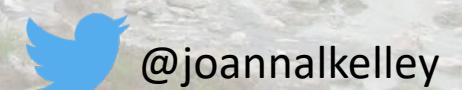
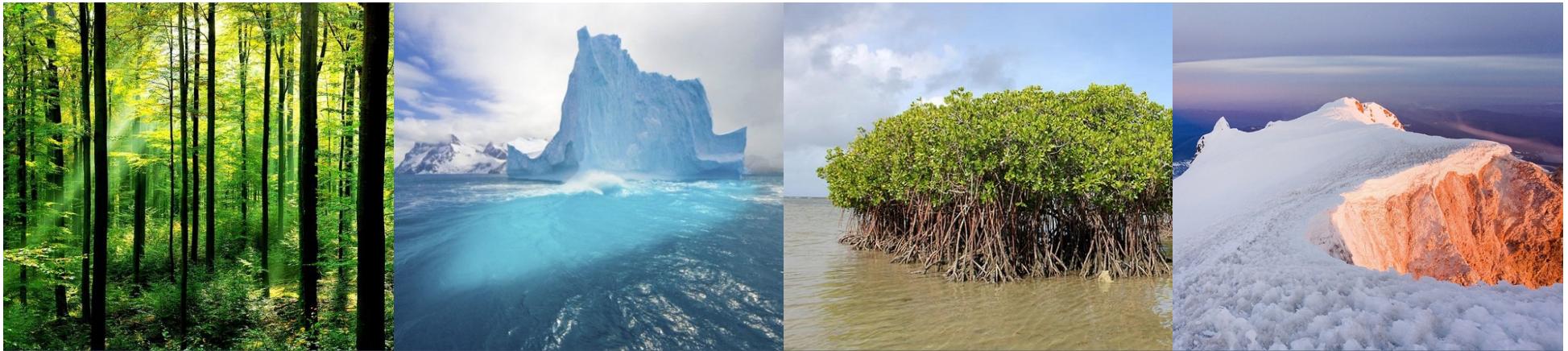


Gene expression, transcriptomics, and adaptive evolution

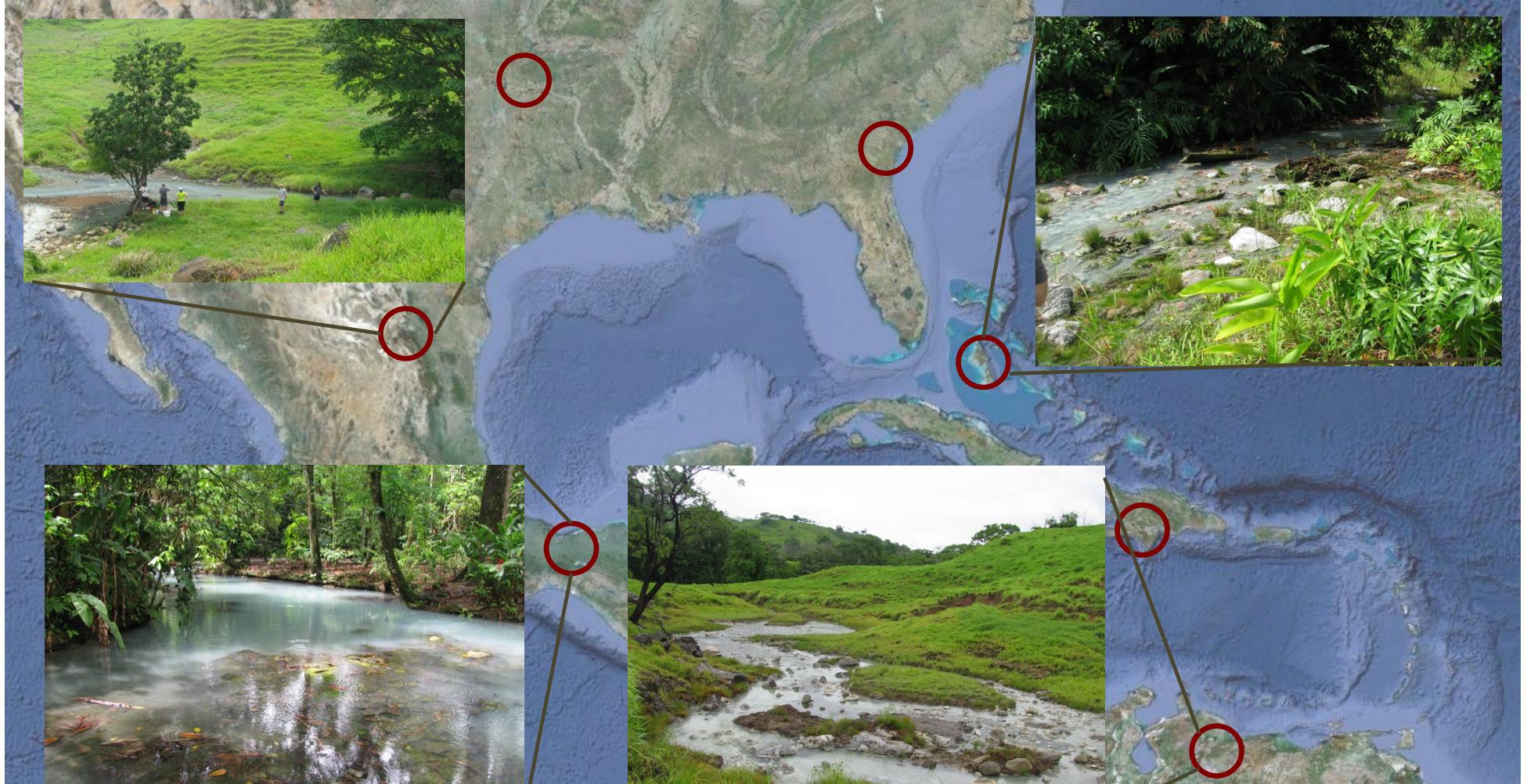
Joanna Kelley
ConGen 2019





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





Hydrogen sulfide is a strong and constant selective pressure

- Naturally occurring in volcanic regions
- Inhibits oxygen transport and cellular respiration
- 100 n micromolar
- Concentrations up to 1100 μM present
- Causes and aggravates hypoxia

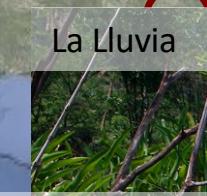
Gambusia affinis
Oklahoma



Gambusia holbrooki
Florida



La Lluvia



Gambusia hubbsi
Bahamas



Limia sulphurophila
Dominican Republic



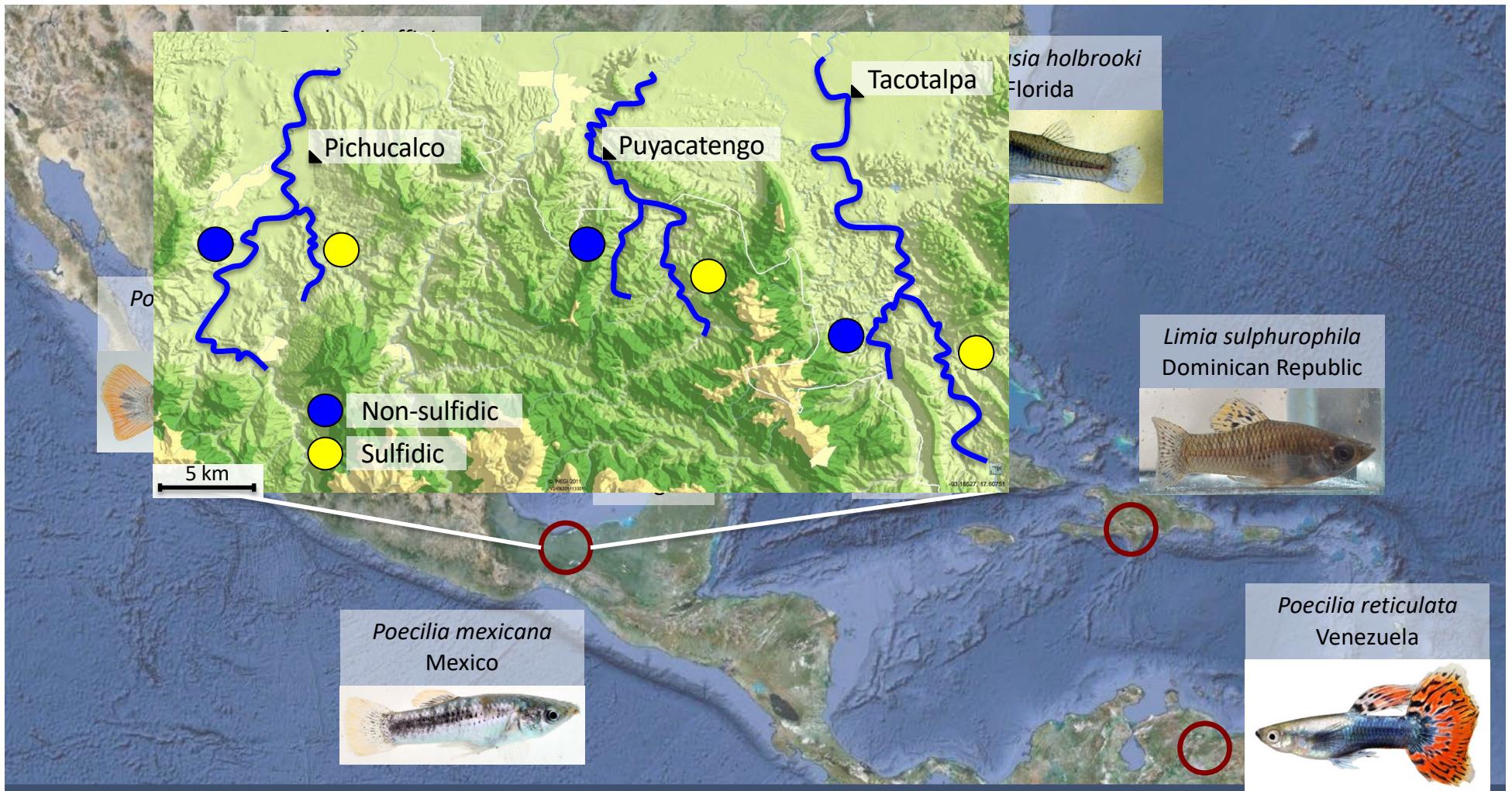
Poecilia mexicana
Mexico



Poecilia reticulata
Venezuela

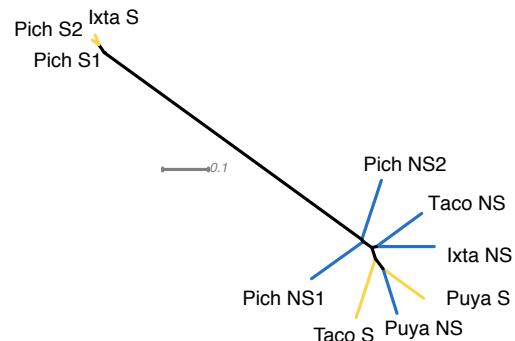
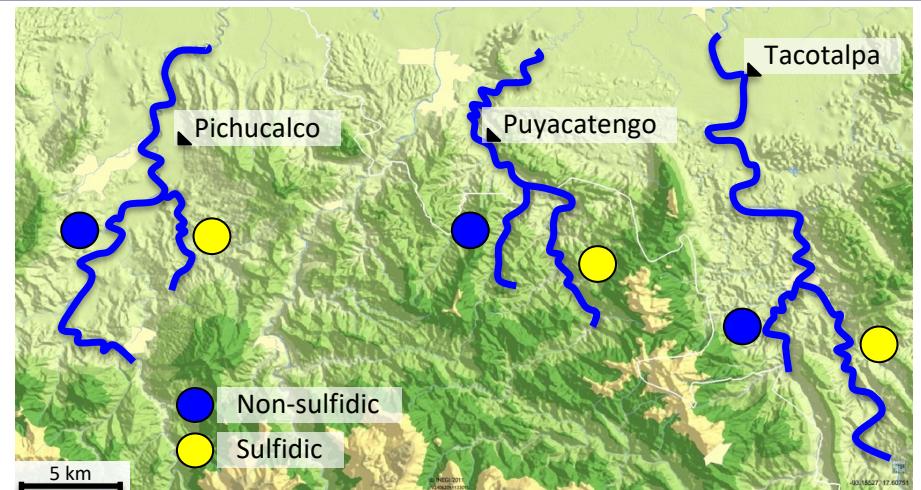
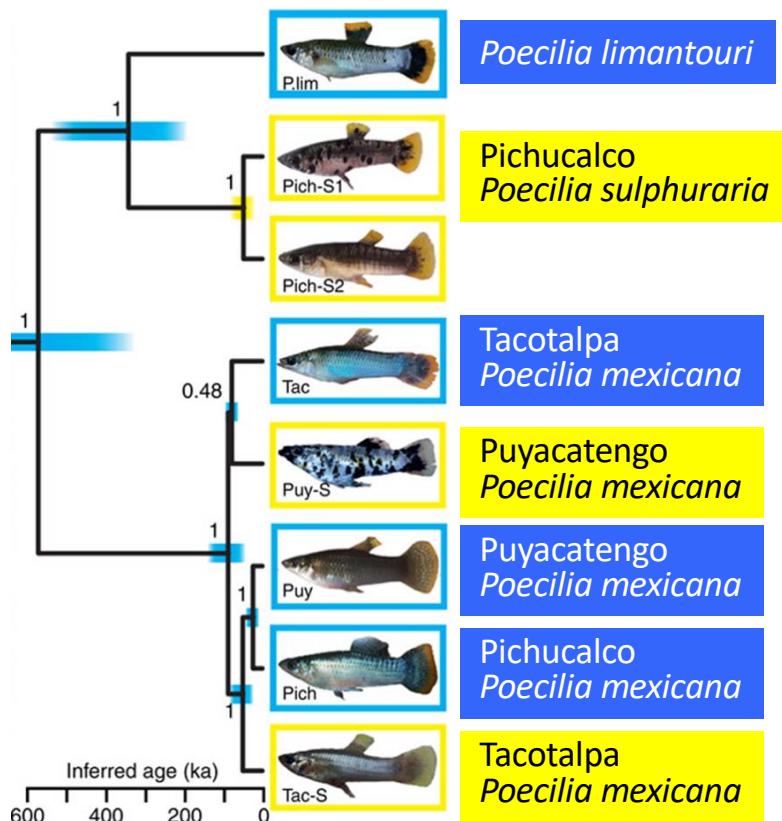


