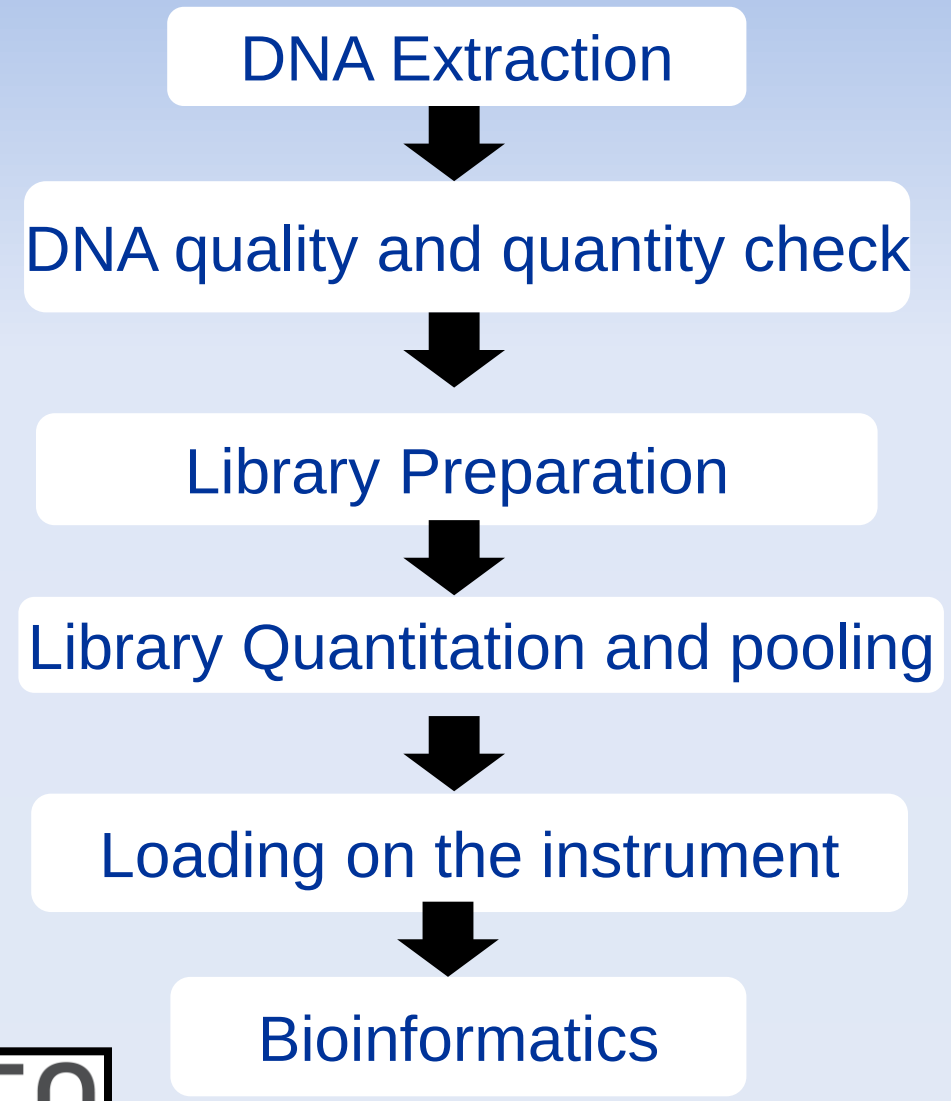
Targeted Sequencing

Removing Sample Prep Bottlenecks Using Access Arrays

Uma Dandekar
Asst. Director UCLA Genoseq Core
David Geffen School of Medicine at UCLA
Los Angeles, CA 90025
webseq@genetics.ucla.edu
310-267-2461

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Next Gen Sequencing Pipeline



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Core Equipment

Next Gen Sequencers

- Roche GS FLX 454
- Roche Junior
- Illumina Miseq

Real Time PCR machines

- Fluidigm Biomark
- ABI 7900HT
- Roche Lightcycler

Other Equipment

- ABI 37030
- PSQ96 Pyrosequencer
- Qiagen TissueLyser
- Agilent Bioanalyzer
- Coulter Counter
- Fluorimeter/spectrophotometer
- Nanophotometer
- Qubit
- Liquid handling robots
- PCR Machines

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Sequencing Timeline

The First Generation

- **1970:Sanger Sequencing**

The Next Gen Sequencers

- **2005 : Roche 454 GS 20**
- **2006: Illumina Genome Analyzer IIx**
- **2008: Applied Biosystems SOLiD 2**
- **2009: Complete Genomics**
- **2010: Ion Torrent**
- **2011: Illumina Miseq**

The Third Gen sequencing

- **Helicos Biosciences**
- **Pacific Biosciences**
- **Oxford Nanopore technologies**

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Roche GS FLX Plus (454)

Capable of 700bp reads in 24 hours

Roche Junior

Capable of 400bp reads in 10 hours

Illumina MiSeq

Capable of 500bp reads in 40 hours

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Library preparation

Most time and labor intensive part of the next-generation sequencing process.

- **Agilent Haloplex**
- **Agilent Sureselect**
- **Nimblegen EZCap**

- **Thunderstorm system from Raindance**

- **Illumina's TruSeq Custom Amplicon**
- **Nextera XT**
- **Roche 454 fusion primer design**
- **Fluidigm Access Array**

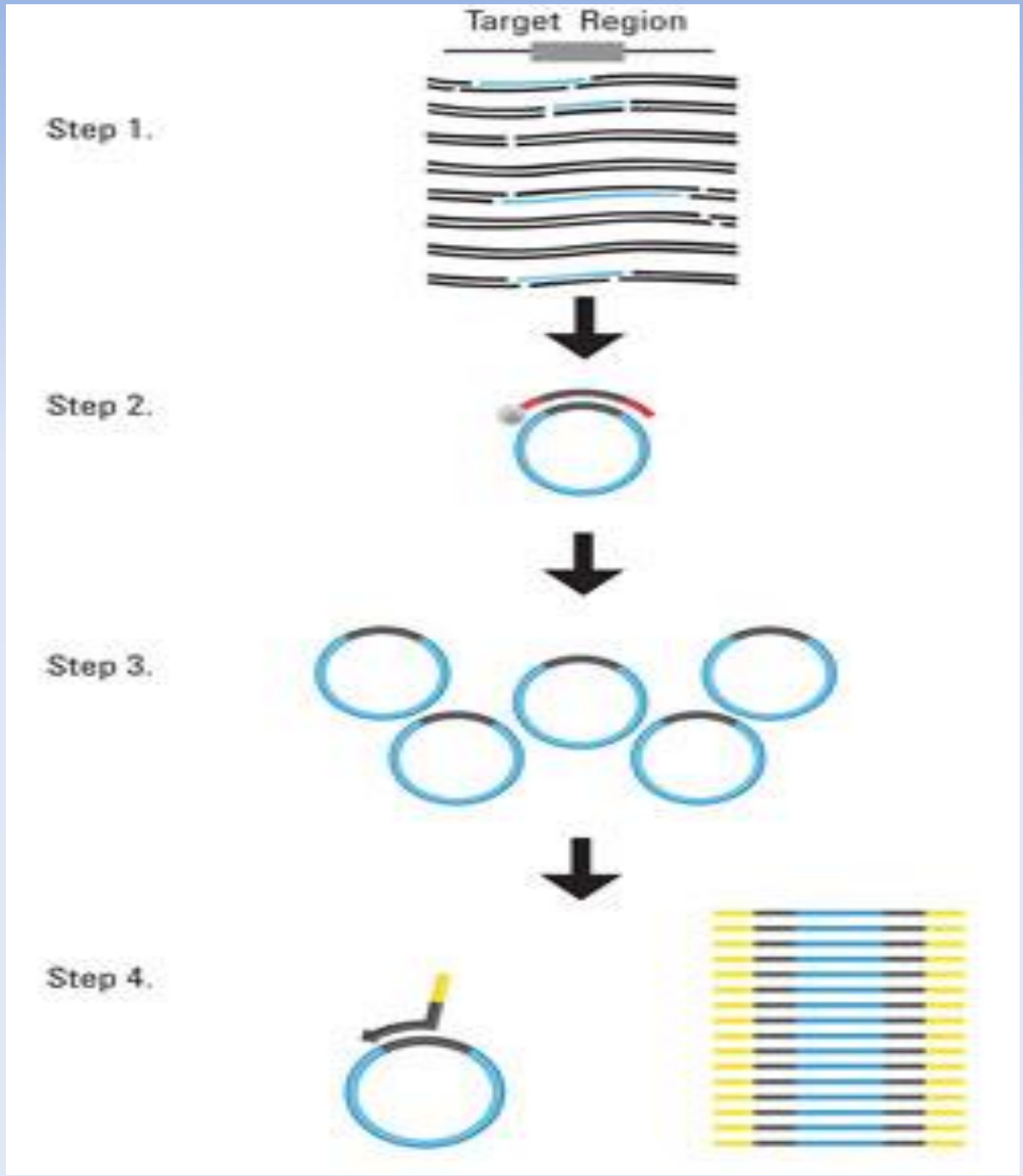
GENOSEQ
UCLA GENOTYPING & SEQUENCING

Library prep comparison

Technology	Sample Number	Size of region	Time required	Ease of Use
Agilent Haloplex	48-96	1-500kb	1-2 days	Long protocol
Agilent Sureselect	16-480	1kb-25Mb-exomes	1-2days	Long Protocol
Nimblegen EZ cap	24-96	100kb-50Mb	Few days	Long Protocol
Raindance	1-96	150bp-10kb	1 day	Short protocol
TrueSeq Custom Amplicon	1-96	2kb-650kb	2 days	Long protocol
Roche Fusion Primers	1-96	200bp-variable	2 days	Very Long Protocol
Access Array	Multiples of 48	24kb-240kb	1 day	Short protocol

