Complex bacterial samples and rRNA removal

Samuel Rulli, Ph.D.
Senior Global Product Manager

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

Background

FastSelect –5S/16S/23S: An introduction

FastSelect –5S/16S/23S: Results

Summary

Agenda

Background

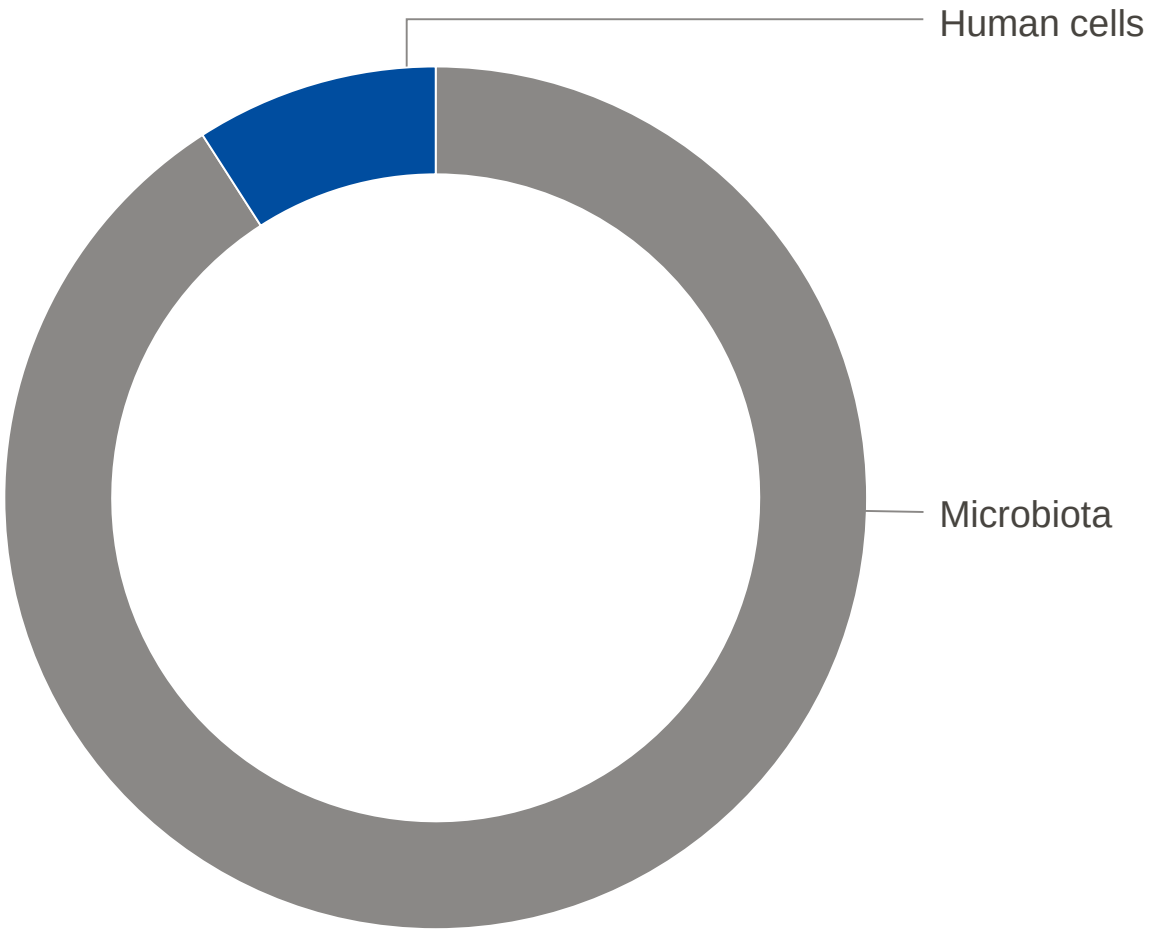
FastSelect –5S/16S/23S: An introduction

FastSelect –5S/16S/23S: Results

Summary

Humans or ‘superorganisms’?

Total number of cells: Human cells versus microbiota



Cellular composition of the ‘superorganism’

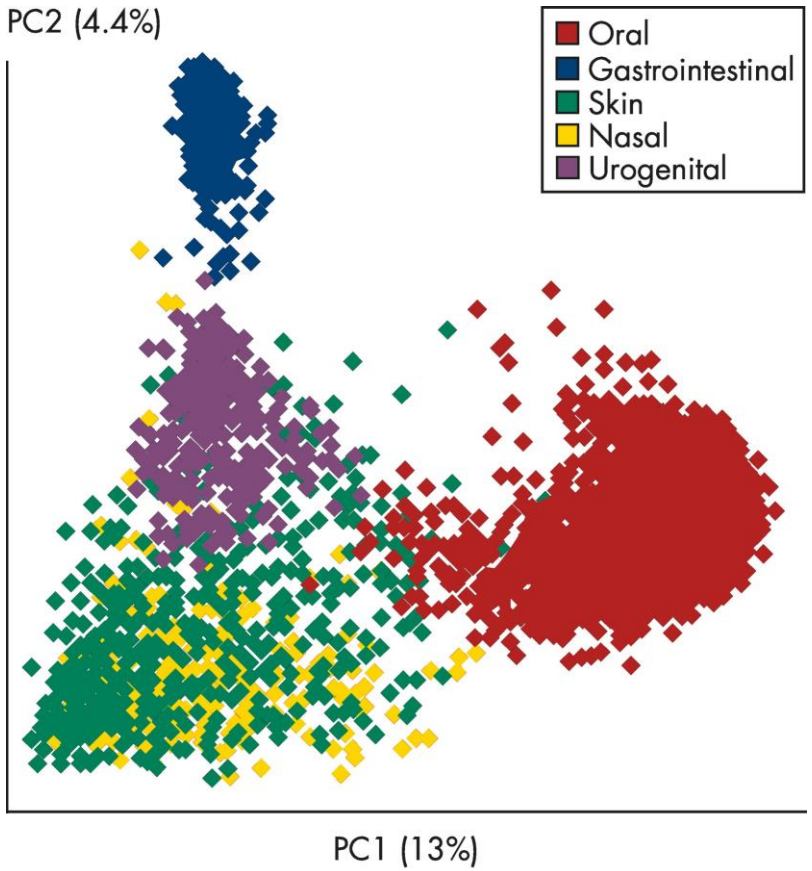
Estimation of the number of microbial cells that live in and on the human body; human cells are outnumbered by a factor of 10

Nomenclature:

- Microbiota are the microbes that live in a specific location, e.g., the human body, the gut, soil, etc.
- Metagenomics is the study of the collection of genomes derived from a specific sample or community
- Metatranscriptomics is the study of the RNA expression of genes from a community sample to interpret the physiological state of that community at that time

Microbiota composition

Microorganisms cluster by body site



Cataloguing efforts by the NIH Human Microbiome Project suggest:

- Around 10,000 organisms live with us
- Around 8 $\times 10^6$ genes constitute this ‘second genome’

Identifying microbiota in healthy individuals revealed:

- Different body sites have unique communities
- Race, age, gender, weight or ethnicity can affect microbiota composition

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature, 486(7402), 207–214.

Environmental microbiomes and metagenomics

The Earth Microbiome Project

Multidisciplinary effort to survey the microbial composition of diverse environments across the globe:

- Aims to process 200,000 samples from different biomes and generate a database of microbes and their gene products

Estimates of bacterial diversity:

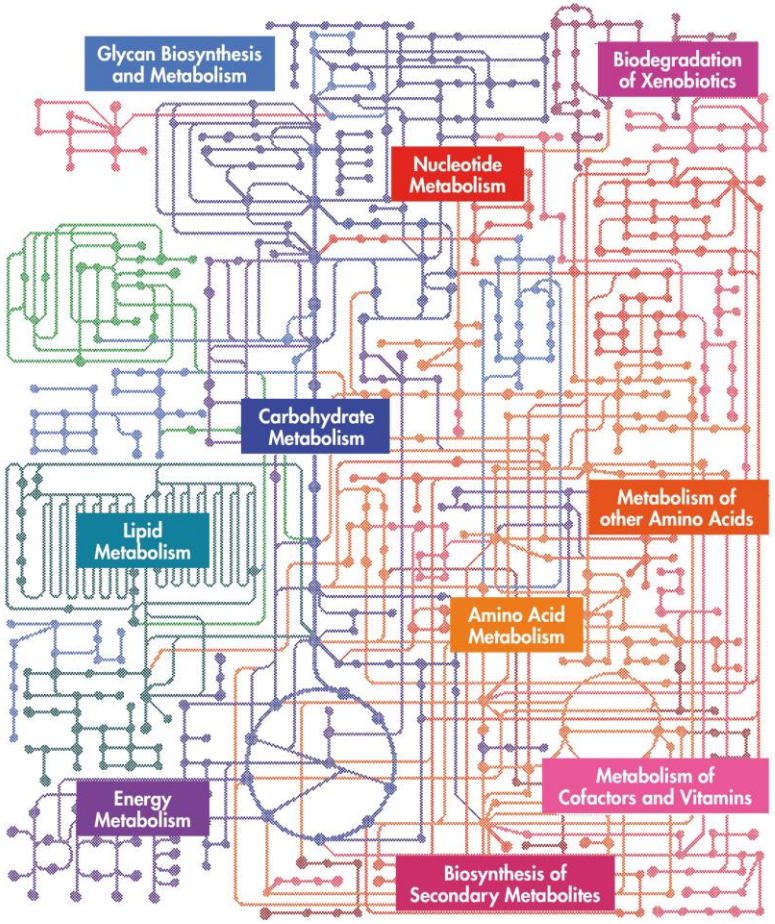
- 160 distinct types of bacteria in 1 ml of ocean water
- 6400–38,000 types of bacteria in 1 gram of soil

These estimates are for bacteria alone; they do not include viruses, archaea or fungi

Source: Curtis, T.P., Sloan, W.T. and Scannell, J.W. (2002) Estimating prokaryotic diversity and its limits. Proc Natl Acad Sci USA 99, 10494–9.

