



RainDance[™]
Technologies

Microdroplet-based PCR Amplification for Large Scale Targeted Sequencing

American Society of Human Genetics (ASHG)

October 22, 2009

James Brayer

Commercial Scientific Application Manager

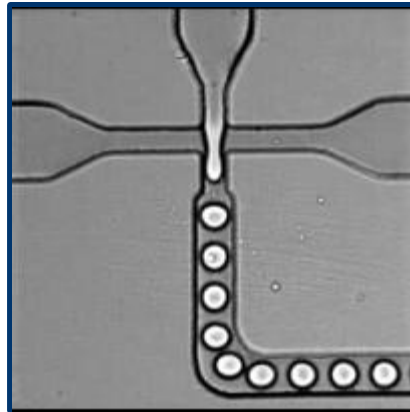


RainDance Technologies

Who we are:

RainDance Technologies is a Provider of **Microdroplet-based Solutions**

The Company's ***RainStorm™*** Technology Produces
Picoliter-sized Droplets at a Rate of **10 million Per Hour**



30 micron (16 pL) droplets: 3,000 /
second

Each droplet is the functional equivalent of an individual test tube or a single well from a microtiter plate and can contain a single molecule, cell or reaction



Targeted Sequencing

Remember why you are doing targeted sequencing

- Your research has brought you to these regions of interest
- You want to make sure that you are able to interrogate all the bases within these regions.
 - Missing data is a lost opportunity for discovery
 - Do not lose data in the shadows
 - Filtering repetitive regions



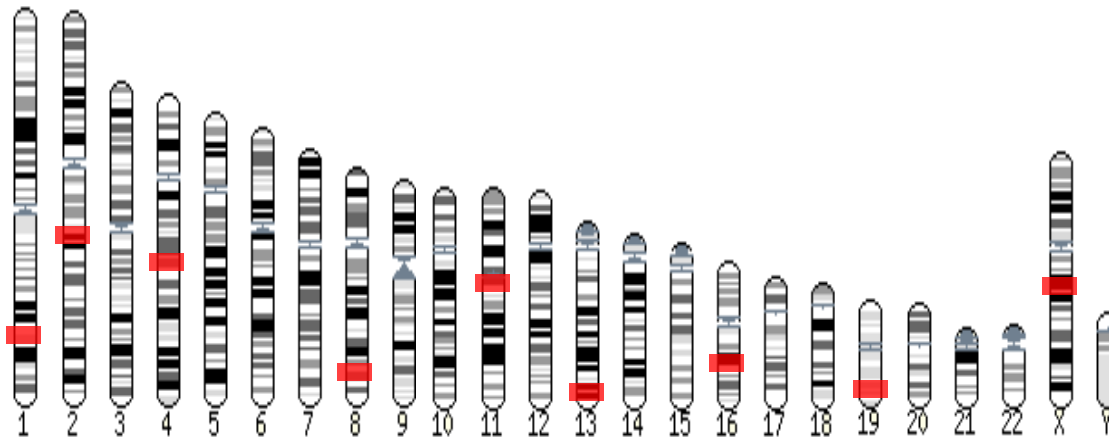


Attributes for Selecting a Platform for Targeted Selection

1. Simplicity of the Workflow
2. Capable of Targeting All Sequences of Interest
3. Uniformity of Sequence Representation (Bias)
4. Specificity of Sequence Enrichment
5. Reproducibility Across Technical & Sample Replicates
6. Accurate Sequence Representation



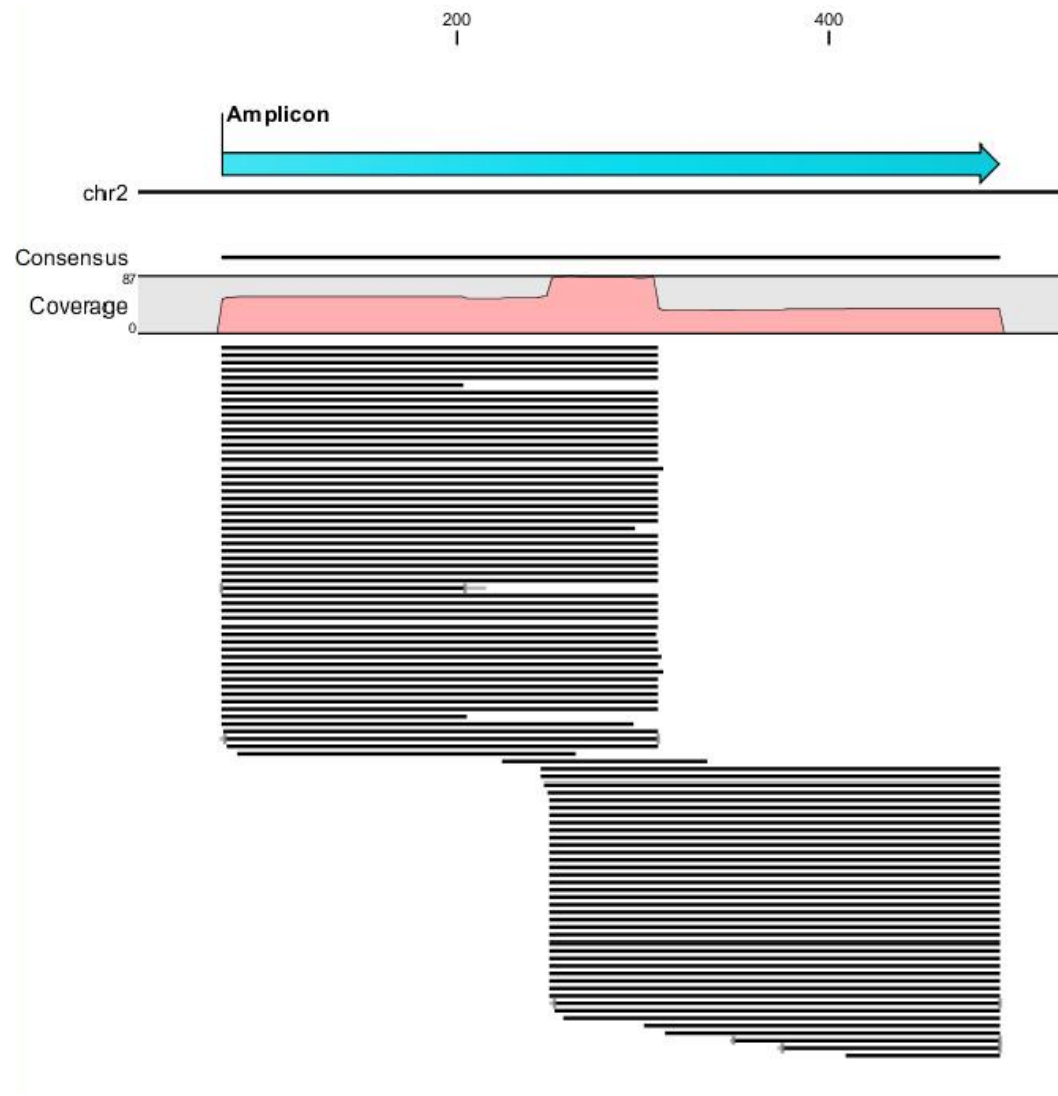
The RainDance Targeted Sequencing Solution



- Polymerase Chain Reaction (PCR) to amplify regions of interest

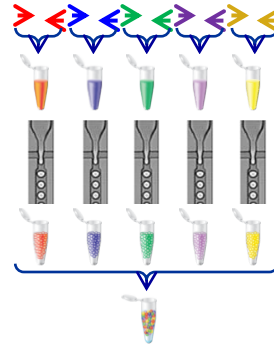
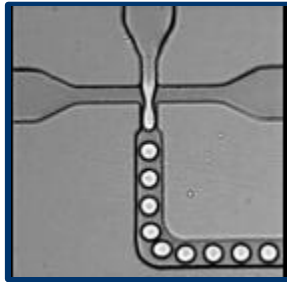


