Welcome to the PHC Webinar Series

This lecture on **"Next-Generation Sequencing for the Clinical Laboratory"** is given by Karl V. Voelkerding, MD, FCAP

Your host is Jill Kaufman, PhD.
For comments about this webinar or suggestions for upcoming webinars, please contact Jill Kaufman at jkaufma@cap.org

THE WEBINAR WILL BEGIN MOMENTARILY. ENJOY!

© 2010 College of American Pathologists. All rights reserved.

Karl Voelkerding, MD, FCAP

- Professor of Pathology at the University of Utah
- Medical Director for Genomics and Bioinformatics at the ARUP Laboratories
- Past President of the Association for Molecular Pathology
- Board certified in Clinical and Molecular Genetic Pathology
- His research interests include translation of nucleic acid based technologies into diagnostics with a current focus on complex genetic analyses by next generation sequencing

© 2010 College of American Pathologists. All rights reserved.

Place sub-brand here

Next-Generation Sequencing for the Clinical Laboratory

Karl V. Voelkerding, MD, FCAP

July 20, 2011

Disclaimer

The College does not permit reproduction of any substantial portion of the material in this Webinar without its written authorization. The College hereby authorizes attendees of the CAP Webinar to use the pdf presentation solely for educational purposes within their own institutions. The College prohibits use of the material in the Webinar – and any unauthorized use of the College's name or logo – in connection with promotional efforts by marketers of laboratory equipment, reagents, materials, or services.

Opinions expressed by the speaker are the speaker's own and do not necessarily reflect an endorsement by CAP of any organizations, equipment, reagents, materials or services used by participating laboratories.

© 2010 College of American Pathologists. All rights reserved.

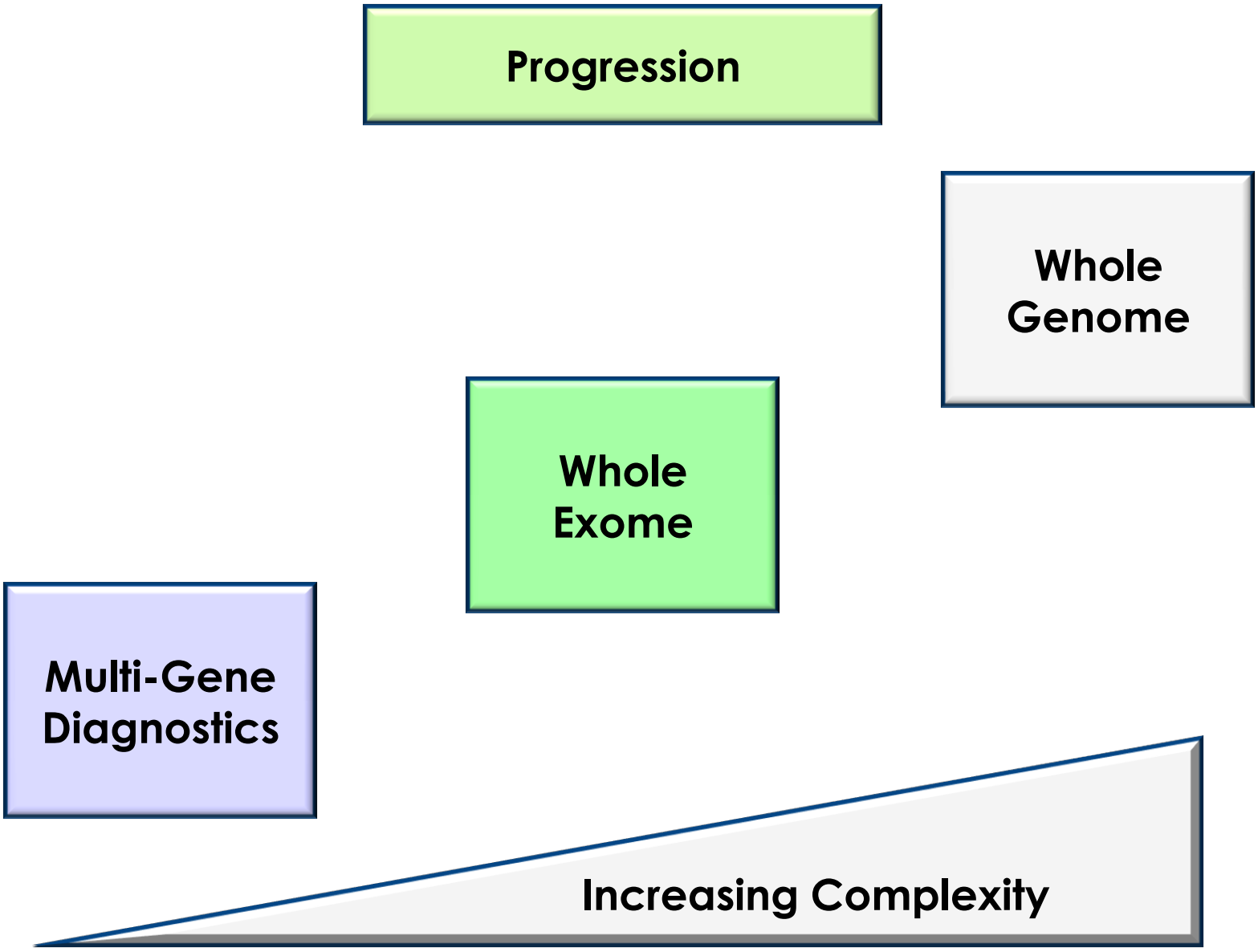
Disclosure

- I have nothing to disclose.

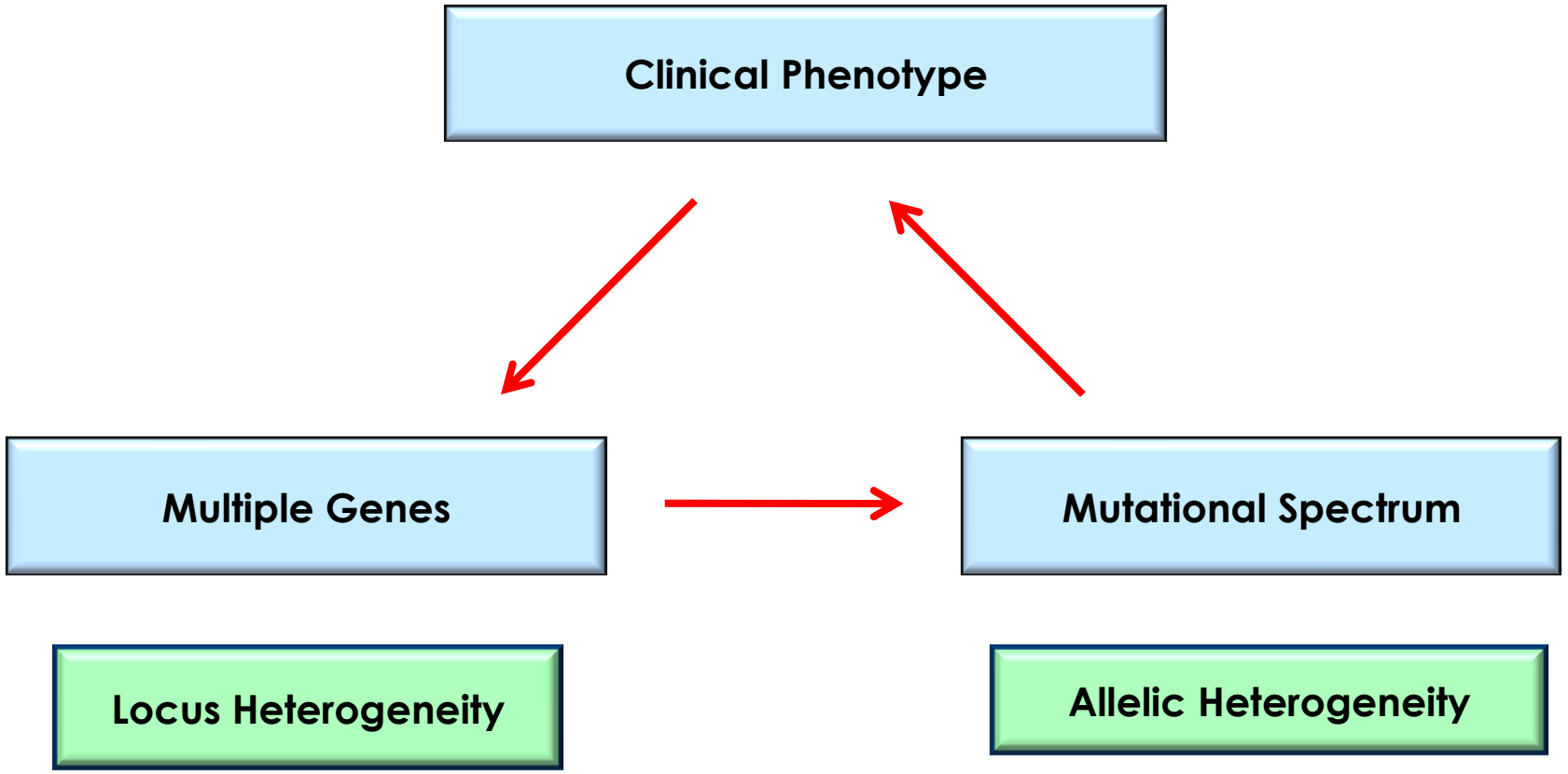
© 2010 College of American Pathologists. All rights reserved.

Outline

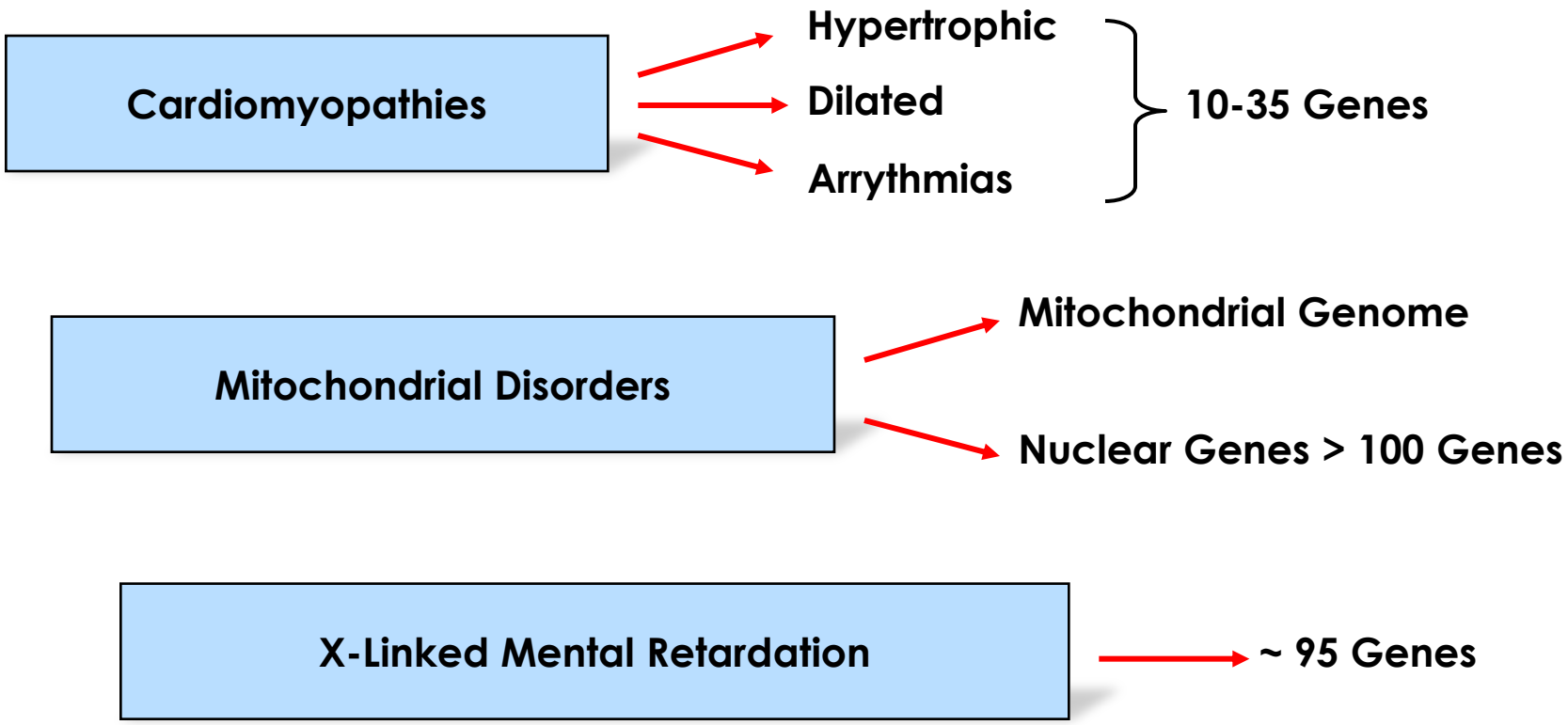
- **Progression: Gene Panels to Genomes**
- **Next Generation Sequencing Technology**
- **Bioinformatics**

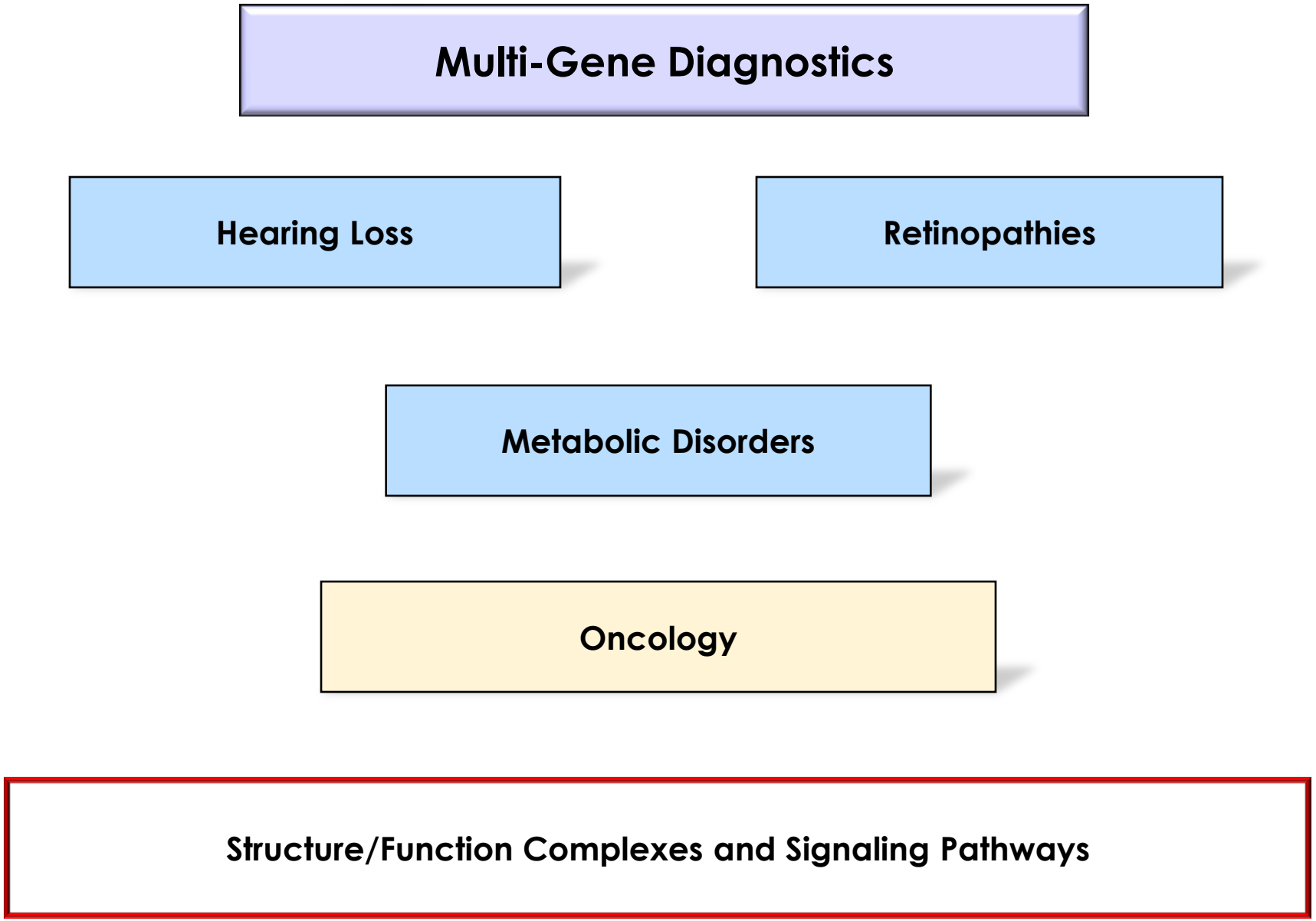

Progression
Whole Genome
Whole Exome
Multi-Gene Diagnostics
Increasing Complexity

Multi-Gene Diagnostics

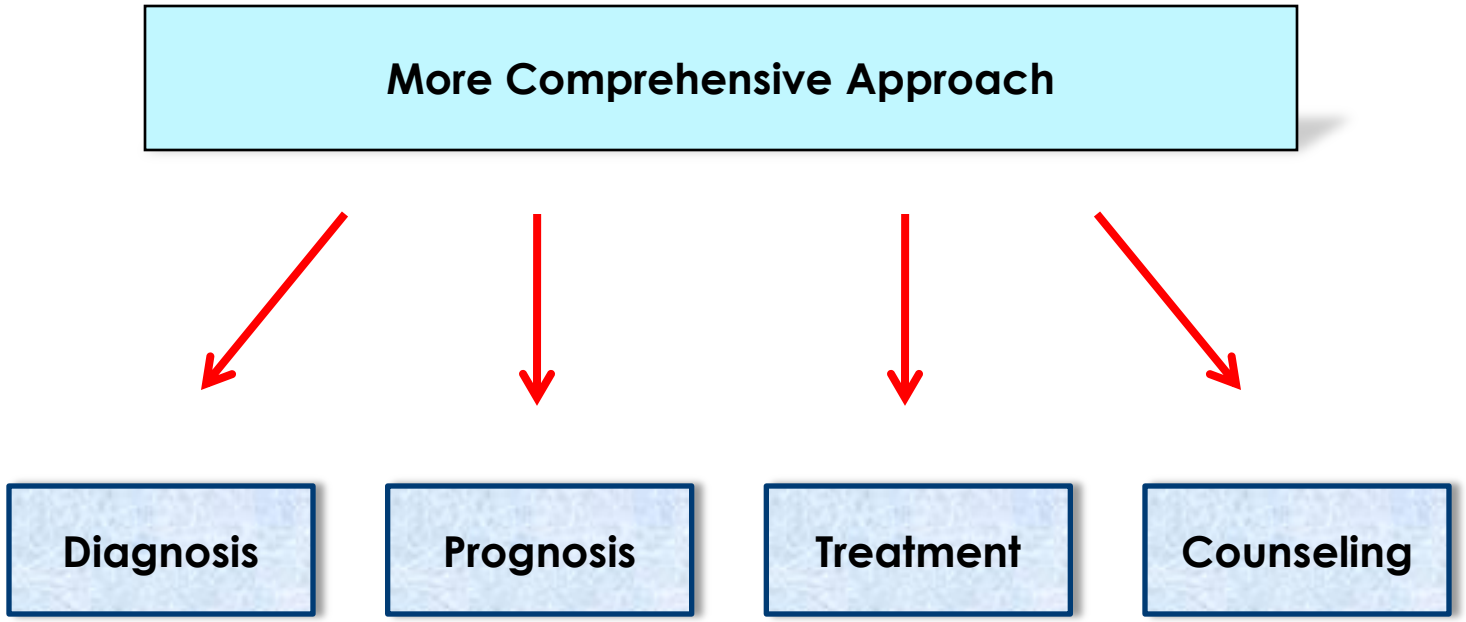


Multi-Gene Diagnostics




Multi-Gene Diagnostics
Hearing Loss
Retinopathies
Metabolic Disorders
Oncology
Structure/Function Complexes and Signaling Pathways

