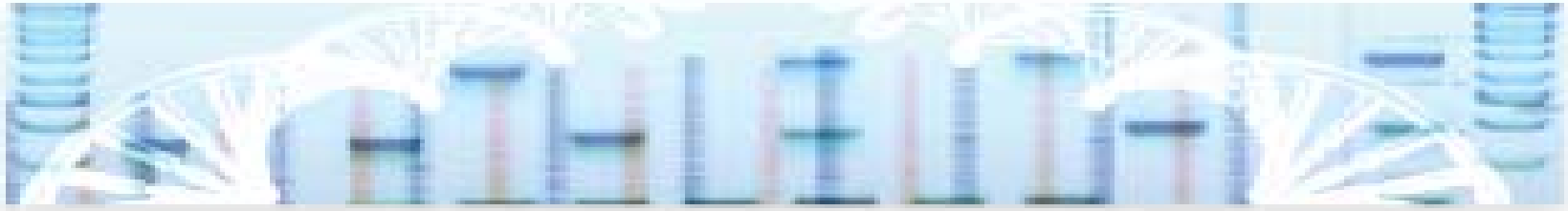
Customized Strategies for Targeted Genotyping, Expression and Methylation Studies in Human, Mitochondrial and Non-Human Genomes.

BeadXpress Symposium, June 22, 2009

Tera Eerkes, Ph.D. – iGenix, inc.

iGenix, inc.

iGenix, inc.



High Quality Laboratory and Informatics Services for:

- -High throughput genotyping
- -Gene expression
- -Methylation studies
- -Copy Number Variation

Focus on project design and data analysis

Key Issues Addressed:

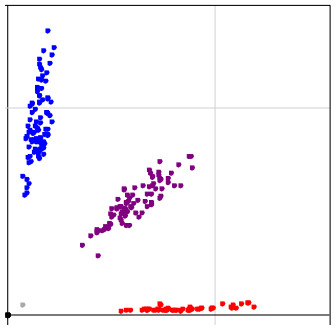
- The BeadXpress is a highly functional and flexible platform.
- At least 95% probe design conversion for all assays.
- > 99% accuracy.
- > 99% reproducibility.
- Costs per sample: **$3-$100** depending on assay.
- Rapid throughput and turn-around.

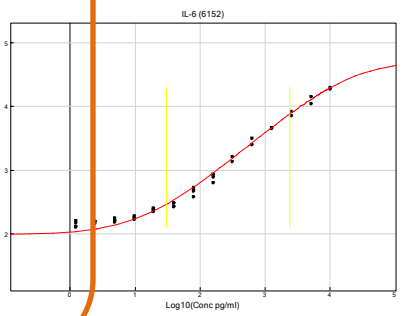
Project Design Dictates All

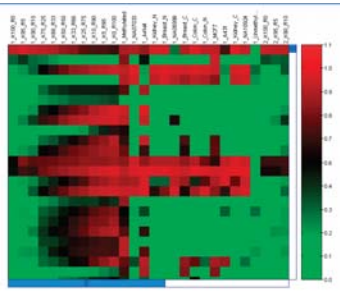
Flexible Solution Structure:

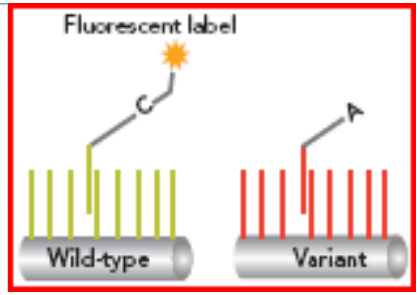
Proper Project/Probe Design

Genotyping | Expression | Methylation | ASPE







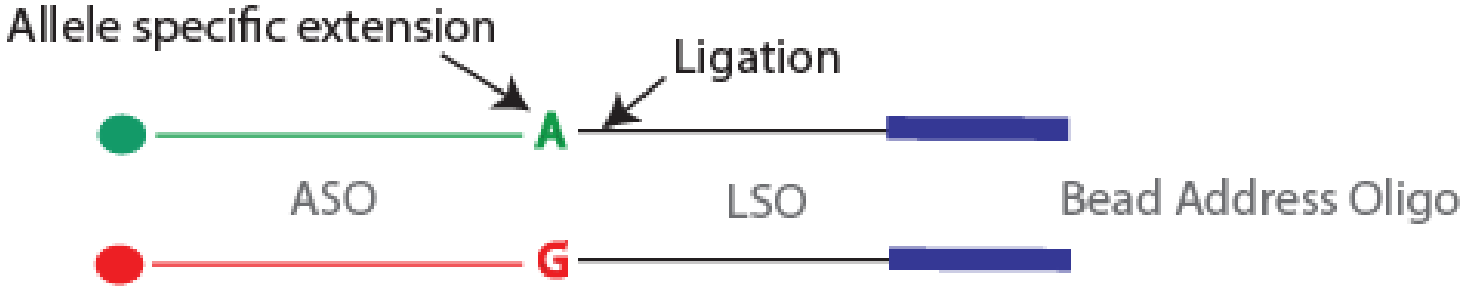


Illumina Veracode

Veracode GoldenGate Basics:

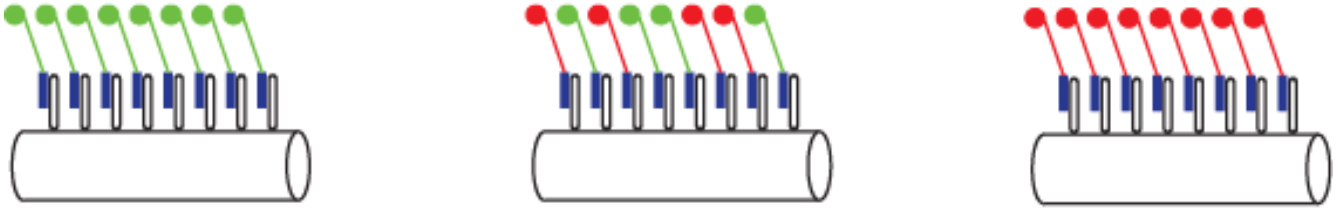
Design Probes for DNA Template

Extend, Ligate, Amplify and Label



Veracode Bead Hybridization

Hybridize to Veracode Beads, via Bead Address Oligo



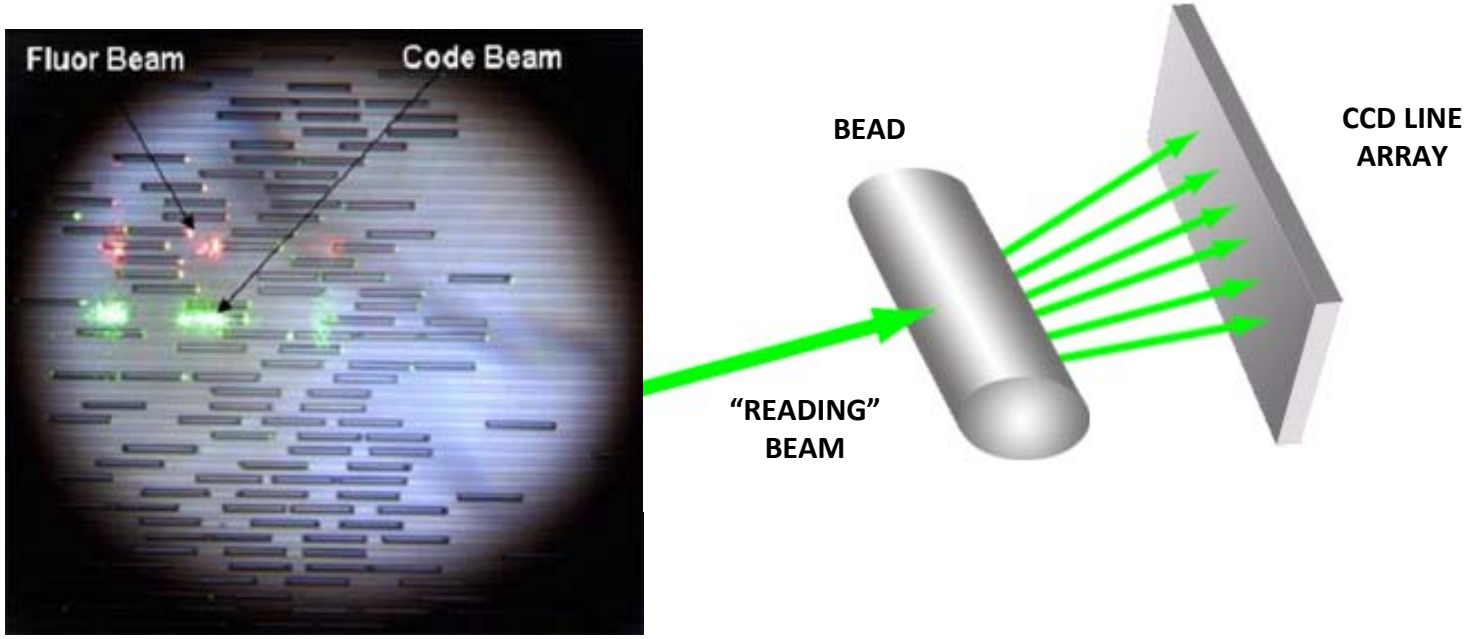
Measure and Quantitate Fluorescence

All Cy3,
Sample 1 = TT

Mixed
Sample 2 = TC

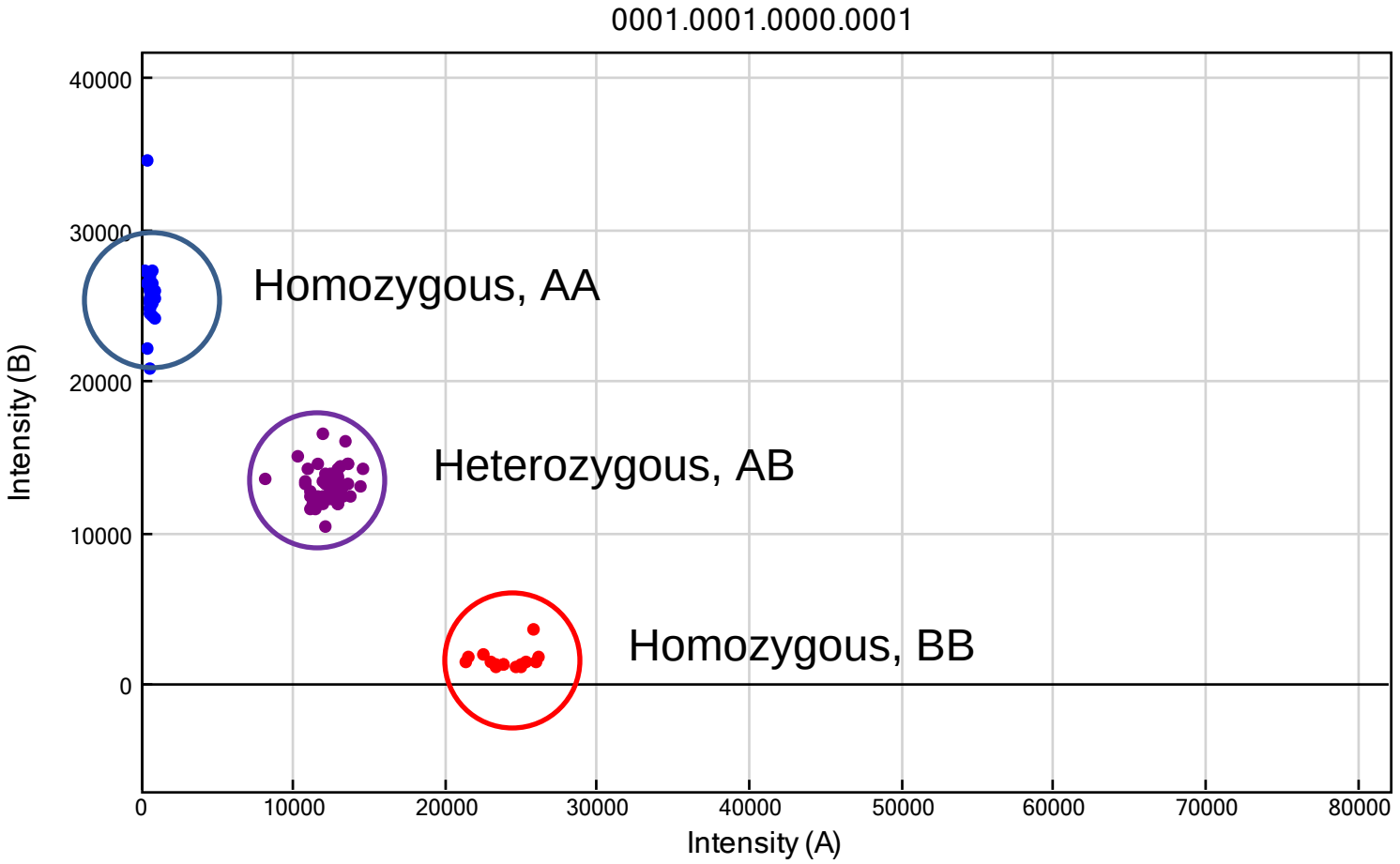
All Cy5
Sample 3 = CC

BeadXpress Measurements



Used with permission of Illumina Inc.

Raw Cartesian Data



Results:

- Automated genotype calling.
- Custom report generation.
- QC/QA statistics built in.

GGGT 384 demo_FinalReport

	A	B	C	D	E	F	G	H
1	[Header]							
2	BSGT Version	1.0.10						
3	Processing Date	6/6/2009 13:45						
4	Content		GS0000376-OPA.opa					
5	Num SNPs	376						
6	Total SNPs	376						
7	Num Samples	95						
8	Total Samples	96						
9	[Data]							
10	SNP Name	Sample ID	Allele1 -	Allele2 -	GC Score	GT Score	Chr	Position
11	rs1397354	GS0015000-DNAA01-NA12753	G	G	0.916	0.916	2	2.15E+08
12	rs649593	GS0015000-DNAA01-NA12753	G	G	0.894	0.894	1	37635225
13	rs911903	GS0015000-DNAA01-NA12753	A	G	0.8604	0.8604	1	46982589
14	rs1981635	GS0015000-DNAA01-NA12753	C	C	0.9407	0.9407	4	10353900
15	rs898249	GS0015000-DNAA01-NA12753	A	C	0.9108	0.9108	8	21183933

Genotyping Case Study:

- Needed fine mapping of regions with high LOD scores with 96 custom probes in Bovine.

-Had a bioinformatics person in the lab, but not experienced with probe design.

-iGenix provided multiple rounds of probe design and analysis.

-Last minute “save” when discovered that 50% proposed SNPs were likely sequencing errors.

Probe Design Process:

iGenix Probe Design and Conversion

Conversion rates based on GenTrain Scores (>.25)		
	Literature based studies	**iGenix**
Human Genome	91%	**>99%**
Non-Human	77%	**>90%**

Proprietary process that includes complete design of entire probe set.

Analysis extends that recommended by Illumina tech. support, especially for non-human genomes.

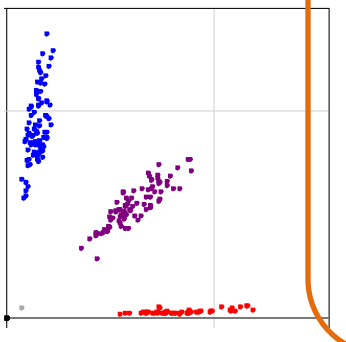
iGenix GoldenGate Assay Statistics:

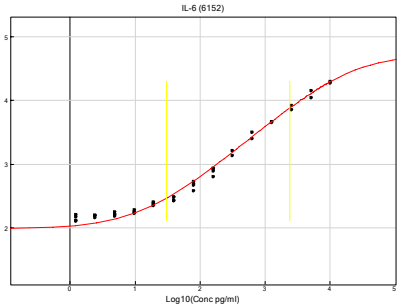
96-plex, 480 samples	
	Genotyping
Conversion	>99%
Reproducibility	99.90%
Cost/sample	$28.00
DNA Input	100-250ng

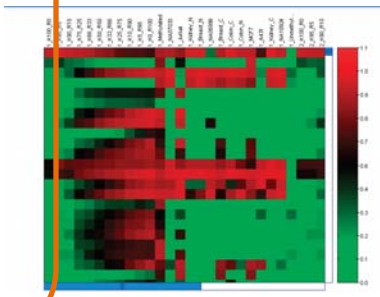
Case Study 2 - Expression:

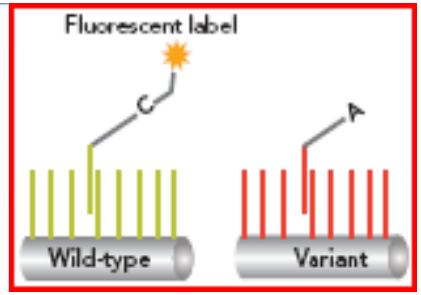
Proper Project/Probe Design

Genotyping | Expression | Methylation | Allele-Specific



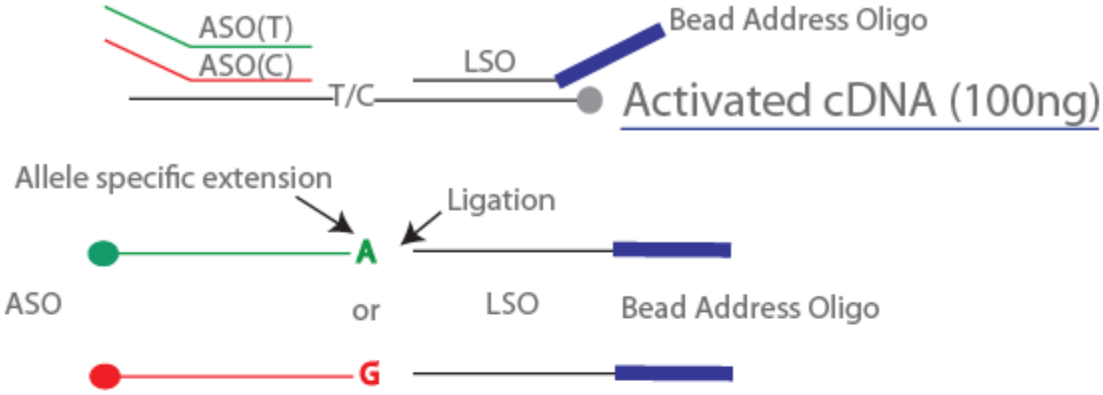




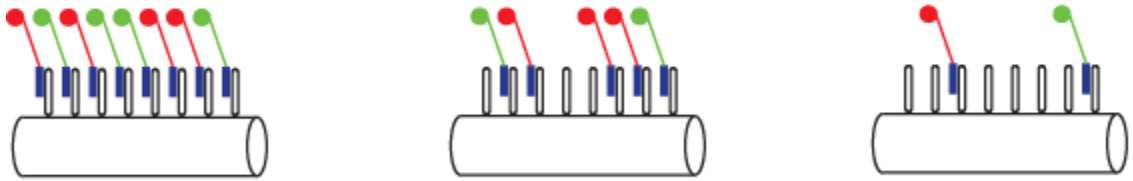


Illumina Veracode

Expression (DASL):



Hybridize to Veracode Beads, Based on Address Oligo



Measure and Quantitate Fluorescence

Project Parameters:

- Compare FFPE and flash-frozen Prostate tumor Samples in 4 assays.