Complexity and function of genomic content

Function of microbiome enables individual survival

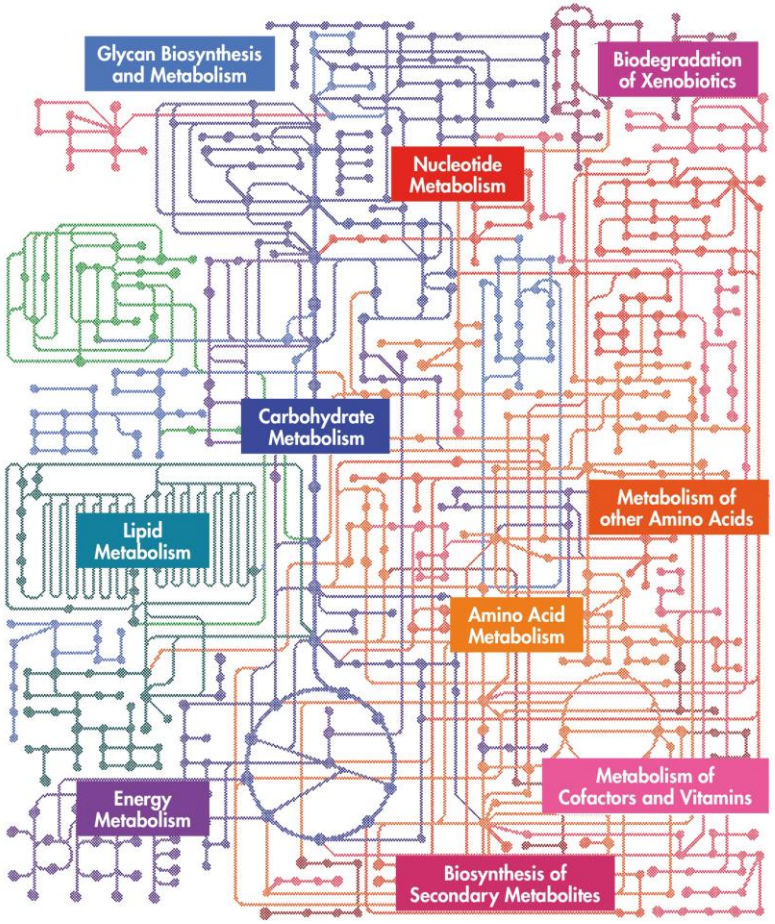


Each organism has developed its:

- Genetic content for its own survival in a specific environment
- Metabolism tuned to local nutrient sources
- Virulence factors for stable colonization
- Antibiotic resistance genes to metabolize toxins

Complexity and function of genomic content

Function of microbiome enables individual survival



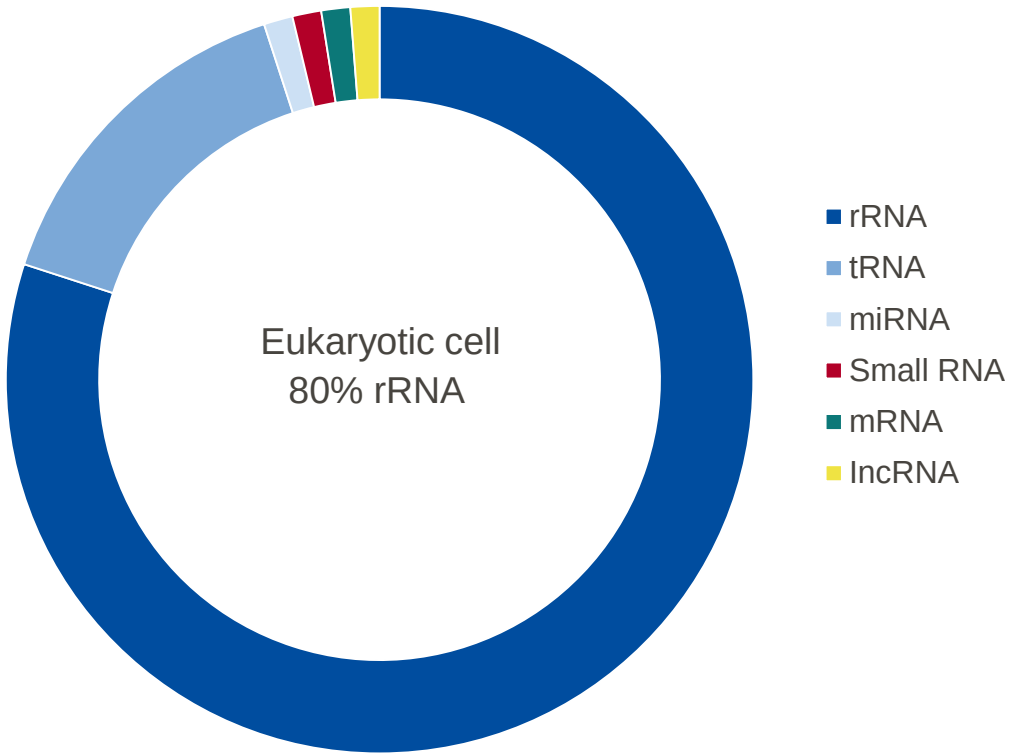
Understanding the microbiota

- DNA – who is or was there
- RNA – who is alive and what they are doing
- RNA can be used to identify:
  - Metabolic status
  - Expression of virulence factors
  - Expression of antibiotic resistance
  - Both host and microbiota

Library enrichment/depletion strategies: Whole transcriptome NGS

Enrichment or depletion is necessary to maximize reads from the RNAs of interest

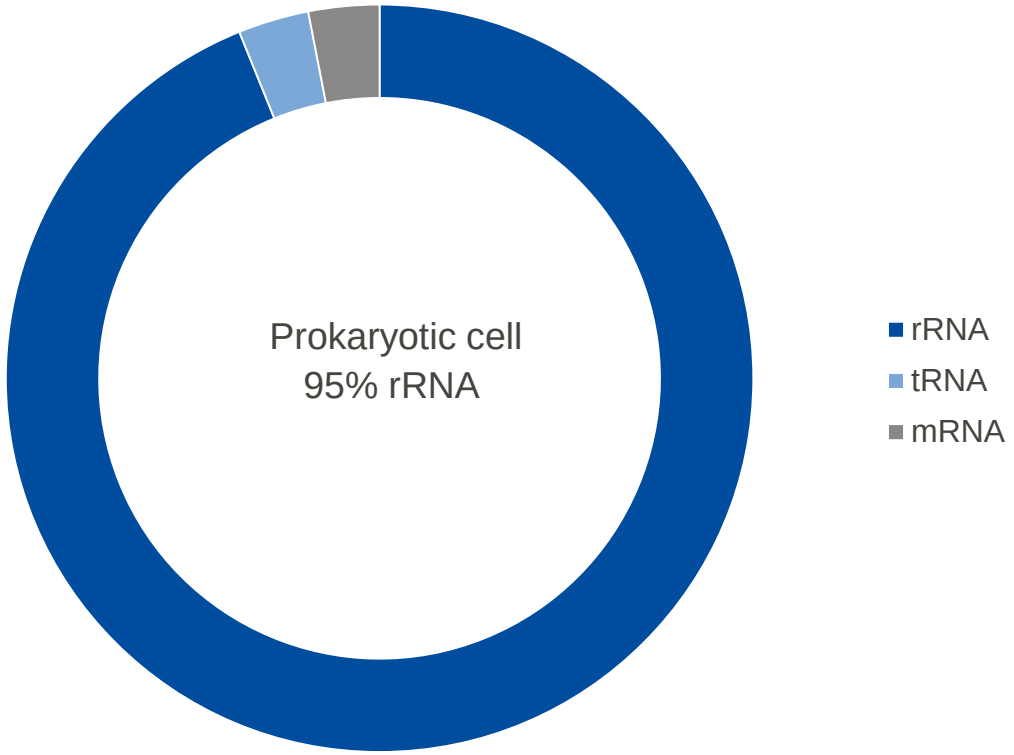
- Typical RNA composition in a eukaryotic cell: >80% rRNA
- Highly abundant transcripts consume a lot of reads
- Enrichment or depletion is used to obtain more reads from the RNAs of interest, such as:
  - mRNA
  - lncRNA



Library enrichment/depletion strategies: Whole transcriptome NGS

Enrichment or depletion is necessary to maximize reads from the RNAs of interest

- Typical RNA composition in a prokaryotic cell: 95% rRNA
- Highly abundant transcripts consume a lot of reads
- Enrichment or depletion is used to obtain more reads from the RNAs of interest, such as:
  - mRNA
  - lncRNA

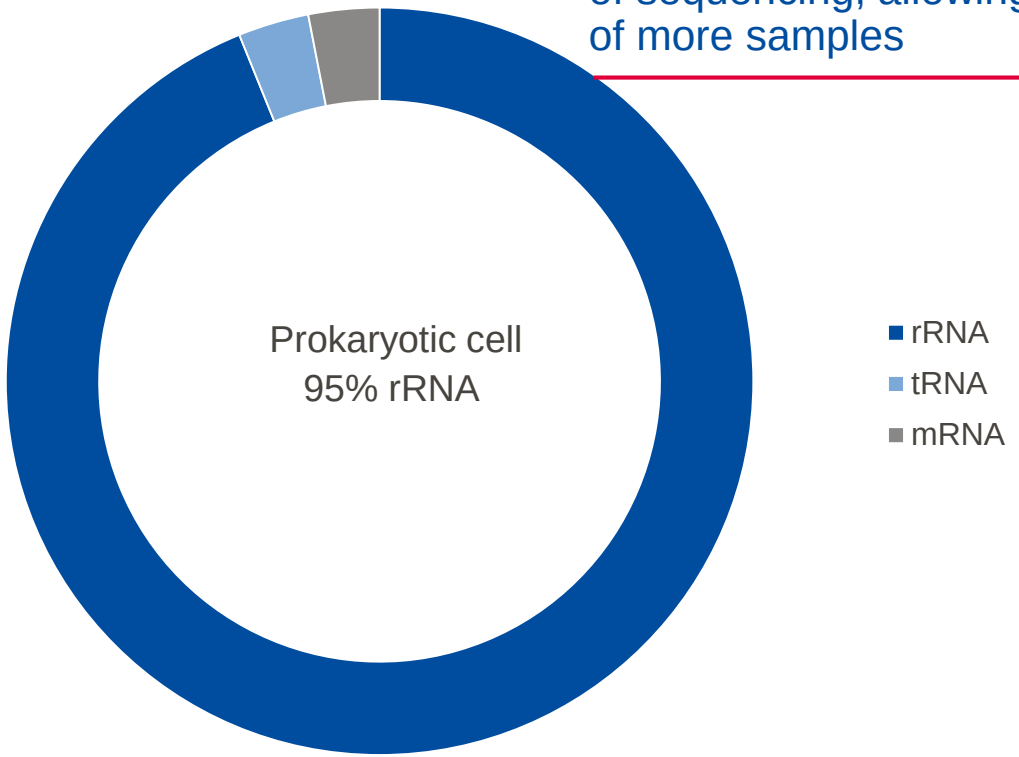


Library enrichment/depletion strategies: Whole transcriptome NGS

Enrichment or depletion is necessary to maximize reads from the RNAs of interest

- Typical RNA composition in a prokaryotic cell: 95% rRNA
- Highly abundant transcripts consume a lot of reads
- Enrichment or depletion is used to obtain more reads from the RNAs of interest, such as:
  - mRNA
  - lncRNA

