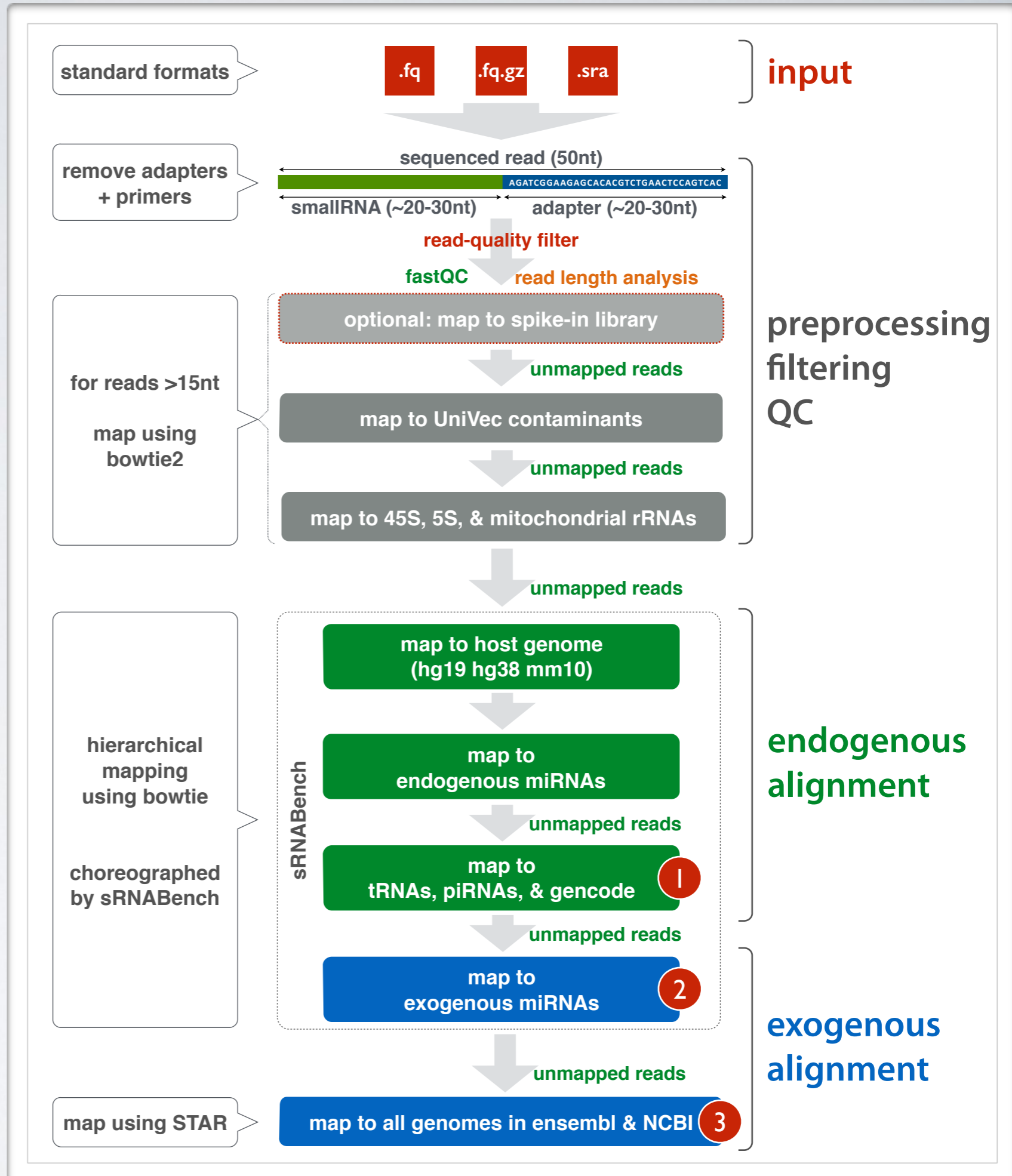


# exceRpt

- automatic pre-processing and QC of sequence reads
- absolute quantitation by quantification of exogenous spike-in sequences
- explicit rRNA filtering & QC
- quantify many different smallRNA types
- choice of 3 end-points