Precision of Using PCR for Targeted Selection





RainDance Sequence Enrichment Solution Reduces Bias



Genomic DNA Template Control

- Droplet Volume
- Genomic DNA Concentration

Primer Library Manufacture

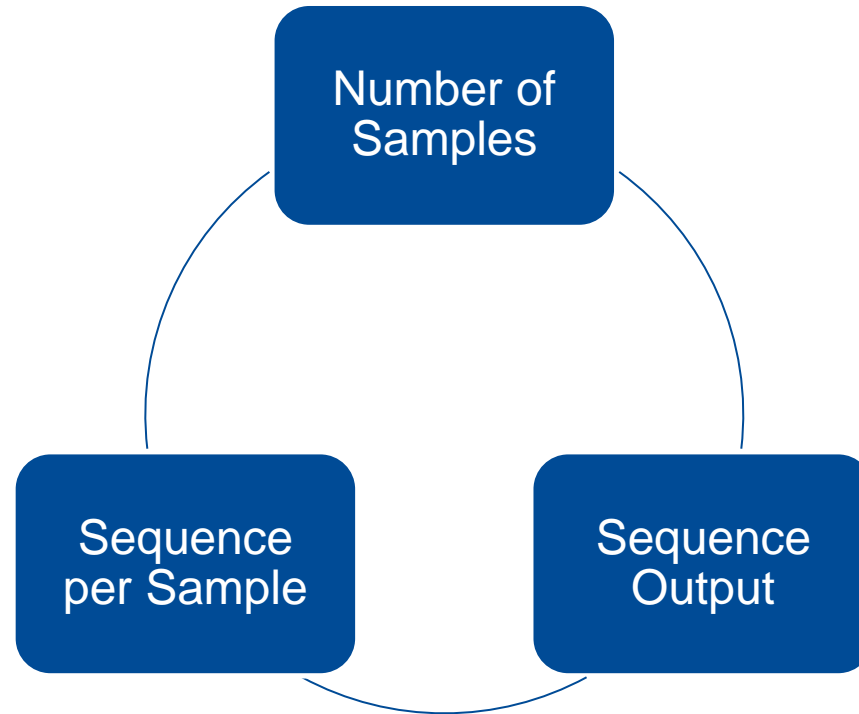
- Droplet Volume
- Primer Concentration
- Droplet Representation

PCR Thermal Cycle Profile

- Consistent amplicon yields



Properly Power your Experiments



- How much sequencing can you do?
- How much sequencing do you need?
 - Coverage and Selection bias
- How many samples do you need to look at to see your biology?



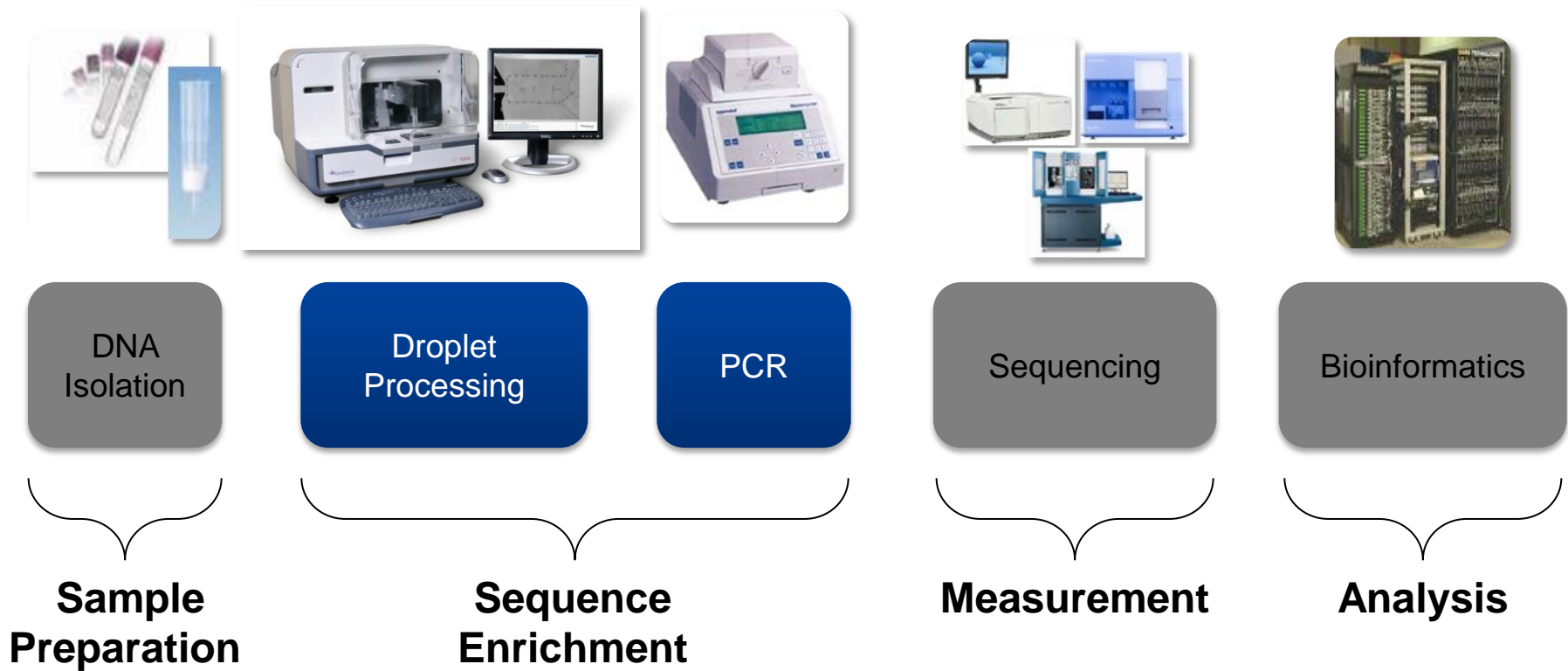
RainDance Sequence Enrichment Solution



The RDT1000 provides the specificity, efficiency and accuracy required for large-scale resequencing studies