Fish Hydrogen sulfide tolerant species
the family Poeciliidae have colonized many the Springs



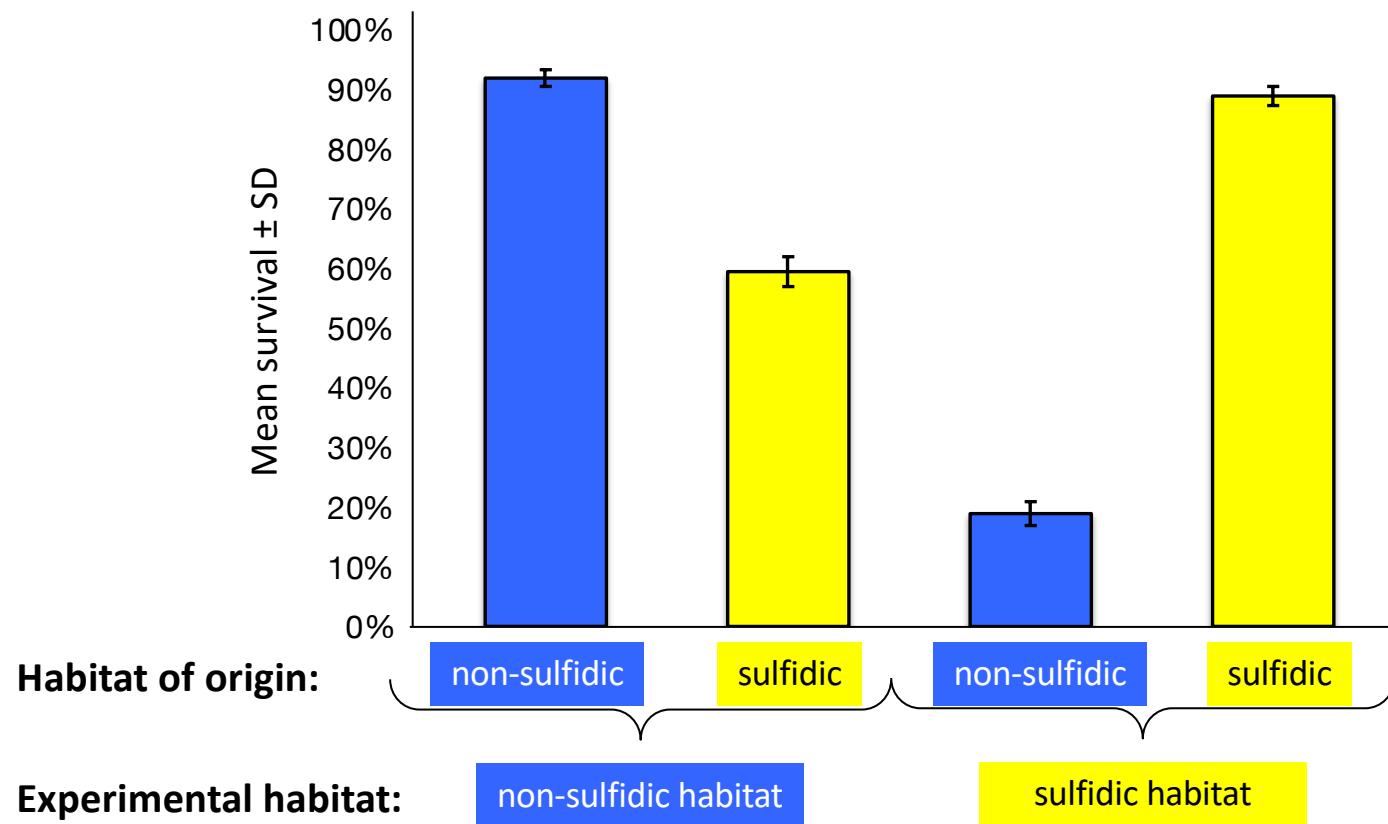
Fish in the family Poeciliidae have colonized many H₂S-rich springs

Independent colonization of sulfide springs within southern Mexico



Pfenninger et al. 2013, *Nat Commun*

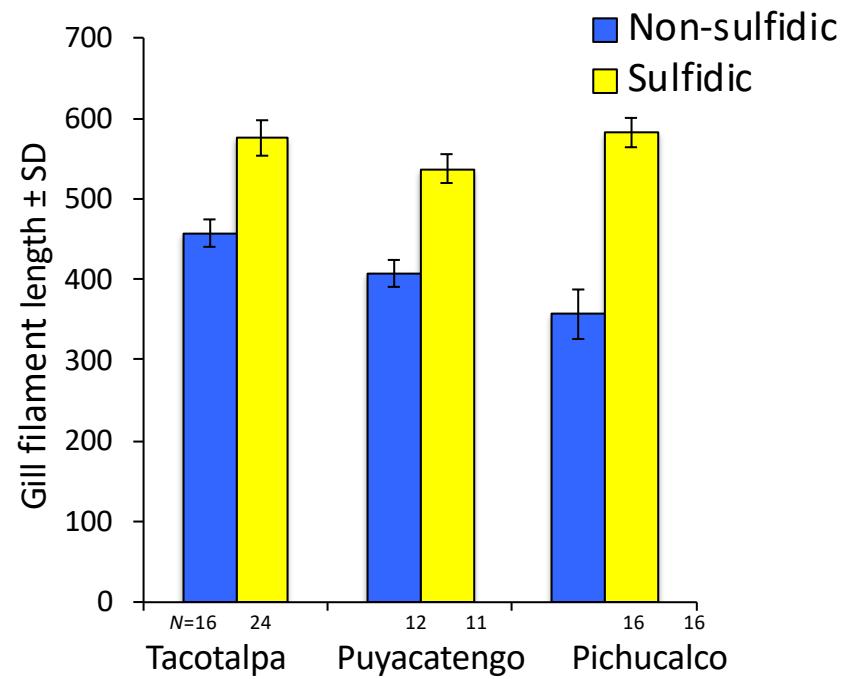
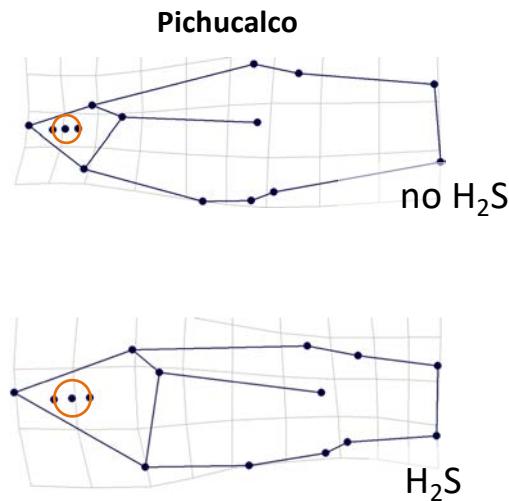
Populations are locally adapted



Plath et al. 2013 *Evolution*

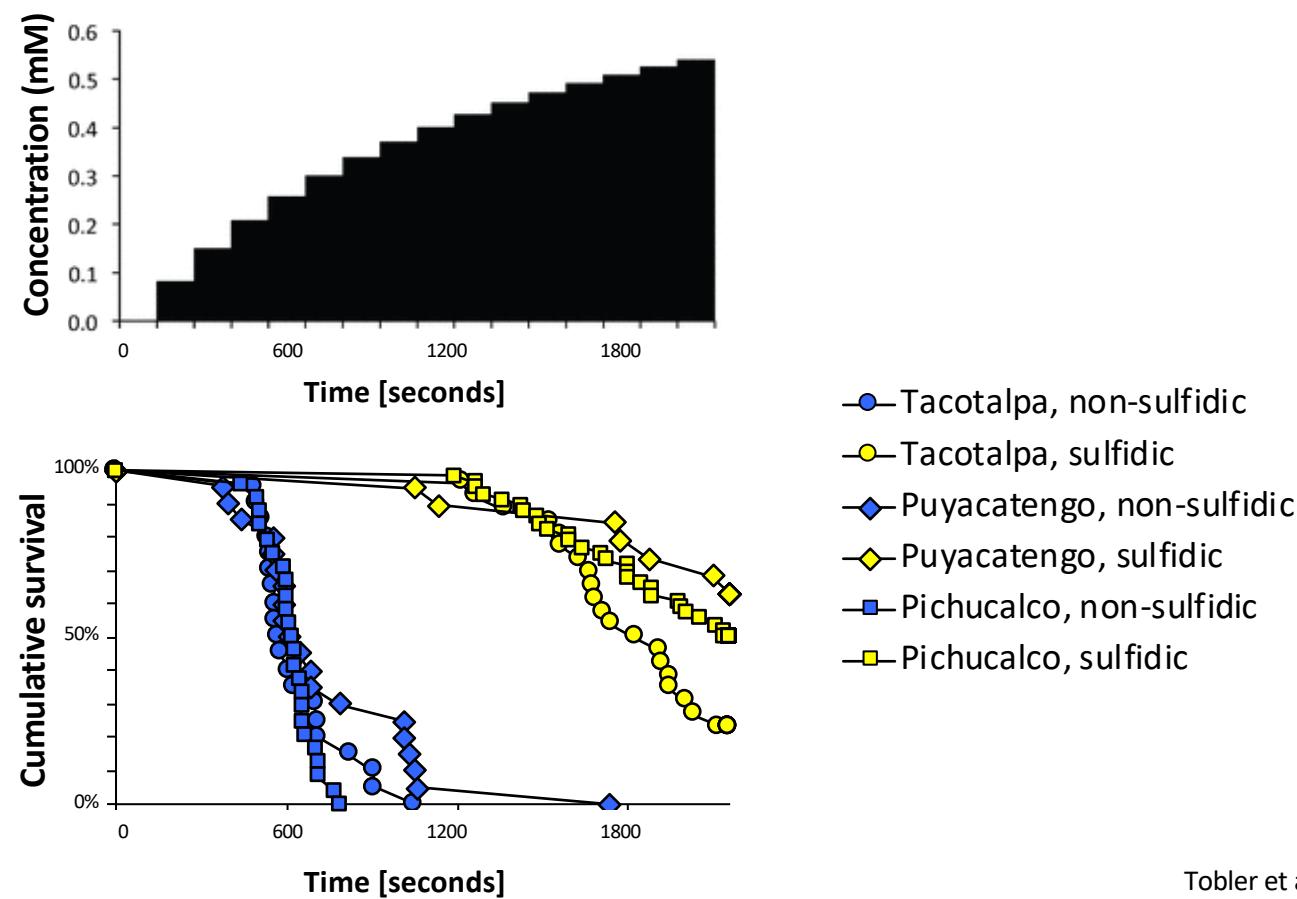
Convergent evolution of large heads and gills

Increase in head size is
correlated with an
increase in total gill
filament length



