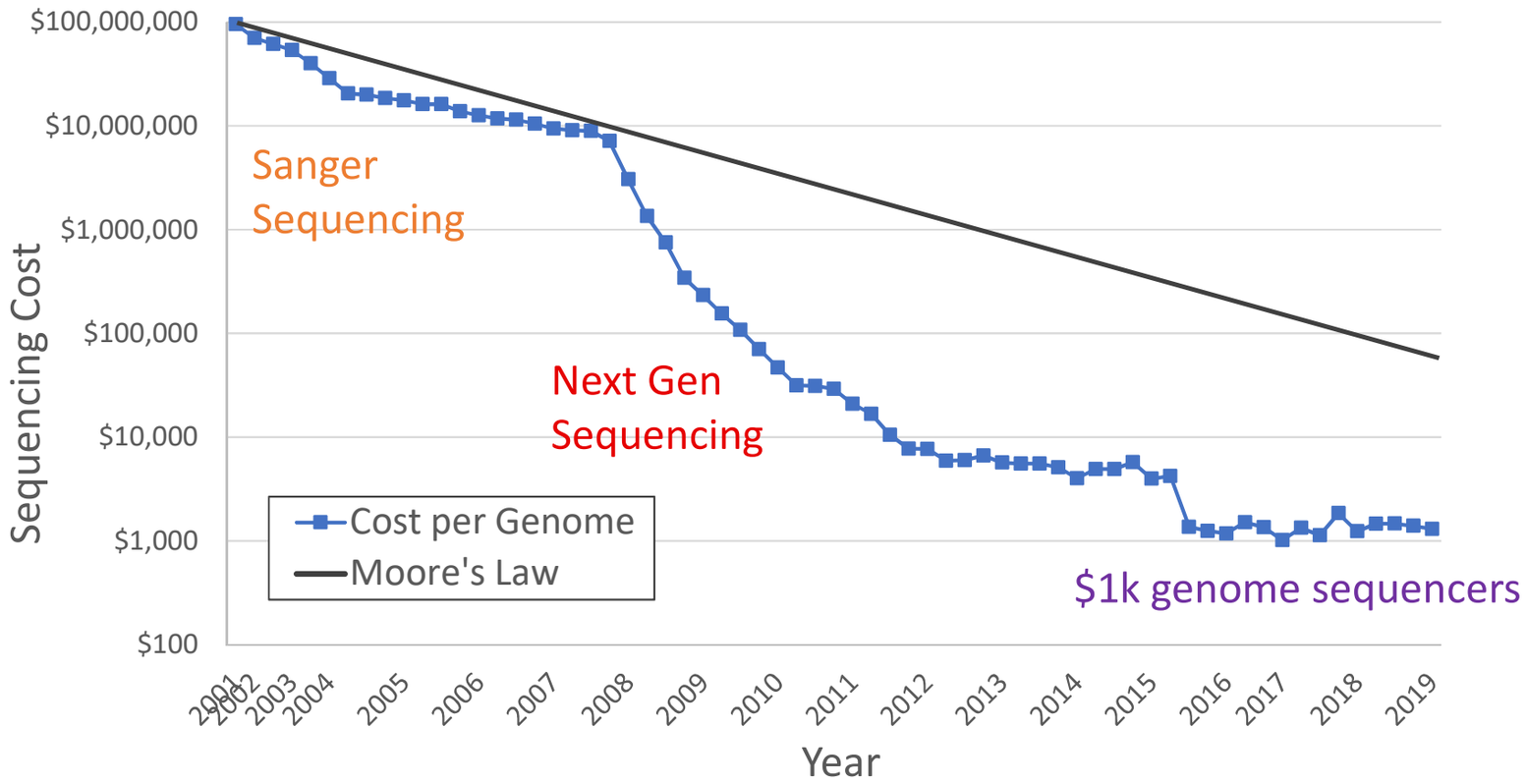
LoopSeq™ Synthetic Long Read Sequencing and Its Applications:

From Microbiome to Transcriptome and Beyond

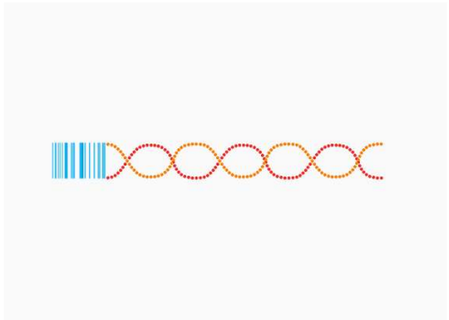
Indira Wu, PhD

The Era of Next Gen Sequencing



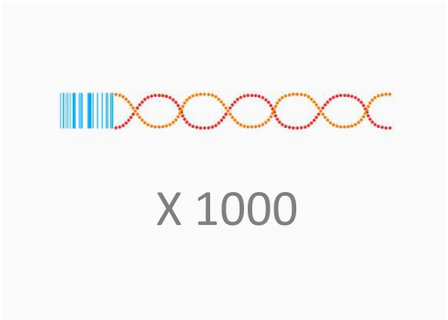
LoopSeq Single-molecule Long Read Sequencing

Attach



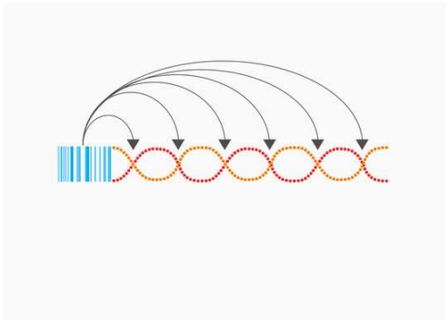
A single unique barcode attaches per strand of DNA

Amplify



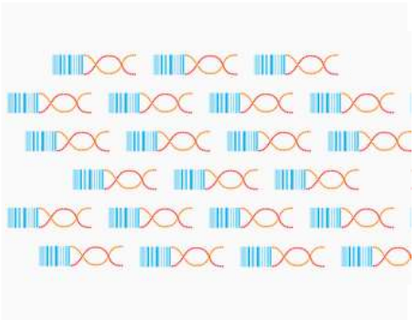
Each molecule with unique barcode is amplified using PCR

Distribute



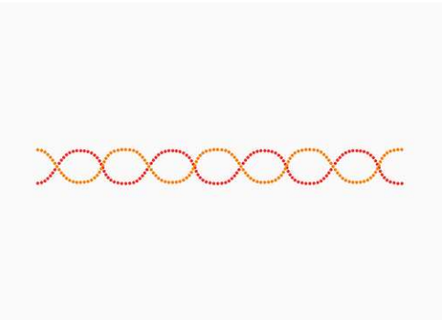
Each molecular copy has its barcode copied and pasted randomly within the molecule

Sequence



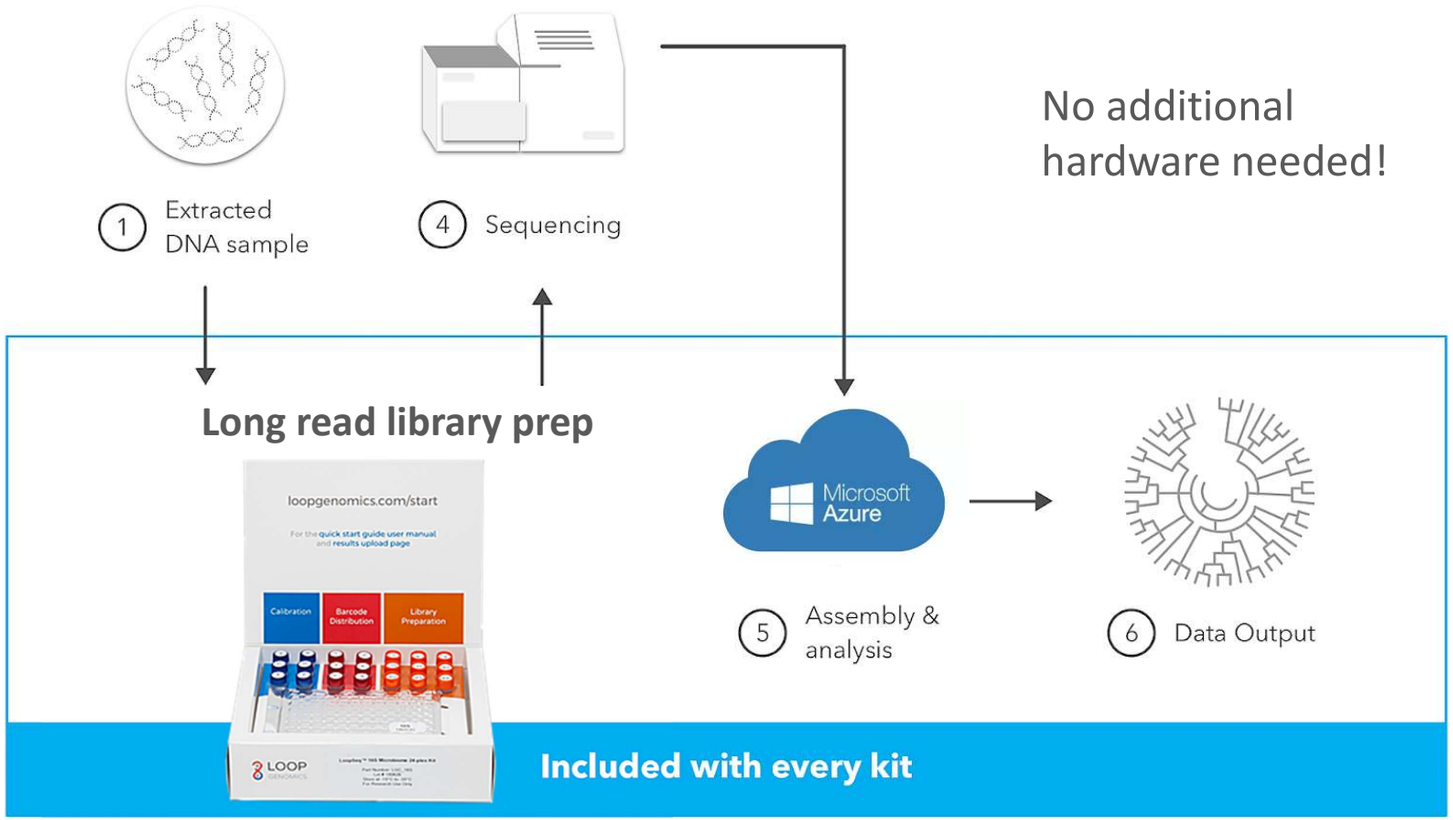
Short reads with barcode and the adjacent segment are sequenced

Assemble

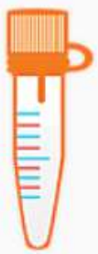


Short reads that share the same barcode are combined into full-length long reads

End-to-end Long Read Sequencing Solution



High Resolution, Quantitative Long Read Sequencing



Multiplex Workflow

Pool all samples into a single tube

Reduced Hands-on Time

Full-length Target

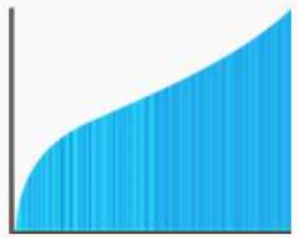
Read continuous sequence from your target

Longer Sequence Length

Identification & Classification

Higher resolution target identification and organism classification

Higher Data Resolution



True Quantification

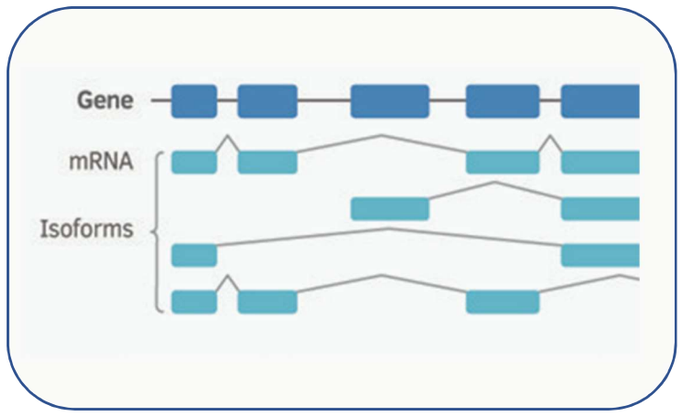
UMI based quantification with no PCR bias