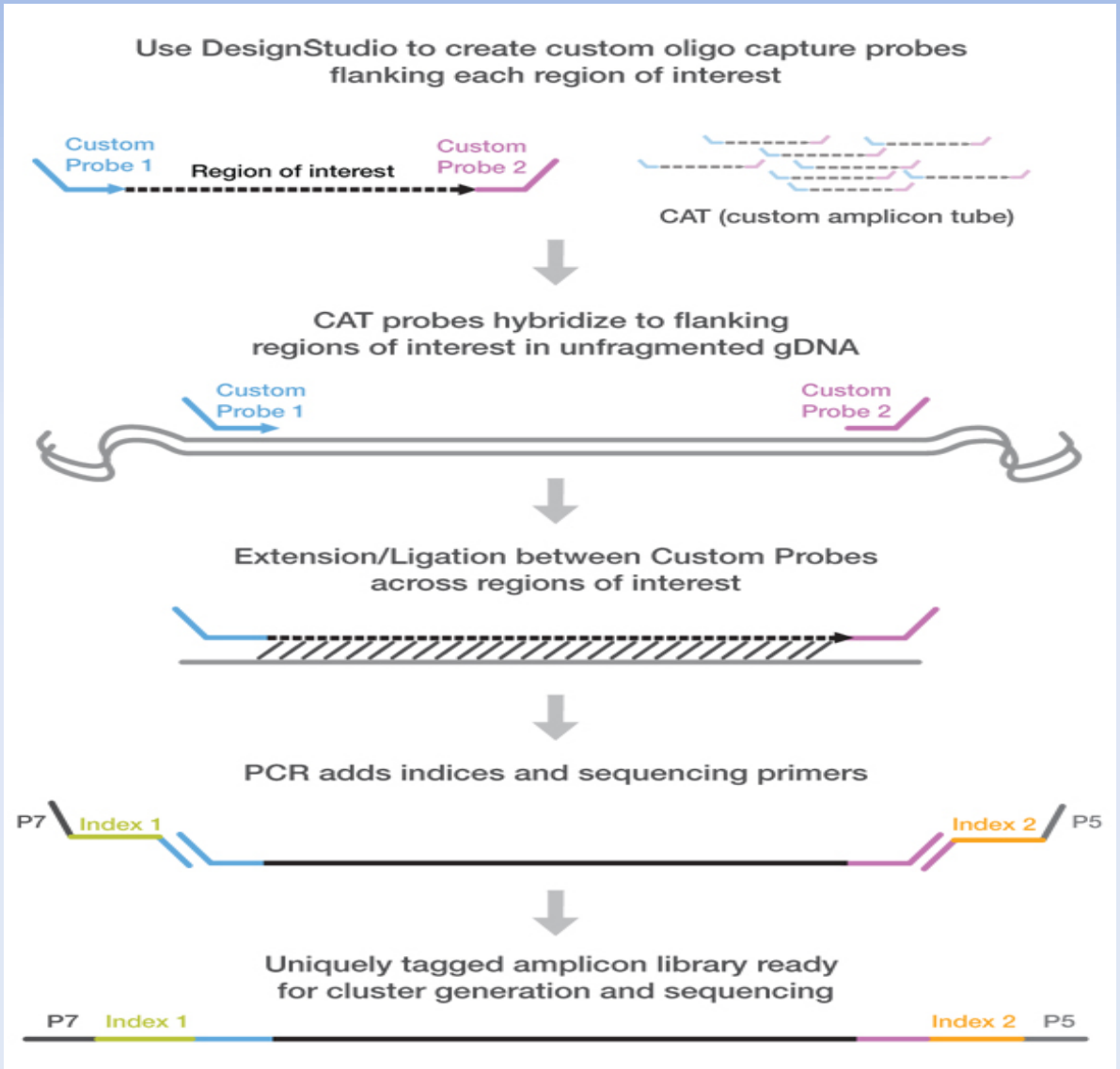
GENOSEQ
UCLA GENOTYPING & SEQUENCING

Agilent Haloplex workflow



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Illumina TSCA workflow

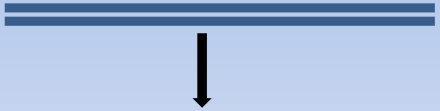


GENOSEQ
UCLA GENOTYPING & SEQUENCING

Roche Fusion primer Amplicon workflow

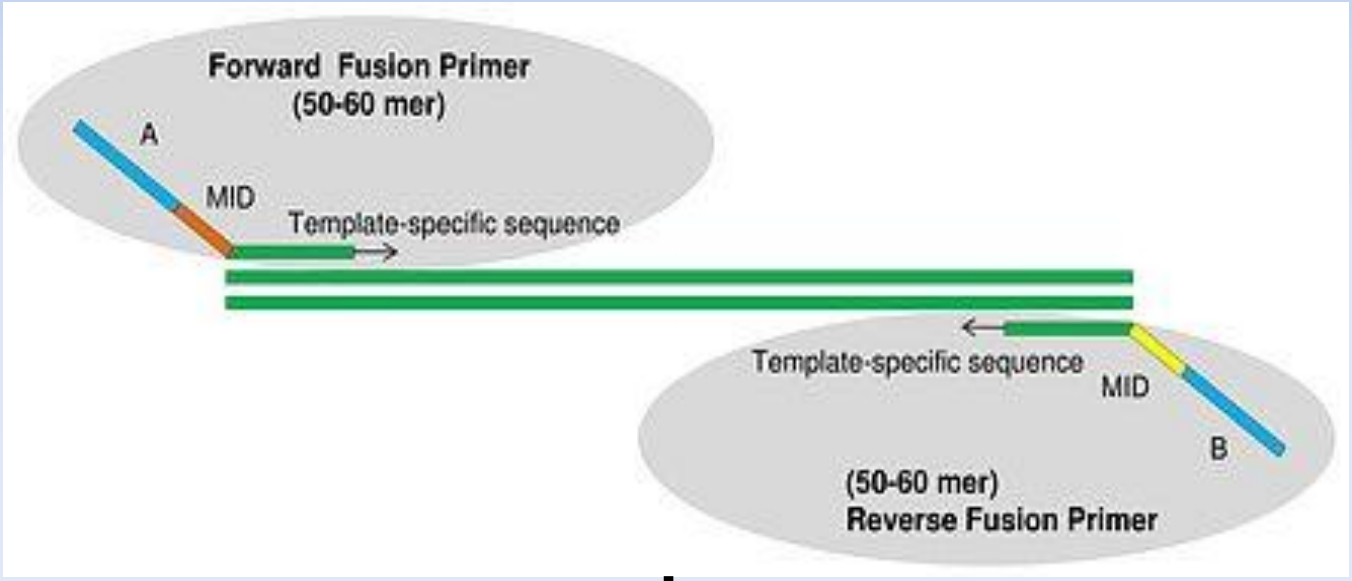
Design and synthesize 50-60 mer fusion primers

↓



Region of interest

↓



↓

1st round PCR amplification with long primers

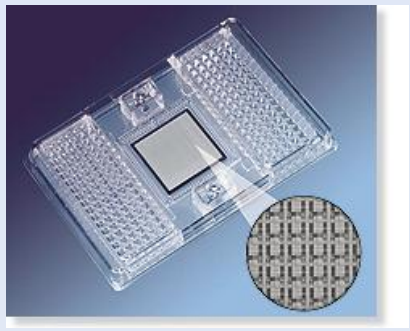
↓

emPCR & Sequencing

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Fluidigm BioMark

- This instrument employs integrated fluidic circuits at nanoliter volumes.
- It can be used for Access Array, Single Cell Gene Expression, Digital PCR and SNP Genotyping
- 48.48 – 48 samples x 48 assays
- 96.96 – 96 samples x 96 assays
- 12.765 Digital array
- 48.770 Digital array



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Fluidigm Biomark

Any Sequencer
Any Application

Capture per Sample	Capture Per Array
Amplicon Tagging*	24kb
Multiplexed Amplicon Tagging*	240kb
Long Range PCR	480kb

GENOSEQ
UCLA GENOTYPING & SEQUENCING

2-Step Access Array
4-Primer Amplicon Tagging Protocol

Order Validated Primers → Load samples & primers on Access Array → PCR amplification

Harvest PCR products → Pool Barcoded PCR products → Next Gen sequencer pipeline

GENOSEQ
UCLA GENOTYPING & SEQUENCING

2-Step Access Array
4-Primer Amplicon Tagging Protocol

Order Validated Primers

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Primer Design Service

Fluidigm

Access Array Assays Design Order Form

Multiplex Assays

Contact and Order Information					
Company/Institution					
Customer Name					
Ship to Address Please include state, country, and postal code					
Date					
Email					
Phone Number					To be completed by Fluidigm
Sales Rep. or Distributor					
Sequencing Platform	Illumina				Received By:
Design Parameters	Optimal/min/max	Optimal	Min	Max	Received Date:
	Amplicon Length	190	100	200	Job ID:
Multiplex Assays	Multiplex assays are for use on Illumina systems and are only provided as non wet-tested assays. Maximum of 480 assays can be multiplexed and distributed across 48 wells (up to 12 assays per plex) according to bioinformatically obtained recommendations. Assays are tagged with CS1/CS2 for use with the 384 Barcode plate (PN 100-3927).				

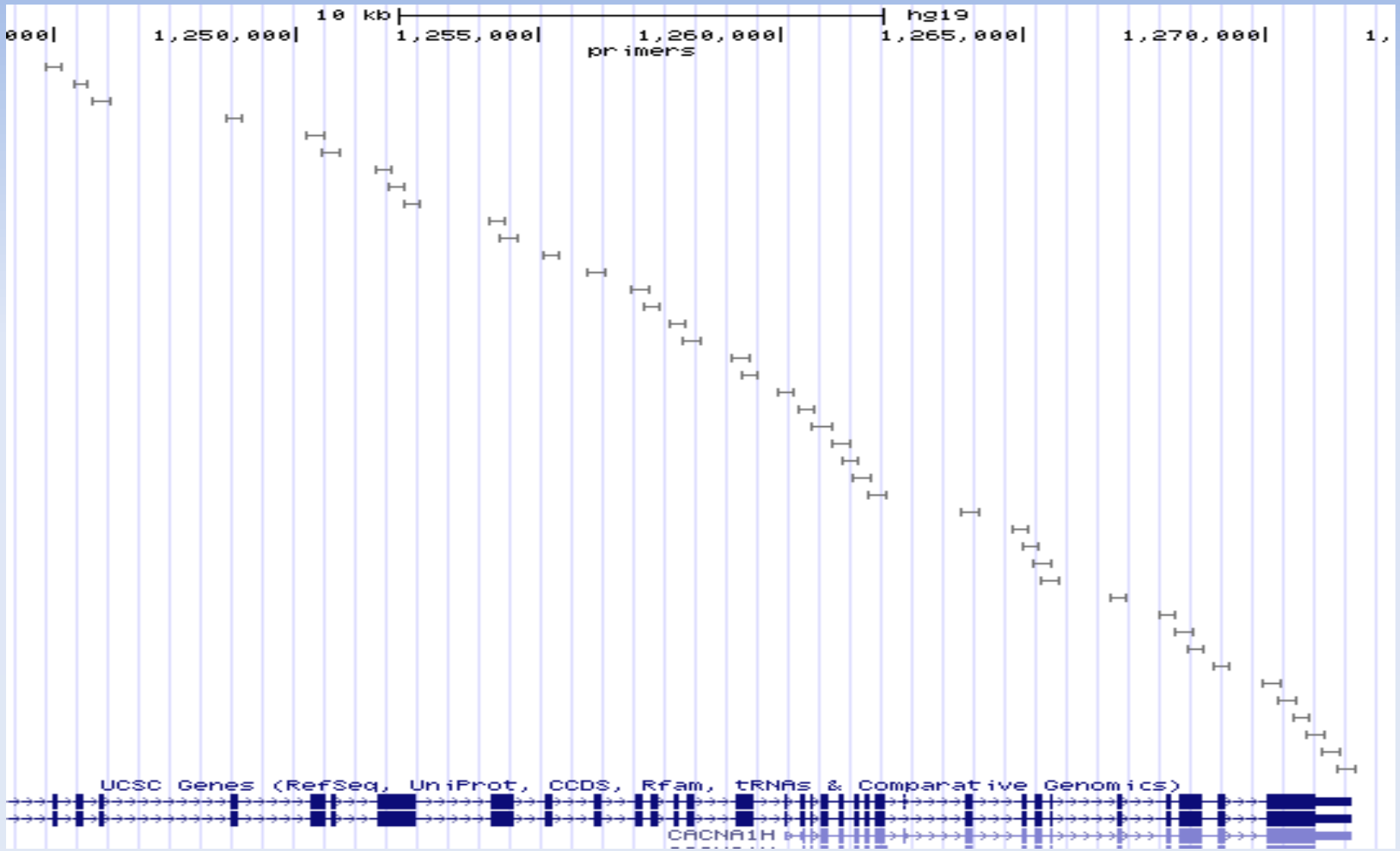
Please email completed design form to: Assay_Design_Group@fluidigm.com

Gene Resequencing: PCR primers will be designed to sequence all transcripts or exons annotated by RefSeq ID, except "retained introns," non-coding transcripts. If design parameters are not specified in the order form, default settings will be used.