Tobler et al. 2011, *Evolution*

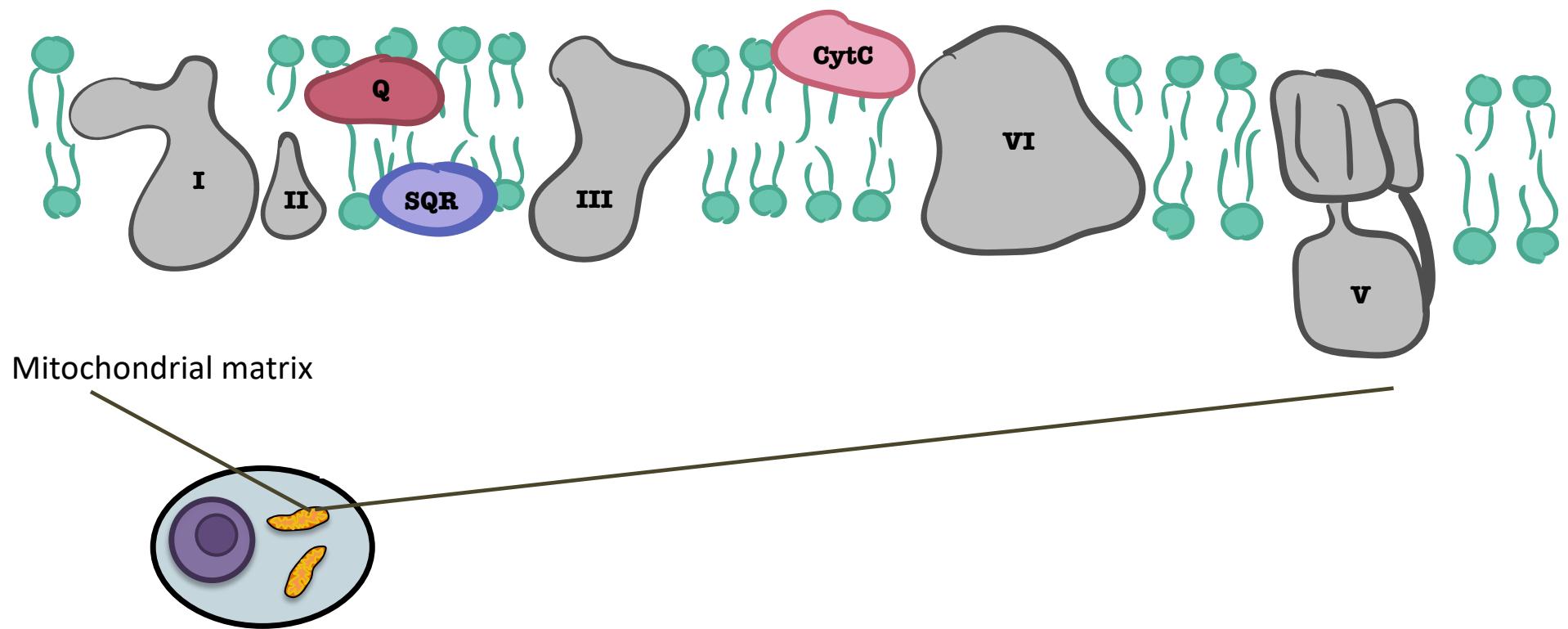
Physiological sulfide tolerance varies between ecotypes



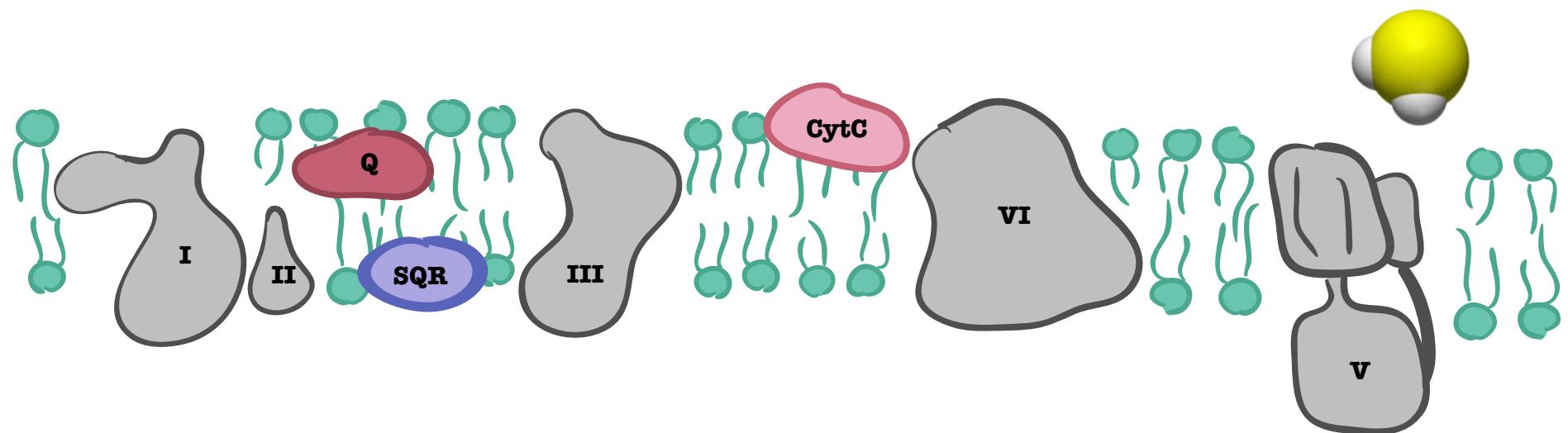
Tobler et al. 2011, *Evolution*

Hydrogen sulfide (H_2S) is a physiochemical stressor

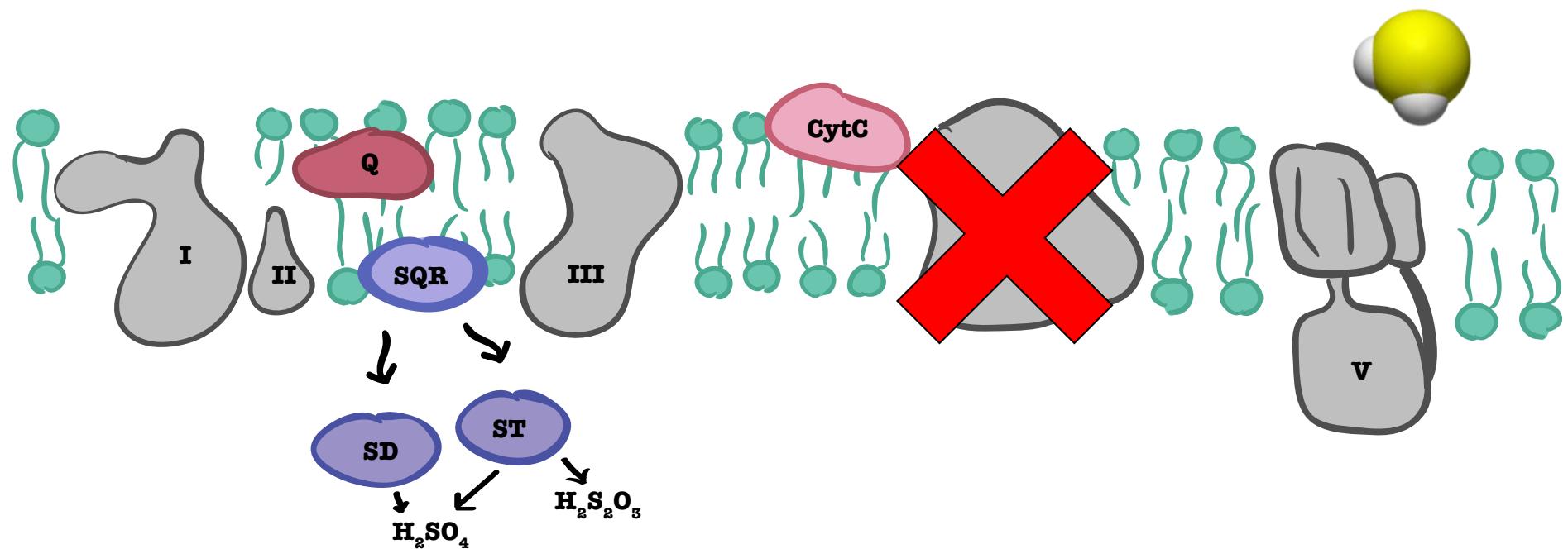
Inter-membrane space



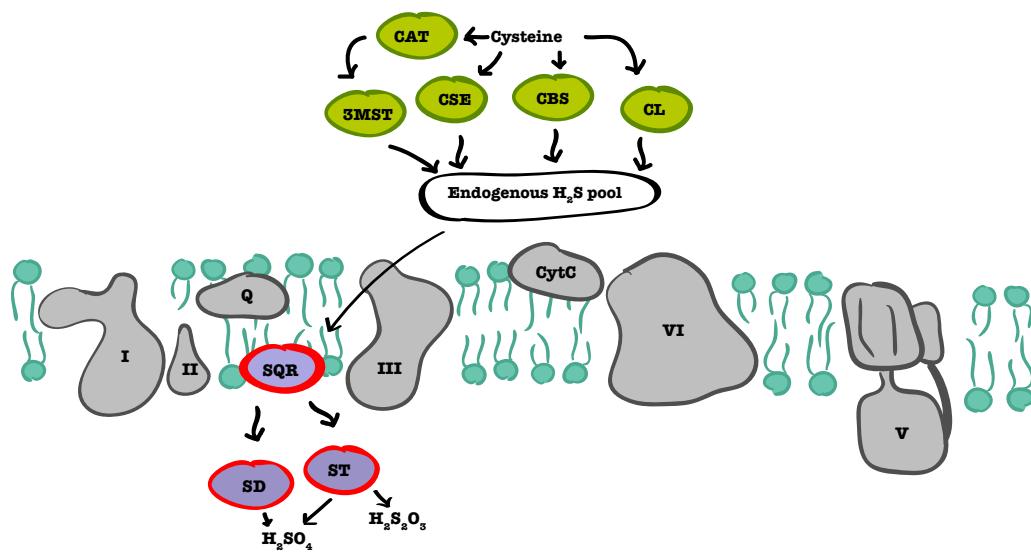
Hydrogen sulfide (H_2S) is a physiochemical stressor



