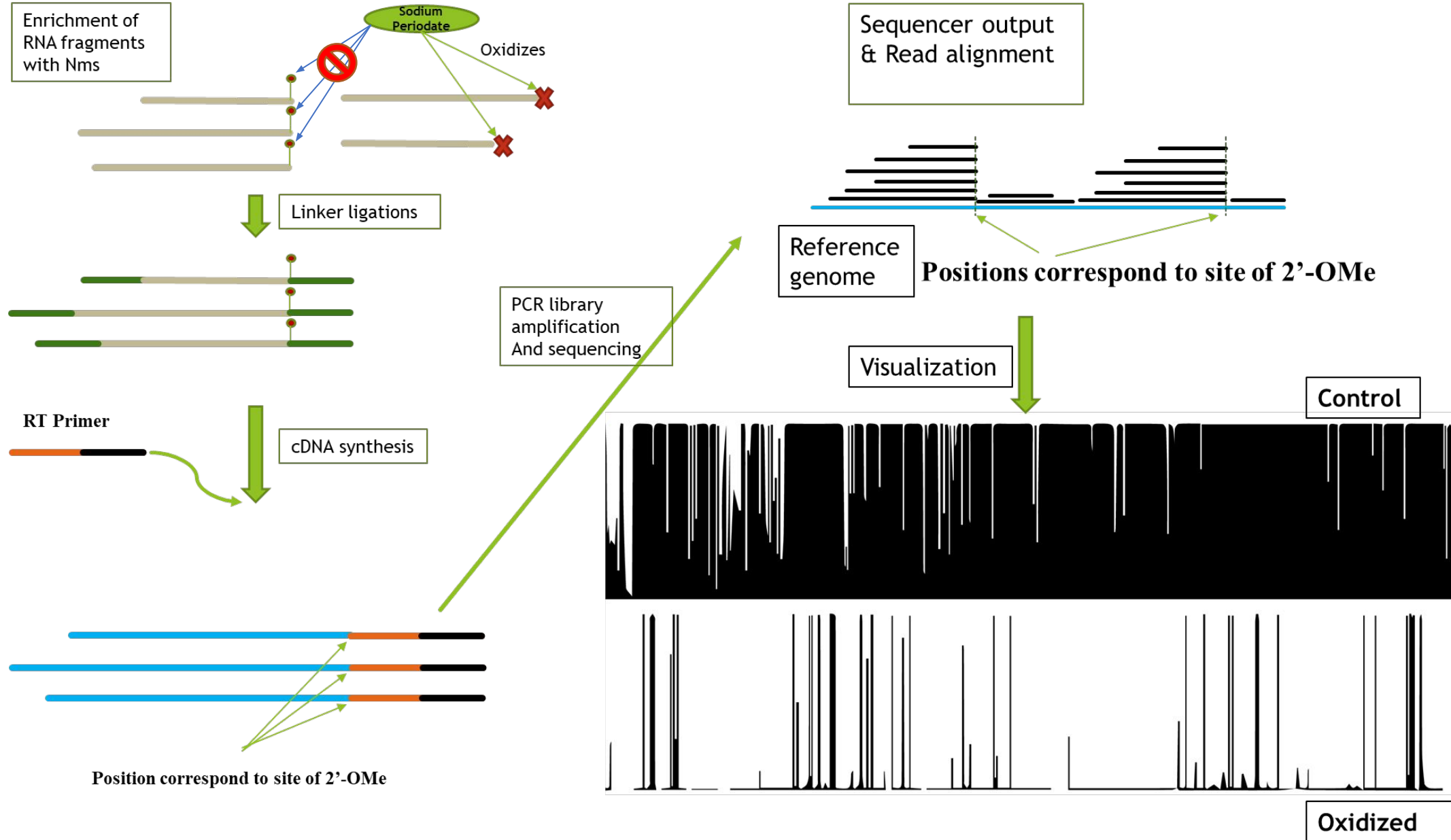


Demonstration of basic RibOxi-seq concept



RNA structure after 3' and 5' linker ligations

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

cDNA structure

Randomer for deduplication 3'-linker sequence

3'-**TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAG**NNNN----insert----**TAGTGC**GACATCCGTGGTAGT**ACTGTCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG**-5'

Complementary to partial Illumina i5 PCR primer In-line barcode for filtering out RT mispriming Identical to partial Illumina i7 PCR primer

Sequencing Output (example using 75x75 Paired end sequencing)

Read1:

@ Sequence_identifier+i7_index+i5_index
NNNN_SOME_INSERT_AND_MAYBE_READ_THROUGH_ADAPTER
+
_____QUALITY_OF_THE_READ_____

Read2:

@Sequence_identifier+i7_index+i5_index
GTCATTGATGGTGCCTACAGCGTGAT_SOME_INSERT_AND
READ_THROUGH_ADAPTER
+
_____QUALITY_OF_THE_READ_____

1. *cutadapt* paired-end mode remove read-through

adapters:

Read1 output:

Read2 output:

```
@ Sequence_identifier+i7_index+i5_index
NNNN_SOME_INSERT
+
_____QUAL
```

```
@Sequence_identifier+i7_index+i5_index
GTCATTGATGGTGCCTACAGCGTGAT_SOME_INSERT
+
_____QUALITY_OF_THE_REA
```

