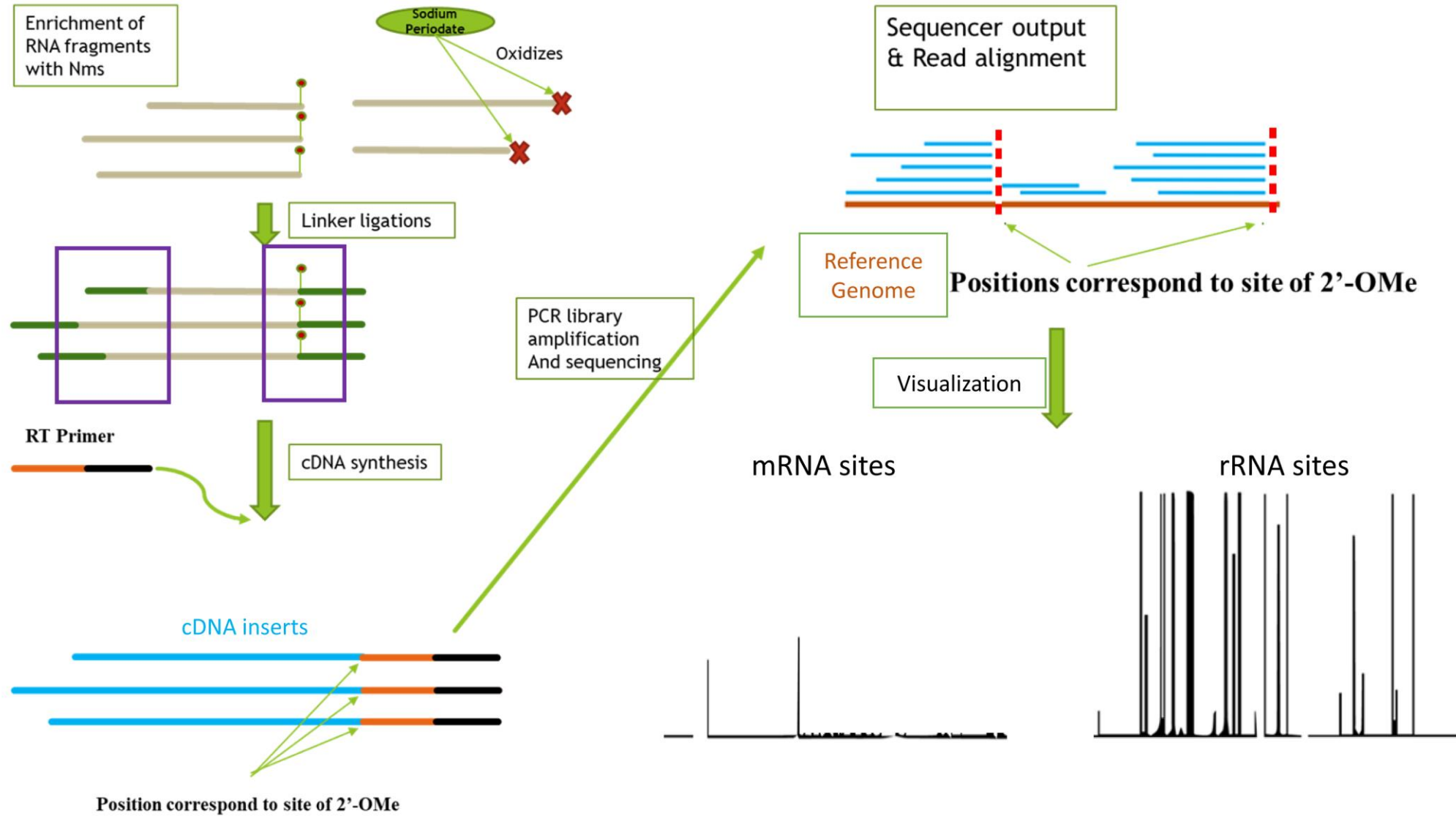


Demonstration of basic RibOxi-seq concept



RNA structure after 3' and 5' linker ligations

5'-/Biosg/**ACACUCUUCCCUACACGACGCUCUCCGAUCU**NNNN-----insert-----**ATCACGCTGTAGGCACCATCAATGACAG**/SpC3/- 3'

cDNA library structure

Randomer for deduplication

3'-linker sequence

3'-**TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAGA**NNNN----insert----**TAGTGC****GACATCCGTGGTAGTTACTG****TCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG**-5'