Better Sample Insight

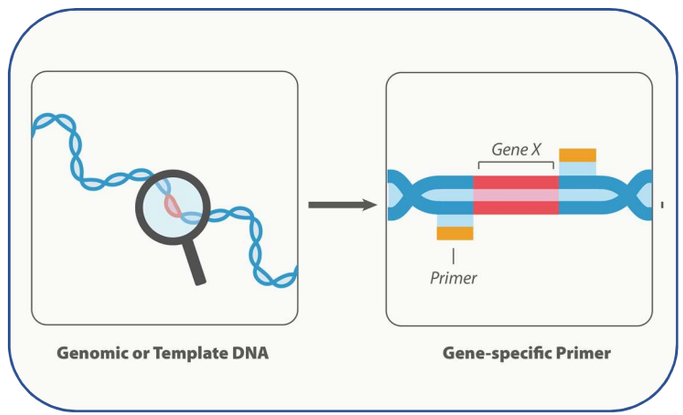
LoopSeq Long Read Applications

Micro/Mycobiome

RNA

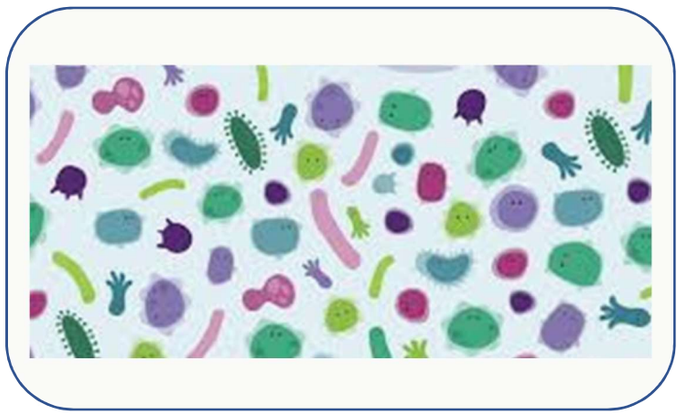


DNA



LoopSeq Long Read Applications

Micro/Mycobiome



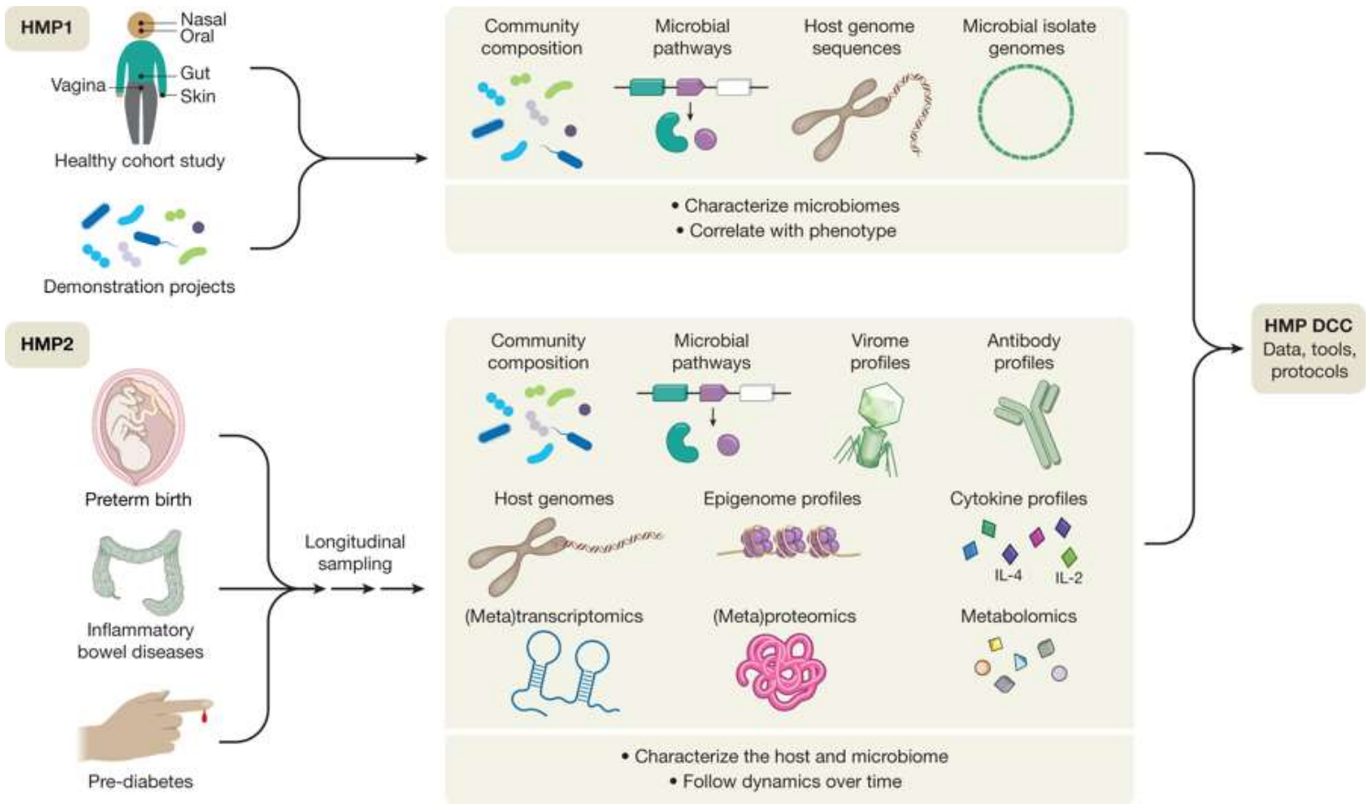
RNA

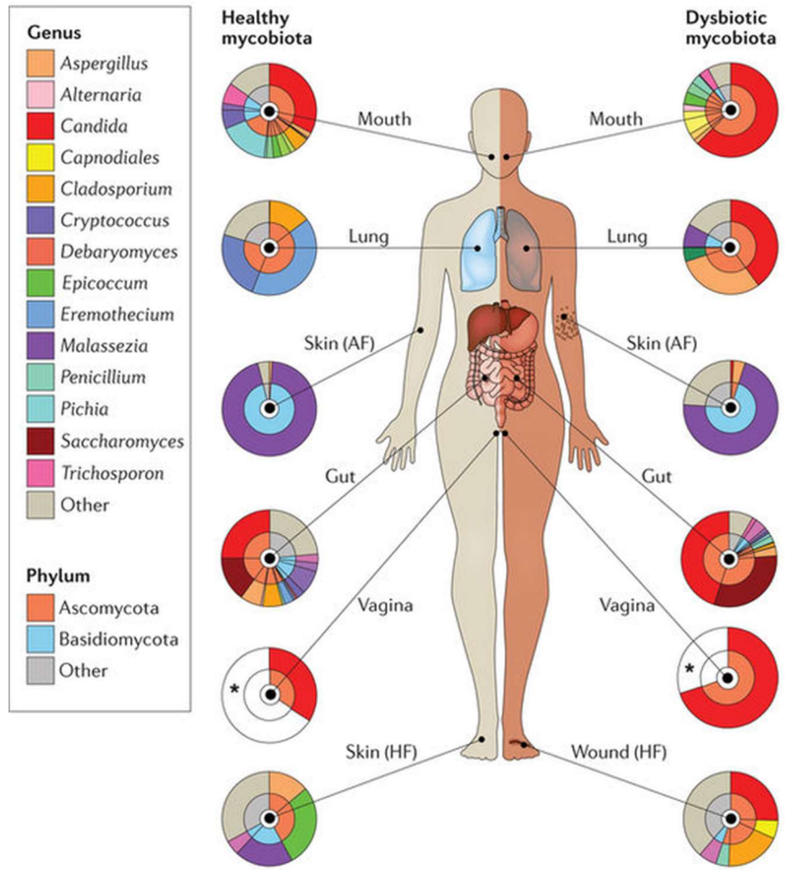
DNA

Single-molecule counting, long read 16S, 18S, or 18S-ITS sequencing

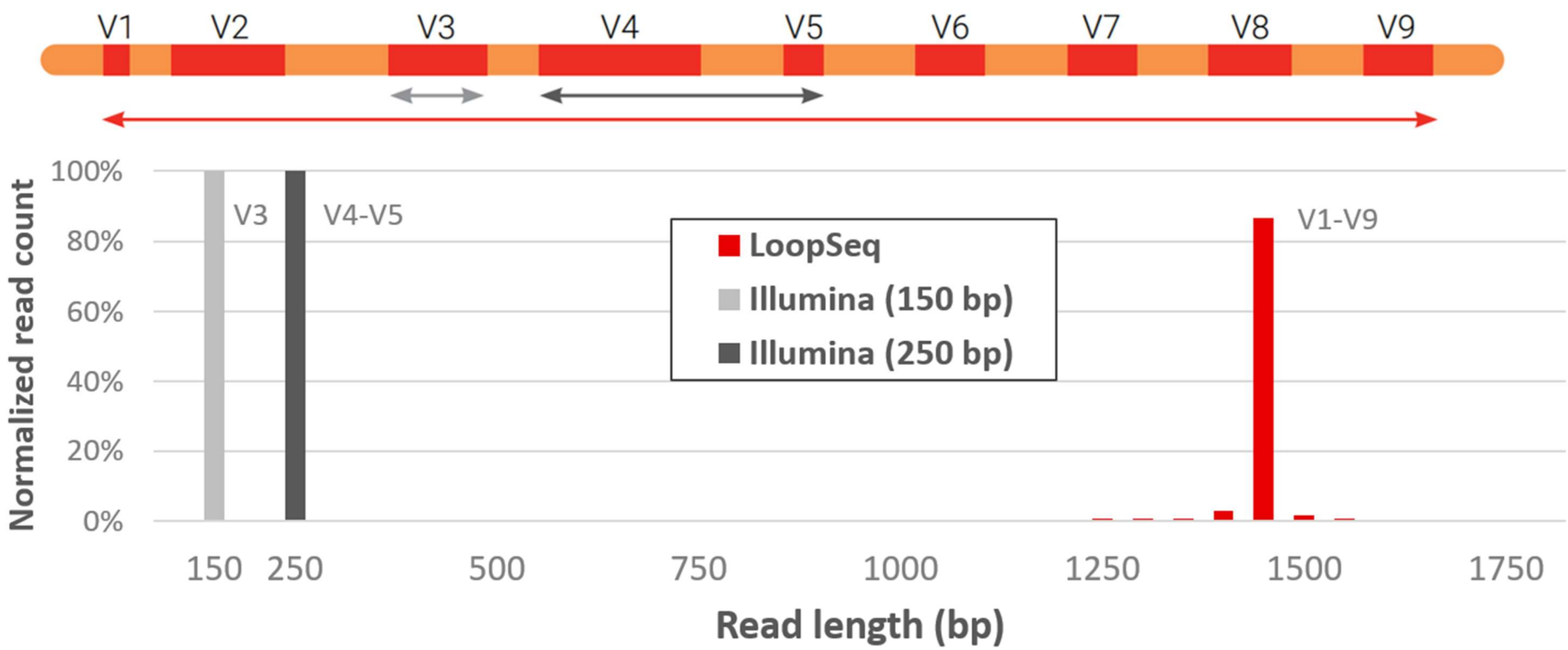
Linking Human Microbiome to Diseases

The Integrative Human Microbiome Project

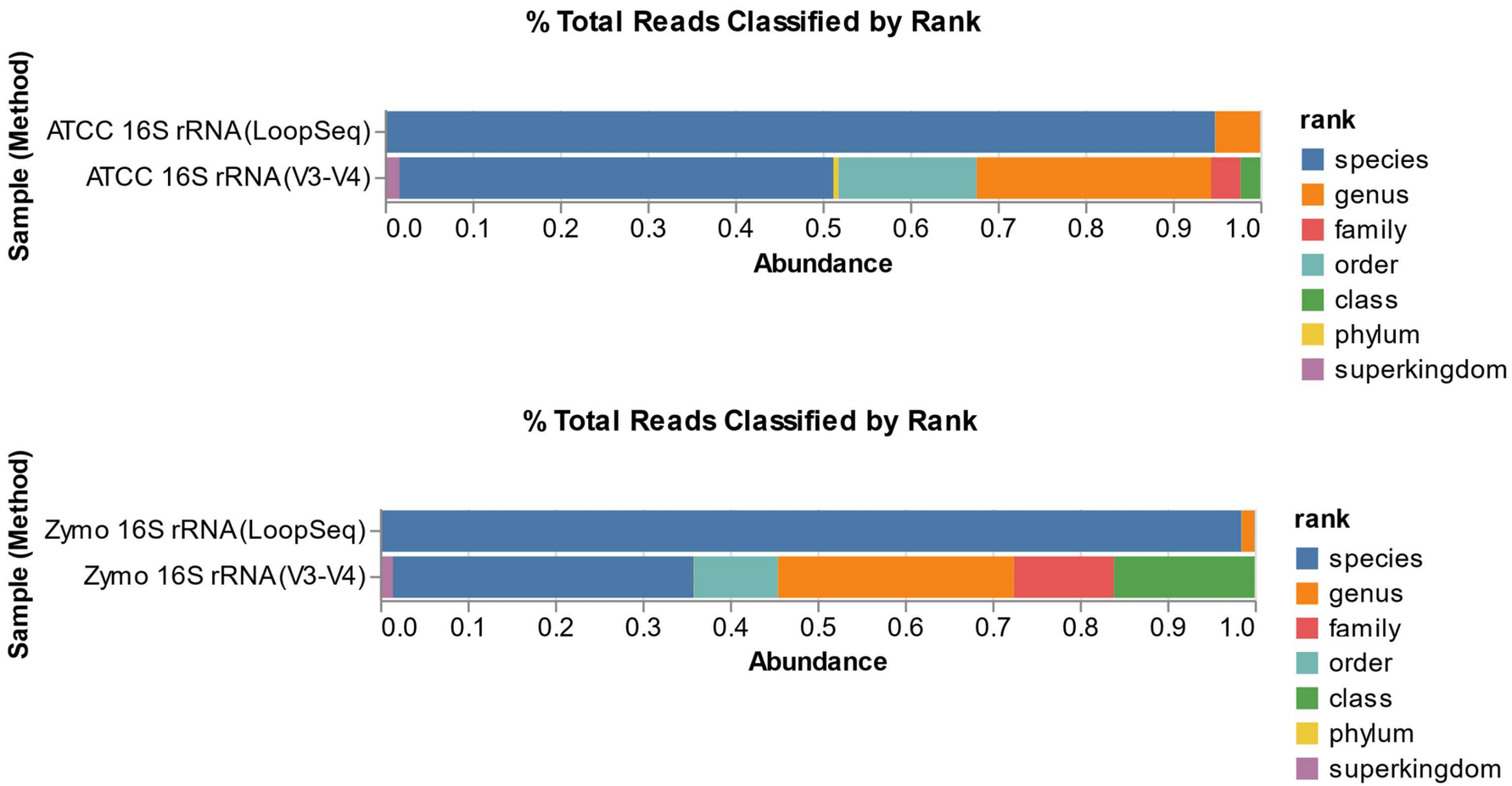




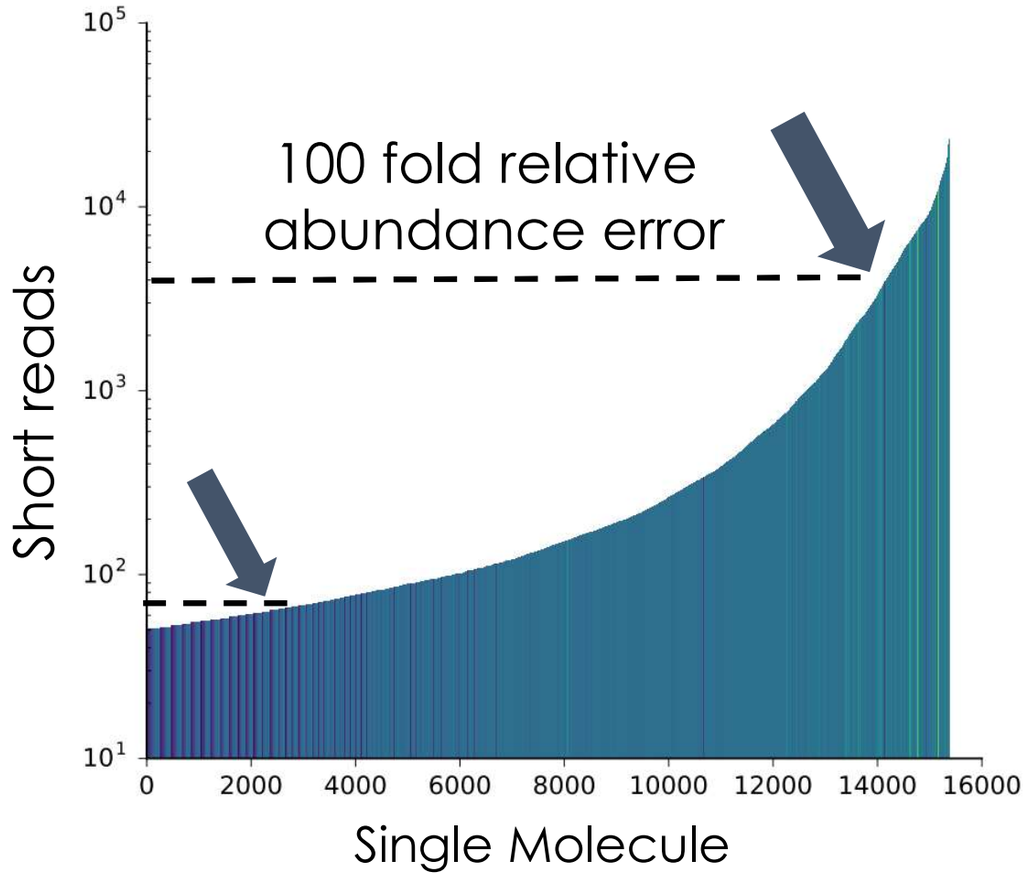
LoopSeq vs. Short Read Microbiome Sequencing



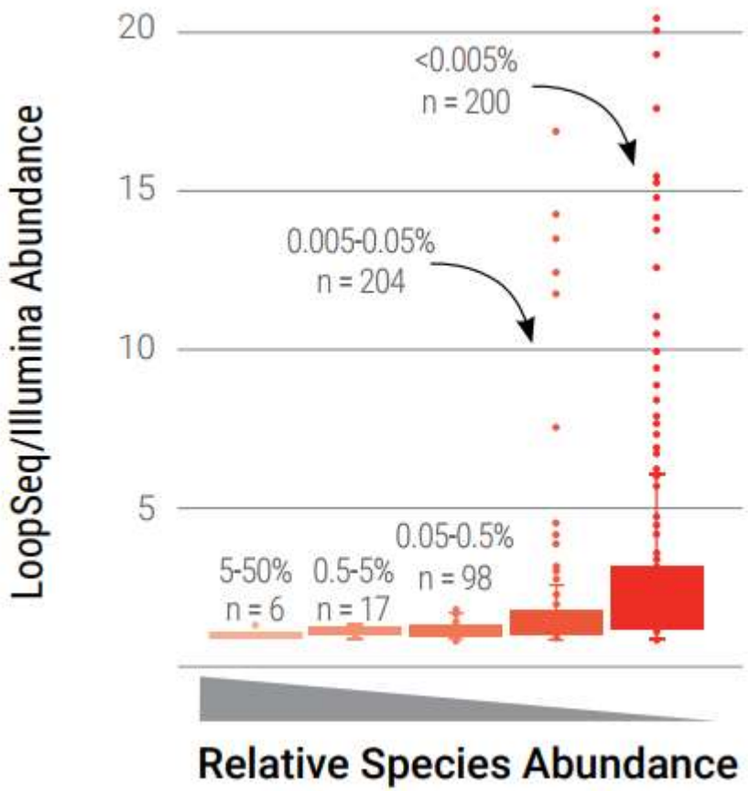
LoopSeq High Resolution Microbiome Classification

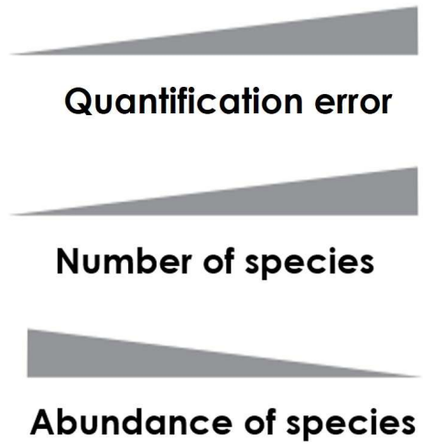


No Sample Preparation Bias

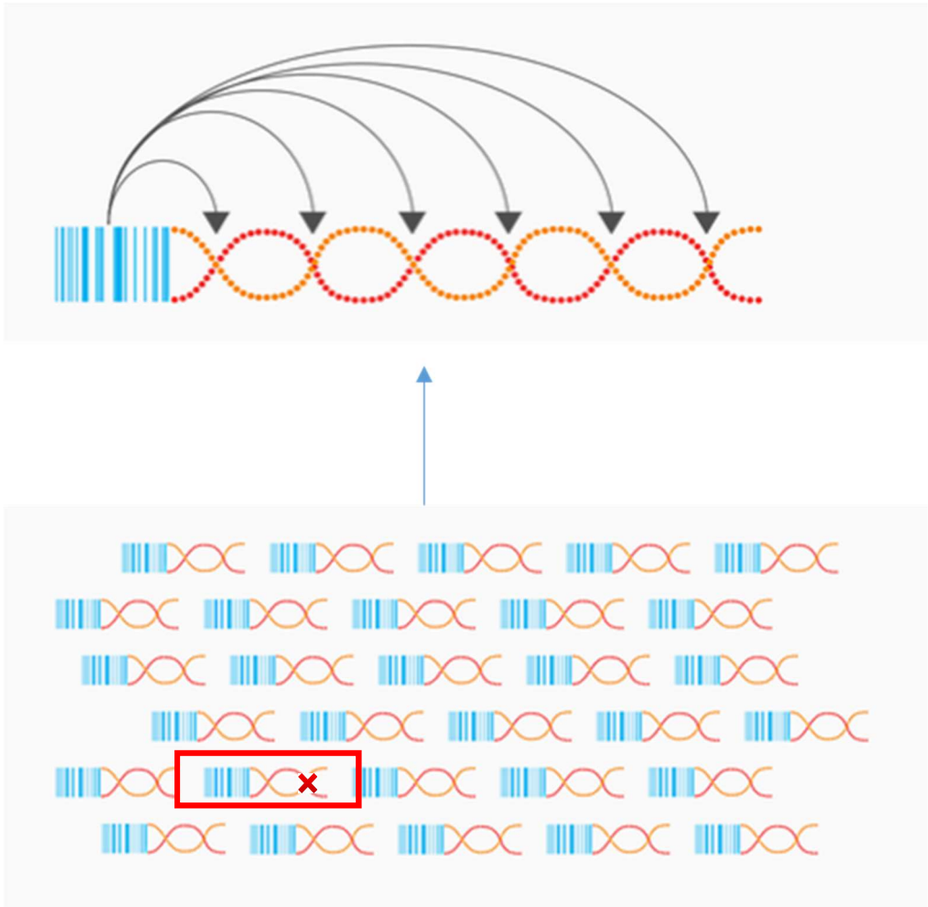


Microbiome Profiling Without Bias





Accurate Reading of the Microbiome



SEQUENCING ERRORS PER MILLION BP

43	1639
LoopSeq Complete	Illumina

Sequencing Errors Confound Analysis