RainDance Targeted Sequencing Workflow



Simplicity of the RainDance Sequence Enrichment Workflow



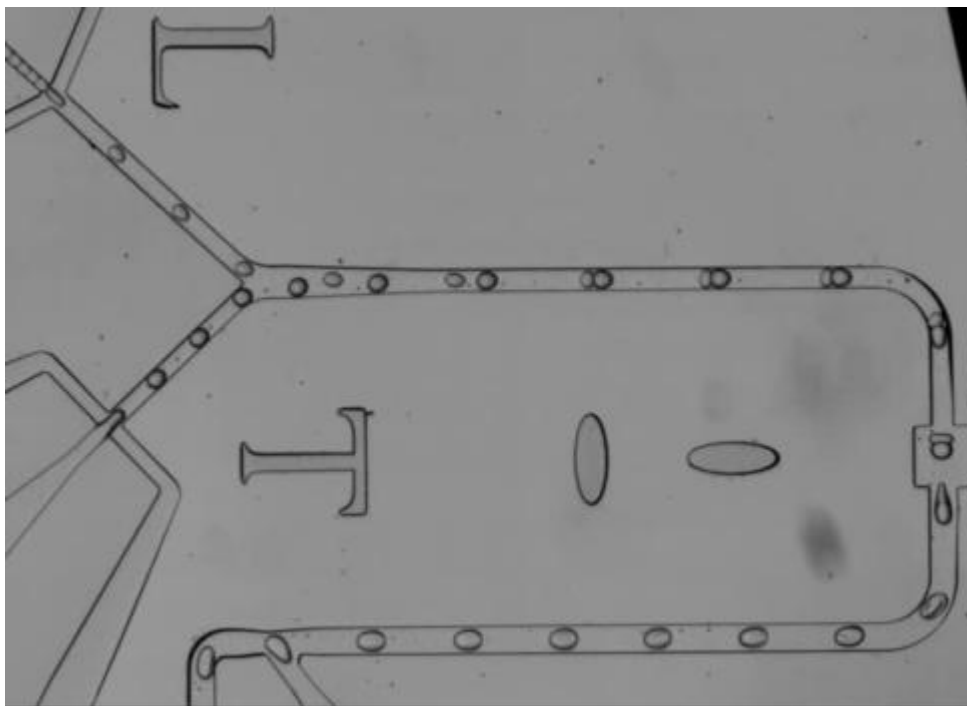
RDT 1000 Droplet-Based PCR



Primer
Library



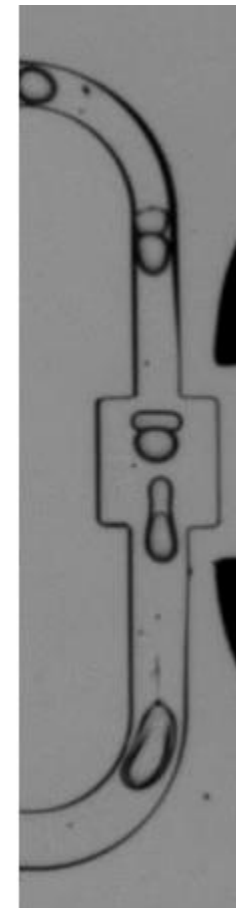
Genomic DNA
Template Mix



1 million droplets
45 minutes per run



PCR Library



Droplet
Merge Area



RainDance Custom Primer Library Design



Customer

Submit Gene List



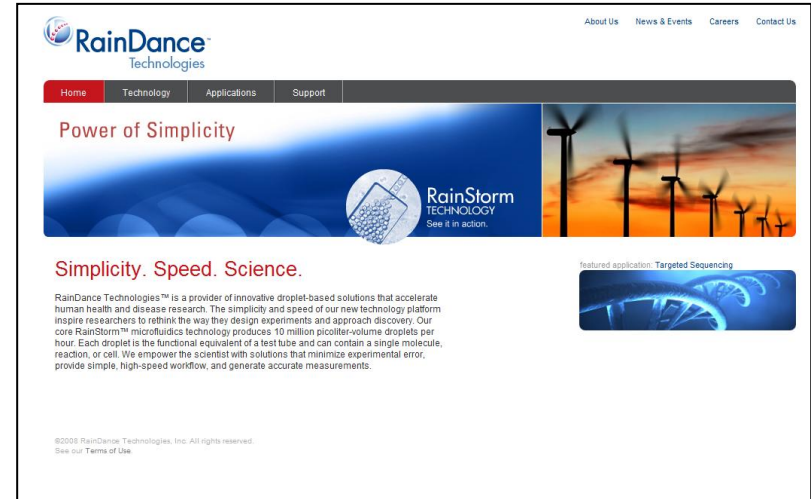
- RefSeq ID
- Gene ID
- Genomic Coordinates



Primer Library

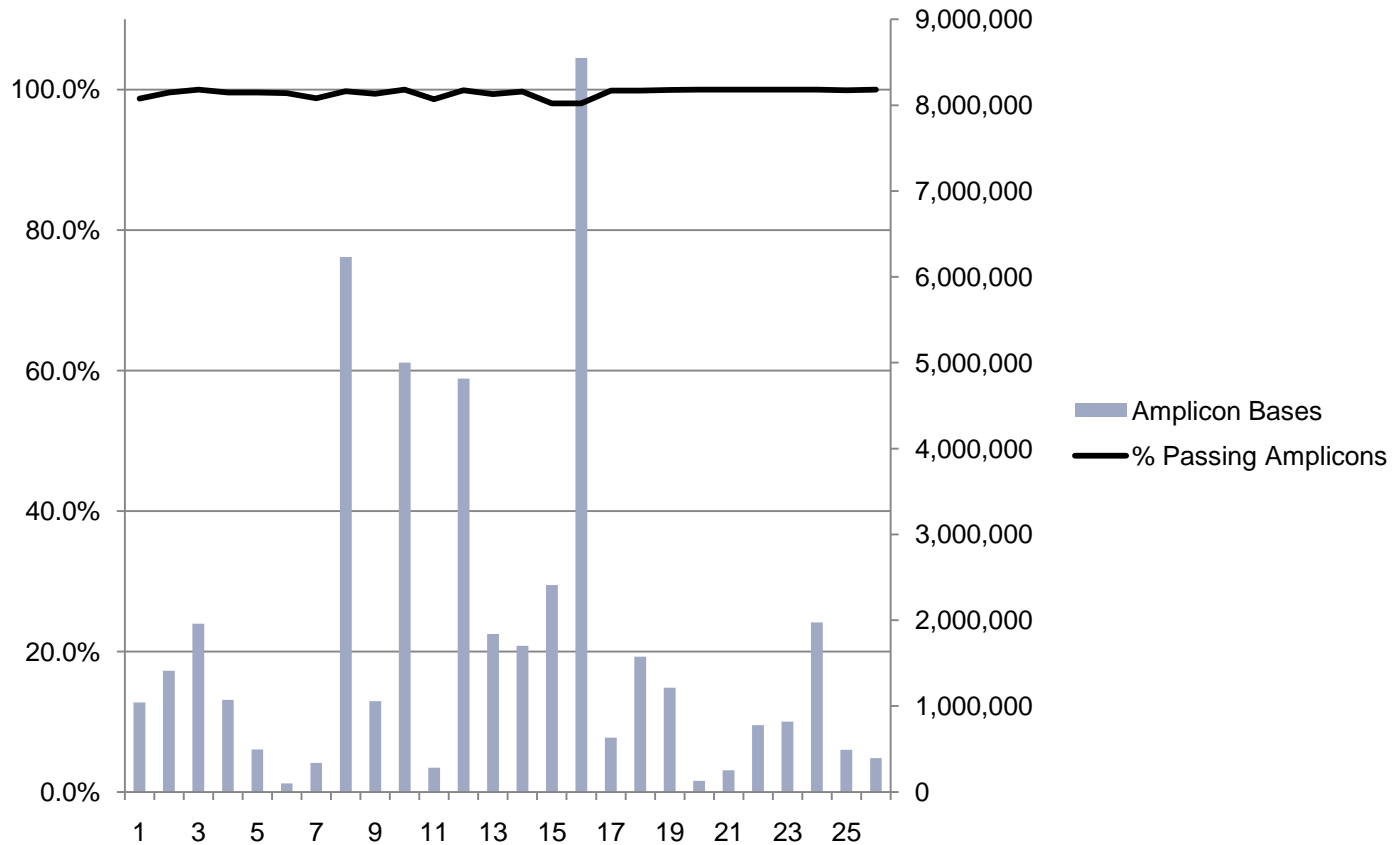


1. Primer Design
2. Oligo Synthesis
3. Primer Library Production





RainDance Primer Library Design Pipeline Performance



99.5% success rate with RainDance primer design pipeline



Genotype Concordance with Diverse HapMap Samples

Kelly A. Frazer, Ph.D.

Eric J. Topol, M.D.