Removing rRNA will increase sensitivity and decrease the cost of sequencing, allowing analysis of more samples

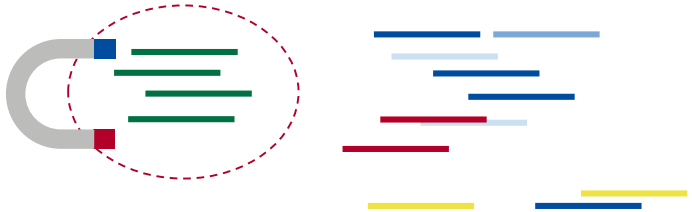


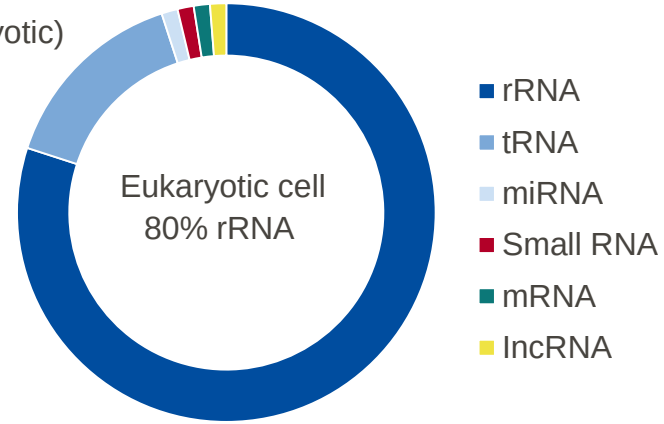
Library enrichment/depletion strategies: Whole transcriptome NGS

Enrichment/depletion strategies

Poly(A) enrichment

- Not useful for fragmented samples
- Not useful for prokaryotic samples



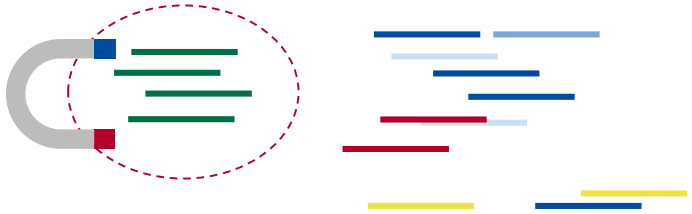


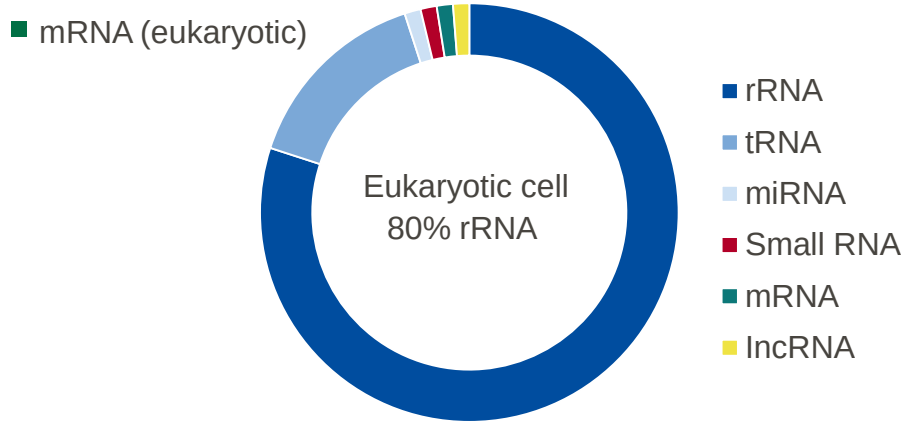
Library enrichment/depletion strategies: Whole transcriptome NGS

Enrichment/depletion strategies

Poly(A) enrichment

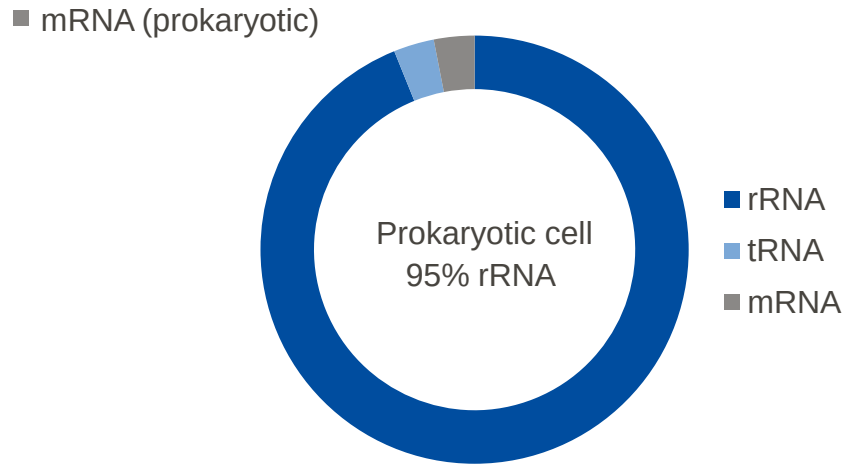
- Not useful for fragmented samples
- Not useful for prokaryotic samples





rRNA depletion

- The “catch all” method



rRNA depletion is the only choice for meta-transcriptomics.

Agenda

Background

FastSelect –5S/16S/23S: An introduction

FastSelect –5S/16S/23S: Results

Summary

NEW

FastSelect –5S/16S/23S: An overview

What is FastSelect –5S/16S/23S?

- Fragmentation and pan-bacterial (5S/16S/23S) rRNA depletion module

Number of reactions:

- 24, 96 and 384

How does it work?

- Inhibits reverse transcription of its specific targets

Coverage:

- Designed to block community level cDNA synthesis of 5S, 16S and 23S rRNA
- Designed against SILVA 16S sequences (nearly 600,000 unique entries), SILVA 23S sequences (nearly unique 170,000 entries) and 5S rRNA Database (over 7,200 unique entries)
- Theoretically blocks >95% cDNA synthesis of all 5S, 16S and 23S rRNA sequences
  - In practice, results will vary, based on the exact composition of the sample

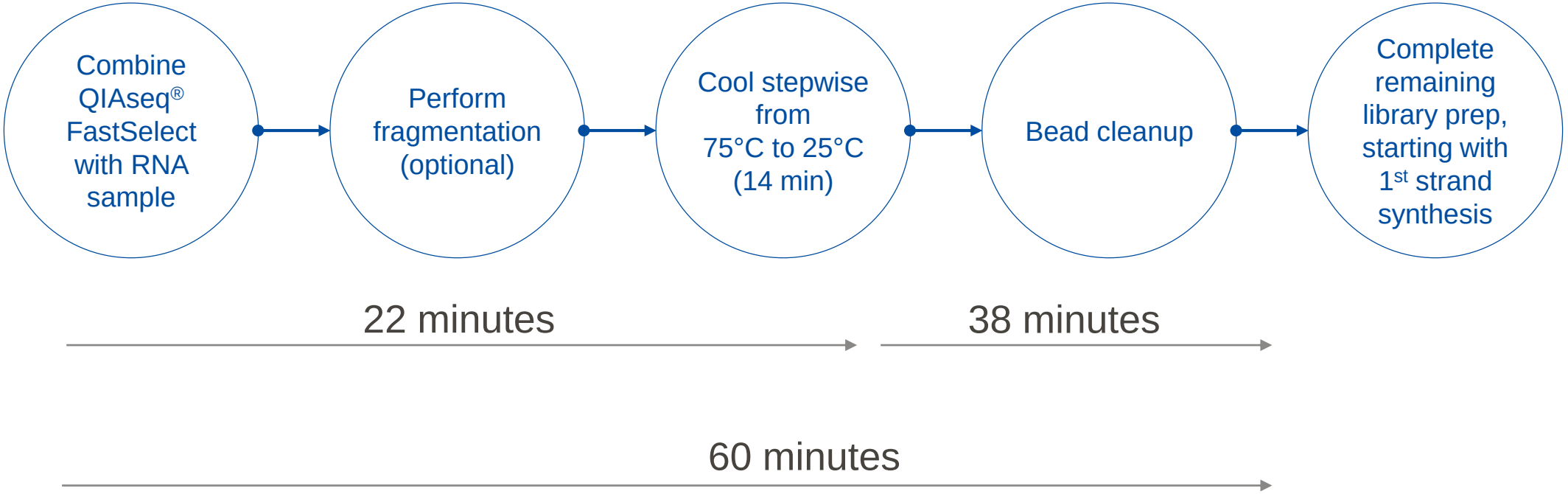
Total RNA input:

- 20 ng to 1 µg

Tested RNA library prep kit compatibility:

- QIAseq Stranded Total RNA Lib Kit (QIAGEN Group), TruSeq® Stranded (Illumina®, Inc), NEBNext® Ultra II Directional (New England Biolabs, Inc)
- FastSelect is compatible with most RNA library prep kits

FastSelect –5S/16S/23S: Workflow



FastSelect –5S/16S/23S: What's in the box?