LoopSeq™ 16S

Name	Readcount (% of classified reads)
Bacillus subtilis	7467 (19.78%)
Listeria monocytogenes	5882 (15.58%)
Staphylococcus aureus	5120 (13.56%)
Salmonella enterica	4390 (11.63%)
Lactobacillus fermentum	4297 (11.38%)
Escherichia coli	4262 (11.29%)
Pseudomonas aeruginosa	1683 (4.46%)
Enterococcus faecalis	826 (2.19%)

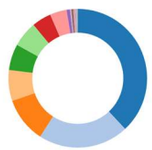
0

false positives

V3-V4 16S

Name	Readcount (% of classified reads)
Bacillus subtilis	805716 (14.03%)
Enterococcus faecalis	447102 (7.79%)
Pseudomonas aeruginosa	224488 (3.91%)
Lactobacillus fermentum	152620 (2.66%)
Salmonella enterica	128760 (2.24%)
Escherichia coli	126806 (2.21%)
Staphylococcus aureus	95496 (1.66%)
Listeria monocytogenes	82992 (1.45%)
Bacillus licheniformis	16722 (0.29%)
Staphylococcus warneri	7558 (0.13%)
Enterococcus faecium	7336 (0.13%)
Staphylococcus epidermidis	5492 (0.1%)
Carnobacterium maltaromaticum	3144 (0.05%)
Staphylococcus pseudintermedius	1954 (0.03%)
Escherichia marmotae	1892 (0.03%)
(Remaining)	10931 (0.19%)

