation Medicine,
2011 AMP Meeting

Hybridization Capture for Targeted Sequencing

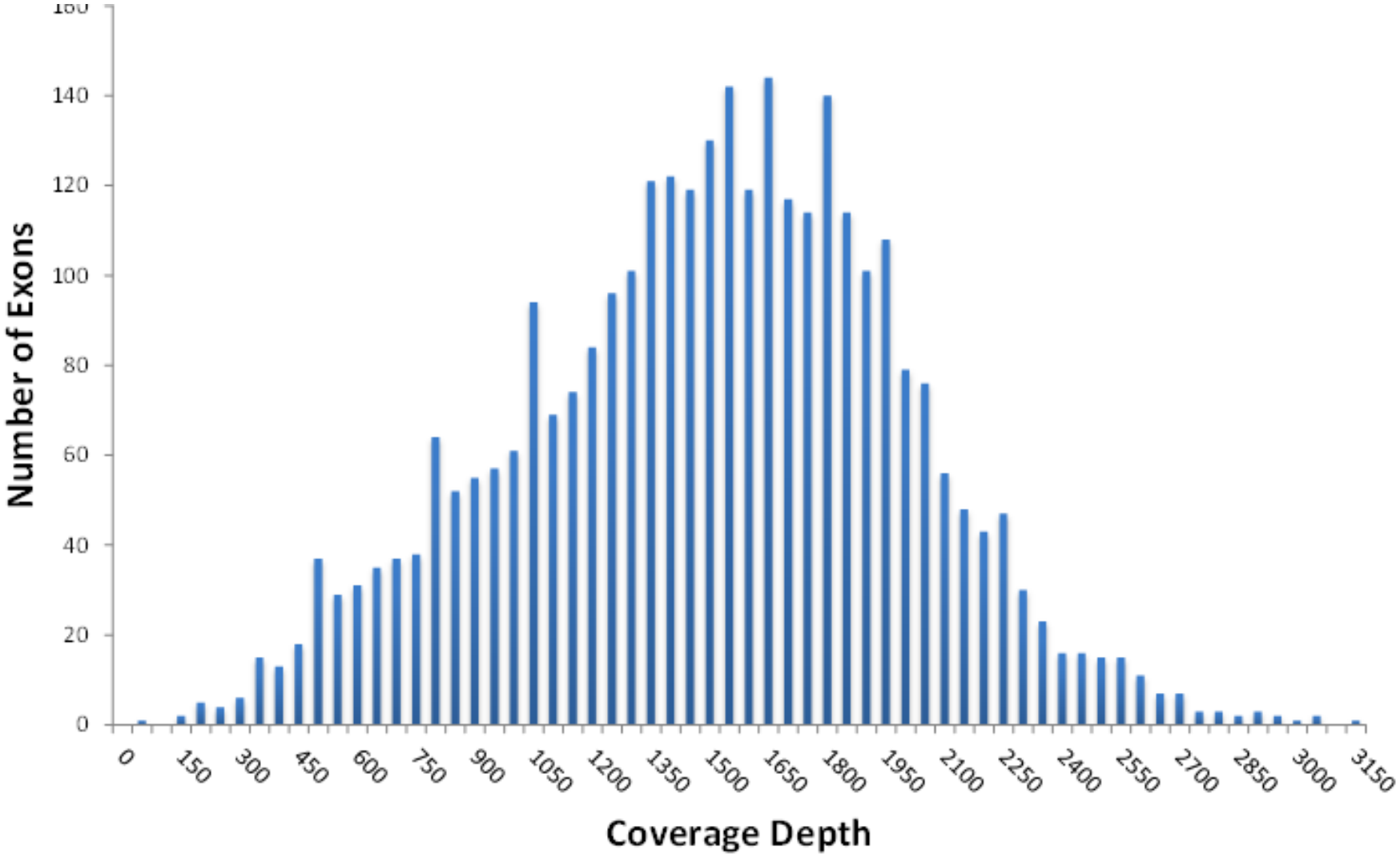
- Hybrid capture probes define test content and are empirically designed
- Hybrid capture probe pool may be RNA or DNA probes
- Hybrid capture probe coverage of targeted sequencing area must be validated under a standard set of process conditions
- Hybrid capture probe pool performance is empirically validated under a standard set of conditions
  - Standard amount of library DNA is used
  - Standardized ratio of probe qty to library qty is defined under standard capture conditions (e.g., buffer, time and temperature)
  - QC metrics defined for acceptable captured library- typically process yield
  - Process controls are typically well-characterized libraries previously successfully captured and sequenced
- Lessons Learned:
  - Manufactured quality of hybridization probe quality varies; new lots must be validated using process controls
  - Condition-sensitive hybridization capture process variables must be well controlled
  - Good genomic sequencing library construction nearly always results in good quality captured sequencing libraries

Optimization of Target Coverage



Foundation Medicine,
2011 AMP Meeting

Sequenced Target Coverage



Foundation Medicine,
2011 AMP Meeting

Sequencing Process- Lessons Learned

- Regardless of platform, standard loading conditions should be used that optimize flow cell performance
- Sample multiplexing must be balanced with required depth of coverage for each specimen
- Molecular barcodes for sample multiplexing must be pre-validated for sequencing performance
- Process controls and sequencing control loading must be consistent with interpretation requirements for process failures (e.g., lane distribution and qty loaded)
- Paired end sequencing processes must be consistent with intended test sensitivity and specificity performance
- Depth of coverage used must be validated to meet claimed assay sensitivity and specificity performances
- Acceptable QC metrics such as sequencing error rate, library complexity, unique on-target coverage and read fractions must be validated to achieve claimed performance metrics