Kit sizes:

- 24 (cat. no. 335925)
- 96 (cat. no. 335927)
- 384 (cat. no. 335929)

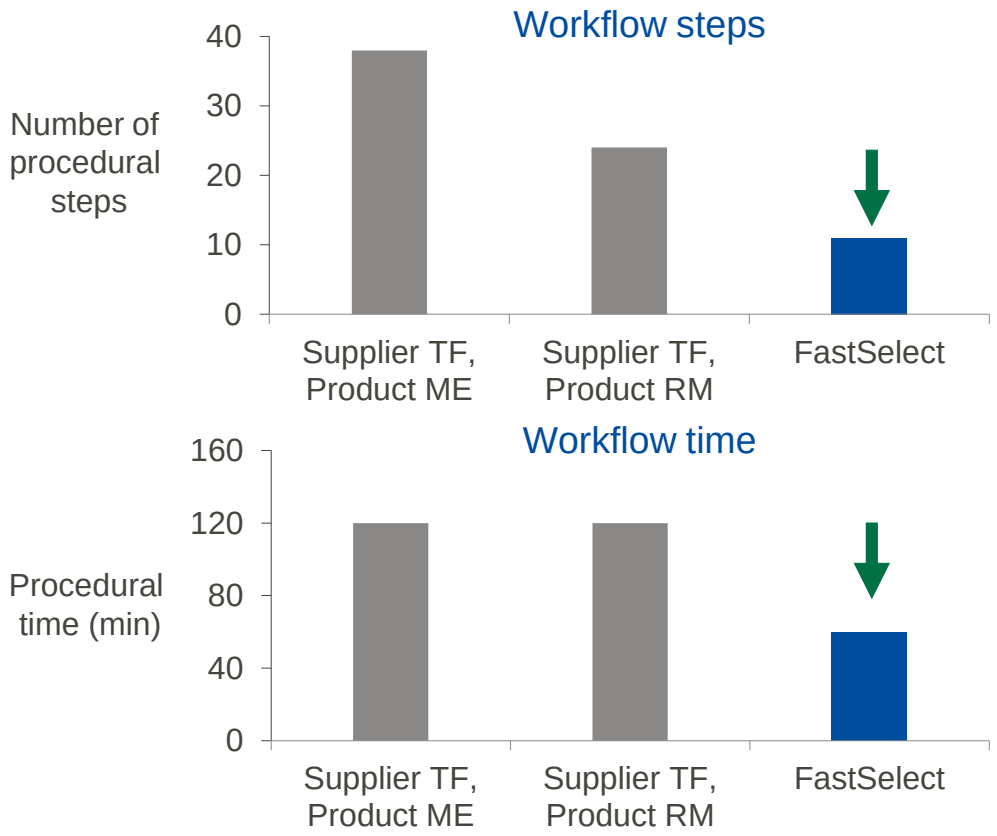
Kit components and purpose:

- FastSelect 5S/16S/23S: rRNA removal reagent
- FastSelect FH Buffer: Fragmentation and/or hybridization buffer
- Nuclease-free Water
- QIAseq Beads
- QIAseq Bead Binding Buffer

NEW

QIAseq FastSelect –5S/16S/23S: rRNA removal

FastSelect offers an attractive, streamlined workflow versus kits from another supplier.



Supplier TF, Product RM

Bead preparation

(7 steps)

Selective hybridization

(4 steps)

Removal of rRNA

(5 steps)

RNA purification

(8 steps)

Library Prep

FastSelect –5S/16S/23S

Add FastSelect reagent and ramp down temperature

(1 step)

Bead cleanup

(11 steps)

Library Prep

Agenda

Background

FastSelect –5S/16S/23S: An introduction

FastSelect –5S/16S/23S: Results

Summary

FastSelect –5S/16S/23S robustly removes rRNA from single-species samples

Experimental overview

- Sample: 1 µg and 100 ng, DH5α *E. coli* total RNA (Thermo Fisher)
- Depletion: No depletion; FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect –5S/16S/23S robustly removes rRNA from single-species samples

Experimental overview

- Sample: 1 µg and 100 ng, DH5α *E. coli* total RNA (Thermo Fisher)
- Depletion: No depletion; FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to rRNA, FastSelect —5S/16S/23S
E. coli K12, 1 µg, R1	2.01
E. *coli* K12, 1 µg, R2	1.88
E. coli K12, 100 ng, R1	0.55
E. coli K12, 100 ng, R2	2.57

FastSelect –5S/16S/23S robustly removes rRNA from single-species samples

Experimental overview

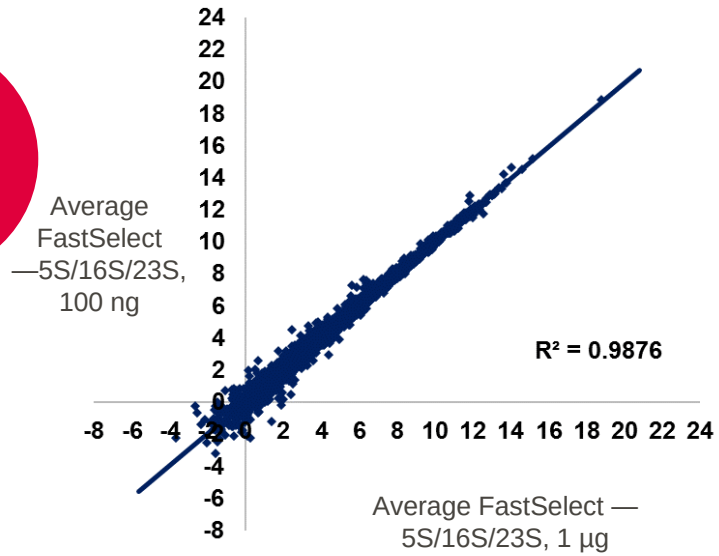
- Sample: 1 µg and 100 ng, DH5α *E. coli* total RNA (Thermo Fisher)
- Depletion: No depletion; FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to rRNA, FastSelect —5S/16S/23S
E. coli K12, 1 µg, R1	2.01
E. *coli* K12, 1 µg, R2	1.88
E. coli K12, 100 ng, R1	0.55
E. coli K12, 100 ng, R2	2.57

98% rRNA removed

Strong gene expression correlation between inputs



FastSelect –5S/16S/23S robustly removes rRNA from single-species samples

Experimental overview

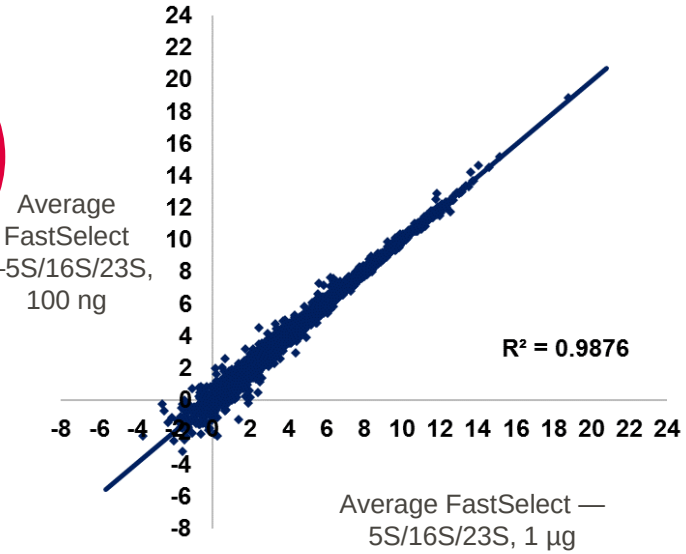
- Sample: 1 µg and 100 ng, DH5α *E. coli* total RNA (Thermo Fisher)
- Depletion: No depletion; FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to rRNA, FastSelect —5S/16S/23S
E. coli K12, 1 µg, R1	2.01
E. *coli* K12, 1 µg, R2	1.88
E. coli K12, 100 ng, R1	0.55
E. coli K12, 100 ng, R2	2.57

98% rRNA removed

Strong gene expression correlation between inputs



- FastSelect efficiently removes more than 95% of rRNA
- Gene expression values from FastSelect-treated samples, even at different RNA input amounts, are highly correlative (Log2 RPKM > 0.3)

FastSelect –5S/16S/23S: Robust rRNA removal from gut bacterial communities

Experimental overview

- Sample: 100 ng, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect –5S/16S/23S
- Library prep: QIAseq Stranded Total RNA Lib Kit
- Sequencing: NextSeq 550 (2 x 75 bp)
- Mapping: CLC Genomics Workbench

FastSelect –5S/16S/23S: Robust rRNA removal from gut bacterial communities

Experimental overview

- Sample: 100 ng, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect –5S/16S/23S
- Library prep: QIAseq Stranded Total RNA Lib Kit
- Sequencing: NextSeq 550 (2 x 75 bp)
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

	Percentage of reads mapped to bacterial rRNA (total)	
Sample	**No treatment**	**FastSelect –5S/16S/23S**
ATCC Gut Microbiome (12 bacteria)	96.35	12.32