Multi-Gene Diagnostics



Multi-Gene Diagnostics

Technical Options

Sanger Sequencing of Individual Genes

Scanning and Sequencing

Multi-Gene Resequencing Microarrays

Next Generation Sequencing

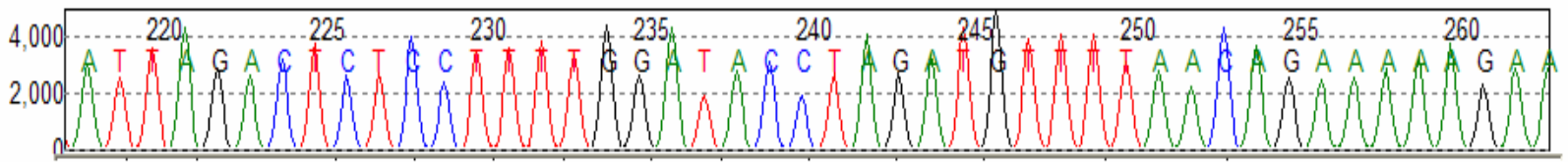
Outline

- Progression: Gene Panels to Genomes
- **Next Generation Sequencing Technology**
- Bioinformatics

Sanger Sequencing

PCR followed by Cycle Sequencing with dNTPs/ddNTPs

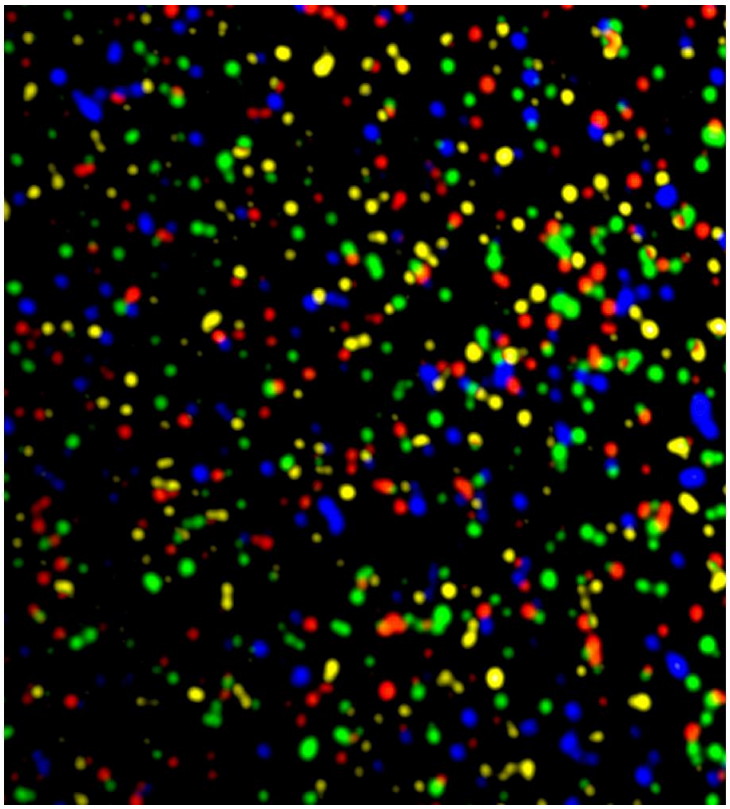
Electrophoretic separation of chain termination products

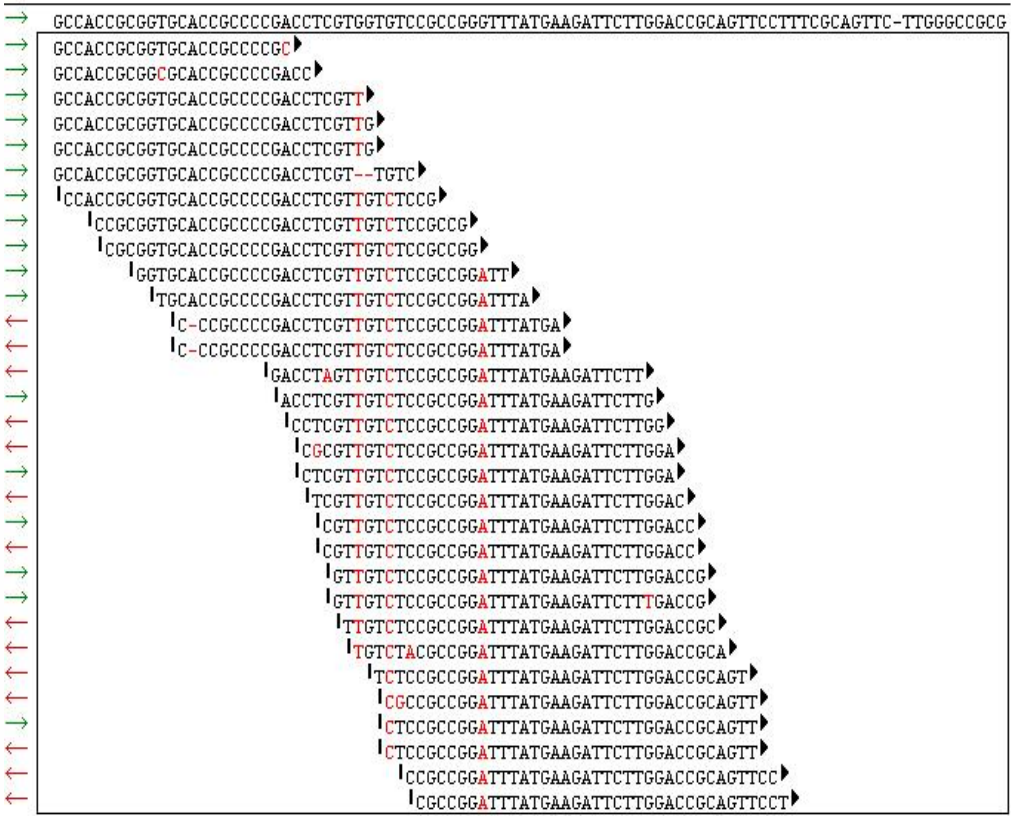


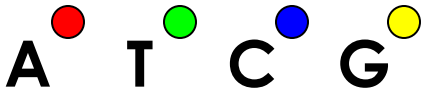
Next Generation Sequencing

Sequence DNA fragment library *in situ* in a flow cell

Massively parallel configuration







Process

Genomic DNA or Enriched Target Genes

Fragmentation

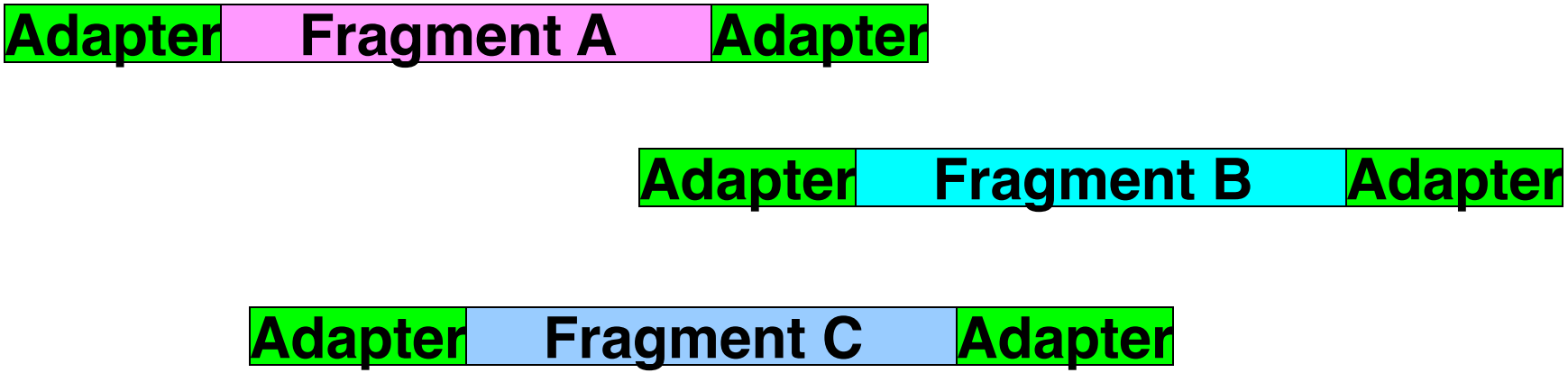
- Nebulization
- Sonication
- Acoustic Wave

150-500 bp

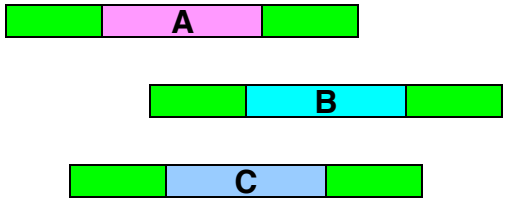
End Repair and Adapter Ligation

+/- PCR

"Fragment Library"



Process

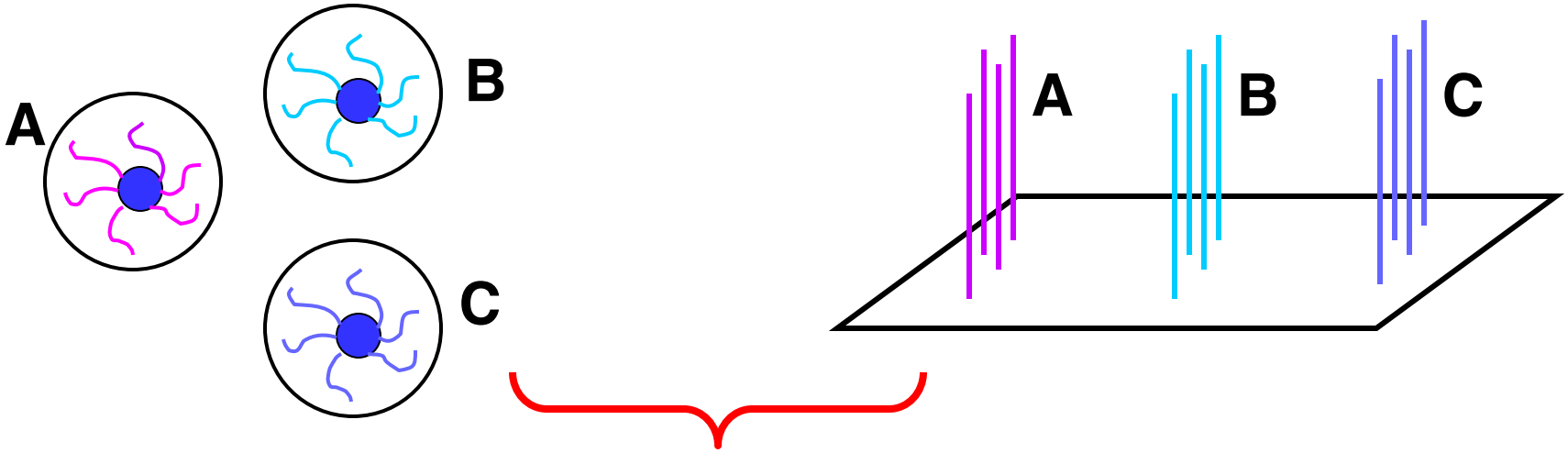


"Fragment Library"

Clonal Amplification of Each Fragment

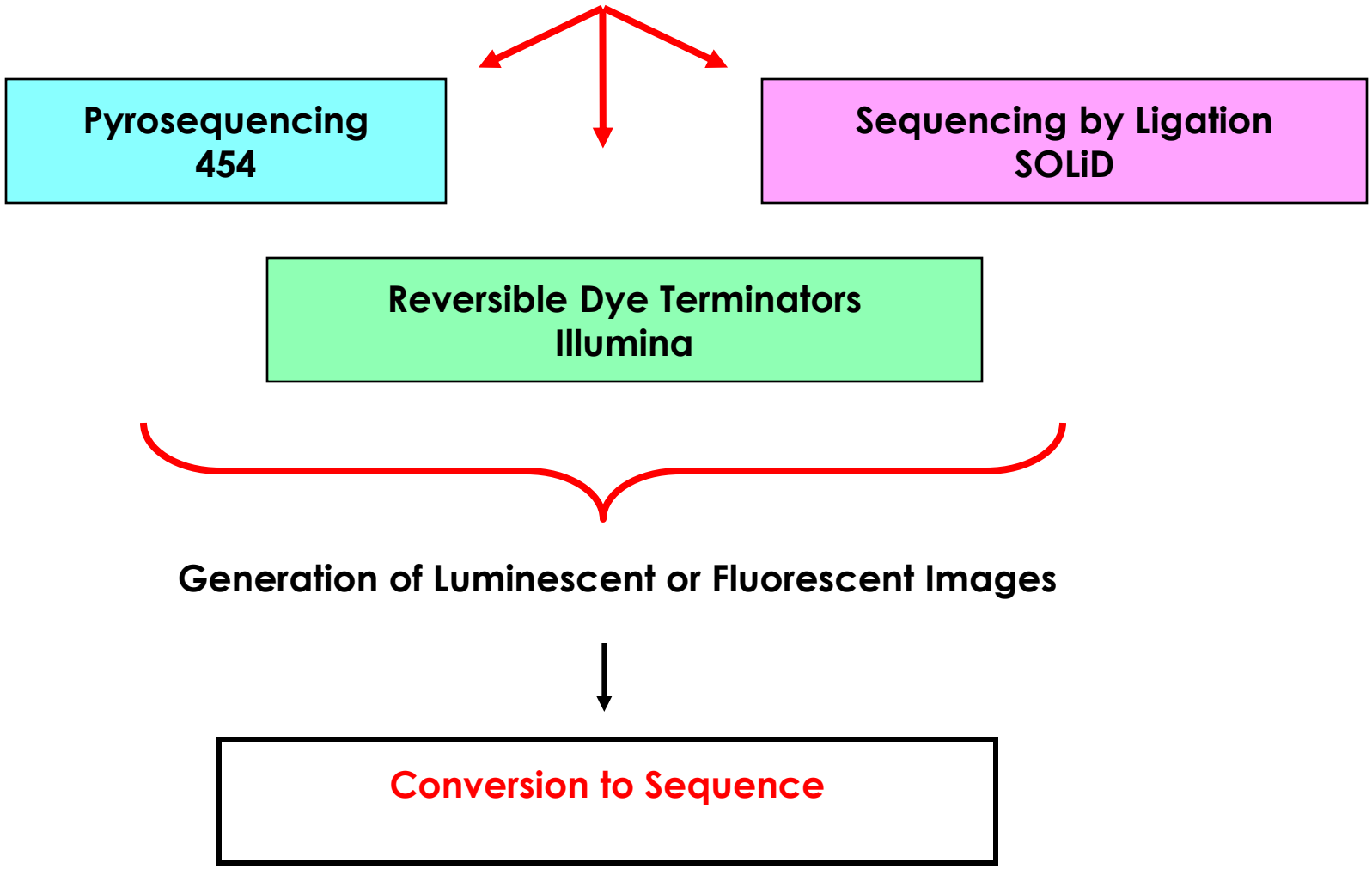
Emulsion Bead PCR

Surface Clusters

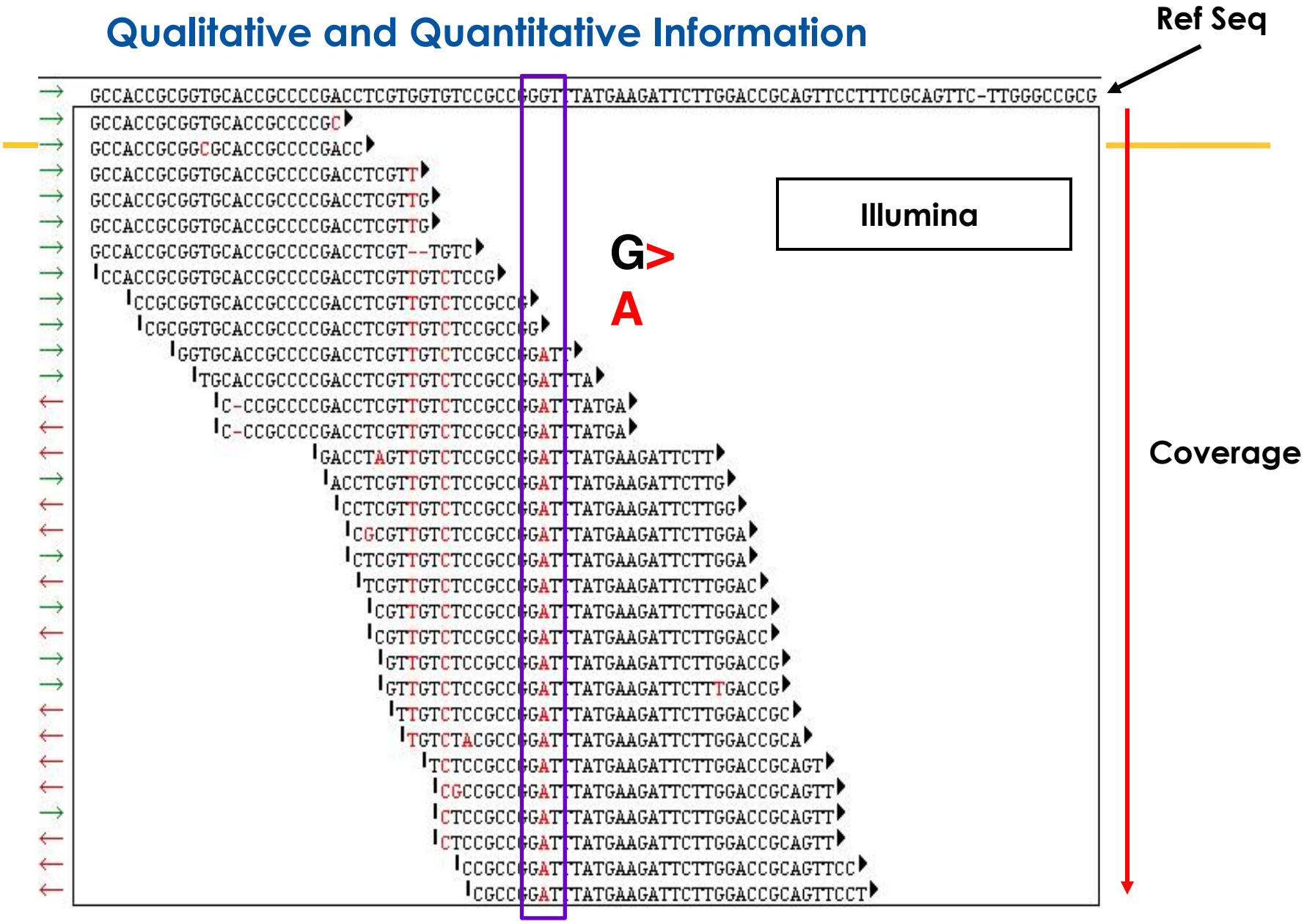


Sequencing of Clonal Amplicons in a Flow Cell

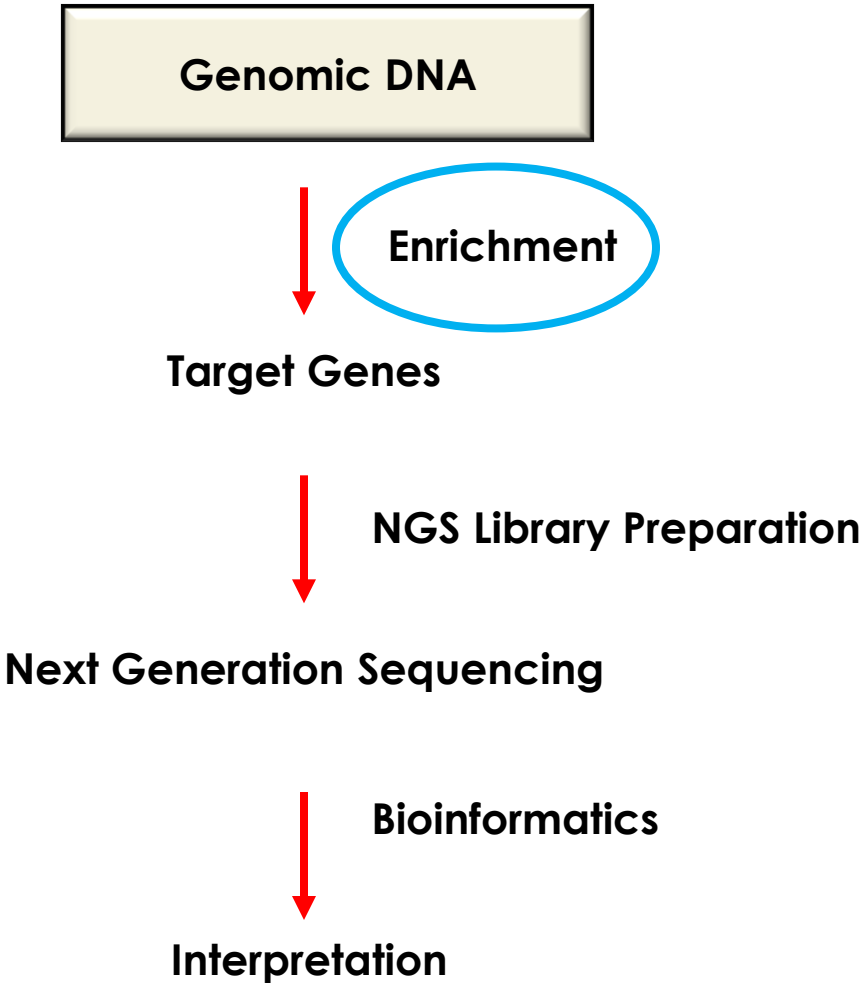
Process


Sequencing of Clonal Amplicons in a Flow Cell
Pyrosequencing
454
Sequencing by Ligation
SOLiD
Reversible Dye Terminators
Illumina
Generation of Luminescent or Fluorescent Images
Conversion to Sequence