Target #	NCBI Gene Symbol	RefSeq ID (preferred)	Upstream TSS	5' UTR (Y/N)	3' UTR (Y/N)	Notes (optional)
1	TP53	NM_001126115	500	Y	N	[Example]
2	NRAS	*	0	N	N	[Example]
3	WT1	NM_024426	0	N	N	[Example]

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Amplicons viewed via UCSC Genome Browser

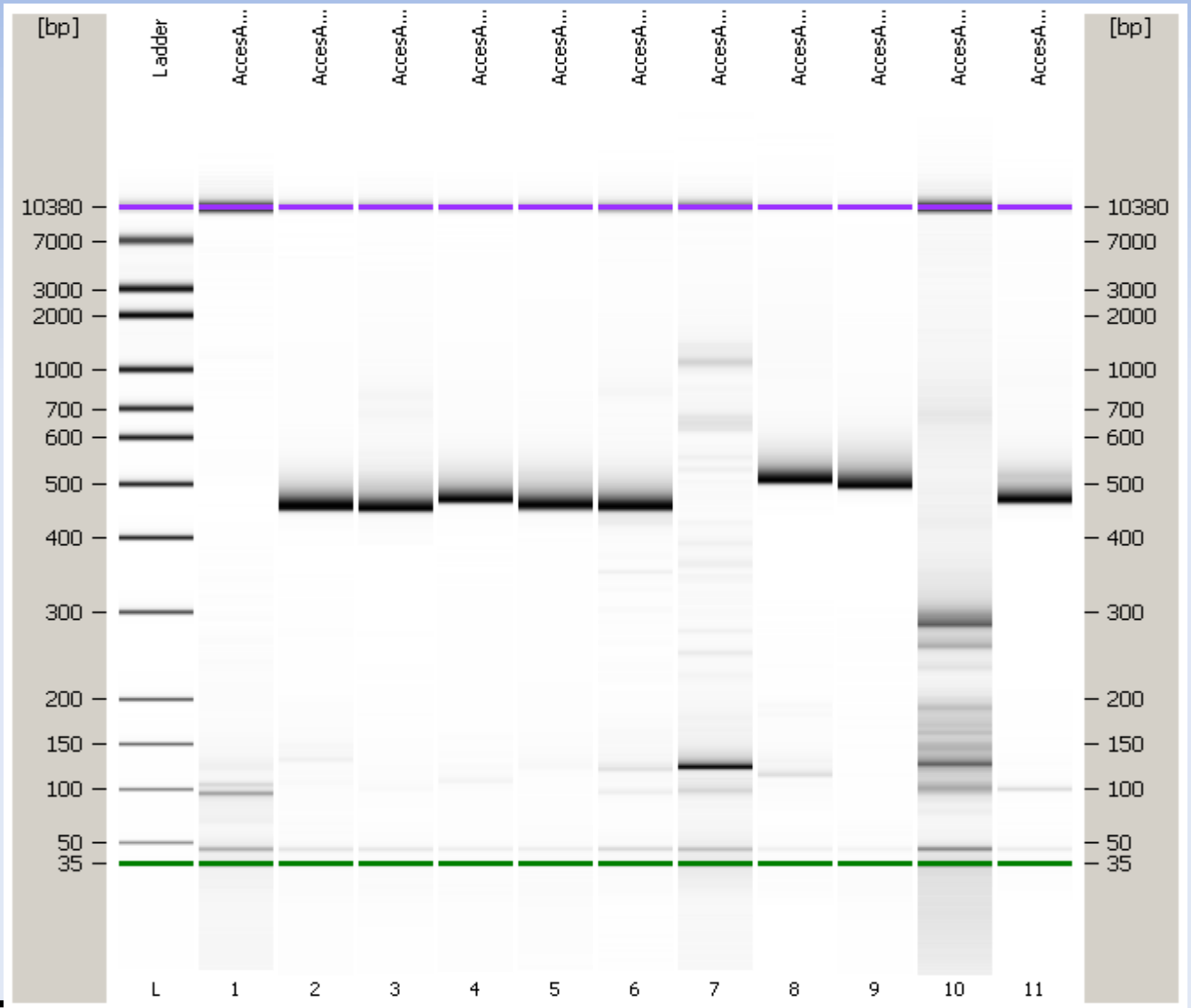


GENOSEQ
UCLA GENOTYPING & SEQUENCING

Examples of failed assays

Target Assays					
Target	Assay ID	Assay Name	Length	GC	Flag
CACNA1H	AAA0087457	CACNA1H_1	377	80	One primer sits in the repeat region
CACNA1H	AAA0087459	CACNA1H_2	359	82	
CACNA1H	AAA0087432	CACNA1H_3	379	86	
CACNA1H	AAA0087447	CACNA1H_4	318	85	One primer sits in the repeat region
CACNA1H	AAA0087419	CACNA1H_5	357	78	One primer sits in the repeat region
CACNA1H	AAA0087427	CACNA1H_6	330	76	
CACNA1H	AAA0087423	CACNA1H_7	351	66	One primer sits in the repeat region
CACNA1H	AAA0087418	CACNA1H_8	315	61	One primer sits in the repeat region
CACNA1H	AAA0087435	CACNA1H_9	385	64	
CACNA1H	AAA0087439	CACNA1H_10	343	64	
CACNA1H	AAA0087448	CACNA1H_11	385	61	
CACNA1H	AAA0087429	CACNA1H_12	384	61	
CACNA1H	AAA0087449	CACNA1H_13	349	61	

Examples of failed assays



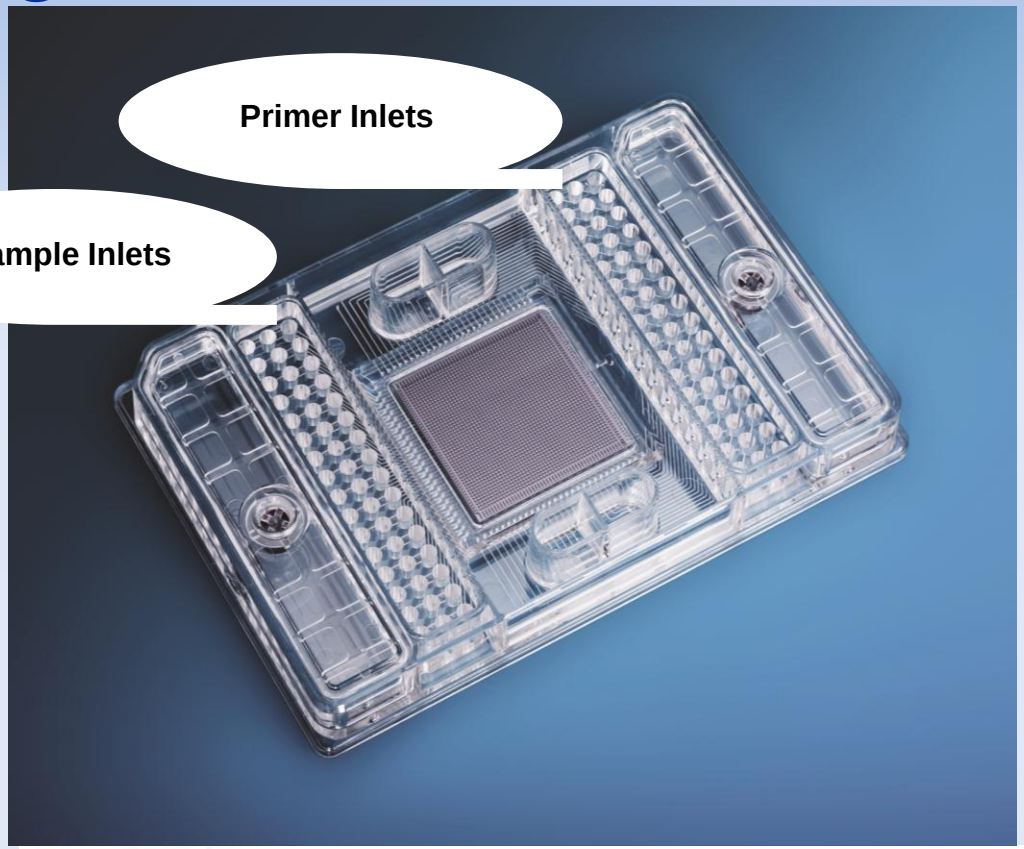
GENOSEQ
UCLA GENOTYPING & SEQUENCING

2-Step Access Array
4-Primer Amplicon Tagging Protocol

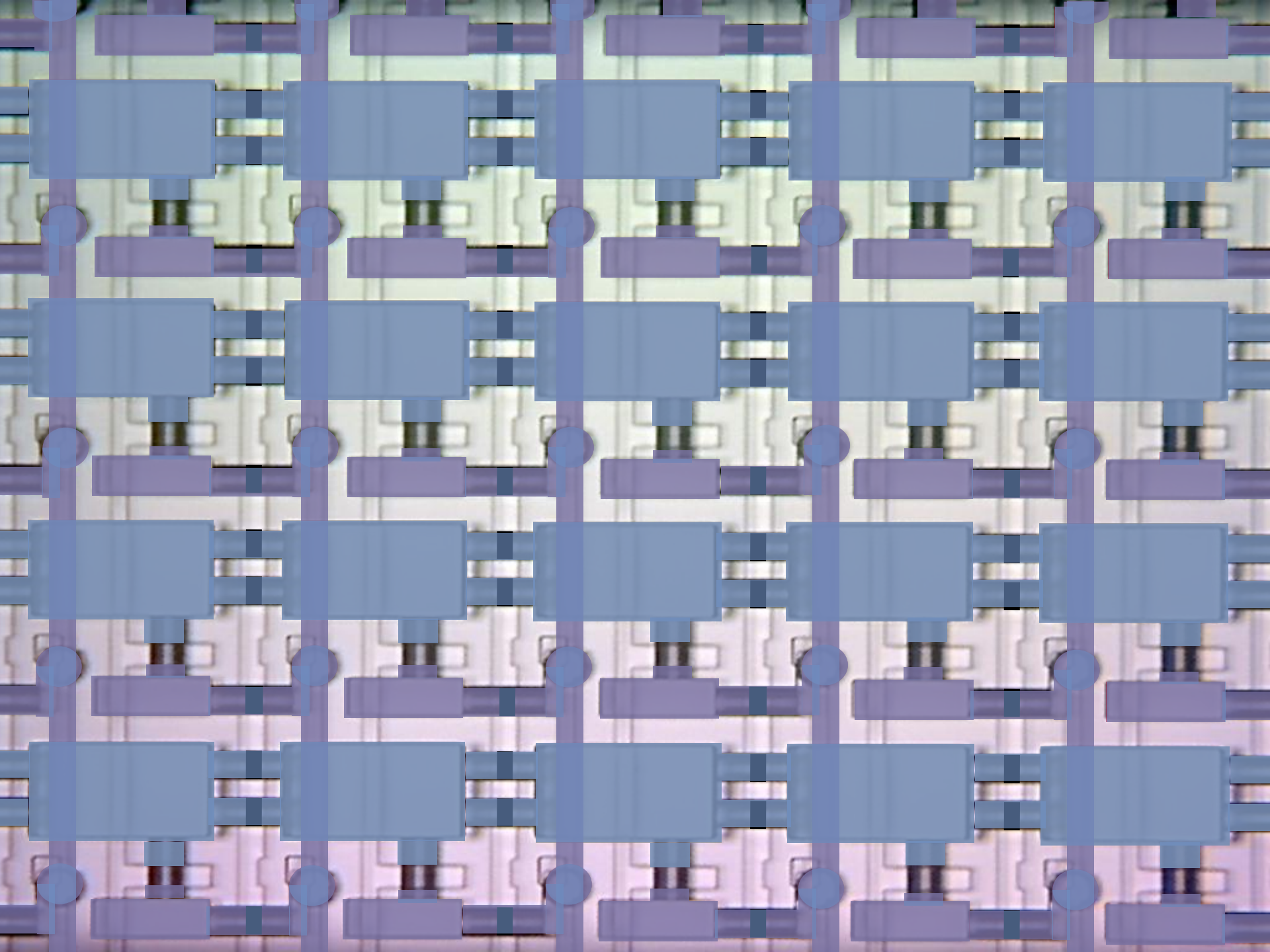
Load samples & primers on Access Array

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Access Array IFC
Integrated Fluidic Circuit