At high concentrations H₂S blocks cytochrome c oxidase

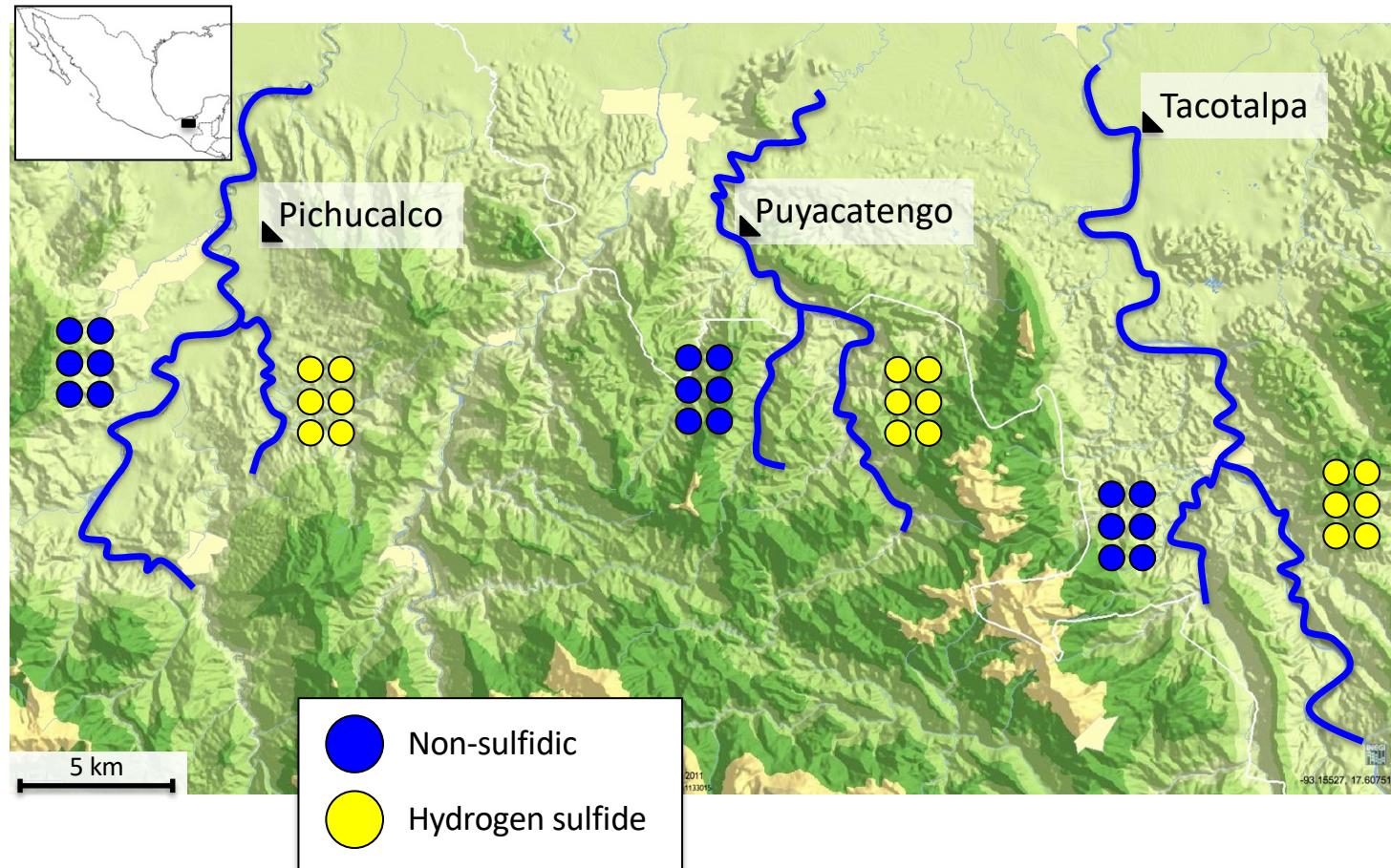


Predictions based on biochemical and physiological mechanisms



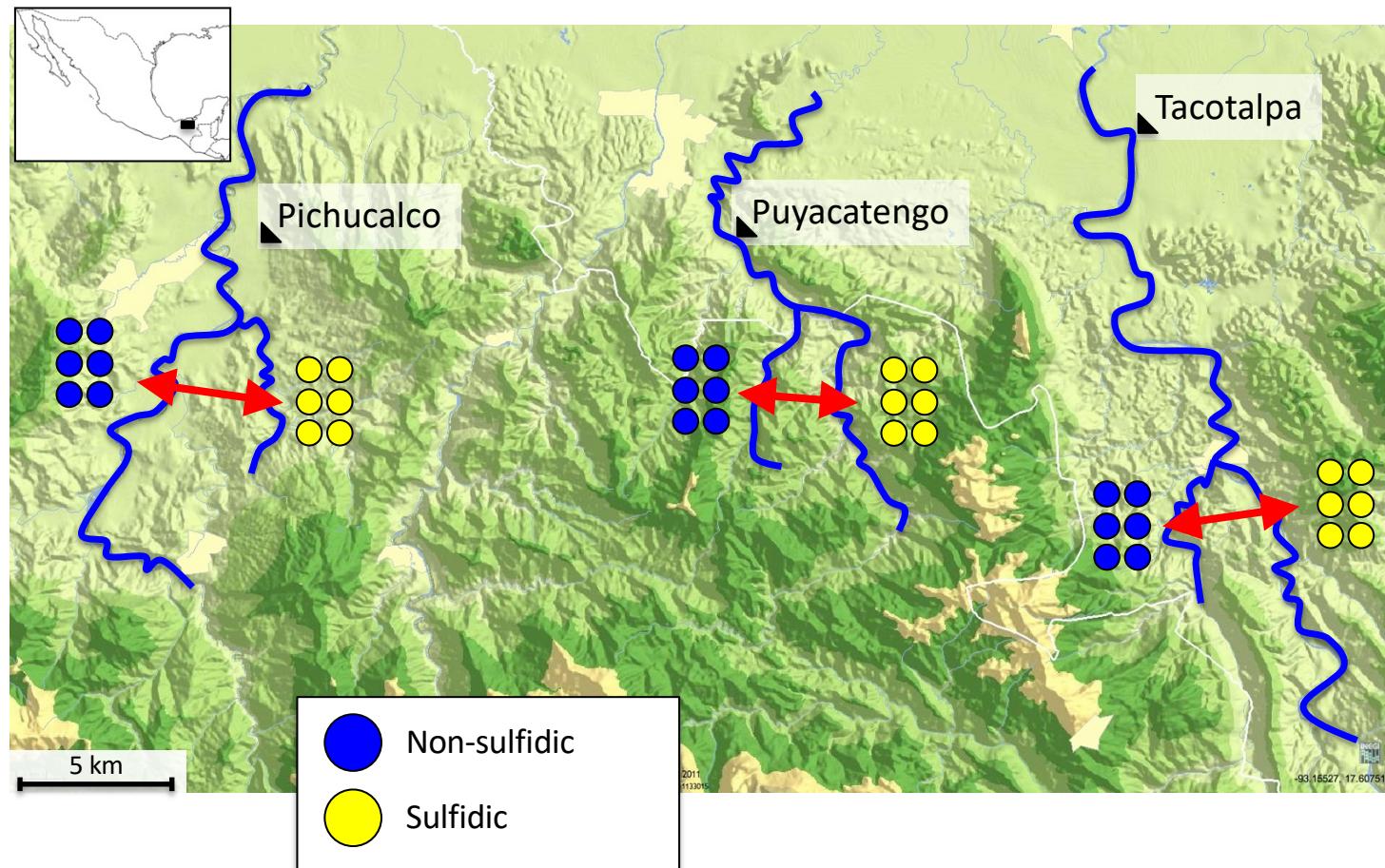
- Modification of cytochrome c oxidase (COX), a major target of H₂S toxicity
- Up-regulation of enzymes detoxifying H₂S
- Down-regulation of enzymes producing H₂S within the body
- Hundreds of other molecular targets have been described based on biomedical research

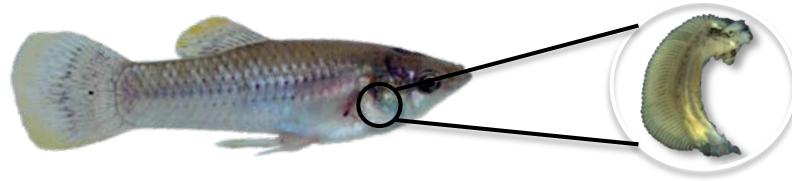
Sample individuals from pairs of sulfidic and non-sulfidic environments to identify convergent changes in expression



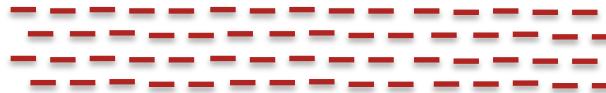


Identify differentially expressed genes between ecotypes





RNA-sequencing



Quantify transcript abundance



Identify allelic variants

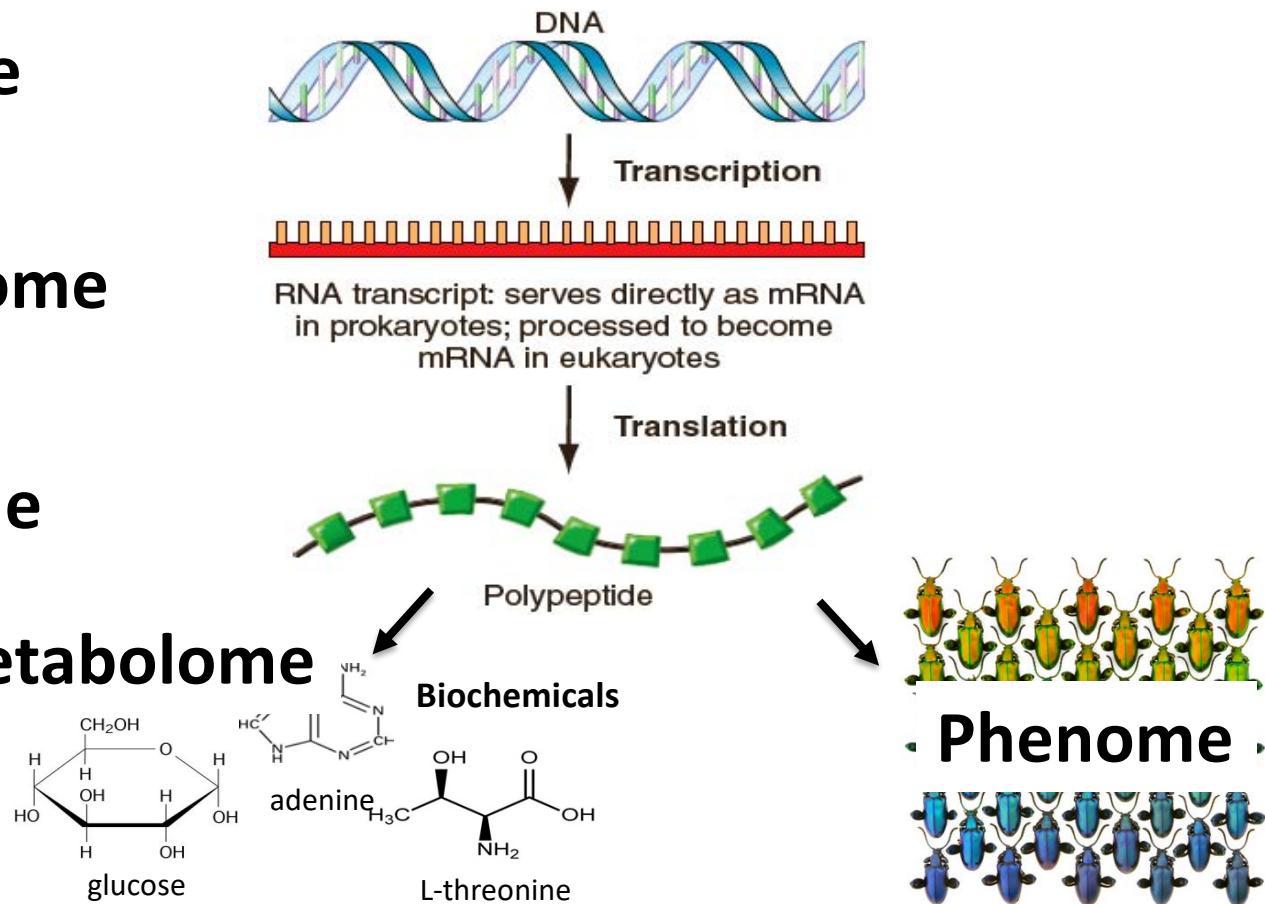
GATGCTCGAA	TATAGCTACCGACGA	TACTGAT
GATGCTCGAA	TATCGCTACCGACGA	TACTGGT
GATGCTGGAA	TATAGCTACCGACGA	TACTGAT
GATGCTGGAA	TATCGCTACCGACGA	TACTGCT

Functional annotation

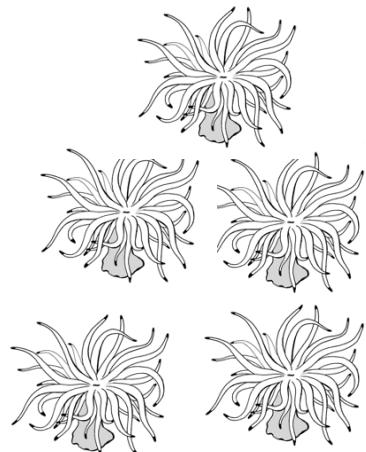


OMICS! approaches

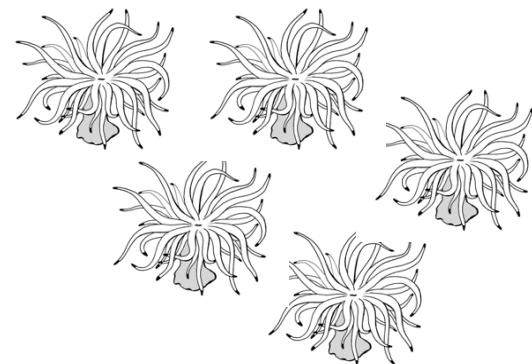
Genome
Transcriptome
Proteome
Metabolome



Why study gene expression differences among individuals and populations?



