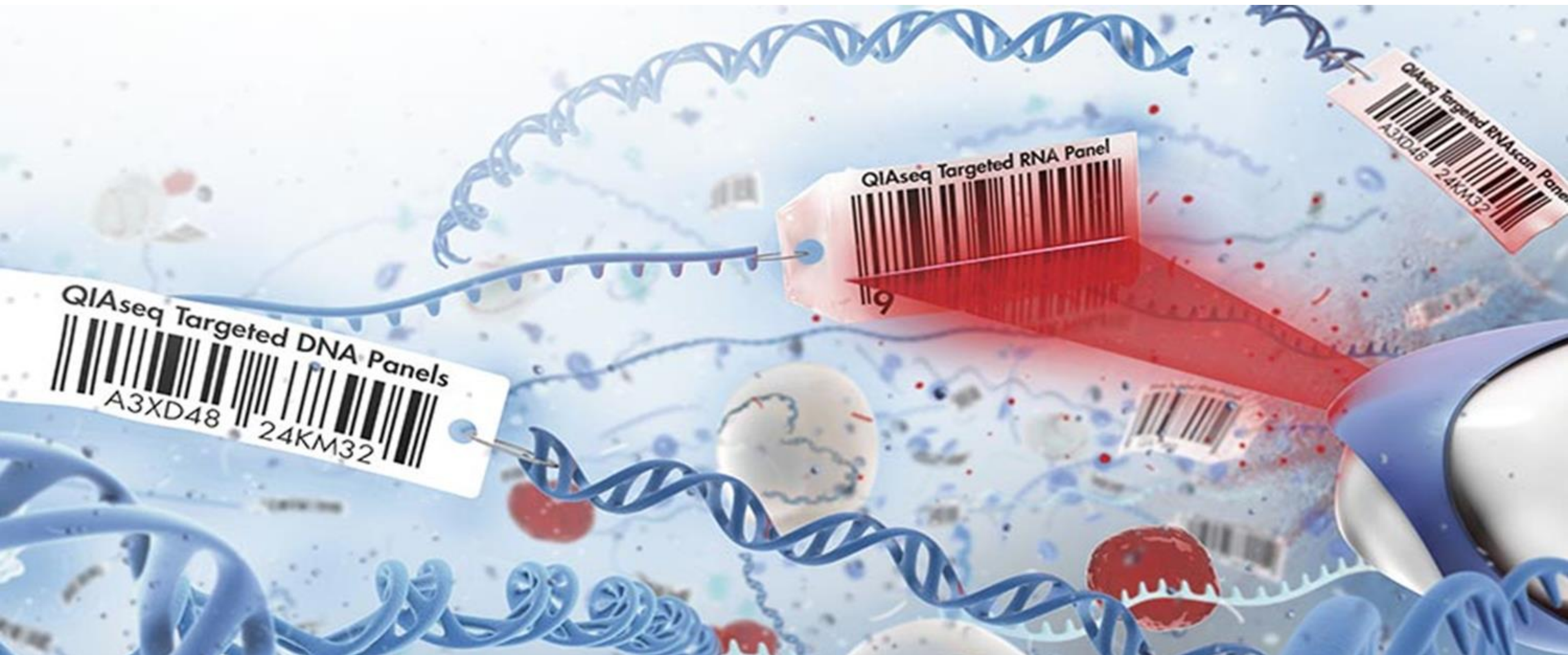
QIAGEN

Introduction to unique molecular indices technology



Samuel Rulli, Ph.D

Global Product Manager, RNA-seq Technology & Applications

Legal disclaimer

QIAGEN products shown here are intended for molecular biology applications. These products are not intended for the diagnosis, prevention or treatment of a disease.

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. QIAGEN kit handbooks and user manuals are available at **www.qiagen.com** or can be requested from QIAGEN Technical Services or your local distributor.

Agenda

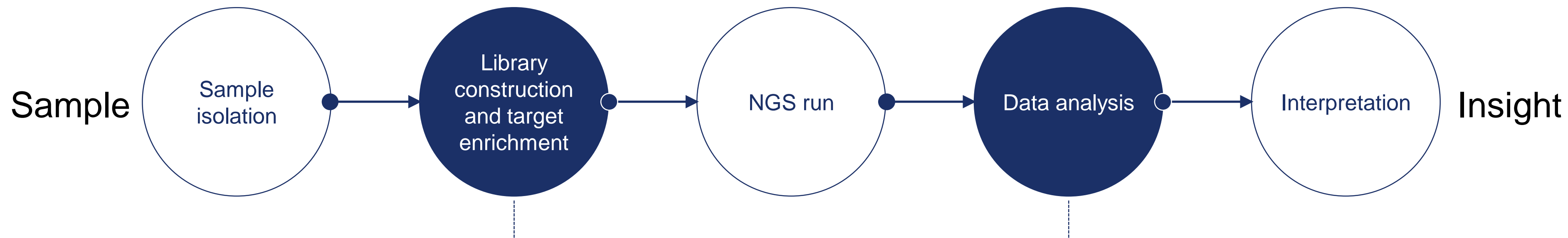
1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. Single primer extension (SPE) versus PCR for capturing information during library construction
4. UMI and SPE in action – gene expression analysis
5. miRNA quantification and biofluids
6. DNA variant analysis and novel gene fusions discovery
7. Summary and questions

Agenda

1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. Single primer extension (SPE) versus PCR for capturing information during library construction
4. UMI and SPE in action – gene expression analysis
5. miRNA quantification and biofluids
6. DNA variant analysis and novel gene fusions discovery
7. Summary and questions

QIAseq targeted NGS – a complete Sample to Insight solution

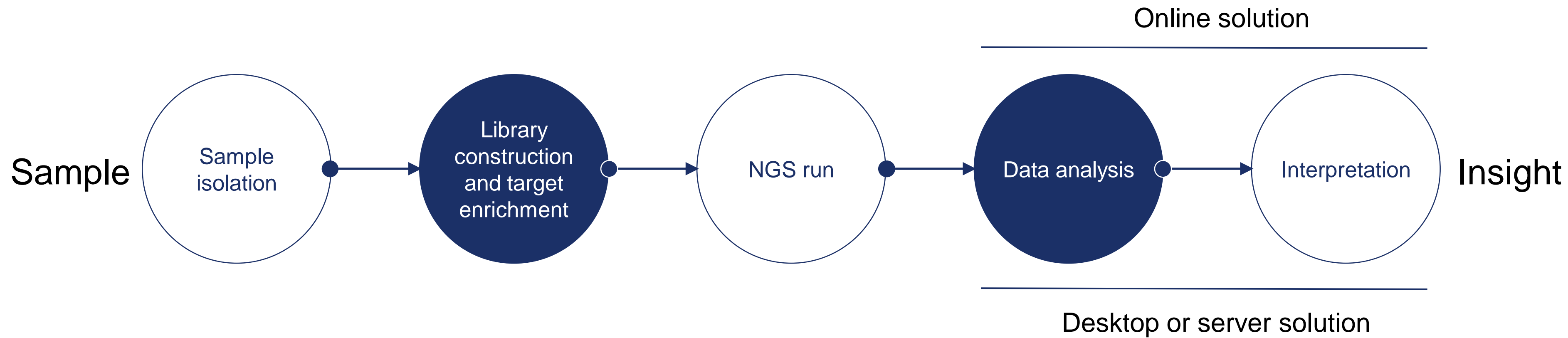
One workflow, several outputs



Workflow	Output
DNA panels for variant analysis	Mutations, indels, copy number variants
RNA panels for differential gene expression	Gene expression levels
RNA panels for fusion profiling	Fusions
miRNome system for miRNA expression	miRNA levels

QIAseq targeted NGS – a complete Sample to Insight solution

One workflow, several outputs



Online and desktop/server solutions are not only optimized for UMIs but also for each chemistry and application.

What is targeted sequencing?

Targeted sequencing involves PCR or amplicon-based capture of molecular information

- Multiplex-based PCR is often used for library construction
- Multiplex PCR has limitations which can cause sequencing problems
  - PCR bias
  - Sequencing artifacts (PCR errors)
  - Uniformity issues due to PCR
- Targeted sequencing should be quantitative, accurate and uniform

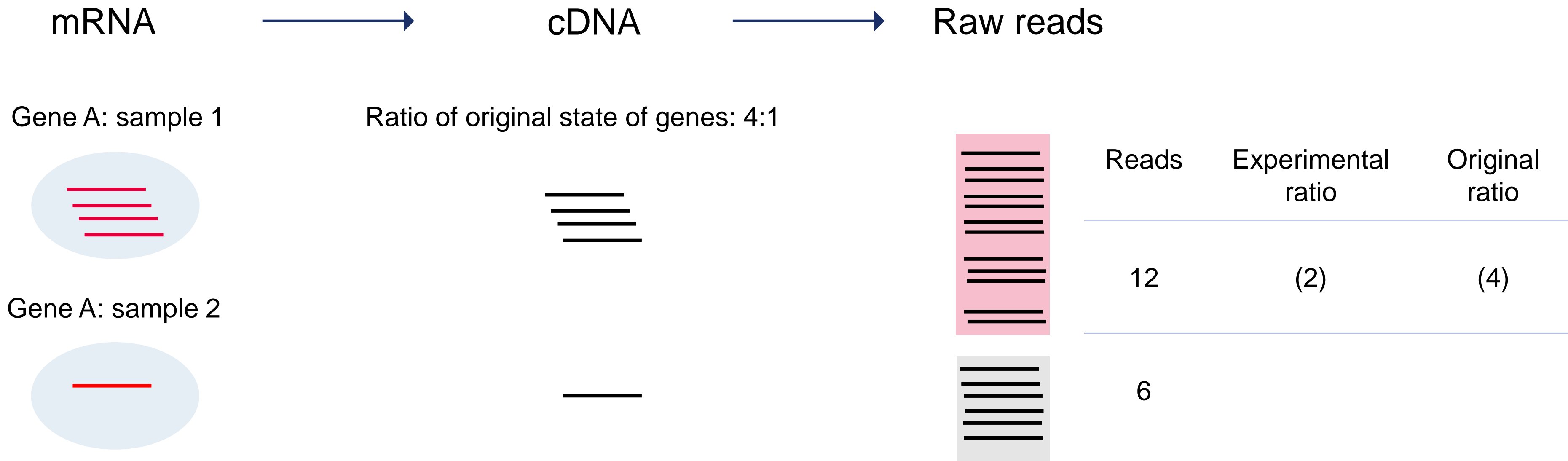
We need a way to monitor each capture event through sequencing and ensure best practices in library construction.