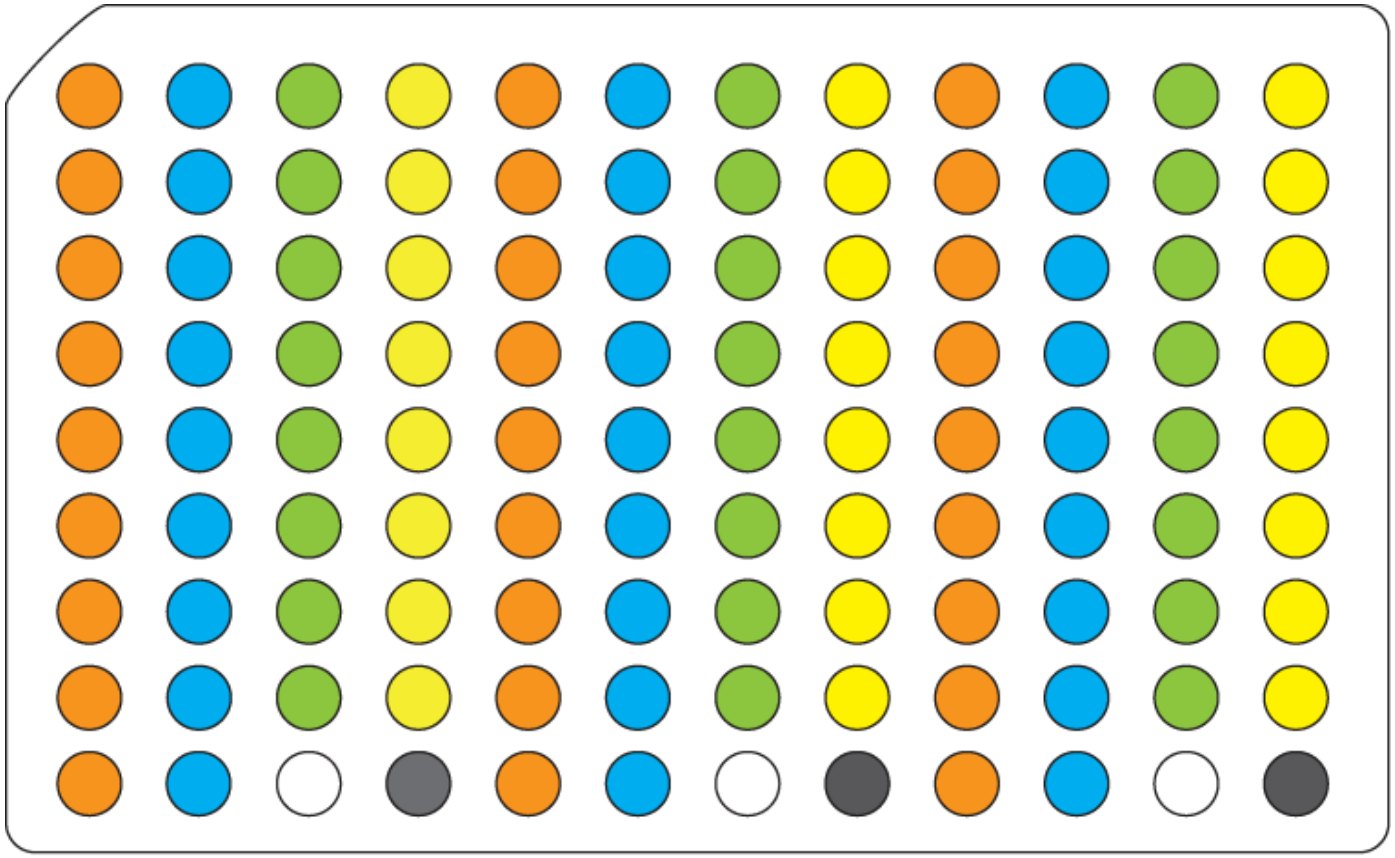
DASL (VeraCode)

DASL (ArrayReader)

TaqMan RT-qPCR

Genome Analyzer

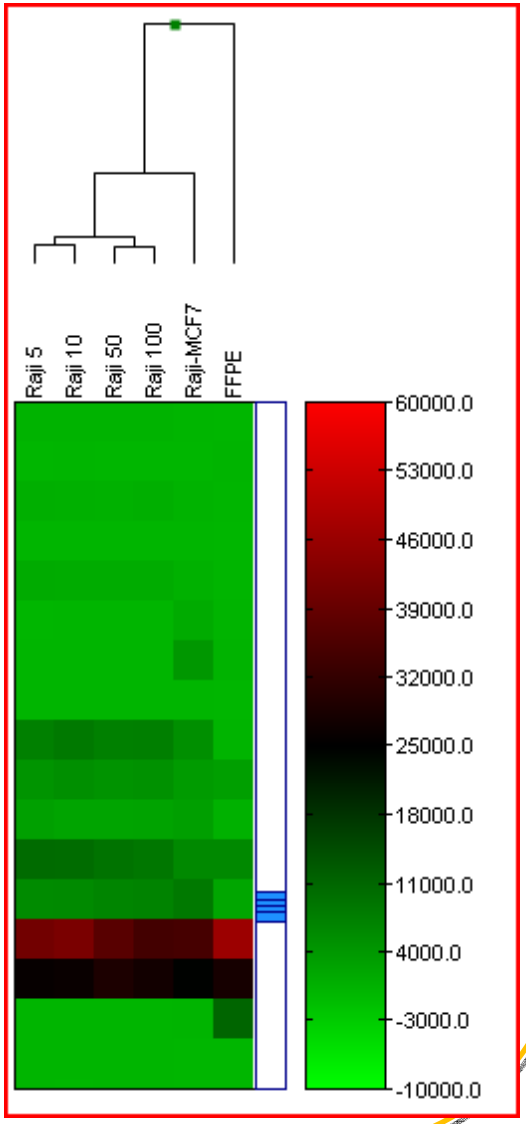
Experimental Set-up on VeraCode:



- Intact Commercial RNA
- Degraded Commercial RNA
- Flash Frozen Prostate RNA
- FFPE Prostate RNA

Customized Result Format:

Gene Probe	Sample 100.AVG_Signal	Sample 100.DiffScor	Sample 100.Diff Pval
7A5	2468.135	0	1
A1BG	466.4581	-0.6025261	0.8704571
A1CF	64.06222	0.07042723	0.9839143
A26A1	323.7717	0.6025261	0.8704571
A26B1	464.6429	0	1
A26C3	49.74313	0.6025261	0.8704571
A2BP1	1448.12	-0.101726	0.9768489
A2M	-5.468216	-0.1956222	0.9559557
A2ML1	94.81216	0	1
A3GALT2	18.31459	-0.09781112	0.9777299
A4GALT	1922.817	0.1956222	0.9559557



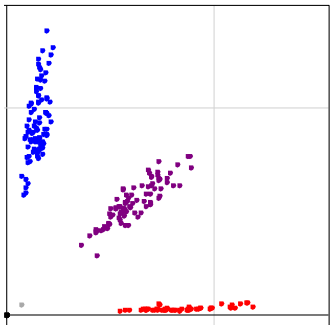
iGenix DASL Assay Statistics:

96-plex, 480 samples	
	Expression
Conversion	>95%
Reproducibility	98.40%
Cost/sample	$48.00
DNA Input	100-250ng

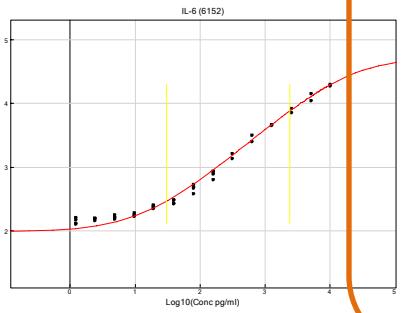
Case Study 3:

Proper Project/Probe Design

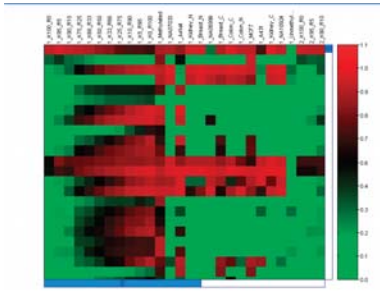
Genotyping



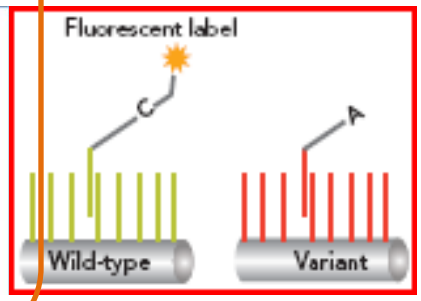
Expression



Methylation

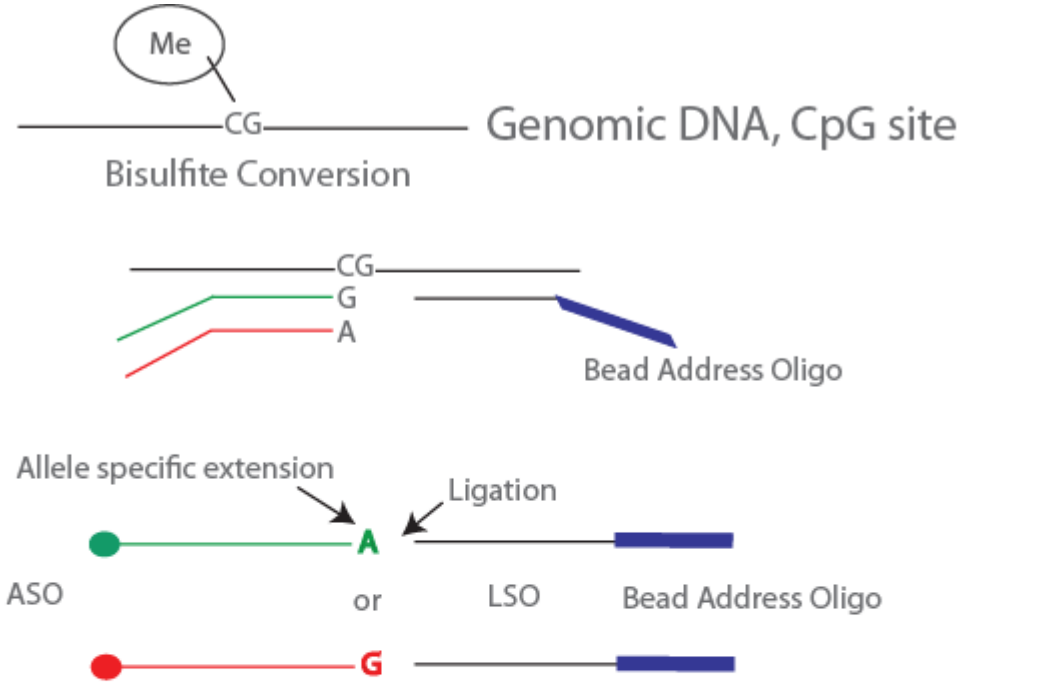


Allele-Specific

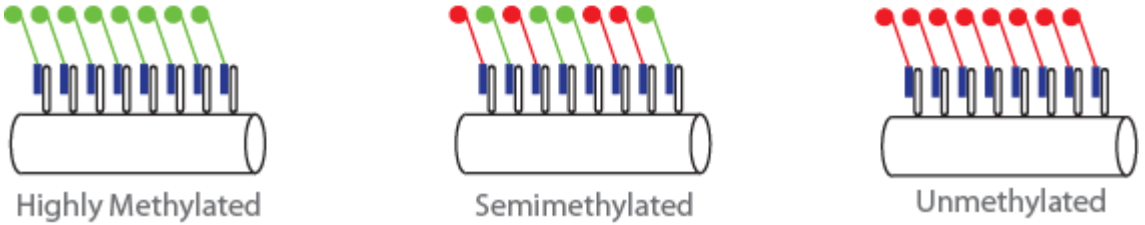


Illumina Veracode

Methylation:



Hybridize to Veracode Beads, Based on Address Oligo



Measure and Quantitate Fluorescence

Client Project Parameters:

- Find differentially methylated candidate regions in mouse.

- Didn’t know a-priori, which regions to target.

- iGenix recommended an automated strategy to use annotated promoter regions on the X to identify good CpG island candidates for their screen.

Design Challenges for Methylation:

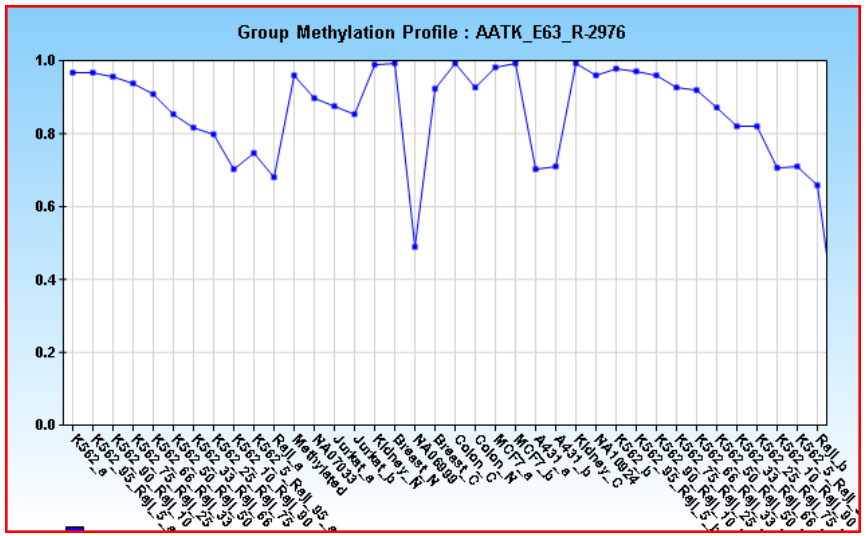
-2807 CpG sites across 97 loci

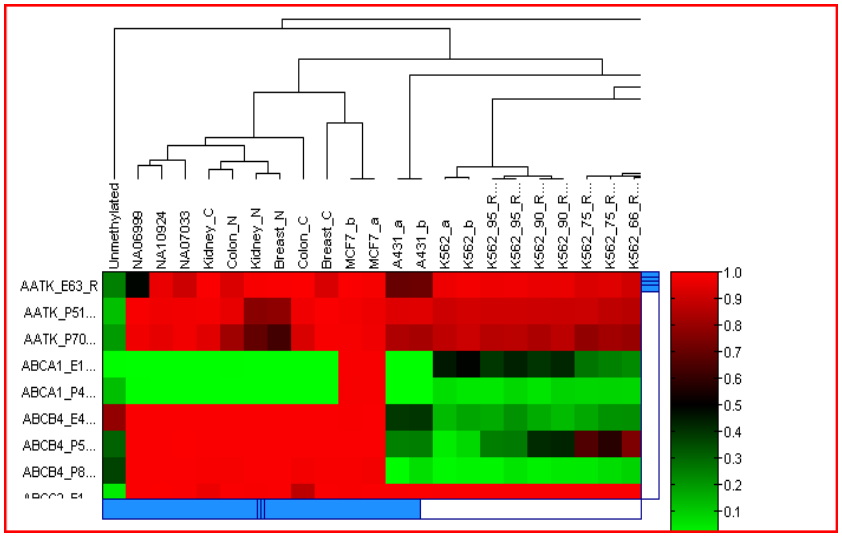
- 153 High-Quality candidate sites
  - -~80% were overlapping other CpG sites
  - -~20% remaining were low quality
- Finalized with 96 high-quality sites

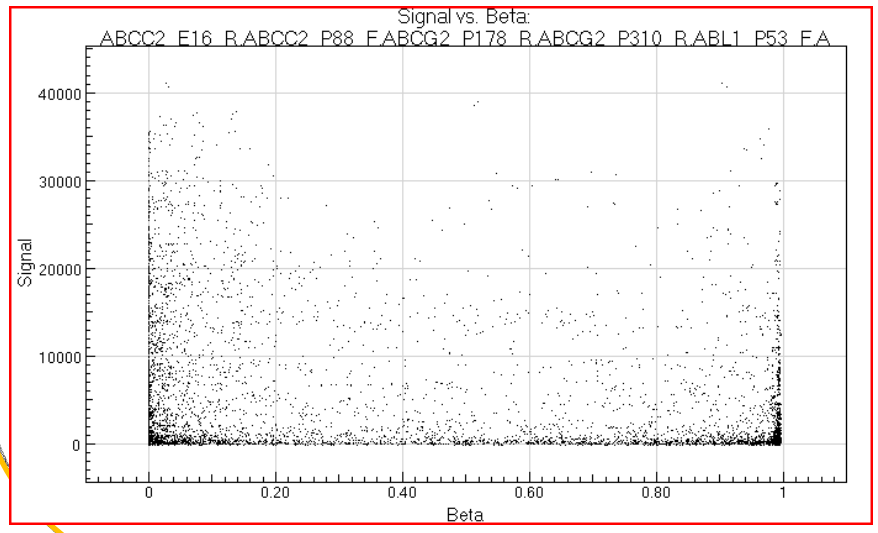
iGenix Methylation Assay Statistics:

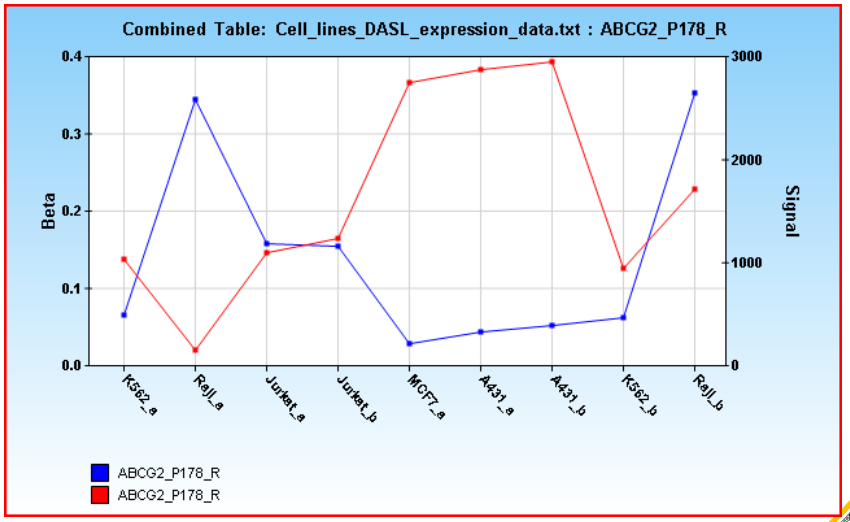
96-plex, 480 samples	
	Methylation
Conversion	>95%
Reproducibility	97.10%
Cost/sample	$56.00
DNA Input	250-500ng

Customized Analysis and Results:





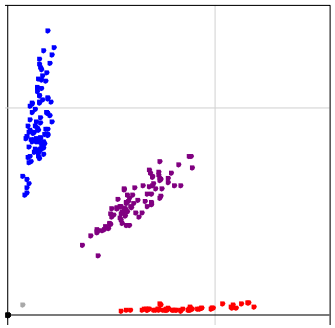




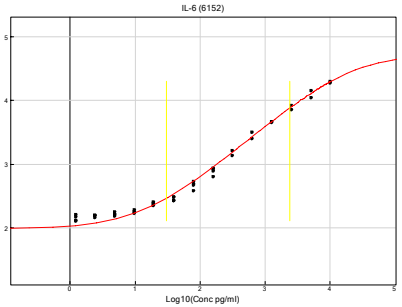
Case Study 4:

Proper Project/Probe Design

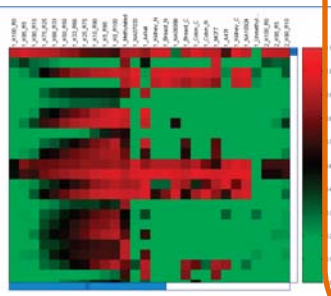
Genotyping



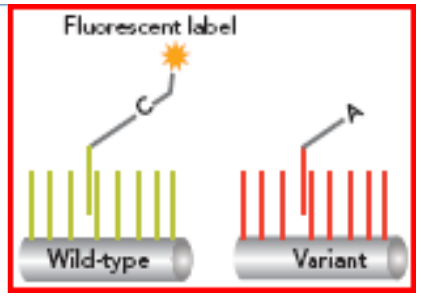
Expression



Methylation



ASPE



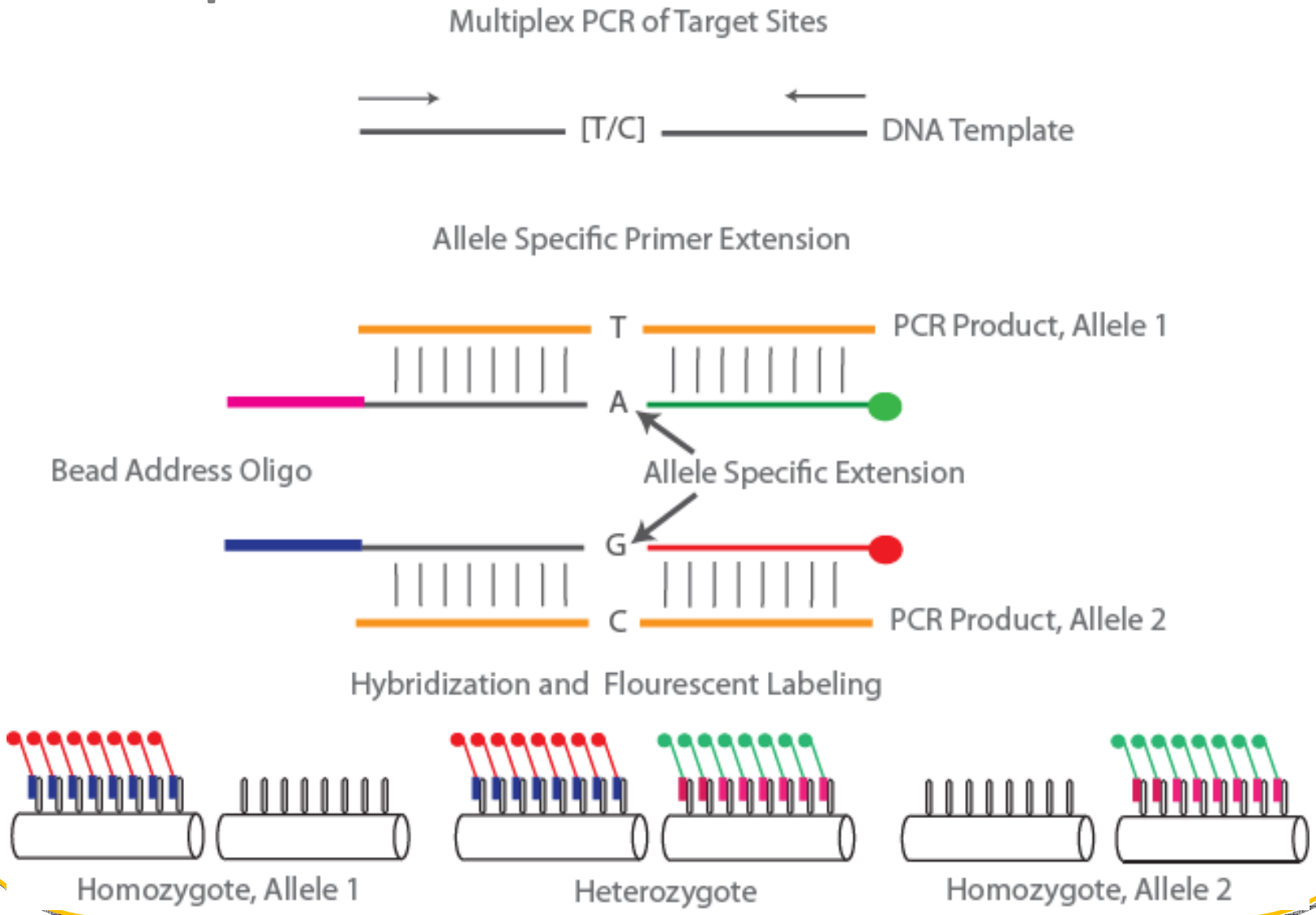
Illumina Veracode

ASPE assay uses:

- Small number of sites (2-24), reduces costs greatly **($2-$3/sample**).

- Can assay low-complexity DNA that does not work for the standard assays.

ASPE process



ASPE Considerations

- Need more time up front – manually optimize the multiplex of SNPs.

- Do your own bead kitting, but once you've done it, the assay is strong and robust and is much less expensive to run, especially for large sample sizes.

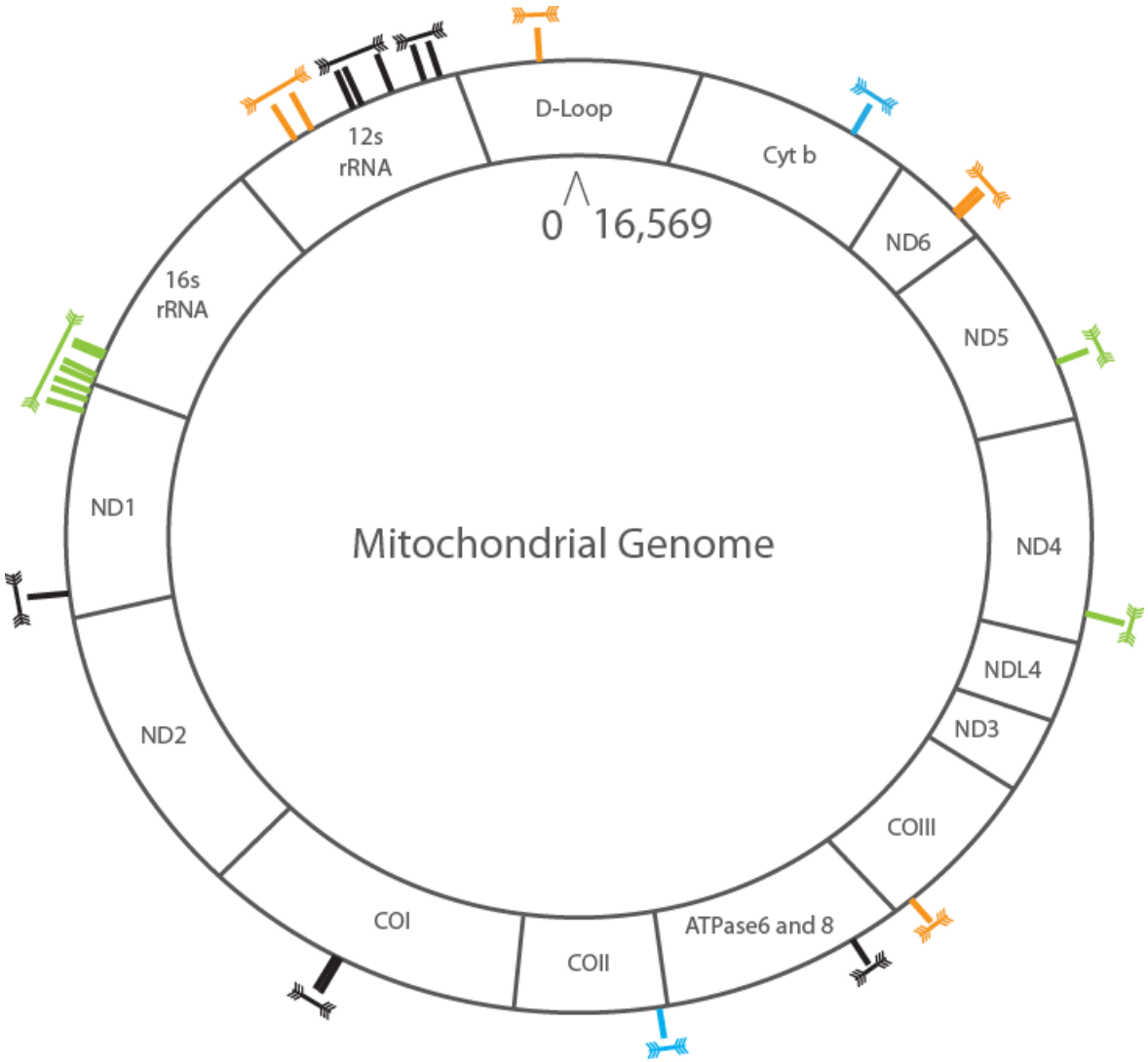
Client Project Parameters:

- Need a quick assay for 24 mutations and polymorphisms in the mtDNA which have not previously been screened.

-3000 samples to screen, did the first 2 mutations by hand. Took two people a year.

- We are designing an ASPE multiplex strategy to run through all 3000 samples in 2 months, for ~$10 a sample.

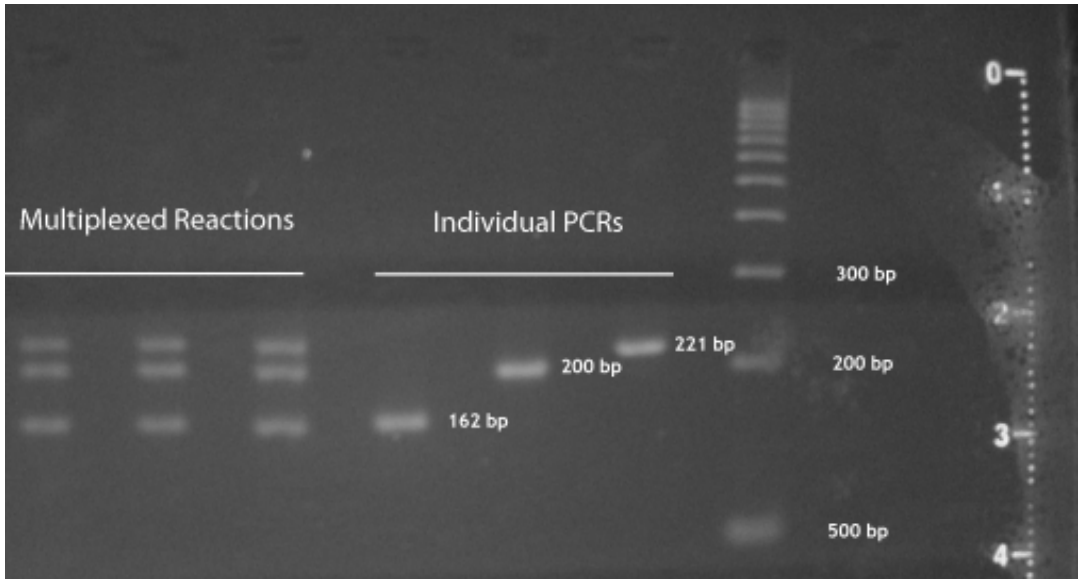
Design Strategy for ASPE:



ASPE Results:

- Currently can type 16 of the 24 in 3 multiplexes.
- 3000 samples, process in 2 months.
- $10/sample

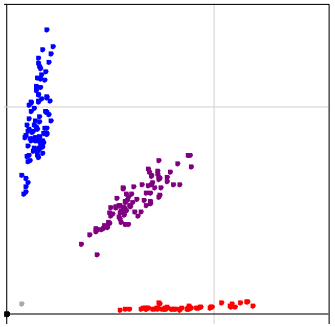
= 48,000 genotypes difficult/impossible to type with standard method.



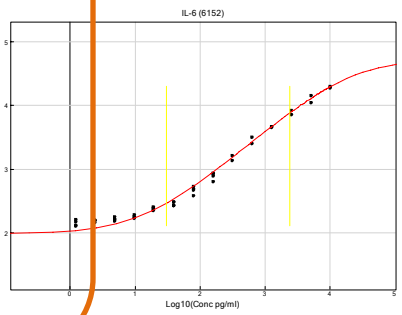
Case Study 5:

Proper Project/Probe Design

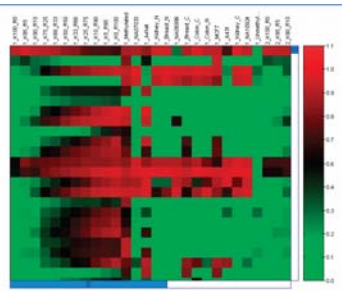
Genotyping



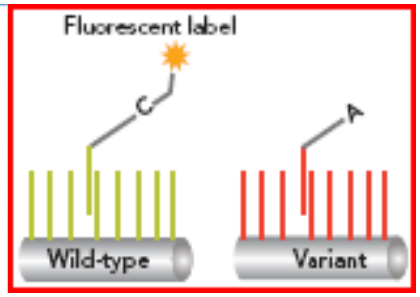
Expression



Methylation



Allele-Specific



Illumina Veracode

The Question:

-Can we screen for Copy Number Variation (CNV) affordably?

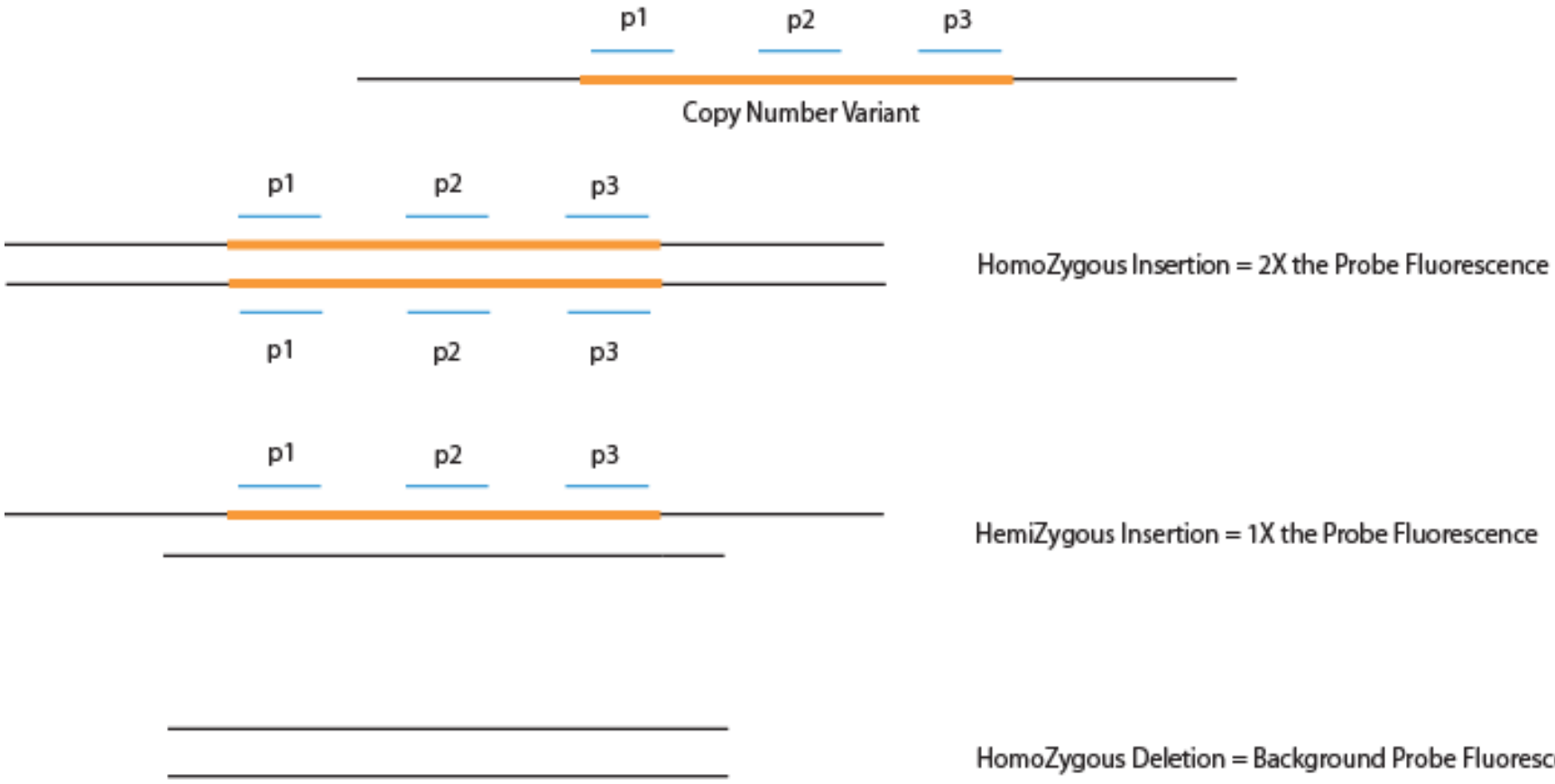
Preliminary Results:

-CNVs could be typed for < $60/sample with ~98% accuracy.