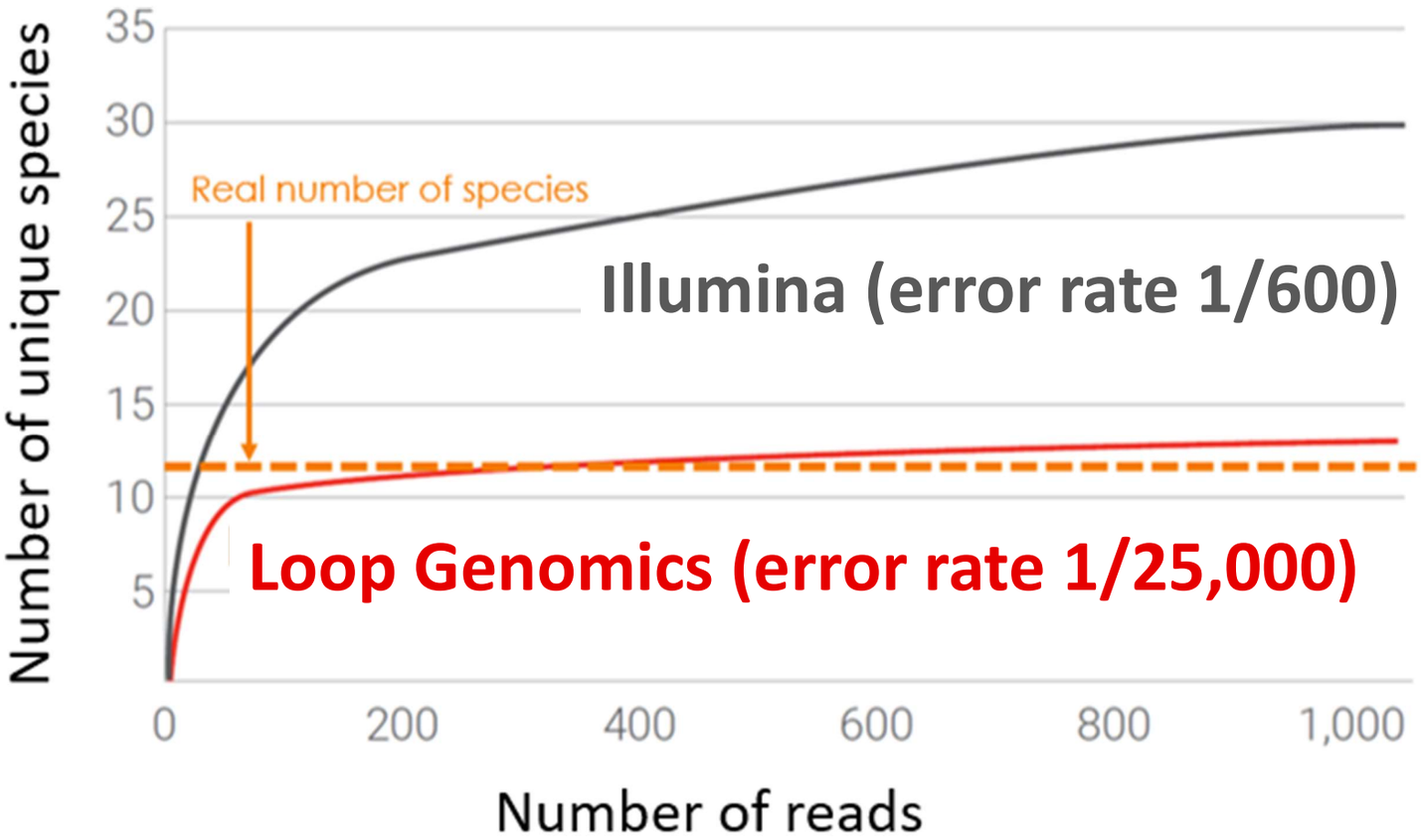
show more / less



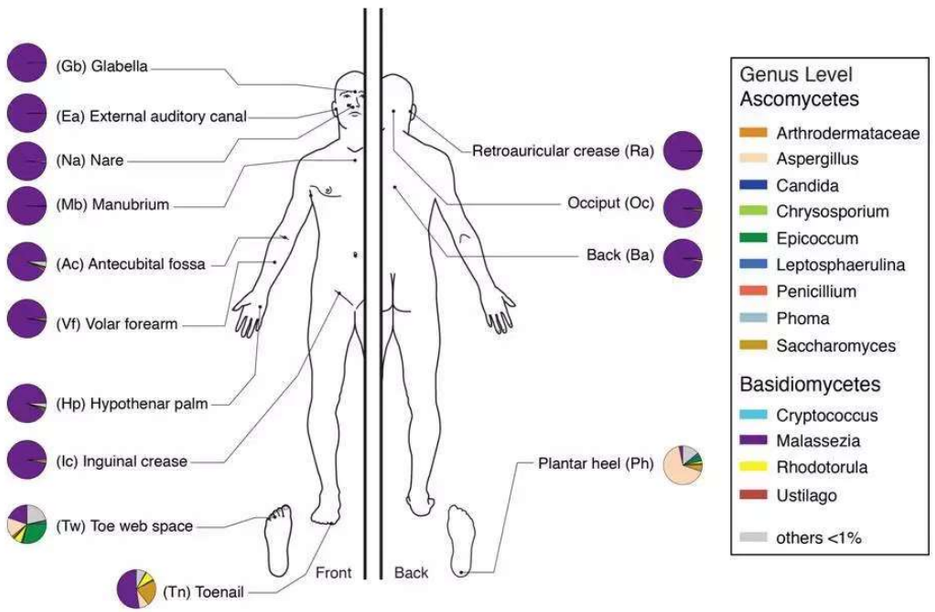
23

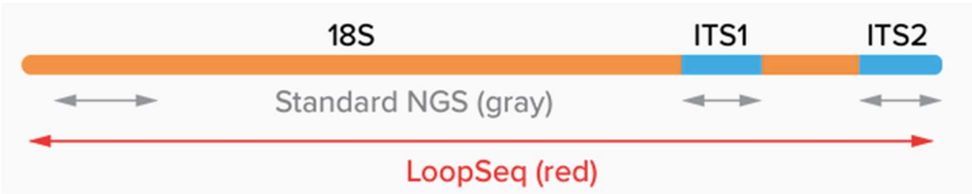
false positives

Reduce False Species Identification



LoopSeq High Resolution Mycobiome Sequencing





A single 2500bp, phased long-read with three markers

Broad range, high resolution classification

Supercharge discovery with contigs queried against three major databases.

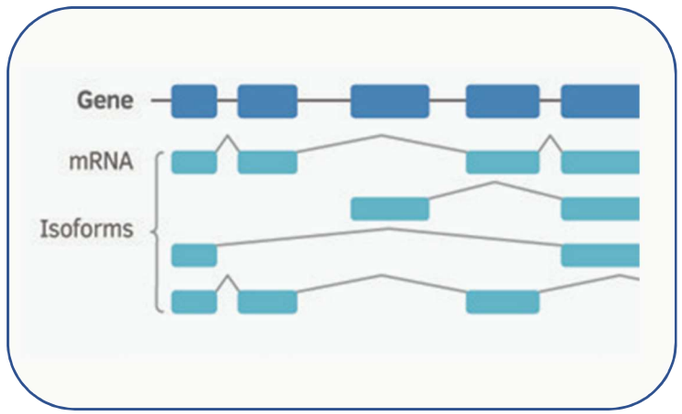
EnsemblFungi · unite community · silva high quality ribosomal RNA databases

LoopSeq Long Read Applications

Micro/Mycobiome

RNA

DNA

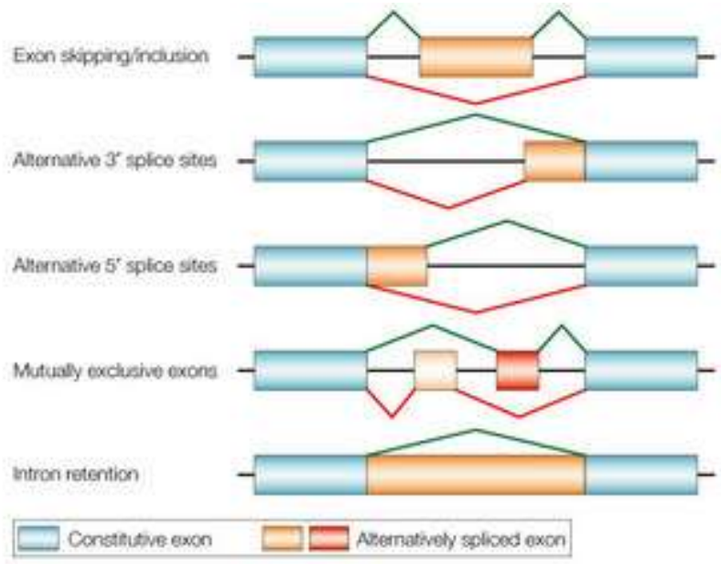


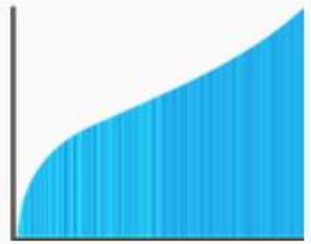
Single-molecule transcript counting, long read mRNA sequencing

LoopSeq Transcriptome Sequencing

Multiplex workflow

Pool all samples into a single tube.





True quantification

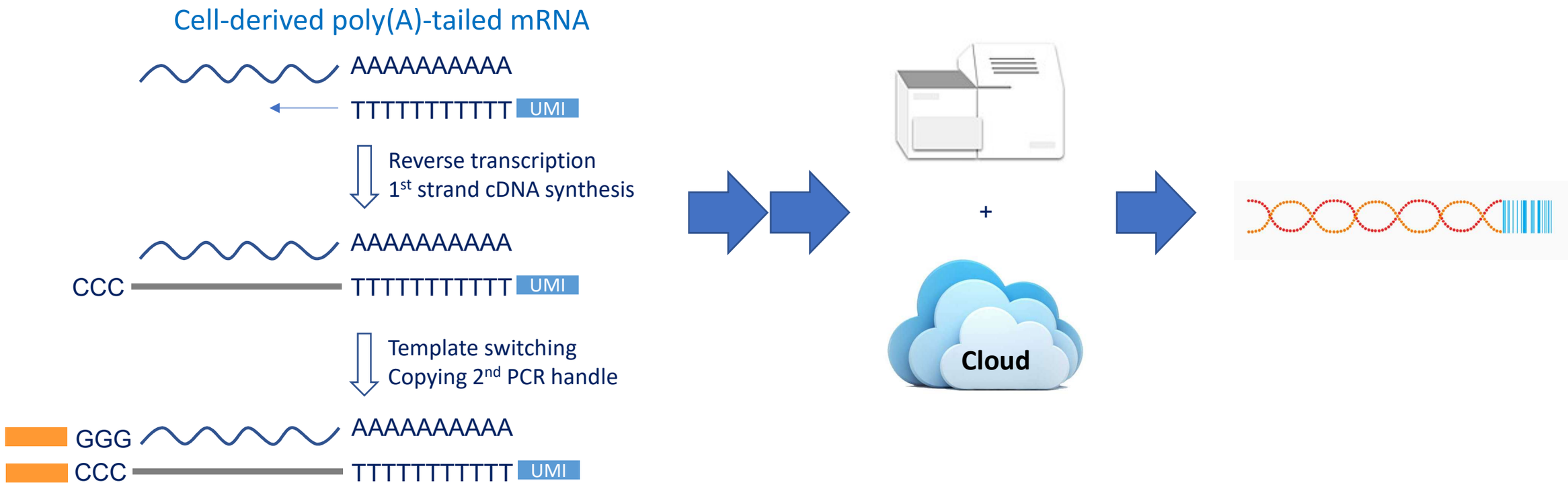
Barcode based quantification with no PCR bias.