Agenda

1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. Single primer extension (SPE) versus PCR for capturing information during library construction
4. UMI and SPE in action – gene expression analysis
5. miRNA quantification and biofluids
6. DNA variant analysis and novel gene fusions discovery
7. Summary and questions

Principles of unique molecular indices (UMIs)

PCR duplicates, amplification bias and optical duplications introduce errors in quantification



Principles of unique molecular indices (UMIs)

Targeted RNA-seq is a "read-based" approach to understanding gene expression

How do we go from "reads" to counting transcripts?

Five replicates of one transcript

All five reads have the same UMI

Counted as one

Five unique transcripts of a gene

Five sequencing reads have five different UMIs

Counted as five

Principles of unique molecular indices (UMIs)

Targeted RNA-seq is a “read-based” approach to understanding gene expression

How do we go from “reads” to counting transcripts?

Five replicates of one transcript

All five reads have the same UMI

Counted as one

Five unique transcripts of a gene

Five sequencing reads have five different UMIs

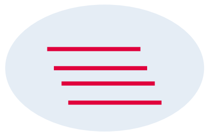
Counted as five

Principles of unique molecular indices (UMIs)

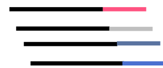
Molecular indices enable accurate quantification

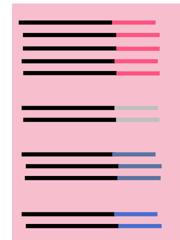
mRNA → cDNA → UMI reads

Gene A: sample 1

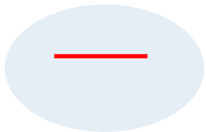


Ratio of original state of genes: 4:1





Gene A: sample 2



Reads	Experimental ratio	Original ratio	UMI	UMI ratio
12	(2)	(4)	4	(4)
6			1	

Tag each gene with UMIs; count unique molecules, not reads.

Principles of unique molecular indices (UMIs)

Each capture event is archived with a UMI

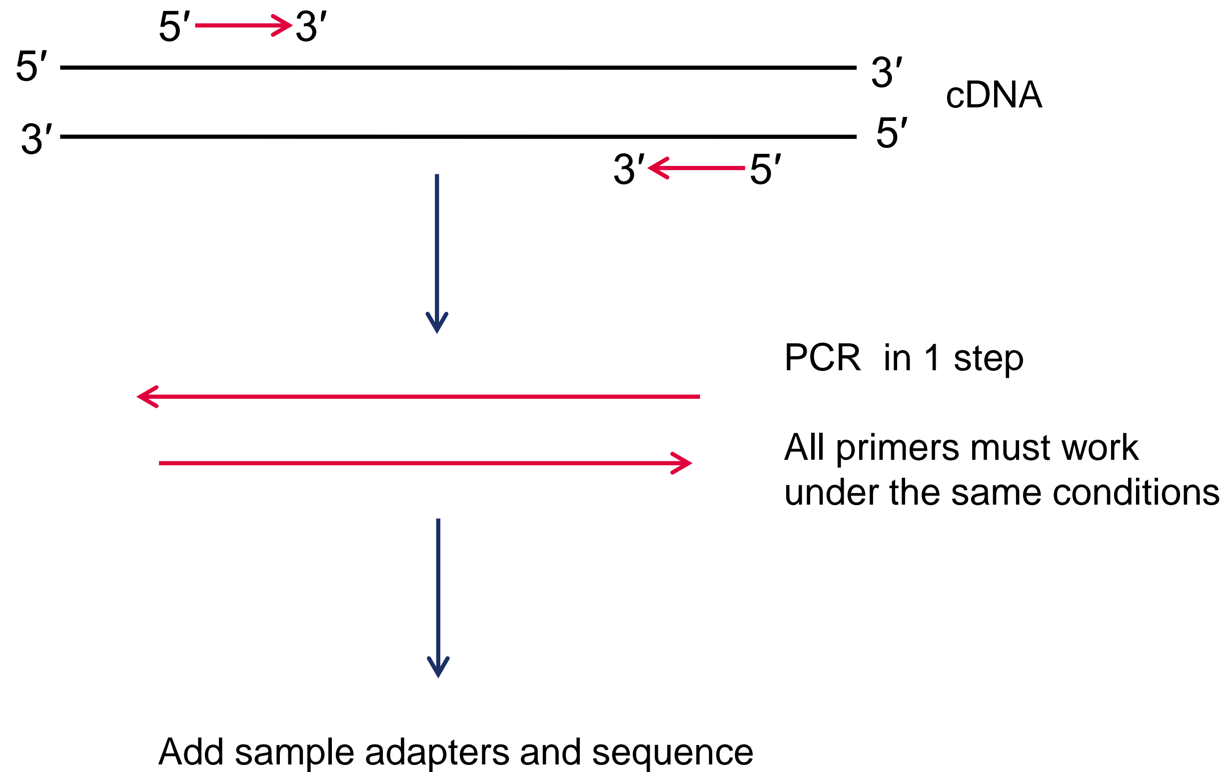
QIAGEN uses 12 random base UMIs.

Agenda

1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. **Single primer extension (SPE) versus PCR for capturing information during library construction**
4. UMI and SPE in action – gene expression analysis
5. miRNA quantification and biofluids
6. DNA variant analysis and novel gene fusions discovery
7. Summary and questions

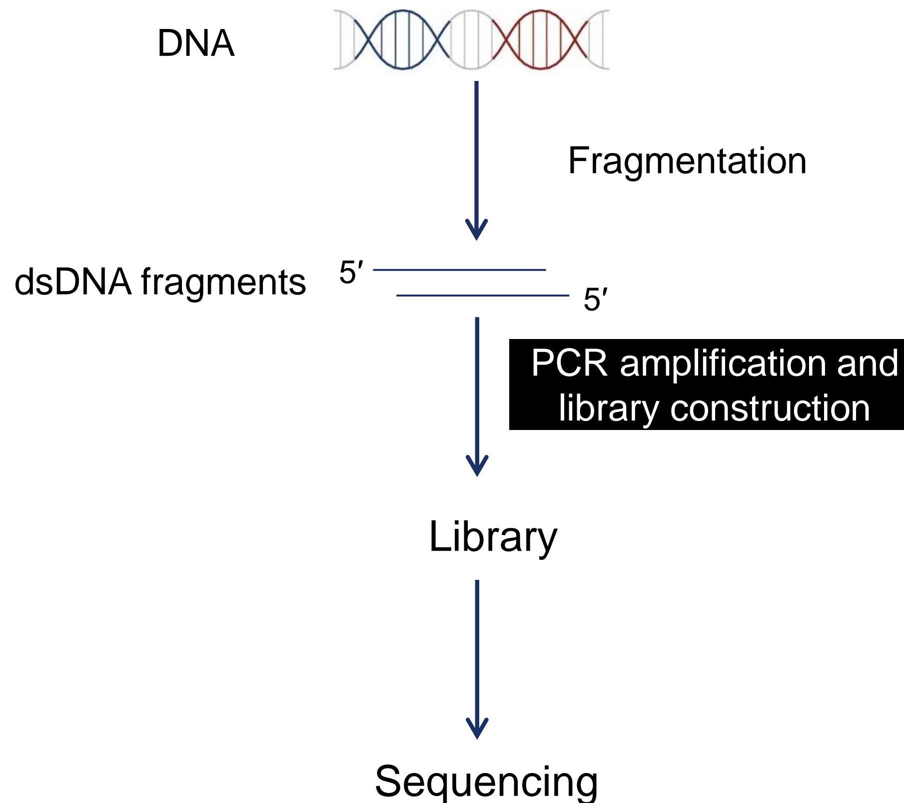
Single primer extension

PCR-based capture



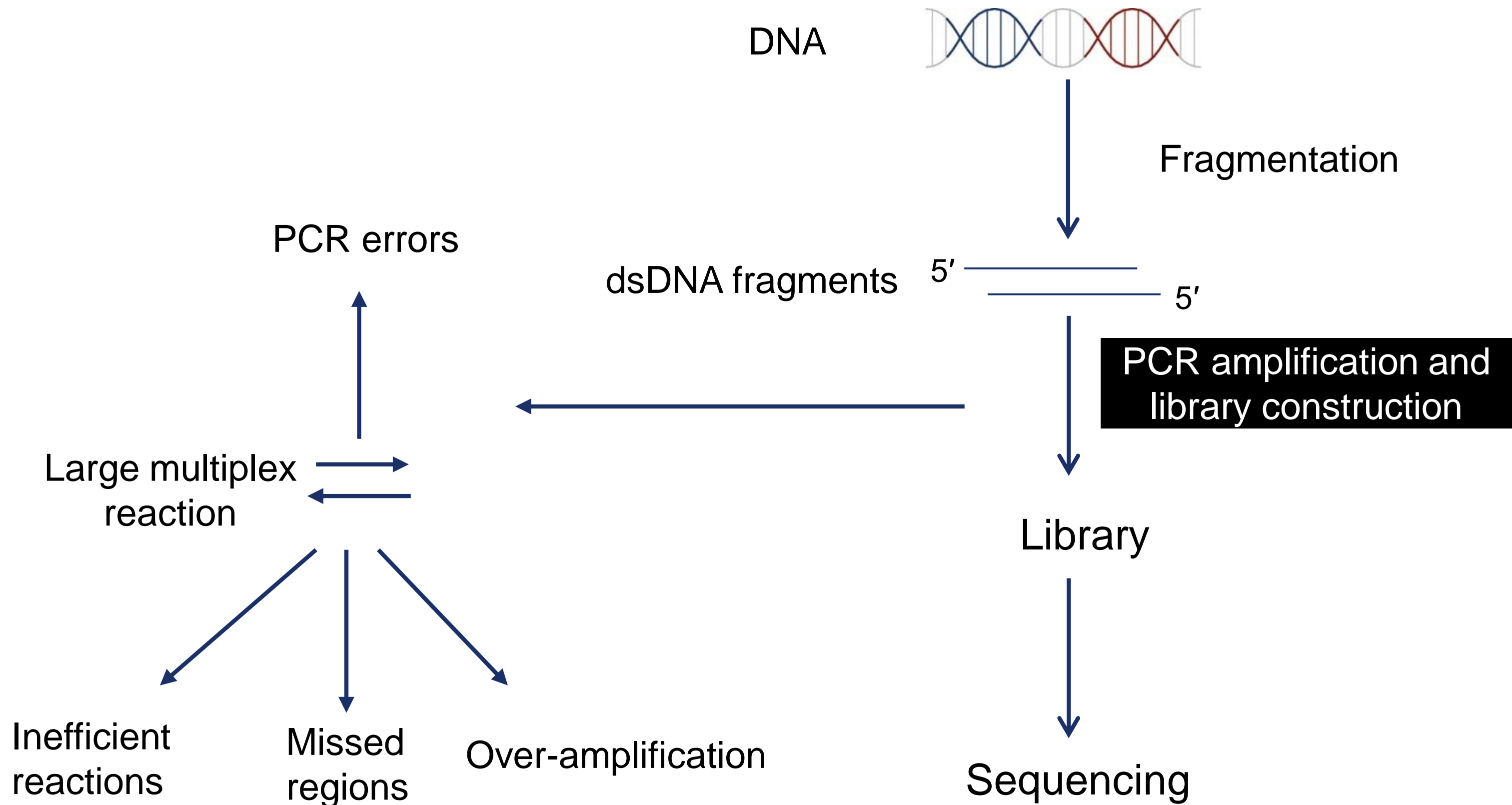
Single primer extension

PCR amplification is required for targeted sequencing, but is a “black box”