GENOSEQ
UCLA GENOTYPING & SEQUENCING



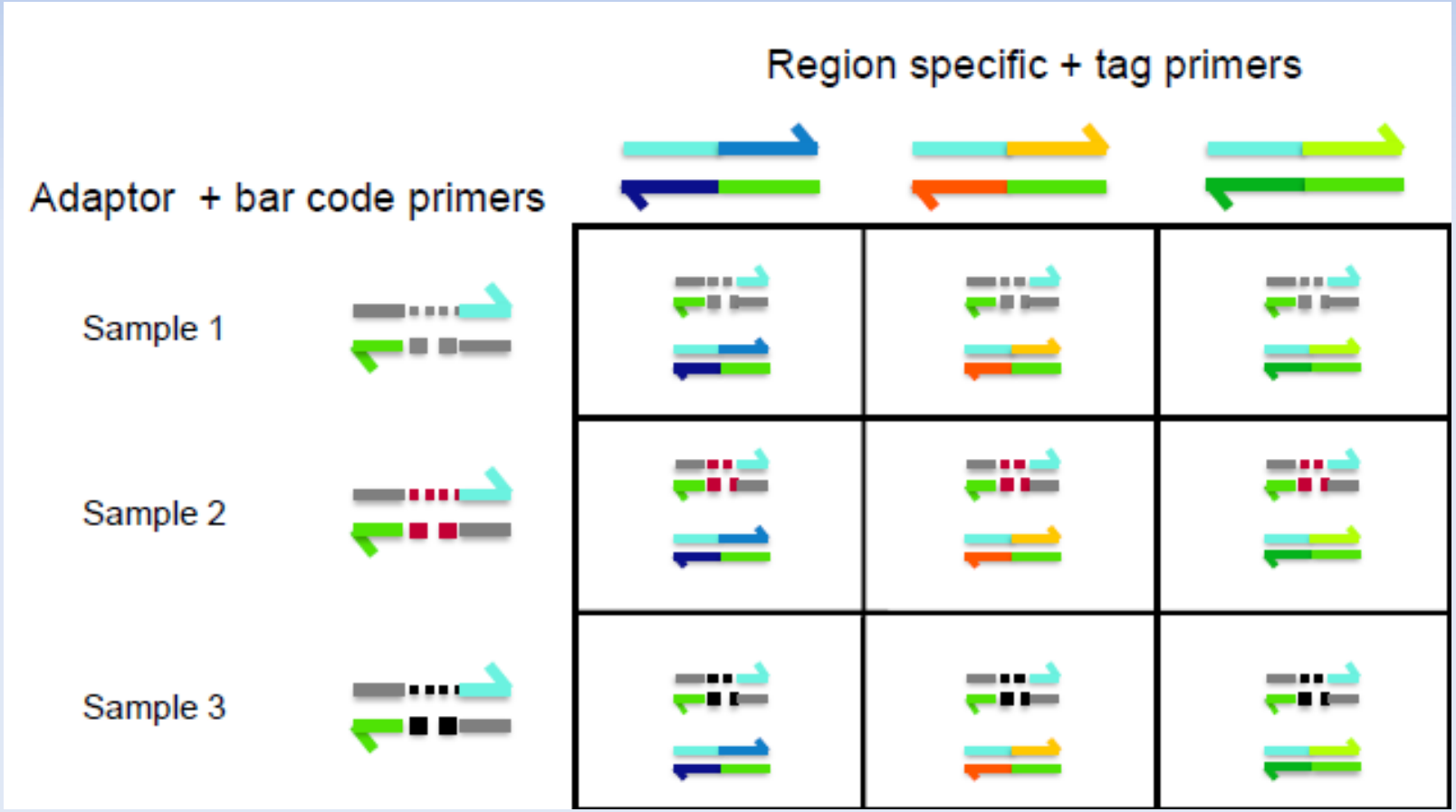
2-Step Access Array
4-Primer Amplicon Tagging Protocol

PCR amplification

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Access Array Primer Set up

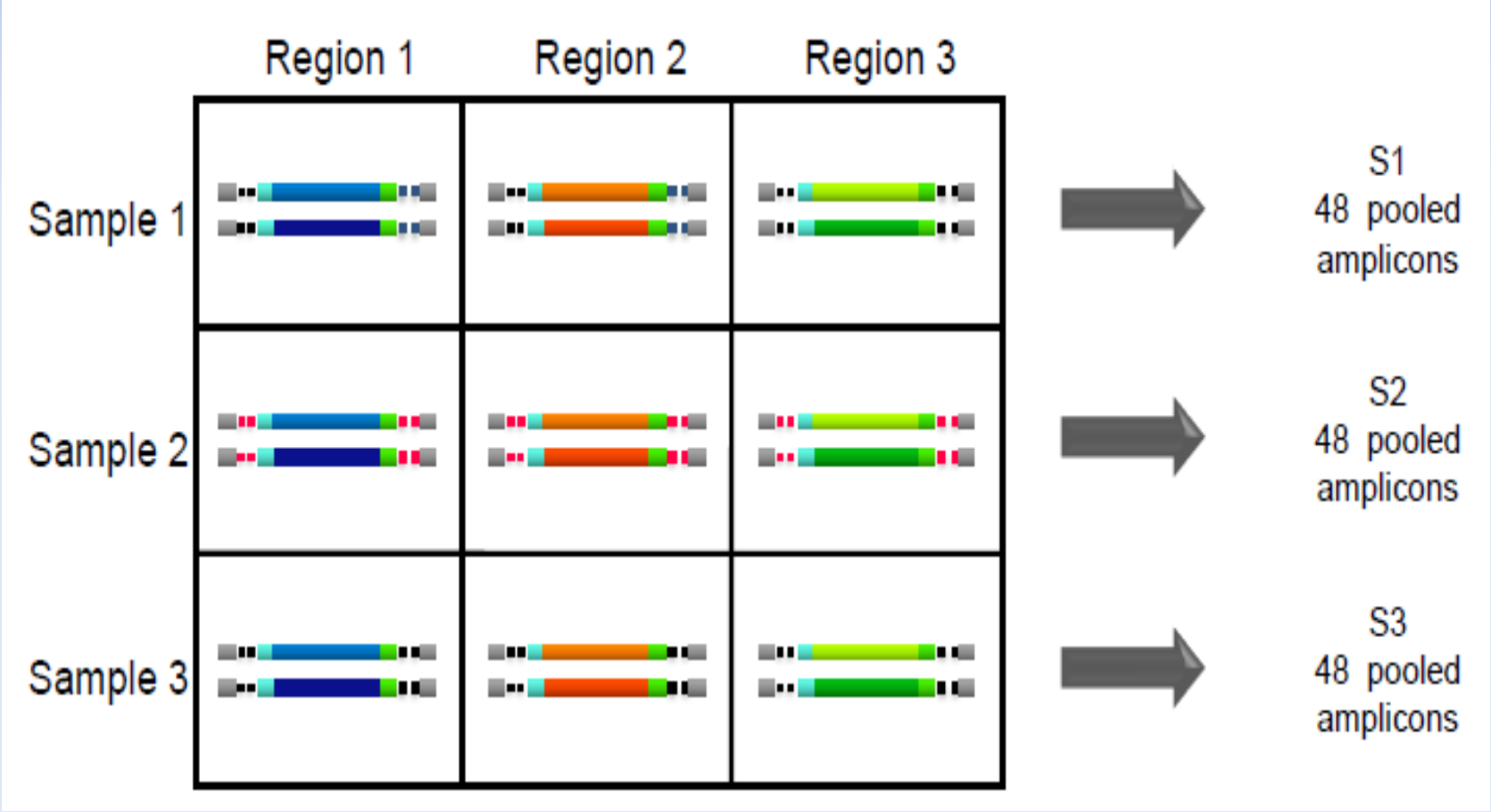
After loading the 48.48 Access Array, each reaction well contains a unique combination of TS primer pairs and Samples.



