

# **Gene expression & transcriptomics: Adaptive evolution**

Joanna Kelley

ConGen2018



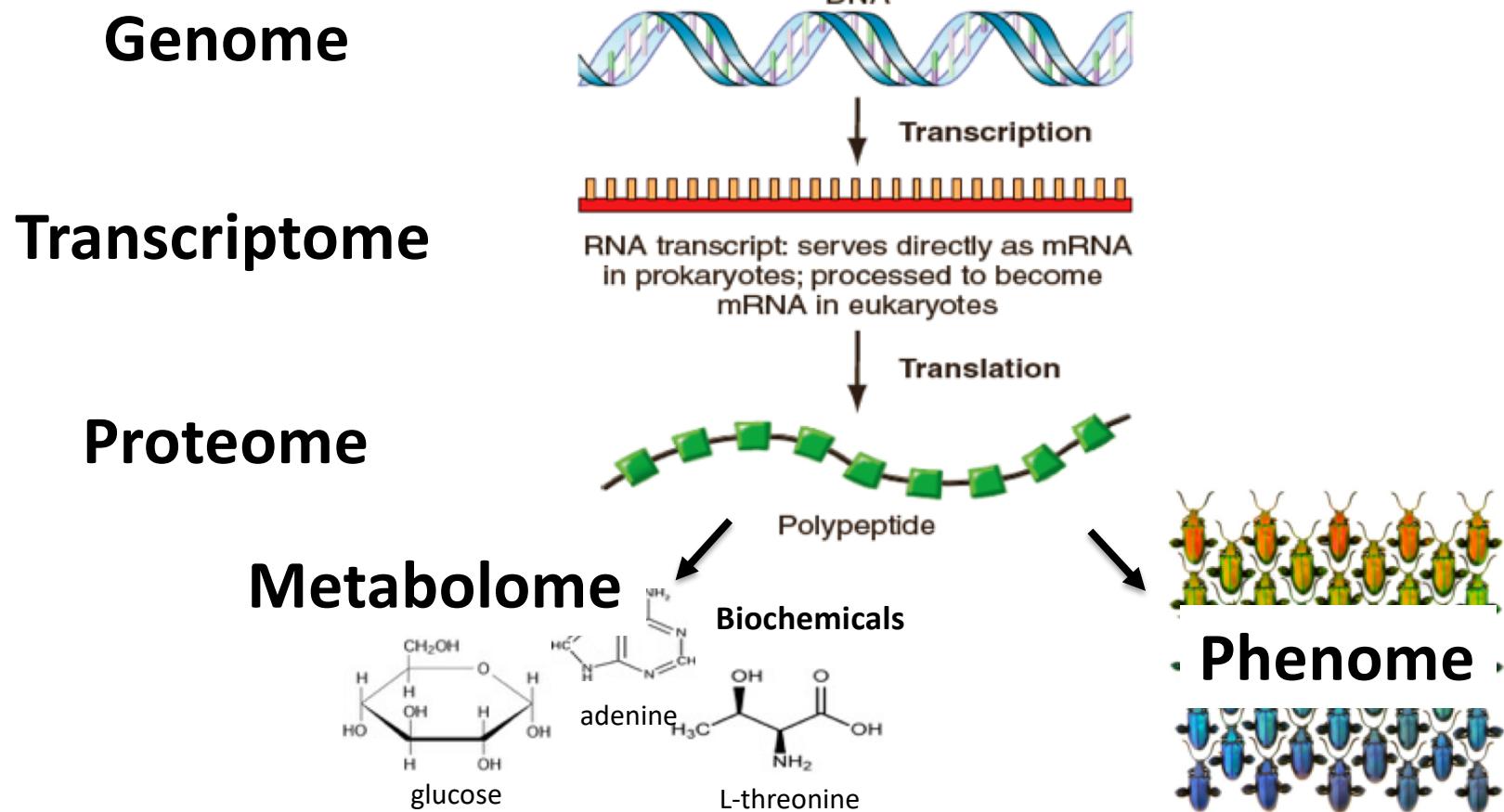


How do organisms diverge and adapt to the wide-range of environments they encounter?

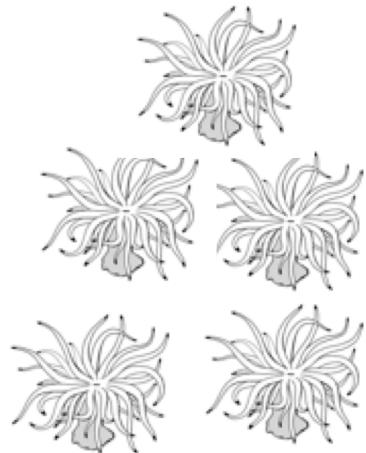


D. Scott Taylor

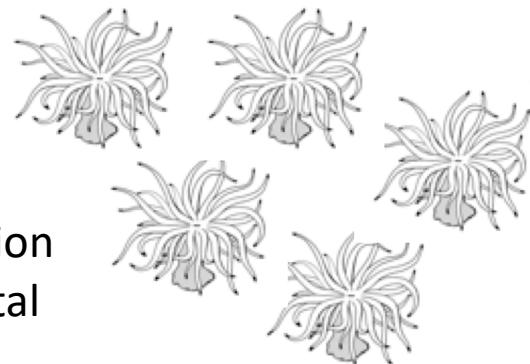
# OMICS! approaches



# Why study gene expression differences among individuals and populations?



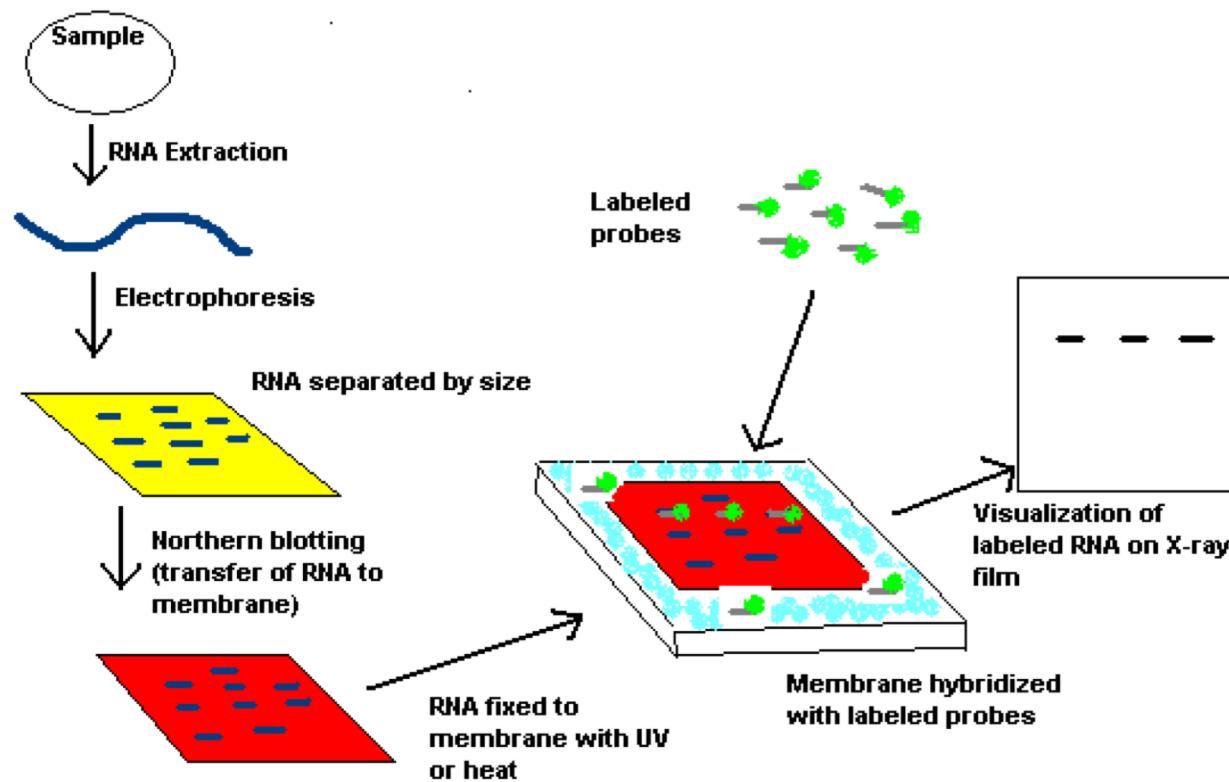
- Insights into the molecular basis of phenotypic diversity
- Interpretation of patterns of expression variation in response to environmental conditions, disease, etc.
- Possible management decisions on how and where to manage or transplant populations



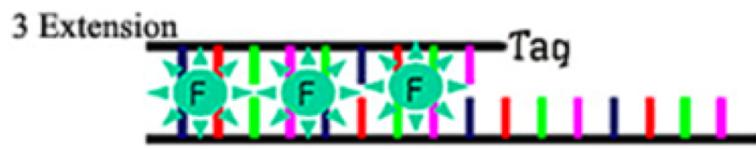
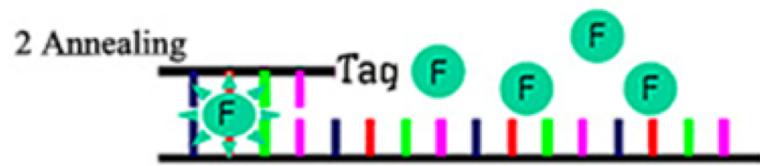
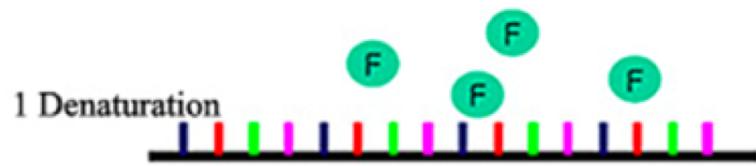
# Approaches to measuring gene expression

- Single/few gene studies
  - Northern blots
  - qPCR
- Transcriptome (everything that is transcribed at a single time point in a specific tissue/cell)
  - Microarrays
  - RNA-sequencing (RNA-seq)
    - poly-A+
    - Ribo-minus

# Northern blot to measure gene expression

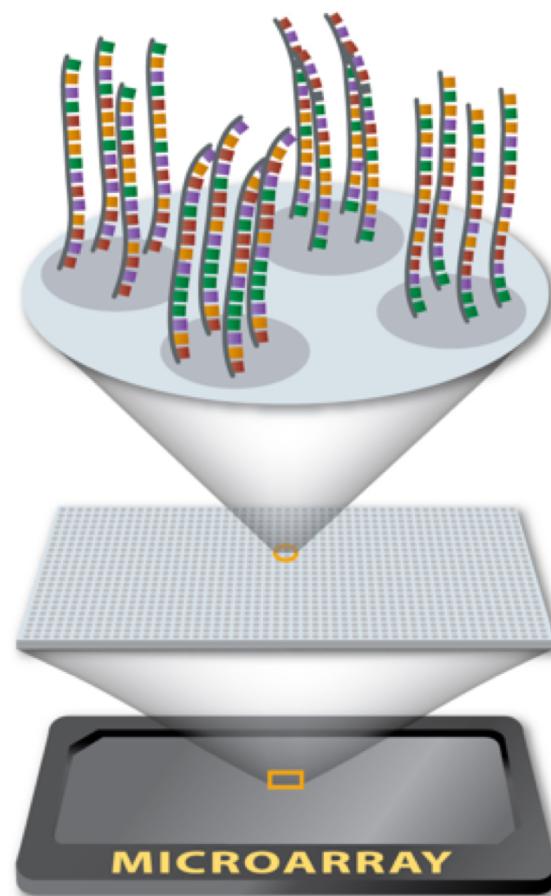
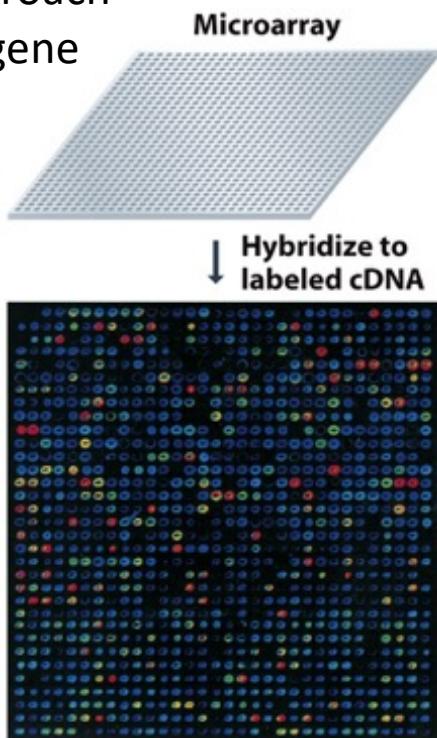