85% rRNA removed

FastSelect –5S/16S/23S: Robust rRNA removal from gut bacterial communities

Experimental overview

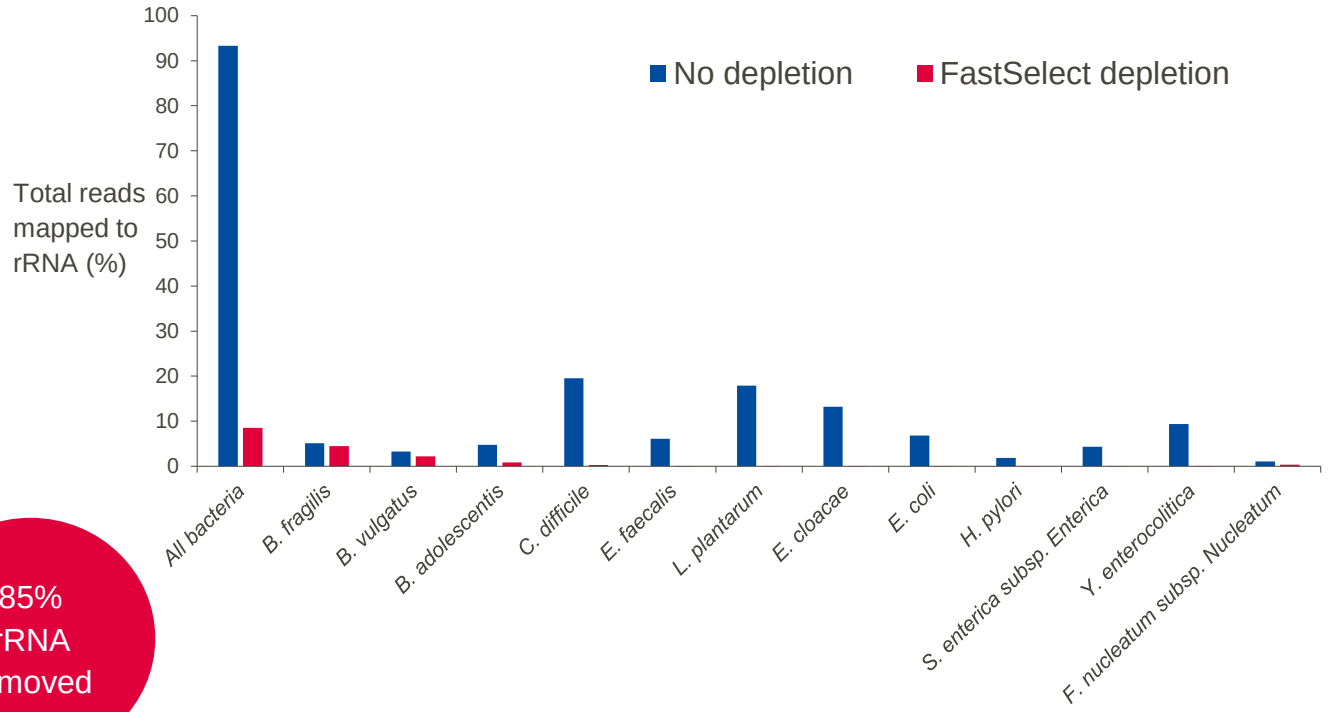
- Sample: 100 ng, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect –5S/16S/23S
- Library prep: QIAseq Stranded Total RNA Lib Kit
- Sequencing: NextSeq 550 (2 x 75 bp)
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No treatment	FastSelect –5S/16S/23S
ATCC Gut Microbiome (12 bacteria)	96.35	12.32

85% rRNA removed

Robust depletion of rRNA from individual species

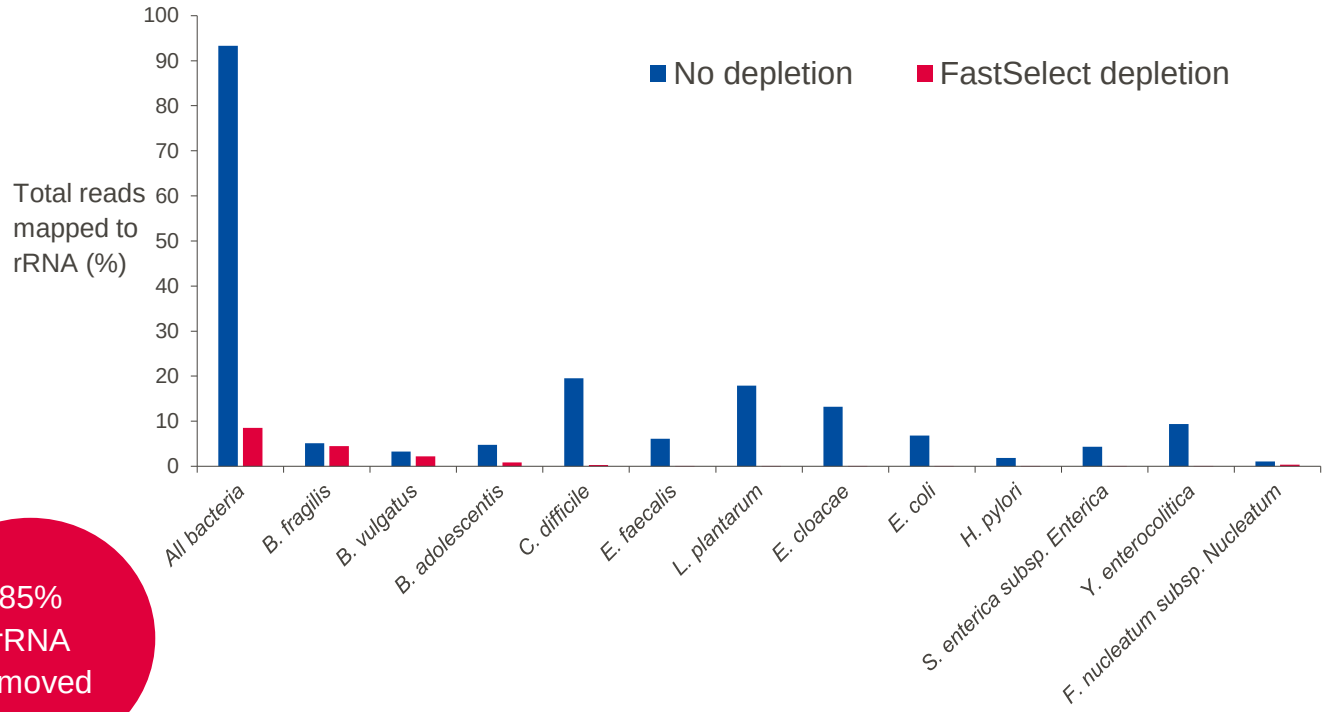


FastSelect –5S/16S/23S: Robust rRNA removal from gut bacterial communities

Experimental overview

- Sample: 100 ng, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect –5S/16S/23S
- Library prep: QIAseq Stranded Total RNA Lib Kit
- Sequencing: NextSeq 550 (2 x 75 bp)
- Mapping: CLC Genomics Workbench

Robust depletion of rRNA from individual species



FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No treatment	FastSelect –5S/16S/23S
ATCC Gut Microbiome (12 bacteria)	96.35	12.32

85% rRNA removed

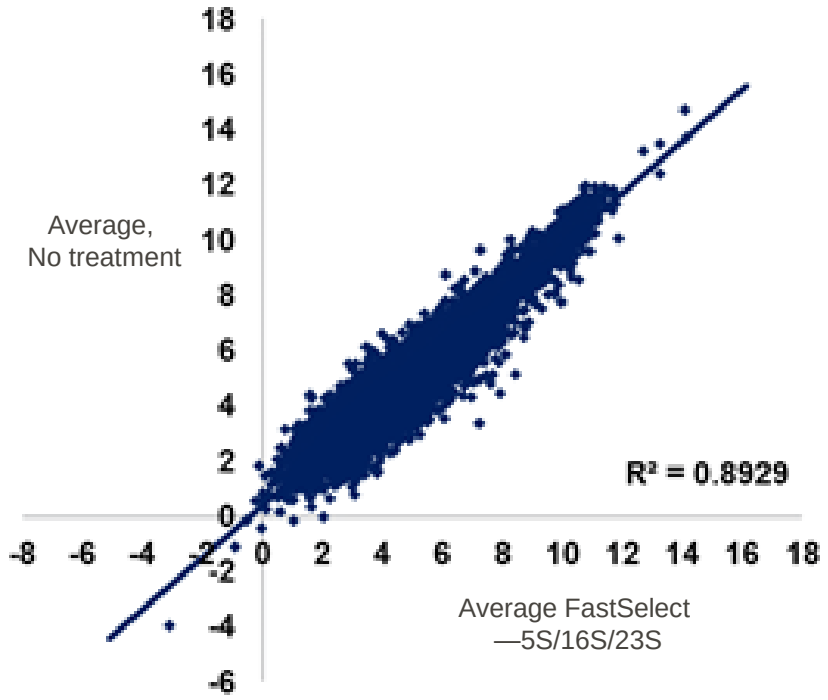
- FastSelect removes more than 85% of all rRNA
- When individual species are mapped, FastSelect removes rRNA from a broad range of species

FastSelect –5S/16S/23S does not alter gene expression in gut bacterial communities

Experimental overview

- Sample: 100 ng, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect –5S/16S/23S
- Library prep: QIAseq Stranded Total RNA Lib Kit
- Sequencing: NextSeq 550 (2 x 75 bp)
- Mapping: CLC Genomics Workbench

Strong gene expression correlation



- FastSelect removes more than 85% of all rRNA
- When individual species are mapped, FastSelect removes rRNA from a broad range of species

FastSelect –5S/16S/23S: Robust rRNA removal from skin and oral bacterial communities

Experimental overview

- Sample: 100 ng, 20 Strain Even Mix Whole Cell Material (ATCC) + Skin Microbiome Whole Cell Mix (ATCC) + Oral Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect –5S/16S/23S: Robust rRNA removal from skin and oral bacterial communities

Experimental overview

- Sample: 100 ng, 20 Strain Even Mix Whole Cell Material (ATCC) + Skin Microbiome Whole Cell Mix (ATCC) + Oral Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No Treatment	FastSelect – 5S/16S/23S
ATCC 3 Mix (28 bacteria), R1	94.81	16.97
ATCC 3 Mix (28 bacteria), R2	94.71	14.45