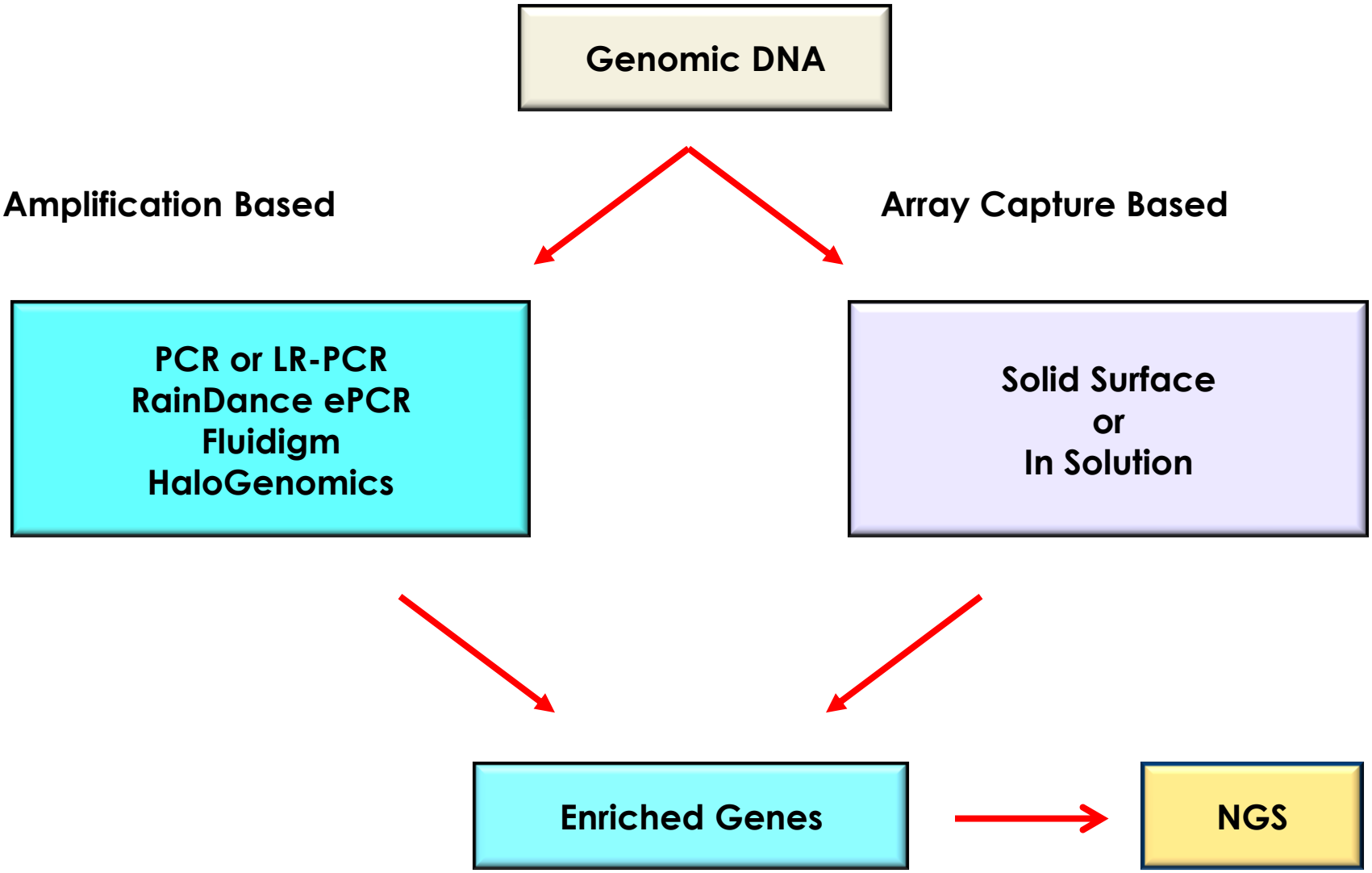
Qualitative and Quantitative Information



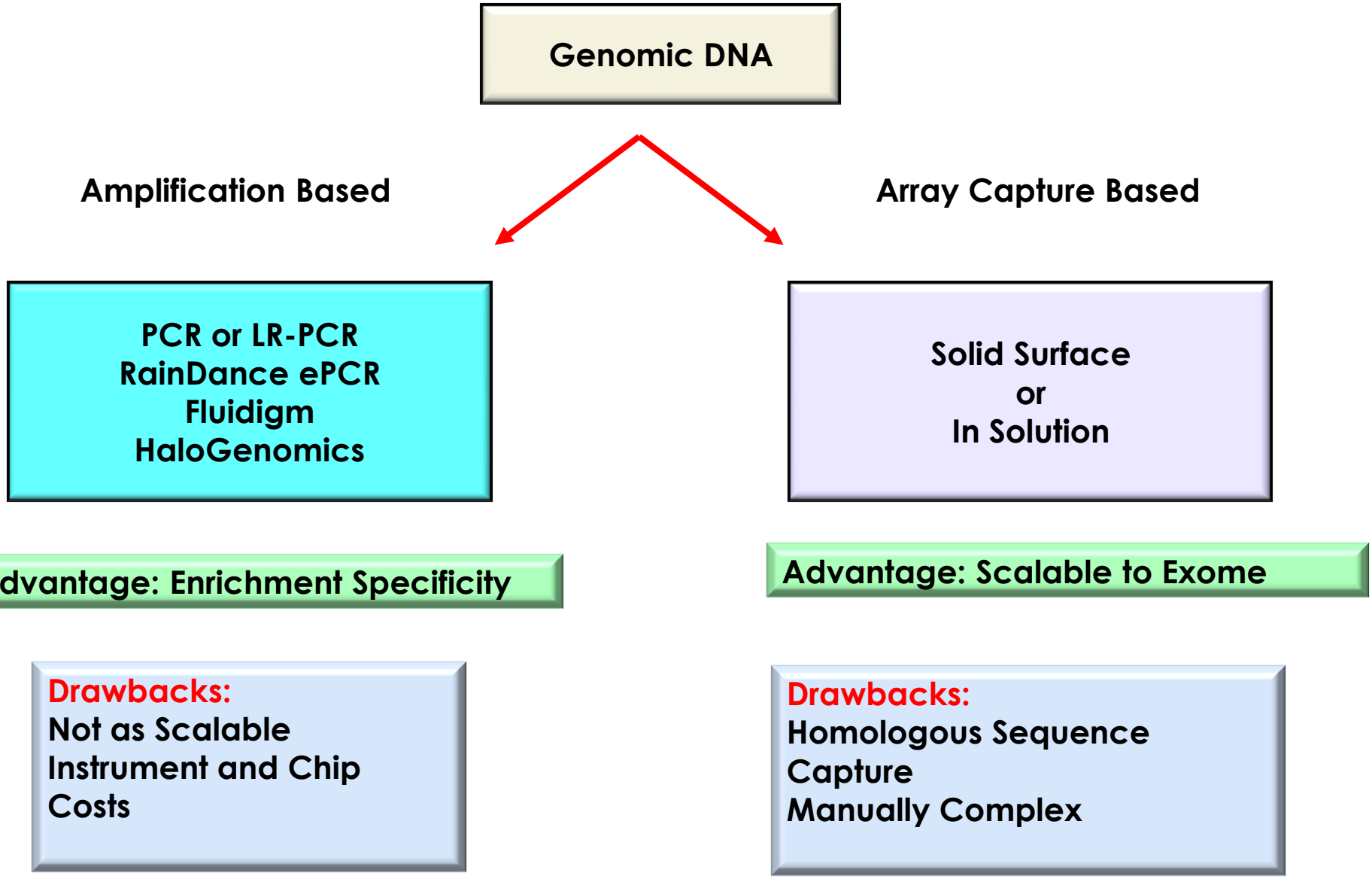
Multi-Gene Diagnostics



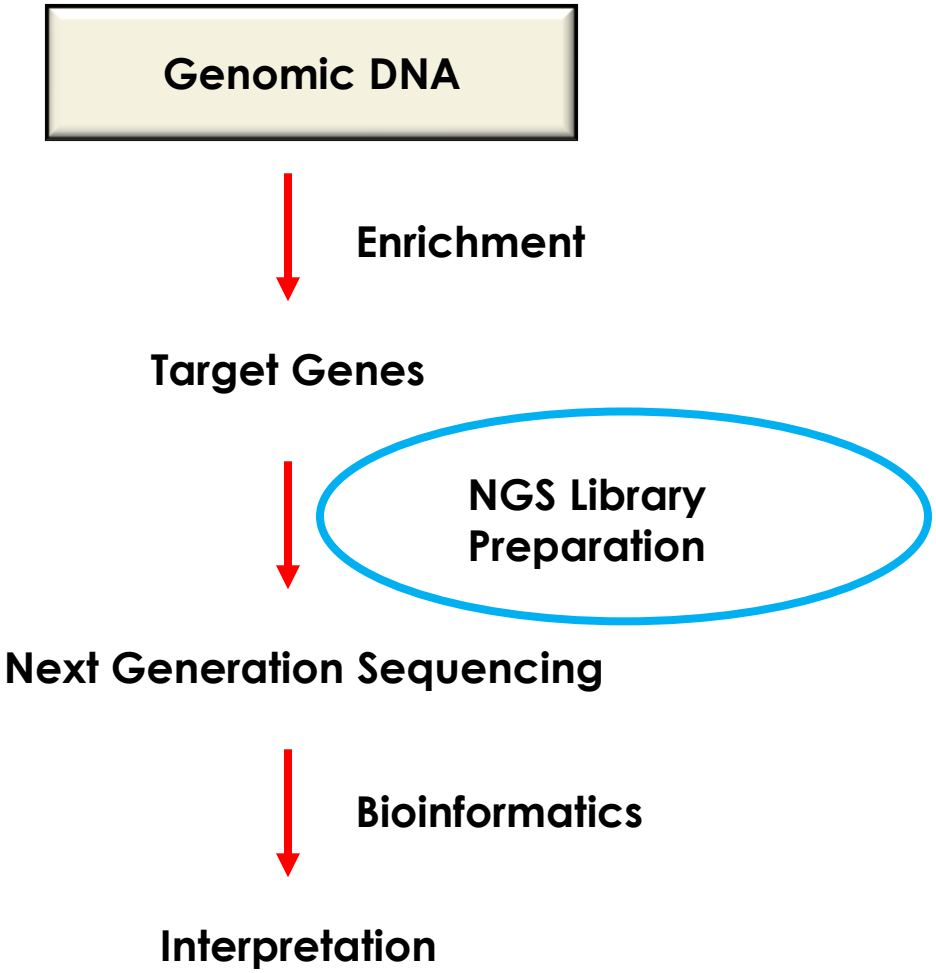
Gene Enrichment Approaches



Gene Enrichment Approaches



Multi-Gene Diagnostics



NGS Library Preparation

Genomic DNA

↓

Fragmentation

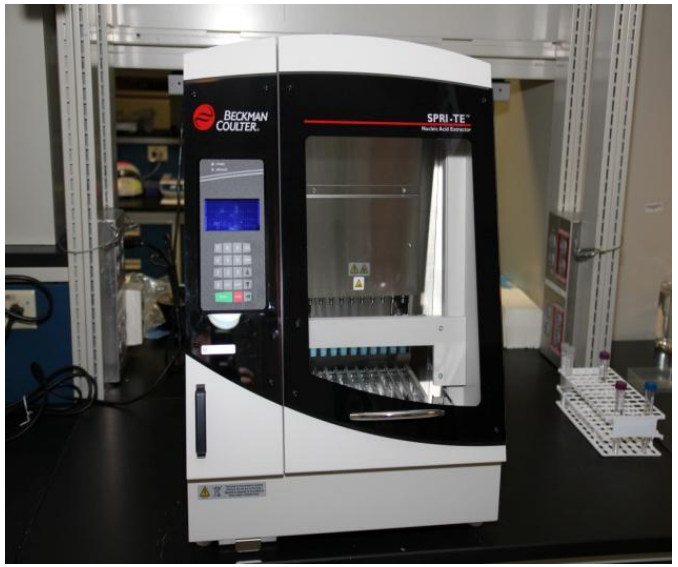
↓

Fragment End Repair

↓

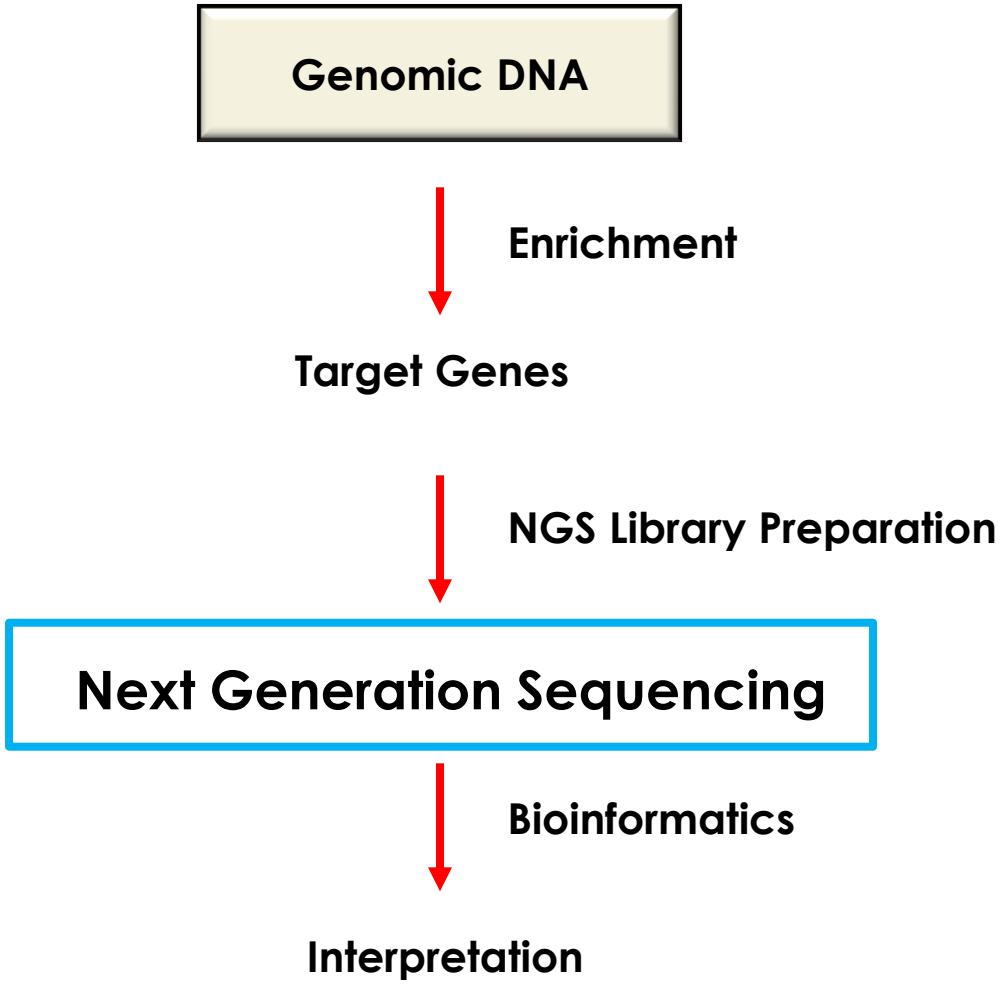
Adapter Ligation

Automation

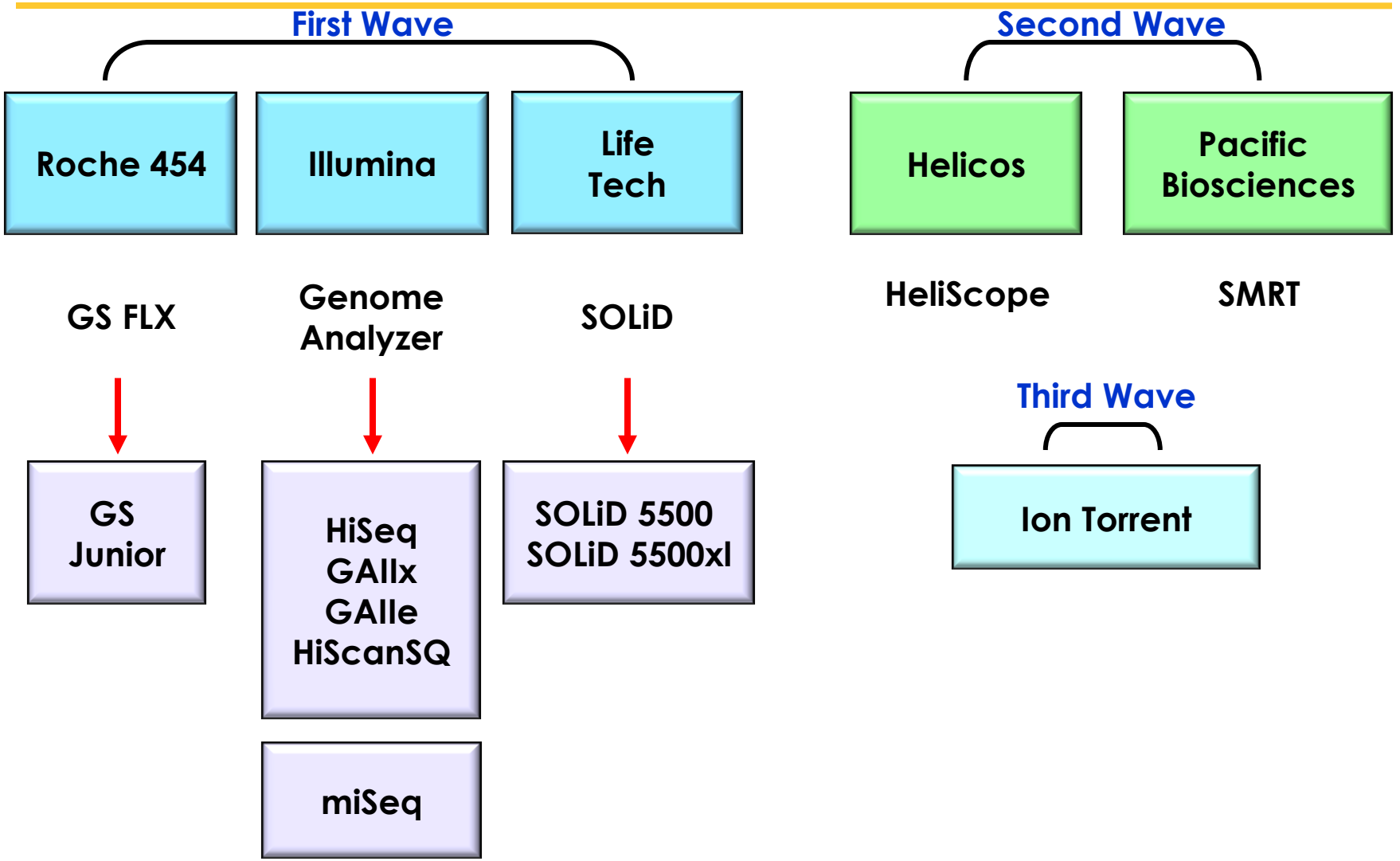


Beckman SPRI-TE
1-10 Samples

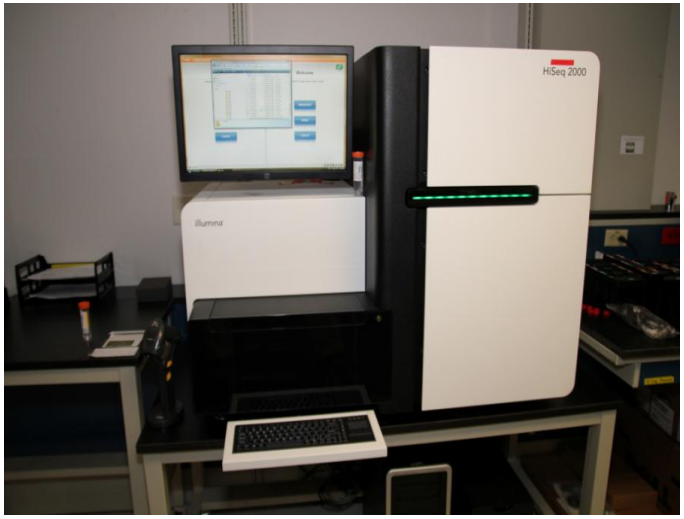
Multi-Gene Diagnostics



Next Generation Sequencing Technology



Illumina HiSeq 2000



Advantage: High Throughput

Independent Flow Cells
8 Lanes per Flow Cell

Drawbacks:
Batching
Sample Coordination

Multiple Panel Samples per Lane

1- 3 Exome(s) per Lane

1- 2 Genome per 8 Lanes

New Platforms
Lower Throughput - Faster TAT
"Random Access"