# qPCR on target genes

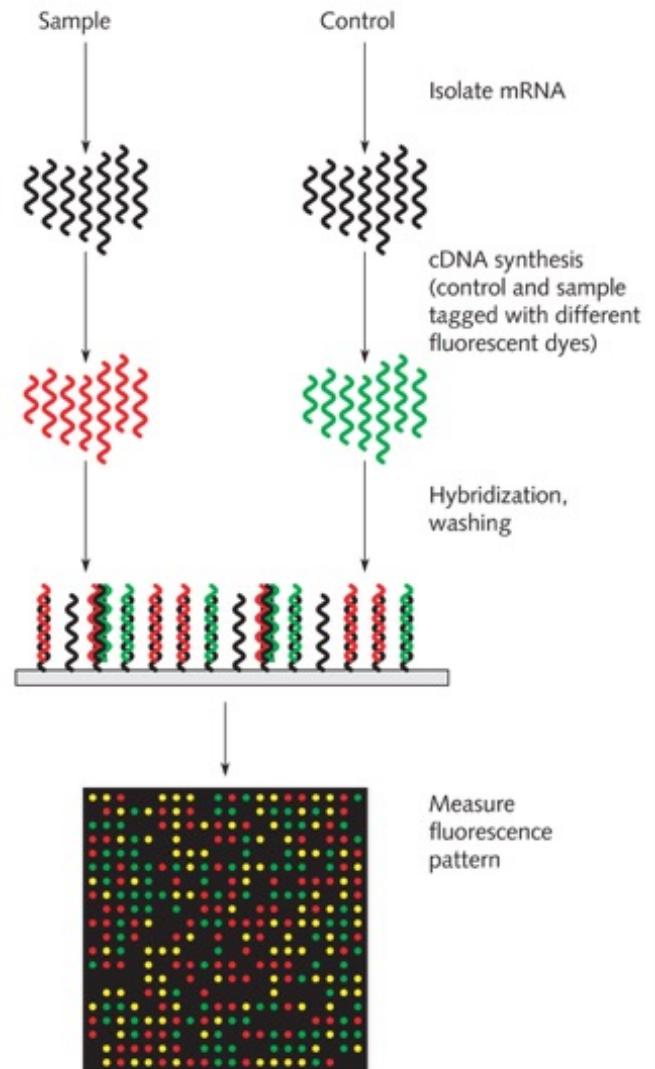


# Microarrays

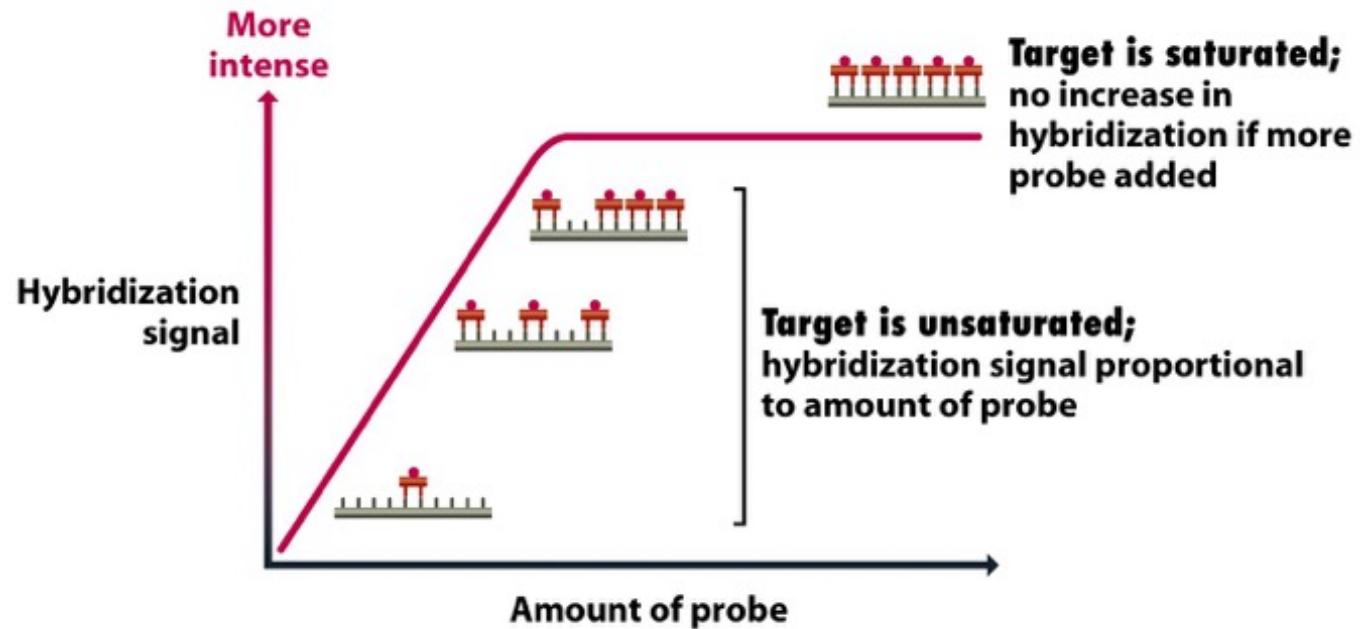
Microarrays:  
Array based approach  
to measuring gene  
expression