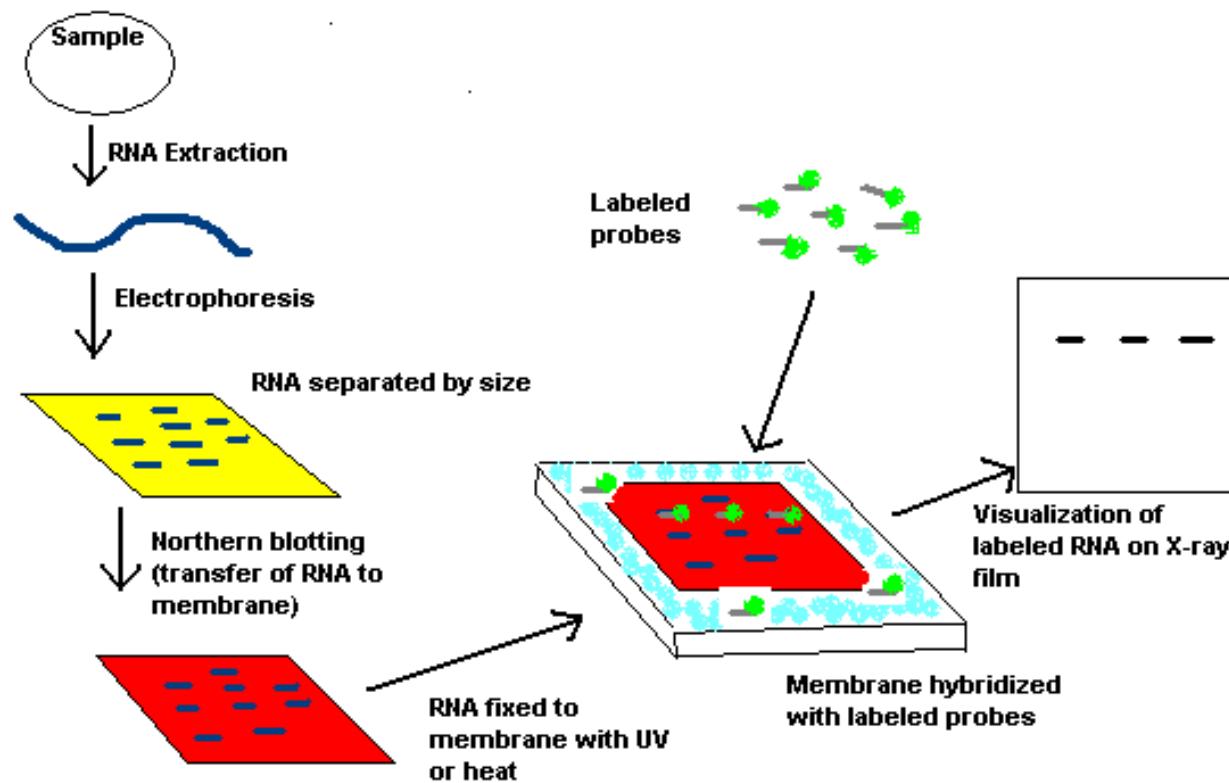
- Insights into the molecular basis of phenotypic diversity
- Interpretation of patterns of expression variation in response to environmental conditions, disease, etc.
- Possible management decisions on how and where to manage or transplant populations



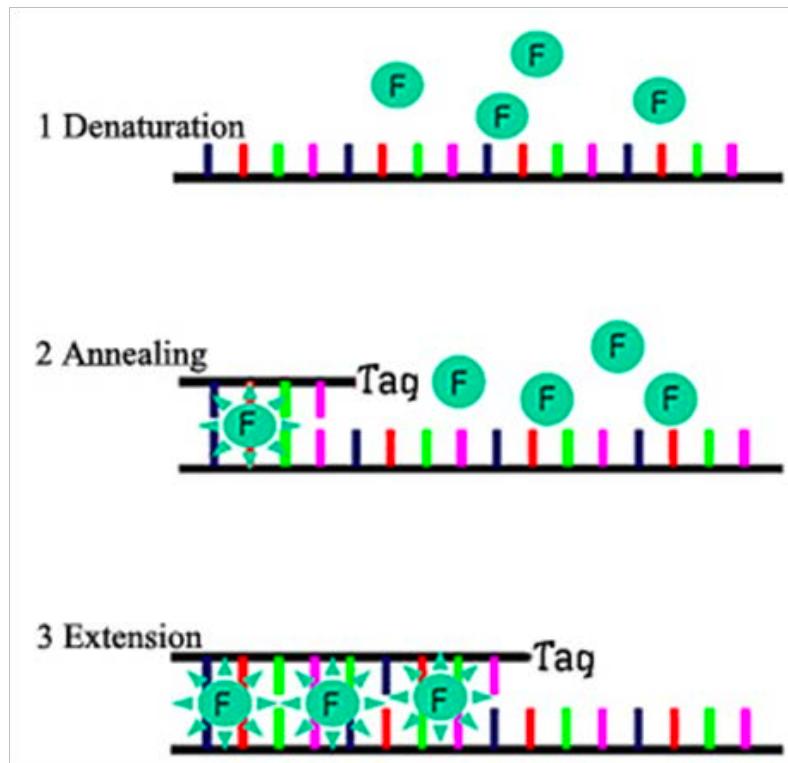
Approaches to measuring gene expression

- 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

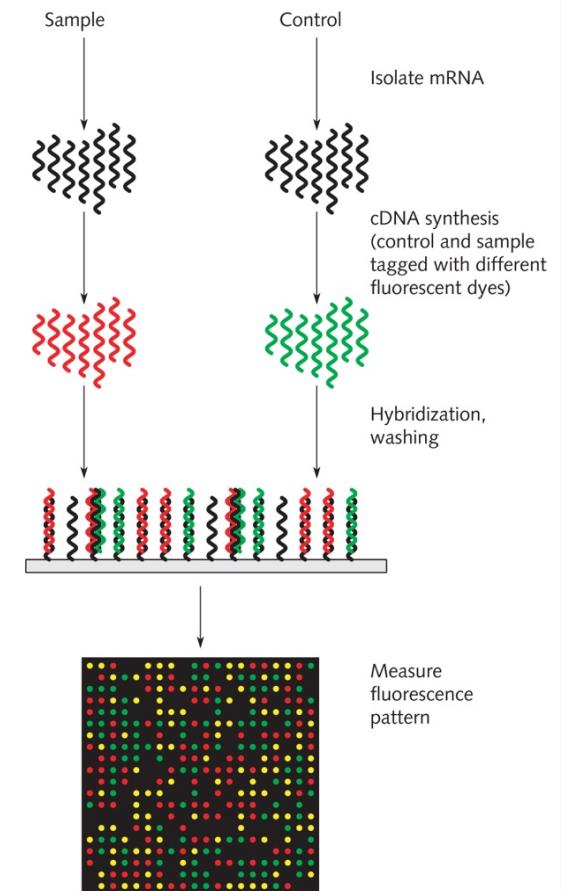
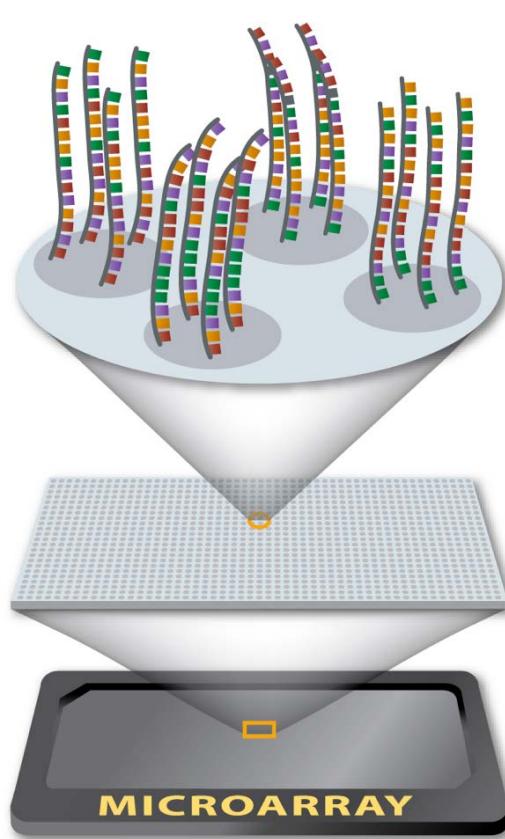
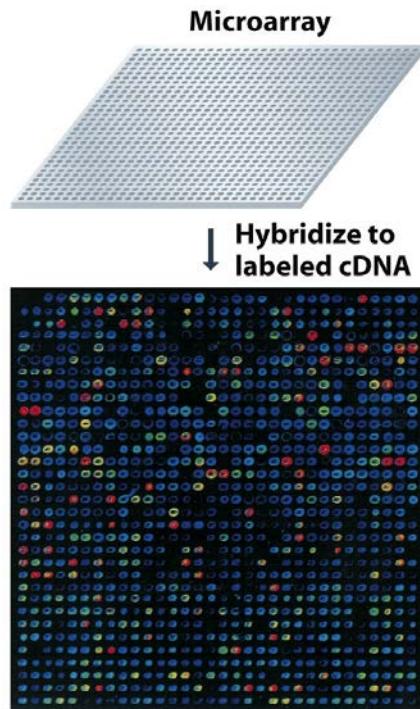