Single primer extension

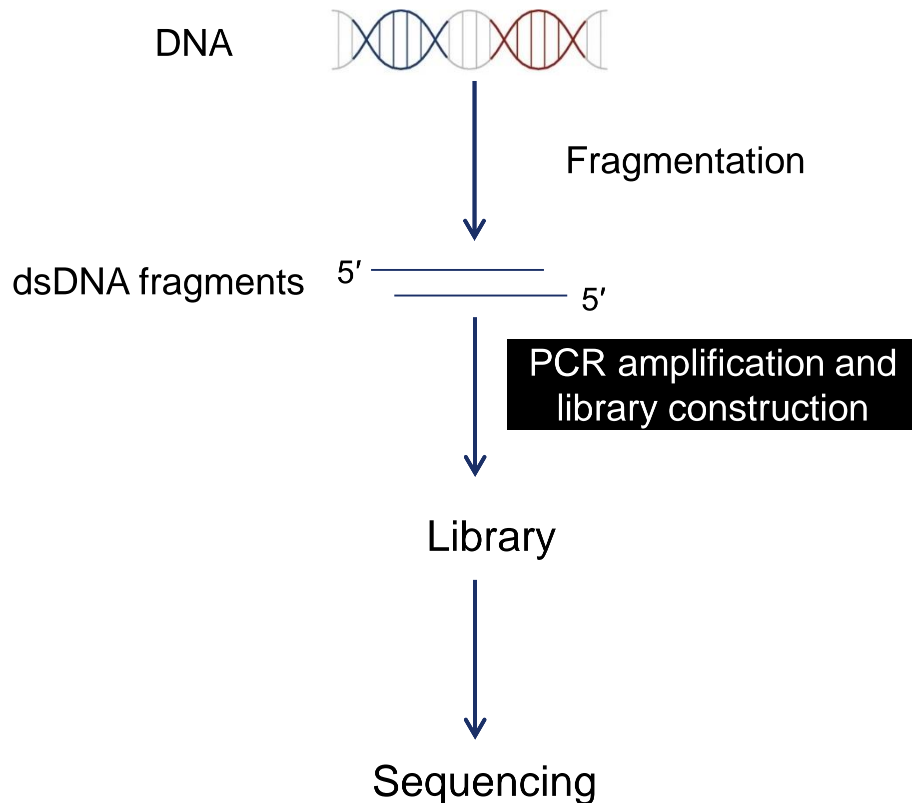
PCR amplification is required for targeted sequencing, but is a “black box”



Quality of data is influenced by the unknowns of the black box: PCR errors, PCR coverage and PCR failure.

Single primer extension

The black box presents several challenges

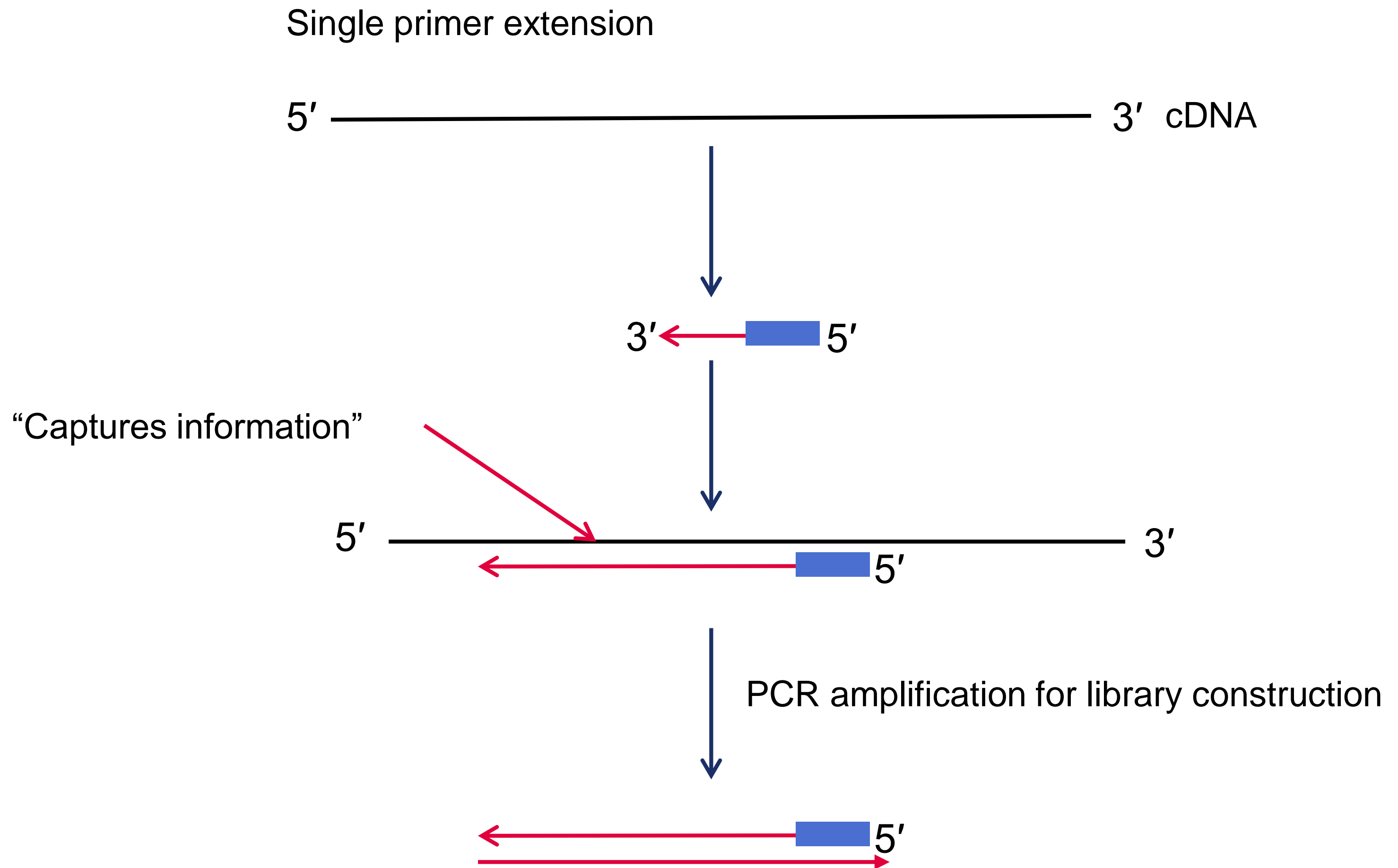


- Unintended PCR errors
- Missed regions due to high G/C content

Is the detected low-frequency variant a true variant or a PCR error?

Single primer extension

The solution: SPE with UMIs followed by PCR with common primers



Single primer extension

Advantages

Requires only 1 region for primer design

- Unlocks entire transcriptome, genome and fusion genes
- Fifty-percent lower primer amount allows for greater multiplexing

Ability to adopt to G/C rich and “hard-to-PCR” regions

- Can sequence everything

Uniform reaction

- Uniform library construction – uniform sequencing

The strategy works very well on challenging sample types

- FFPE
- Fragmented
- Low quality

Agenda

1 Introduction: QIAseq targeted sequencing

2 Principles of unique molecular indices (UMIs)

3 Single primer extension (SPE) versus PCR for capturing information during library construction

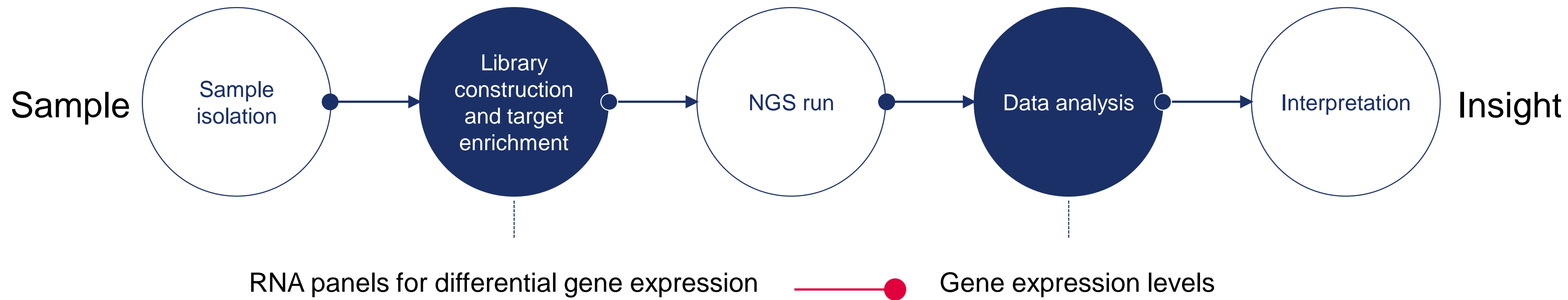
4 UMI and SPE in action – gene expression analysis