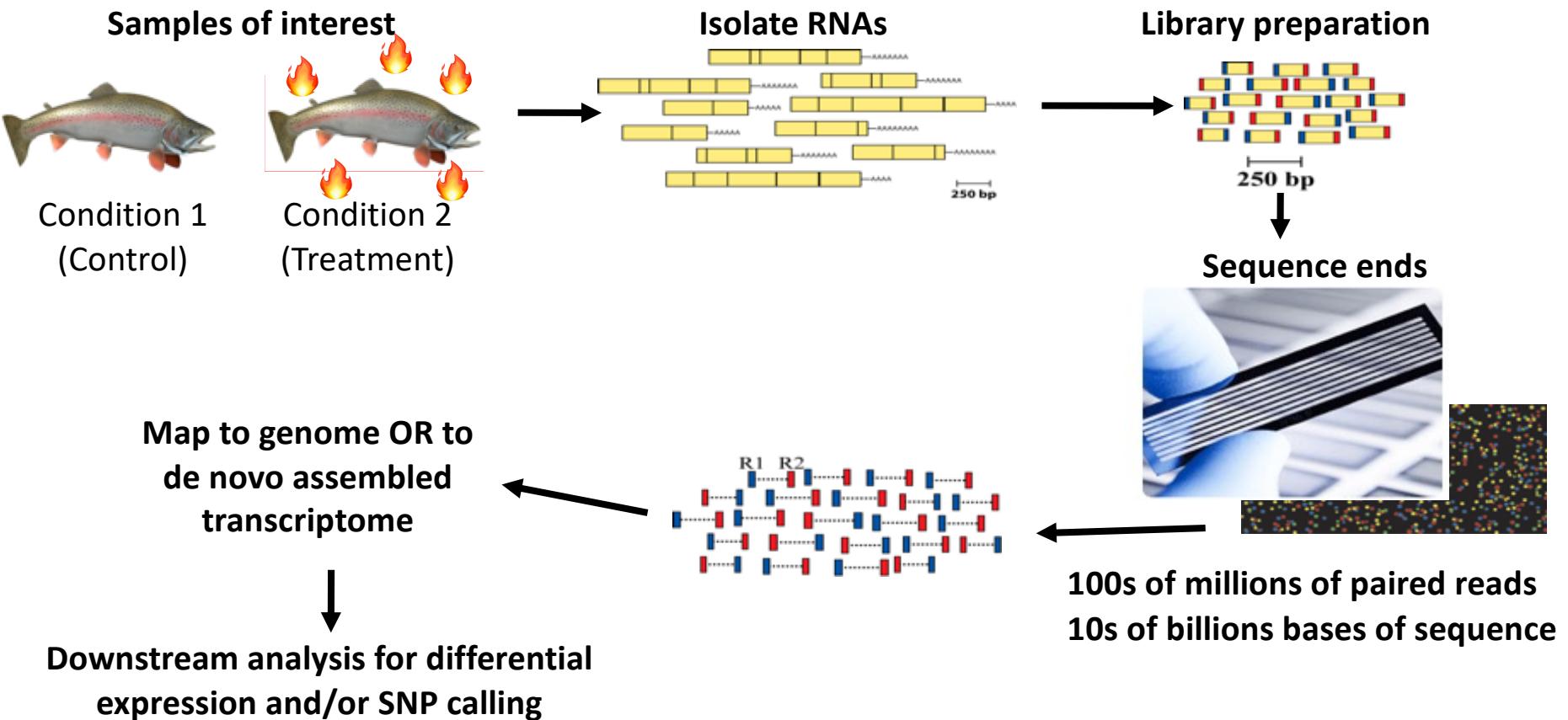
# Microarrays



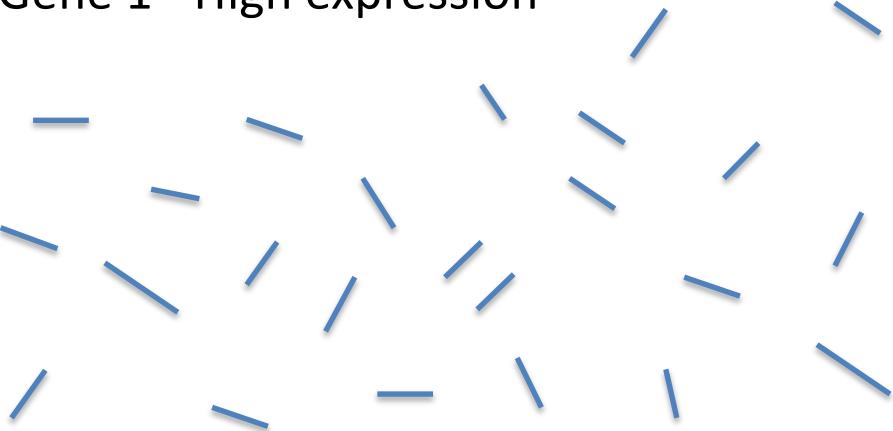
# Microarrays



# RNA sequencing experiment



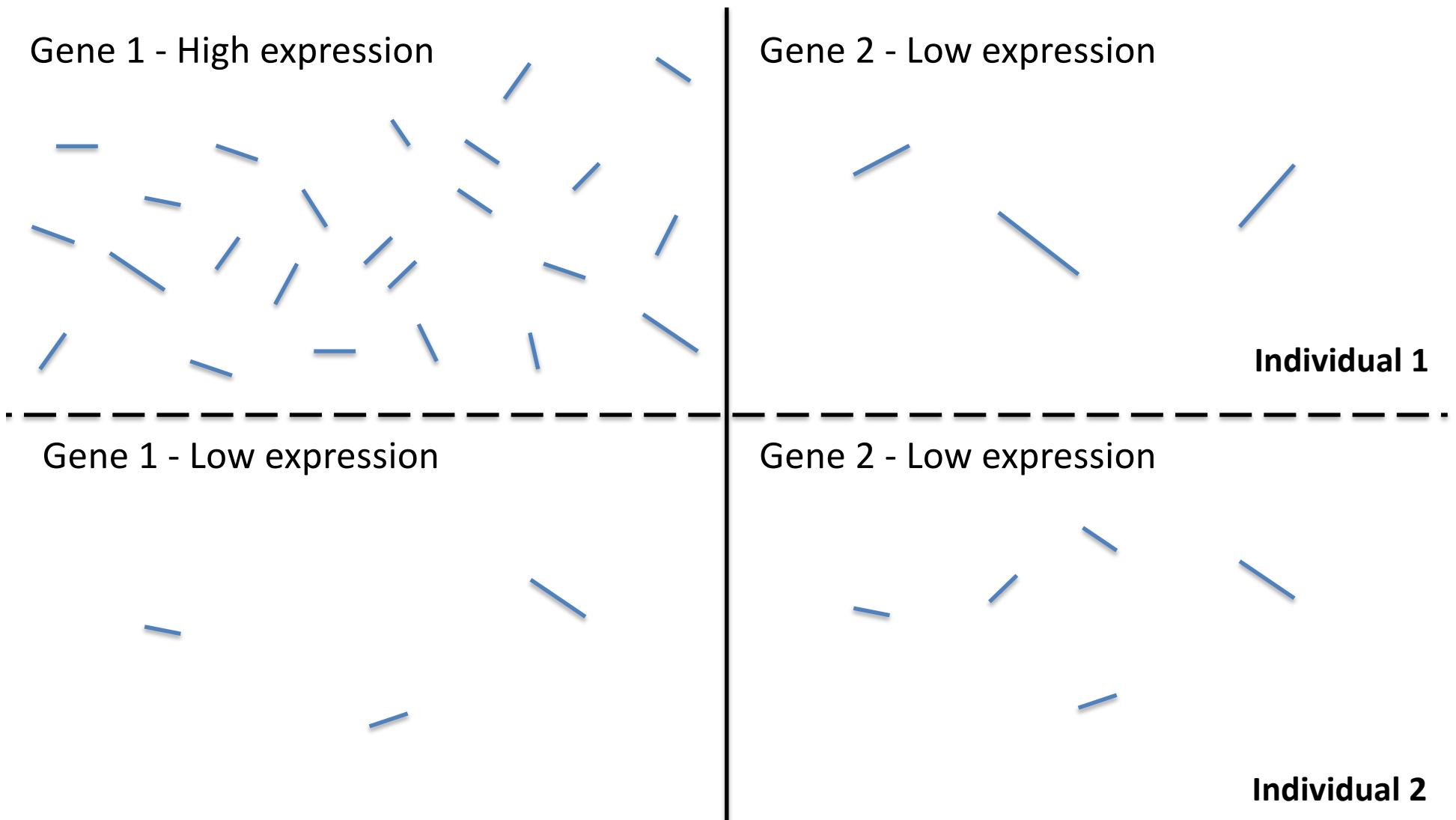
Gene 1 - High expression



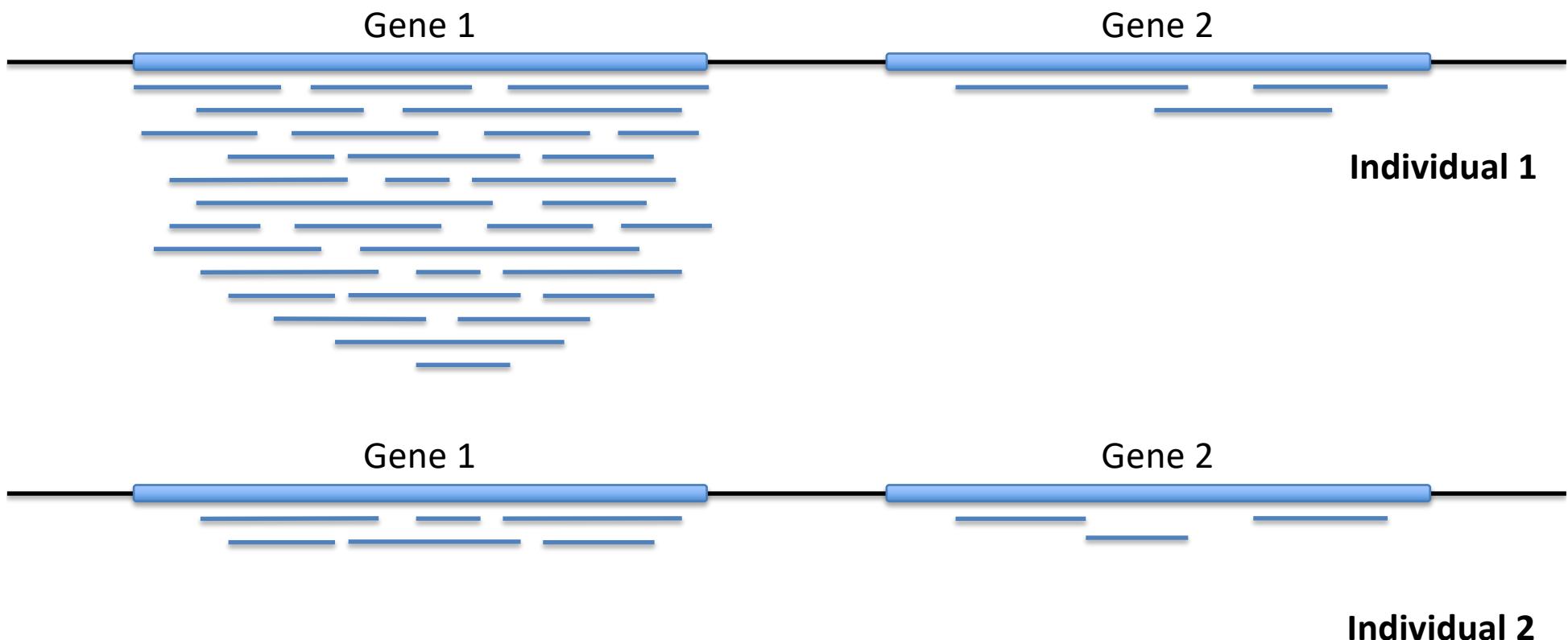
Gene 2 - Low expression



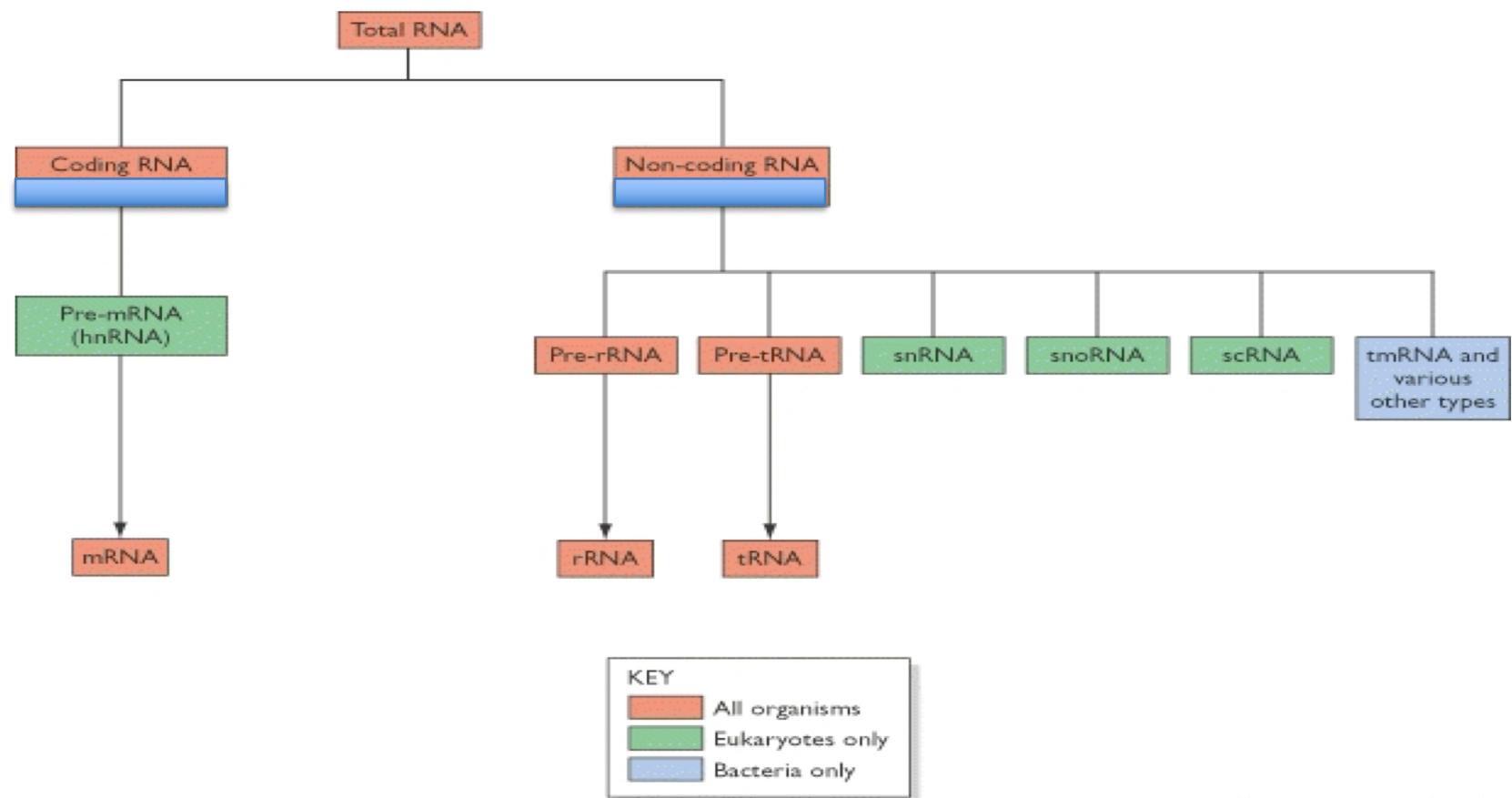
Individual 1



# RNA-Seq reads pile up higher on genes that are highly expressed



# Components of total RNA



# Extracting total RNA

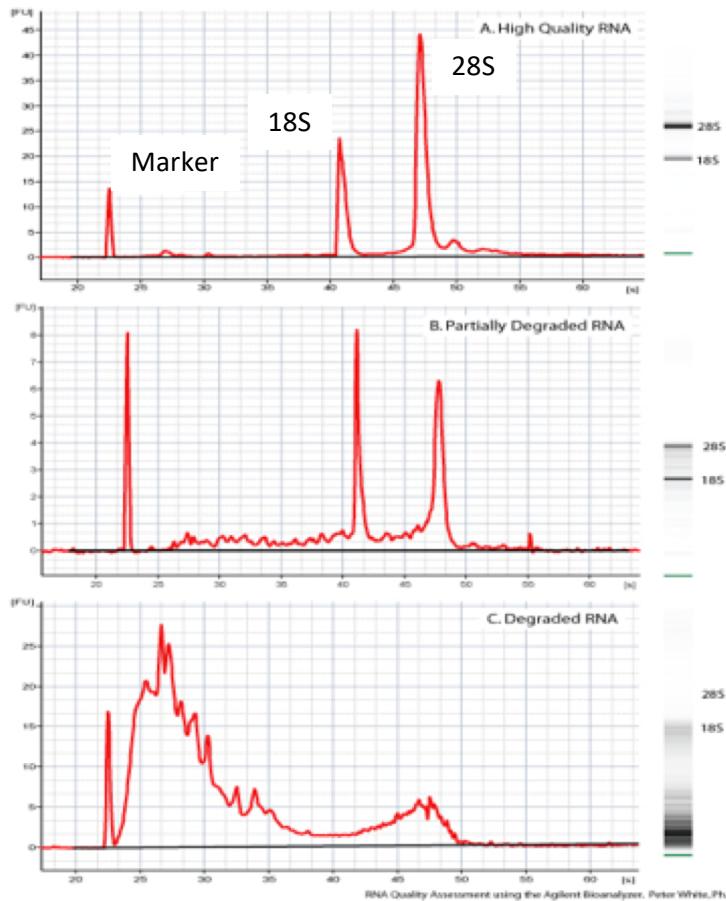
What would total RNA look like if we ran it on a gel?

High quality RNA?

Slightly degraded RNA?

Degraded RNA?