Reduced Hands-on Time

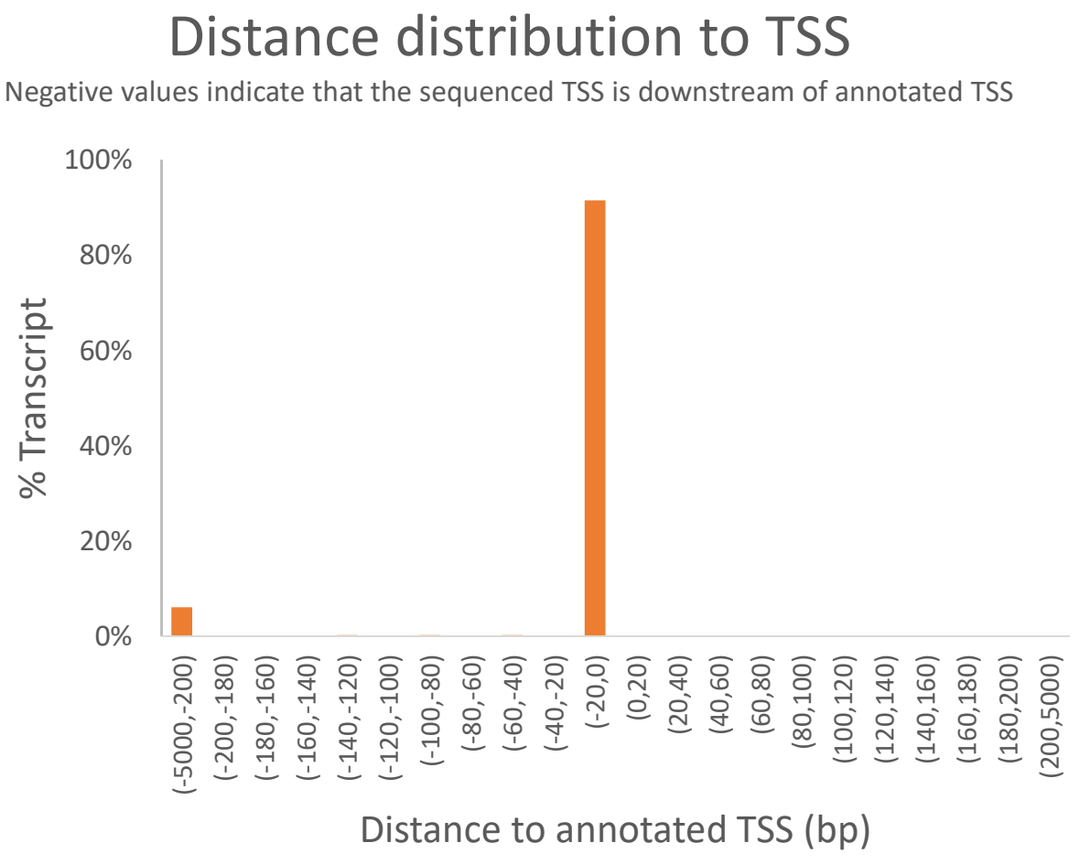
Isoform Mapping

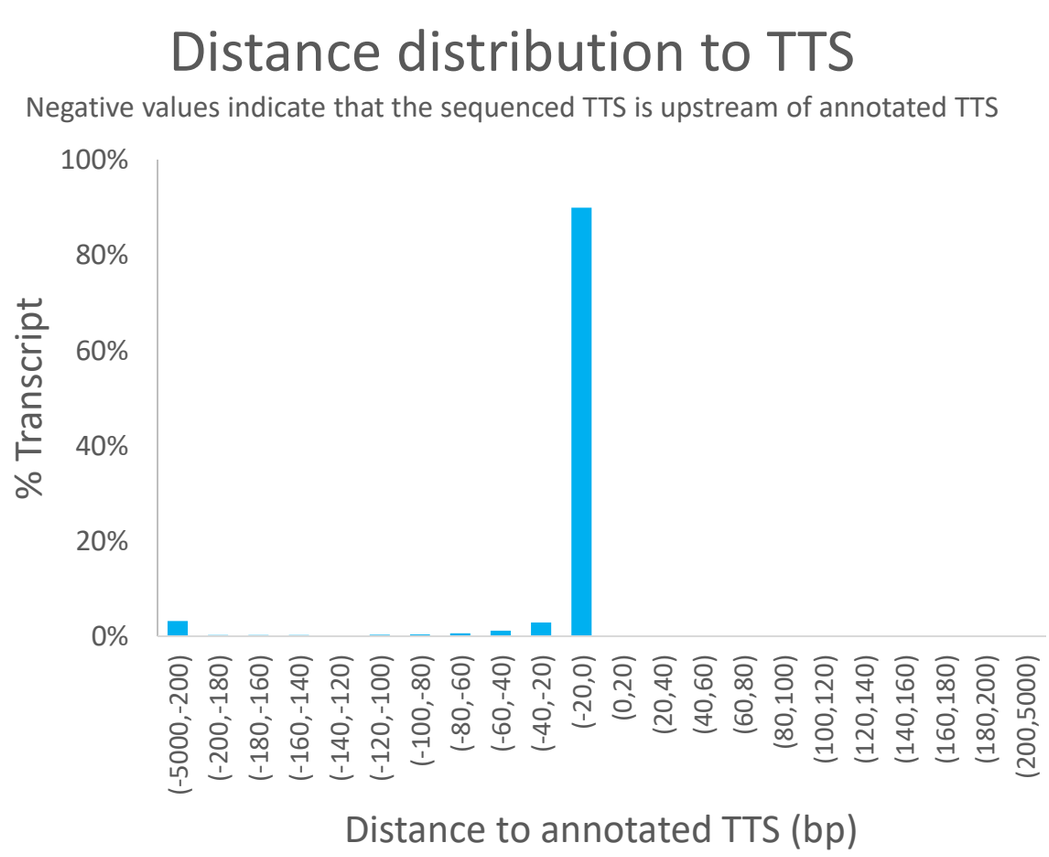
UMI-based Transcript Counting

LoopSeq UMI-tagged Library Prep



Full-length Coverage of Synthetic Standard

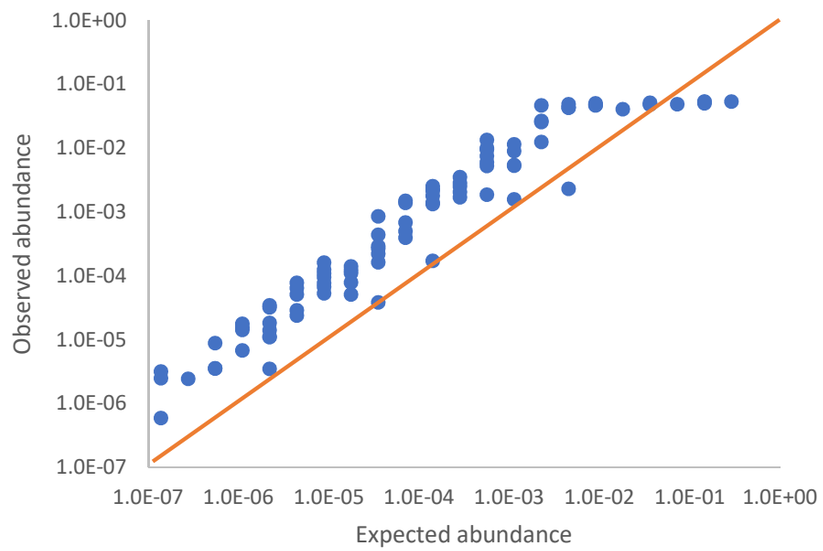




Accurate Quantification of Synthetic Standard

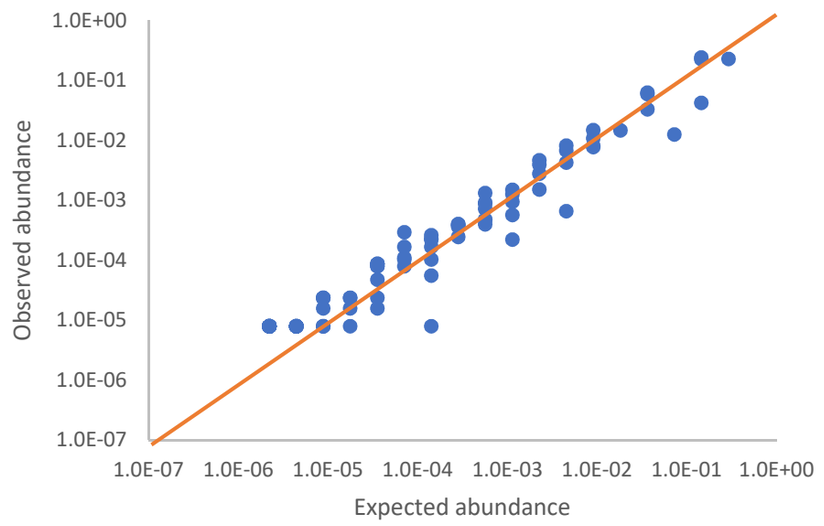
TruSeq RNA Prep Kit v2 + ERCC

- Short reads
- Quantification via coverage not accurate

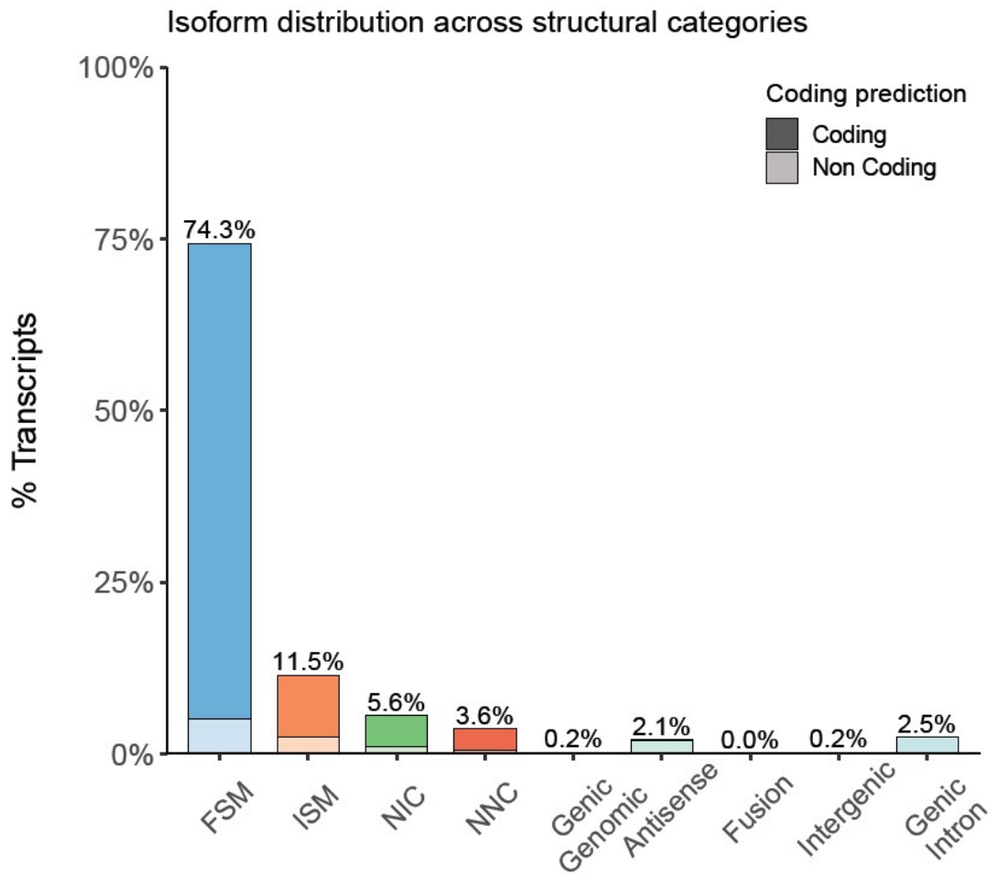


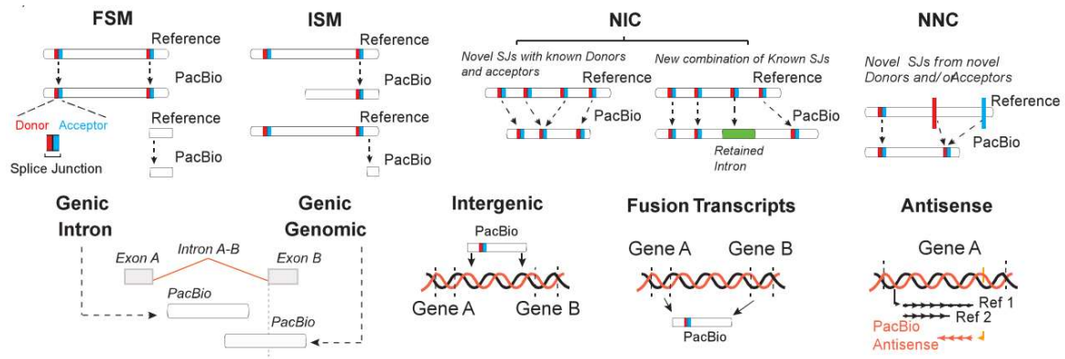
LoopSeq Transcriptome + ERCC

- Synthetic long reads from short reads