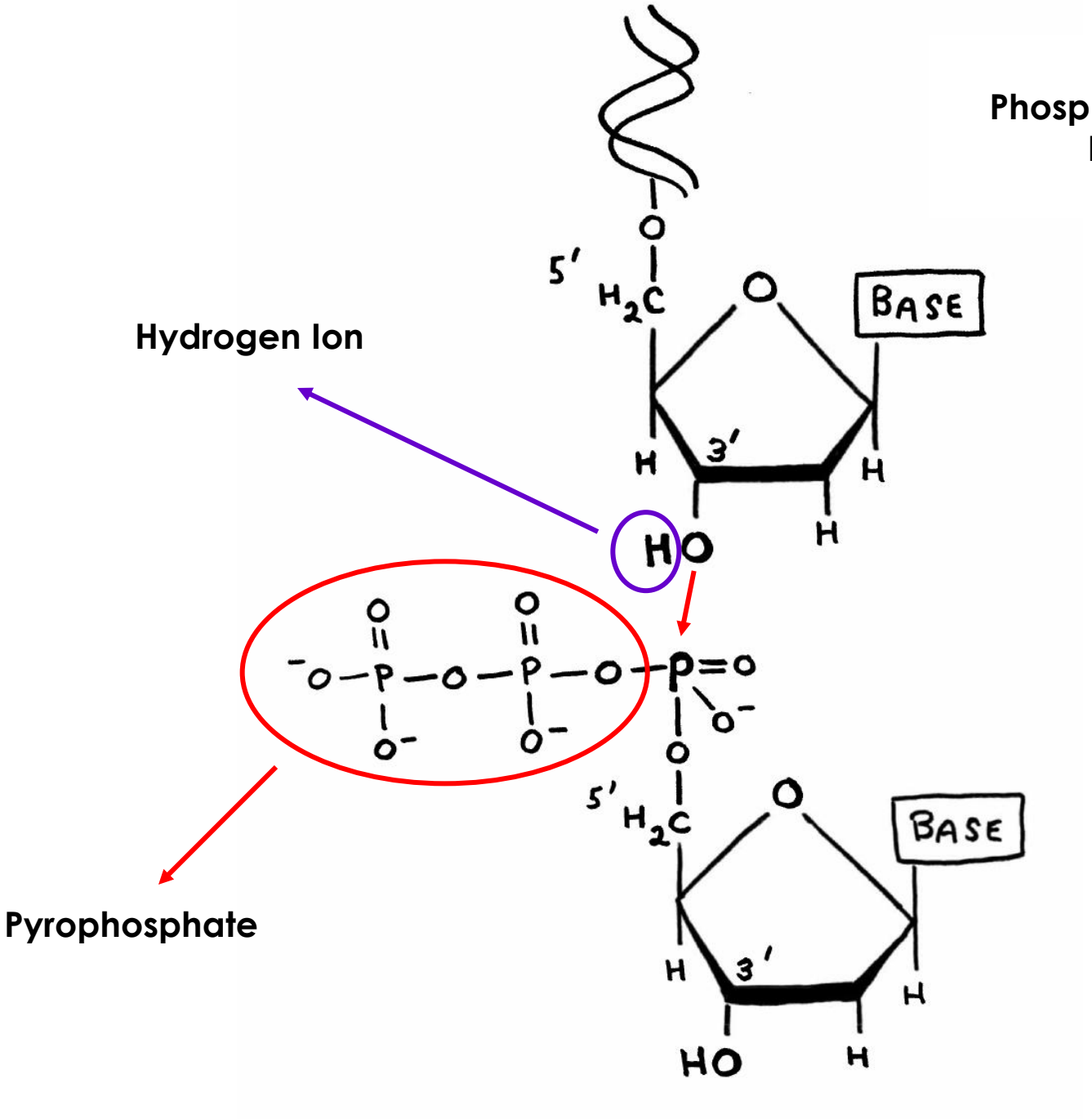
Illumina miSeq

Ion Torrent PGM

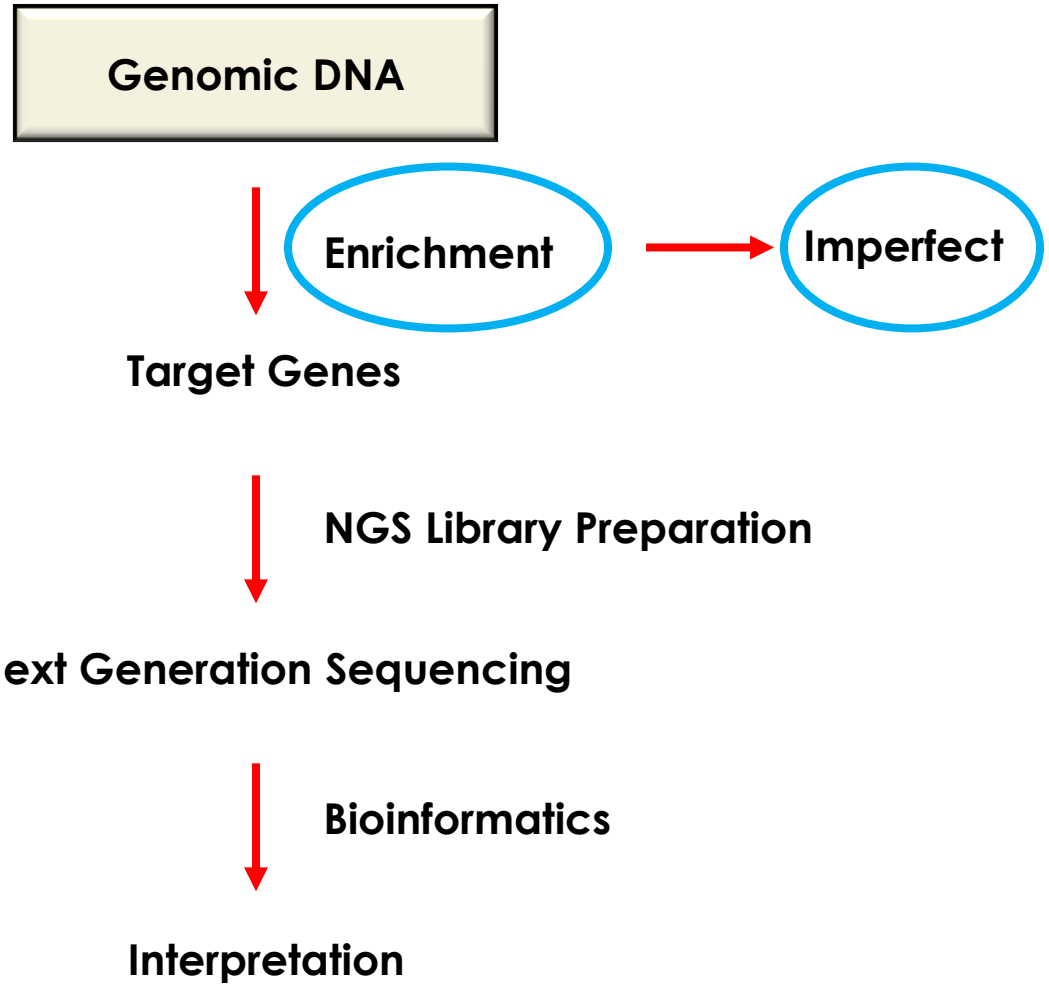
Reversible Dye Terminators

Monitors H+ Release

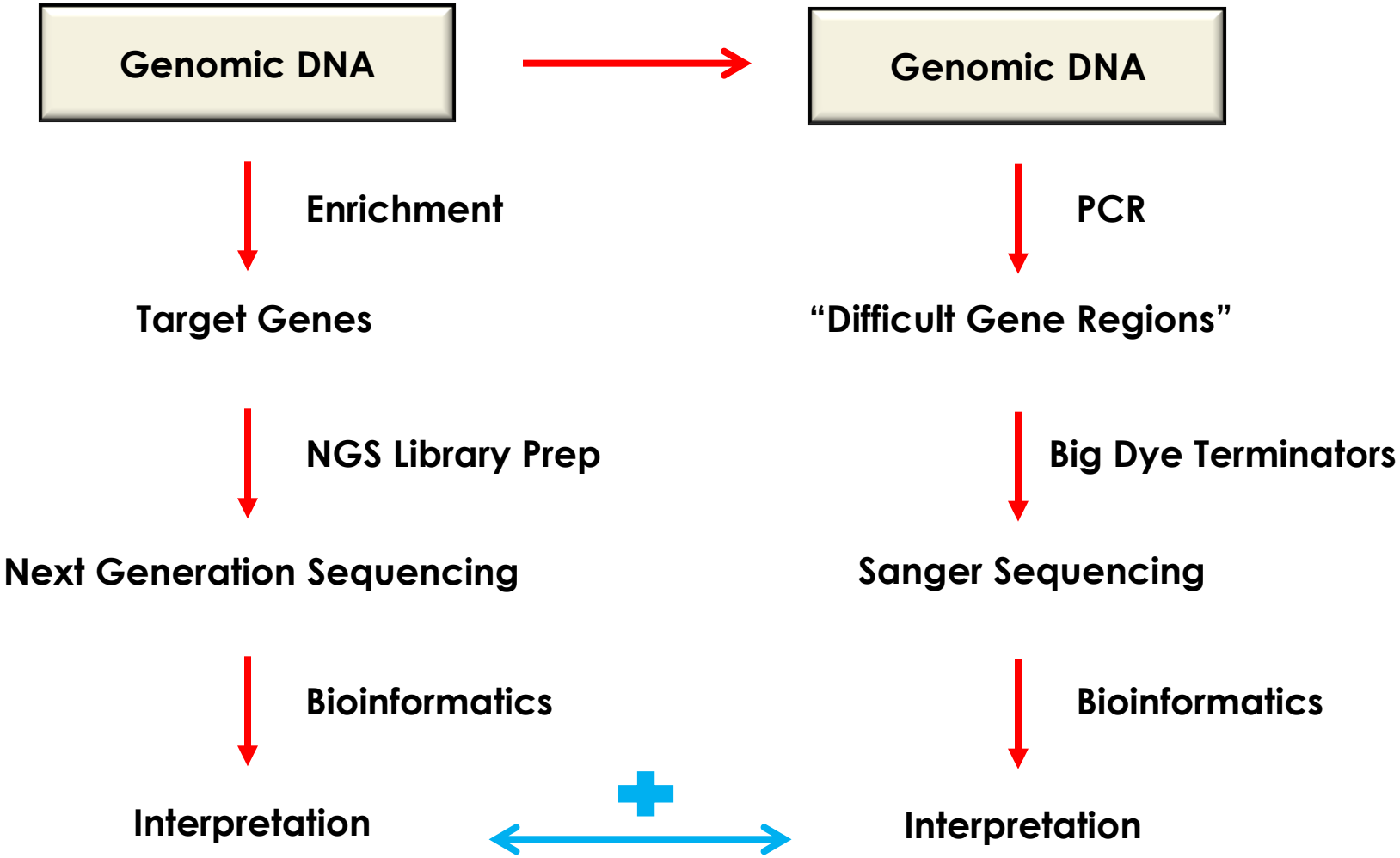
Phosphodiester Bond Formation



Multi-Gene Diagnostics

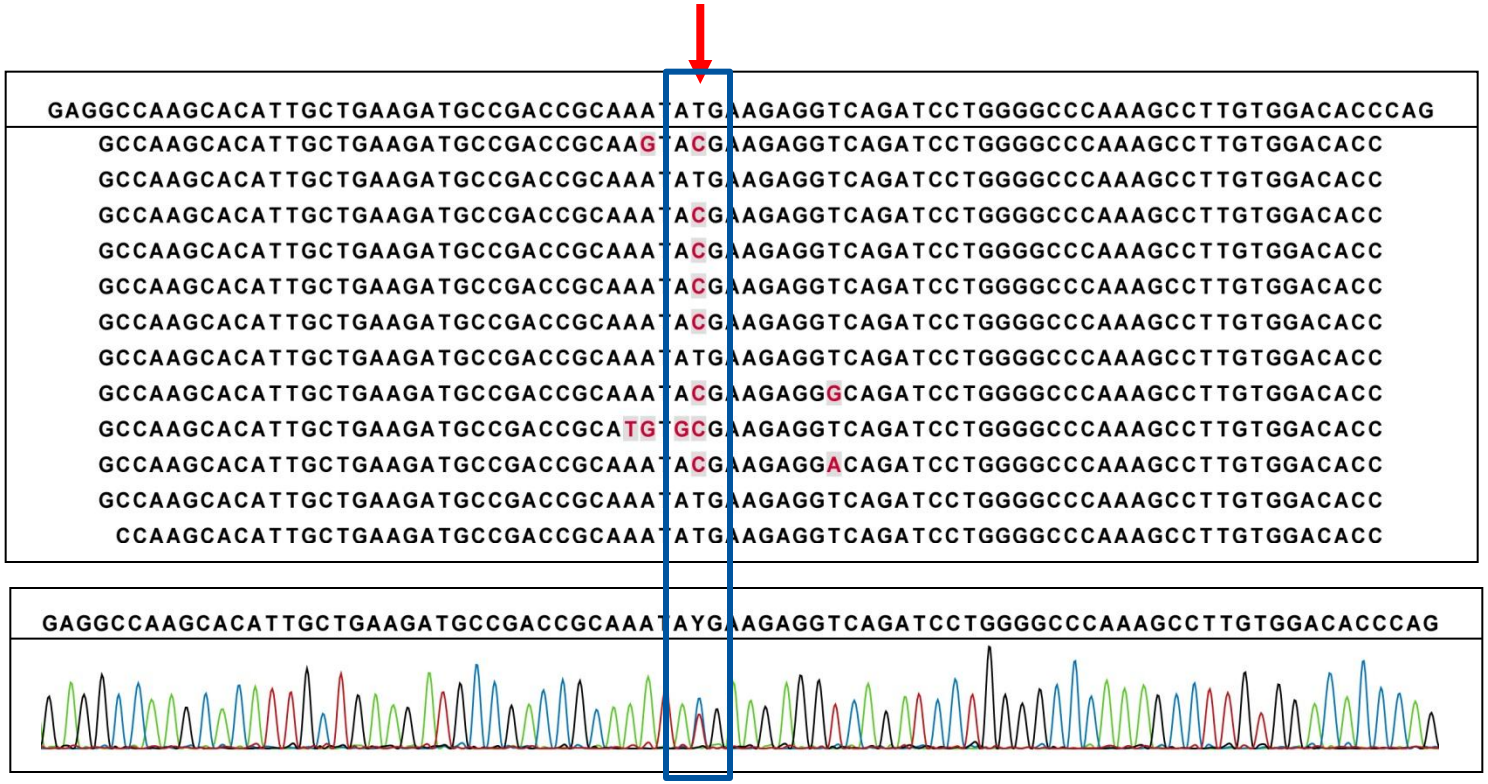


Multi-Gene Diagnostics – Parallel Testing



Re-Sequencing – Cardiomyopathy Genes

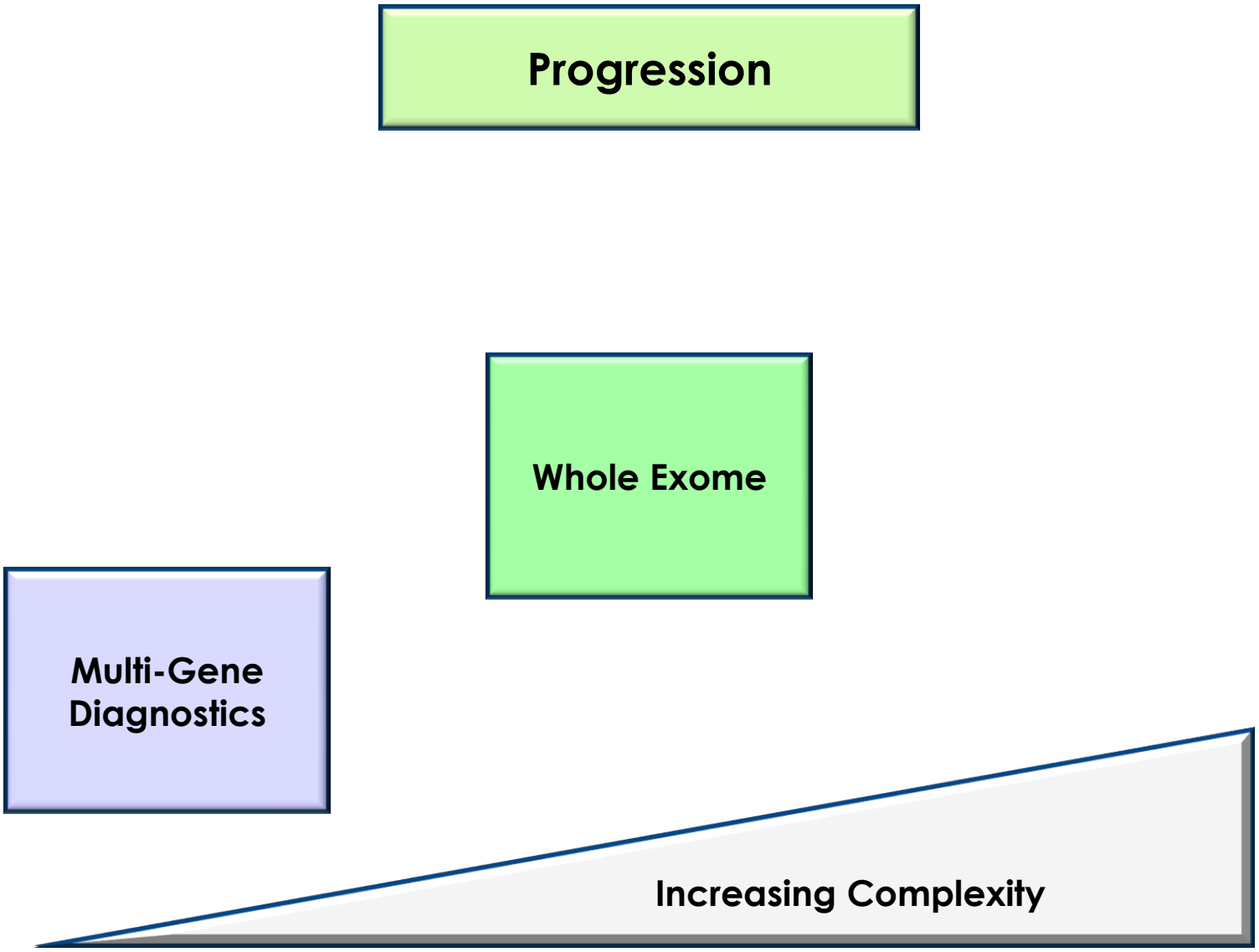
Variant g.34142190T>C in *TPM1*



Reference

LR-PCR 47%

Sanger


Progression
Whole Exome
Multi-Gene Diagnostics
Increasing Complexity

Human Exome

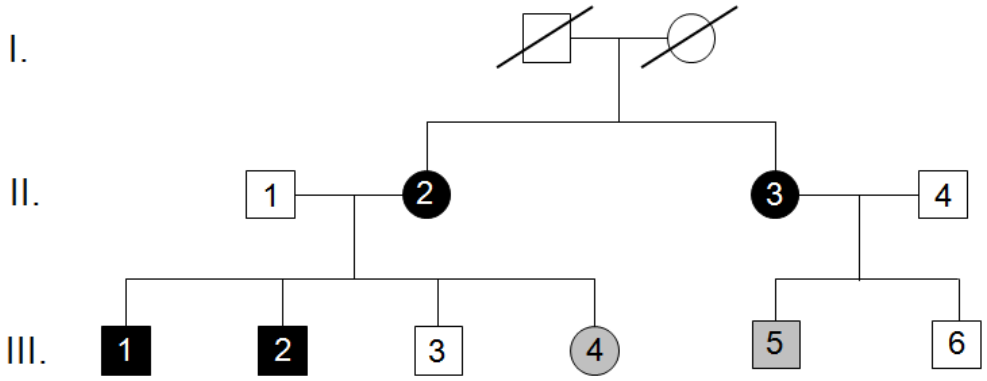
"Journey to the Center of the Genome"