SCRIPPS GENOMIC MEDICINE

A COLLABORATION OF SCRIPPS HEALTH AND THE SCRIPPS RESEARCH INSTITUTE



UNIVERSITY of CALIFORNIA, SAN DIEGO

MEDICAL CENTER MOORES CANCER CENTER



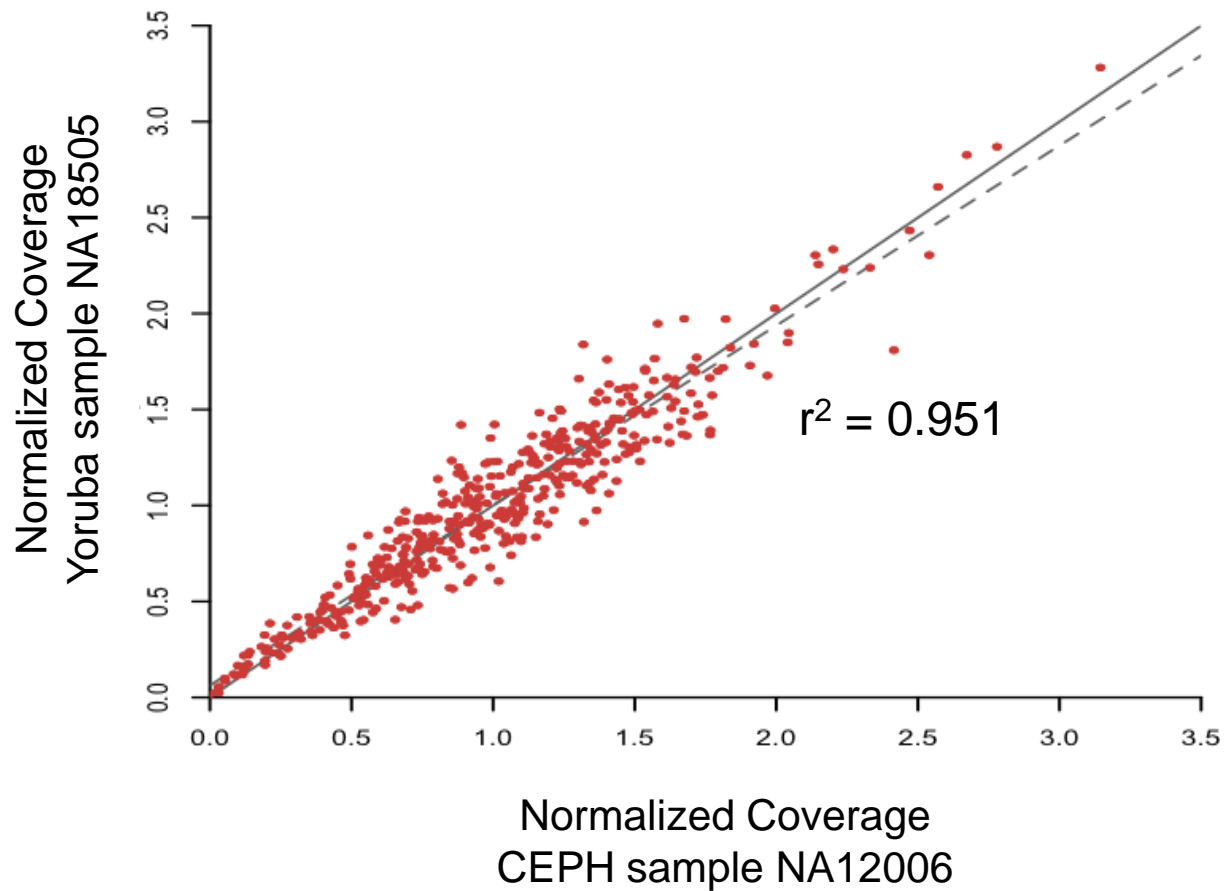
Genotype Concordance with Diverse HapMap Samples: Experimental Design

Samples	3 caucasian	technical replicate
	3 african	influence of unknown DNA variant in PCR
Targets	29 genes in ENCODE	well annotated region sequence available in 5/6 samples
	8 TRP channel genes	ability to target related sequences
	11 genes in venous thrombosis	
	435 exons - 457 amplicons	varying amplicon length and GC content

- 457 amplicons (172 kb)
- Six indexed samples per lane on Illumina GAI (36 bp reads)



Genotype Concordance with Diverse HapMap Samples: Reproducibility

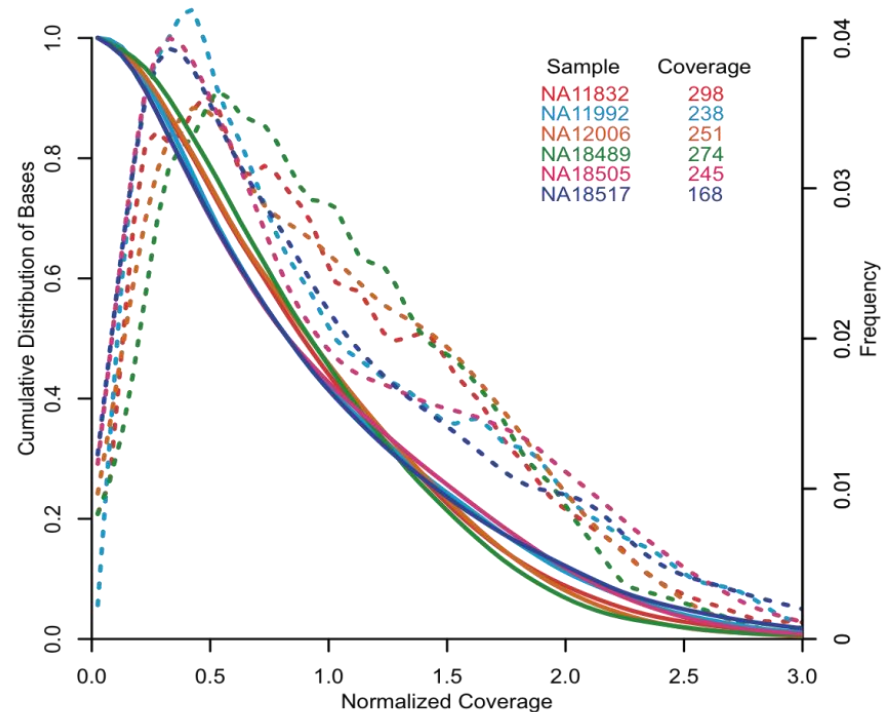




Genotype Concordance with Diverse HapMap Samples: Uniformity of Coverage

Normalized coverage – divided the observed coverage of each base by the mean coverage of all targeted bases for each indexed sample

- 89.6% of all bases covered by at least $\frac{1}{4}$ the mean coverage
- 99.6% of all bases covered by at least one read
- Ran all 6 samples in a single lane of an Illumina GAI





Genotype Concordance with Diverse HapMap Samples: Accuracy

- 99.1% concordance with HapMap
- 50% discordance due to HapMap mistakes
 - (sequencing traces checked)
- <0.1% due to allelic bias
- Not sensitive to the presence of unknown variants
 - No difference between ENCODE and non-ENCODE region
 - No difference between Caucasian and African samples

The RDT 1000 provides the specificity, efficiency and accuracy required for population-based resequencing studies