- Single molecule quantification



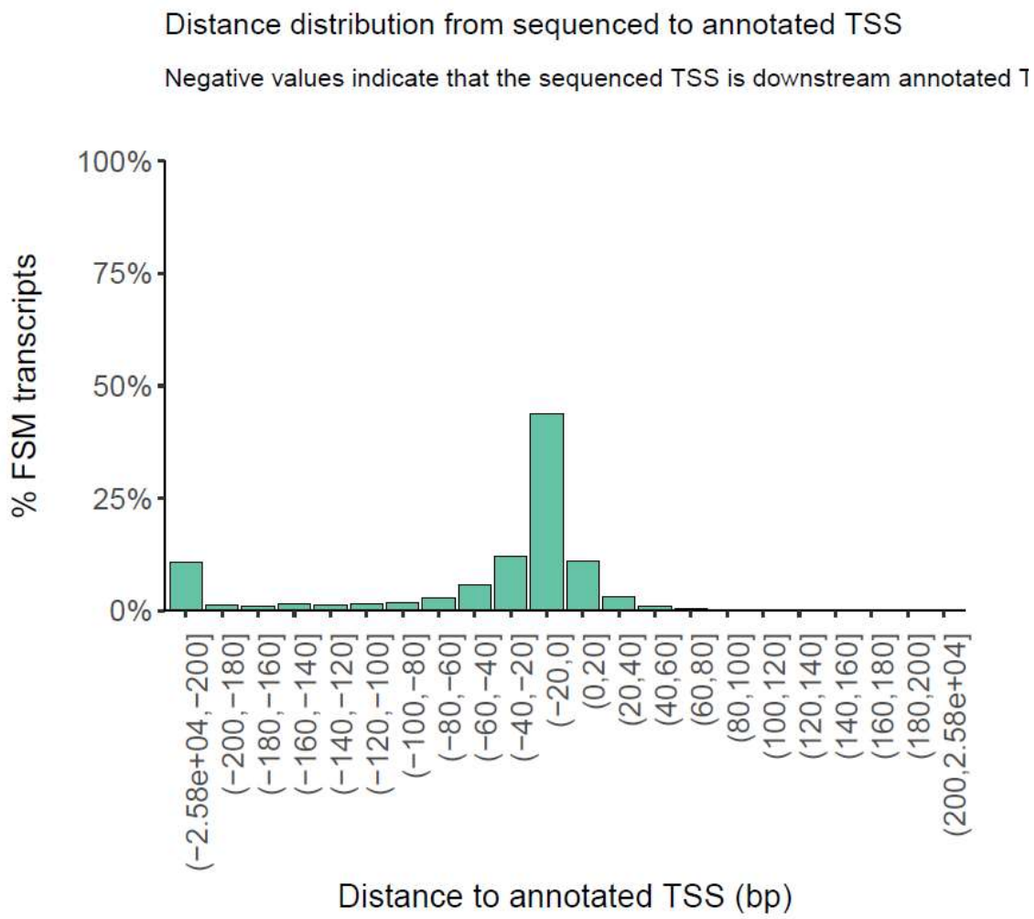
Full-length Coverage of Transcripts

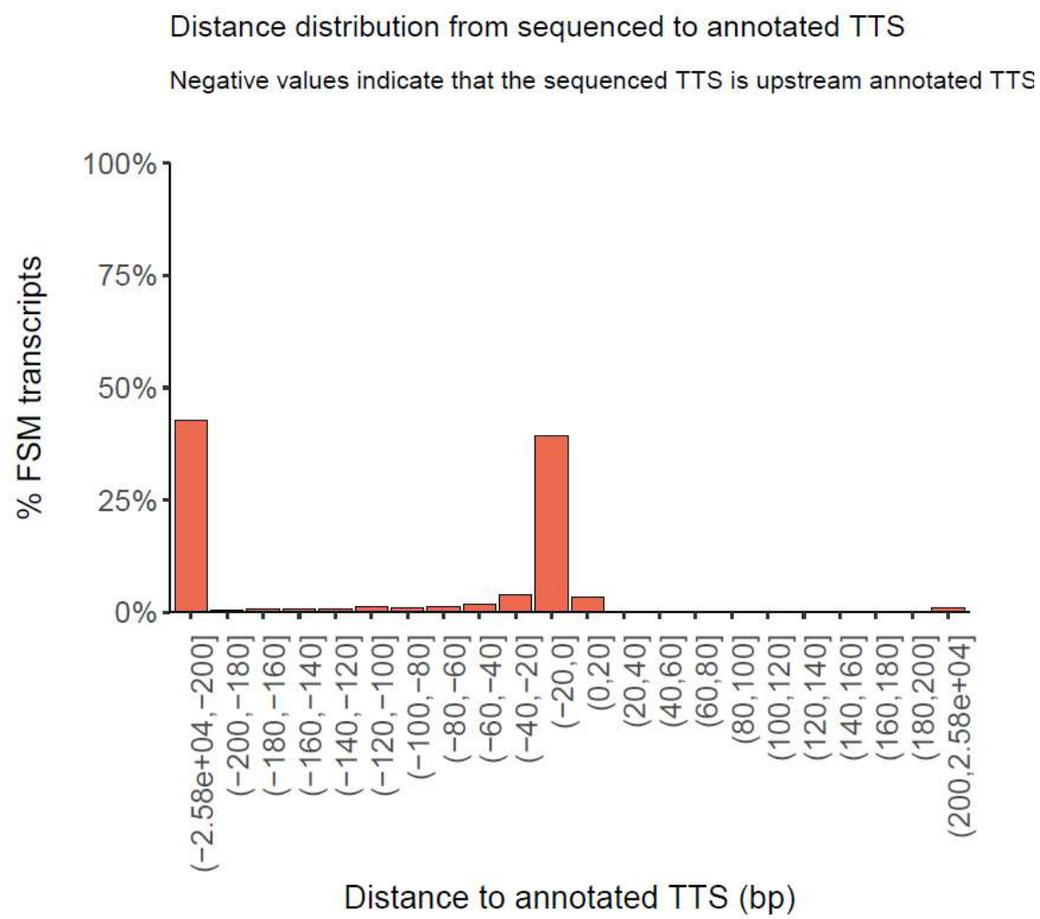




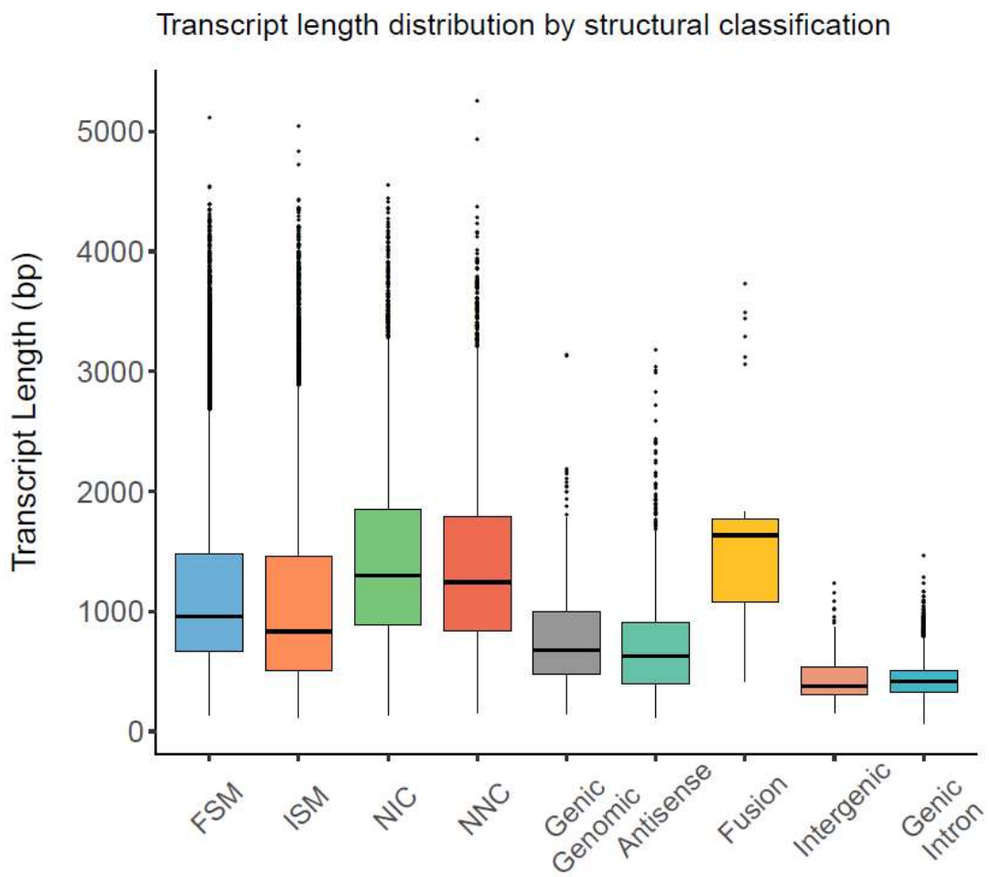
Abbreviation	Description
FSM	Full splice match
ISM	Incomplete splice match
NIC	New combo of known splice junctions Novel splice junction with known donor/acceptor
NNC	Novel splice junction from novel donor/acceptor
Genic genomic	Intron + Exon
Antisense	Antisense to annotated strand
Fusion	Fusion event between two annotated genes
Intergenic	Genomic sequence between annotated genes
Genic intron	Intron only