qPCR on target genes

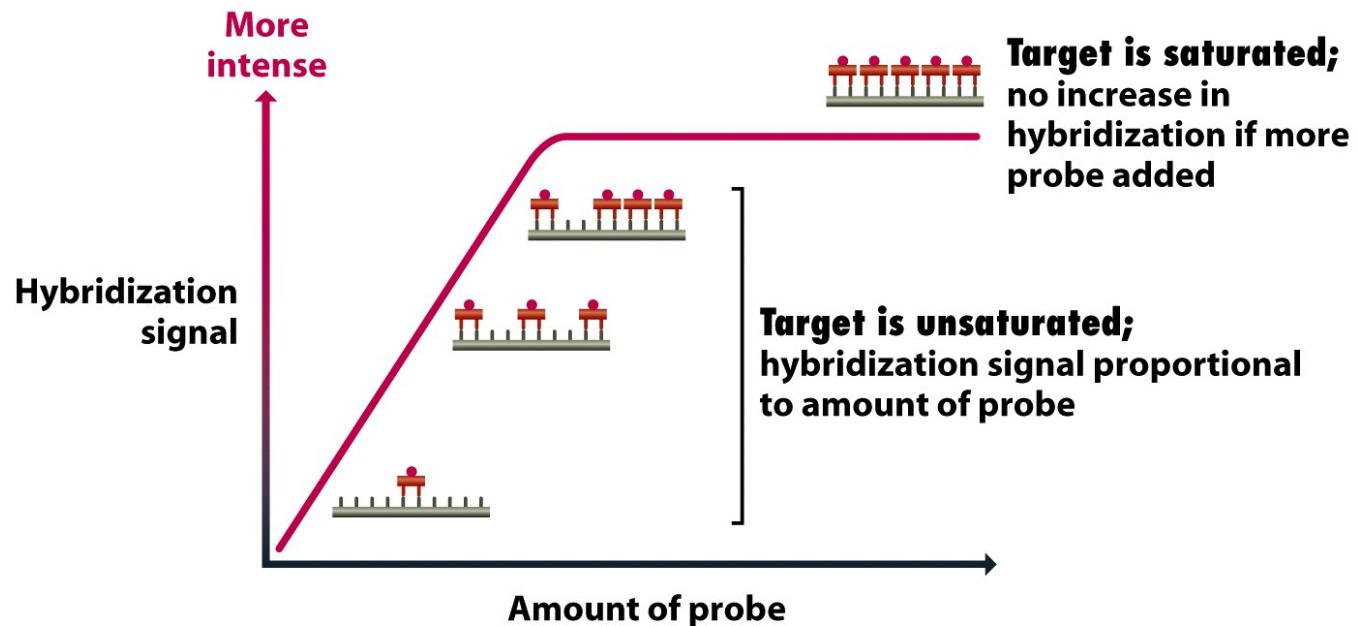


Microarrays

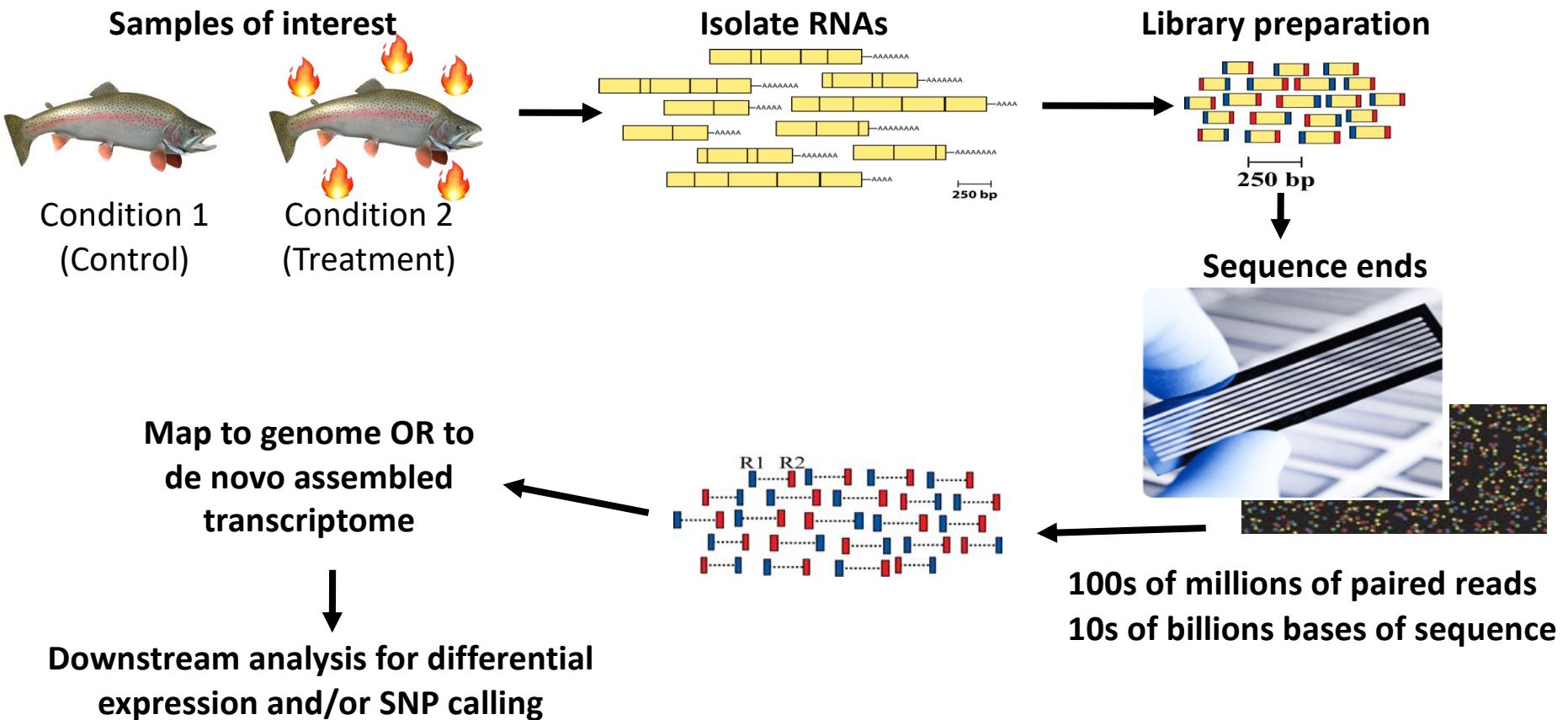
Array based approach
to measuring gene
expression



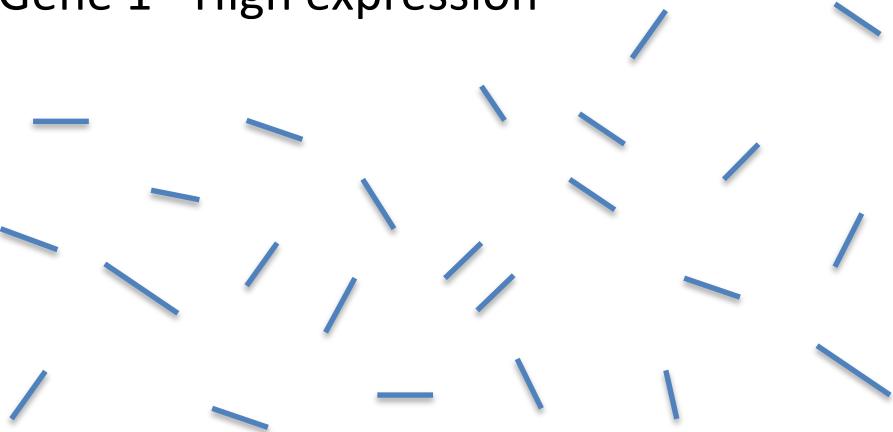
Microarrays have limited range for detecting changes in expression