# Degraded total RNA?

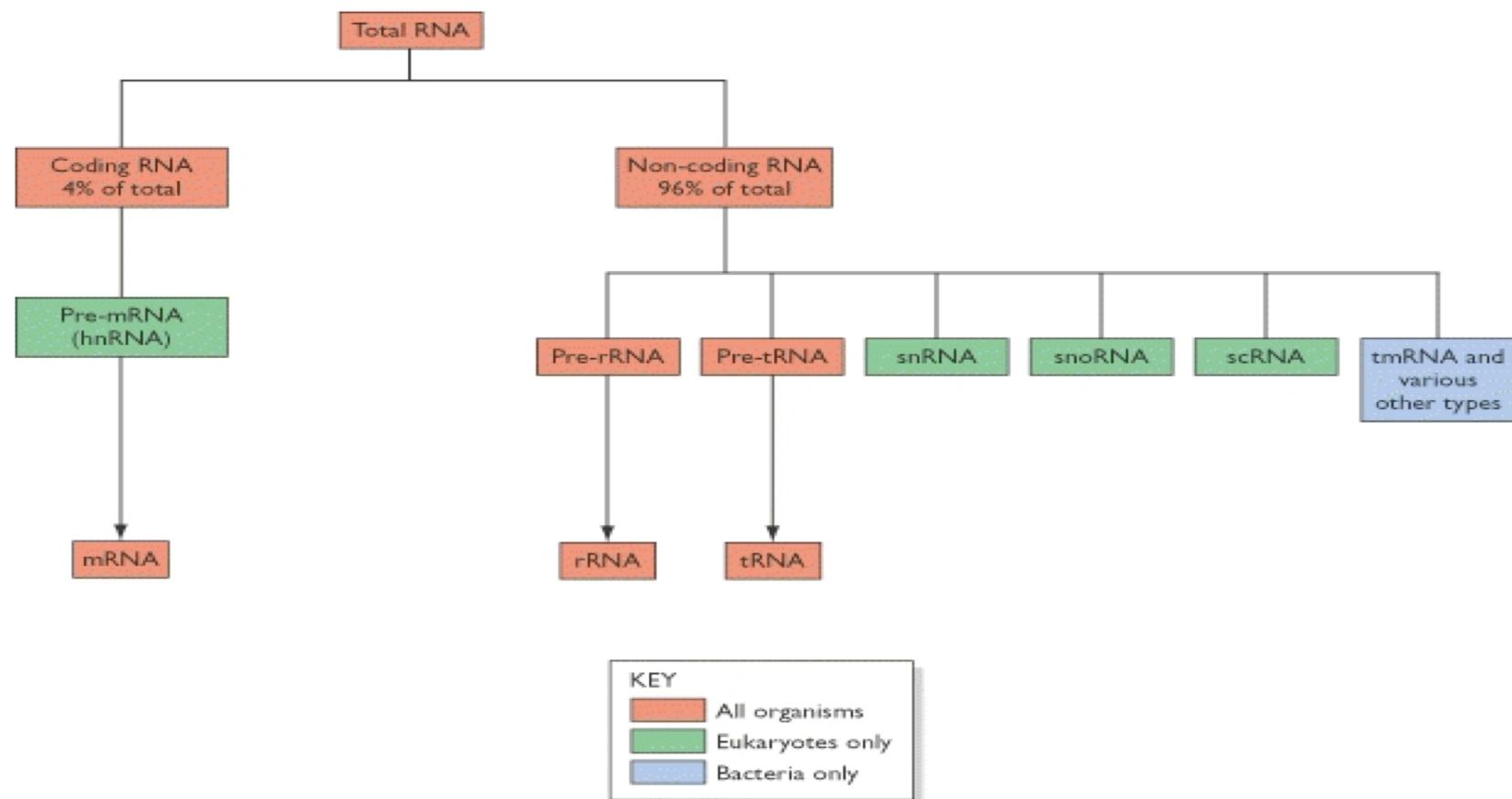


High quality RNA

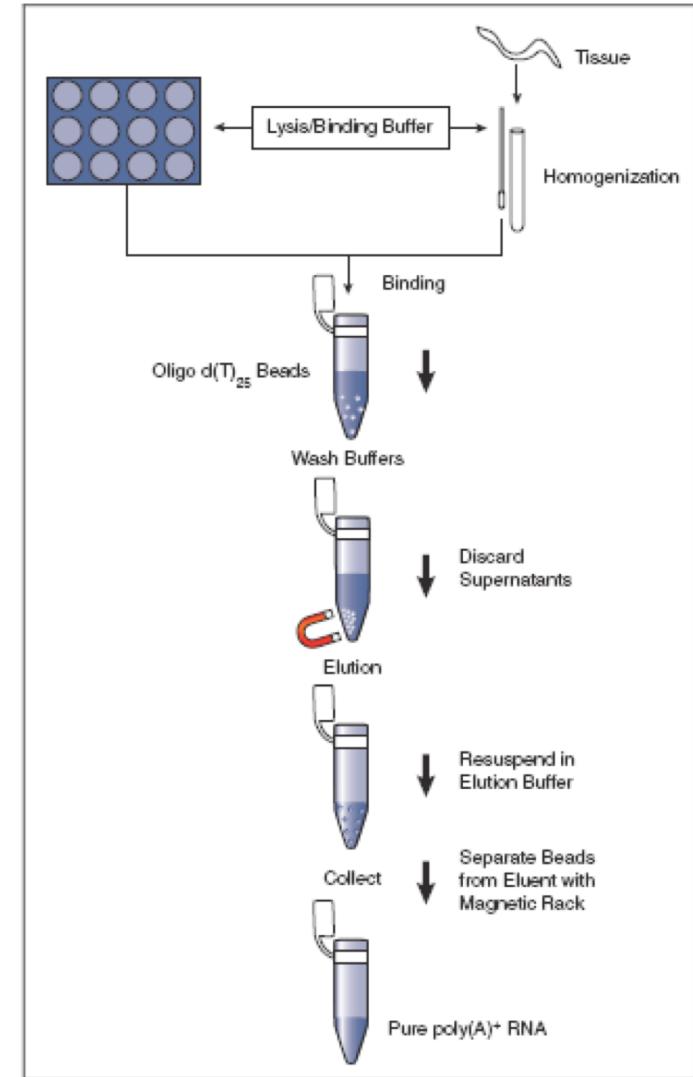
Slightly degraded RNA

Degraded RNA

# How to enrich for the RNA that we are interested in?

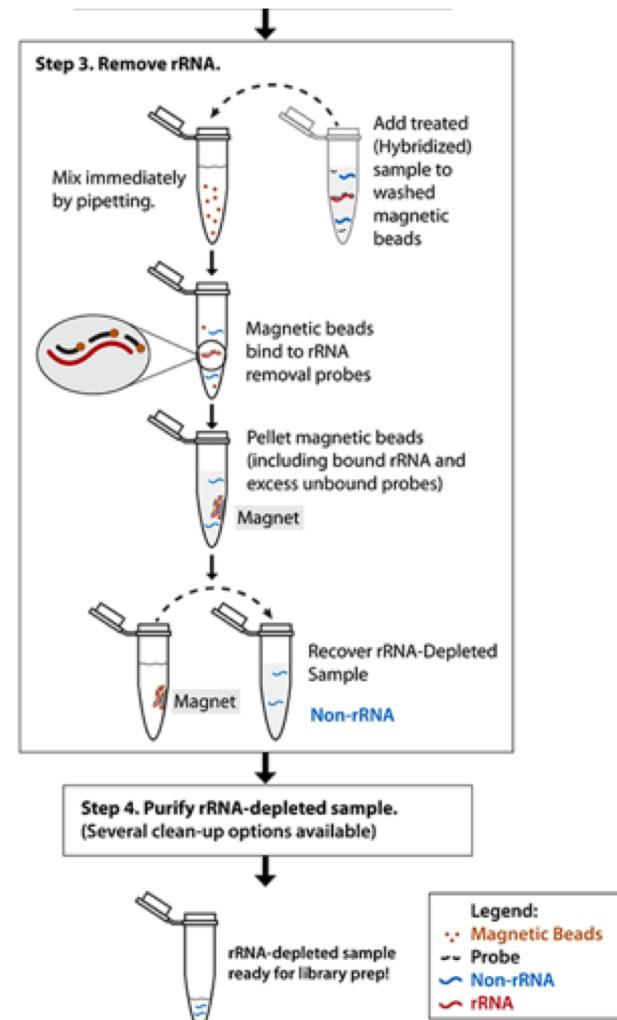
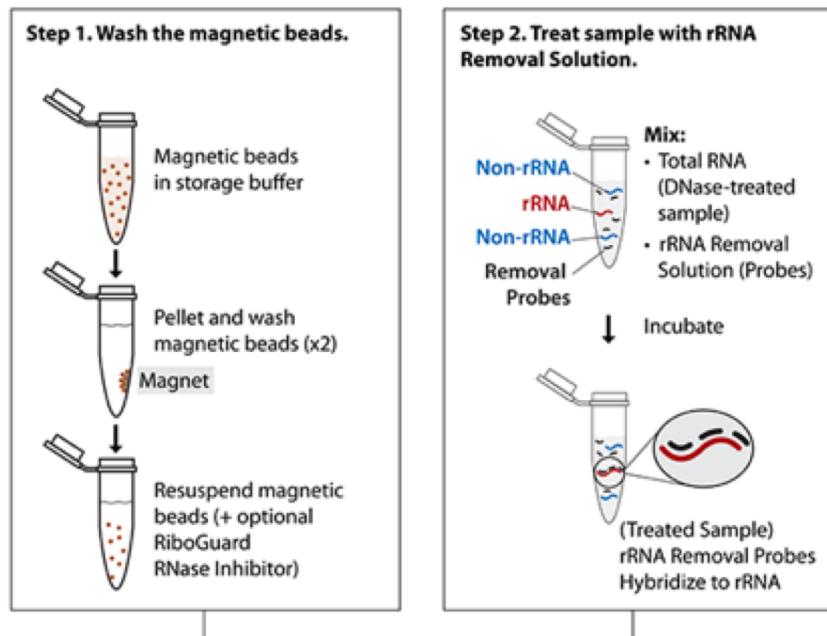