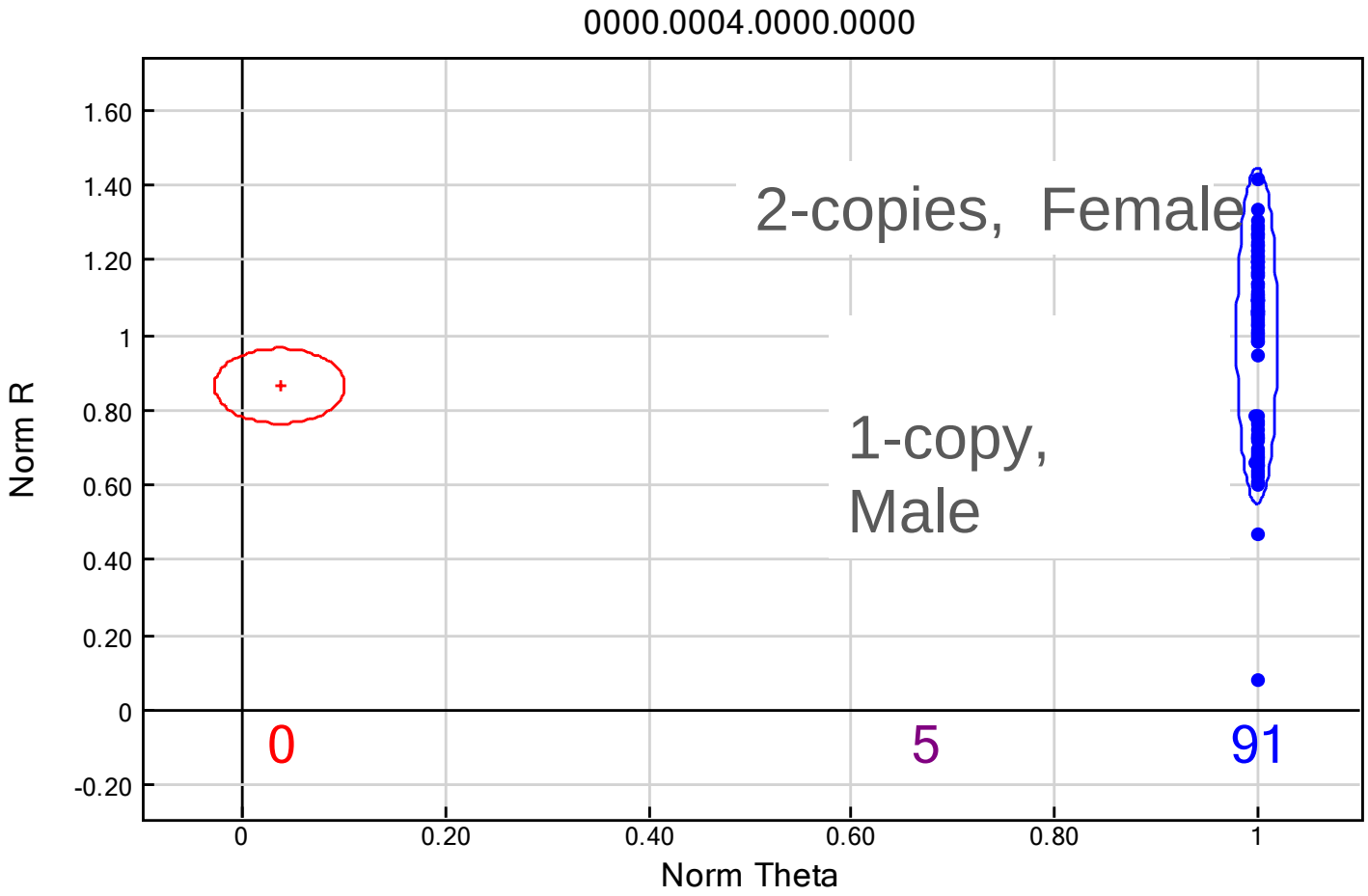
Imagine typing 1000 samples for follow-up validation….

	Affymetrix - Whole Genome SNP Chip	Illumina BeadArray/iScan - 1536 SNPs	Illumina BeadXpress - 384 Loci
Per sample cost	$500	$100	$58
Total Project Cost	$500,000	$100,000	$58,000