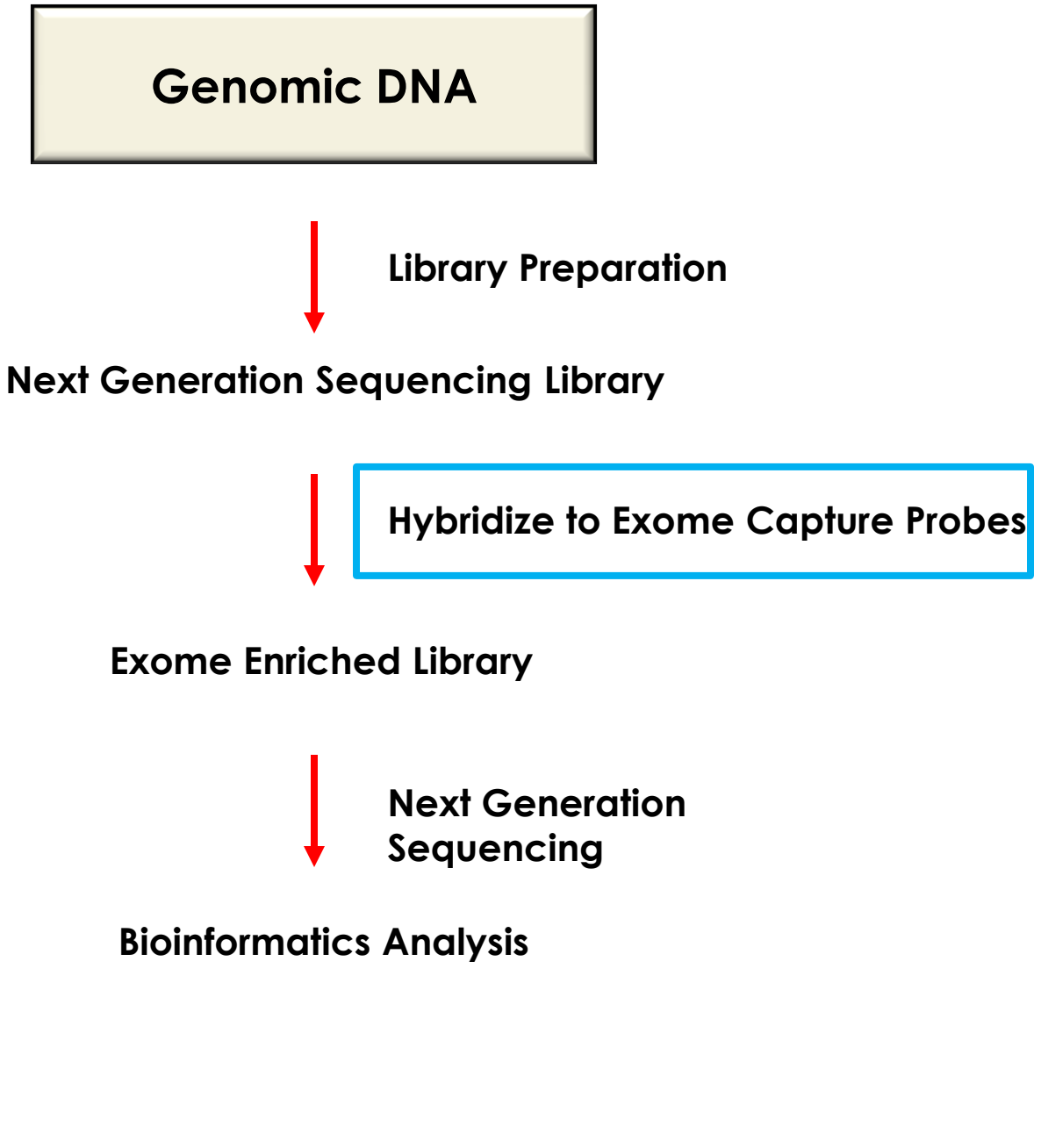
~ 30 Megabases (~ 1% of the genome)

~ 180,000 exons (~ 20,500 genes)

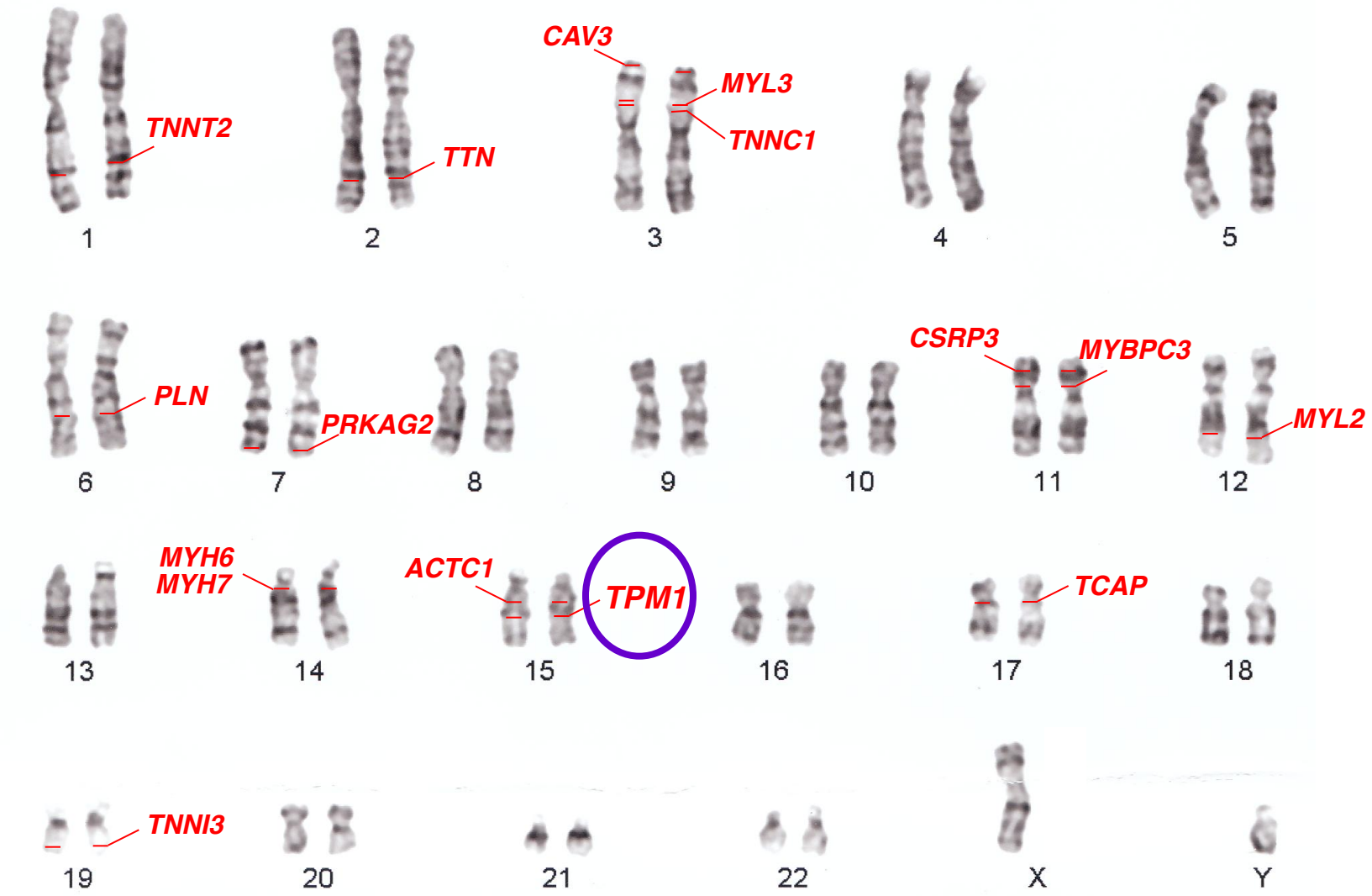
Harbors "Majority" of Mendelian Mutations

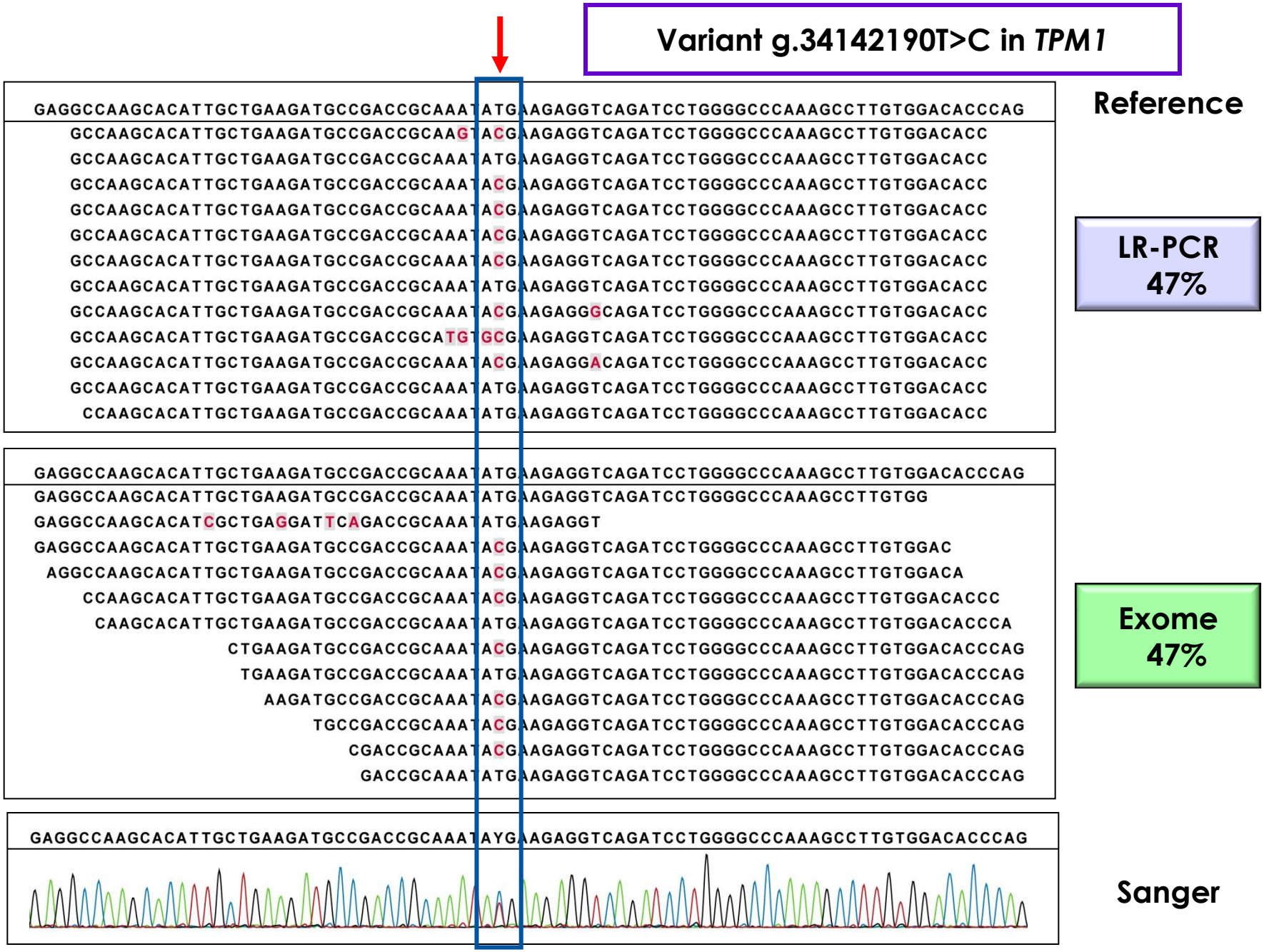
Gene Discovery
~ 40 Publications
July 2011

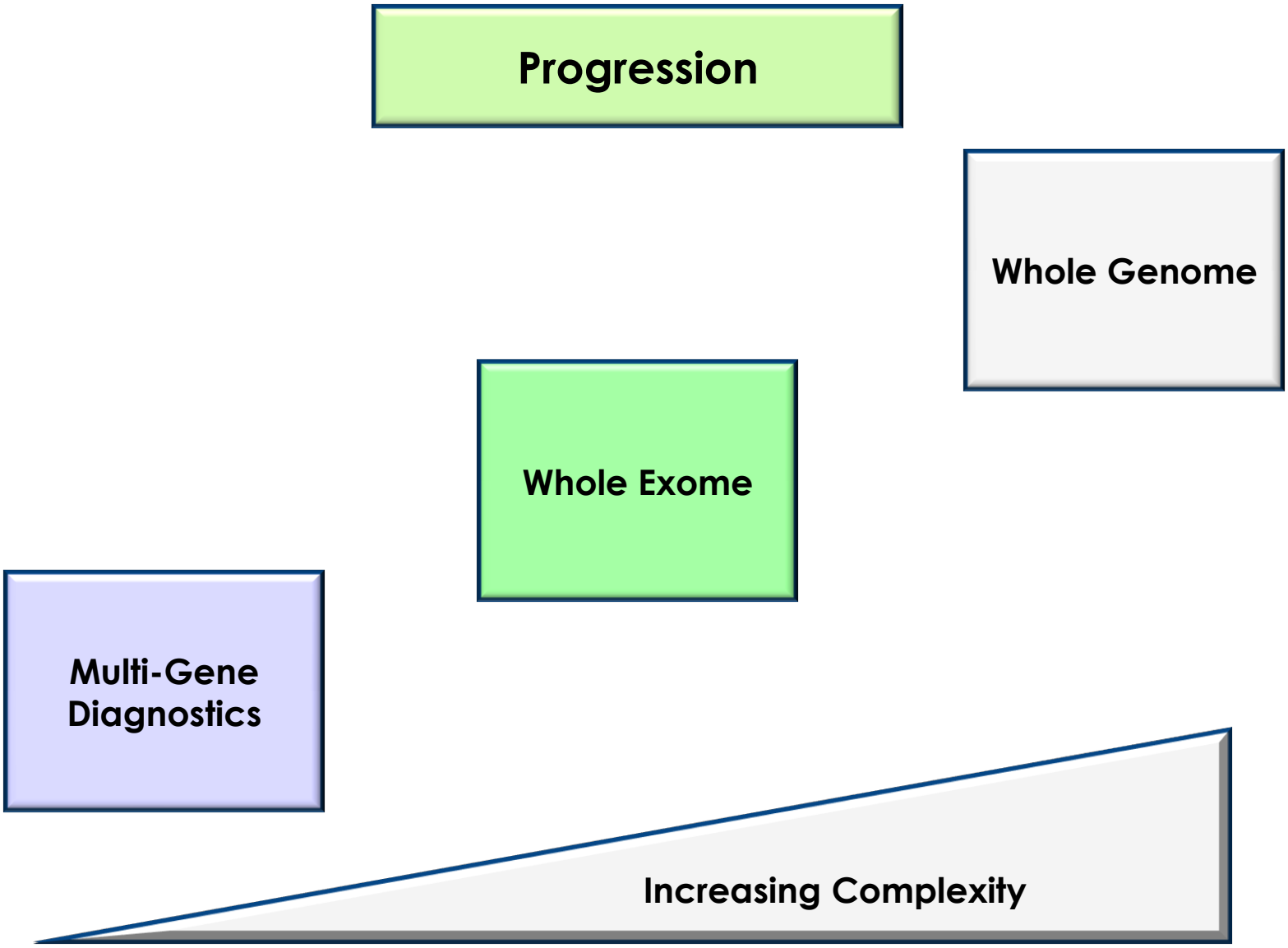



Genomic DNA
Library Preparation
Next Generation Sequencing Library
Hybridize to Exome Capture Probes
Exome Enriched Library
Next Generation Sequencing
Bioinformatics Analysis

Exome Sequencing – Cardiomyopathy Genes




Variant g.34142190T>C in *TPM1*
Reference
LR-PCR 47%
Exome 47%
Sanger


Progression
Whole Genome
Whole Exome
Multi-Gene Diagnostics
Increasing Complexity

