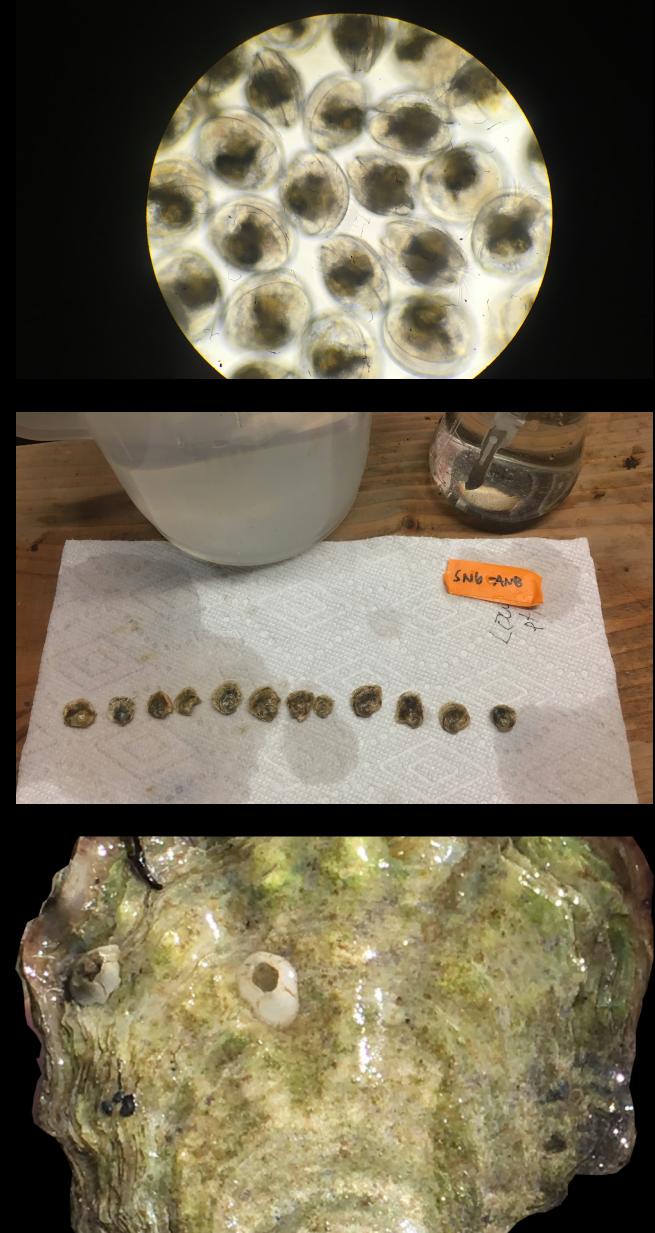


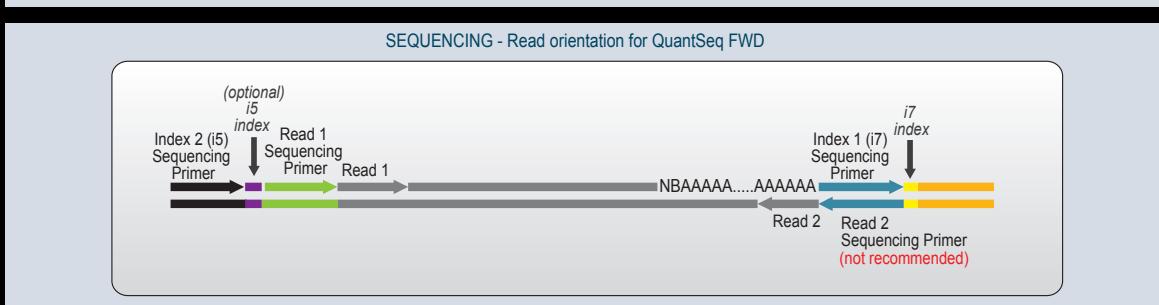
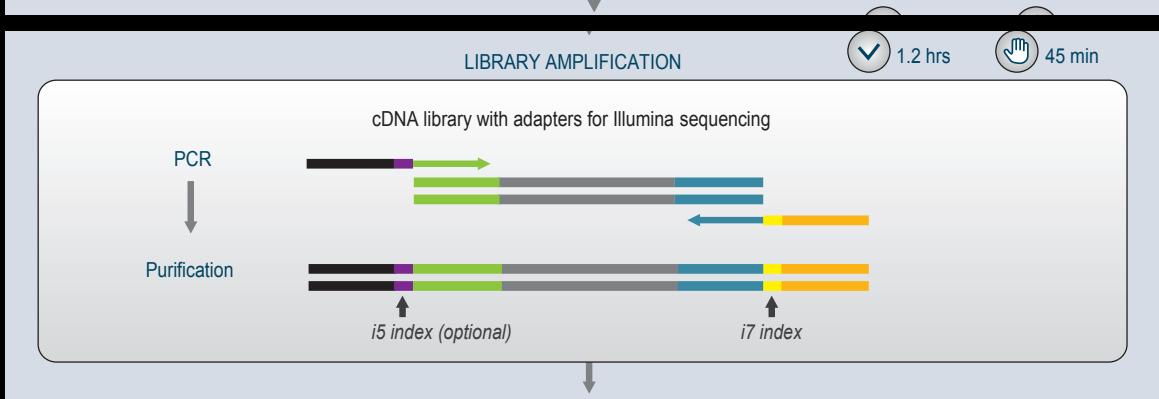
PHYSIOLOGICAL CARRYOVER EFFECTS OF HIGH pCO₂ EXPOSURE