5 miRNA quantification and biofluids

6 DNA variant analysis and novel gene fusions discovery

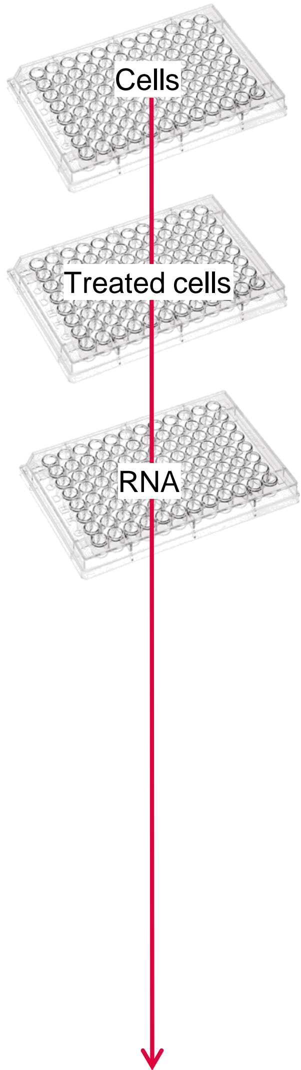
7 Summary and questions

UMI in action – gene expression example



UMI and SPE in action – gene expression example

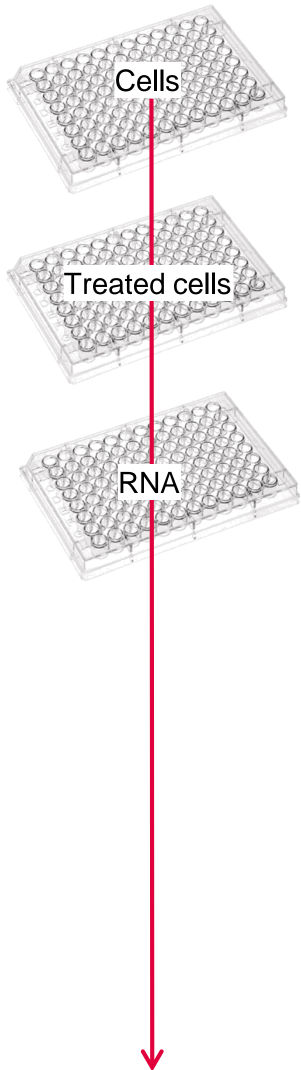
Small molecules – signal transduction application



- Experiment is to identify novel compounds that modulate known signal transduction pathways

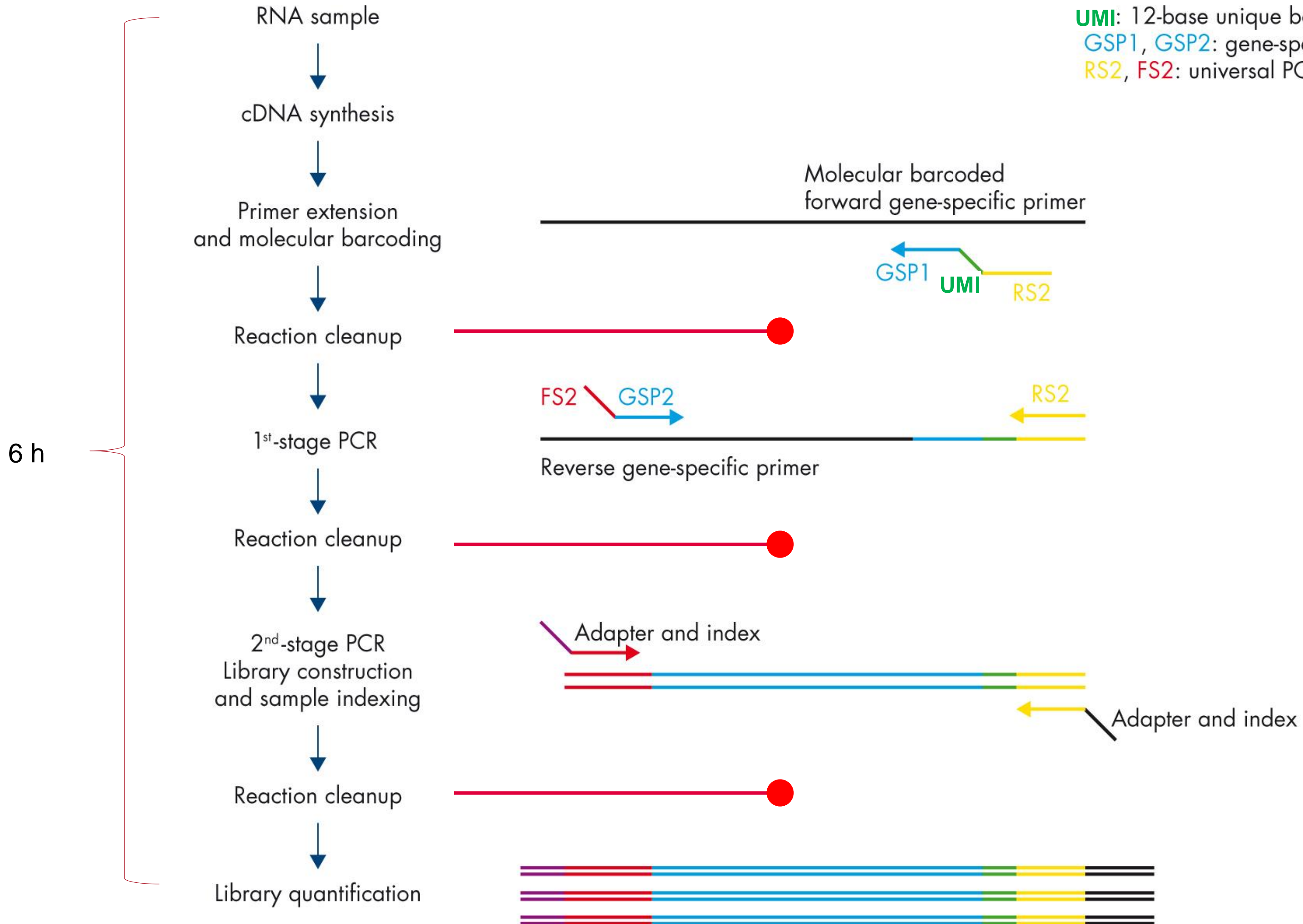
UMI and SPE in action – gene expression example

Small molecules – signal transduction application



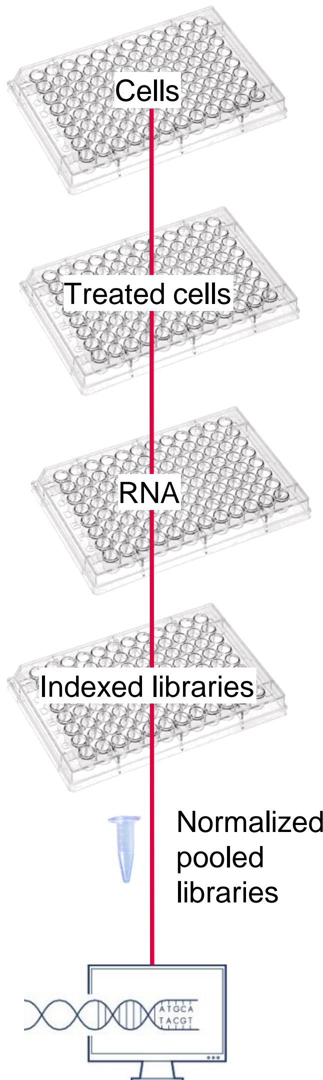
- Treat cells with different chemical inhibitors
- Isolate RNA from cells
- Build library with QIAseq Targeted RNA Panels
- Human Signal Transduction PathwayFinder – 421 targets/10 ng total RNA

UMI and SPE in action – gene expression example



UMI and SPE in action – gene expression example

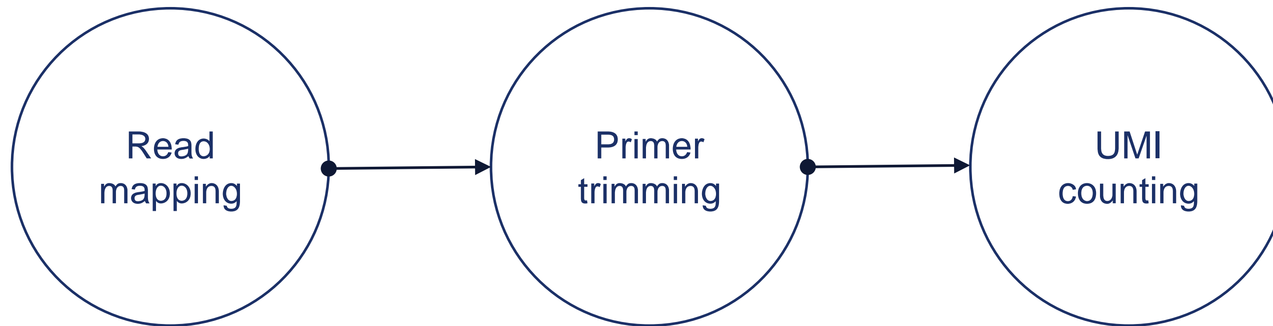
Small molecules – signal transduction application



- HEK293T cells were treated with 90 different chemical inhibitors
- The 421 Signal Transduction Gene QIAseq Panel was interrogated
- In a single day, we accomplished sequence ready libraries from total RNA for 96 samples
- The final libraries were quantified, normalized and pooled
- Prior to loading onto a NextSeq, the denatured libraries were diluted to the appropriate input concentration to generate suitable clusters on the NextSeq
- A parameter of the NextSeq sequencing run: a single 151 bp read with a custom sequencing primer (included in the kit)

UMI and SPE in action – gene expression example

QIAseq Targeted RNA Panel data analysis – automated workflow

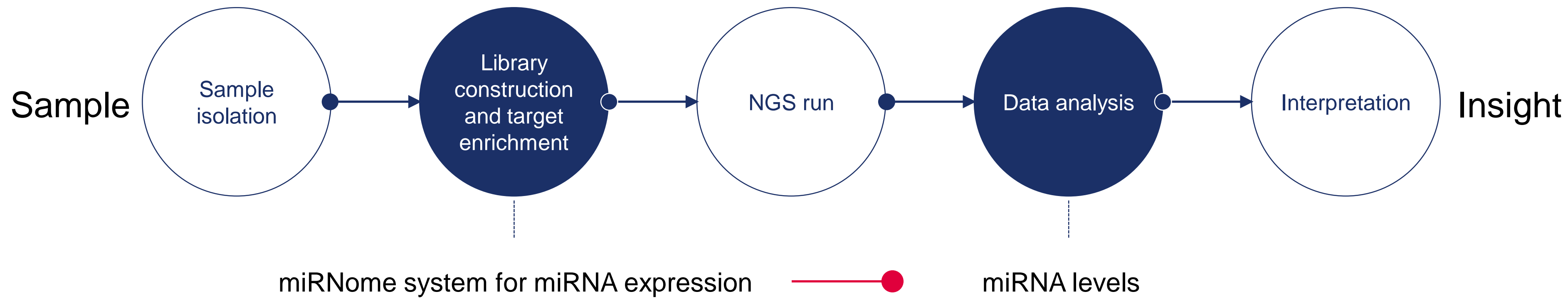


- Read mapping
  - Identify the possible position of the read within the reference genome
  - Align the read sequence to reference sequences
- Primer trimming
  - Remove the primer sequences from the reads
- UMI counting
- Calculation of gene expression