Sequence Analysis Goals

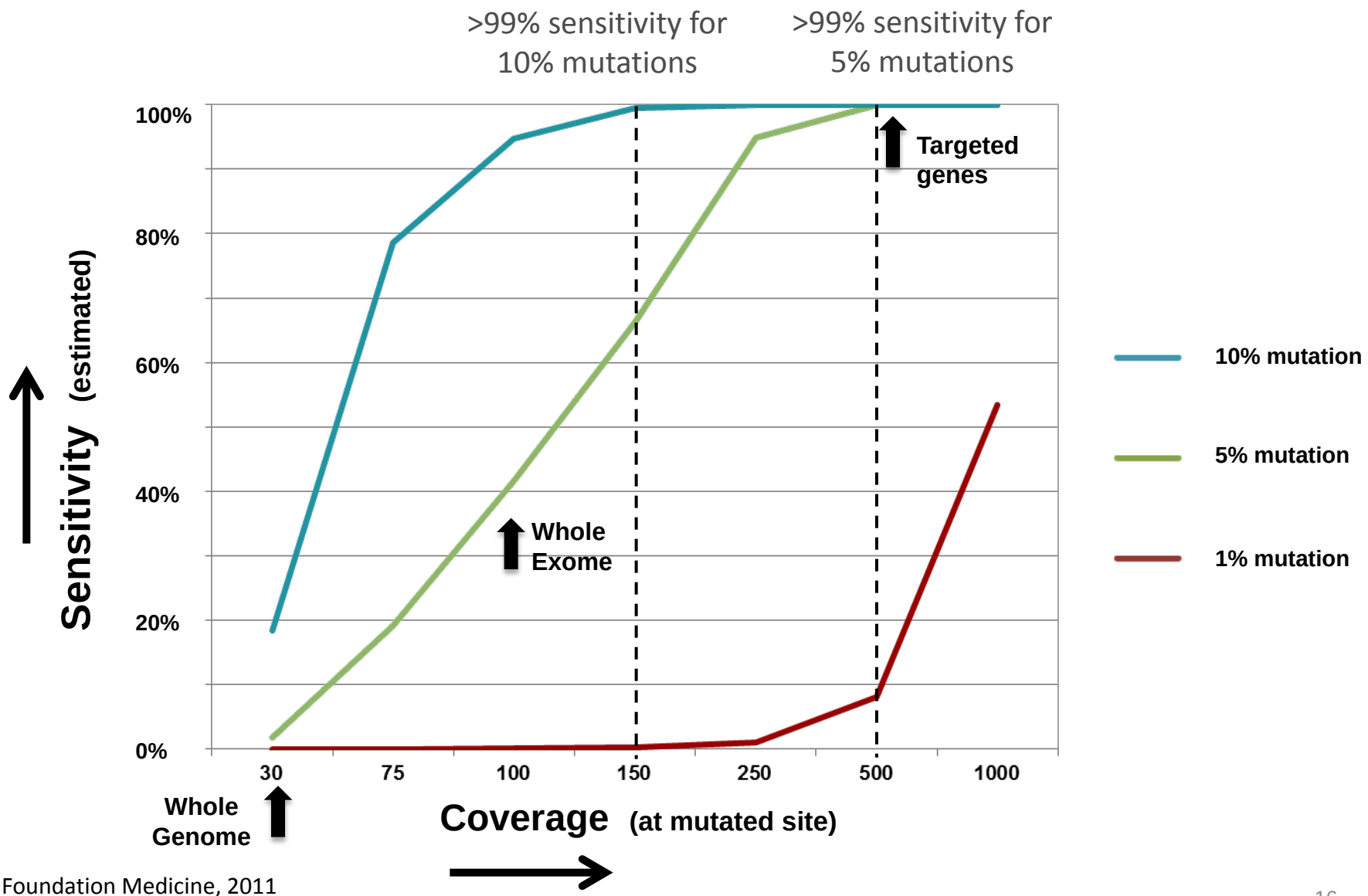
Laboratory process metrics and QC

- **Sample tracking**
- **Hybrid-capture efficiency**
- **Barcode deconvolution**
- **Sequencing yield**
- **Library complexity**
- **Coverage depth**
- **Error rates**
- **Positive & negative controls**

Mutation calling and reporting

- Mutation calling
  - Base Substitutions
  - Insertions & Deletions
  - Copy number alterations
  - Select rearrangements
- Mutation filtering
- Variation databases
  - dbSNP, COSMIC, FMI normals
- Mutation annotation
- Mutation classification

Deep Coverage Improves Mutation Detection



Foundation Medicine, 2011
Agilent Users Meeting, Boston, MA

Empirical Validation of Detection Sensitivity

- Demonstrate assay sensitivity to rare mutations (5-10%) in the sample DNA – i.e. a low-purity, multi-clonal specimen
- Sequence mix of 10 normal, individually sequenced cell-lines from 1000 Genomes project (private het SNPs simulate 5% mutations)

	Mean Seq depth	N	sensitivity	N	sensitivity	N	sensitivity	Estimated PPV
Typical sample	**707**	**121**	**96%**	**63**	**100%**	**173**	**100%**	**99.4%**
	576	114	96%	61	100%	169	100%	99.4%
	474	108	93%	59	100%	167	100%	100%
	403	104	94%	57	100%	156	100%	100%
Borderline sample	**350**	**102**	**92%**	**57**	**98%**	**151**	**100%**	**100%**

Foundation Medicine, 2011
Agilent Users Meeting, Boston, MA

Conclusions

- Pre-analytic sample characterization is important to successful process performance and final interpretation
- NGS processes are complex and require optimization of several component steps, however, the process as a whole must be controlled and validated
- Sensitive, specific, clinical-grade genomic variant calling performance depends on successful NGS process control with defined, validated performance metrics