# Poly-A+ selection

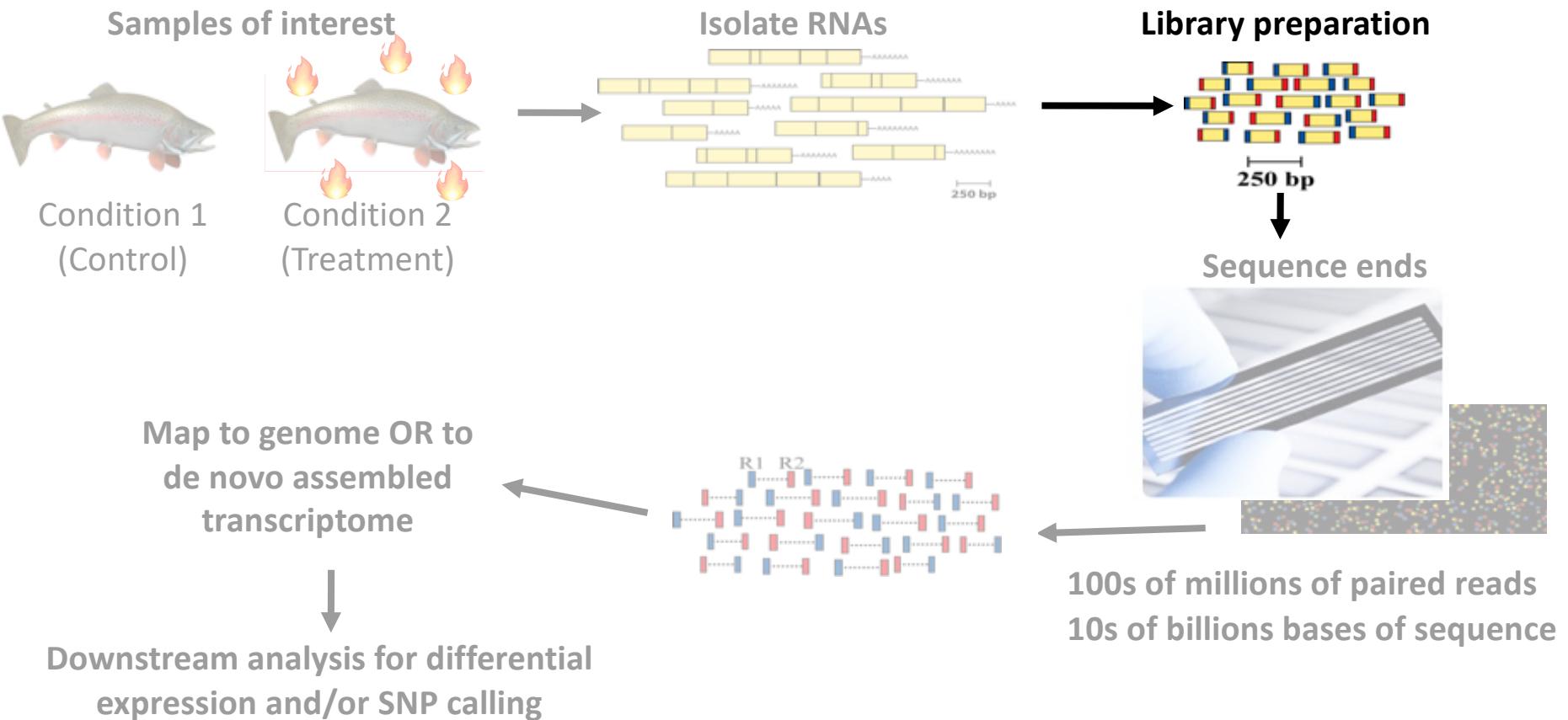


# Ribosomal depletion

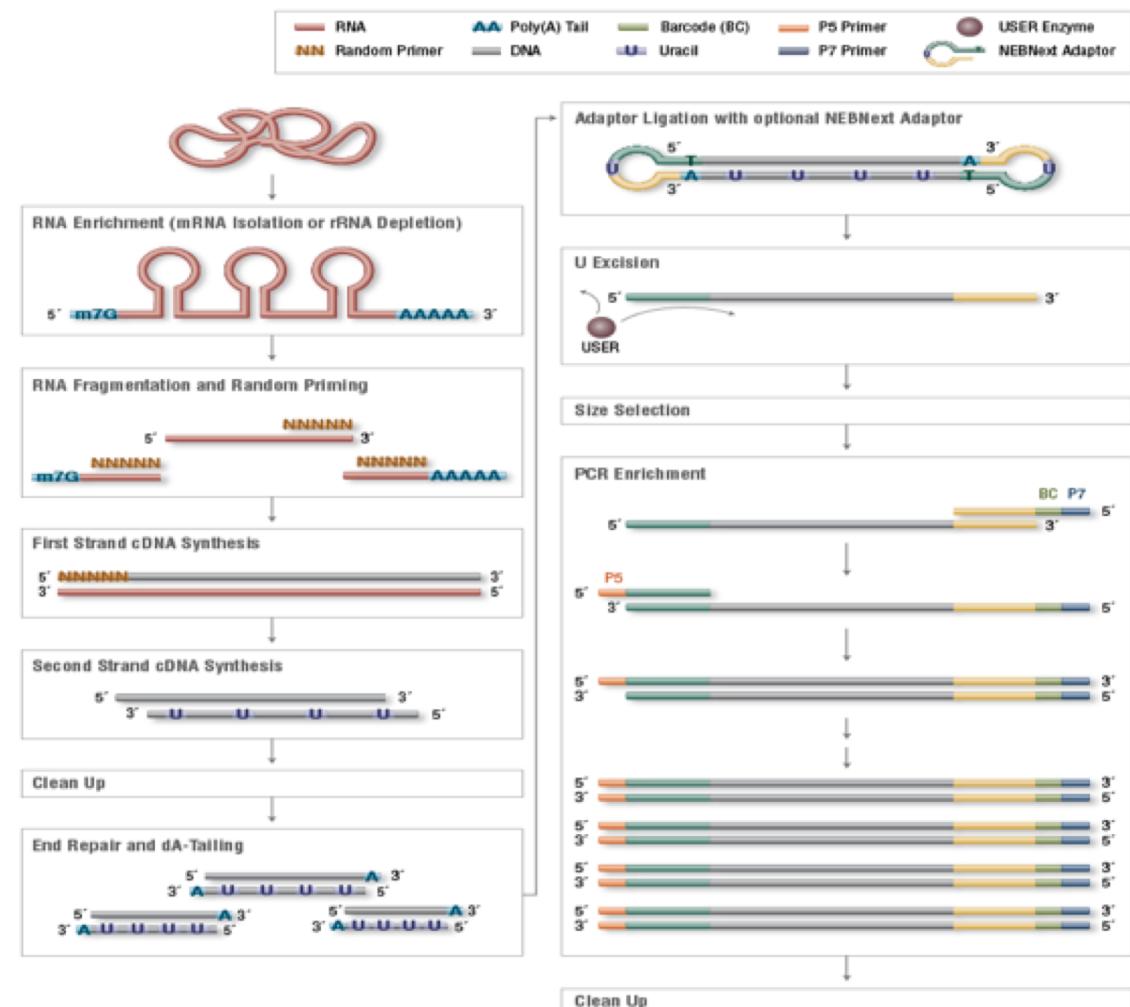
## Ribo-Zero Workflow (4-Steps)



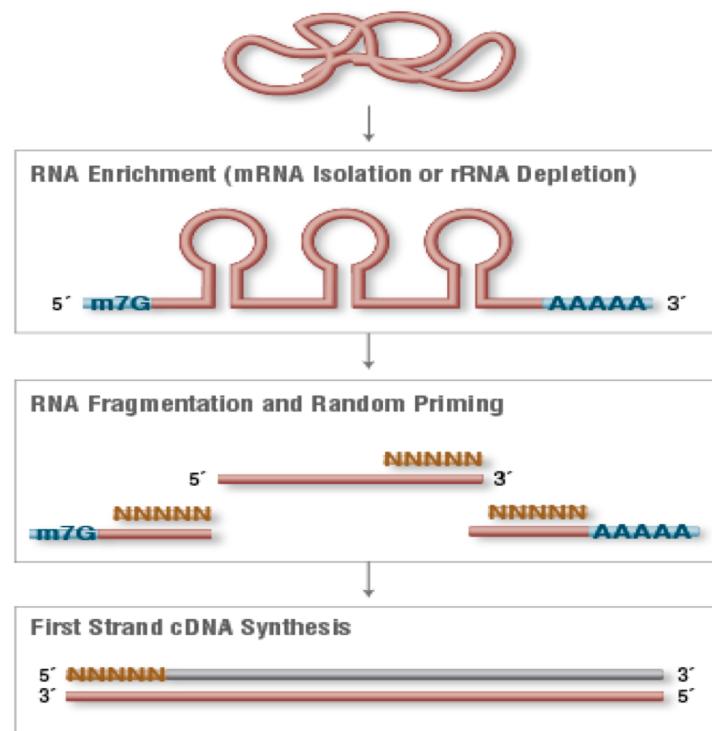
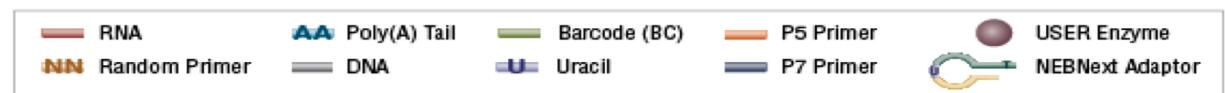
# RNA sequencing experiment