Agenda

1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. Single primer extension (SPE) versus PCR for capturing information during library construction
4. UMI and SPE in action – gene expression analysis
5. **miRNA quantification and biofluids**
6. DNA variant analysis and novel gene fusions discovery
7. Summary and questions

miRNA sequencing with UMIs

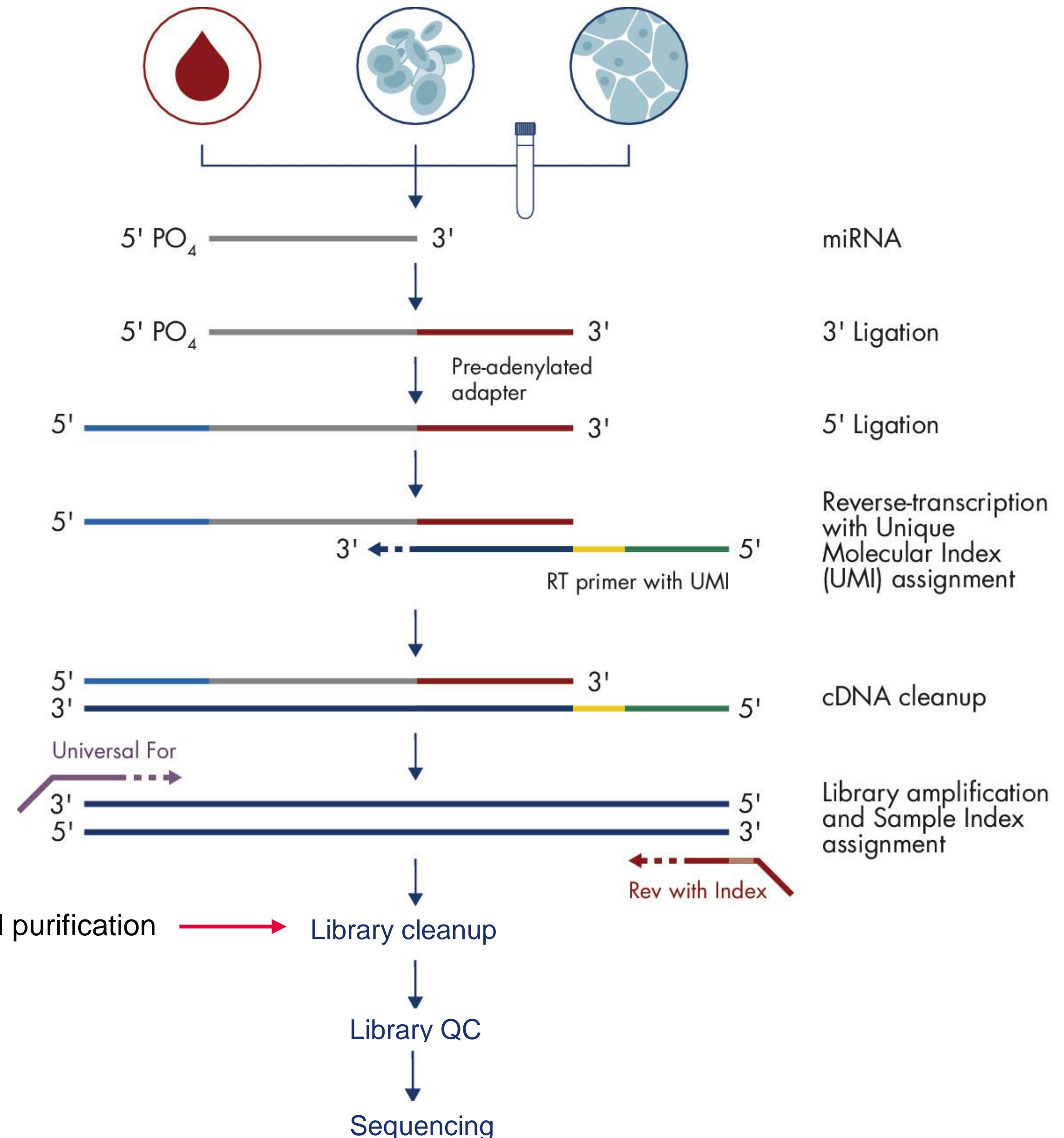


miRNA sequencing with UMIs

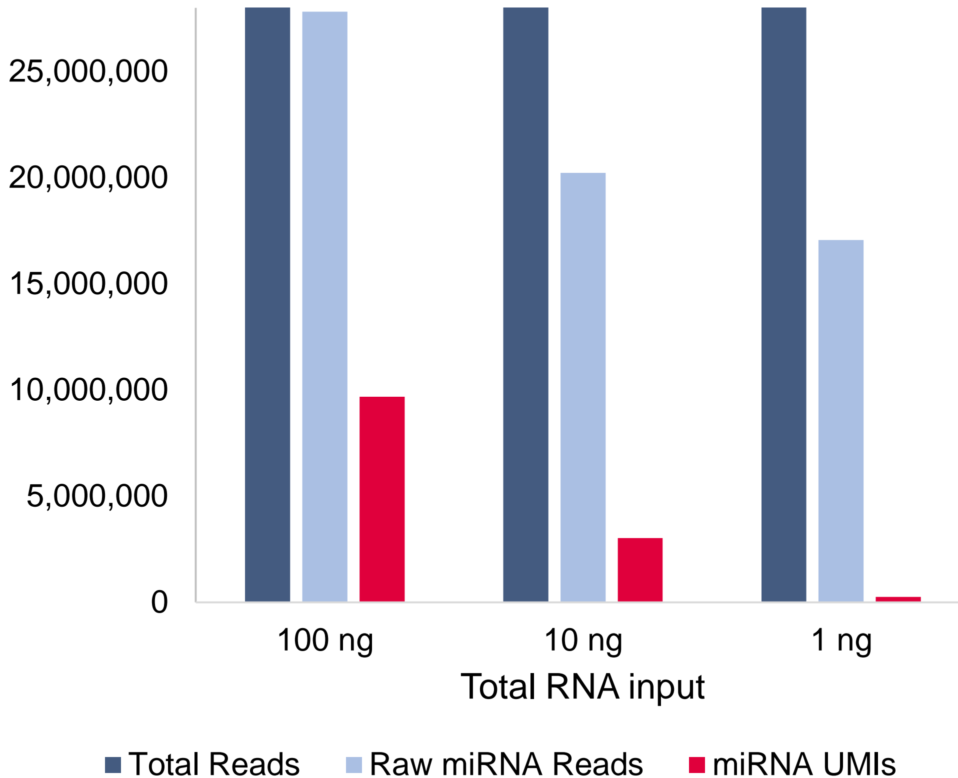
A total of 8 hours to sequencing

- Sample isolation: fluid, tissues, cells (1–2 h)
- 3′ ligation (1.5 h)
- 5′ ligation (1 h)
- RT with UMI assignment (1 h)
- cDNA cleanup (30 min)
- Library amplification and sample Index assignment (1 h)
- Library cleanup (1.5 h)
- Library QC (30 min+)

No gel purification



miRNA sequencing with UMIs



Assessment of raw miRNA reads

- Sequencing of the same miRNA molecule over and over results in an overestimation of miRNA expression
- The lower the RNA input, the lower the effect

Assessment of miRNA UMIs

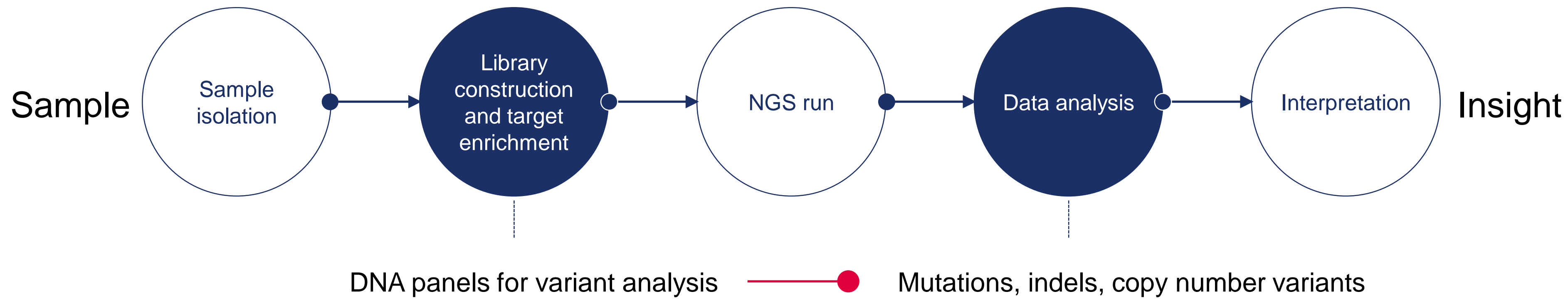
- Individual miRNA molecules are being counted, resulting in a true assessment of miRNA expression
- The lower the RNA input, the more powerful the UMIs

UMIs give a true readout of miRNA expression.