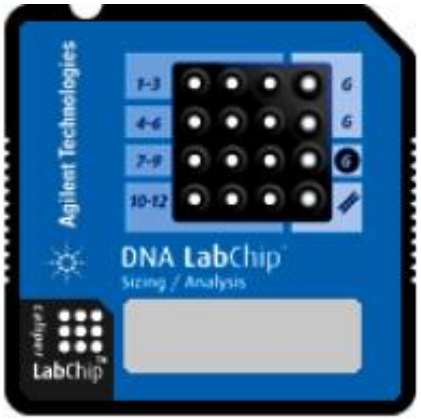
Genomic DNA

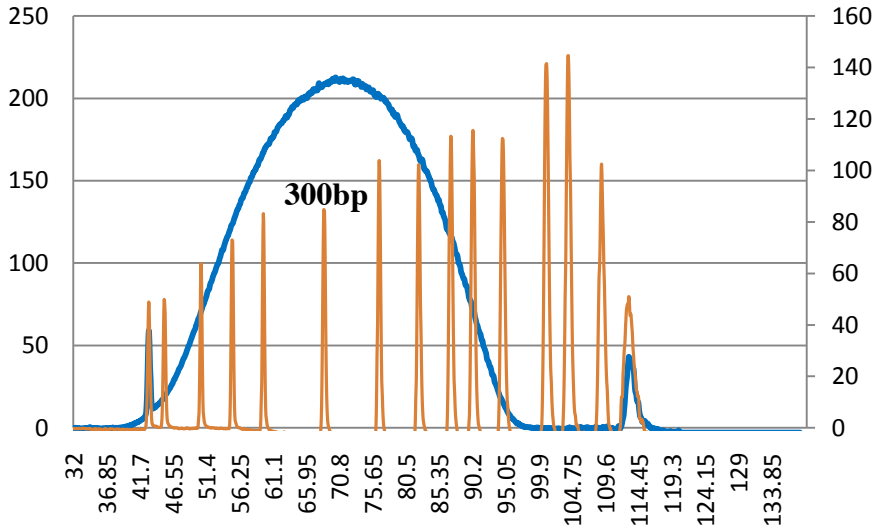
Fragmentation

QC

Process

Covaris
Acoustic Wave





Process

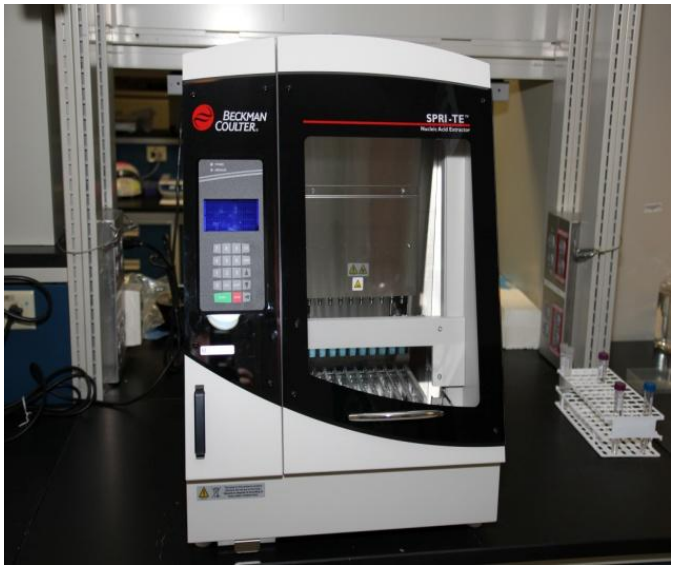
Library Preparation - Illumina

↓

Fragment End Repair

↓

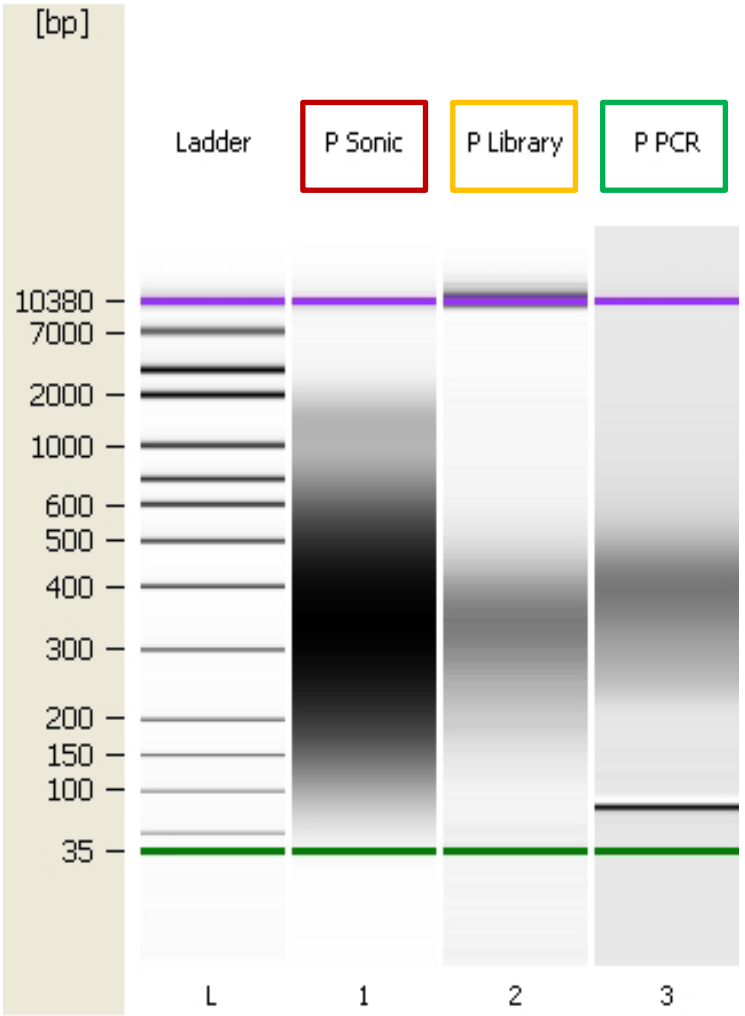
Adapter Ligation



Beckman SPRI-TE

PCR →

QC



BioAnalyzer

Process

PCR Amplified Library

↓

Gel Electrophoresis

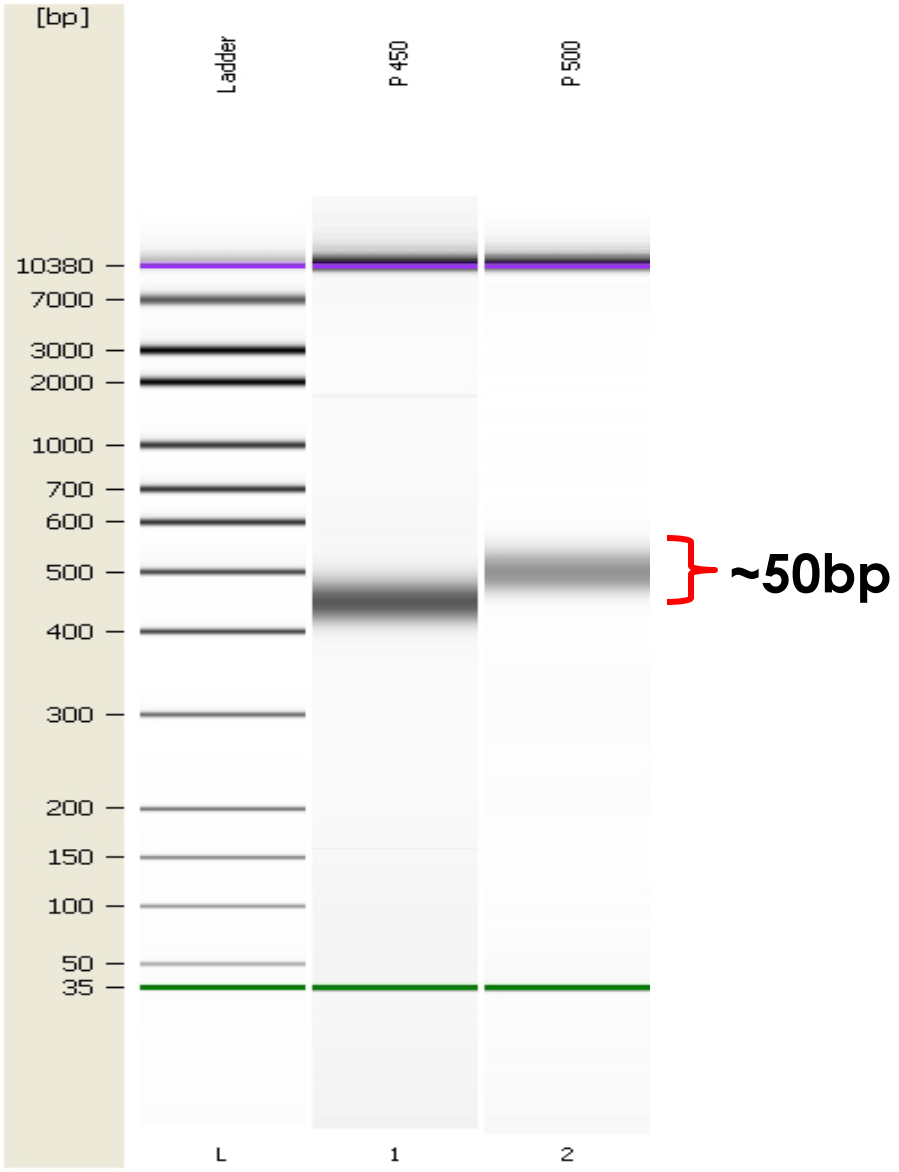
↓

Size Selection

↓

Gel Purification

↳ QC



Process

Gel Purified Library

↓

Quantitative PCR

↓ QC

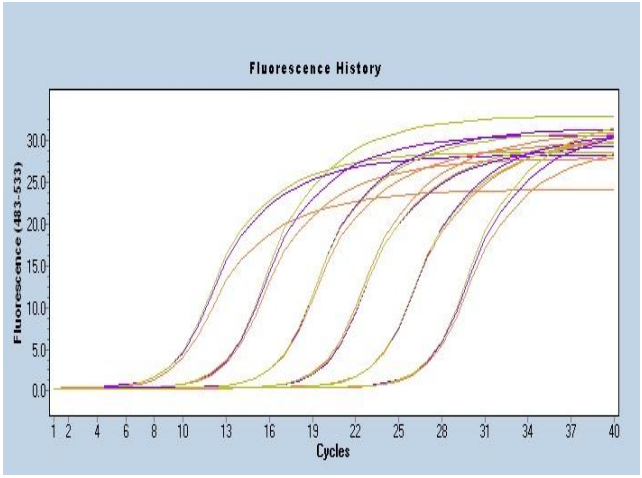
Dilution/Denaturation

↓ 5-7pM

Flow Cell Cluster Generation

↳

Illumina cBot



Process

Sequencing HiSeq 2000

↓

"First Base" Report

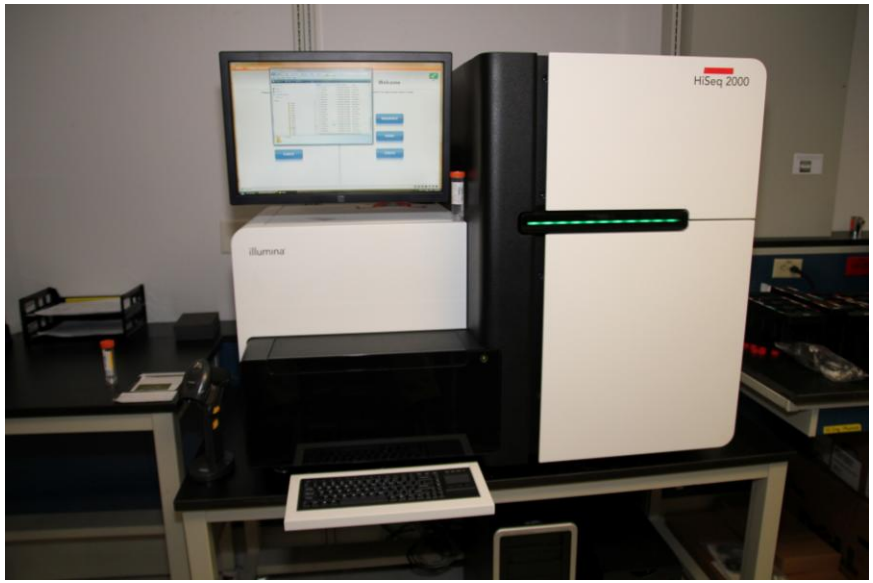
↓ QC

Cluster Densities/Intensities

↓

Real Time Analysis - Cycle 25+

↓ QC



2 X 100bp
Pair End

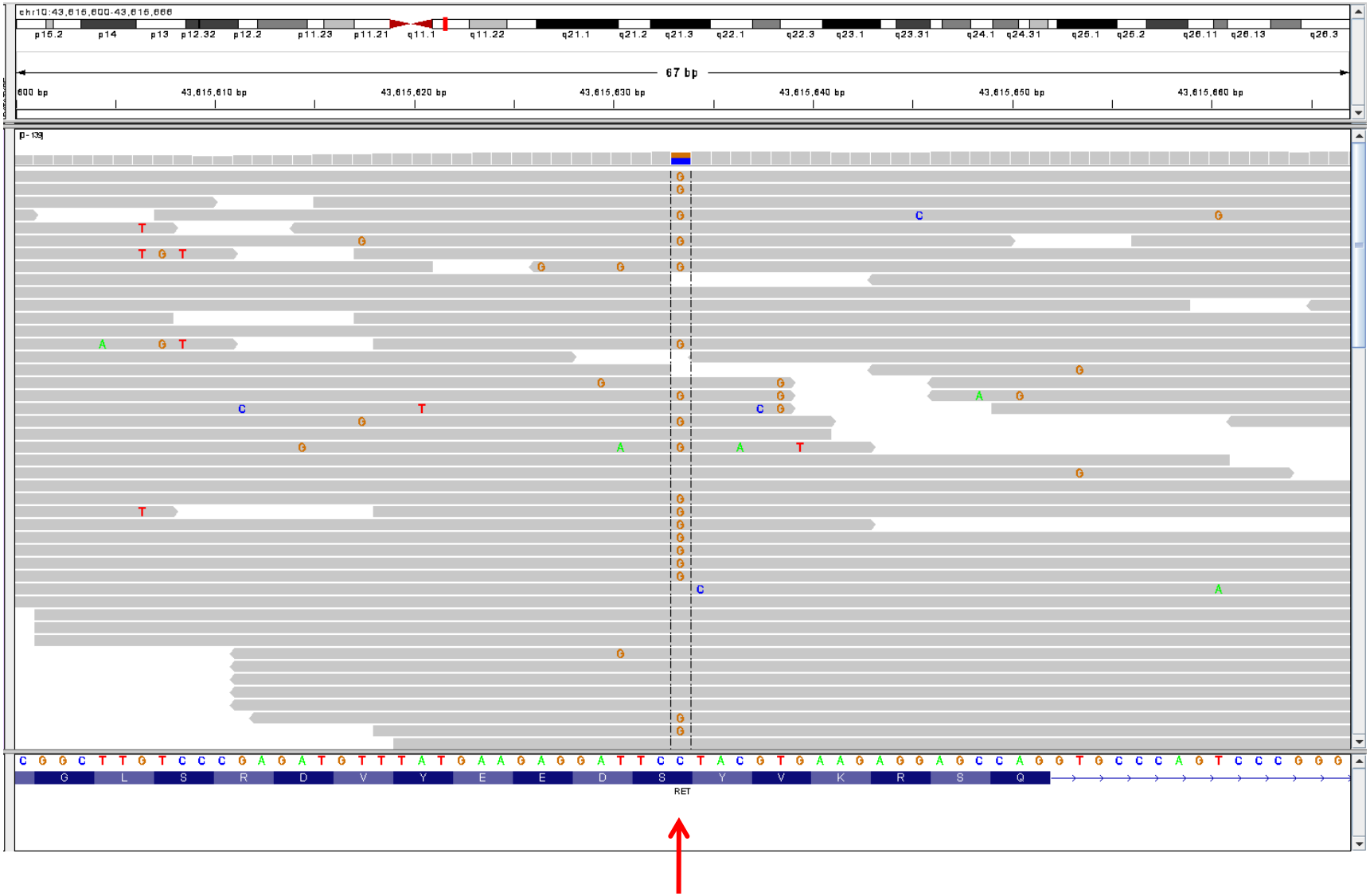
→

Target
~ 100+ Gb

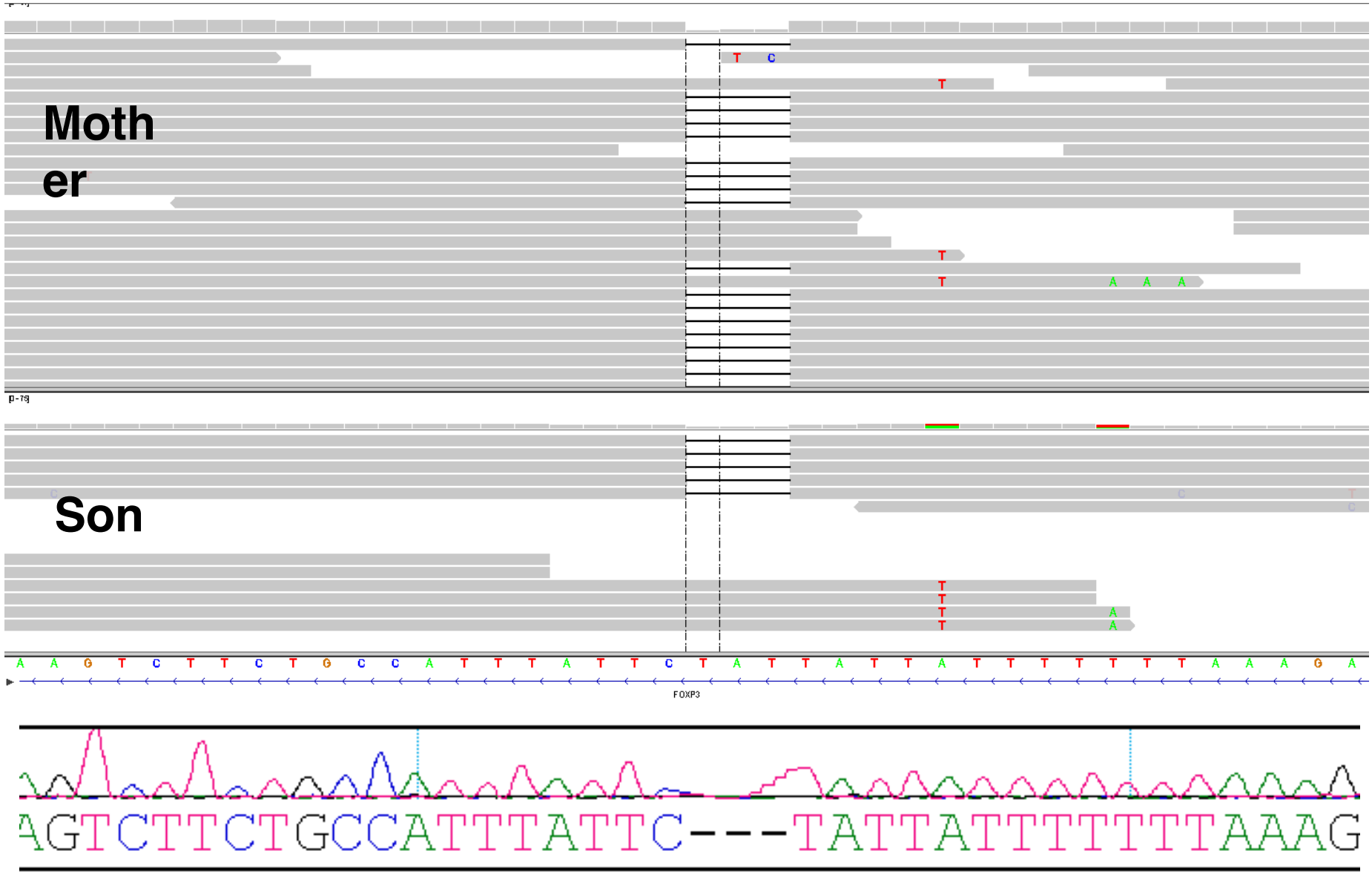
→

FastQ Files

Whole Genome Sequencing
Chr 10: g.43,615,633C>G in *RET*



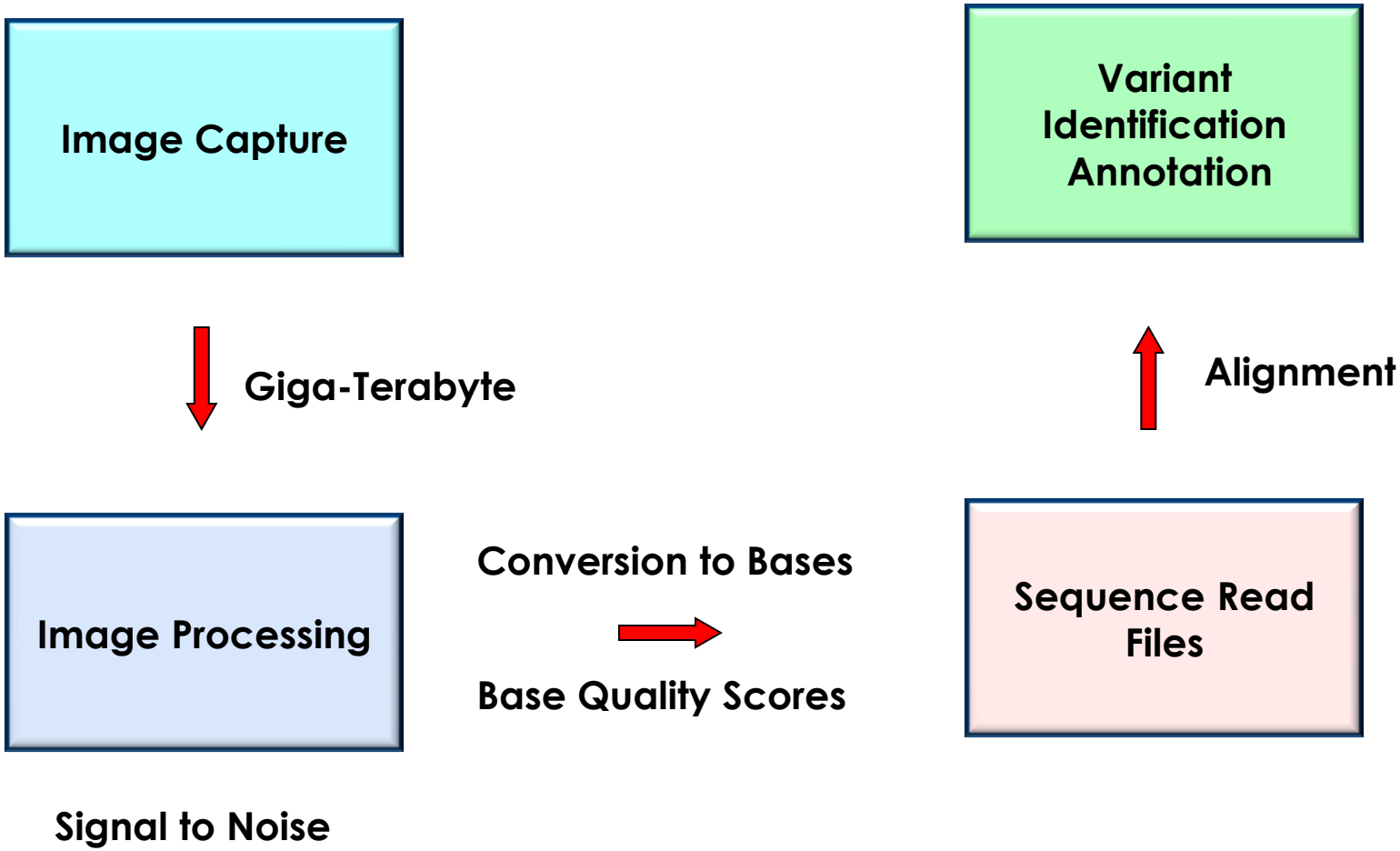