RNA sequencing experiment



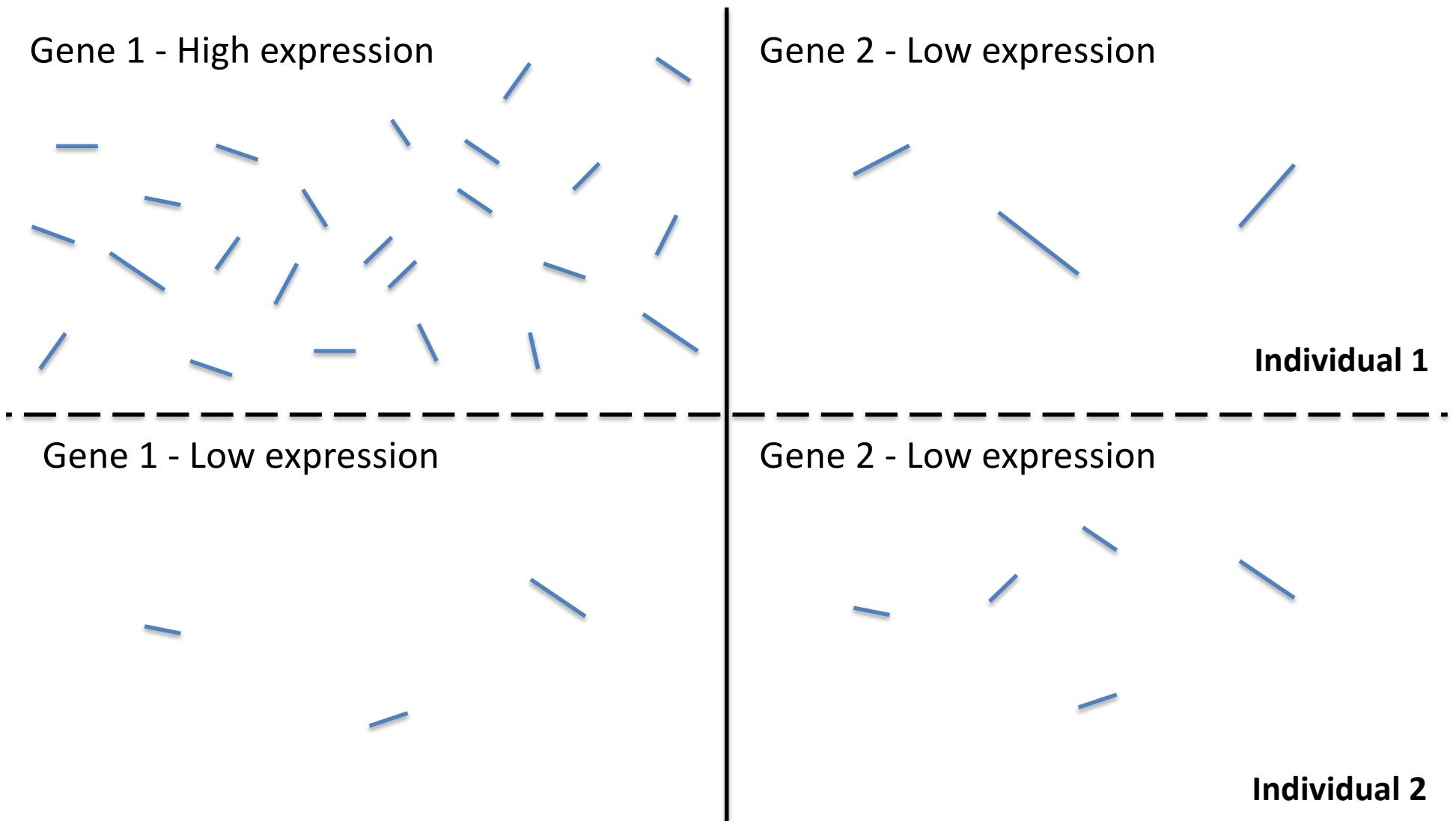
Gene 1 - High expression



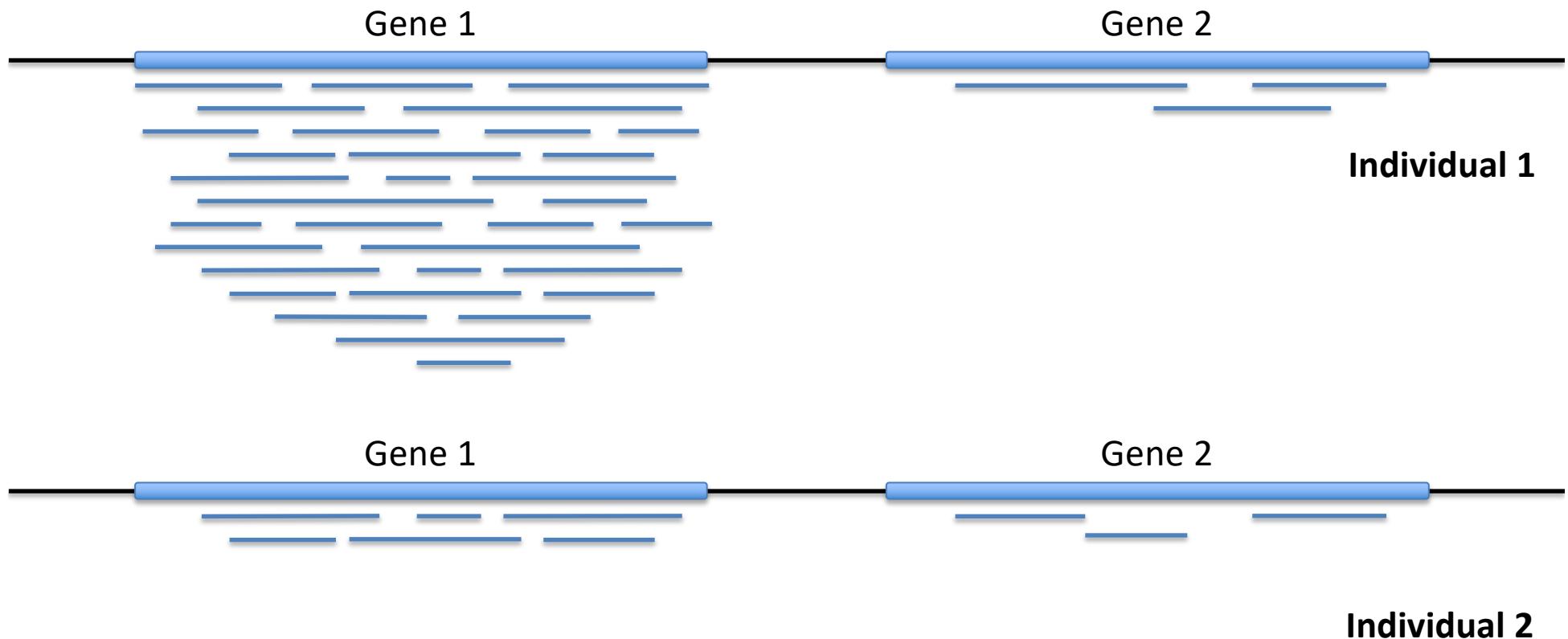
Gene 2 - Low expression



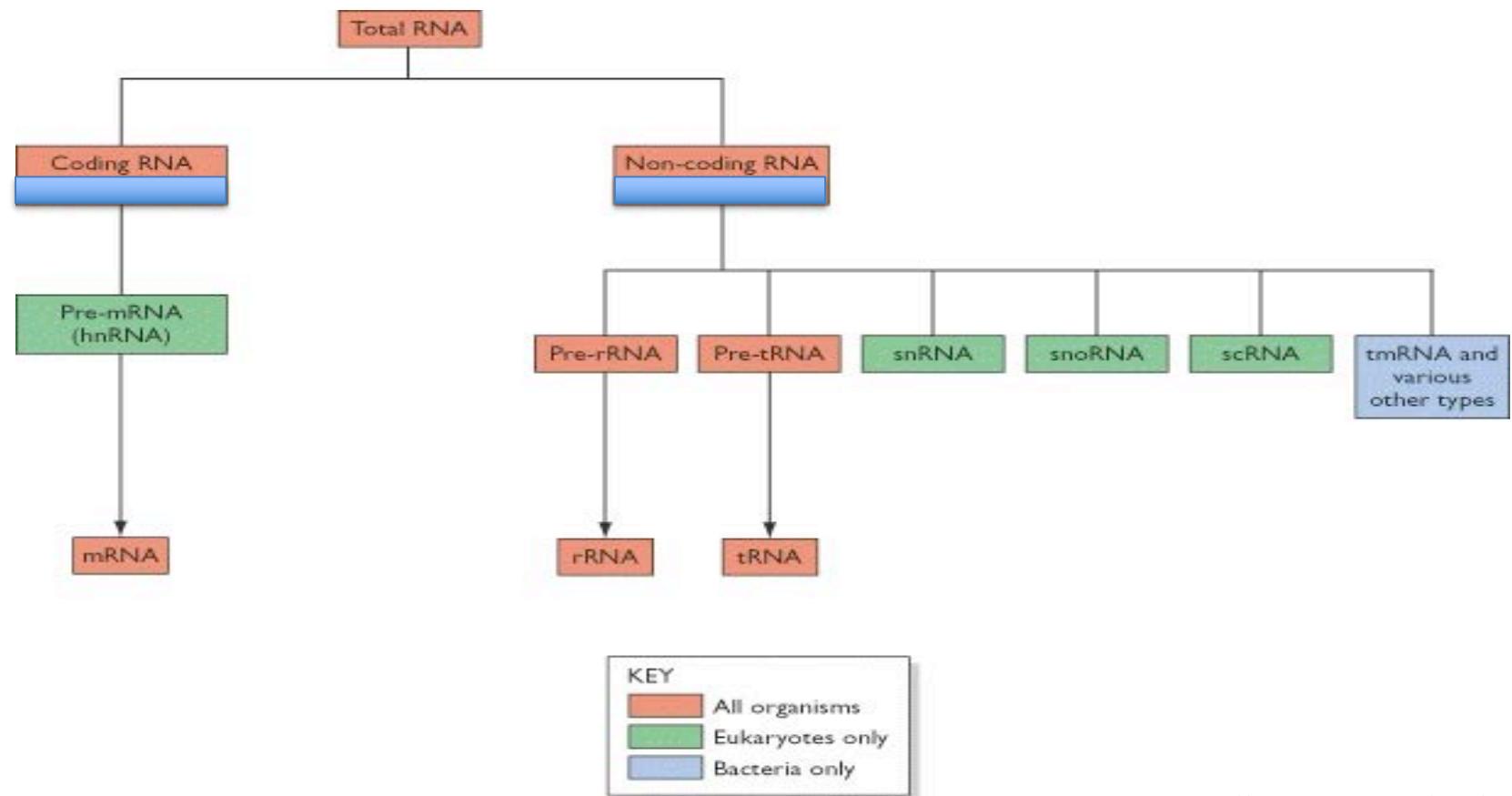
Individual 1



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



Components of total RNA



Extracting total RNA

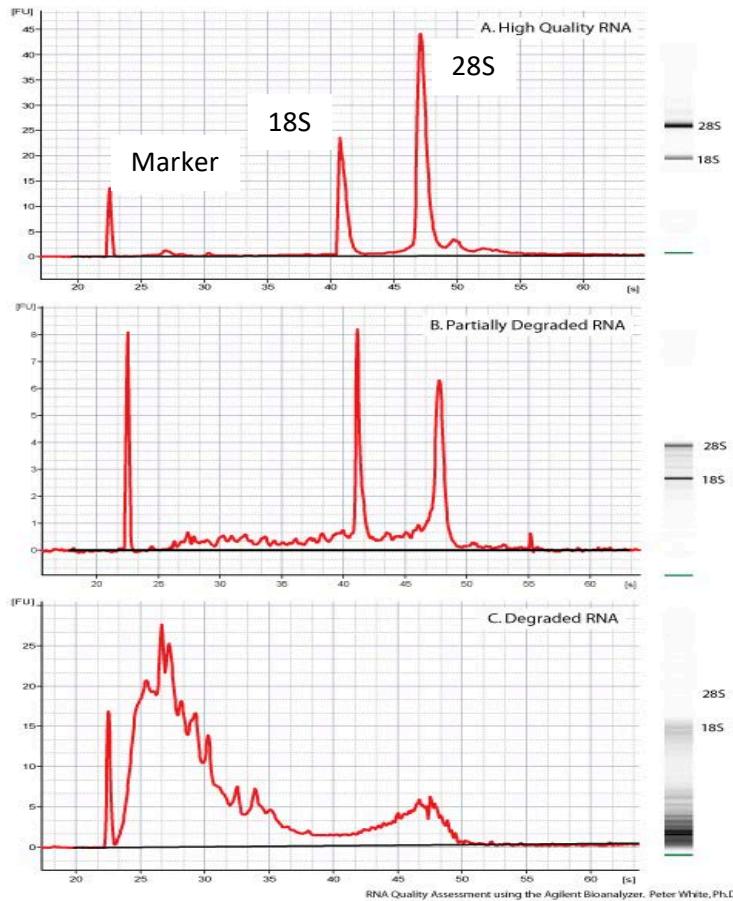
What would total RNA look like if we ran it on a gel?

High quality RNA?

Slightly degraded RNA?

Degraded RNA?

Degraded total RNA?

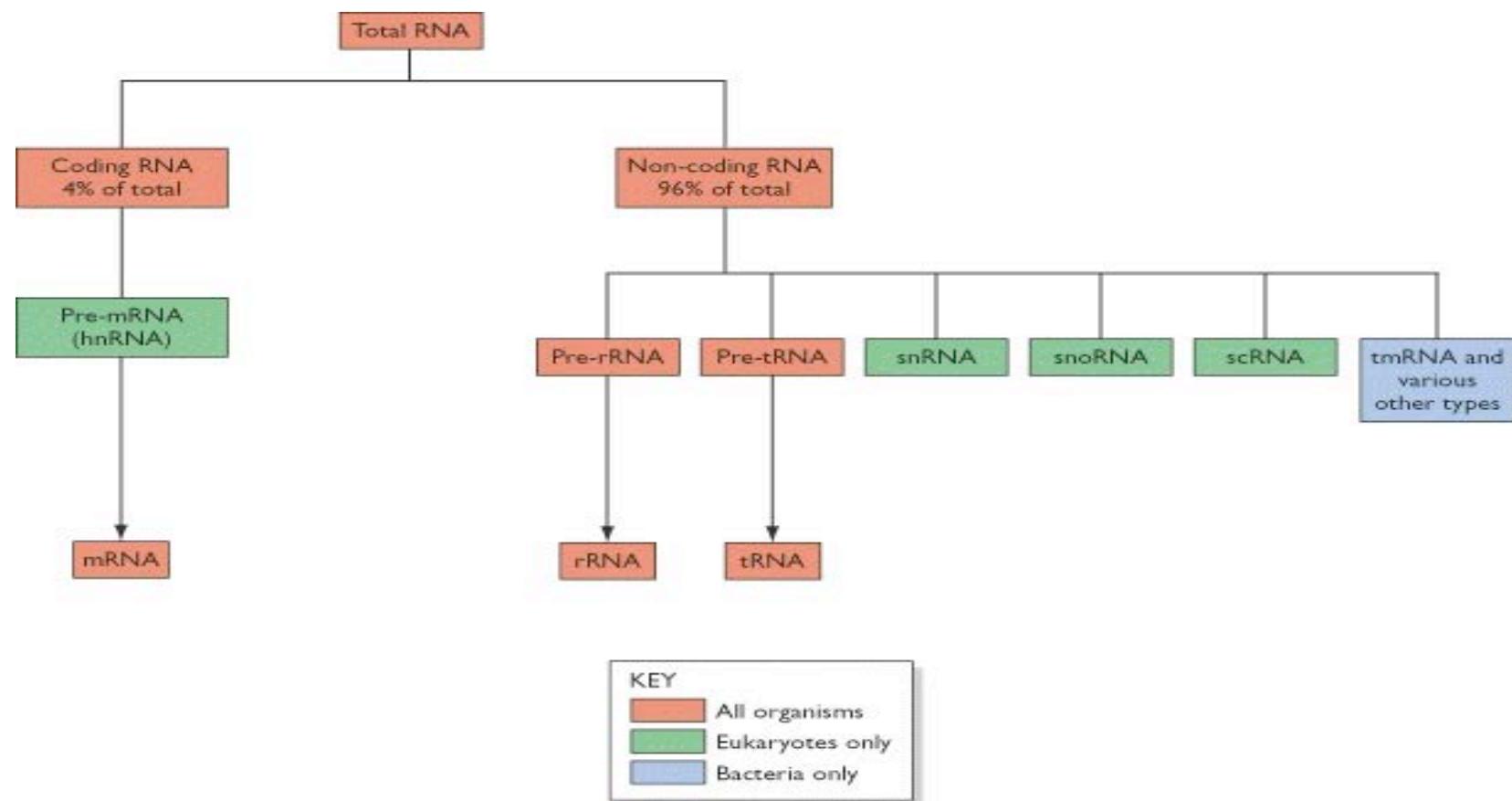


High quality RNA

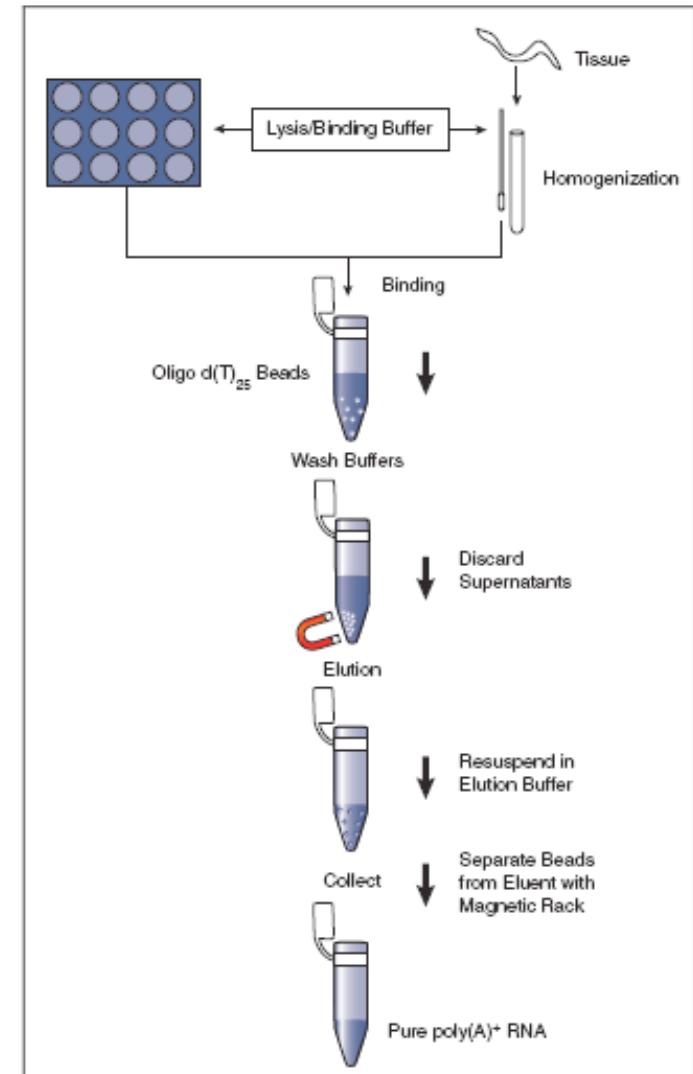
Slightly degraded RNA

Degraded RNA

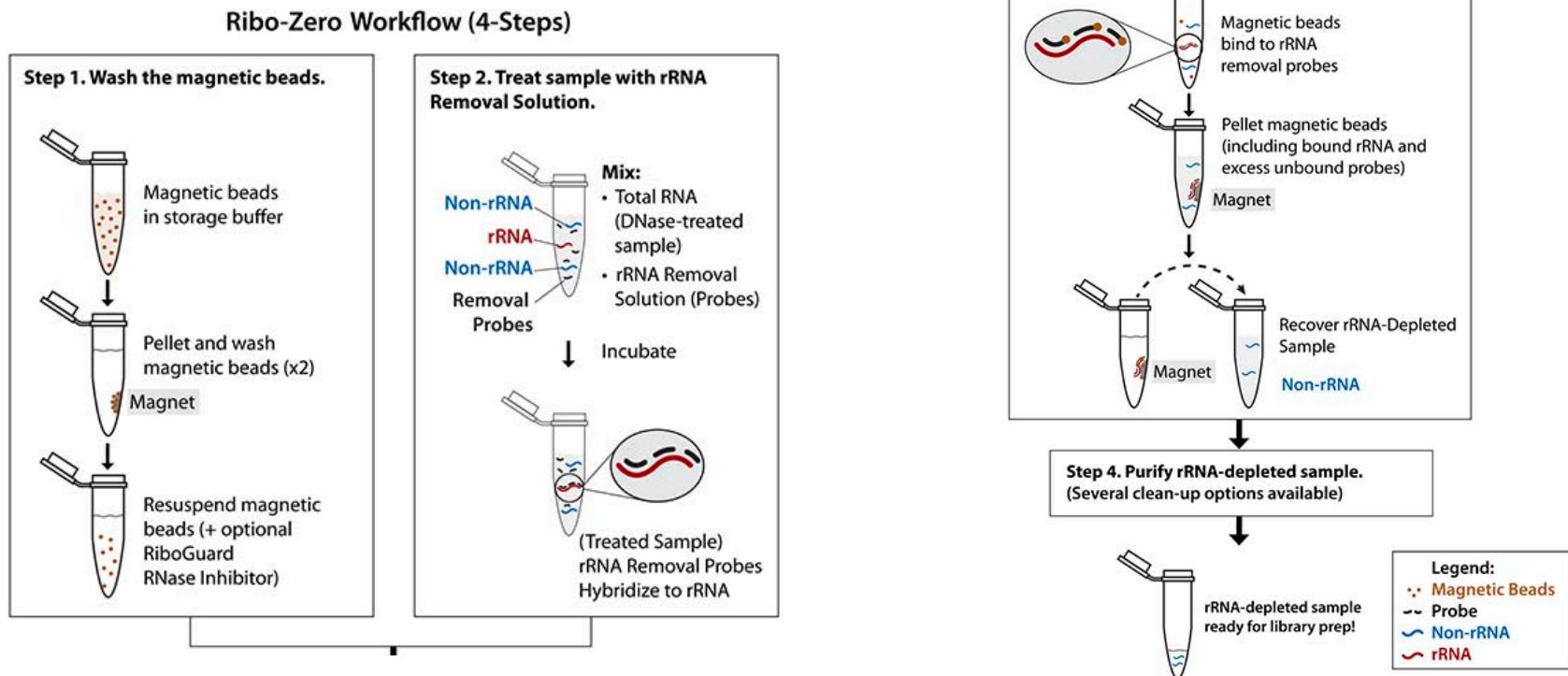
How to enrich for the RNA that we are interested in?