Complementary to partial Illumina i5 PCR primer

In-line barcode for filtering out RT mis-priming

Identical to partial Illumina i7 PCR primer

dsDNA library structure

Forward	5'- Illumina_i5_adapter NNNN----insert---- ATCACGCTGTAGGCACCATCAATGAC Illumina_i7_adapter -3'
Reverse	3'- Illumina_i5_adapter NNNN----insert---- TAGTGC GACATCCGTGGTAGTTACTG Illumina_i7_adapter -5'

Sequencing Output (example using 1*150 single-end sequencing)

Read:

@ Sequence_identifier+i7_index+i5_index
NNNN----INSERT----**ATCACGCTGTAGGCACCATCAATGAC**---MAYBE_READ_THROUGH_INTO_ILLUMINA_PCR_ADAPTER
+
_QUALITY_OF_THE_READ_

1. *cutadapt* used to remove partial linker sequence and read-through adapters, discarding reads without match:

Read output:

```
@ Sequence_identifier+i7_index+i5_index
NNNN----INSERT----ATCACG
+
__QUALITY_OF_THE_READ__
```

2. *move_umi.py* moves randomer sequence from read to read-identifier:

```
@ Sequence_identifier_NNNN+i7_index+i5_index
INSERT----ATCACG
+
__OF_THE_READ__
```

3. *cutadapt* used to remove the barcode portion of linker and discard any read without a perfect match:

```
@ Sequence_identifier_NNNN+i7_index+i5_index
INSERT
+
__OF_TH
```

4. *STAR* alignment to genome of interest

5. *samtools* indexing the bam output

6. *umi_tools* deduplication:

Example aligned read #1:

UMI: AATC

Sequence: GGTTACG

Example aligned read #2:

UMI: CGTA

Sequence: GGTTACG

Not duplicates

Example aligned read #1:

UMI: AATC

Sequence: GGTTACG

Example aligned read #2:

UMI: AATC

Sequence: GGTTACG

Duplicates, keep only 1 copy

7. *bedtools* converts deduplicated bam to bed file format

8. *genomecov* generates genome coverage track of 3'-end only to visualize Nm sites on genome browser

9. *riboxi_bed_parsing.py* counts all 3'-end alignments and generates a tab delimited file listing chromosome number, base position, gene name and counts from the bed file:

chr	base	gene	WT5_R	DKO3	DKO3_R	WT5
chr1	13236	DDX11L1	0	1	0	0
chr1	13238	DDX11L1	0	0	0	1
chr1	14405	WASH7P	0	5	0	1
chr1	14406	WASH7P	0	3	0	0
chr1	14408	WASH7P	0	1	0	1
chr1	14409	WASH7P	0	1	0	0
chr1	14413	WASH7P	0	1	0	1

10. The shinyapp allows more interactive interrogation of data such as filtering and visualization on the fly:



Data Visualization Summary

[Click here for a demonstration on how counts are normalized \(Median of ratios\).](#)

Click any row below to select a gene and fields on the left will be filled automatically.

Show 10 entries

Search:

Please select a row.

chr	base	gene	WT	WT.1	WT.2	DKO	DKO.1	DKO.2	counts_mean
All	All	All	All	All	All	All	All	All	All
226	chr21	8438679 FP236383.3	1764	1431	23	47	711	1052	838
143	chr21	8394469 FP236383.3	1773	1415	23	47	709	1056	837.166666666667
59	chr21	8211434 FP671120.4	1769	1411	23	47	707	1050	834.5
130	chr21	8393532 FP236383.3	1521	1199	9	19	549	918	702.5
213	chr21	8437742 FP236383.3	1523	1196	9	19	551	917	702.5
46	chr21	8210497 FP671120.4	1518	1198	9	19	551	915	701.666666666667
139	chr21	8394112 FP236383.3	1223	1078	24	98	697	902	670.333333333333
222	chr21	8438322 FP236383.3	1236	1070	24	97	694	899	670
55	chr21	8211077 FP671120.4	1226	1060	24	97	701	906	669
53	chr21	8211021 FP671120.4	1299	1139	15	24	554	902	655.5

Showing 1 to 10 of 276 entries

Previous 1 2 3 4 5 ... 28 Next

Download table