Whole Genome Sequencing
Chr X: 3bp deletion in *FOXP3*



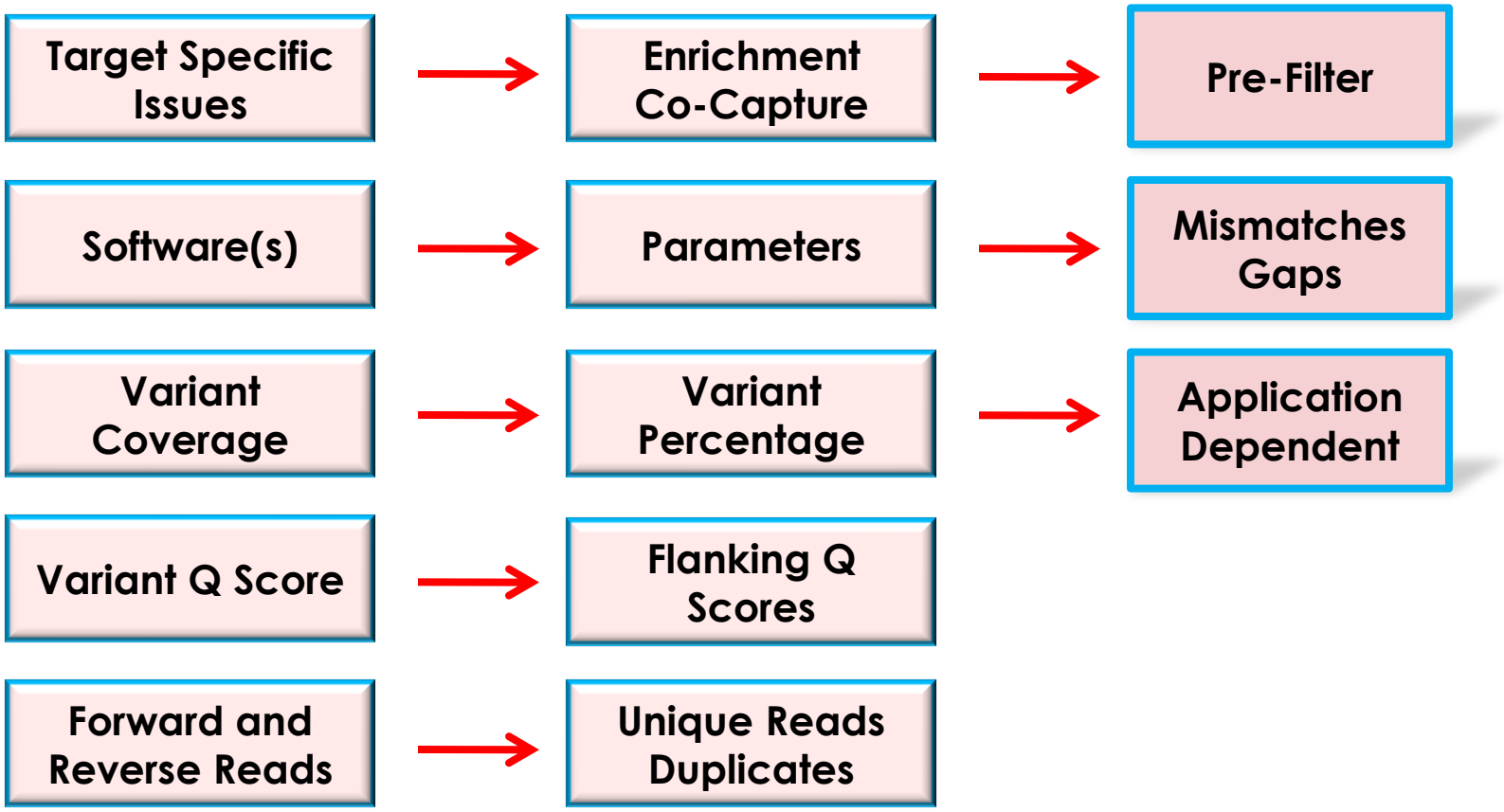
Outline

- Progression: Gene Panels to Genomes
- Next Generation Sequencing Technology
- **Bioinformatics**

Next Generation Sequencing Bioinformatics



Alignment Considerations



Alignment Softwares

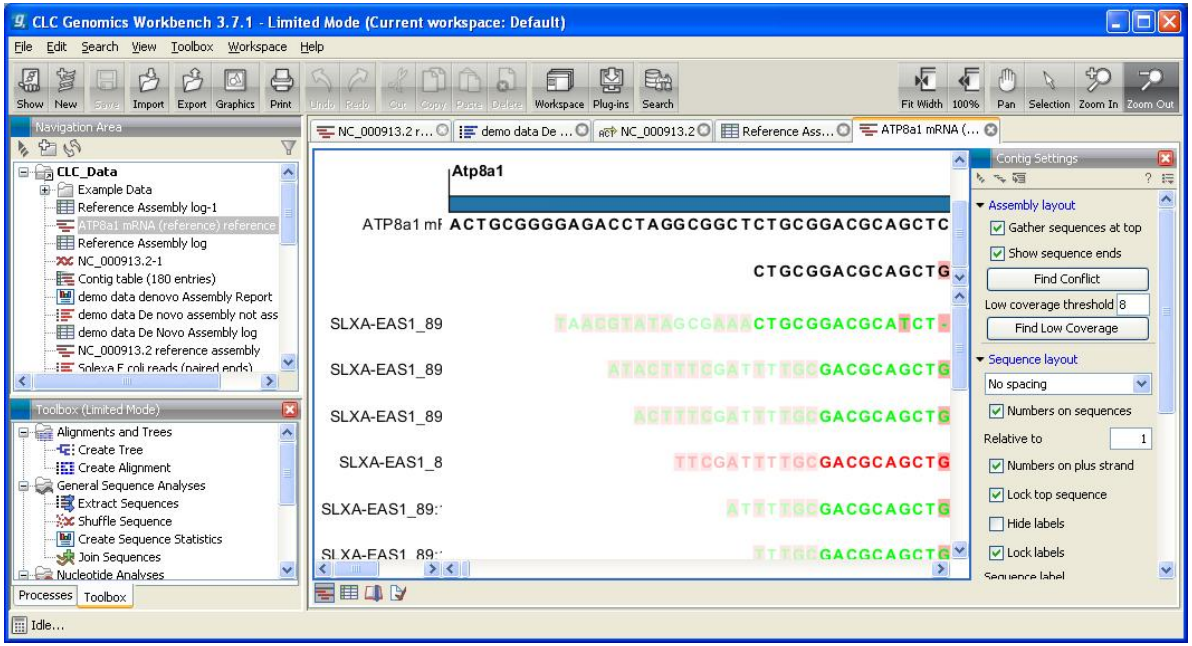
Academic Softwares

- Command Line
- Free
- Genome community support

```
maq assemble [-sp] [-m maxmis] [-Q maxerr] [-r
   hetrate] [-t coef] [-q minQ] [-N nHap]
   out.cns in.ref.bfa in.aln.map 2> out.cns.log
```

Commercial Softwares

- Feature rich user interface
- $$$
- Company support



Alignment Softwares

Commercial Softwares

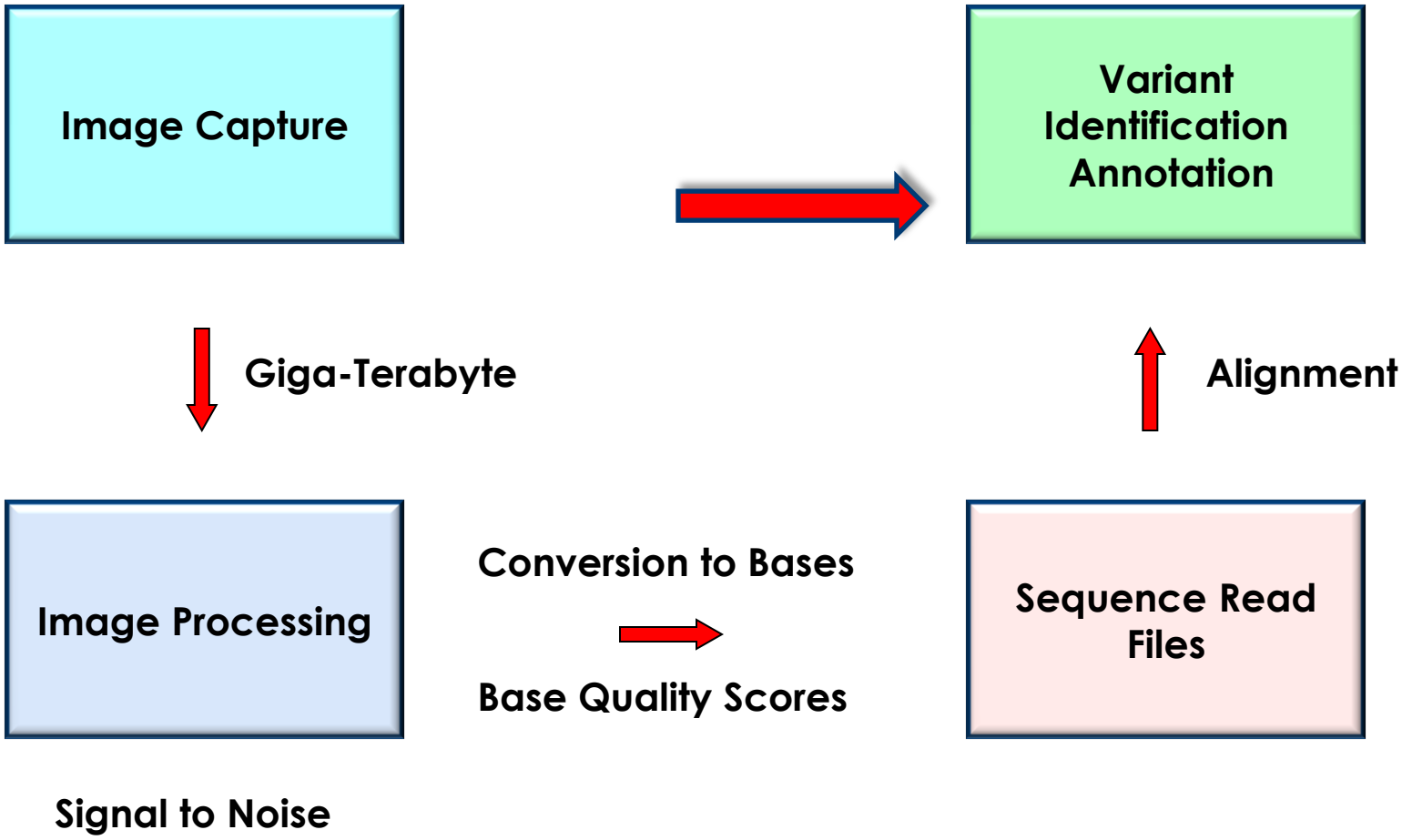
Academic Softwares

BFAST (UCLA)

BWA (Sanger Institute)

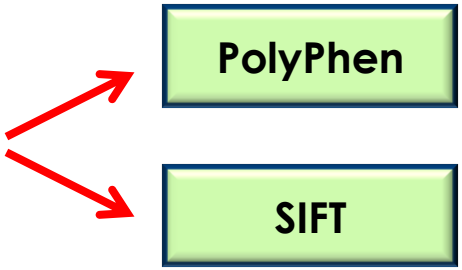
SAMtools (Sanger Institute)

Next Generation Sequencing Bioinformatics



Variant Annotation

- ✓ dbSNP ← 1,000 Genome Project
- ✓ OMIM
- ✓ Human Genome Mutation Database
- ✓ Locus Specific Databases
- ✓ Literature and Internet
- ✓ Functional Prediction Programs