Agenda

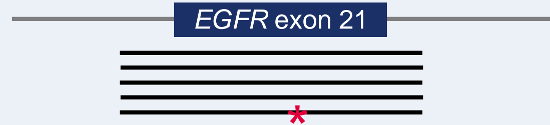
1. Introduction: QIAseq targeted sequencing
2. Principles of unique molecular indices (UMIs)
3. Single primer extension (SPE) versus PCR for capturing information during library construction
4. UMI and SPE in action – gene expression analysis
5. miRNA quantification and biofluids
6. **DNA variant analysis and novel gene fusions discovery**
7. Summary and questions

DNA variant analysis and novel gene fusions discovery



PCR and sequencing errors (artifacts) limit variant calling accuracy

Traditional targeted DNA sequencing



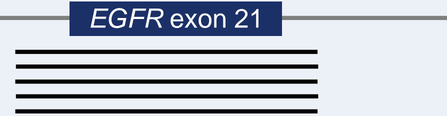
A mutation is seen in 1 out of 5 reads that map to *EGFR* exon 21
Cannot accurately tell whether the mutation is:

- A PCR or sequencing error (artifact/false positive), or
- A true low-frequency mutation

Variant calling based on non-unique reads does not reflect the mutational status of original DNA molecules; applies to a wide range of panels.

Count and analyze single original molecules (not total reads) = digital sequencing

Traditional targeted DNA sequencing

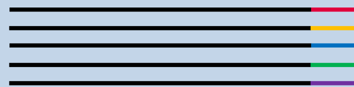


Five reads or library fragments that look exactly the same
Cannot tell whether they represent:

- Five unique DNA molecules, or
- Quintuplets of the same DNA molecule (PCR duplicates)

Add UMIs
before any amplification

Digital sequencing with UMIs



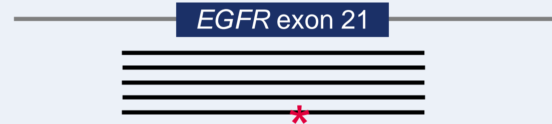
Five unique DNA molecules since 5 UMIs are detected

Quintuplets of the same DNA molecule (PCR duplicates) since 1 UMI is detected

DNA variant analysis and novel gene fusions discovery

Count and analyze single original molecules (not total reads) = digital sequencing

Traditional targeted DNA sequencing

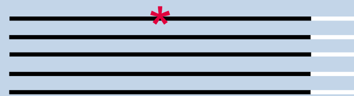


A mutation is seen in 1 out of 5 reads that map to *EGFR* exon 21
Cannot accurately tell whether the mutation is:

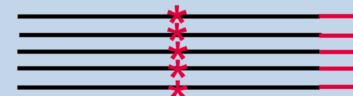
- A PCR or sequencing error (artifact/false positive), or
- A true low-frequency mutation

Add UMIs
before amplification

Digital sequencing with UMIs



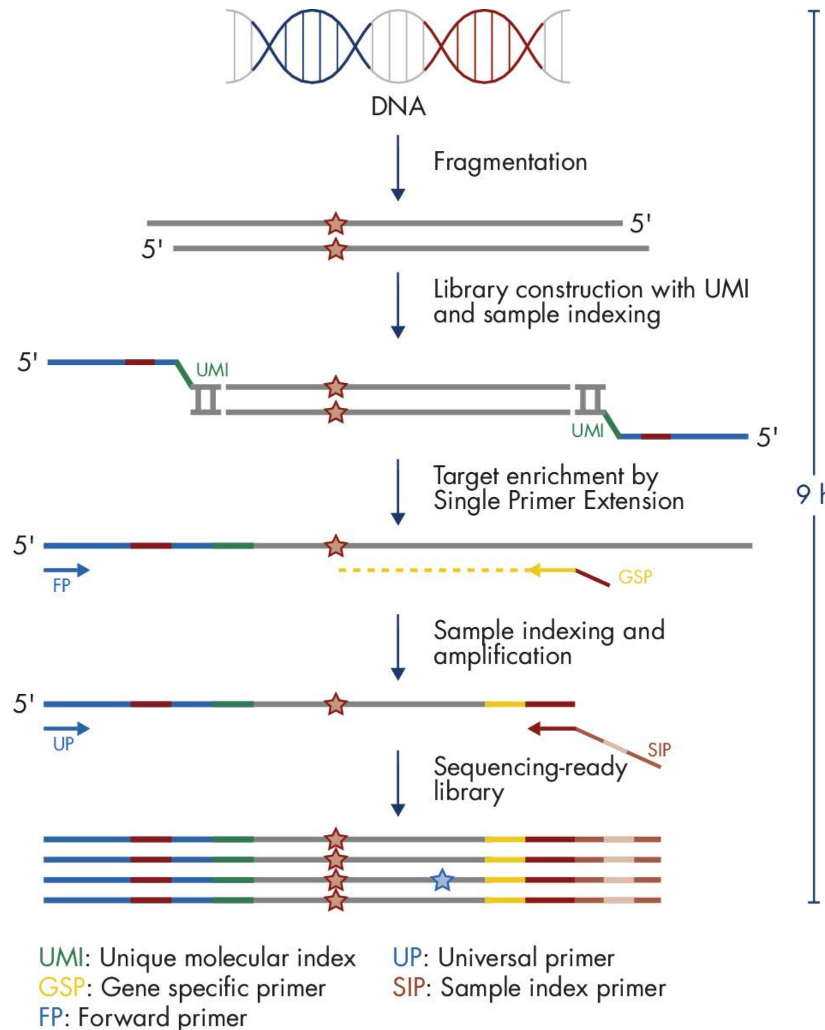
False variant is present in some fragments carrying the same UMI



True variant is present in all fragments carrying the same UMI

DNA variant analysis and novel gene fusions discovery

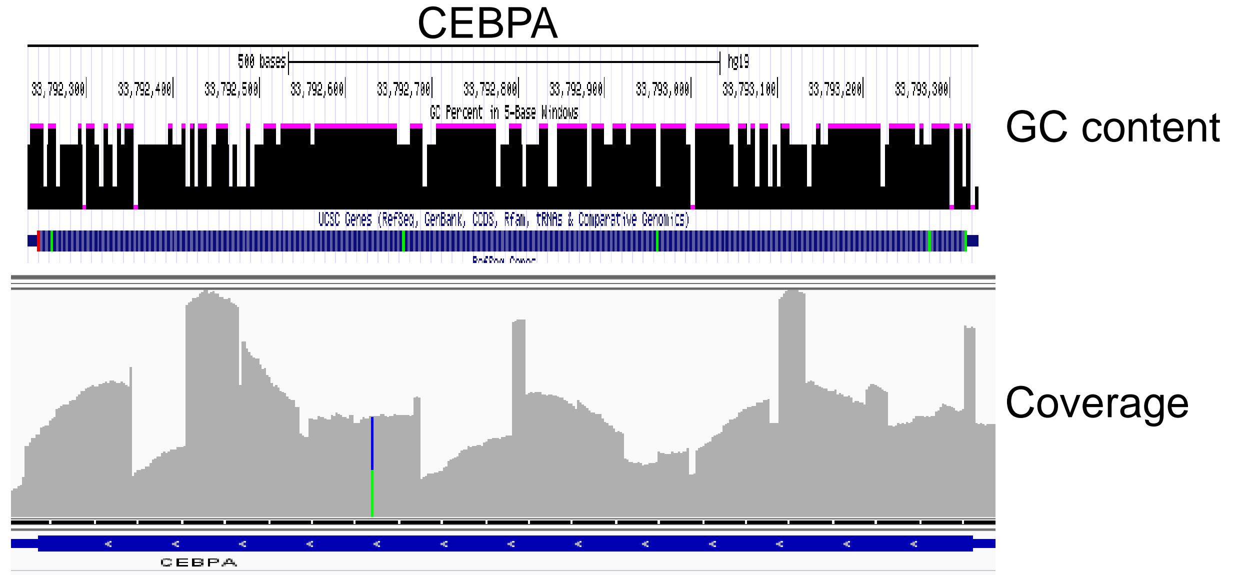
QIAseq Targeted DNA Panel workflow

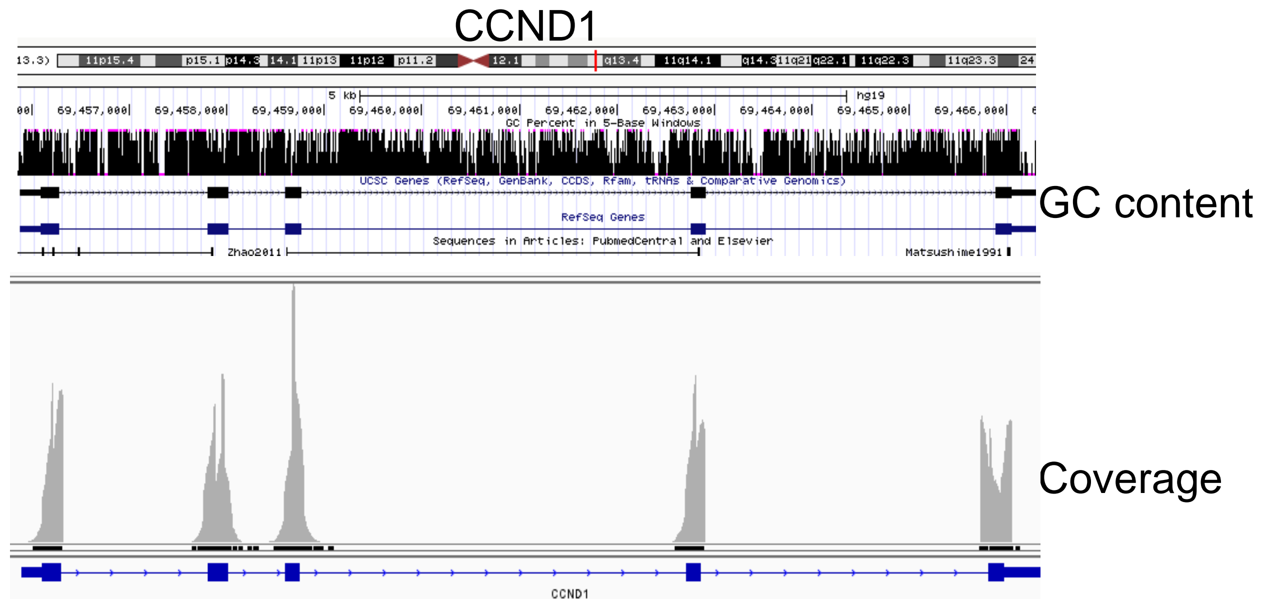


- Low DNA input
- One tube per sample
- Fast processing
- No sonication
- Automatable
- Multiple stop points
- Optimized sample-specific protocols
- Proprietary advanced chemistry
- UMI technology