2. *pear* to merge read1 and read2 into a single read:

```
@ Sequence_identifier+i7_index+i5_index
NNNN_SOME_INSERT_ATCACGCTGTAGGCACCATCAATGACAG
+
_____QUALAER_EHT_FO_YTILAUQ_____
```

3. *move_umi.py* moves randomer sequence from read to read identifier and discard reads lacking **ATCACG**:

```
@ Sequence_identifier_NNNN+i7_index+i5_index
SOME_INSERT_ATCACGCTGTAGGCACCATCAATGACAG
+
_____QUALAER_EHT_FO_YTILAUQ_____
```

4. *cutadapt* removes the rest of the linker sequence:

```
@ Sequence_identifier_NNNN+i7_index+i5_index  
SOME_INSERT  
+  
_____QUAL
```

5. *STAR* alignment to genome of interest

6. *samtools* indexing the bam output

7. *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, discard

- 8. *bedtools* converts deduplicated bam to bed file format
- 9. *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, (+15,base,-15) nucleotide sequence and counts for all samples:

1	chr	base	gene	seq	ms_wt_186_0x	
2	chr13	97190585	Hexb	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		13263
3	chr9	123462160	Lars2	gttggtgcatggtaatcctgctcagtacga		11174
4	chr2	102829412	Cd44	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		10624
5	chr17	39848136	Rn45s	gaagacggtcgaacttgactatctagaggaa		9501
6	chr17	39847806	Rn45s	tcccccaacttcttagagggacaagtggcgt		4412
7	chr17	39846960	Rn45s	gaggatccattggagggaagtctggtgcca		3148
8	chr7	19697484	Apoe	CCAAGTCACACAAGAACTGACGTGAGTGTC		2871
9	chr5	136932984	4933404012Rik	gtctccaaggtgaacagcctctggcacattg		2850
10	chr1	100180179	Cntnap5b	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		2801
11	chr17	39847687	Rn45s	attccgtgggtggtggtgcatggccgttctt		2405
12	chr11	17221313	Wdr92	gtctccaaggtgaacagcctctggcatgttg		2286

Downstream analysis can then be performed

RibOxi-seq Counts Visualization

Sample:

WT WT.1 DKO.1 DKO DKO.2

Base position min mean counts:

50

Gene Symbol:

DLG2

Chromosome:

chr11

Plot

Shaw 10 entries

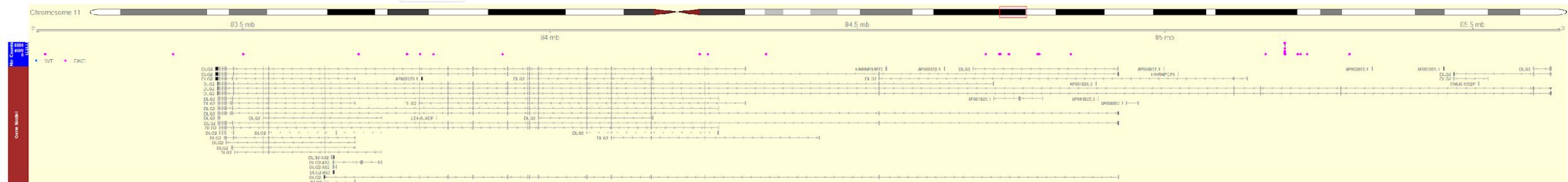
Search:

chr	base	gene	seq	WT	WE1	DKO	DKO.1	DKO.2	counts_mean
All	All	All	All	All	All	All	All	All	All
308 chrJh_g1000220	154835	RNA3-BS5	GAAGACGTCGAACCTGACTATCTAGAGGAA	24211	40345	4C175	57289	39099	40156
150 chrJh_g1000220	110801	MIR6724-3	GAACACGGTCCAACTGACTATCTAGAGCAA	24212	40338	4C176	57265	39059	40151.2
188 chrJh_g1000220	115762	MIR6724-3	GAGAAAGGTTCCATGTGAACAGCAGTTGAC	11435	55412	23168	39187	35956	36933.8
6 chr12	20704171	PDE3A	ATAACGAACGAGACTCTGGCATGCTAACTAG	20743	31833	27045	52255	28240	32023.2
295 chrJh_g1000220	154440	RNA5-BS5	ATAACGAACGAGACTCTGGCATGCTAACTAG	20745	31822	27047	52248	28238	32022.2
137 chrJh_g1000220	110468	MIR6724-3	ATAACGAACGAGACTCTGGCATGCTAACTAG	20744	31820	27043	52237	28234	32015.6
177 chrJh_g1000220	114972	MIR6724-3	CCSACCTGGGTATAGGGCGAAAGACTAATC	12958	25445	17525	20006	22370	19881.6
305 chrJh_g1000220	154752	RNA5-BS5	CCTTTGTACACACCGCCGTCGCTACTACCG	3404	10550	15105	32183	17377	16942.2
147 chrJh_g1000220	110780	MIR6724-3	GCTTTGTACACACCGCCGTCGCTACTACCG	3484	15554	15106	32174	17373	16939.2
296 chrJh_g1000220	154539	RNA5-BS5	GATTGA3CAATAACAGGTCTGTGATGCCCTT	5866	12446	11240	17202	12875	12025.8

Showing 1 to 10 of 339 entries

Download Table

Previous 1 2 3 4 5 ... 31 Next

 Download picture

Select a site:

- 85195149
- 60195147
- 88195129
- 86195127

Download

File Download plot