1 2 3



# effect of filtering

- what happens to alignments when we individually remove upstream libraries?
- gencode < RE
- UniVec / RE have largest effect on exogenous genome alignments

input	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
successfully_clipped	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
failed_quality_filter	0.0%	10.3%	10.3%	10.3%	10.3%	10.3%	10.3%	10.3%	10.3%
failed_homopolymer_filter	0.0%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
calibrator	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
UniVec_contaminants	7.0%	0.0%	6.4%	6.4%	6.4%	6.4%	6.4%	6.4%	6.4%
rRNA	4.2%	3.8%	0.0%	3.8%	3.8%	3.8%	3.8%	3.8%	3.8%
reads_used_for_alignment	88.8%	85.8%	83.3%	79.5%	79.5%	79.5%	79.5%	79.5%	79.5%
genome	65.2%	62.3%	62.5%	58.8%	58.8%	58.8%	58.8%	58.8%	58.8%
miRNA_sense	52.3%	47.9%	47.9%	48.5%	48.5%	48.5%	48.5%	48.5%	48.5%
miRNA_antisense	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%
tRNA_sense	0.4%	0.4%	0.4%	0.0%	0.4%	0.4%	0.4%	0.4%	0.4%
tRNA_antisense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
piRNA_sense	0.1%	0.2%	0.2%	0.1%	0.0%	0.1%	0.1%	0.1%	0.1%
piRNA_antisense	0.1%	0.1%	0.1%	0.1%	0.0%	0.1%	0.1%	0.1%	0.1%
gencode_sense	4.6%	4.8%	4.5%	3.9%	3.8%	0.0%	3.7%	3.7%	3.7%
gencode_antisense	3.9%	5.8%	4.6%	3.5%	3.5%	0.0%	3.4%	3.4%	3.4%
repetitiveElement_sense	1.5%	1.8%	2.3%	1.3%	1.2%	4.7%	0.0%	1.2%	1.2%
repetitiveElement_antisense	1.2%	1.2%	1.3%	1.0%	1.0%	4.2%	0.0%	1.0%	1.0%
circularRNA_sense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
circularRNA_antisense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
input_to_miRNA_exogenous	20.7%	19.8%	18.4%	18.4%	18.4%	18.7%	19.4%	18.4%	18.4%
miRNA_exogenous_sense	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
input_to_exogenous_genomes	20.7%	19.8%	18.4%	18.4%	18.4%	18.6%	19.4%	18.4%	18.4%
exogenous_genomes	1.4%	2.0%	1.3%	1.3%	1.3%	1.4%	1.7%	1.3%	1.3%
	SRR822433_1_noQualFilter	SRR822433_2_noUniVec	SRR822433_3_noRiboRNA	SRR822433_4_notRNA	SRR822433_5_nopiRNA	SRR822433_6_noGencode	SRR822433_7_noRE	SRR822433_8_noCirc	SRR822433_FullPipeline

# mappability

- what happens if we run the smallRNA libraries through the pipeline as if they were reads?
- miRNAs include all species, but most are highly conserved
- 1/20 piRNAs are rRNA

input	100.0%	100.0%	100.0%	100.0%	100.0%
clipped	100.0%	100.0%	100.0%	100.0%	100.0%
failed_quality_filter	0.0%	0.0%	0.0%	0.0%	0.0%
failed_homopolymer_filter	0.3%	0.3%	0.3%	0.0%	0.0%
calibrator	NA%	NA%	NA%	NA%	NA%
UniVec_contaminants	0.0%	0.0%	0.0%	0.0%	0.0%
rRNA	0.1%	0.2%	0.1%	5.0%	0.0%
reads_used_for_alignment	99.6%	99.5%	99.6%	95.0%	100.0%
genome	75.5%	75.5%	73.9%	94.9%	84.2%
miRNA_sense	32.5%	32.4%	27.7%	0.0%	0.0%
miRNA_antisense	0.6%	0.6%	0.3%	0.0%	0.0%
tRNA_sense	0.0%	0.0%	0.0%	0.2%	84.2%
tRNA_antisense	0.0%	0.0%	0.0%	0.0%	0.0%
piRNA_sense	0.0%	0.0%	0.0%	94.6%	0.0%
piRNA_antisense	0.0%	0.0%	0.0%	0.1%	0.0%
gencode_sense	7.8%	7.7%	8.3%	0.0%	12.3%
gencode_antisense	4.9%	4.9%	5.8%	0.0%	3.2%
repetitiveElement_sense	8.1%	7.9%	7.5%	0.0%	0.5%
repetitiveElement_antisense	8.9%	9.1%	7.4%	0.0%	25.8%
circularRNA_sense	0.0%	0.0%	0.0%	0.0%	0.0%
circularRNA_antisense	0.0%	0.0%	0.0%	0.0%	0.0%
input_to_miRNA_exogenous	18.7%	18.7%	20.8%	0.0%	0.0%
miRNA_exogenous_sense	13.6%	13.6%	14.9%	0.0%	0.0%
input_to_exogenous_genomes	1.4%	1.4%	1.4%	0.0%	0.0%
exogenous_genomes	1.2%	1.2%	1.4%	0.0%	0.0%
	miRBase21_hg19_FullPipeline -	miRBase21_hg38_FullPipeline -	miRBase21_mm10_FullPipeline -	piRNAs_hg38_FullPipeline -	tRNAs_hg38_FullPipeline -