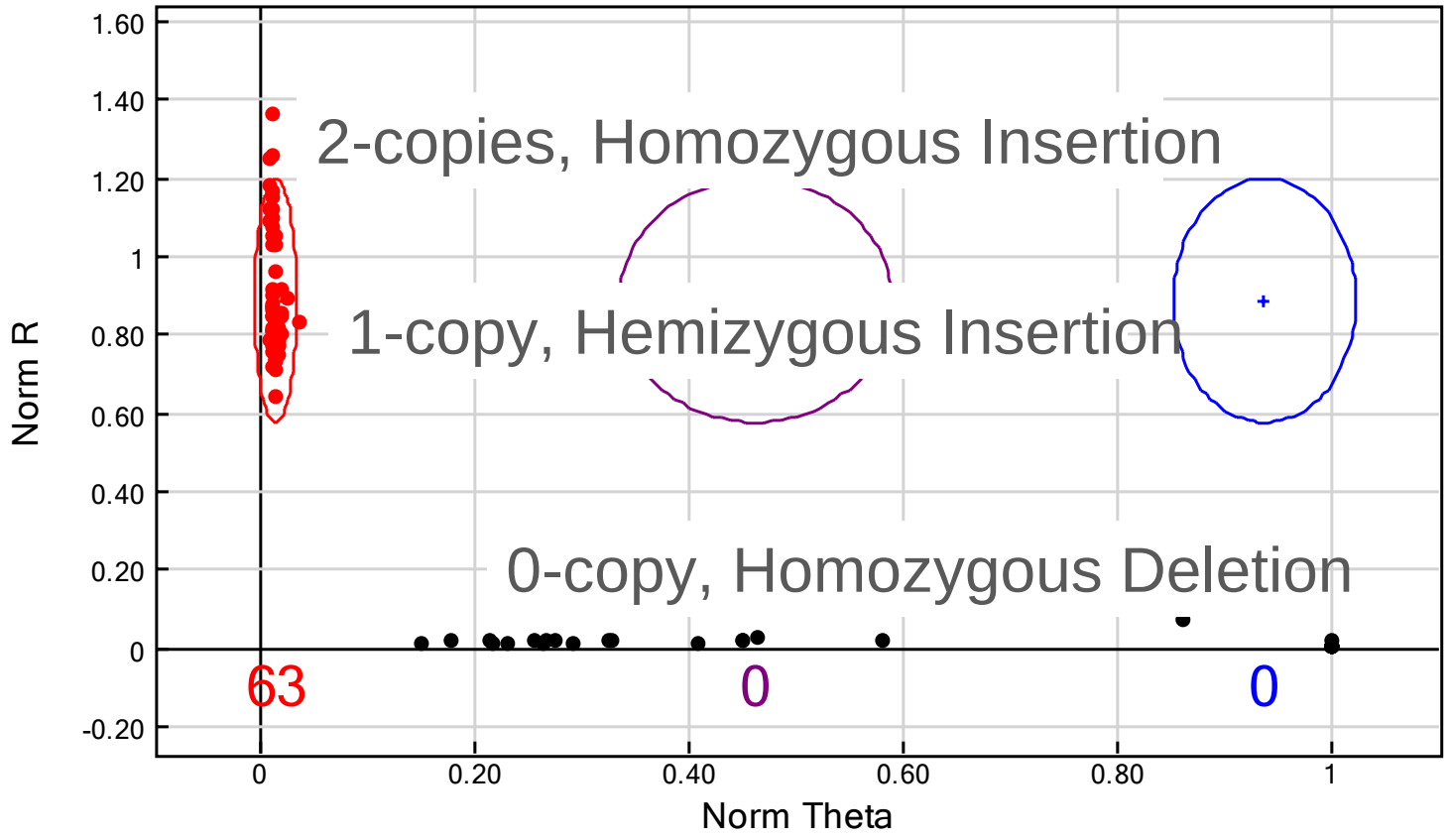
Probe Design: CNV Genotyping



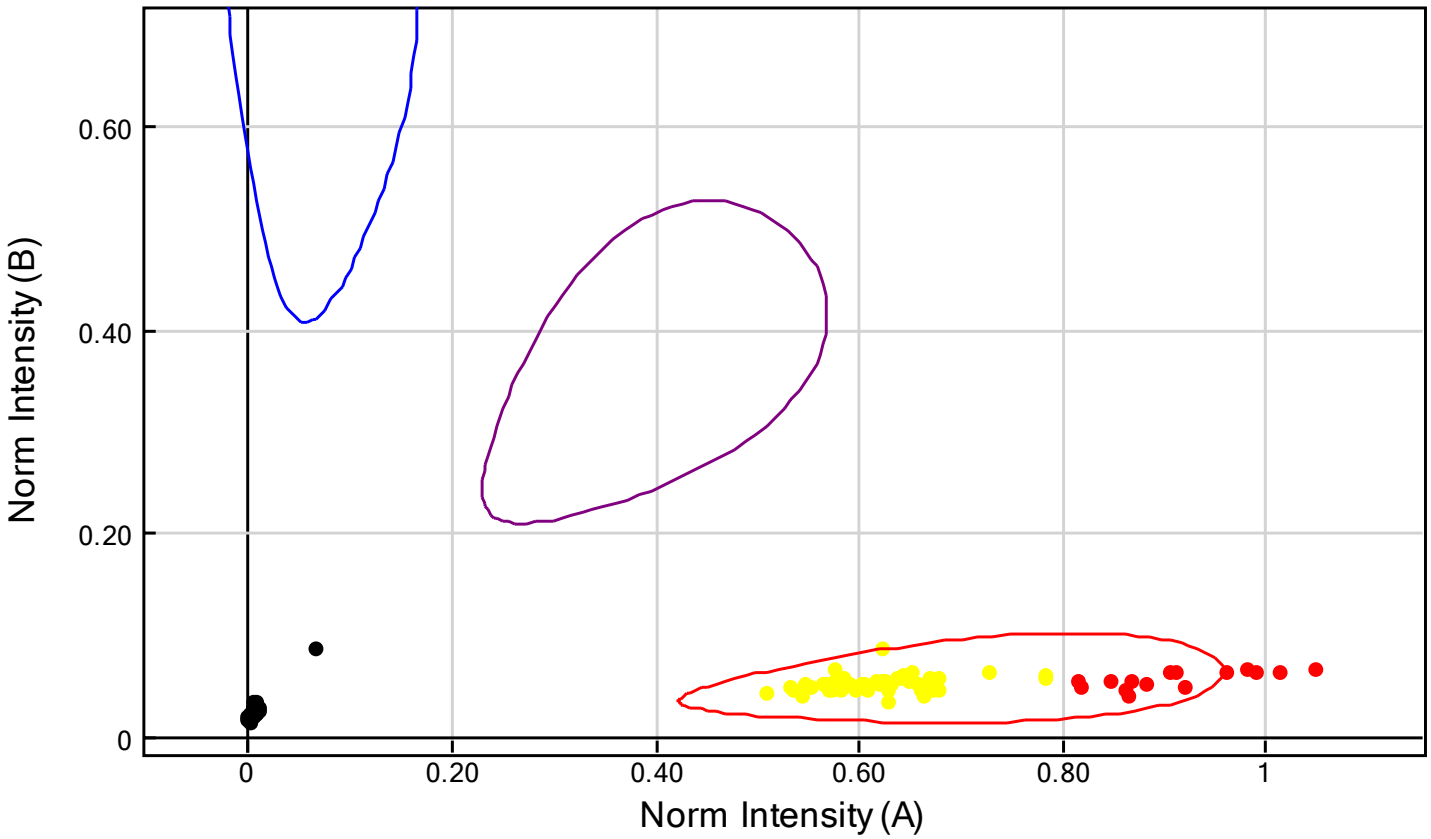
Normalized Polar Data: X chr.



Norm. Polar Data: Putative CNV



Same Individuals Across Probes?



CNV Project Design – Proof of Principle

- 384 probes
  - -40% within or near 10 known or putative common CNVs.
  - -60% placed in flanking single copy DNA on SNPs.
  - -~ 1 probe per 1.5Kb,
  - - Average = 5.7/CNV.

-Typed in 90HapMap-CEPH individuals

Recent SNP-based CNV Studies

Study	Type	Analysis
Newman et. al, 2006	Focused	Manual Typing
Carlson et. al, 2006	Focused	Manual Detection
McCarrol, et. al, 2006	Genome-Wide	Pedigree / Manual Det. and Typing
Oosting, et. al, 2007	Genome-Wide	Auto. Normalization Det.
Shen, et. al, 2008	Genome-Wide	Auto. Segmentation Det.
Kidd, et. al, 2008	Genome-Wide	Sequence resolution Det. And Typing
Cooper, et. al, 2008	Genome-Wide	Auto. HMM/SCIMM Clust. Typing
McCarrol, et. al, 2008	Genome-Wide	Birdsuite 4-stage Det. And Typing

Oosting et. al Method

- Normalize between red and green channels (quantile).

- Median centering of sample intensity to 1, using high-quality het. probes.

- Median centering of probes and adjust to copy number 2 using normal samples.

- Found 94-95% correlation with their method when compared to BAC array and GeneChip analyses.

Oosting et. al, 2007, *Gen. Res*. 17:368-376.

Adjustments to Oosting method

- Did not throw out “poor samples” with bad genotyping calls.

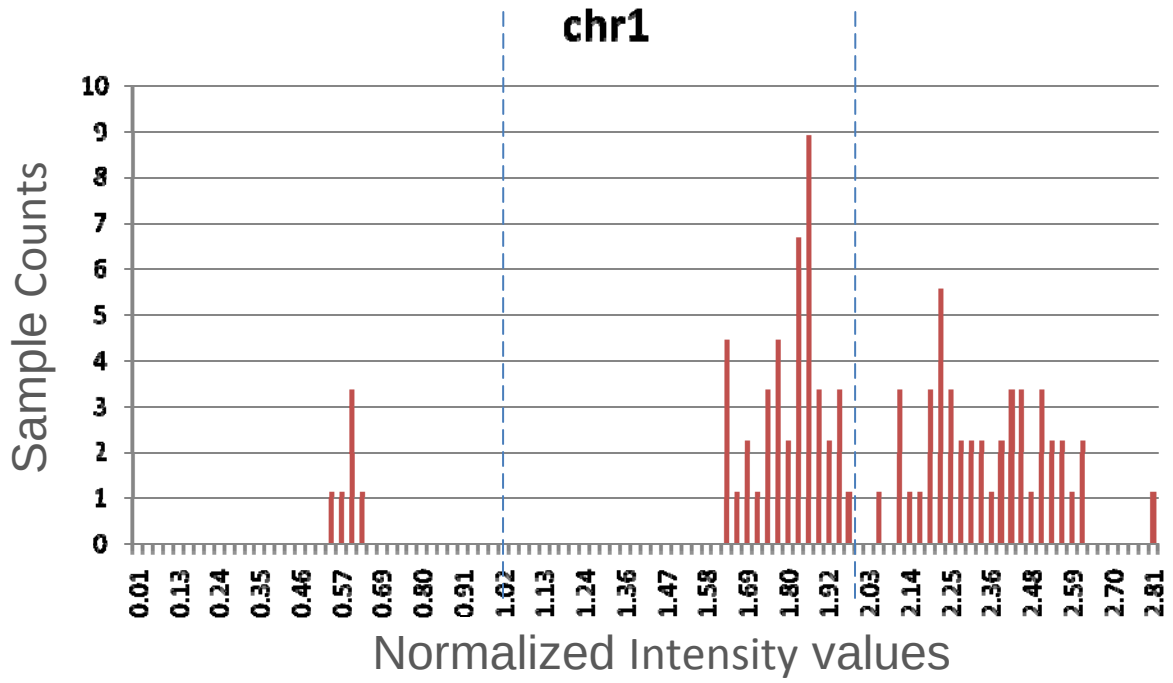
- Did not assume a-priori tumor or normal status - assumed all samples were normal.

- Did not smooth data with lambda - averaged sign. Intensity across the “inside probes”.

- Used iGenix heuristic for clustering/calling.

Distributions/Clustering

- Two rule heuristic clustering algorithm
  - -Look for trend shift between clusters
  - -Maximize separation between clusters



Preliminary Results:

Method	Chr	Start	End	Correct Calls	% Corres.
McCarrol et. al.	1	149,574,996	149,581,773	88/88	100.0
	15	74,678,296	74,682,830	86/88	97.7
	22	37,688,119	37,699,807	87/88	98.8
Cooper et. al.	1	149,572,945	149,583,429	47/47	100.0
	7	97,039,956	97,047,292	47/47	100.0
	8	51,193,494	51,200,884	46/47	97.8
	10	70,950,995	70,961,085	43/47	91.4
				Average:	98.0

Report Format

		Sample	NA06985	NA06991	NA06993	NA06994	NA07000	NA07029	NA07345	NA07348
1	149572945	149583429	2.555087	2.626644	2.205595	2.522691	1.640233	1.747417	1.916181	0.601293
		Call	2	2	2	2	1	1	1	0

iGenix Process:

1. Assay design, 2-6 weeks, depending on complexity.
2. We order array and reagents, ~4-6 weeks.
3. Send DNA samples, 10 uL with 50ng/uL.
4. Run samples 200-400/week.
5. Data interpretation, calling, analysis.

Contact Tera Eerkes

teerkes@igenixinc.com

To inquire about the promotion for this symposium.