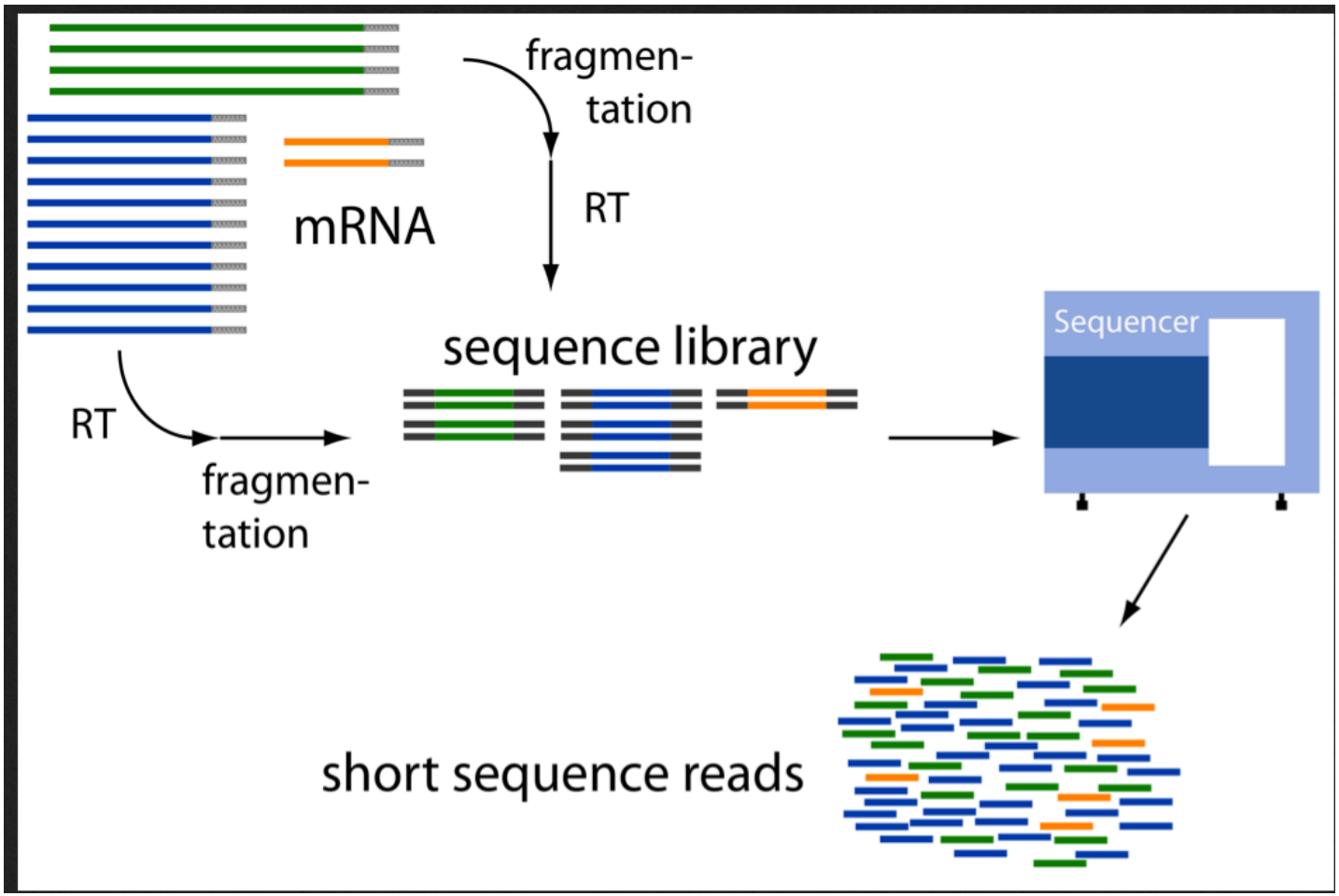
RNA-seq introduction

RNA-seq data analysis

Johan Reimegård | 13-May-2019

How are RNA-seq data generated?