RainDance[™]
Technologies

Comparison to Hybridization Capture



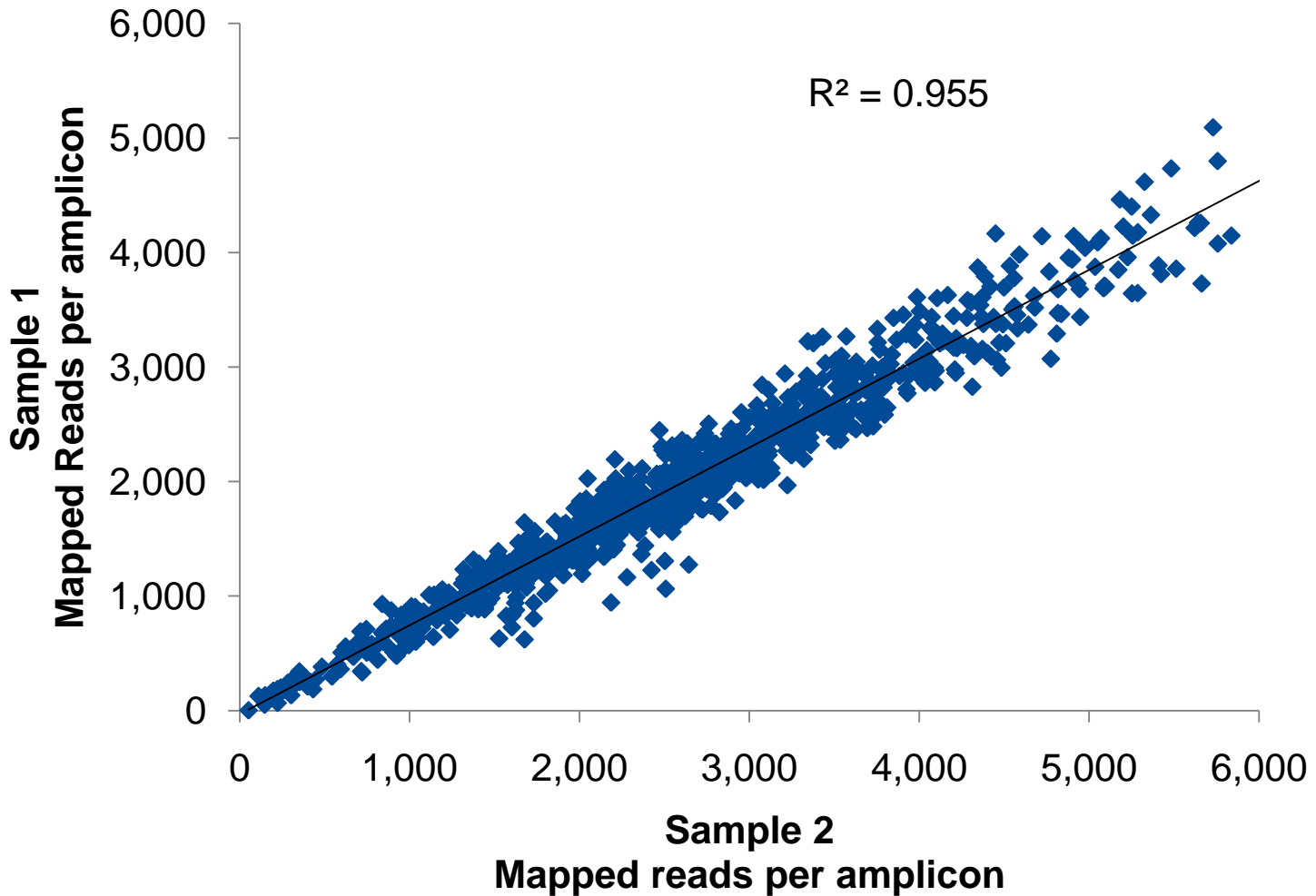
Targeted Resequencing of 200kb Locus

- 200 kb locus of chromosome 9 (*CDKN2A+B*)
- 979 tiled amplicons
 - Average amplicon length 450 bp
 - Average tile overlap 200 bp
- Coriell sample NA15501
- Analysis of 36bp single end reads on Illumina GAI
- Comparison to solution-phase hybridization capture¹