GENOSEQ
UCLA GENOTYPING & SEQUENCING

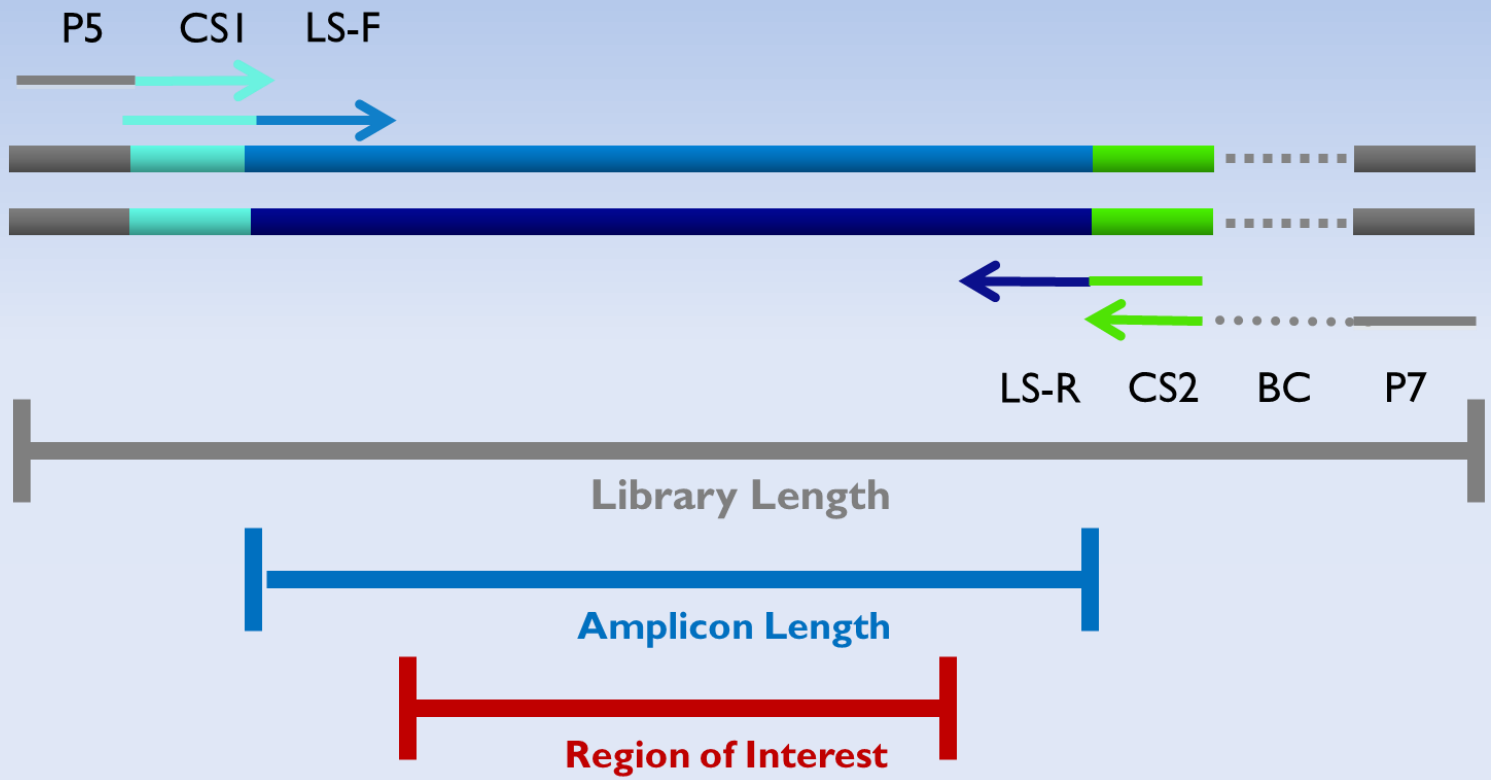
Access Array Primer Set up

Resultant PCR products.



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Illumina Amplicon



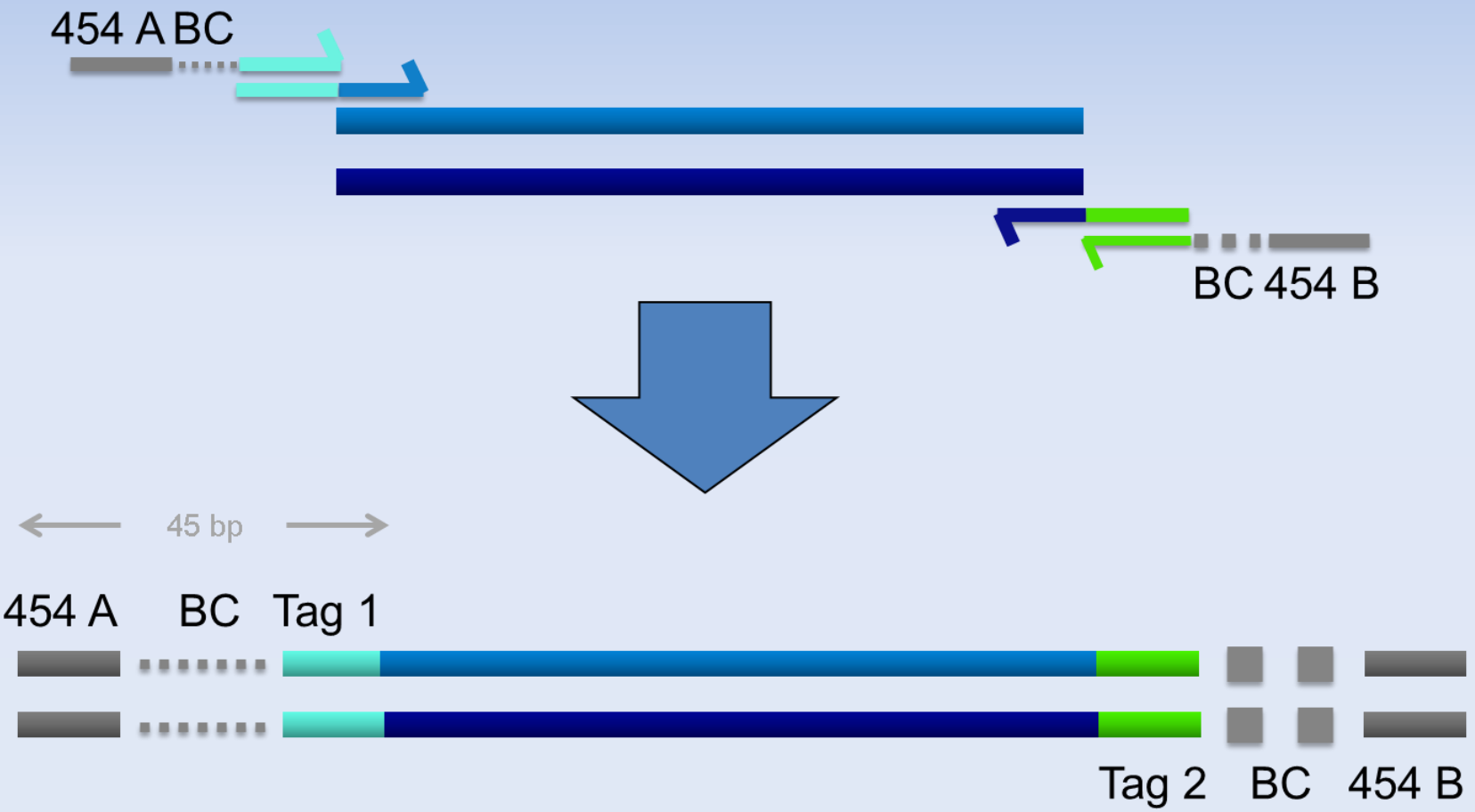
KEY

PE1 : Illumina Adaptor 1
CS1 : Consensus Seq. 1 Linker
LS-F : Loci Specific Forward Primer

PE2 : Illumina Adaptor 2
CS2 : Consensus Seq. 2 Linker
LS-R : Loci Specific Reverse Primer