Sampling process

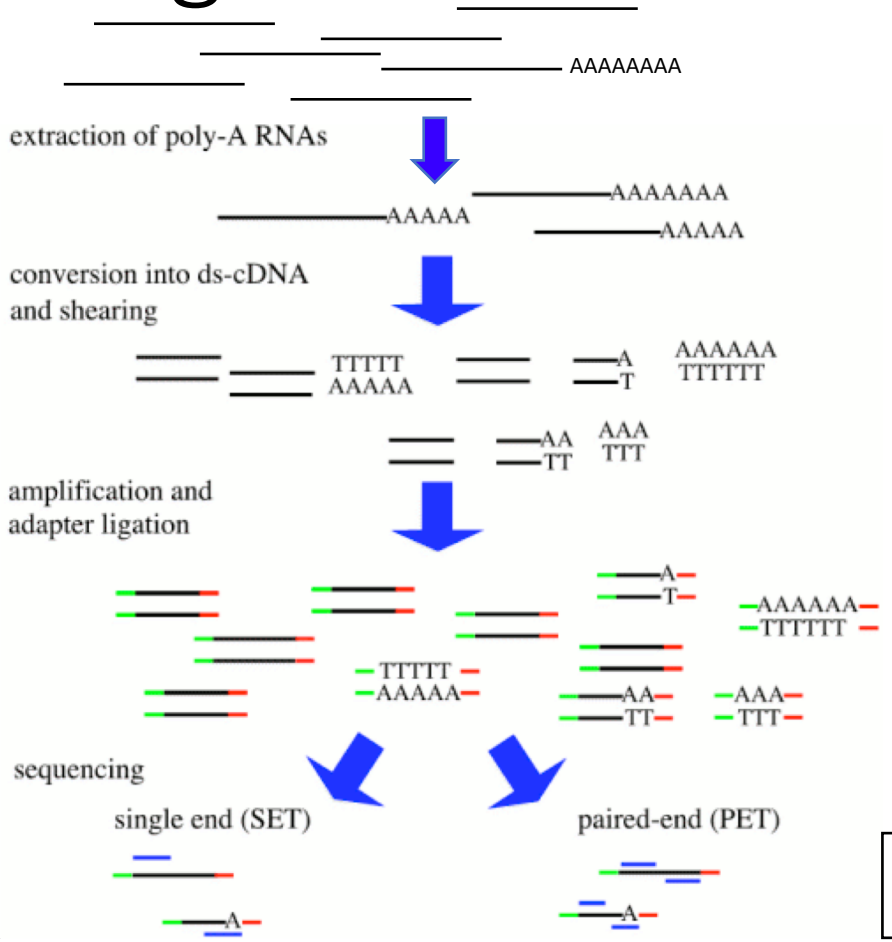
Depending on the different steps you will get different results

RNA->

enrichments ->

library ->

reads ->



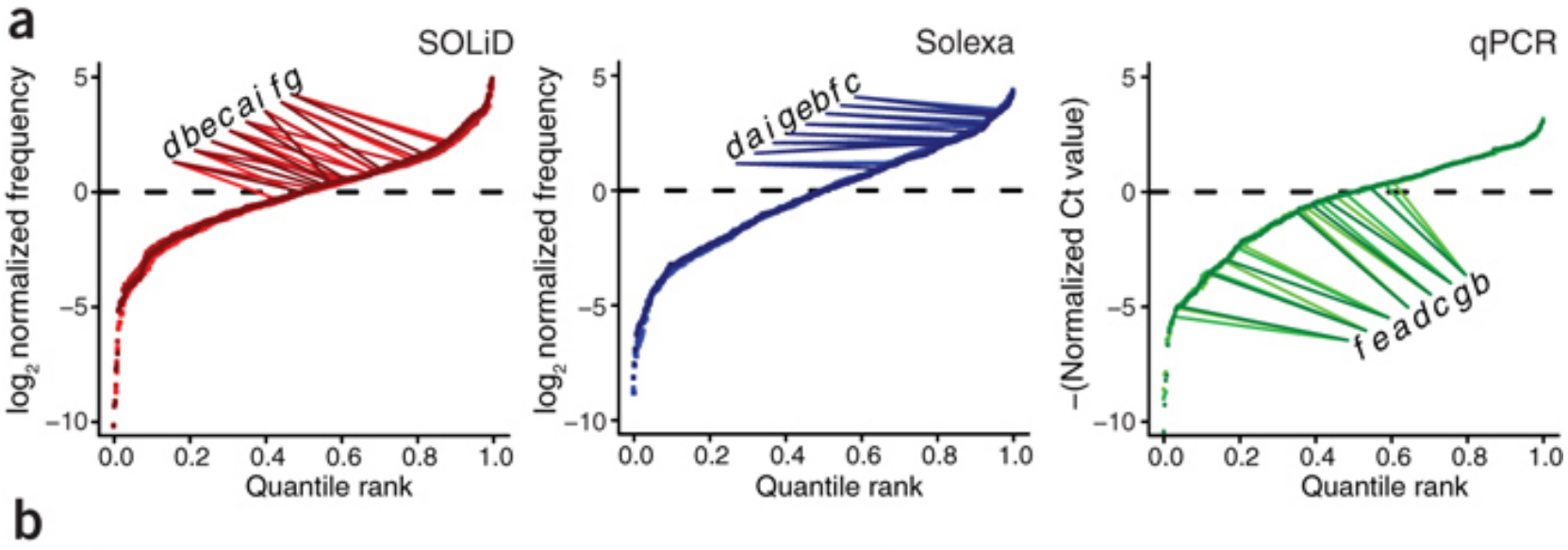
PolyA	(mRNA)
RiboMinus	(- rRNA)
Size <50 nt	(miRNA)
…..	