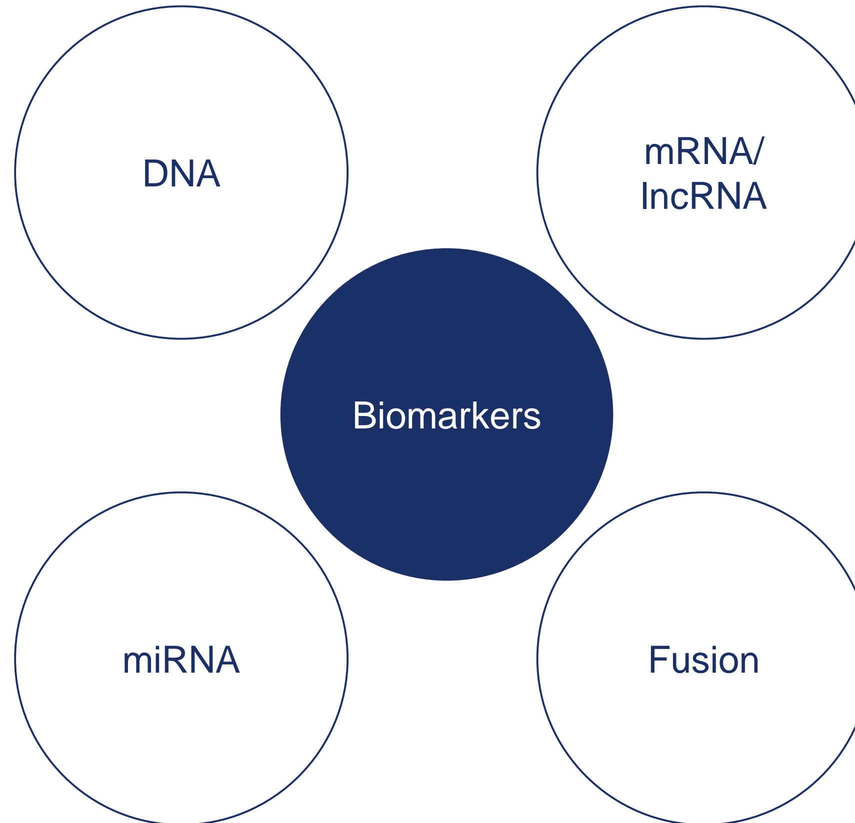
DNA variant analysis and novel gene fusions discovery

The proprietary SPE PCR chemistry used in QIAseq enables efficient coverage of regions high in GC content





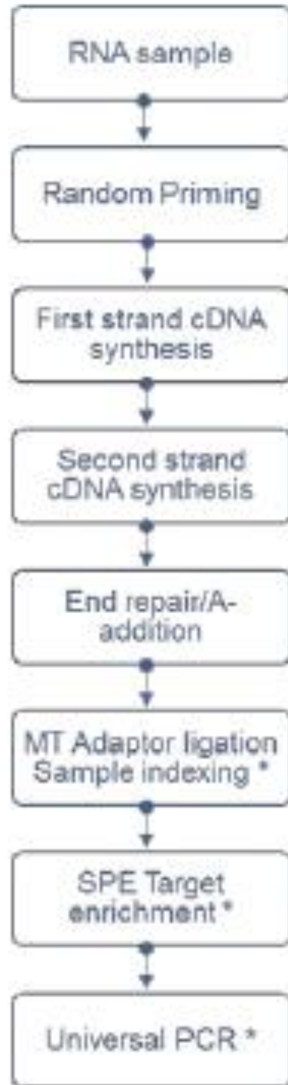
DNA variant analysis and novel gene fusions discovery



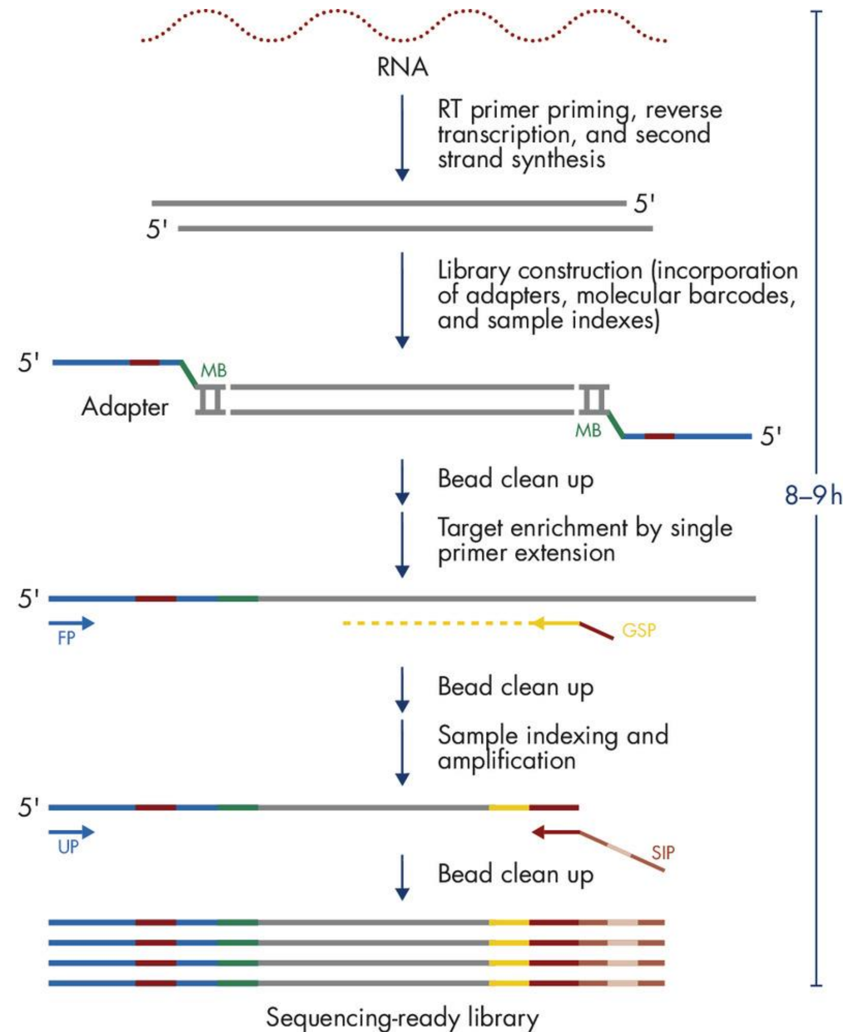
QIAseq Targeted RNAscan Panels

- Unique molecular index
- Known fusion genes (validation)
- Unknown partners (discovery)

DNA variant analysis and novel gene fusions discovery



* QIAseq beads cleanup

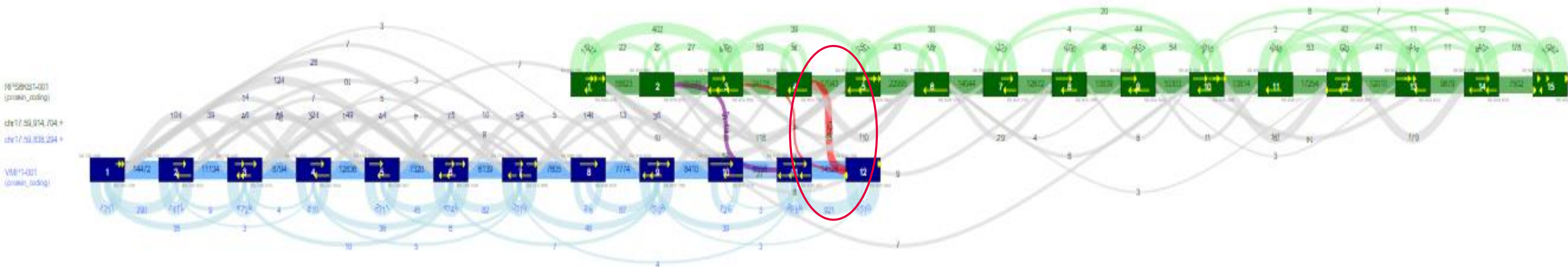


DNA variant analysis and novel gene fusions discovery

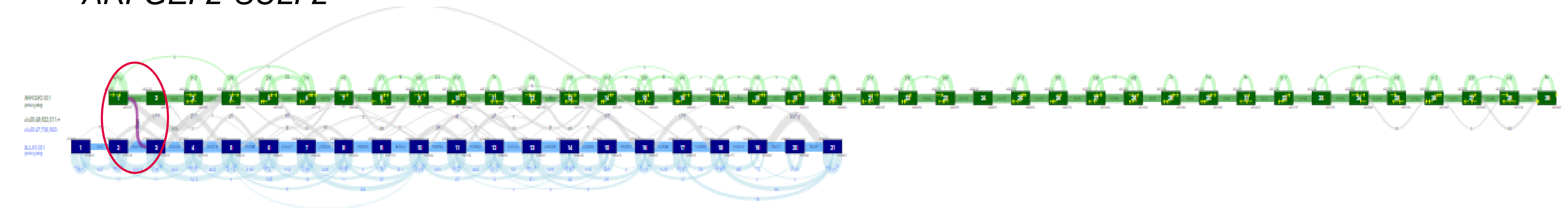
QIASeq Targeted RNAscan is a RNA target enrichment method that allows verification of known fusions and discovery of novel fusions with next-generation sequencing (NGS)

1. NCOA3--SULF2, *score*: 3806.00
2. RPS6KB1--VMP1, *score*: 1136.00
3. ARFGEF2--SULF2, *score*: 184.00
4. SULF2--PRICKLE2, *score*: 56.00
5. CLTC--VMP1, *score*: 6.00