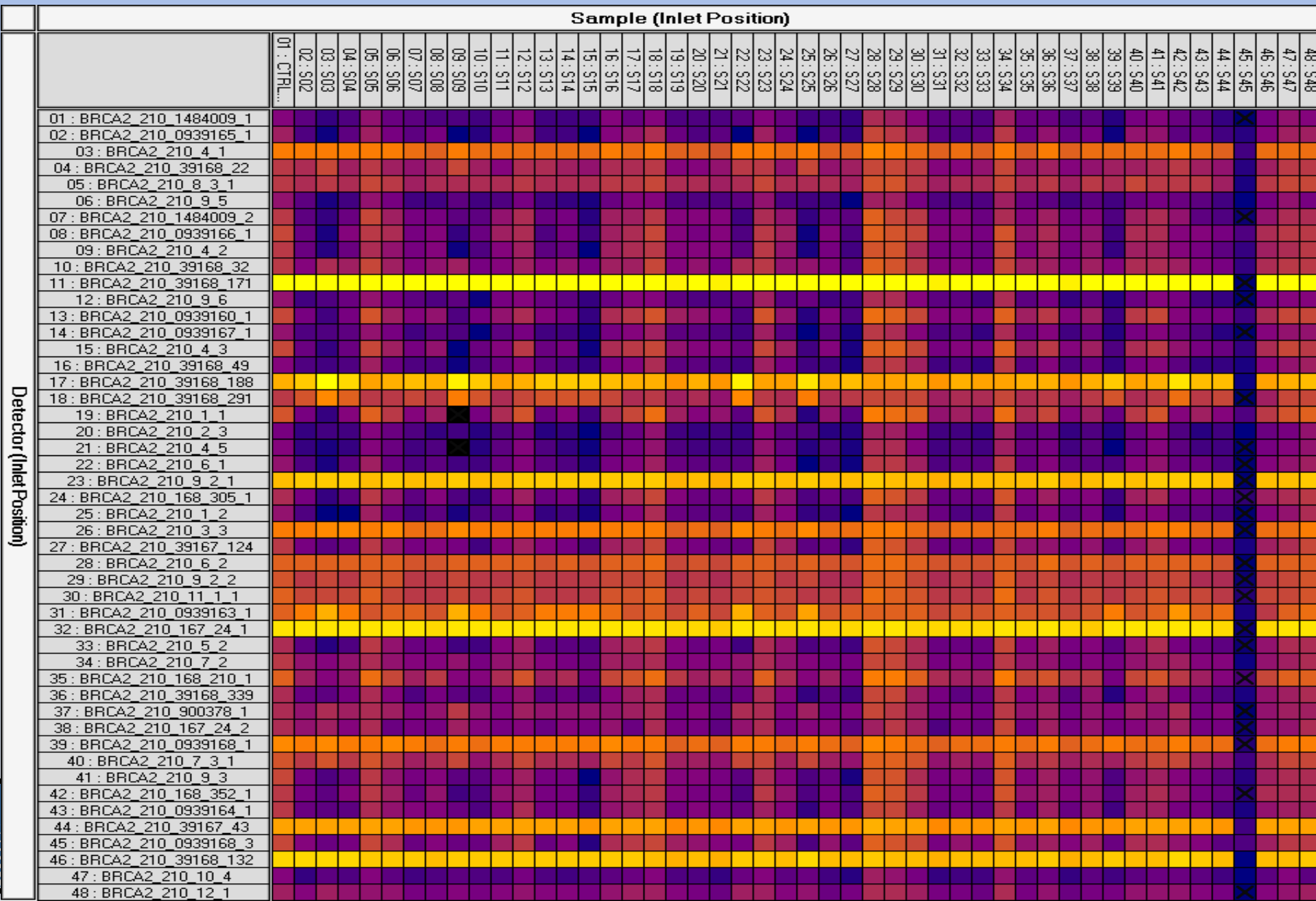
GENOSEQ
UCLA GENOTYPING & SEQUENCING

Roche Amplicon Tagging

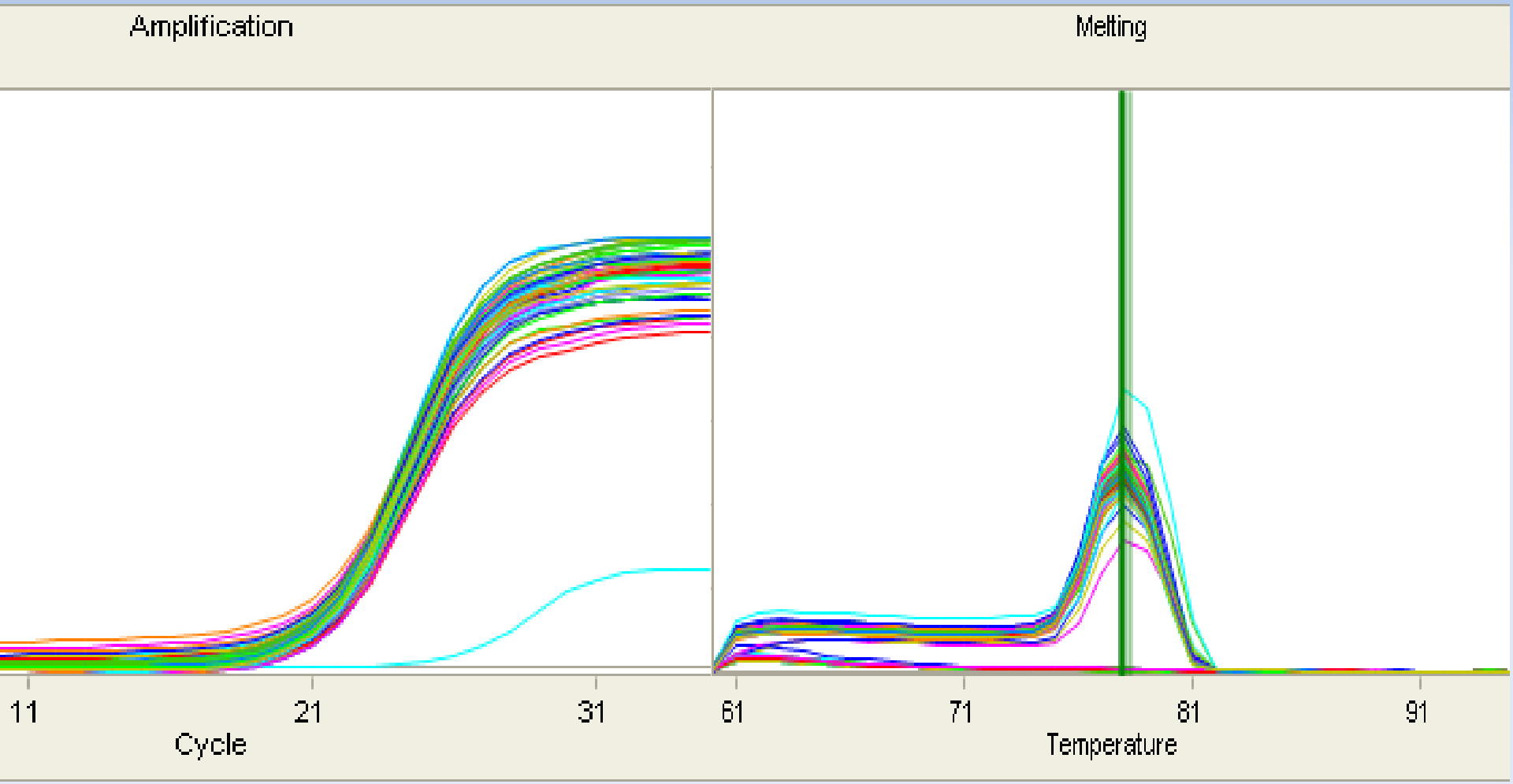


GENOSEQ
UCLA GENOTYPING & SEQUENCING

Real-Time Heat Map-BRCA Panel



Example Melt Curve Data BRCA2_210_6_2

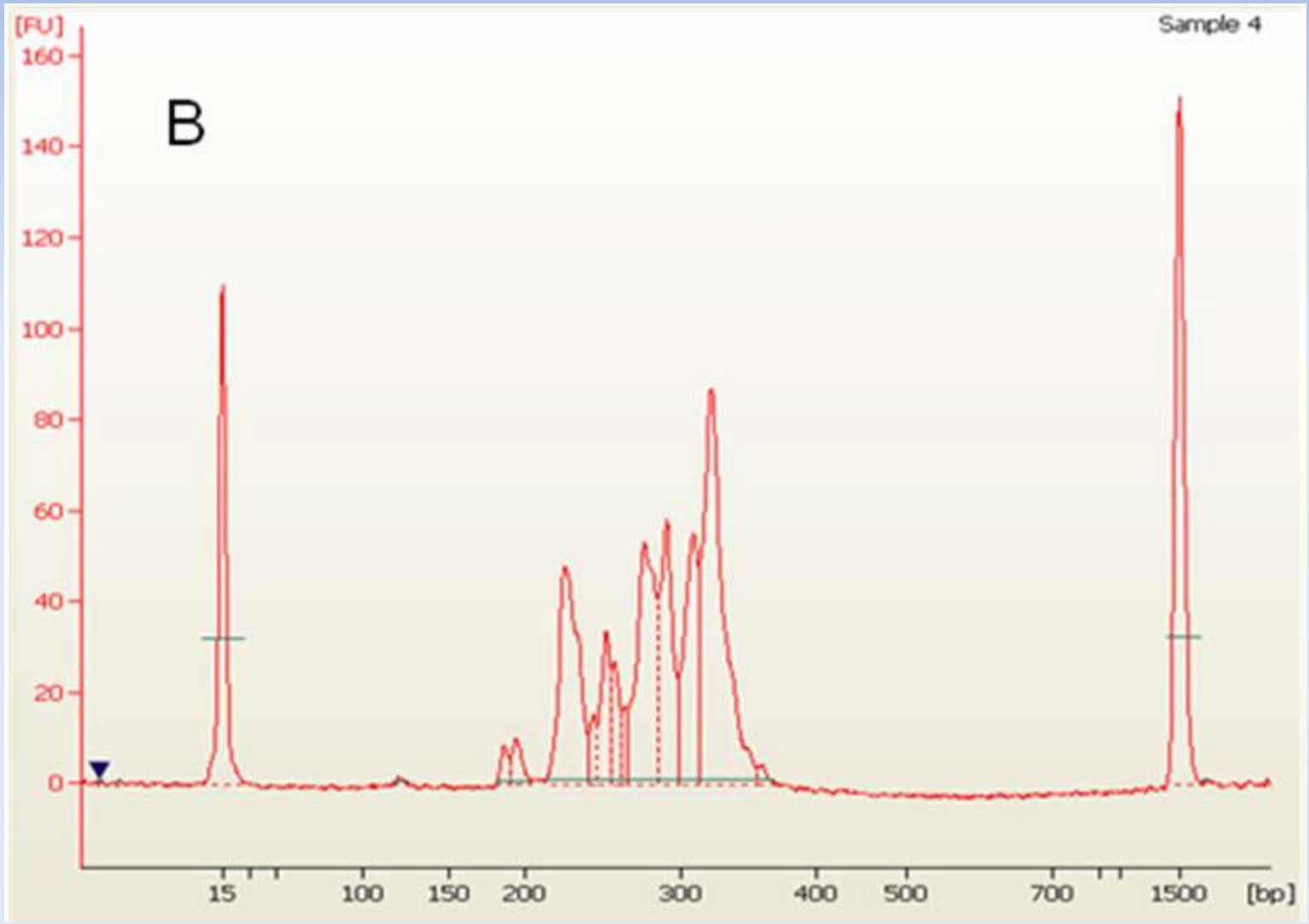


GENOSEQ
UCLA GENOTYPING & SEQUENCING

2-Step Access Array
4-Primer Amplicon Tagging Protocol

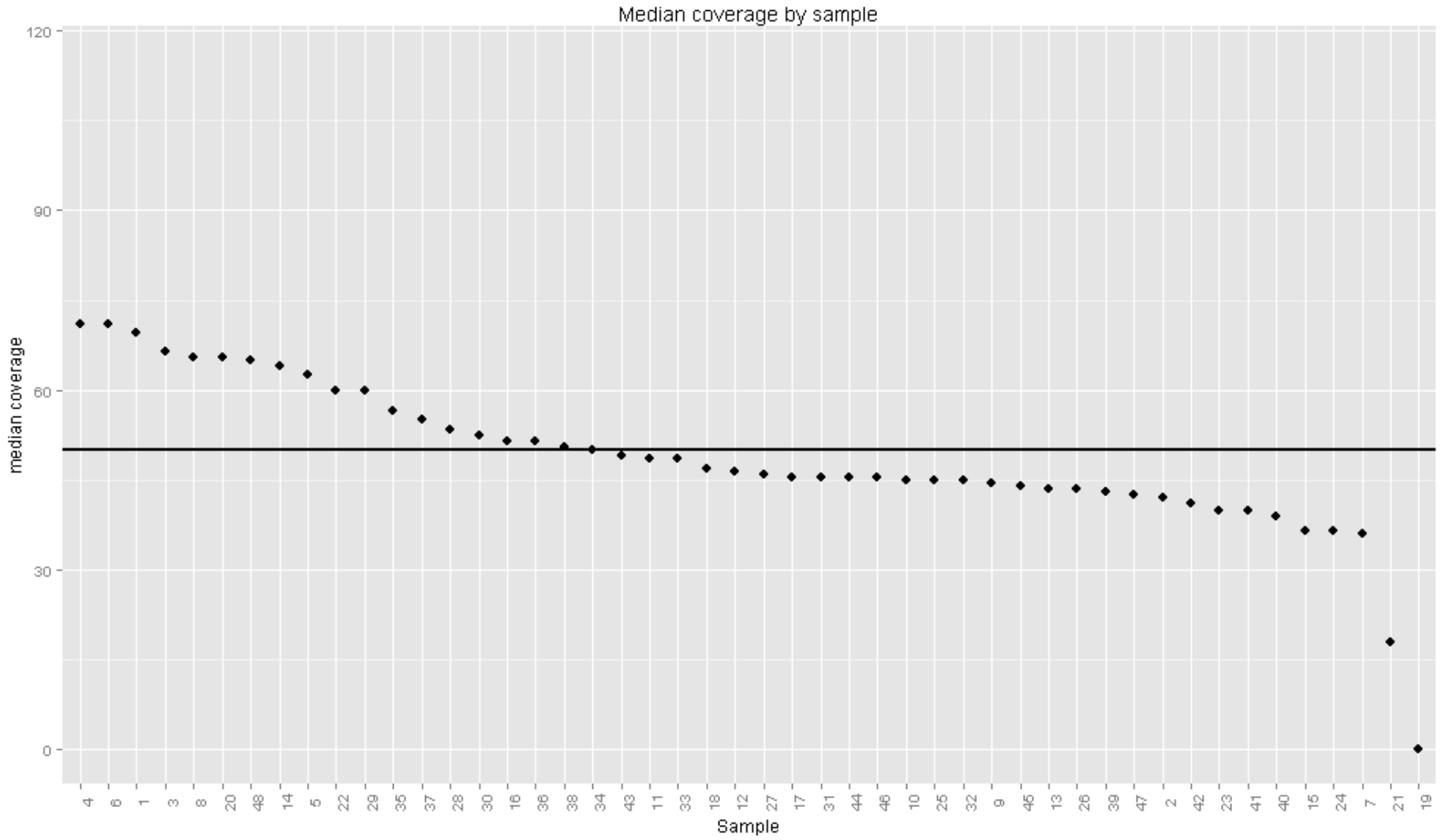
GENOSEQ
UCLA GENOTYPING & SEQUENCING

Harvested PCR product pool



GENOSEQ
UCLA GENOTYPING & SEQUENCING

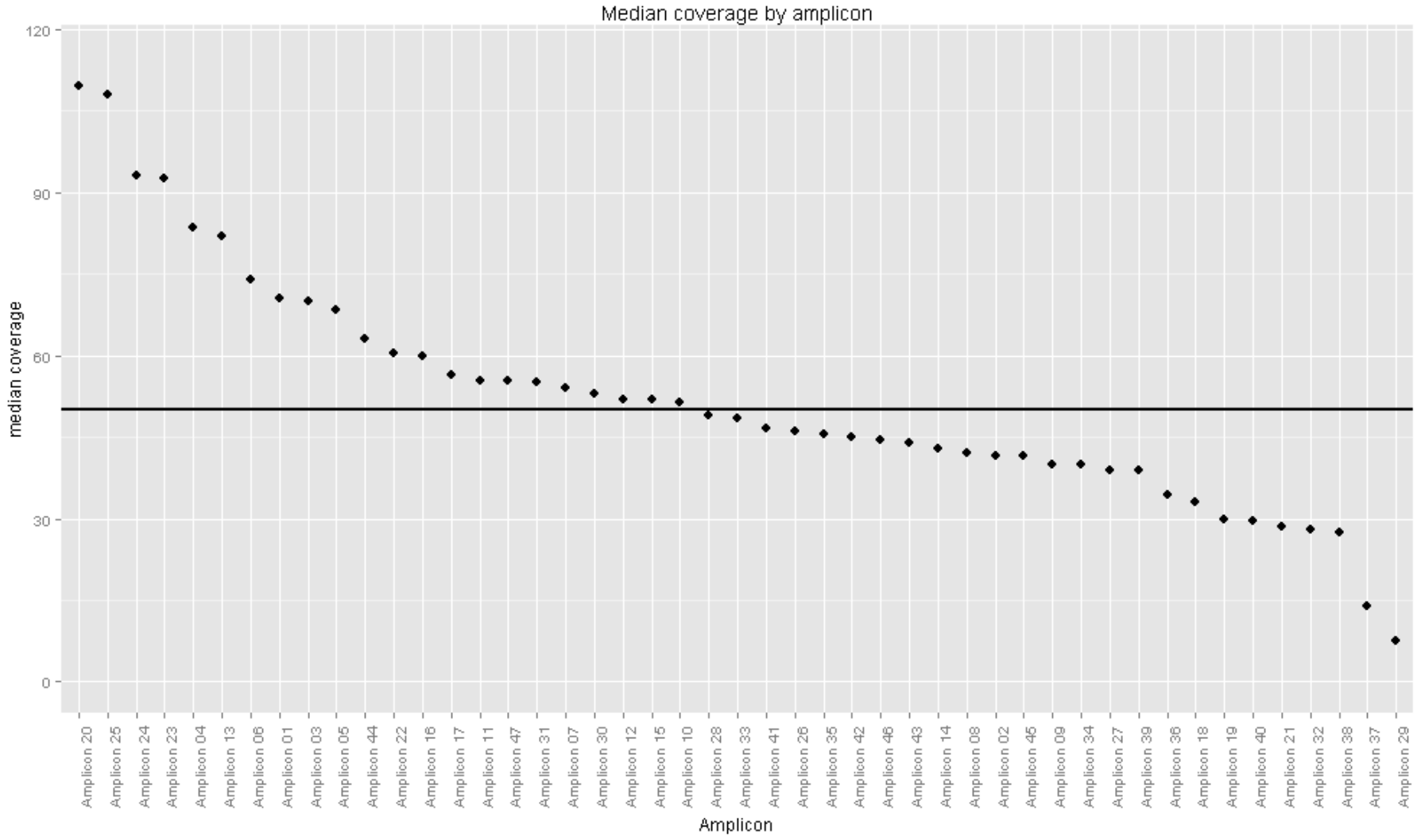
Sample Coverage



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Courtesy: J.Gleeson UCSD

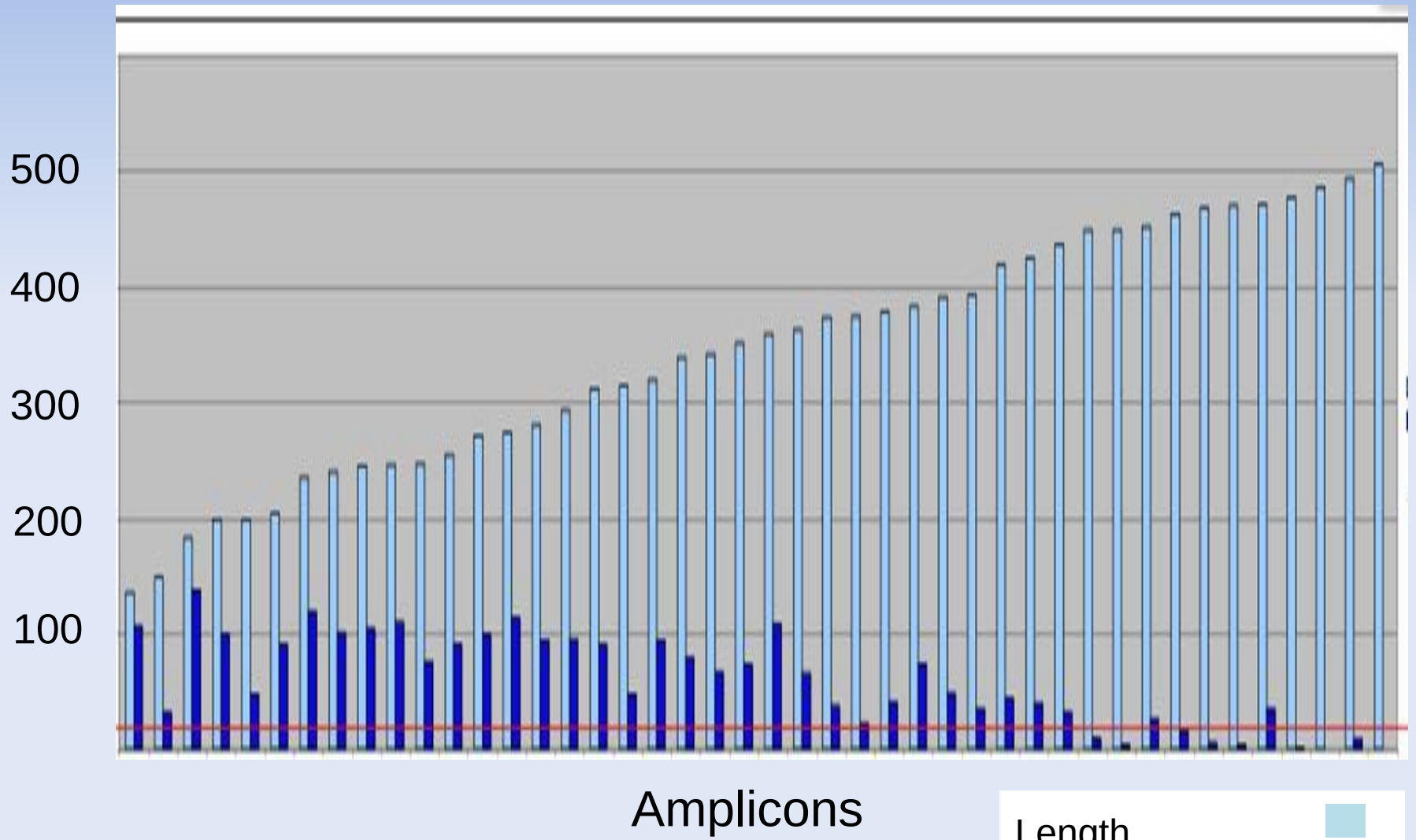
Amplicon coverage



GENOSEQ
UCLA GENOTYPING & SEQUENCING

Courtesy: J.Gleeson UCSD

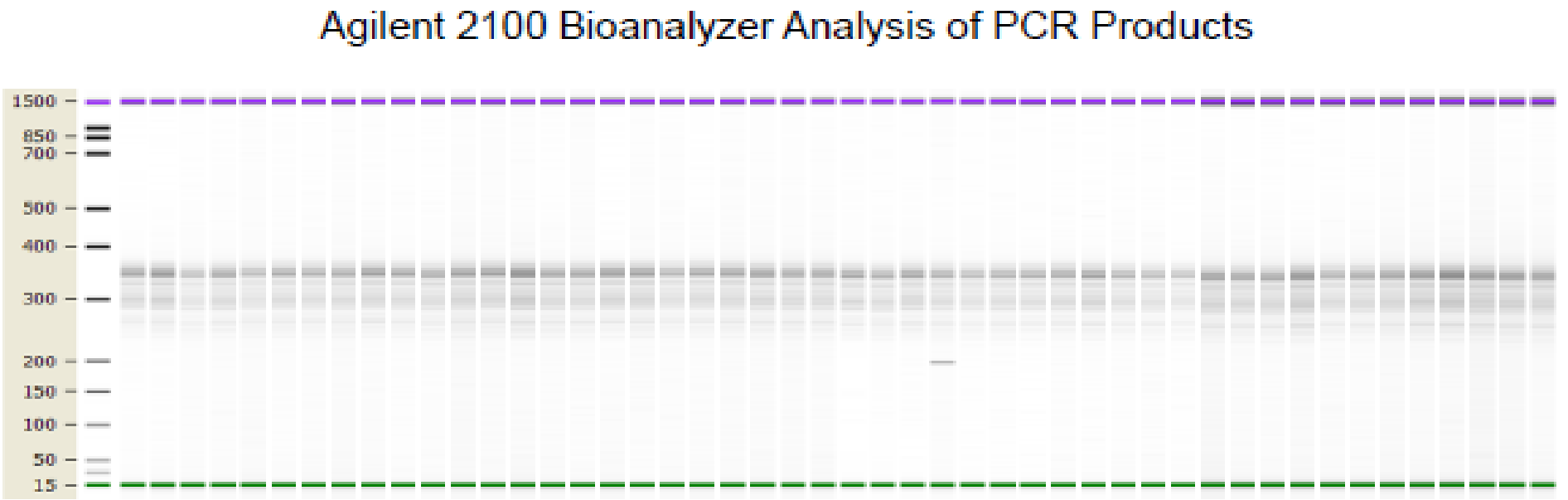
Amplicon coverage

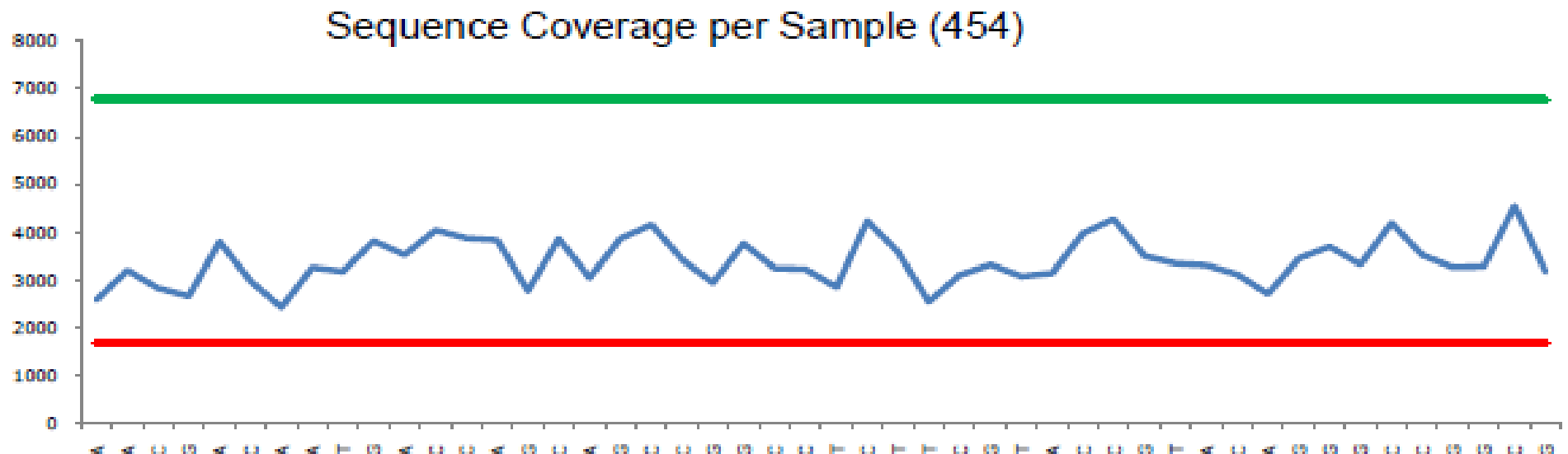


GENOSEQ
UCLA GENOTYPING & SEQUENCING

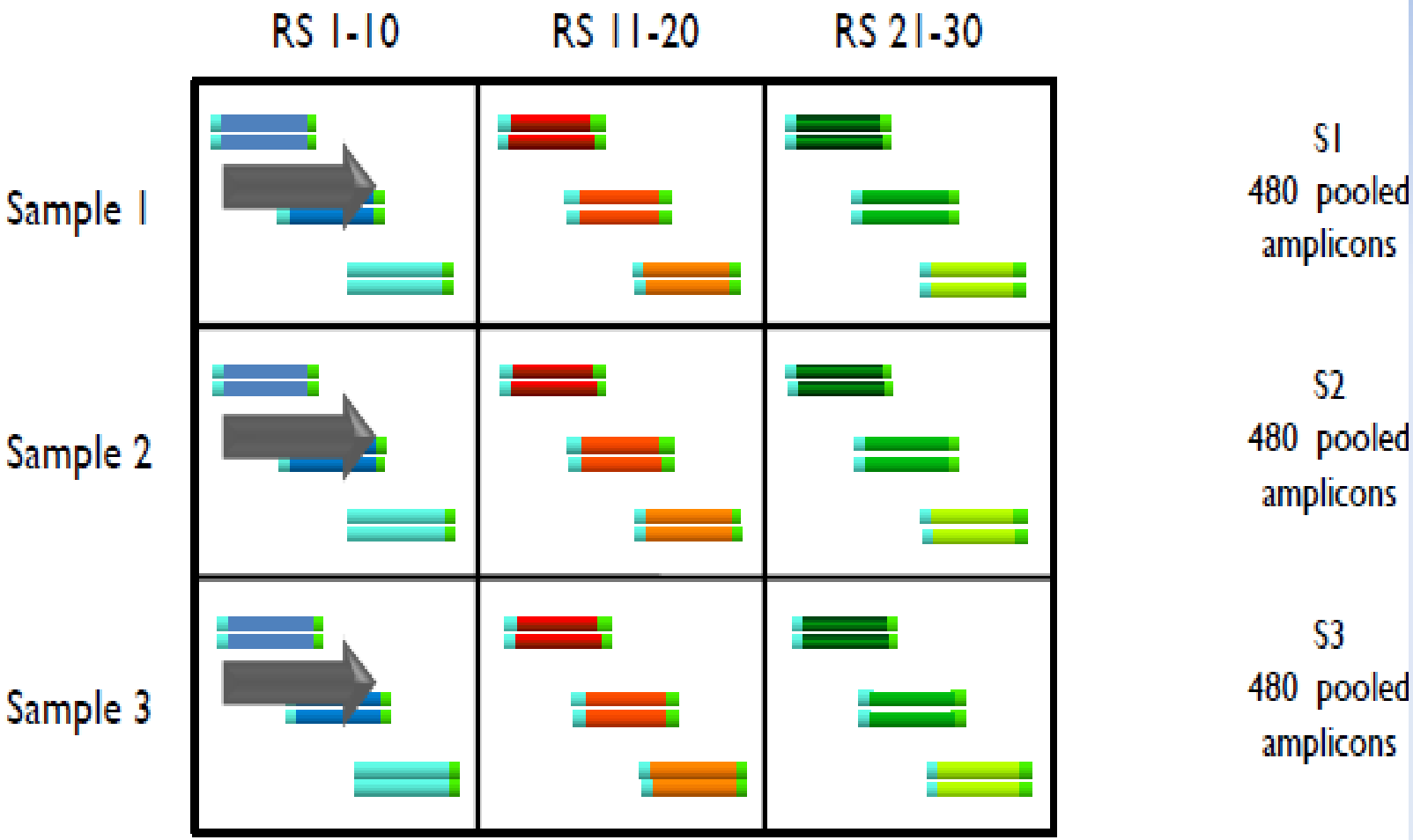
Courtesy: J.Gleeson UCSD

Sample Uniformity and Size





Ten primers per inlet



GENOSEQ
UCLA GENOTYPING & SEQUENCING

High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing

Jan Halbritter, Katrina Diaz, Moumita Chaki, et al. *J Med Genet* 2012 49: 756-767