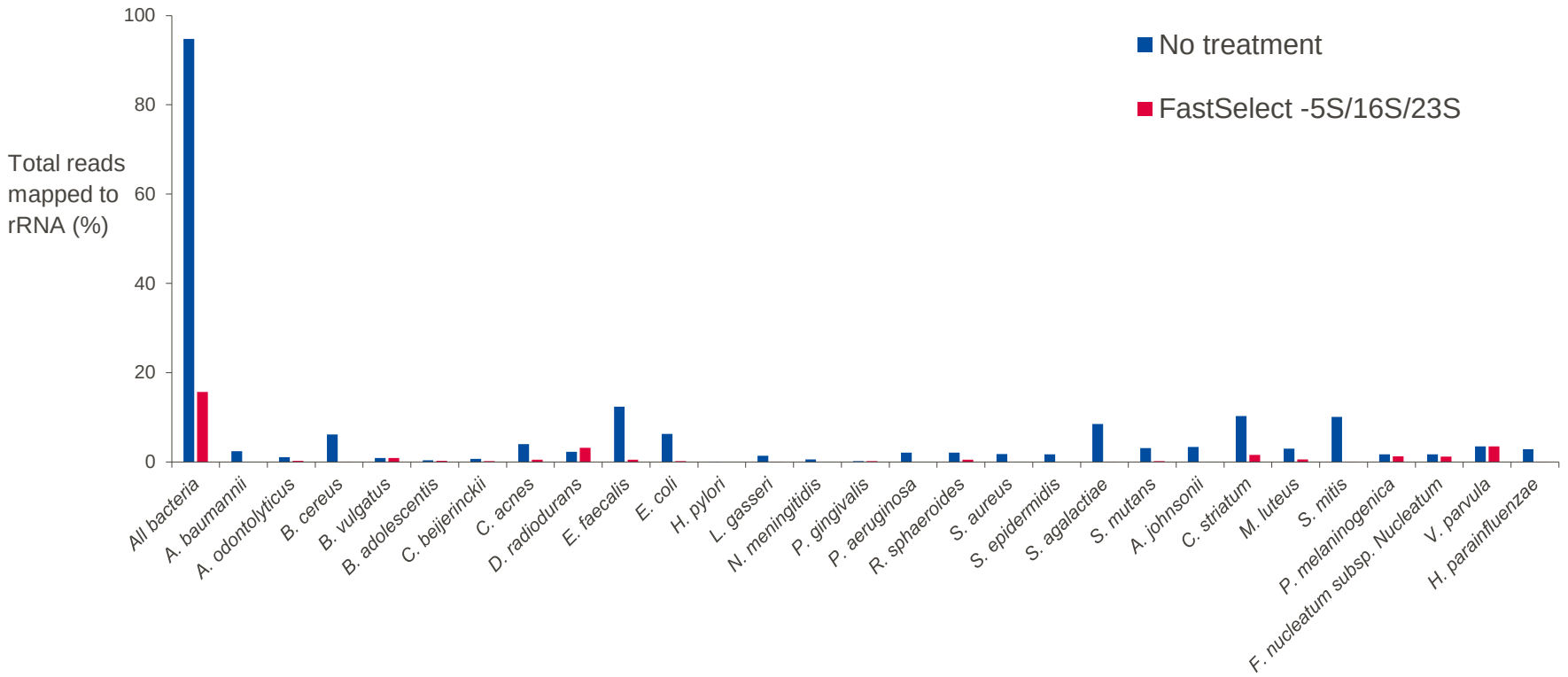
85% rRNA removed

FastSelect –5S/16S/23S: Robust rRNA removal from skin and oral bacterial communities

Experimental overview

- Sample: 100 ng, 20 Strain Even Mix Whole Cell Material (ATCC) + Skin Microbiome Whole Cell Mix (ATCC) + Oral Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect robustly depletes rRNA (individual species)

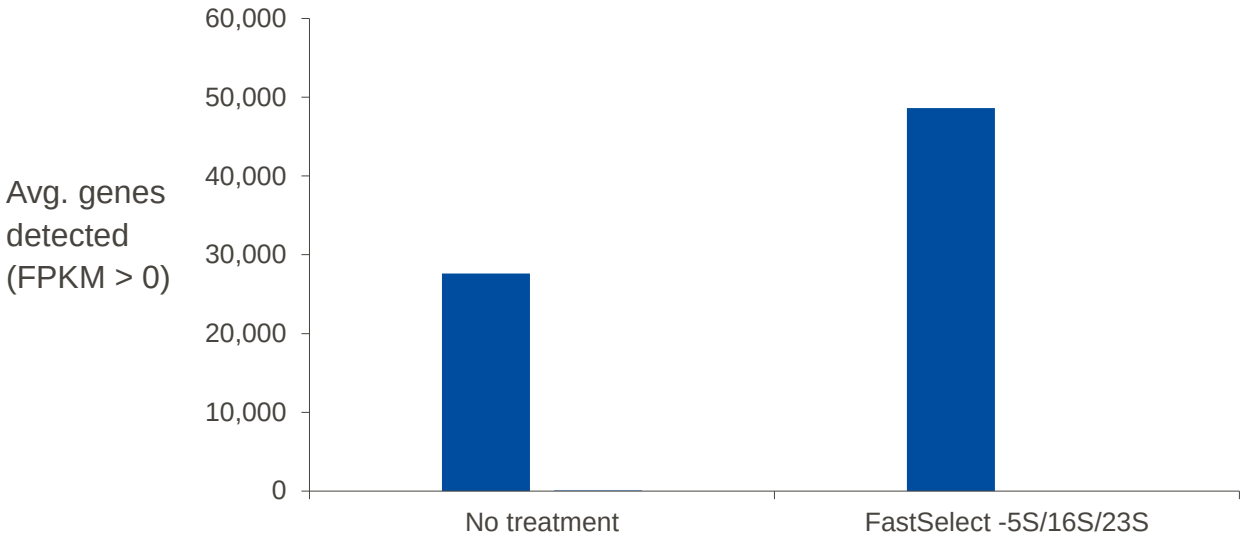


FastSelect –5S/16S/23S: Robust rRNA removal from skin and oral bacterial communities

Experimental overview

- Sample: 100 ng, 20 Strain Even Mix Whole Cell Material (ATCC) + Skin Microbiome Whole Cell Mix (ATCC) + Oral Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect increases detected genes



- FastSelect efficiently removes rRNA, freeing up substantial read budget.
- In turn, this read budget enables a dramatic increase in the number of genes detected

FastSelect –5S/16S/23S dramatically outperforms the other supplier’s kits

Experimental overview

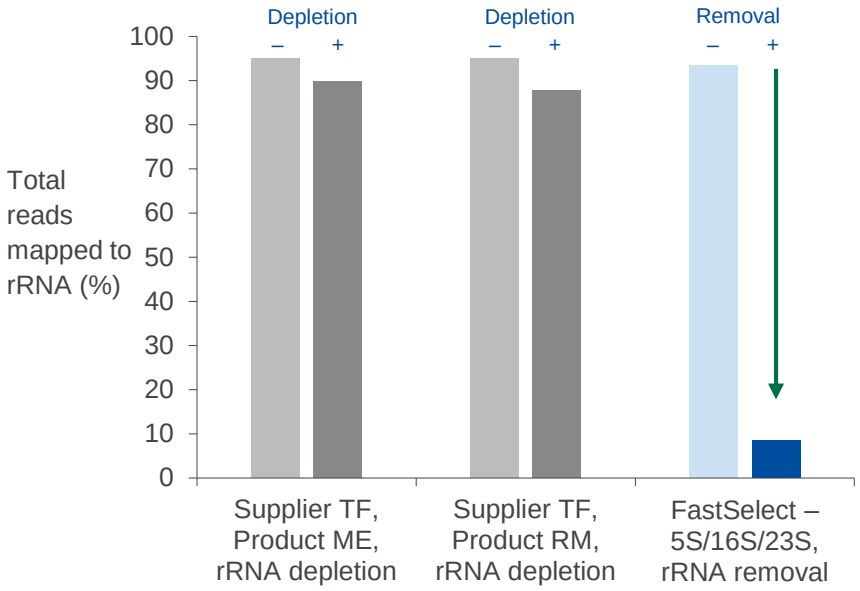
- Sample: 1 µg, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, Supplier TF – Product ME, Supplier TF – Product RM, FastSelect – 5S/16S/23S
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect –5S/16S/23S dramatically outperforms the other supplier's kits

Experimental overview

- Sample: 1 µg, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, Supplier TF – Product ME, Supplier TF – Product RM, FastSelect – 5S/16S/23S
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect robustly removes rRNA, while the other supplier's kits do not

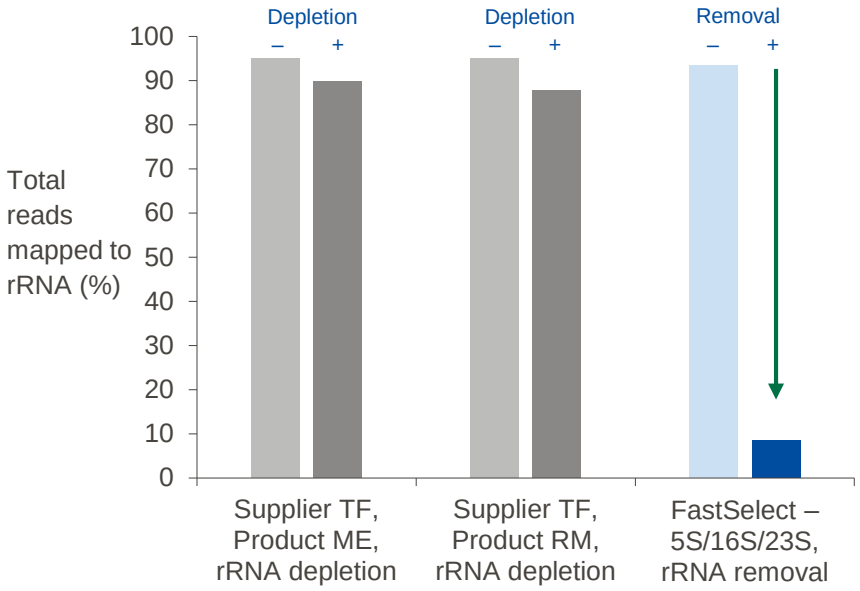


FastSelect –5S/16S/23S dramatically outperforms the other supplier's kits

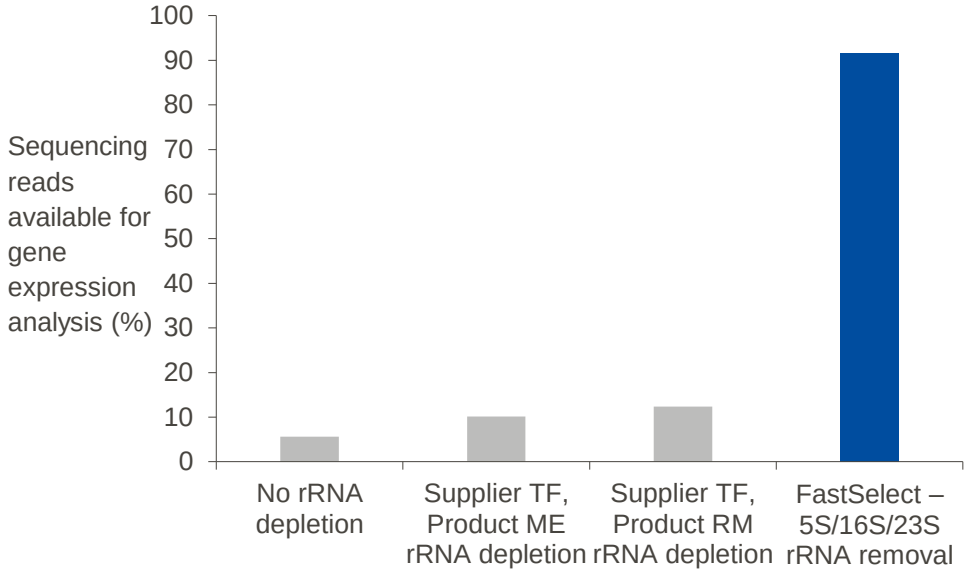
Experimental overview

- Sample: 1 µg, Gut Microbiome Whole Cell Mix (ATCC)
- Depletion: No depletion, Supplier TF – Product ME, Supplier TF – Product RM, FastSelect – 5S/16S/23S
- Library prep: QIAseq Stranded
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect robustly removes rRNA, while the other supplier's kits do not