¹Gnirke, et. al. Nat Biotechnol. 2009 Feb;27(2):182-9

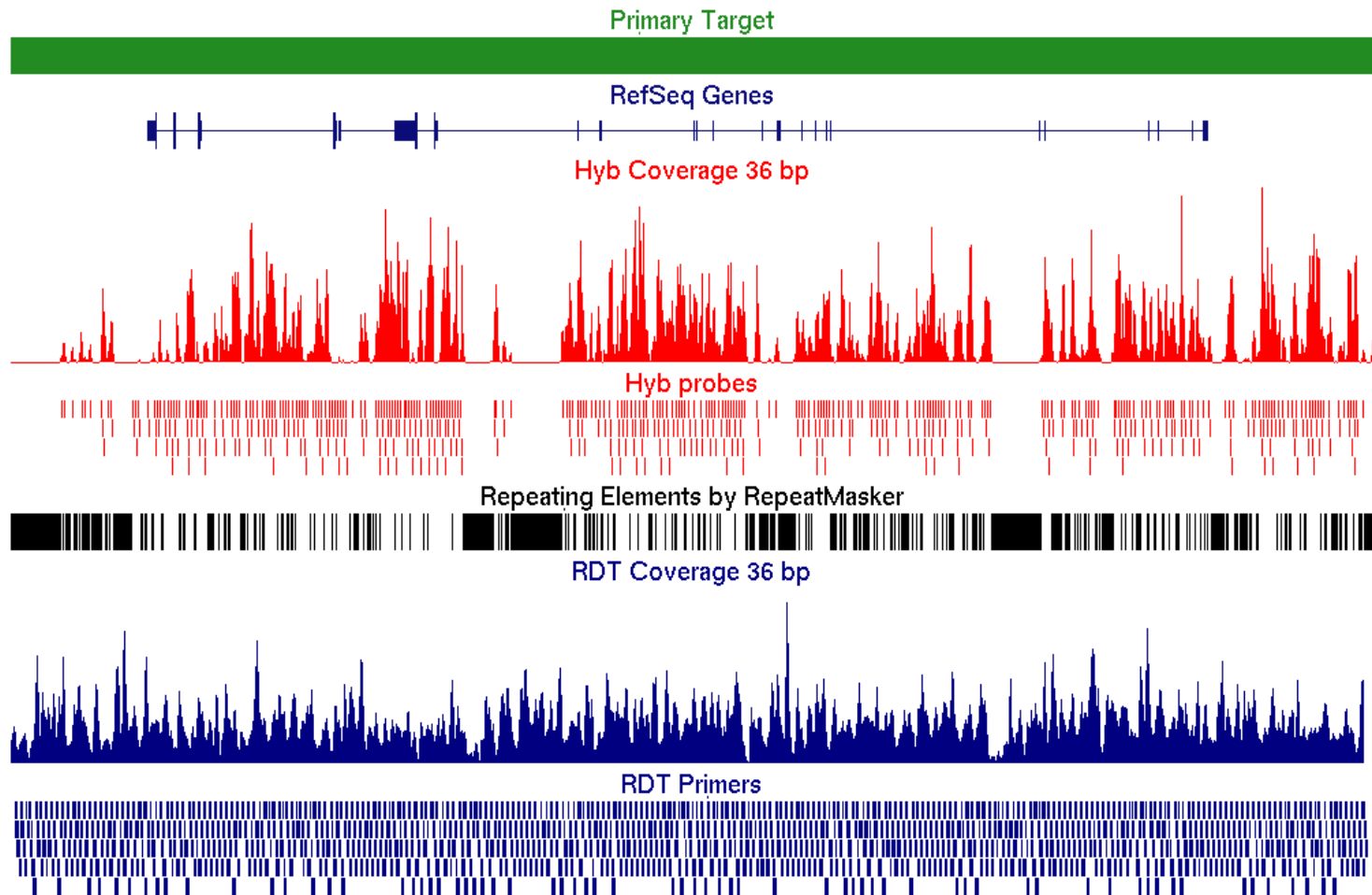


Reproducibility of Coverage with RDT 1000





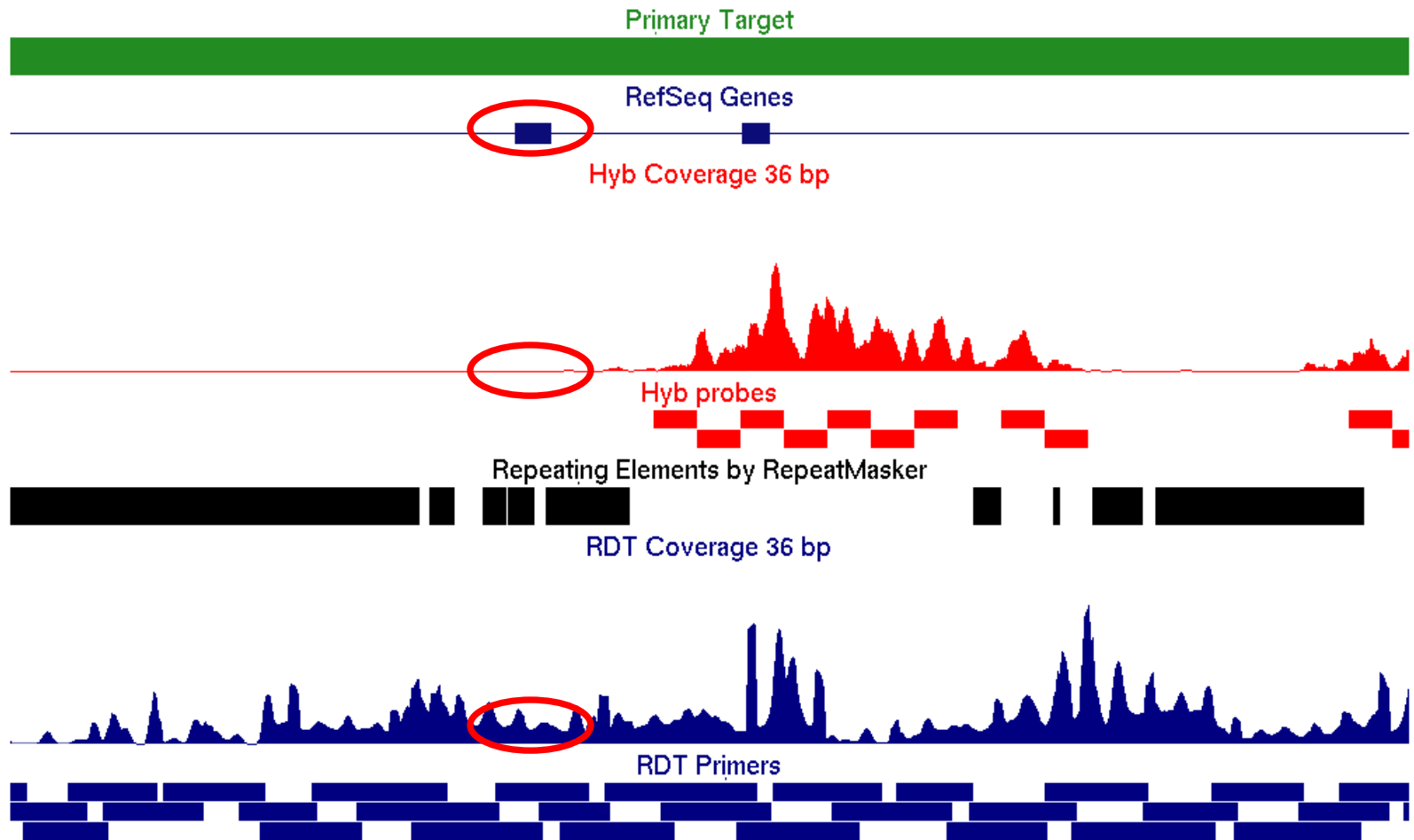
Targeted Resequencing of 200kb Locus



RainDance Coverage: 99.8%
Hybridization Coverage: 63.7%



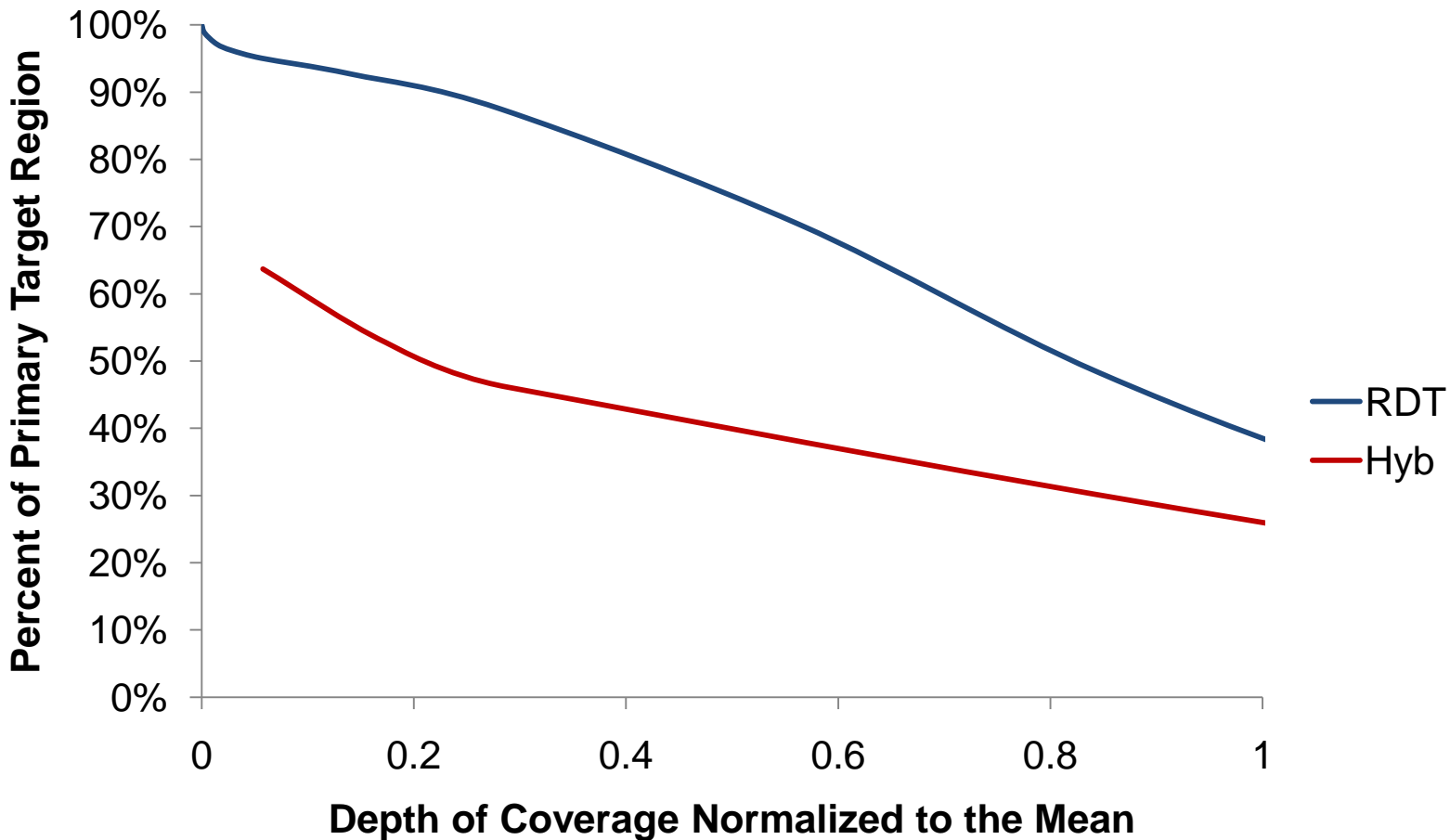
Targeted Resequencing of 200kb Locus



RainDance covers regions that hybridization misses



Targeted Resequencing of 200kb Locus



RainDance Coverage: 93% at 0.2X mean
Hybridization Coverage: 49% at 0.2X mean



Longer Reads Improve Uniformity

Primary Target



RefSeq Genes



Repeating Elements by RepeatMasker



RDT Coverage 36 bp



RDT Coverage 100 bp

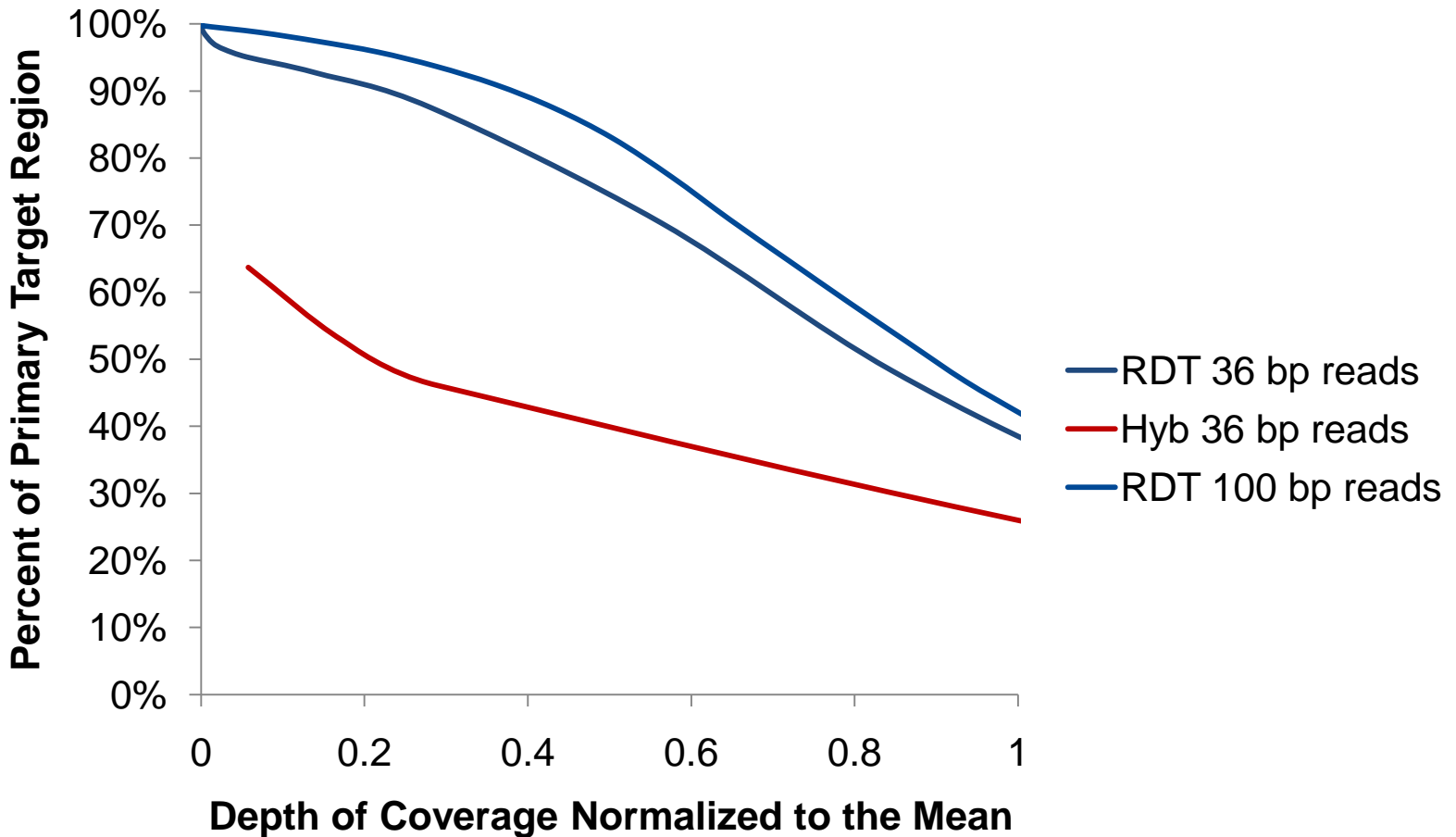


RDT Primers





Targeted Resequencing of 200kb Locus



100 bp reads increases RDT coverage to 97% of primary target region within 0.2X mean