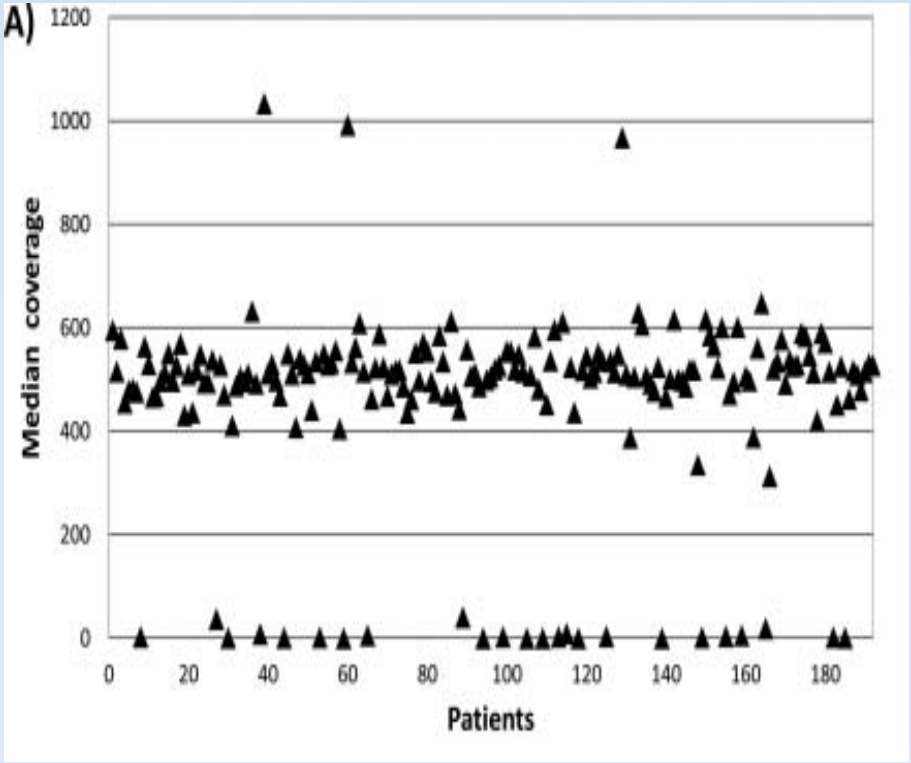
Multiplex approach using 10 primer pairs per single primer inlet

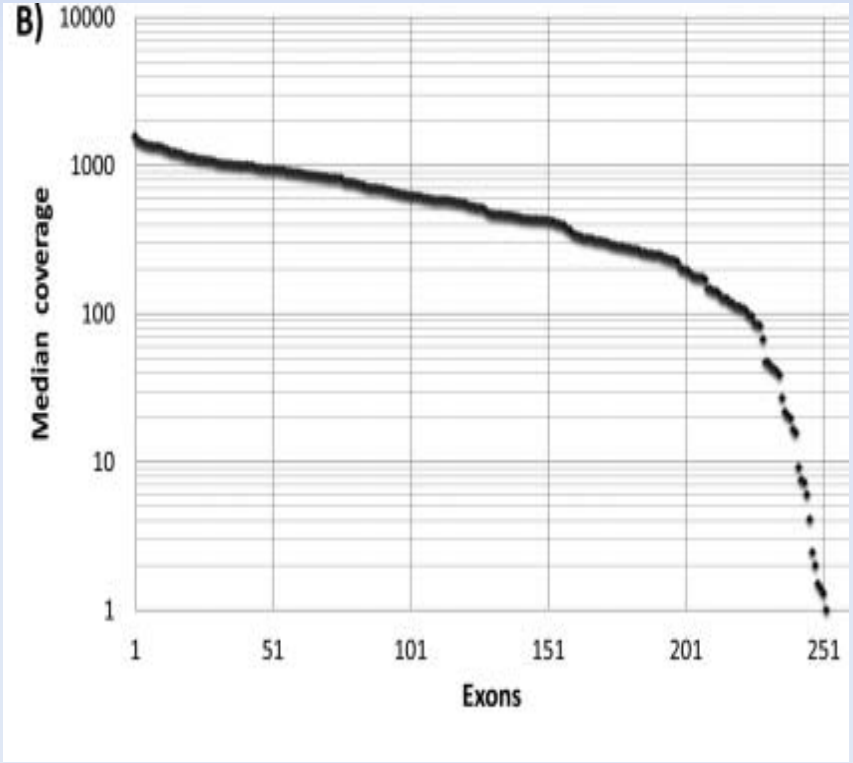
- 192 patient samples, 475 amplicons (251 coding exons)
- Amplified $48 \times 480 = 23\,040$ different amplicons per array.
- Used four Fluidigm access arrays.
- Sequencing done on one lane of a flow-cell on a HiSeq2000

GENOSEQ
UCLA GENOTYPING & SEQUENCING

High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing

Jan Halbritter, Katrina Diaz, Moumita Chaki, et al. *J Med Genet* 2012 49: 756-767





GENOSEQ
UCLA GENOTYPING & SEQUENCING

High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing

Jan Halbritter, Katrina Diaz, Moumita Chaki, et al. *J Med Genet* 2012 49: 756-767

Author's Conclusions

- NGS platforms do generate sequencing base call errors.
- Sequence quality scores decline with read length.
- PCR artifacts, especially shorter products generated with unspecific primers present within the multiplex pools, can mimic false variants after alignment.
- To rule out false positives always confirm potential mutations by Sanger sequencing using the original stock DNA sample as template.
- Fluidigm/NGS approach include occasional insufficient coverage due to failed exon amplification, or because of the failure of some low-quality DNA samples.