RPS6KB1-VMP1



ARFGEF2-SULF2



Agenda

1 Introduction: QIAseq targeted sequencing

2 Principles of unique molecular indices (UMIs)

3 Single primer extension (SPE) versus PCR for capturing information during library construction

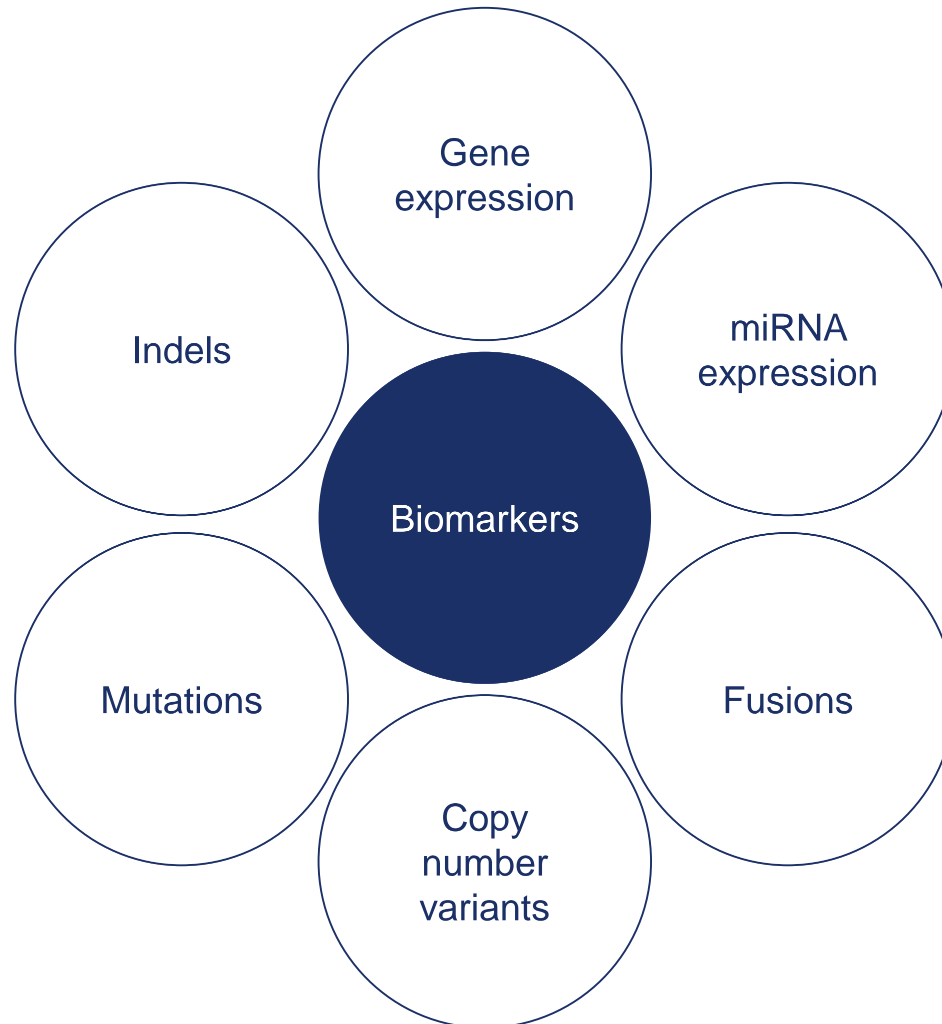
4 UMI and SPE in action – gene expression analysis

5 miRNA quantification and biofluids

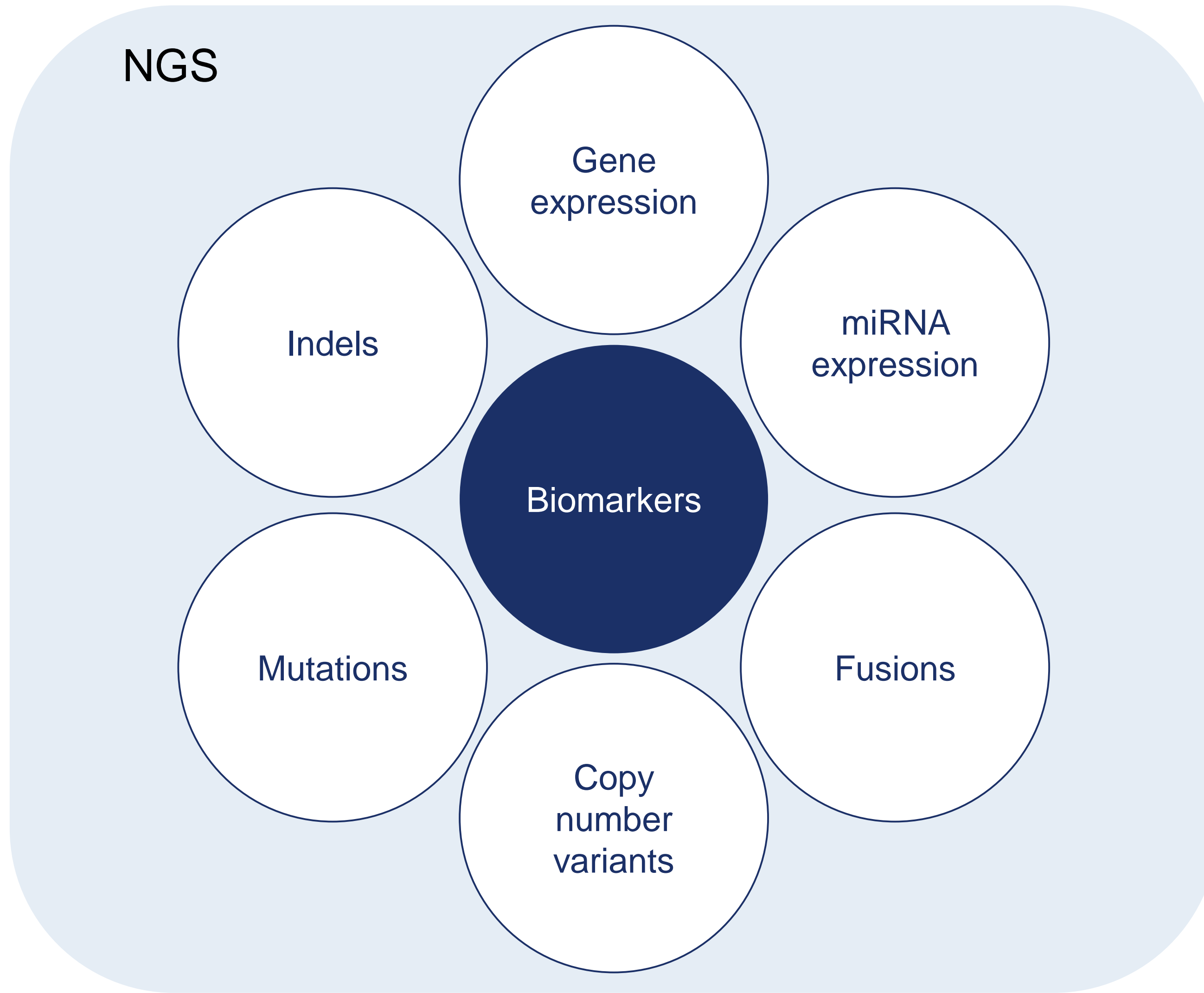
6 DNA variant analysis and novel gene fusions discovery

7 **Summary and questions**

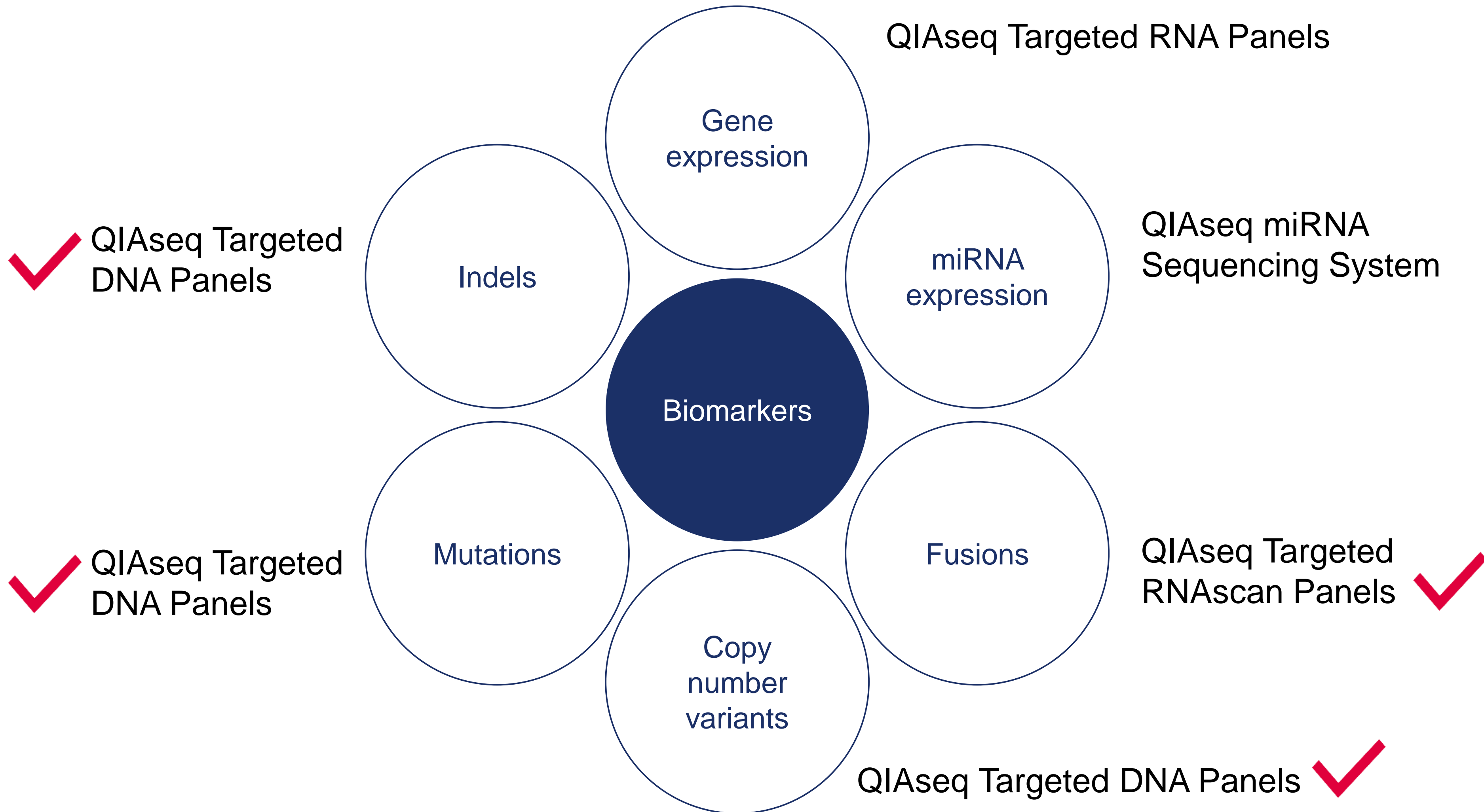
Biomarkers come in many flavors



NGS can be used for all of them



QIAseq solutions to detect all biomarkers using NGS



QIAGEN

Thank you for your attention!

Questions? **qiawebinars@qiagen.com**

Contact QIAGEN: **www.qiagen.com/about-us/contact/global-contacts/**