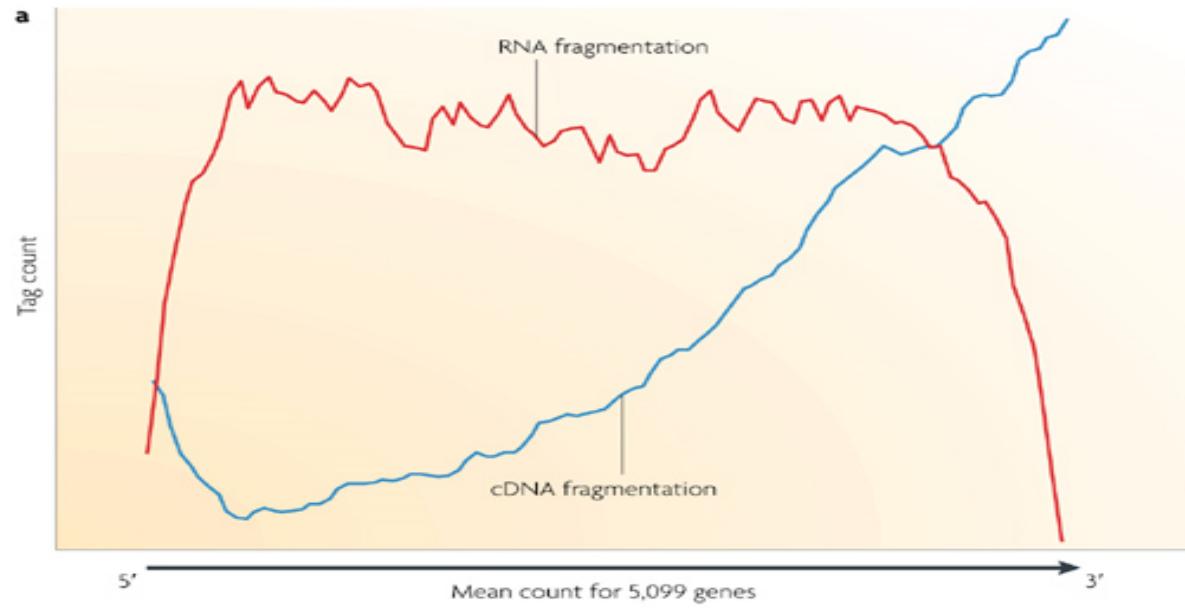


# RNA-seq Library preparation process

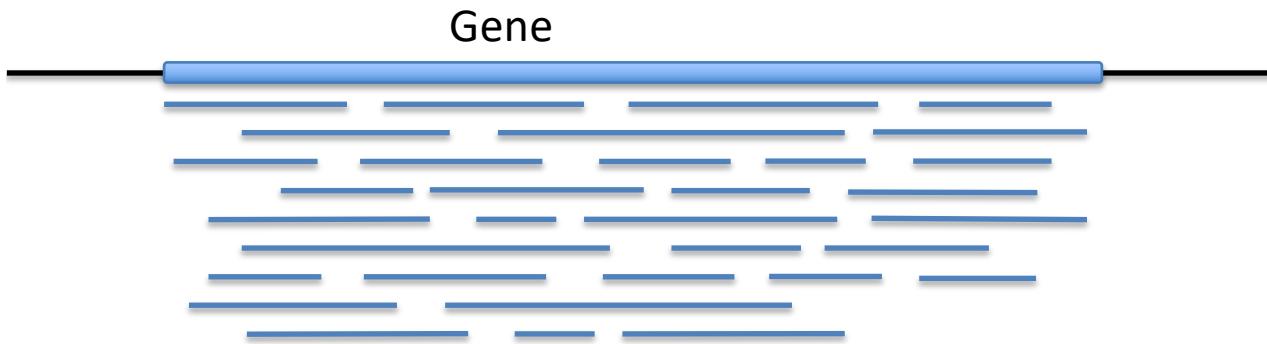


# Steps in the library prep process





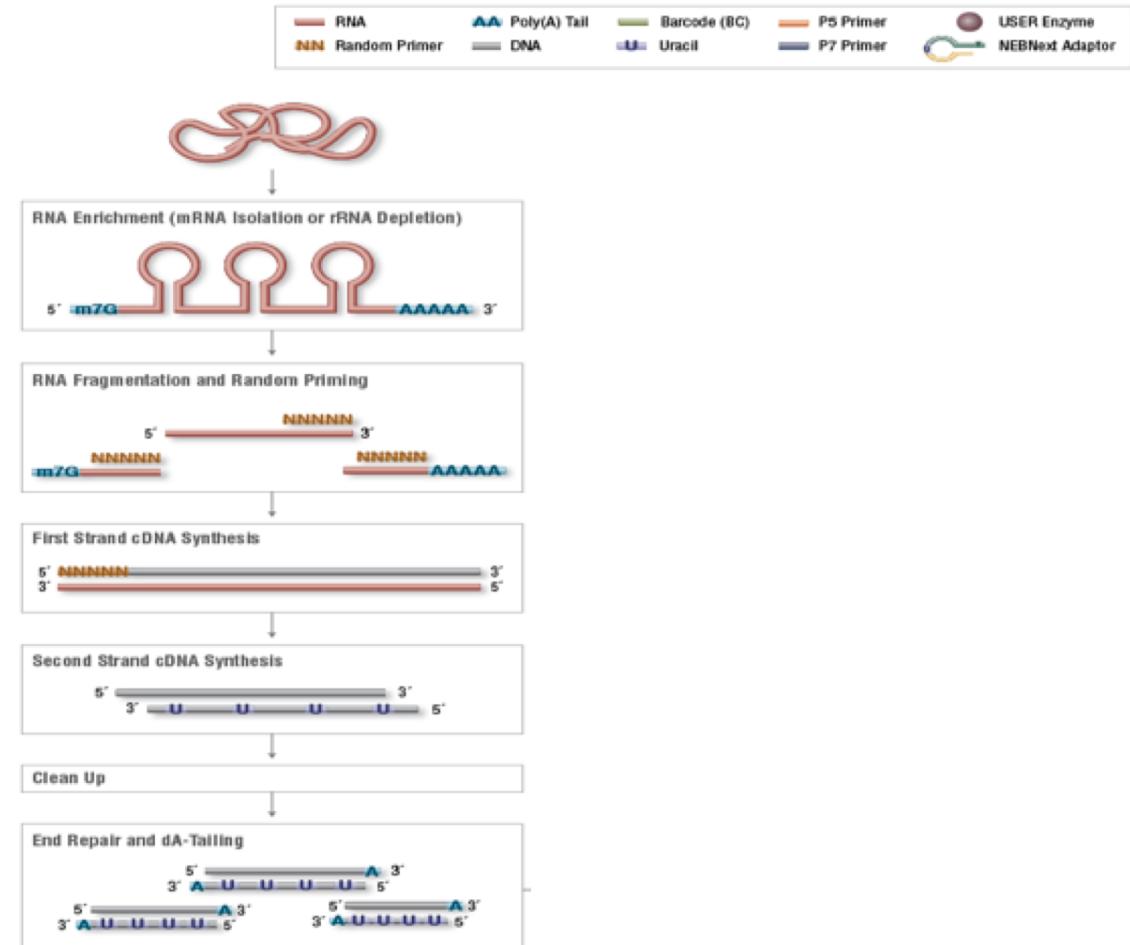
**Fragmentation** of oligo-dT primed cDNA (blue line) is more biased towards the 3' end of the transcript. RNA fragmentation (red line) provides more even coverage along the gene body, but is relatively depleted for both the 5' and 3' ends.



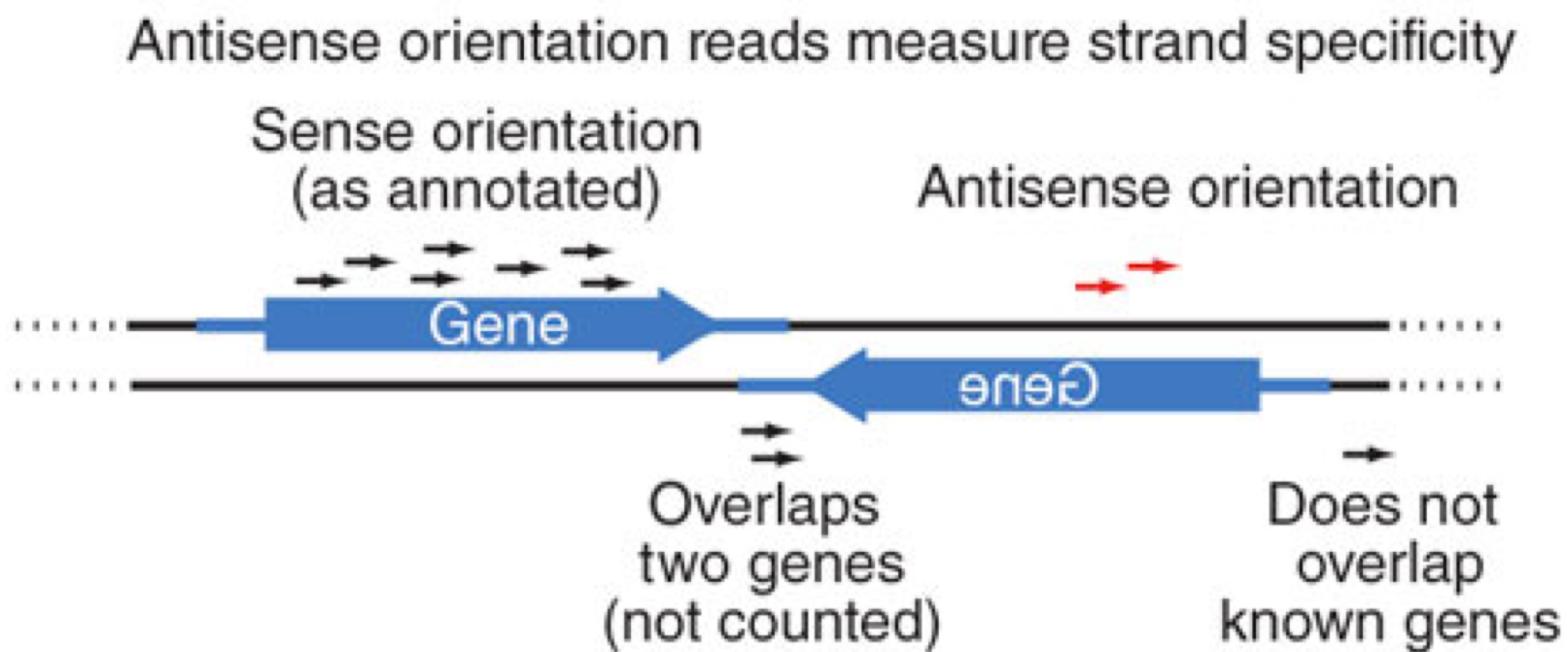
Wang et al 2009

# RNA-seq Library preparation process

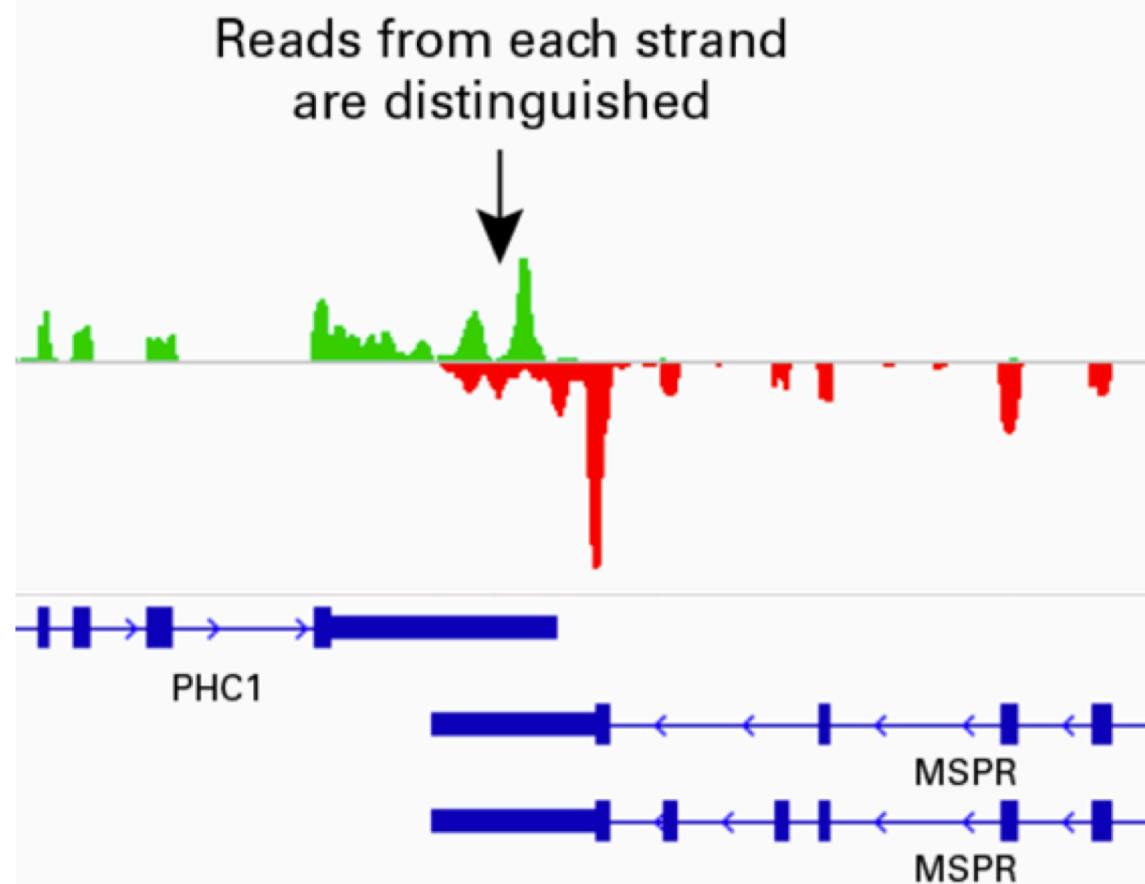
Directional  
(Stranded) vs  
non-directional