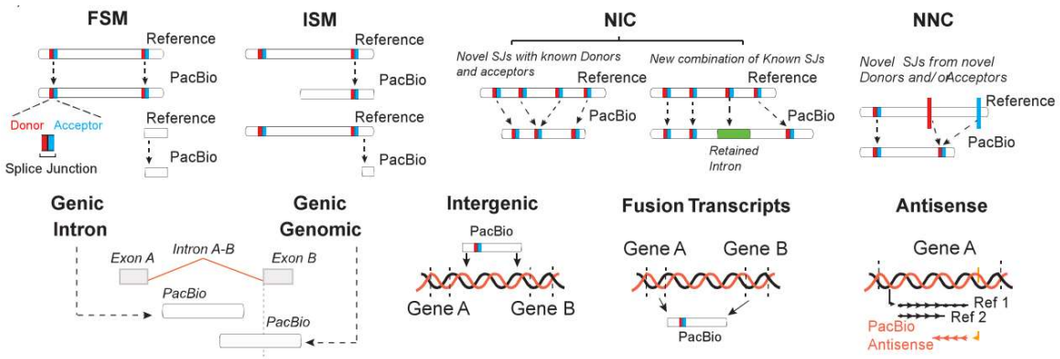
Full-length Coverage of Transcripts





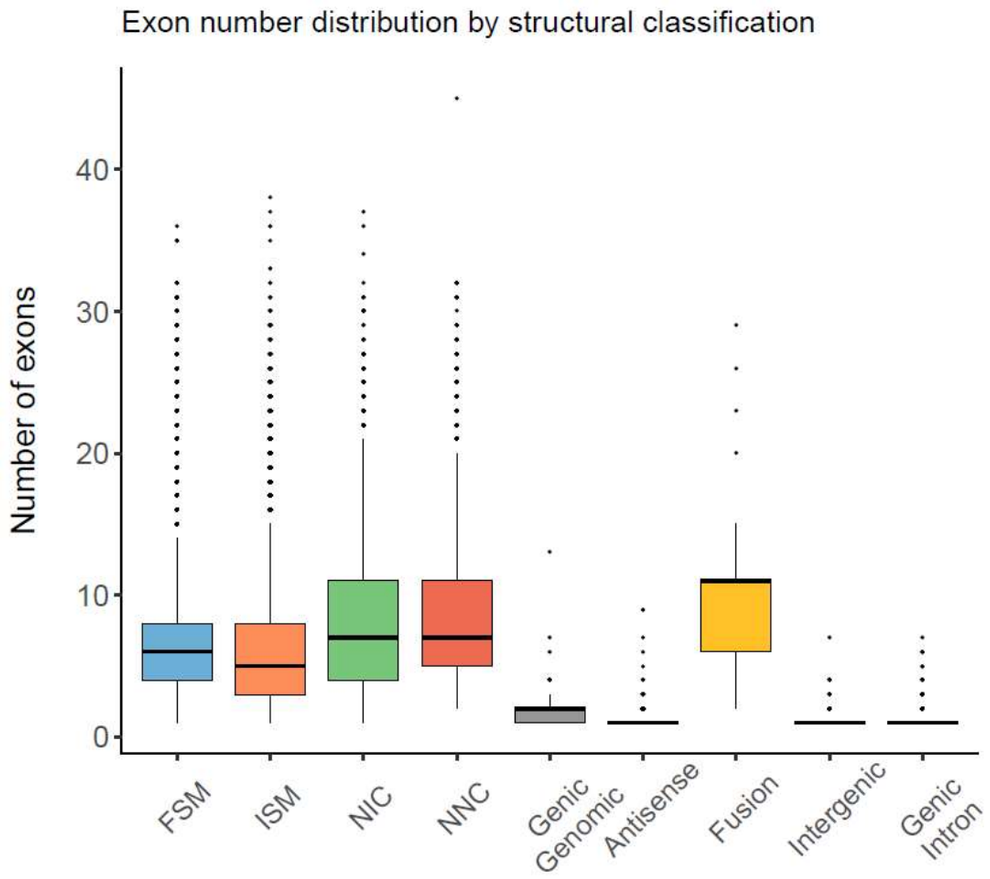
Long Transcripts for Isoform Discovery

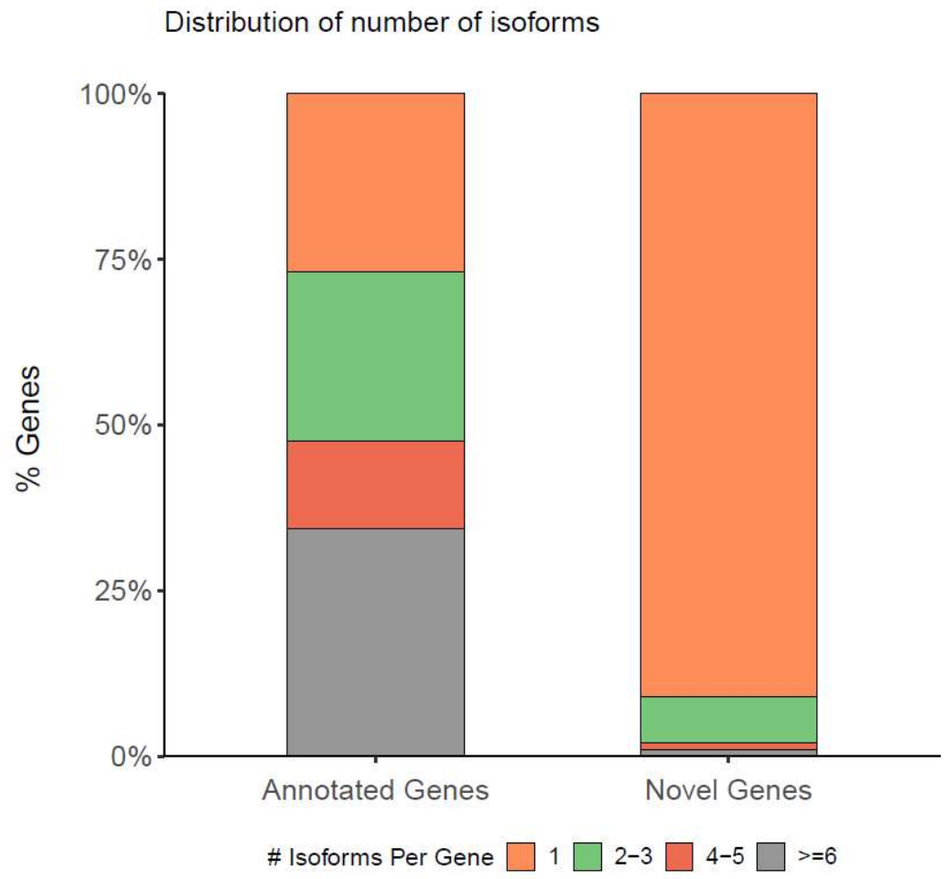




Abbreviation	Description
FSM	Full splice match
ISM	Incomplete splice match
NIC	New combo of known splice junctions Novel splice junction with known donor/acceptor
NNC	Novel splice junction from novel donor/acceptor
Genic genomic	Intron + Exon
Antisense	Antisense to annotated strand
Fusion	Fusion event between two annotated genes
Intergenic	Genomic sequence between annotated genes
Genic intron	Intron only

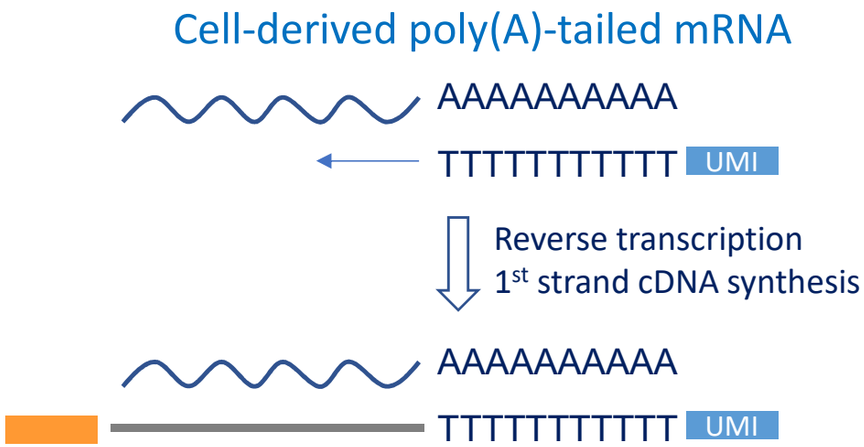
Long Transcripts for Isoform Discovery

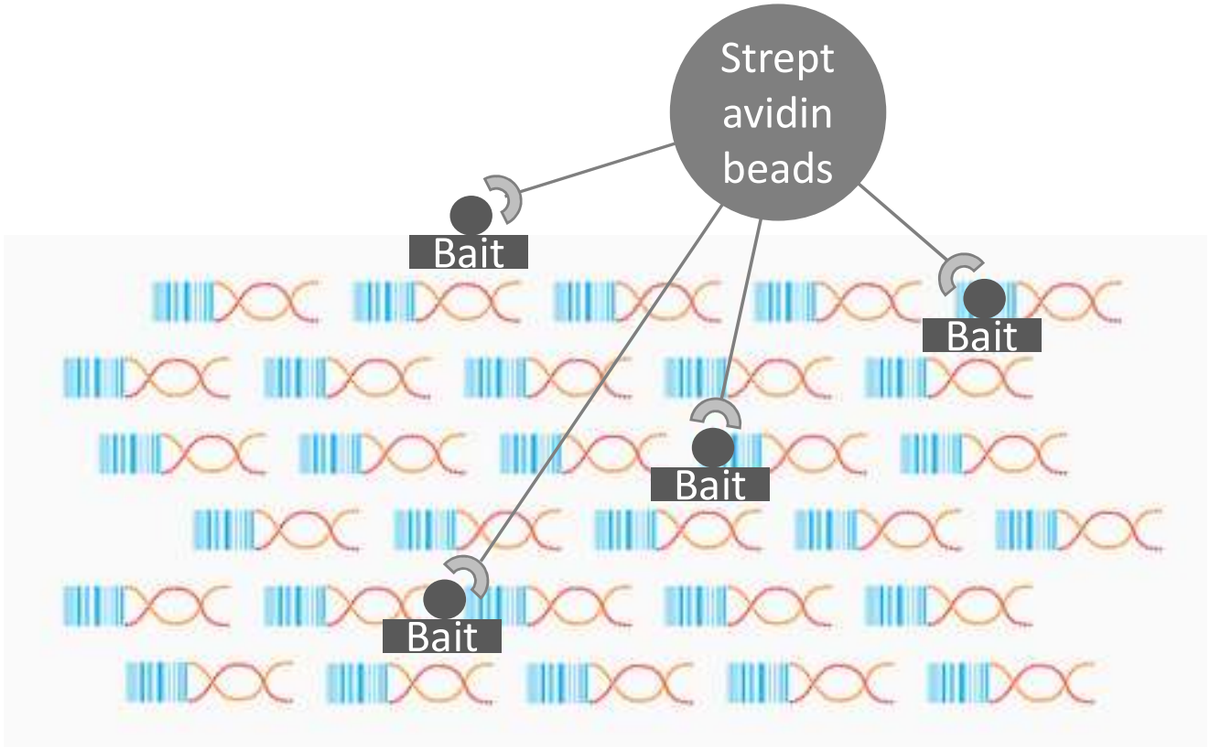




Many Full-length Transcript Applications

LoopSeq Library
+ Probe Capture*





***Probe capture on long-reads for targeted applications**

Target Enrichment Focuses Sequencing Power

LoopSeq Transcriptome + IDT Pan-Cancer Panel

Sample	Long Reads	% reads aligned	% selected bases
+ Capture	64,081	0.9998	0.86
No Capture	117,097	0.9999	0.02

LoopSeq Transcriptome + Twist Custom 800kb

Sample	Long Reads	% reads aligned	% selected bases
+ Capture	68,508	0.9999	0.74
No Capture	117,097	0.9999	0.01

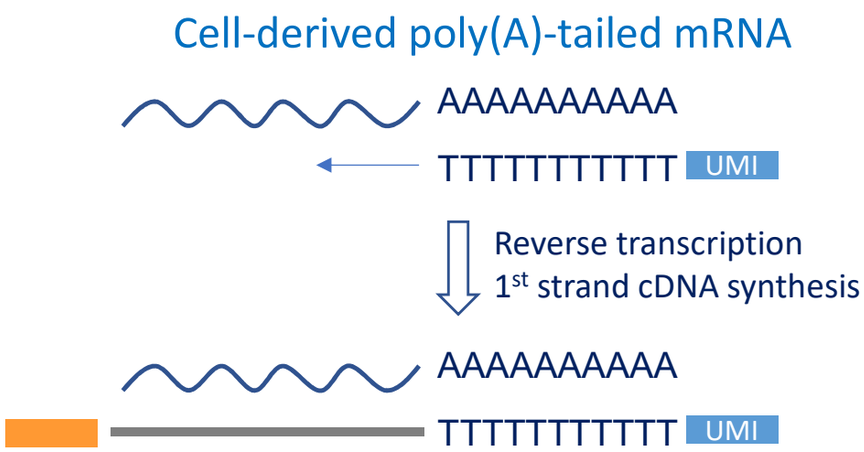
Isoform Discovery Case Study – EIF4A2

- 1.9kb transcript length, up to 12 exons
- Isoforms
  - Protein coding
  - Missing exon
  - Retained intron

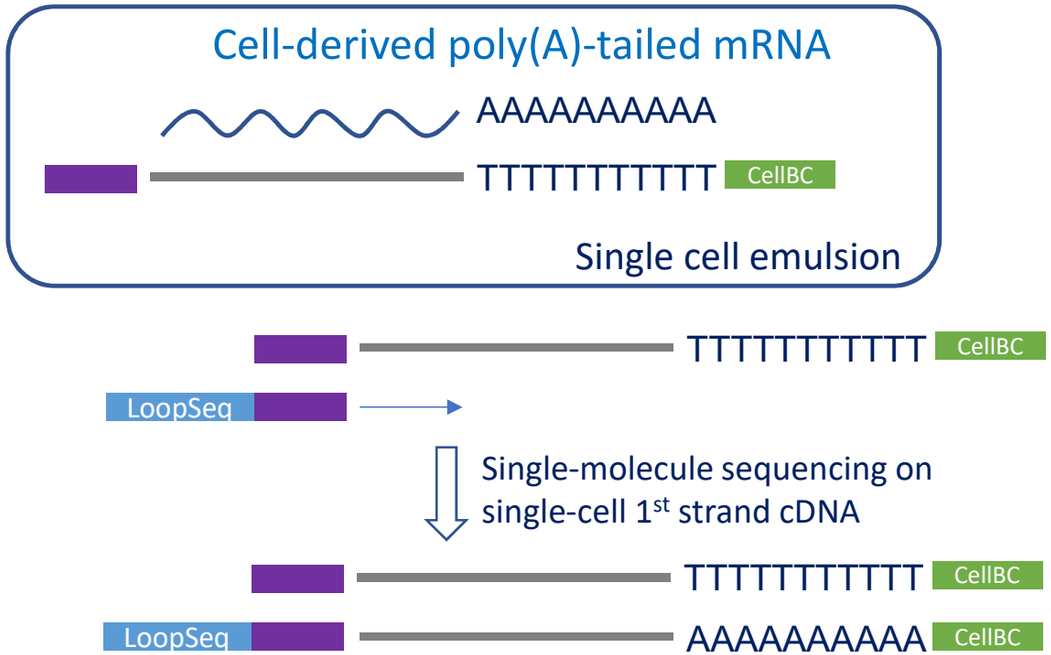
Name	Ref ID	Ref length	Count	Name	Ref ID	Ref length	Count
EIF4A2-201	ENST00000323963	1886	322	EIF4A2-213	ENST00000465792	734	2
EIF4A2-202	ENST00000425053	1977	32	EIF4A2-214	ENST00000466362	545	11
EIF4A2-203	ENST00000426808	1751	21	EIF4A2-218	ENST00000475653	909	23
EIF4A2-204	ENST00000429589	1005	3	EIF4A2-219	ENST00000485101	5327	32
EIF4A2-205	ENST00000440191	1889	92	EIF4A2-220	ENST00000486805	590	1
EIF4A2-206	ENST00000441007	480	10	EIF4A2-222	ENST00000492144	993	11
EIF4A2-207	ENST00000443963	1729	93	EIF4A2-223	ENST00000494445	629	6
EIF4A2-209	ENST00000461021	575	1	EIF4A2-224	ENST00000495049	576	2
EIF4A2-210	ENST00000465032	578	18	EIF4A2-225	ENST00000496382	460	11
EIF4A2-212	ENST00000465267	1054	2	EIF4A2-227	ENST00000498746	579	14