QUANTSEQ 2020 PROJECT

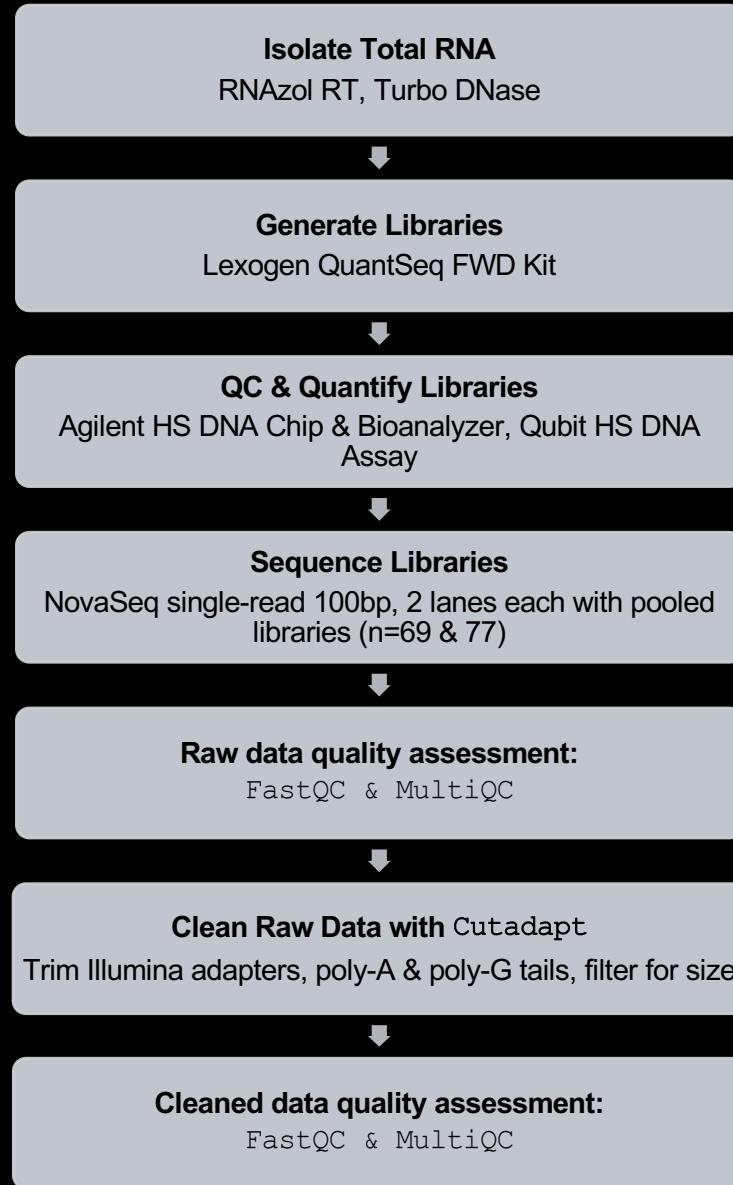
- Does/how does high pCO₂ exposure affect adult gill tissue gene expression?
- Does parental low pH exposure affect offspring gene expression
 - ~2wk-old larvae upon release
 - 15month-old juveniles at end of field trial in Port Gamble Bay, WA
- → Important: offspring never exposed to different pCO₂ conditions



QuantSeq / TagSeq Library Prep



Goal: generate high quality gene expression data (RNASeq)



Adult ctenidia
After pCO_2 exposure

Larvae, whole body, pooled by “pulse”
Upon release from brood

Juvenile, whole body
~15 month old, after 3-month field trial

My bioinformatics workflow

Part 1: Raw read processing

Goal: generate accurate gene count matrices

Option A:

Align reads to genome & count reads/gene

Align with Bowtie2 or STAR

Generate Bowtie or STAR index from genome

Generate *counts per gene* with featureCounts, genome feature files

My bioinformatics workflow

Part 2: translating reads to gene counts

Option B:

Generate counts without aligning

Salmon in mapping-based mode with selective-alignment

Generate Salmon index from mega-transcriptome. Use genome to create list of “decoy” sequences