## Possible insights from a directional RNAseq experiment

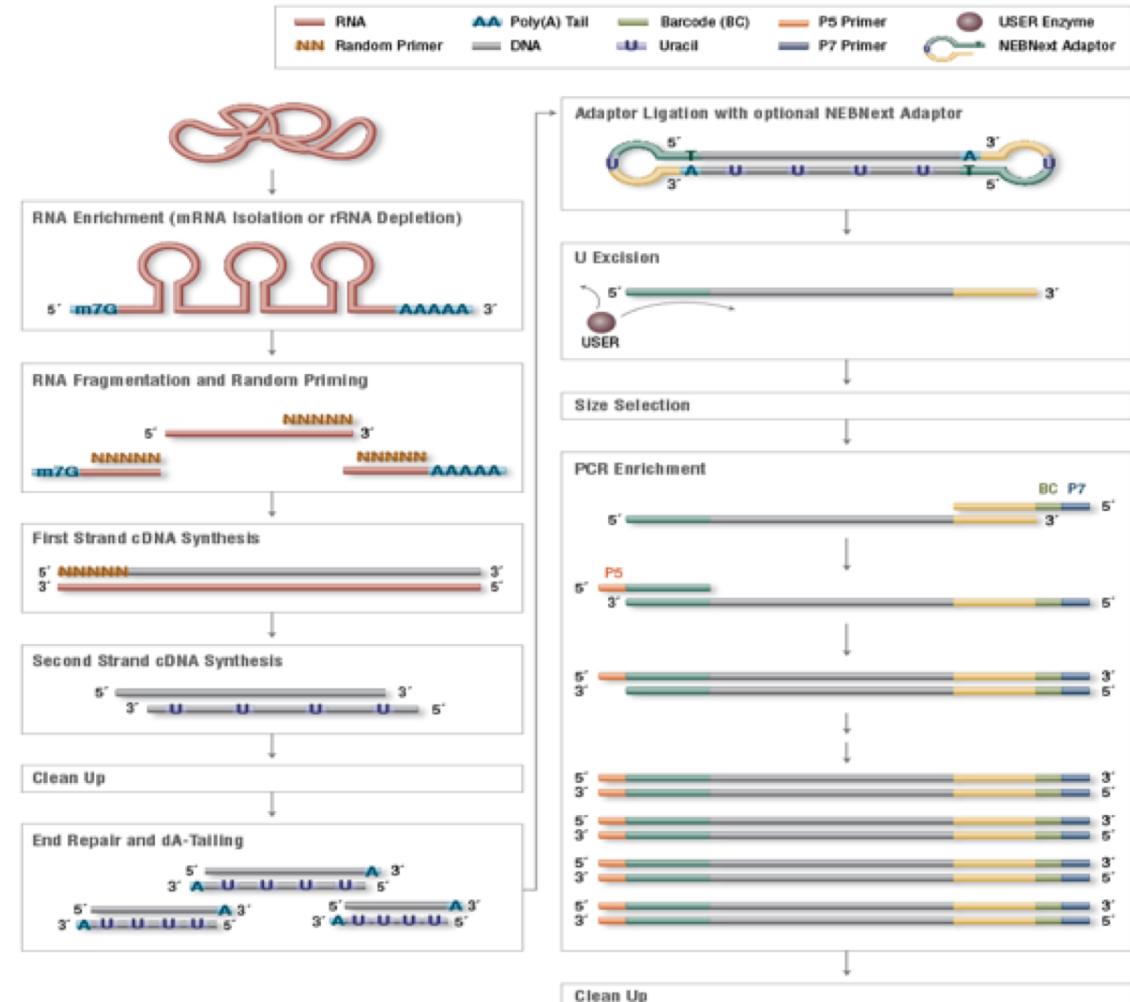


## Possible insights from a directional RNAseq experiment

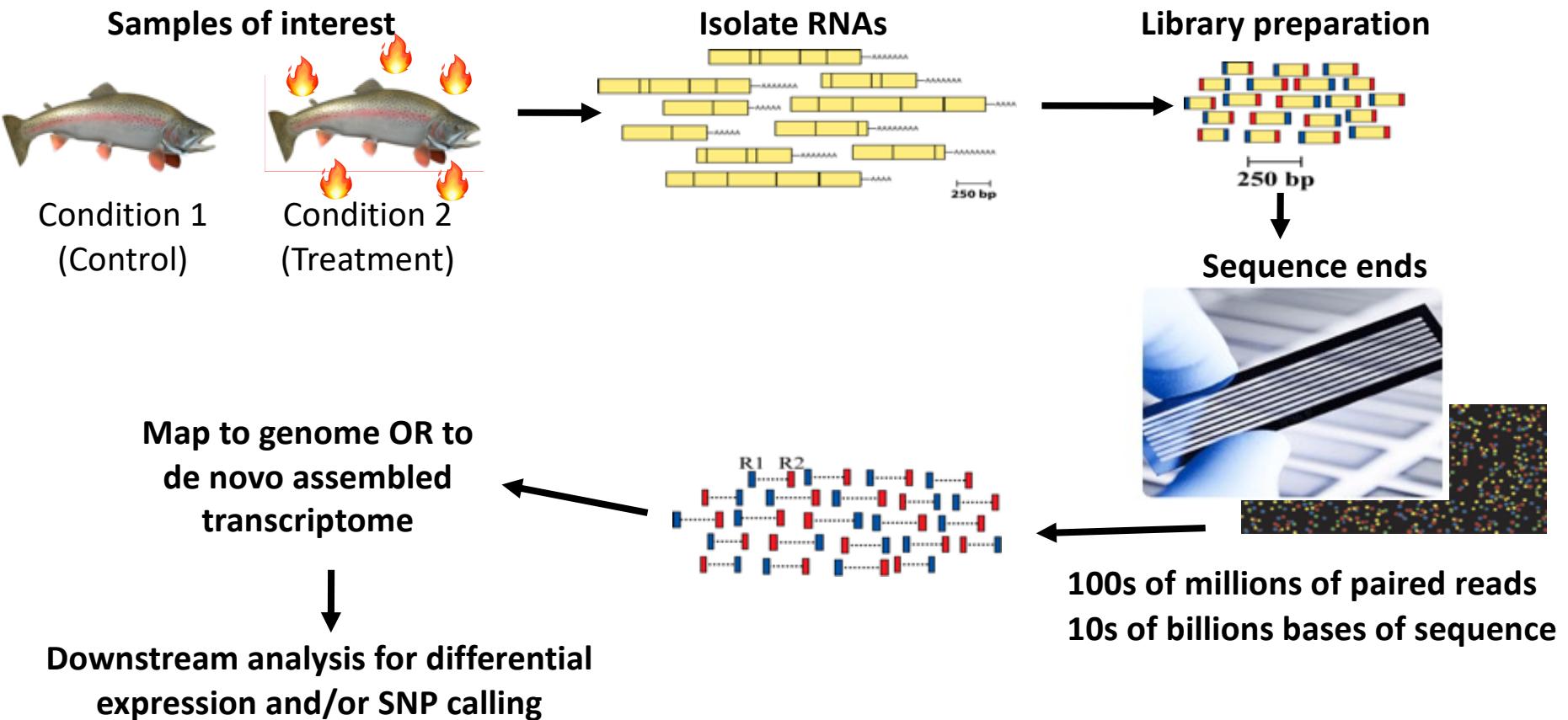


# RNA-seq Library preparation process

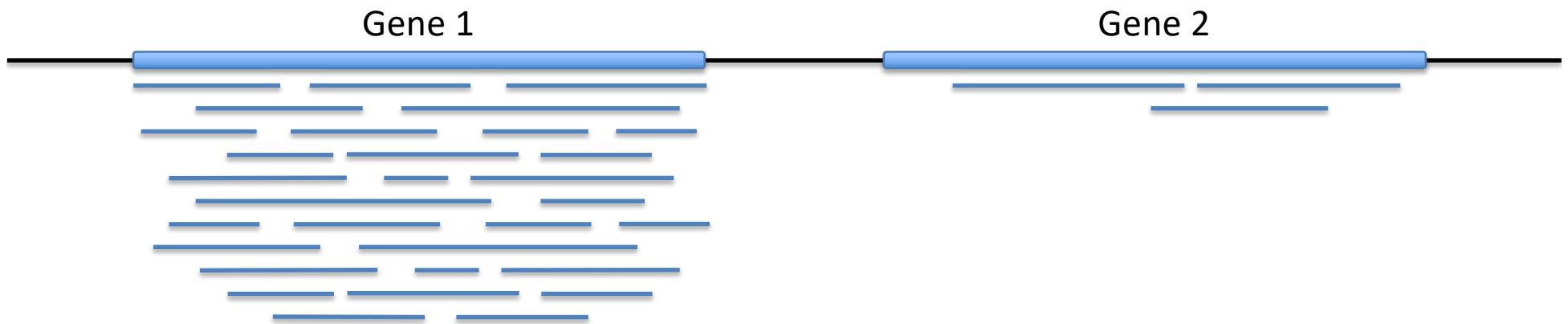
- Adapter ligation
- Barcoding
- Amplification



# RNA sequencing experiment



# Reference genome



# The closest genome? Is it close “enough”?



<http://arthropodgenomes.org/wiki/i5K>



<http://genome10k.soe.ucsc.edu/>

**All depends on your question:**

Could be close enough – but might lose important information  
Advantages – annotations already done

# No reference? No problem!

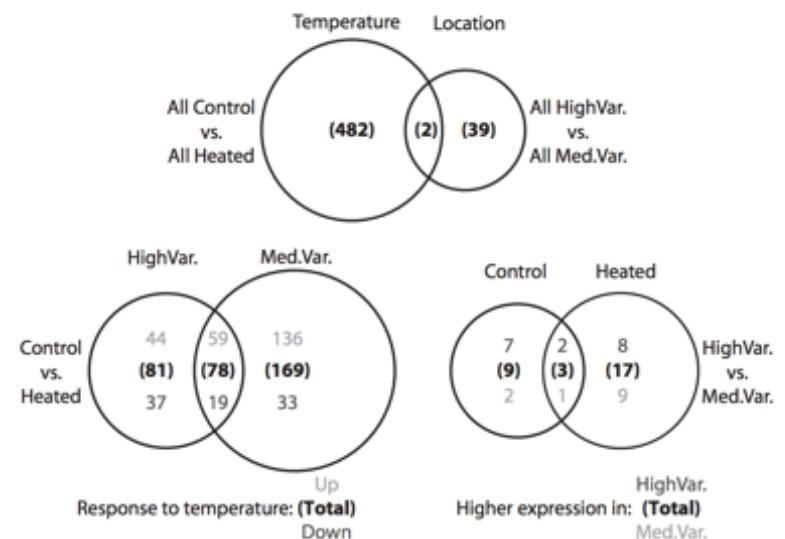
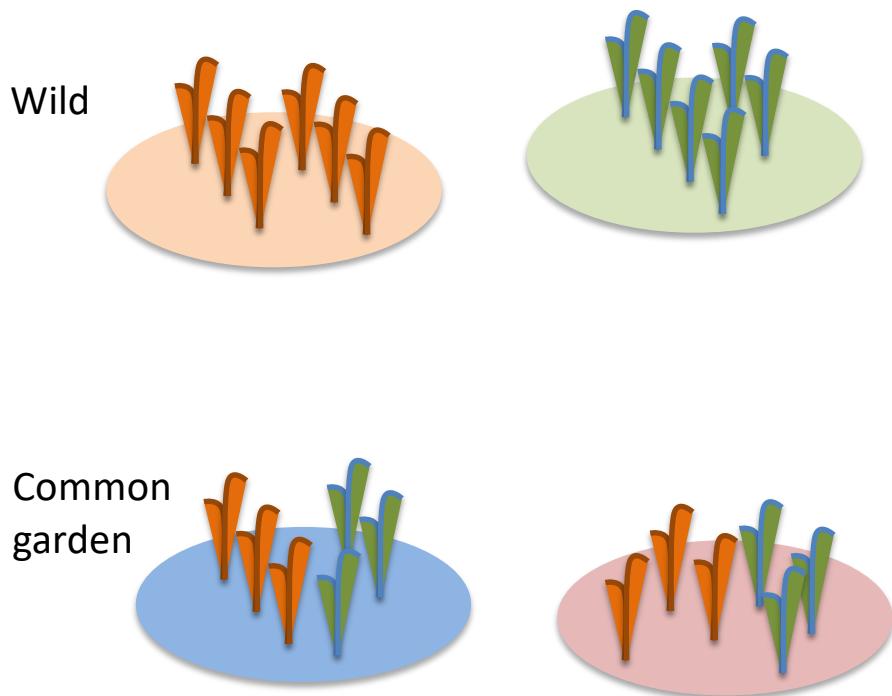
## De novo reference transcriptome assembly!

- Trinity de novo
- Other programs

# RNAseq to learn about climate change

## Genomic basis for coral resilience to climate change

Daniel J. Barshis<sup>1,2</sup>, Jason T. Ladner, Thomas A. Oliver, François O. Seneca, Nikki Taylor-Knowles, and Stephen R. Palumbi



**Fig. 2.** Venn diagram showing the number of differentially expressed genes detected during analysis based on temperature, location, within-location temperature response, and within-treatment location differences. Bold numbers in parentheses represent totals and respective shades of gray denote up- vs. down-regulated or higher in HV vs. MV, respectively.

# RNAseq for management decisions

Research article

Open Access

## Transcriptomic response to heat stress among ecologically divergent populations of redband trout

Shawn R Narum\* and Nathan R Campbell

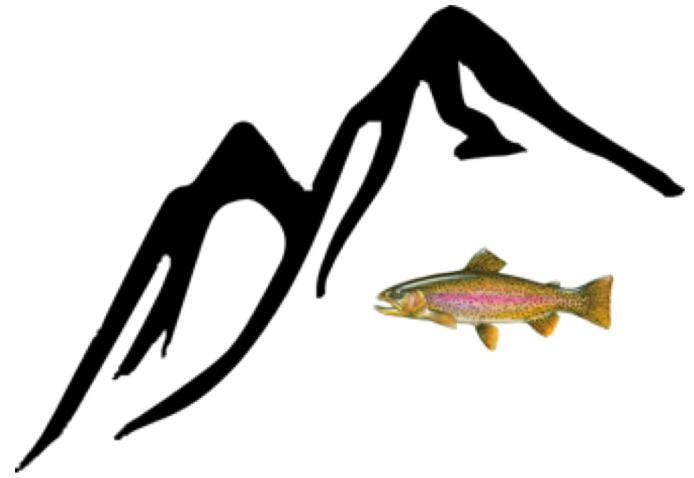
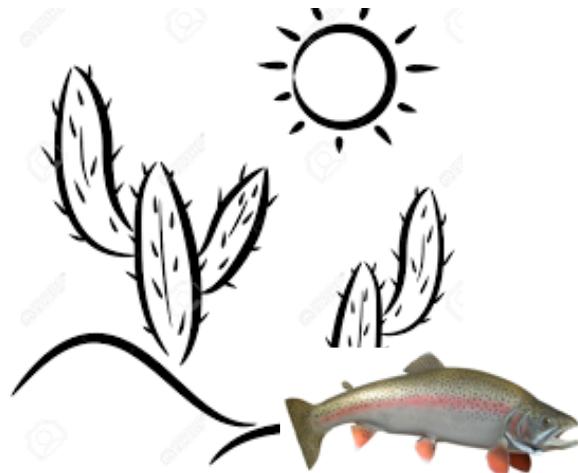
\* Corresponding author: Shawn R Narum [nars@critfc.org](mailto:nars@critfc.org)

▼ Author Affiliations

Columbia River Inter-Tribal Fish Commission, 3059-F National Fish Hatchery Road, Hagerman 83332, ID, USA

For all author emails, please [log on](#).