Size of fragment
Strand specific
5’ end specific
3’ end specific
…..

Single end (1 read per fragment)
Paired end (2 reads per fragment)

Paired end read

Different sequencing teqhniques have different preferences



Sequencing frequency of 472 artificial mircoRNAs in equal abundance

(Figure from Linsen *et al.,* Nature Methods. 2009)

But evens out over longer RNAs

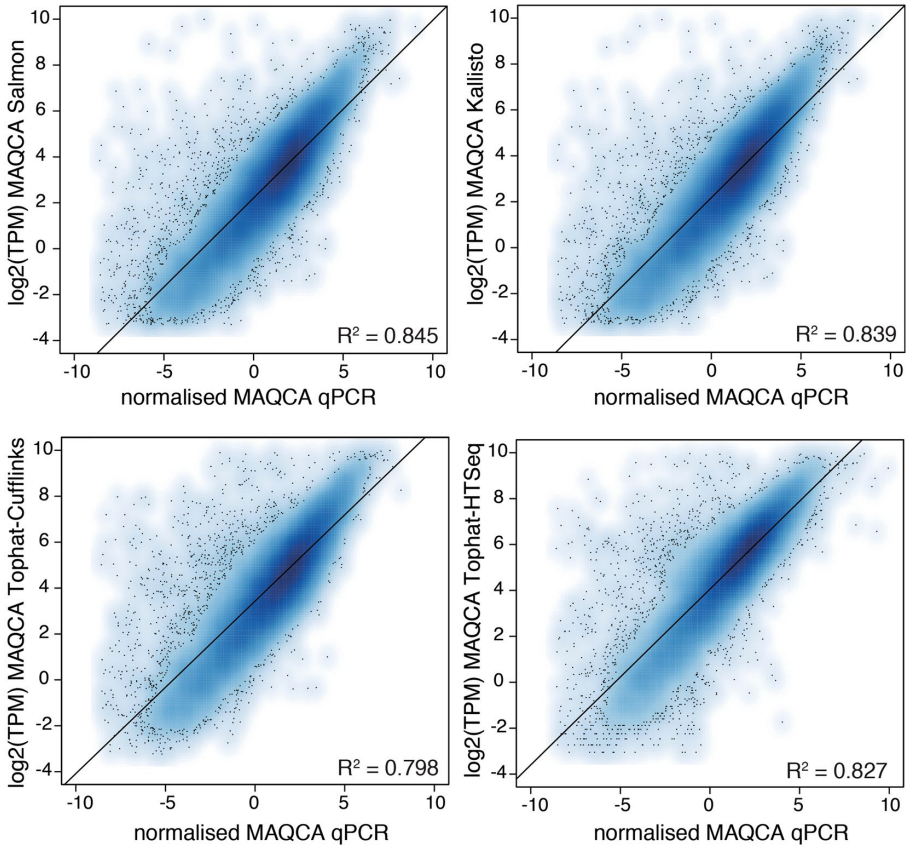
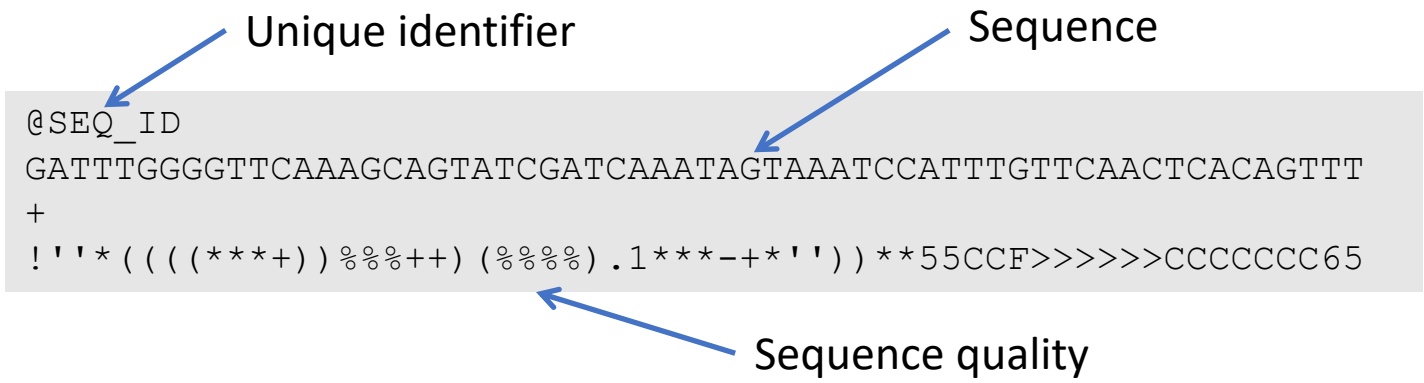


Figure 1. Gene expression correlation between RT-qPCR and RNA-seq data. The Pearson correlation coefficients and linear regression line are indicated. Results are based on RNA-seq data from dataset 1.