RainDance[™]
Technologies

RDT1000 Droplet Statistics

“What can we do with 1 million PCR droplets”



PCR Reactions per Run

2 ug input genomic DNA



- fragment to 2-4 kb
- purification and QC

1.5 ug fragmented Genomic DNA Template Mix (25 uL)



- 20 uL aspiration volume
- 18.8 pL per Genomic DNA Template droplets

- 20,000,000 pL/18.8 pL

1,063,893 Genomic DNA droplets generated



- 85% minimum merge efficiency, > 90% typical
- 27% PCR amplification

244,163 PCR positive droplets

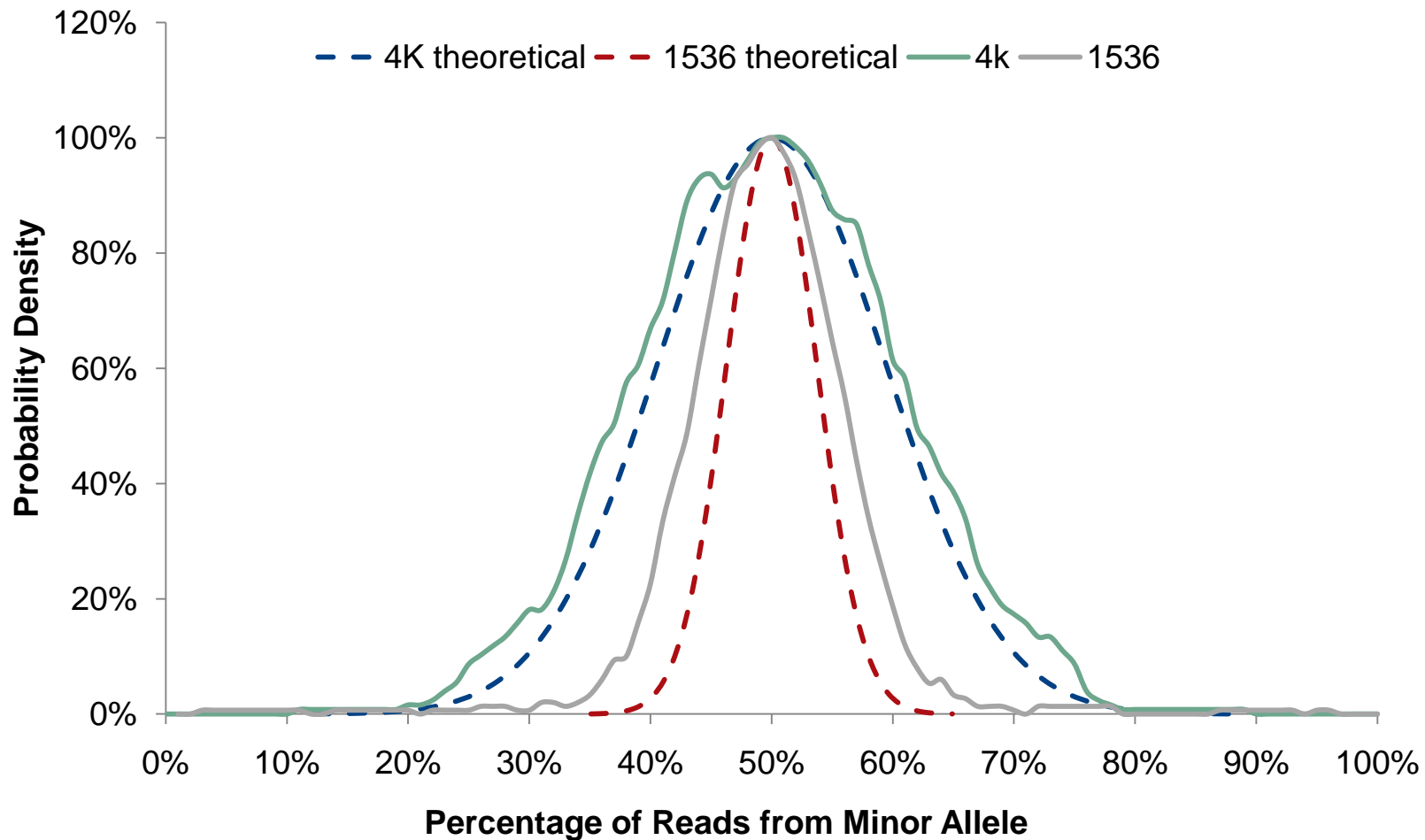


RainDance Droplet Statistics

Assumptions: Droplet Sampling (4,000 Primer Pairs)	
1. n = number of droplets	247,272
2. p = probability (4,000 primers)	1/4,000
3. mean = np	62
4. standard deviation ($\sqrt{np(1-p)}$)	7.86
5. minimum # droplets (99.6% CAM)	38