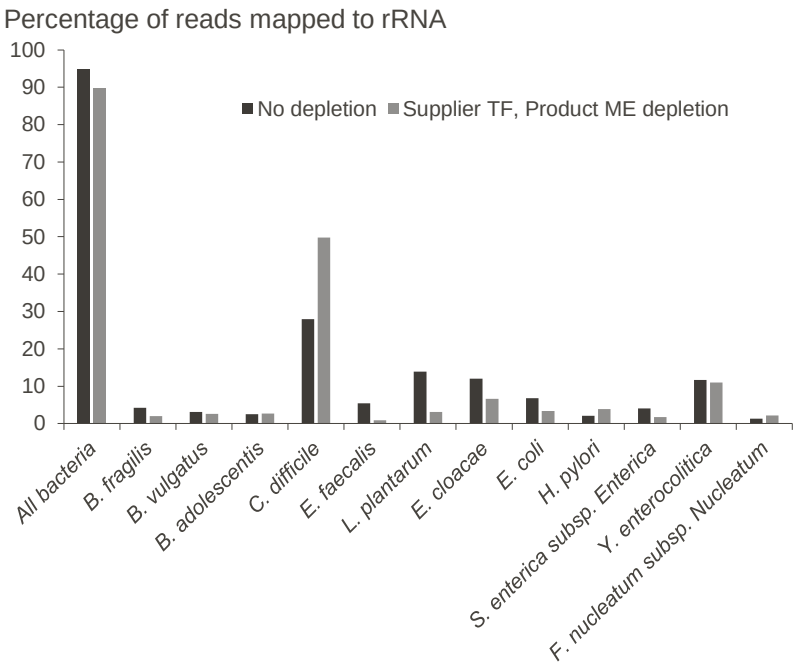
FastSelect frees up reads for gene detection



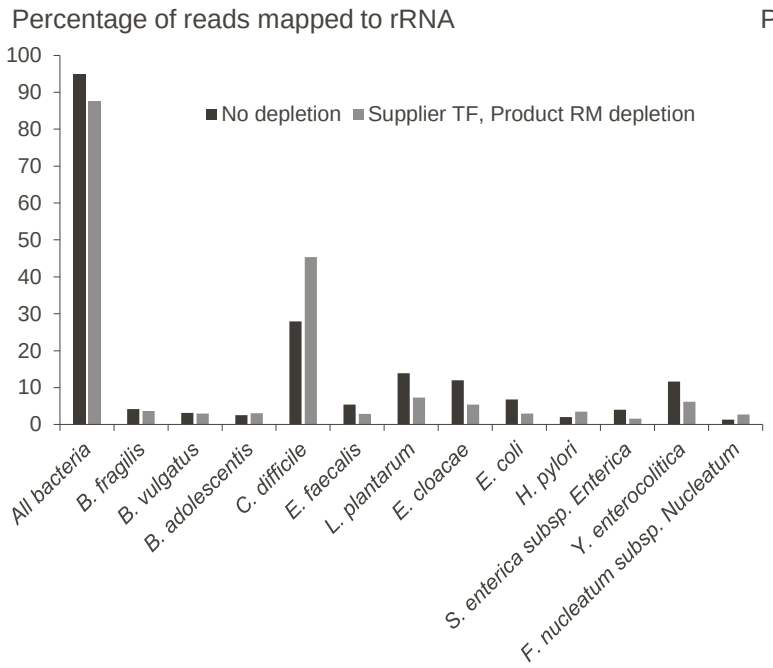
- FastSelect efficiently removes rRNA, while Supplier TF's rRNA depletion products do not
- The robust rRNA removal frees up a substantial amount of sequencing reads (9x compared to the other supplier's kits) for gene expression analysis

FastSelect –5S/16S/23S dramatically outperforms the other supplier's kits

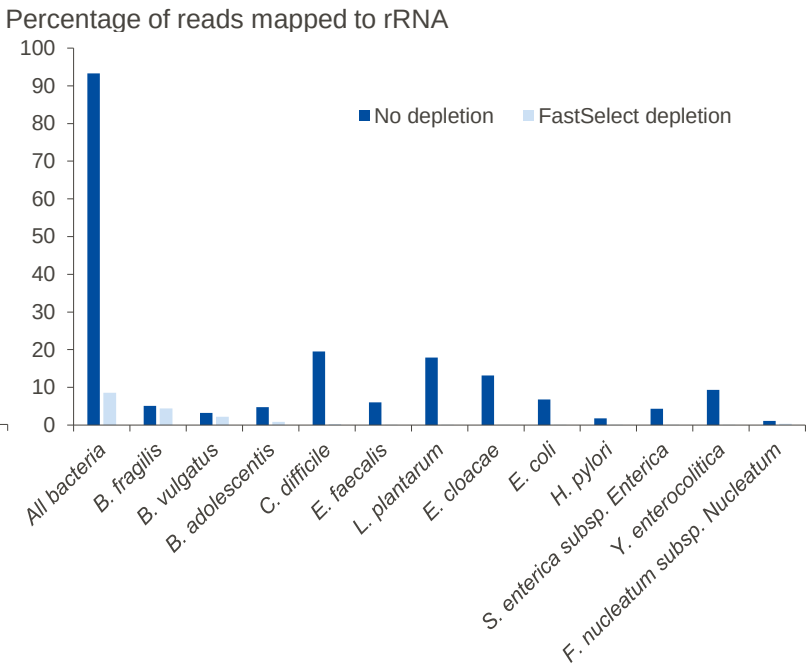
Supplier TF, Product ME



Supplier TF, Product RM



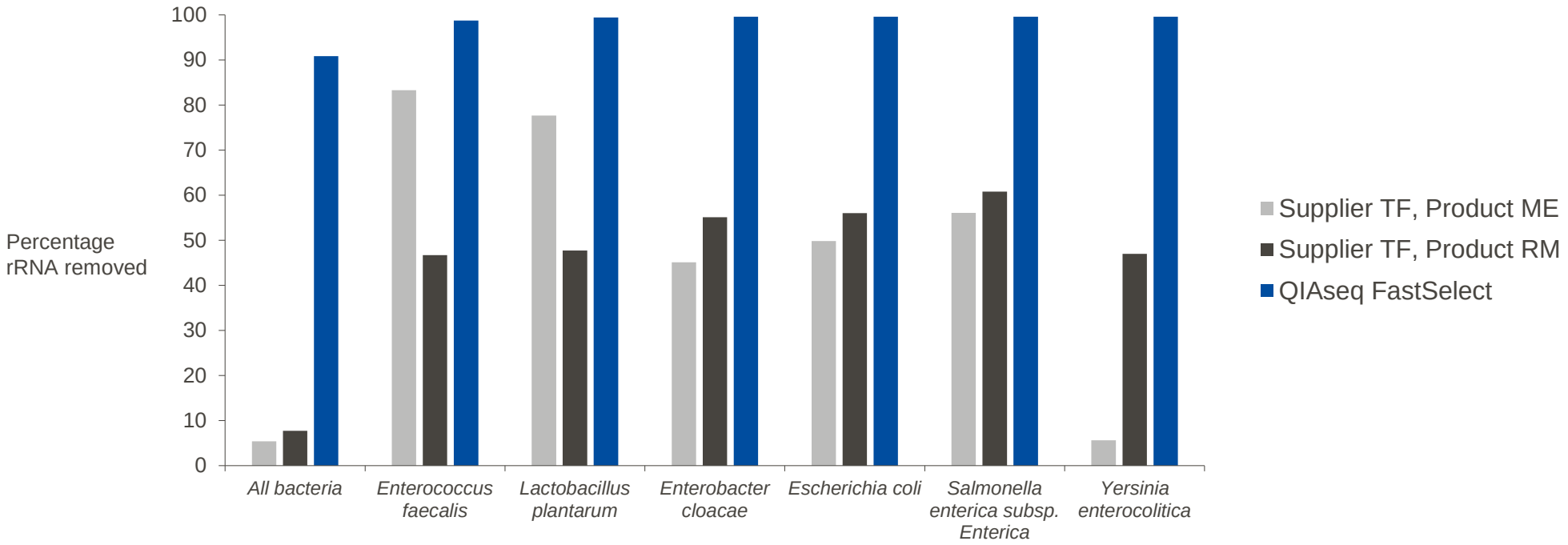
FastSelect –5S/16S/23S



- The three figures depict the percentage of rRNA removed for "all bacteria" and the individual species in the community
- FastSelect efficiently removes rRNA from a broad range of bacterial species, while Products ME and RM from Supplier TF do not

FastSelect –5S/16S/23S dramatically outperforms the other supplier's kits

FastSelect robustly removes rRNA, while the other supplier's kits do not



- When "all bacteria" are analyzed, FastSelect removes greater than 90% of all rRNA, while other supplier's kits remove less than 10%
- When select, individual bacteria are analyzed, the depletion varies for the other supplier's kits, while it remains consistently high for FastSelect

FastSelect is compatible with the QIAseq Stranded Total RNA Lib Kit

Other compatible kits:

TruSeq Stranded (Illumina)

NEBNext Ultra II Directional (NEB)

- FastSelect is compatible with most RNA library prep kits.