Convergence of Chemistry + Bioinformatics

Pipelines

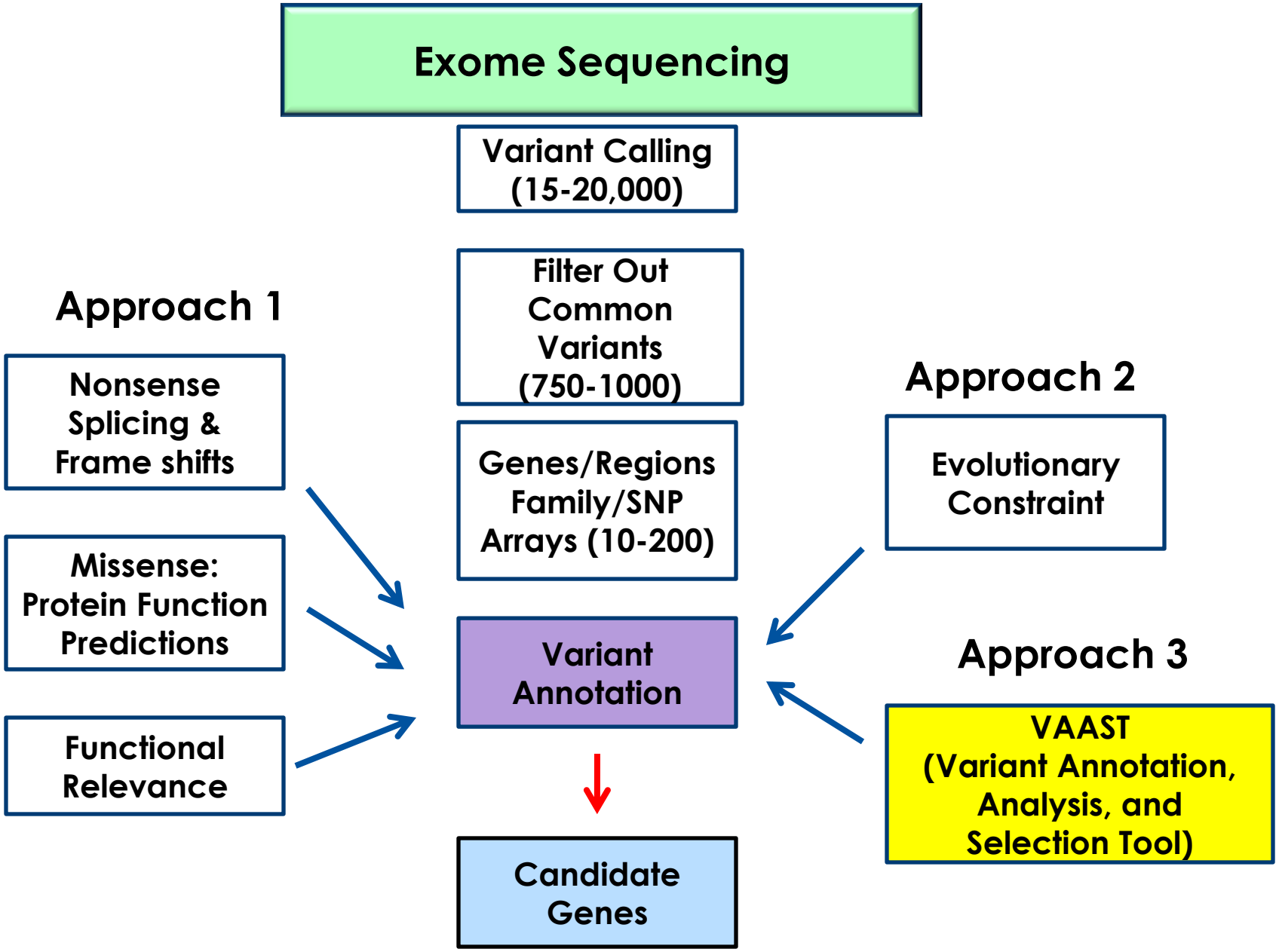
Manually Complex

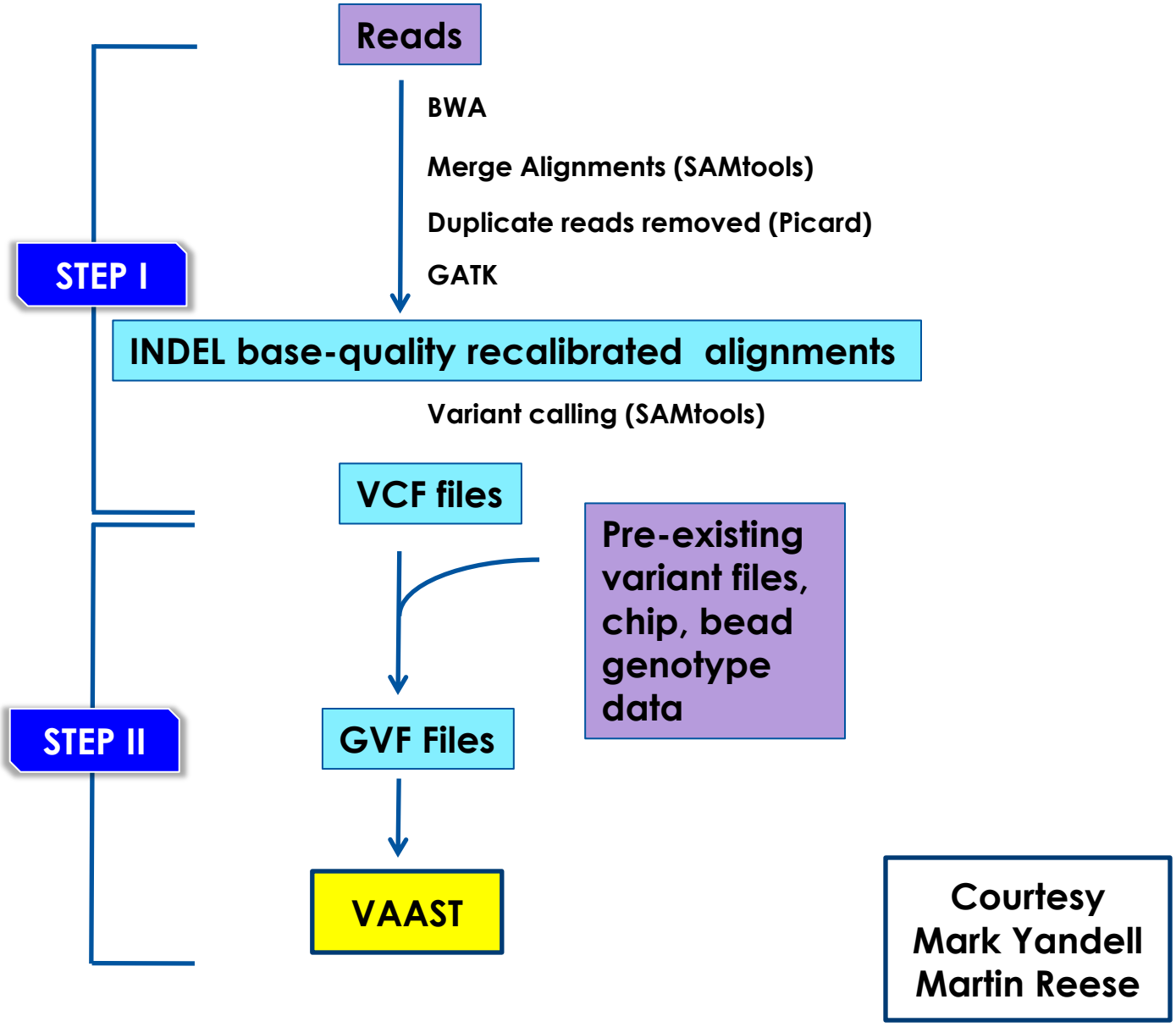
Large Cognitive Component

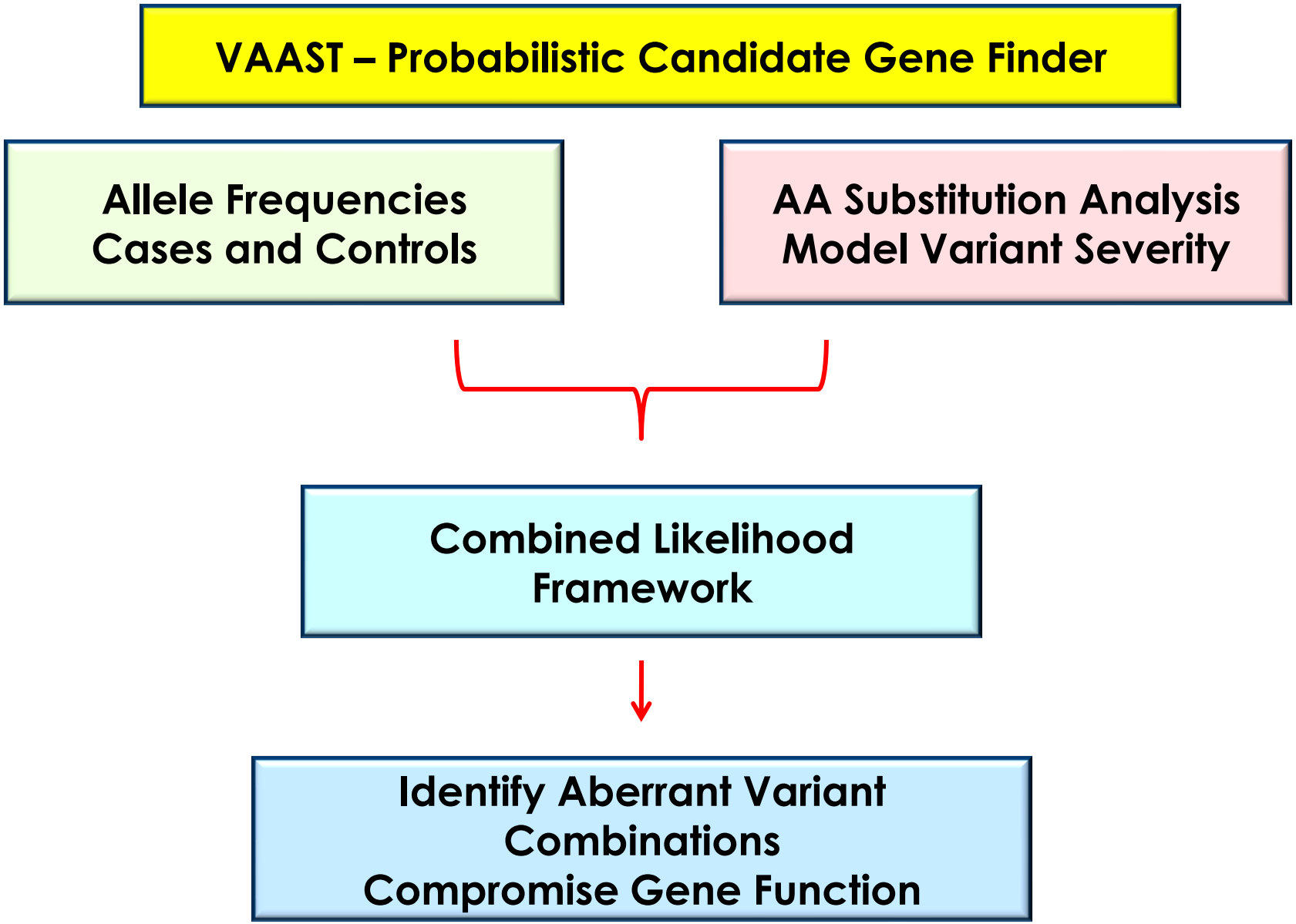
Automation

TCGAAGTCTGCCTAGCTGT

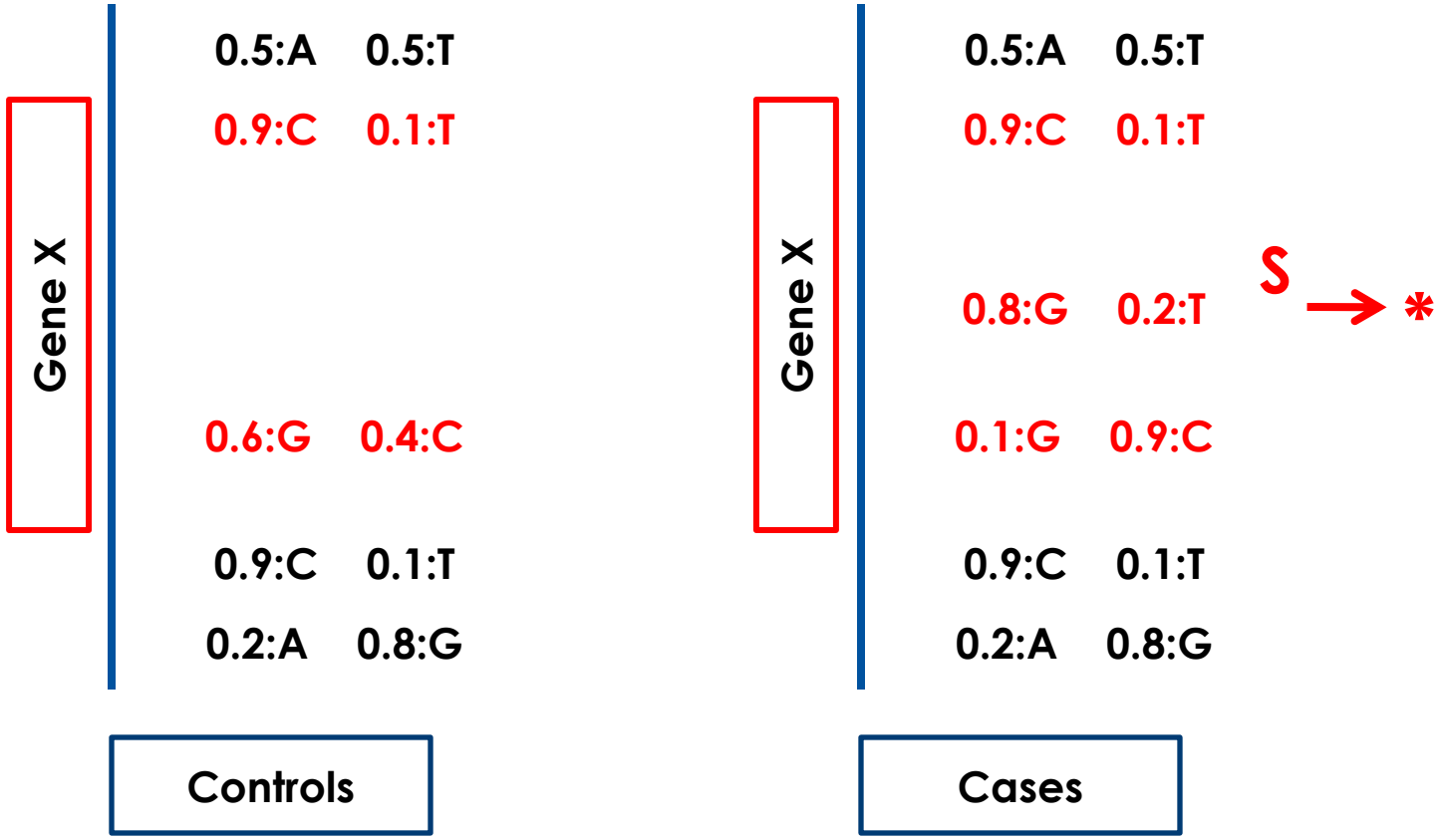
CCGTACGTCTGATGCGTA


Exome Sequencing
Variant Calling (15-20,000)
Filter Out Common Variants (750-1000)
Genes/Regions Family/SNP Arrays (10-200)
Approach 1
Nonsense Splicing & Frame shifts
Missense: Protein Function Predictions
Functional Relevance
Approach 2
Evolutionary Constraint
Approach 3
VAAST (Variant Annotation, Analysis, and Selection Tool)
Variant Annotation
Candidate Genes


Reads
BWA
Merge Alignments (SAMtools)
Duplicate reads removed (Picard)
GATK
STEP I
INDEL base-quality recalibrated alignments
Variant calling (SAMtools)
VCF files
Pre-existing variant files, chip, bead genotype data
STEP II
GVF Files
VAAST
Courtesy
Mark Yandell
Martin Reese

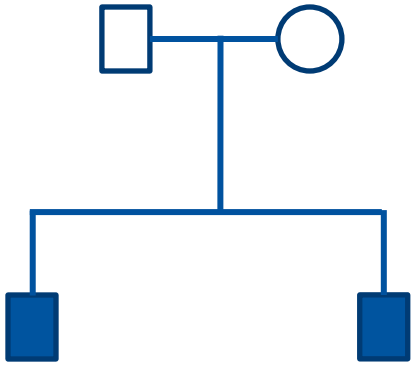

VAAST – Probabilistic Candidate Gene Finder
Allele Frequencies
Cases and Controls
AA Substitution Analysis
Model Variant Severity
Combined Likelihood
Framework
Identify Aberrant Variant
Combinations
Compromise Gene Function

VAAST – Probabilistic Candidate Gene Finder



© 2010 College of American Pathologists. All rights reserved.

VAAST – Miller Kindred Quartet



Miller + PCD

Exome Sequence Data

Two Siblings

= Affected

VAAST

VAAST – Miller Kindred Quartet

Caucasian Only (65 genomes)				
Genome-wide Significant Genes	DHODH		DNAH5	
	Rank	P-Value	Rank	P-Value
17	14	9.93E-07	19	5.79E-05

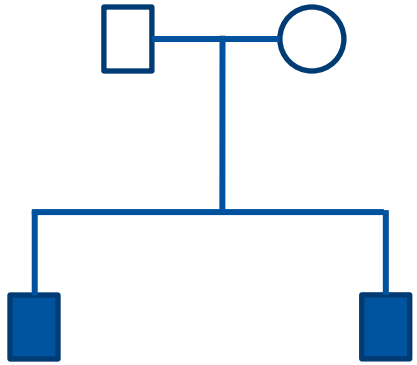
Mixed Ethnicities (189 Genomes)				
Genome-wide Significant Genes	DHODH		DNAH5	
	Rank	P-Value	Rank	P-Value
9	4	7.60E-09	5	1.18E-08

Miller : *DHOD*

PCD: *DNAH5*

GWS Alpha is 2.4 E-6

VAAST – Miller Kindred Quartet



Miller + PCD

Exome Sequence Data

Two Siblings

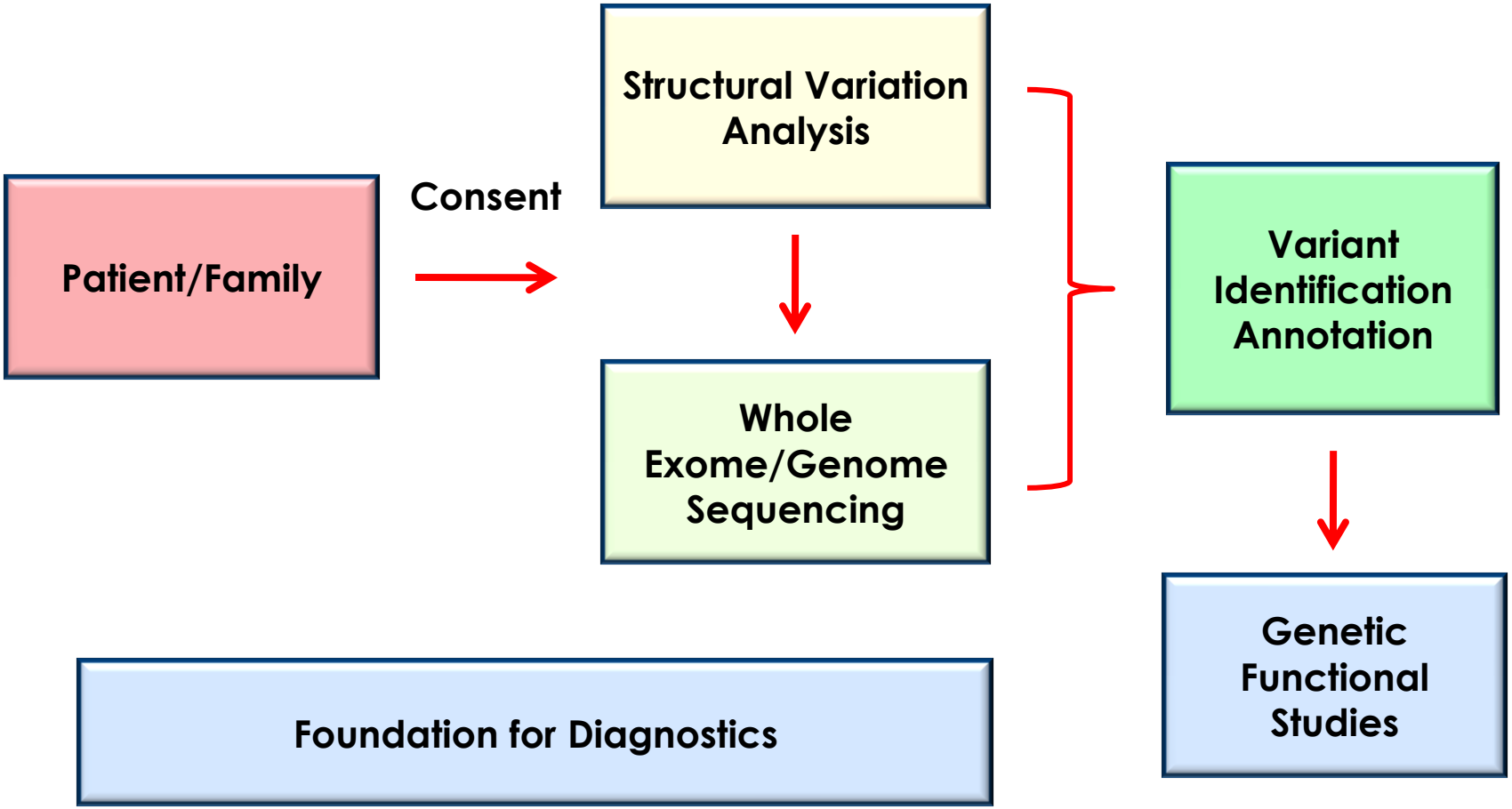
Parents

= Affected

VAAST

VAAST – Miller Kindred Quartet

Genomics Clinical Research Program



© 2010 College of American Pathologists. All rights reserved.

Summary

- **Progression: Gene Panels to Genomes**
- **Next Generation Sequencing Technology**
- **Bioinformatics**

Acknowledgements

Genomics-Bioinformatics
Rebecca Margraf Jacob Durtschi
Emily Coonrod Perry Ridge

U of Utah Genetics
Mark Yandell Lynn Jorde

voelkek@aruplab.com

Omicia
Martin Reese

ARUP Laboratories Institute
for
Clinical and Experimental Pathology

© 2010 College of American Pathologists. All rights reserved.

Next in the Series of Free PHC Webinars

- **How to Have Successful Patient Interactions, Wednesday, August 17th, 11:00-12:00 pm CT**
  - Mary Ann Abrams, MD, MPH & Barbara Savage, MT(ASCP)
- Go to www.cap.org/institute For All Upcoming Webinars!

- Past Webinars Available Now Online at www.cap.org/institute
  - Accountable Care Organizations
  - Whole Genome Analysis as a Universal Diagnostic
  - How to Build and Fund a Financially Viable Molecular Lab
  - Cancer: The Critical Role of Pathology
  - Molecular Markers in Breast Cancer
  - Bethesda System: Integrating Cytology and HPV Molecular Testing
  - Molecular Diagnosis for Lung Cancer Patients
  - Molecular Diagnosis for Colorectal Cancer Patients

© 2010 College of American Pathologists. All rights reserved.

CAP Events of Interest

- Don't Forget to Register for CAP'11 – THE Pathologists' Meeting – September 11 – 14, 2011 held at the Gaylord Texan in Grapevine, Texas!

  – Go to www.cap.org/CAP11 or call 1-800-967-4548. International attendees please call 1-847-996-5891.

For more information go to www.cap.org/CAP11

Tuesday, Sept 12th:

TP120 Breakfast Workshop – Hot Topics in Pathology: What Every Community Pathologist Should Know About Clinical Requests for Molecular Tests (6:30-7:45 am)
Faculty--Samuel K. Caughron, MD, FCAP
Frederick L. Kiechle, MD, PhD, FCAP
Michael S. Brown, MD, FCAP

ST109 Companion Diagnostics for Targeted Therapy in Cancer (2:00-5:30 pm)
Faculty--Sanja Dacic, MD, PhD, FCAP
David Hicks, MD, FCAP
Jeffrey Kant, MD, PhD, FCAP

Wednesday, Sept 13th:

ST110 Direct-to-Consumer Genetic Testing: Staying Ahead of Patients in This Current Trend (8:00-9:00 am)
Faculty--Nazneen Aziz, PhD
Elizabeth A. Mansfield, PhD

ST111 What's in It for Me? Using Technology to Become a Diagnostic Hero (8:00-11:30 am)
Faculty--Kenneth J. Bloom, MD, FCAP
John W. Turner, MD, FCAP

© 2010 College of American Pathologists. All rights reserved.