BMC Genomics 2015, **16**:103 doi:10.1186/s12864-015-1246-5



and F1s



# RNAseq for management decisions

Research article

Open Access

## Transcriptomic response to heat stress among ecologically divergent populations of redband trout

Shawn R Narum\* and Nathan R Campbell

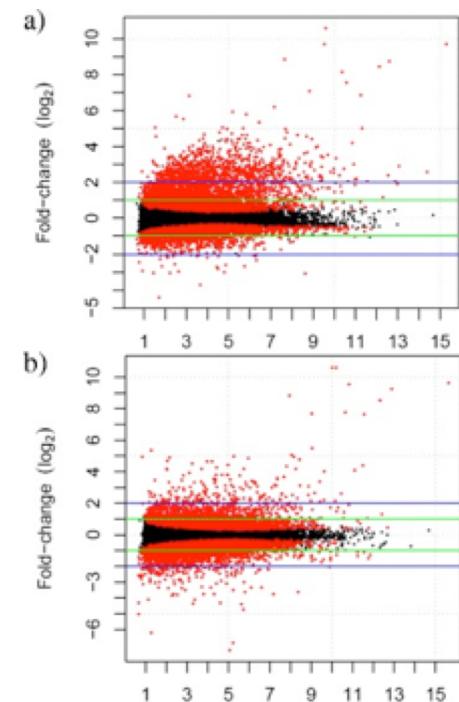
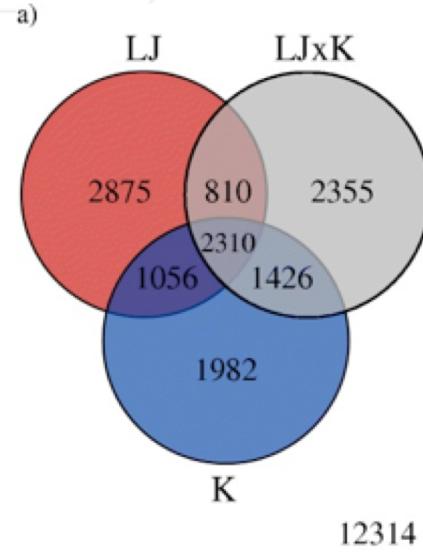
\* Corresponding author: Shawn R Narum [nars@critfc.org](mailto:nars@critfc.org)

▼ Author Affiliations

Columbia River Inter-Tribal Fish Commission, 3059-F National Fish Hatchery Road, Hagerman  
83332, ID, USA

For all author emails, please [log on](#).

BMC Genomics 2015, **16**:103 doi:10.1186/s12864-015-1246-5



# RNAseq for management decisions

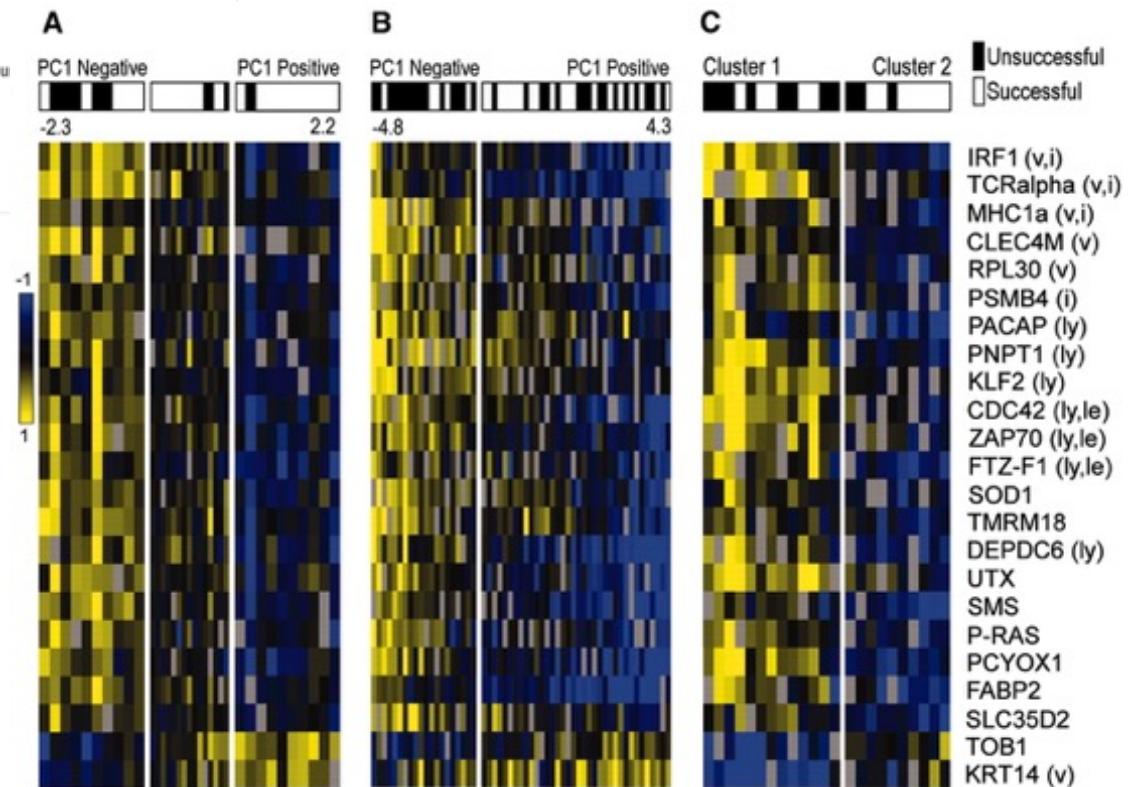
REPORT

## Genomic Signatures Predict Migration and Spawning Failure in Wild Canadian Salmon

Kristina M. Miller<sup>1,2,\*</sup>, Shaorong Li<sup>1</sup>, Karia H. Kaukinen<sup>1</sup>, Norma Ginther<sup>1</sup>, Edd Hammill<sup>3</sup>, Janelle M. R. Cu

\* See all authors and affiliations

Science 14 Jan 2011:  
Vol. 331, Issue 6014, pp. 214-217  
DOI: 10.1126/science.1196901



# SNPs from RNAseq data

## MOLECULAR ECOLOGY

Molecular Ecology (2015) 24, 2310–2323

doi: 10.1111/mec.13165

### INVITED REVIEWS AND SYNTHESES

#### SNP genotyping and population genomics from expressed sequences – current advances and future possibilities

PIERRE DE WIT,\* MELISSA H. PESPENI† and STEPHEN R. PALUMBI‡

Am J Hum Genet. 2013 Oct 3; 93(4): 641–651.  
doi: [10.1016/j.ajhg.2013.08.008](https://doi.org/10.1016/j.ajhg.2013.08.008)

PMCID: PMC3791257

#### Reliable Identification of Genomic Variants from RNA-Seq Data

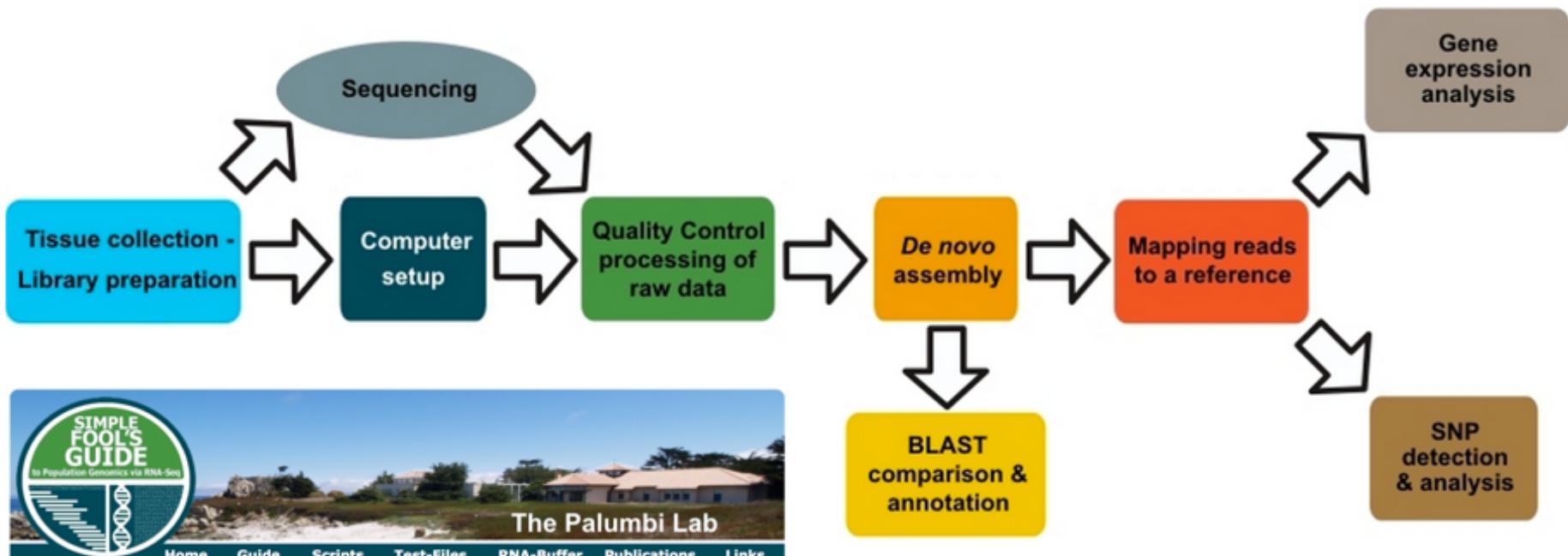
Robert Piskol,<sup>1</sup> Gokul Ramaswami,<sup>1</sup> and Jin Billy Li<sup>1,\*</sup>

### RESEARCH ARTICLE

#### Development of Strategies for SNP Detection in RNA-Seq Data: Application to Lymphoblastoid Cell Lines and Evaluation Using 1000 Genomes Data

Emma M. Quinn , Paul Cormican , Elaine M. Kenny , Matthew Hill, Richard Anney, Michael Gill, Aiden P. Corvin, Derek W. Morris 

Published: March 26, 2013 • DOI: 10.1371/journal.pone.0058815



### The Simple Fool's Guide to Population Genomics via RNA-Seq: An Introduction to High-Throughput Sequencing Data Analysis

This website and accompanying documents are intended as a tool to help researchers dealing with non-model organisms acquire and process transcriptomic high-throughput sequencing data without having to learn extensive bioinformatics skills. It covers all steps from tissue collection, sample preparation and computer setup, through addressing biological questions with gene expression and SNP data.



You may cite this work as follows:

De Wit P, Pespeni MH, Ladner JT, Barshis DJ, Seneca F, Jaris H, Overgaard Therkildsen N, Morikawa M and Palumbi SR (2012) The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Molecular Ecology Resources* **12**, 1058-1067.

# Data analysis (and there are no pipelines)

- Demultiplex
- Look at the data (for example, fastQC)
- Trim adapters
- Trim low quality reads and/or bases

# Data analysis (and there are no pipelines)

- Mapping
- Determining differentially expressed genes