Generate *counts per transcript* with Salmon, decoy-aware mega-transcriptome index

Annotate mega-transcriptome (denovo, v3) via Blast & Uniprot/Swissprot db

Summarize *counts per gene* via Tximport, annotated transcriptome

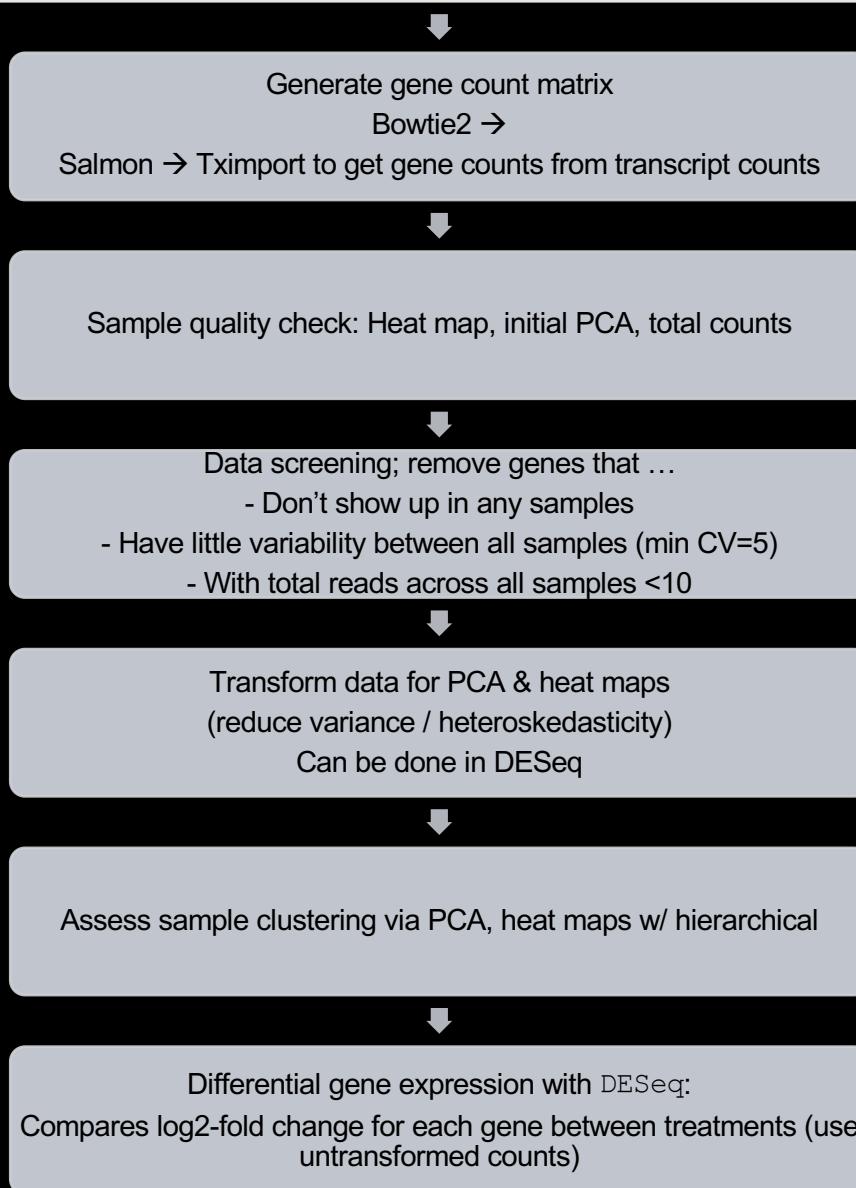
Goal: Identify best read mapping method

I tested 8 of my 144 samples on three different read generation methods:

- Alignment with STAR against Oly genome v081, annotate with GTF of exons with isoform information (from Stringtie)
- Alignment with Bowtie against Oly genome v081; annotate with GFF of genes+2kb downstream
- Selective alignment with Salmon against transcriptome v3 that is “decoy-aware” (decoy sequences from Oly genome); annotate by blastx with Uniprot db

SUMMARY of alignment methods	STAR	Bowtie	Salmon
Total no. input reads	47,995,645	47,995,645	47,995,645
Average % uniquely mapped reads	58.22	41.4	NA
% Multi-mapped reads	4.09	30.4	NA
Overall alignment rate	62.3%	71.8%	39.2%
# aligned reads	30,145,597	33,937,733	19,818,281
# genes ID'd after filtering	59,023	29,775	5,741
# counts in genes (after pre-screening)	18,725,025	18,522,641	5,907,371
% counts in genes / total reads analyzed	39.0%	38.6%	12.3%

Goal: Identify differentially expressed genes among parental treatments, cohorts



My bioinformatics workflow

Part 3: Comparing gene expression among treatments