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Library prep comparison

Technology	Sample Number	Size of region	Time required	Ease of Use
Agilent Haloplex	48-96	1-500kb	1-2 days	Long protocol
Agilent Sureselect	16-480	1kb-25Mb-exomes	1-2days	Long Protocol
Nimblegen EZ cap	24-96	100kb-50Mb	Few days	Long Protocol
Raindance	1-96	150bp-10kb	1 day	Short protocol
TrueSeq Custom Amplicon	1-96	2kb-650kb	2 days	Long protocol
Roche Fusion Primers	1-96	200bp-variable	2 days	Very Long Protocol
Access Array	Multiples of 48	24kb-240kb	1 day	Short protocol

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Summary

- 50-60ng quantity of DNA required
- Wet lab tested primers reduce hands on time
- Short preparation time
- Robust and even amplification
- Ease of Indexing/Barcoding
- Ease of multiplexing

Capture per Sample	Capture Per Array
Amplicon Tagging*	24kb
Multiplexed Amplicon Tagging*	240kb
Long Range PCR	480kb

GENOSEQ
UCLA GENOTYPING & SEQUENCING

Genoseq Core:

- Jeanette Papp Director
- Uma Dandekar Asst.Director
- Hemani Wijesuriya
- Vivian Knight
- Chrissie Bandong
- Calvin Pan

Webseq@genetics.ucla.edu

Fluidigm:

Steven Hoffman

Susan Robelli

GENOSEQ
UCLA GENOTYPING & SEQUENCING