FastSelect –5S/16S/23S: Robust, reproducible results

Experimental overview

- Sample: 1 µg and 100 ng, Gut Microbiome Whole Cell Mix
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded, TruSeq Stranded, NEBNext Ultra II Directional
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect –5S/16S/23S: Robust, reproducible results

Experimental overview

- Sample: 1 µg and 100 ng, Gut Microbiome Whole Cell Mix
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded, TruSeq Stranded, NEBNext Ultra II Directional
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

FastSelect substantially removes rRNA, regardless of the RNA library prep kit used

QIAseq Stranded

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No treatment	FastSelect –5S/16S/23S
1 µg	93.28	8.54
100 ng	92.46	11.94

TruSeq Stranded

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No treatment	FastSelect –5S/16S/23S
1 µg	93.44	19.42
100 ng	92.18	24.11

NEBNext Ultra II Directional

Sample	Percentage of reads mapped to bacterial rRNA (total)	
	No treatment	FastSelect –5S/16S/23S
1 µg	93.17	8.55
100 ng	92.17	8.83

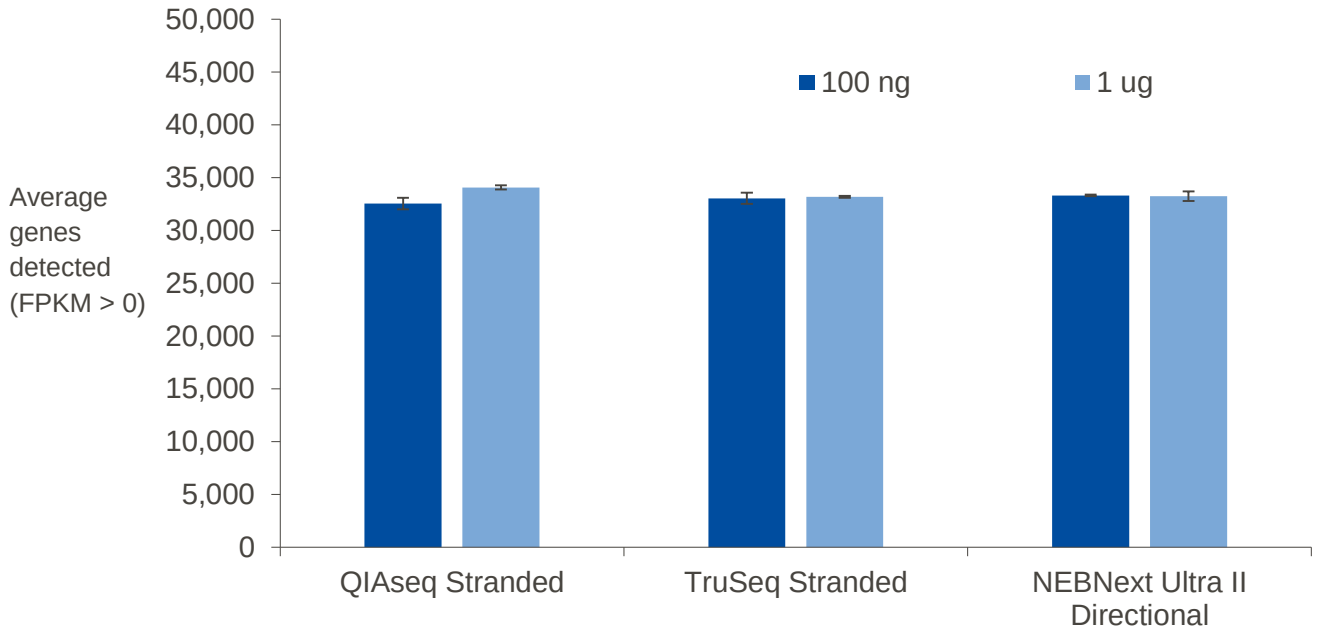
FastSelect –5S/16S/23S is compatible with QIAGEN, Ilumina and NEB Kits.

FastSelect –5S/16S/23S: Robust, reproducible results

Experimental overview

- Sample: 1 µg and 100 ng, Gut Microbiome Whole Cell Mix
- Depletion: No depletion, FastSelect
- Library prep: QIAseq Stranded, TruSeq Stranded, NEBNext Ultra II Directional
- Sequencing: NextSeq 550
- Mapping: CLC Genomics Workbench

Genes detected with each stranded kit



Agenda

Background

FastSelect –5S/16S/23S: An introduction

FastSelect –5S/16S/23S: Results

Summary

FastSelect –5S/16S/23S overview

What is FastSelect –5S/16S/23S? Fragmentation and pan-bacterial (5S/16S/23S) rRNA depletion module

- Number of reactions: 24, 96 and 384

How does it work? Inhibits reverse transcription of its specific targets

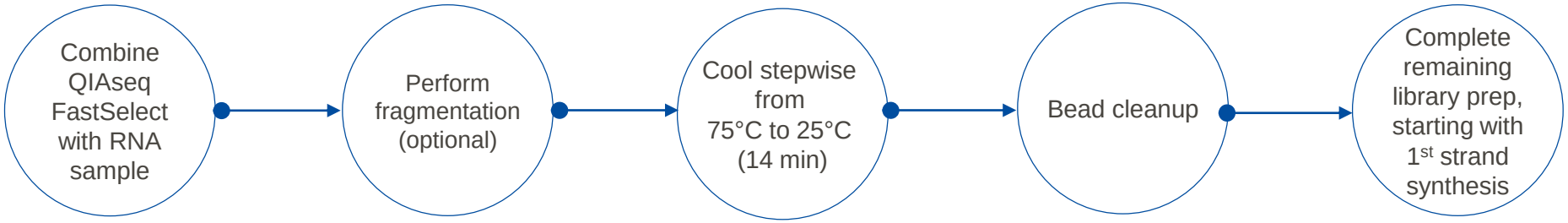
Coverage: Designed to block community level cDNA synthesis of 5S, 16S and 23S rRNA

- Designed against SILVA 16S sequences (nearly 600,000 unique entries), SILVA 23S sequences (nearly unique 170,000 entries) and 5S rRNA Database (over 7,200 unique entries)
- Theoretically blocks >95% cDNA synthesis of all 5S, 16S and 23S rRNA sequences

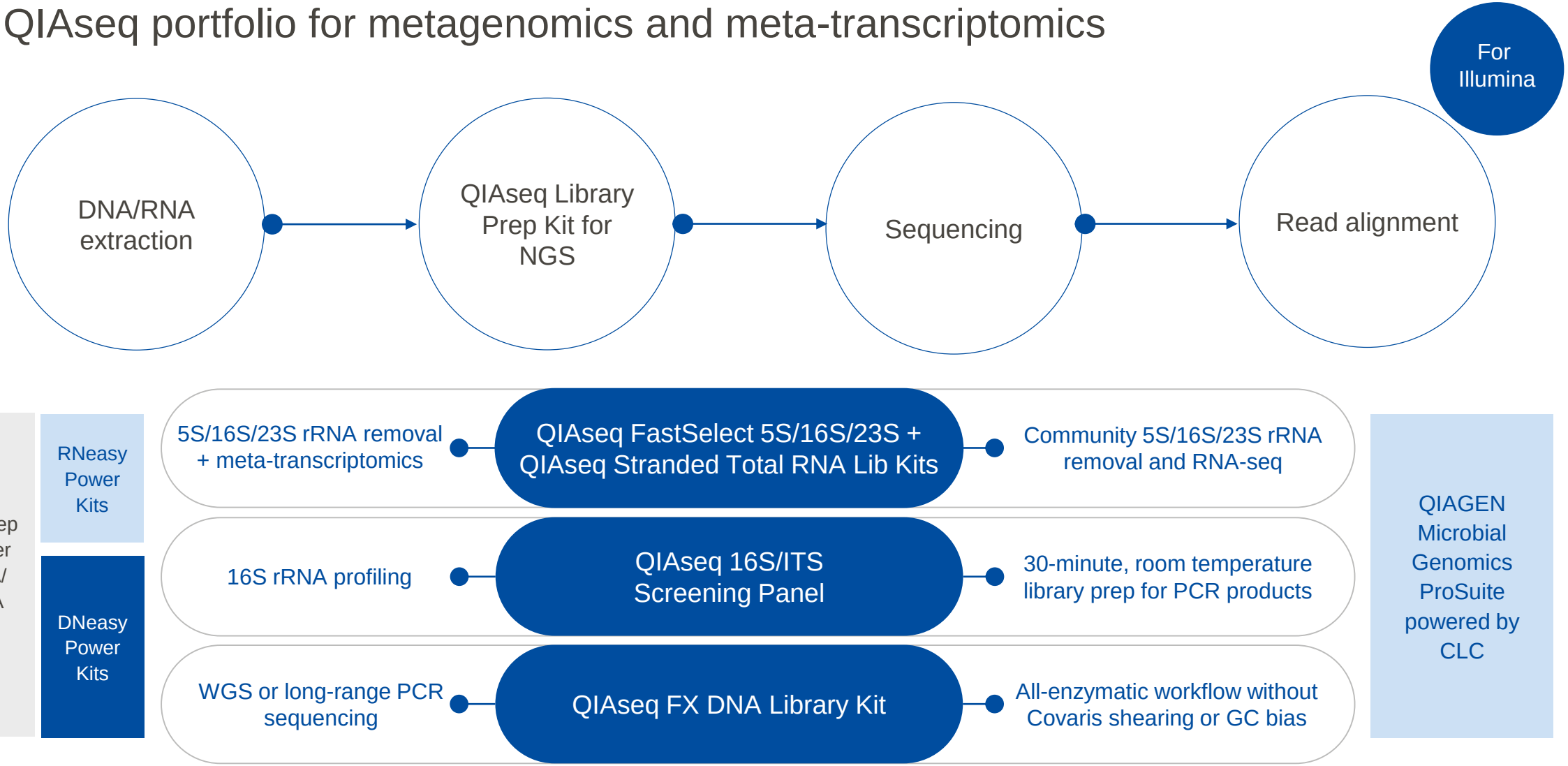
Total RNA input: 20 ng to 1 µg

Kit compatibility: QIAseq Stranded Total RNA Lib Kit (QIAGEN Group), TruSeq Stranded (Illumina, Inc), NEBNext Ultra II Directional (New England Biolabs, Inc)

- FastSelect is compatible with most RNA library prep kits



Full QIAseq portfolio for metagenomics and meta-transcriptomics



Thank you for attending.

Questions?

Samuel Rulli, Ph.D.
Samuel.Rulli@qiagen.com