Validating Model for Heterozygous SNPs





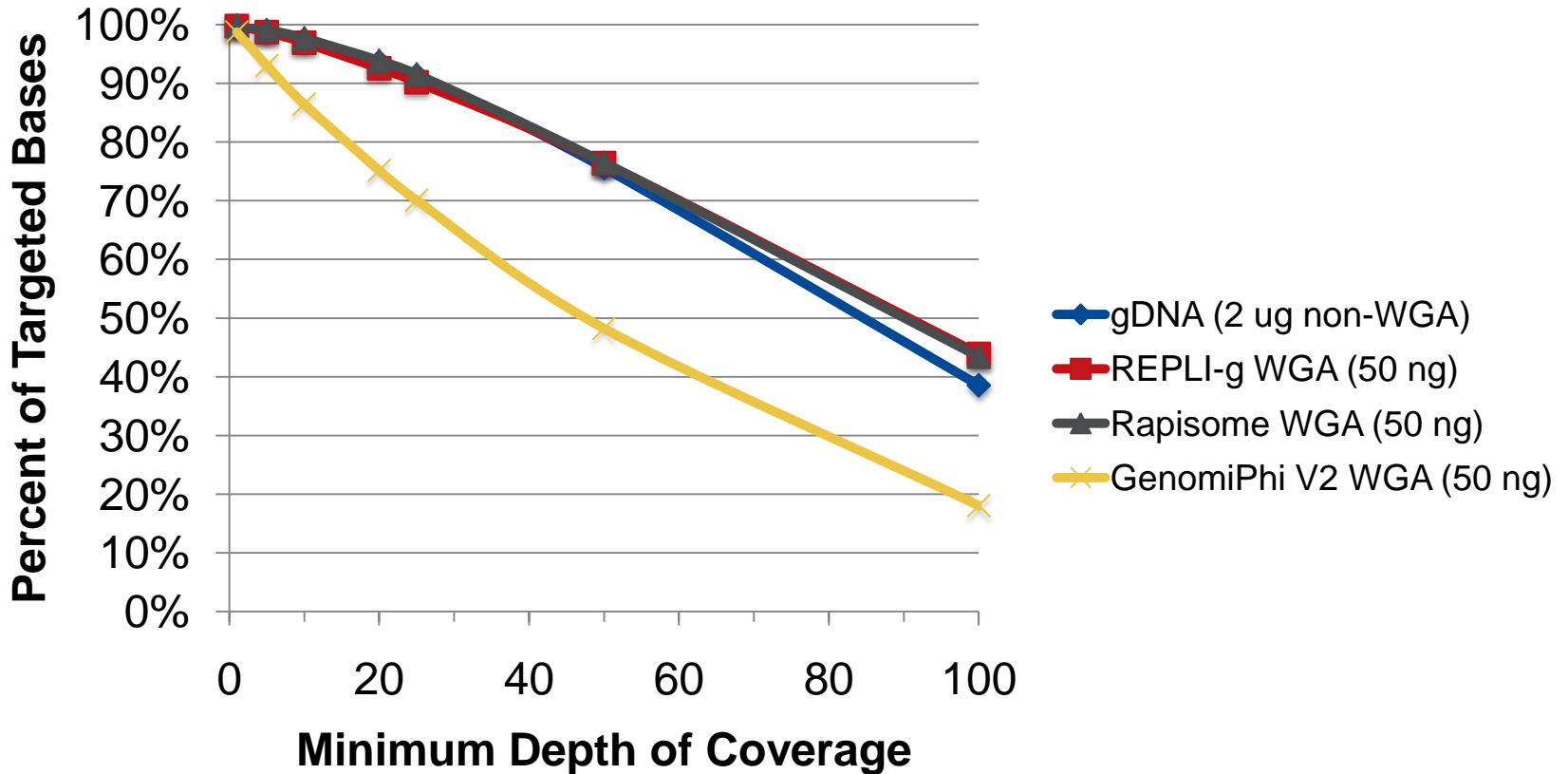
RainDance[™]
Technologies

Whole Genome Amplified DNA



Targeted Sequencing with WGA Template

Established protocol for processing WGA DNA for sequence enrichment application



Performance maintained with WGA Samples



Targeted Sequencing with WGA Template

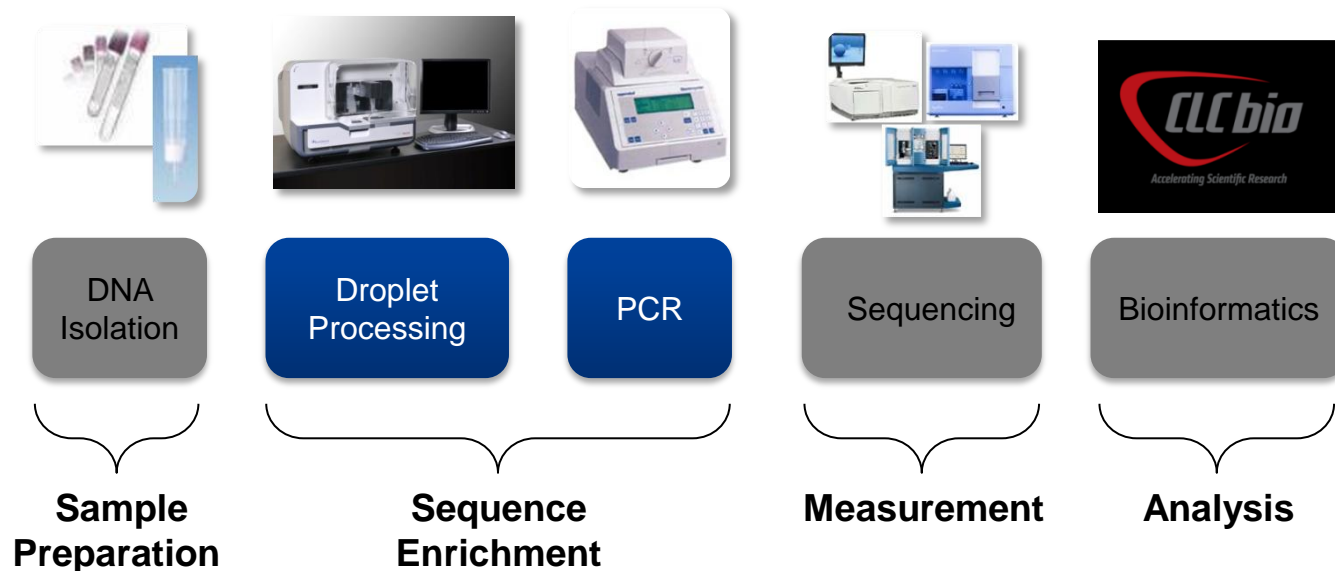
Sample Type	Overall Concordance to HapMap	Homozygous SNP Concordance	Heterozygous SNP Concordance
No amplification	98.7%	99.1%	97.1%
Qiagen Repli-G Mini	98.7%	99.0%	97.1%
BioHelix Corp. Rapisome pWGA Kit	98.7%	99.0%	97.1%
GE illustra GenomiPhi V2 DNA Amplification Kit	98.3%	98.7%	96.4%

Performance maintained with WGA Samples



RainDance Targeted Resequencing Data Analysis

- RainDance & CLC bio partnership delivering optimized software for targeted resequencing
- Expansion of CLC's Genomics Workbench
 - Rapid analysis & reporting for targeted resequencing
 - Support for all sequencing platforms (454, Illumina, SOLiD, Helicos)





RDT 1000 Product Road Map 2010



- **Expanded content**
 - 20,000 amplicons
 - 12 Mb amplicon sequence
- **Methyl-SEQ**
 - High resolution analysis of 5-MeC in targeted regions
- **Ultra-deep resequencing**
 - Increased sensitivity to detect variants in heterogeneous samples

Multiple applications with the same platform



RainDance Targeted Sequencing Solution

- ✓ Simplicity of the Workflow
- ✓ Capable of Targeting All Sequences of Interest
- ✓ Uniformity of Sequence Representation (Bias)
- ✓ Specificity of Sequence Enrichment
- ✓ Reproducibility Across Technical & Samples Replicates
- ✓ Accurate Sequence Representation



RainDance Sequence Enrichment Solution



Simplicity. Speed. Science.