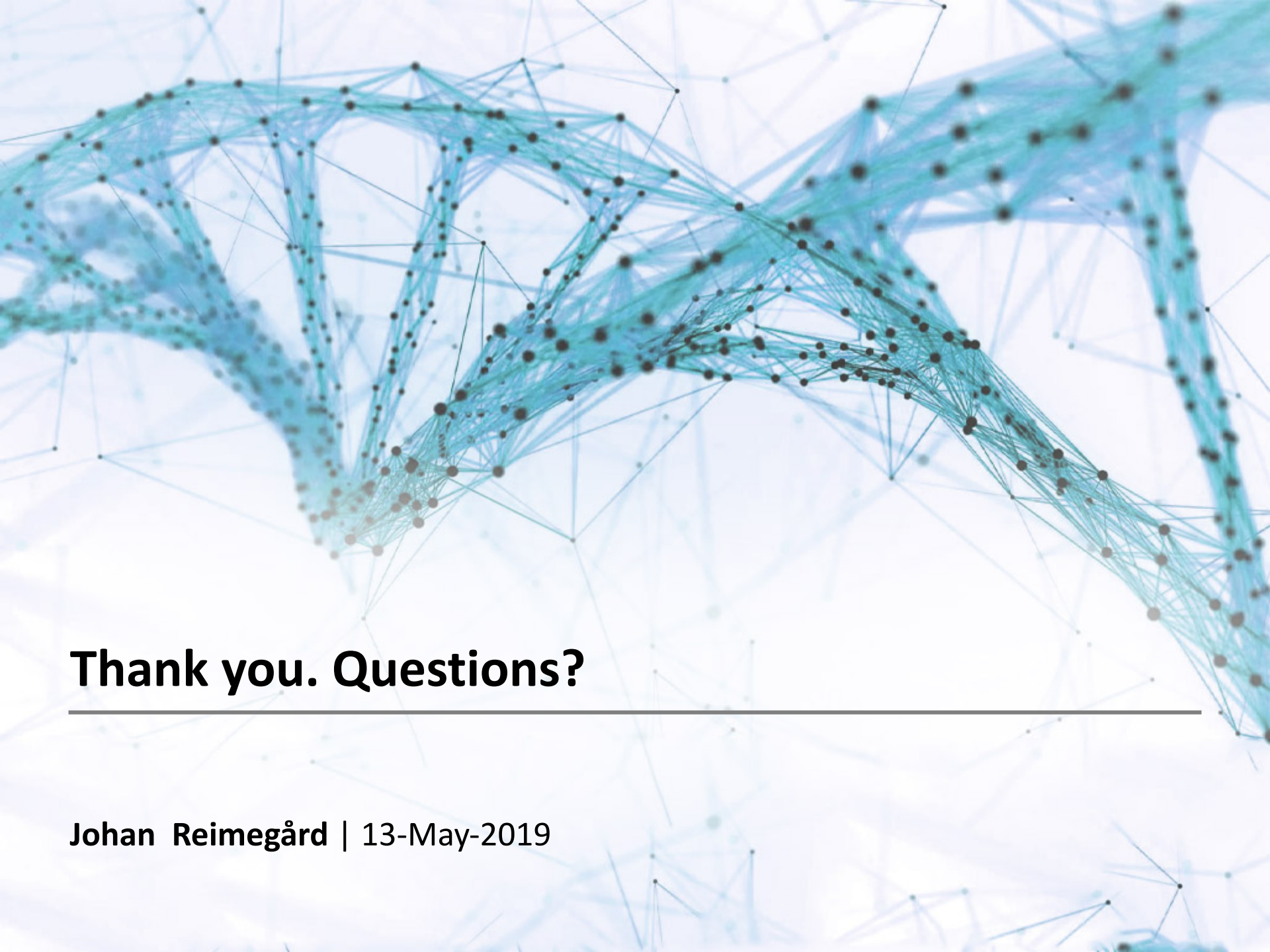
Benchmarking of RNA-sequencing analysis workflows using whole transcriptome RT-qPCR expression data

Fastq – read file format



Paired end data usually in format sampleX_1.fastq and sampleX_2.fastq with same SEQ_ID for both mate pairs, followed by /1 and /2 (or _f and _r)

Phret-score



Thank you. Questions?

Johan Reimegård | 13-May-2019