Many Full-length Transcript Applications

LoopSeq Transcriptome



Single-cell 1st strand cDNA + LoopSeq Long-read

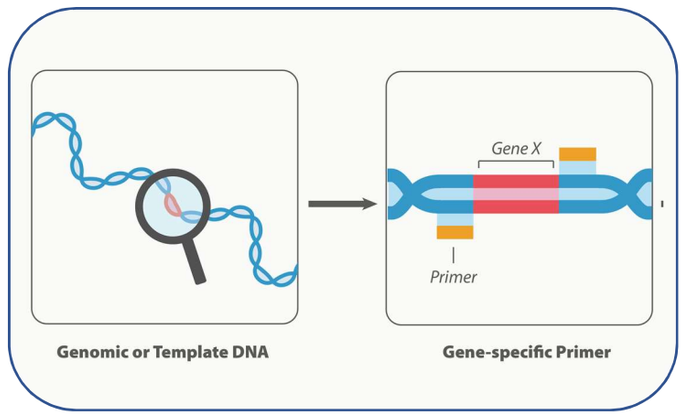


LoopSeq Long Read Applications

Micro/Mycobiome

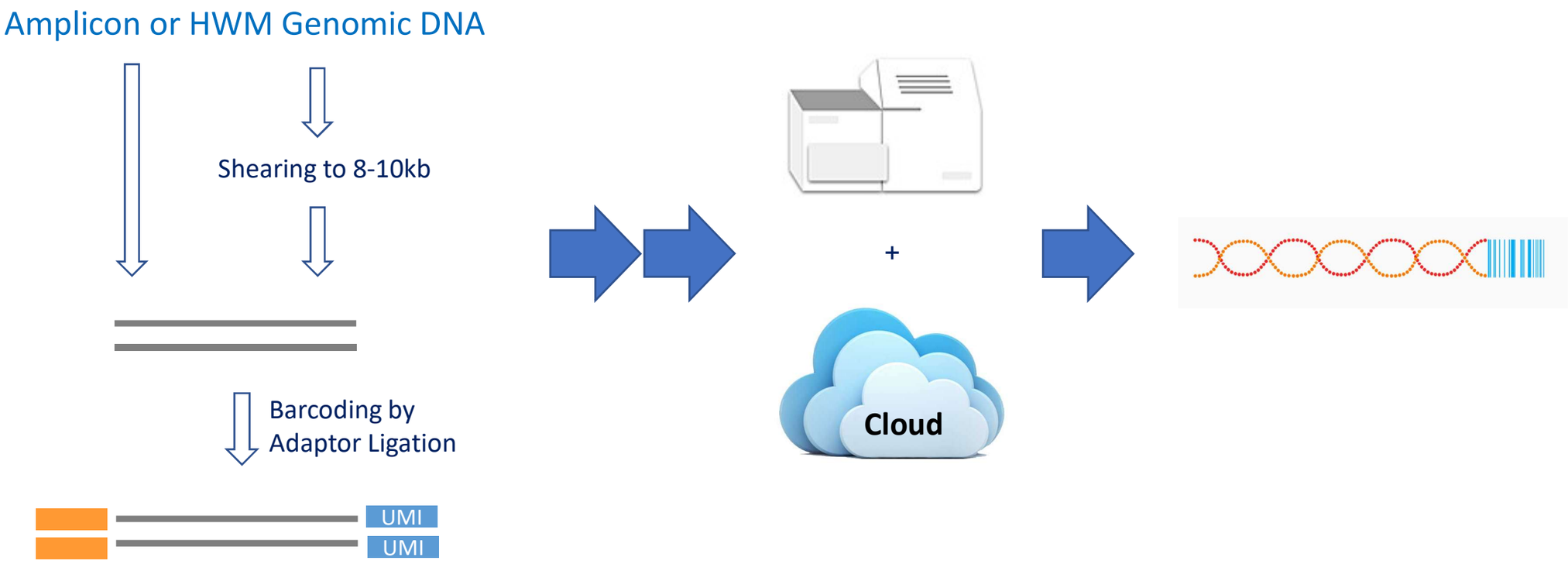
RNA

DNA

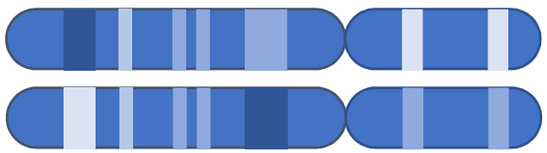


Single-molecule, long read genome or amplicon sequencing

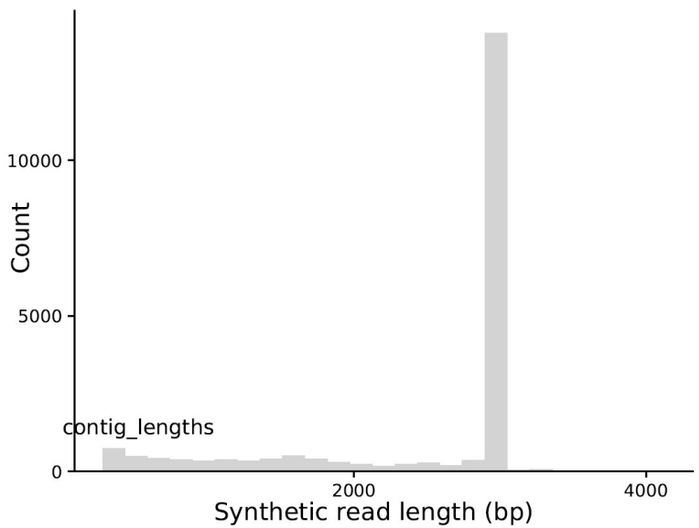
LoopSeq Amplicon or Genome Sequencing



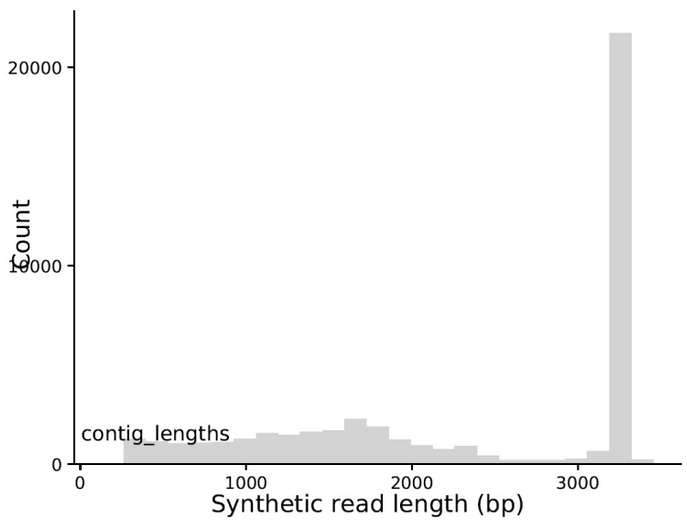
Phased Allele Calls with Single-molecule Sequencing



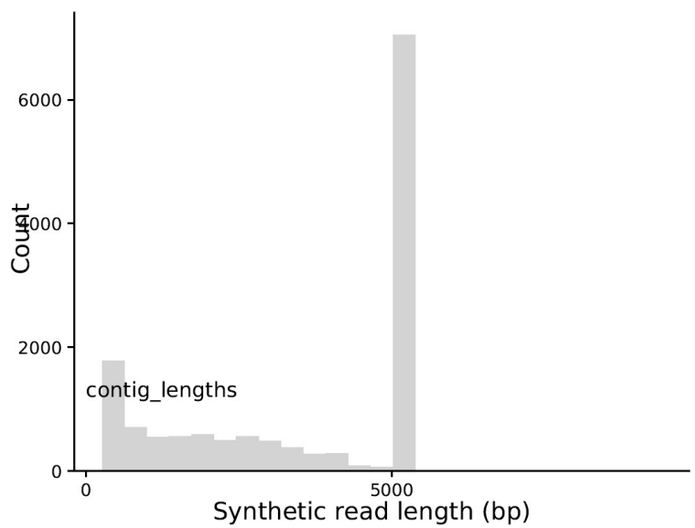
Target length: 2936bp

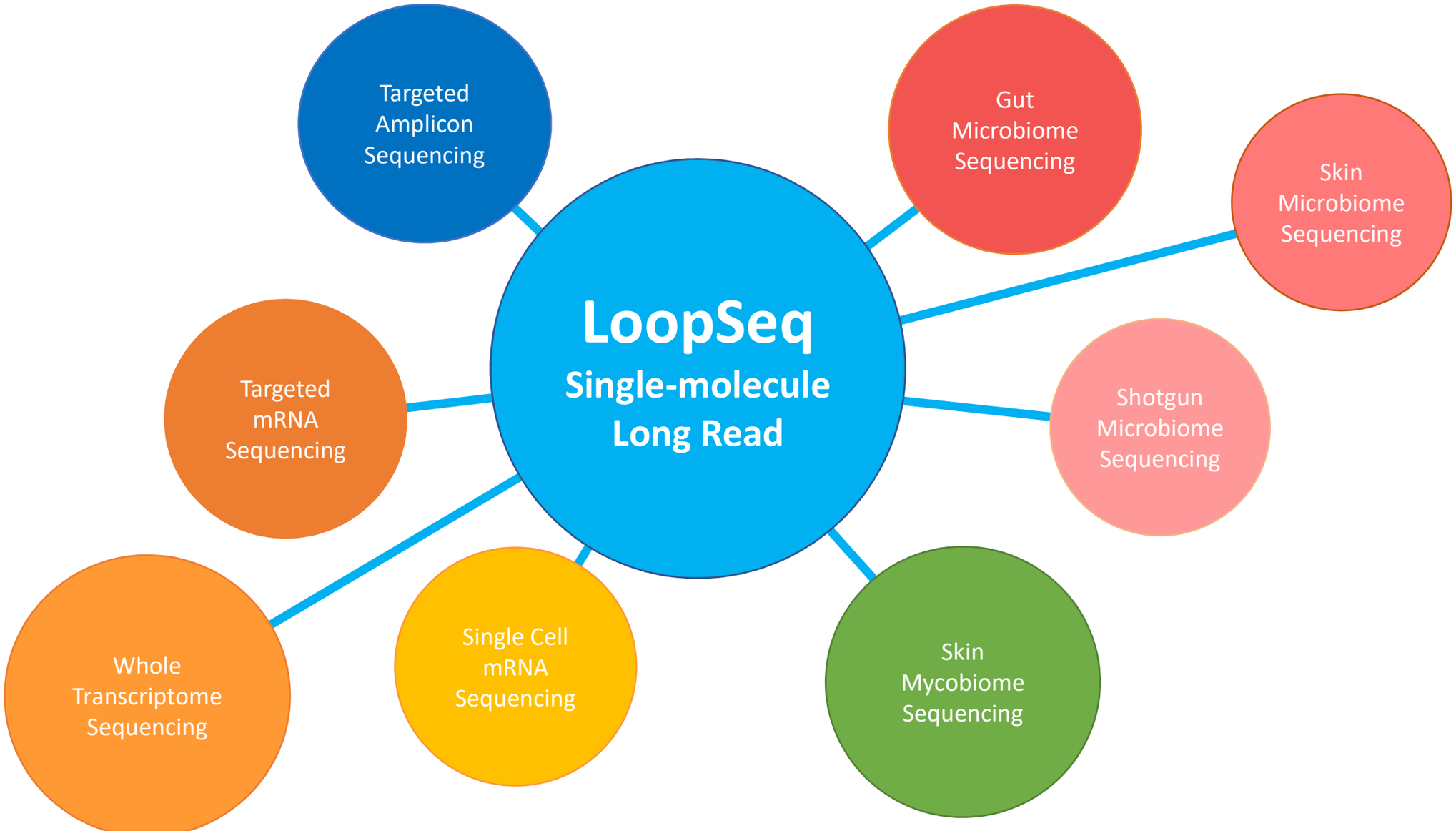


Target length: 3239bp



Target length: 5244bp




LoopSeq
Single-molecule
Long Read
Targeted
Amplicon
Sequencing
Gut
Microbiome
Sequencing
Skin
Microbiome
Sequencing
Targeted
mRNA
Sequencing
Shotgun
Microbiome
Sequencing
Whole
Transcriptome
Sequencing
Single Cell
mRNA
Sequencing
Skin
Mycobiome
Sequencing

LoopSeq™ 16S Microbiome SSC 24-plex Kit
LoopSeq™ 16S & 18S Microbiome SSC 24-plex Kit

LoopSeq™ 16S Low Biomass Microbiome SSC 24-plex Kit
LoopSeq™ 16S & 18S Low Biomass Microbiome SSC 24-plex Kit

LoopSeq™ 16S Read Cloud Microbiome SSC 24-plex Kit
LoopSeq™ 16S & 18S Read Cloud Microbiome SSC 24-plex Kit

LoopSeq™ Mycobiome 18S ITS1 & ITS2 SSC 24-plex Kit

LoopSeq™ Transcriptome 3x8-plex Kit

LoopSeq™ DNA Genome 3x8-plex Kit
LoopSeq™ DNA Amplicon 3x8-plex Kit

Thank you!