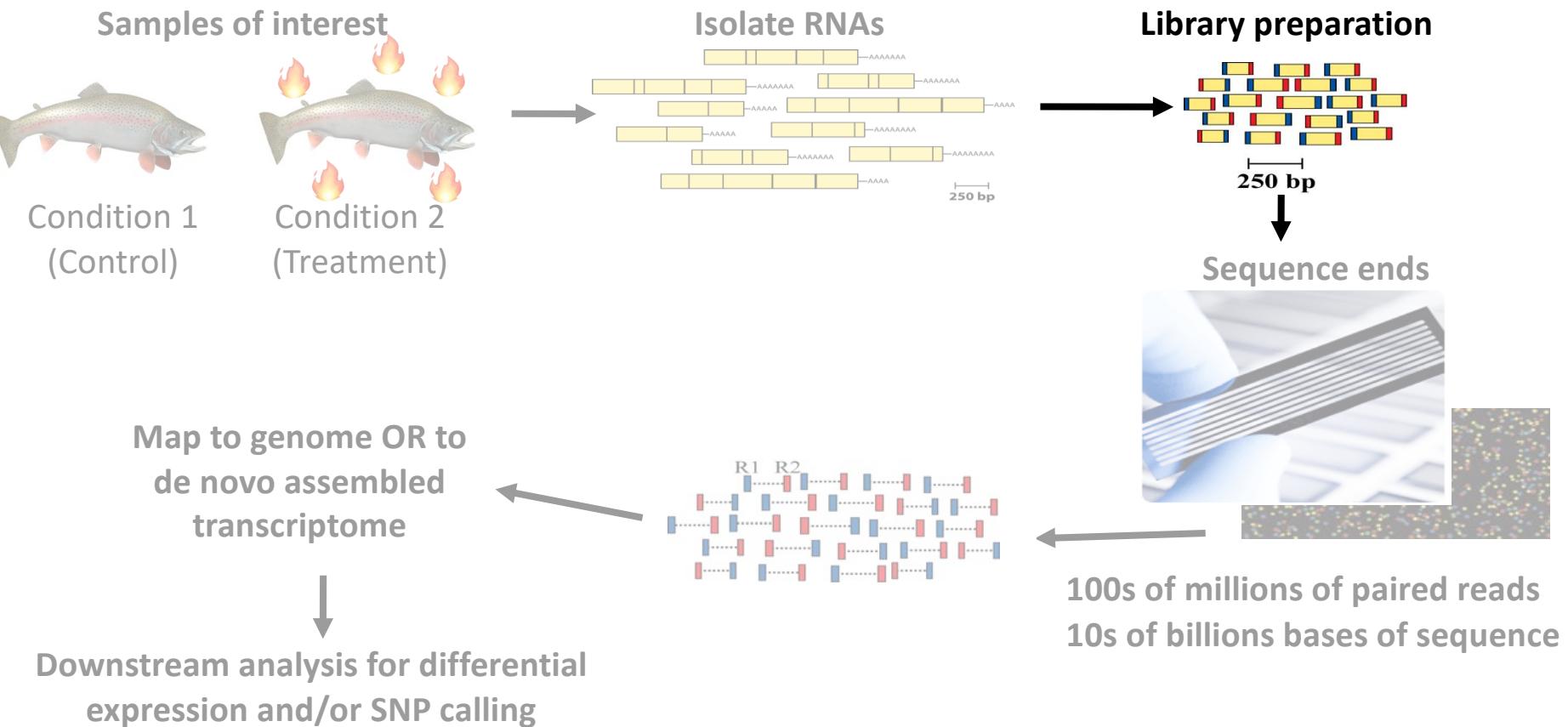
Poly-A+ selection



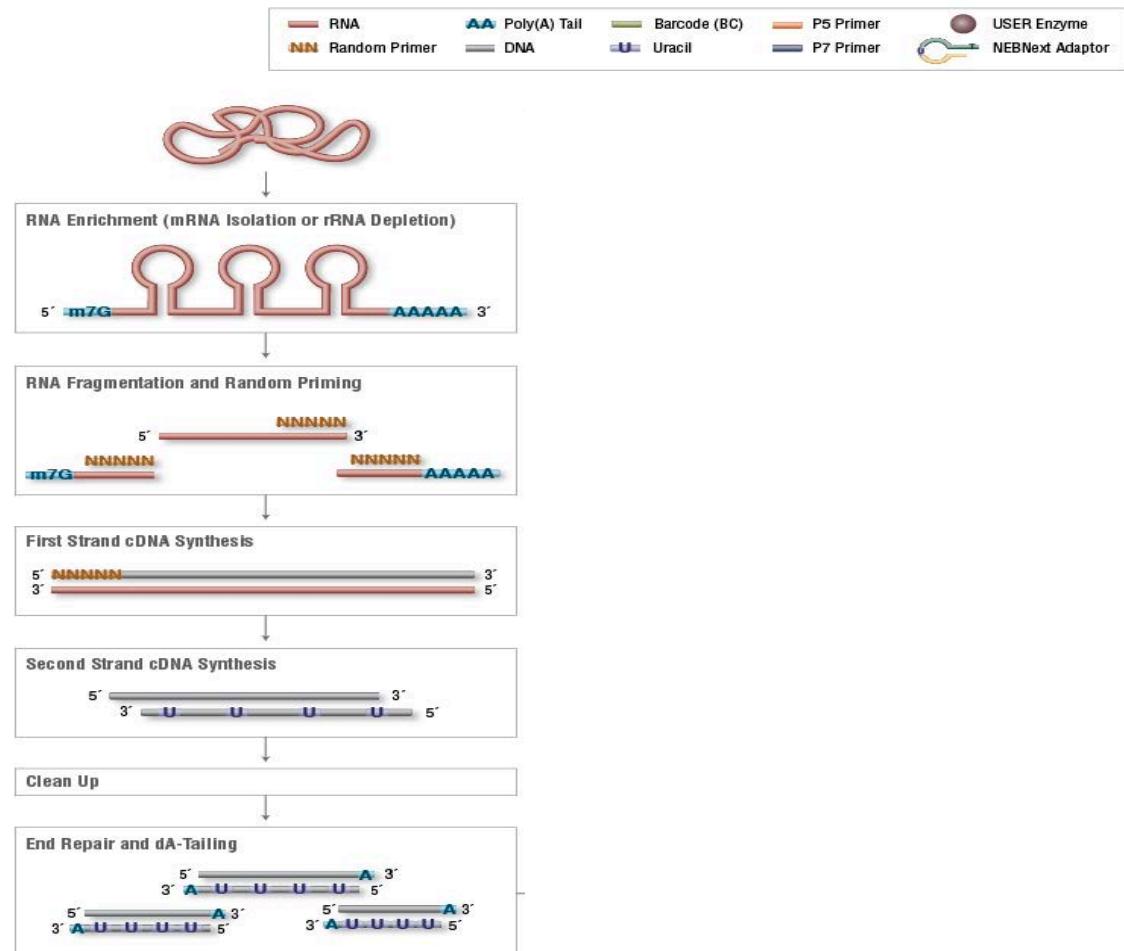
Ribosomal depletion



RNA sequencing experiment

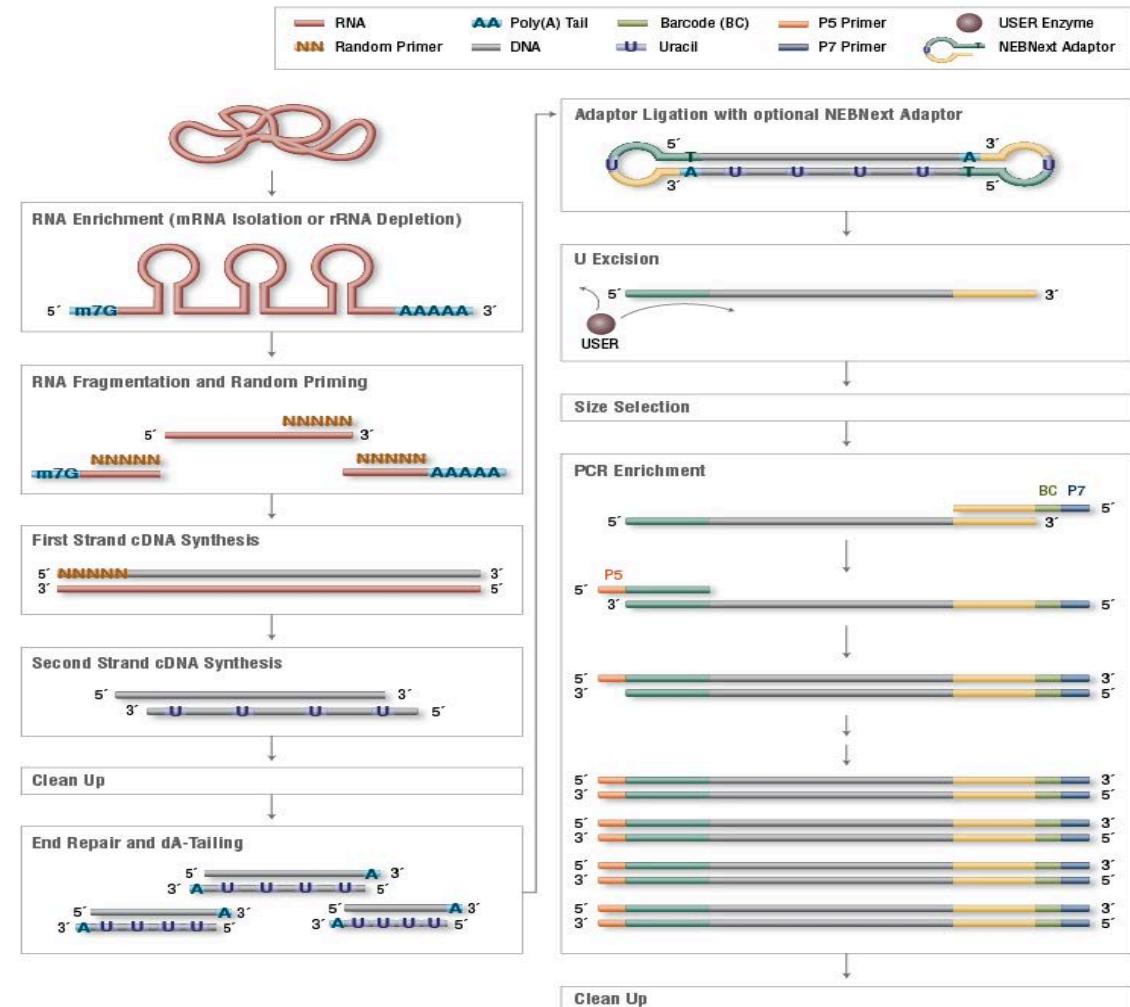


RNA-seq Library preparation process

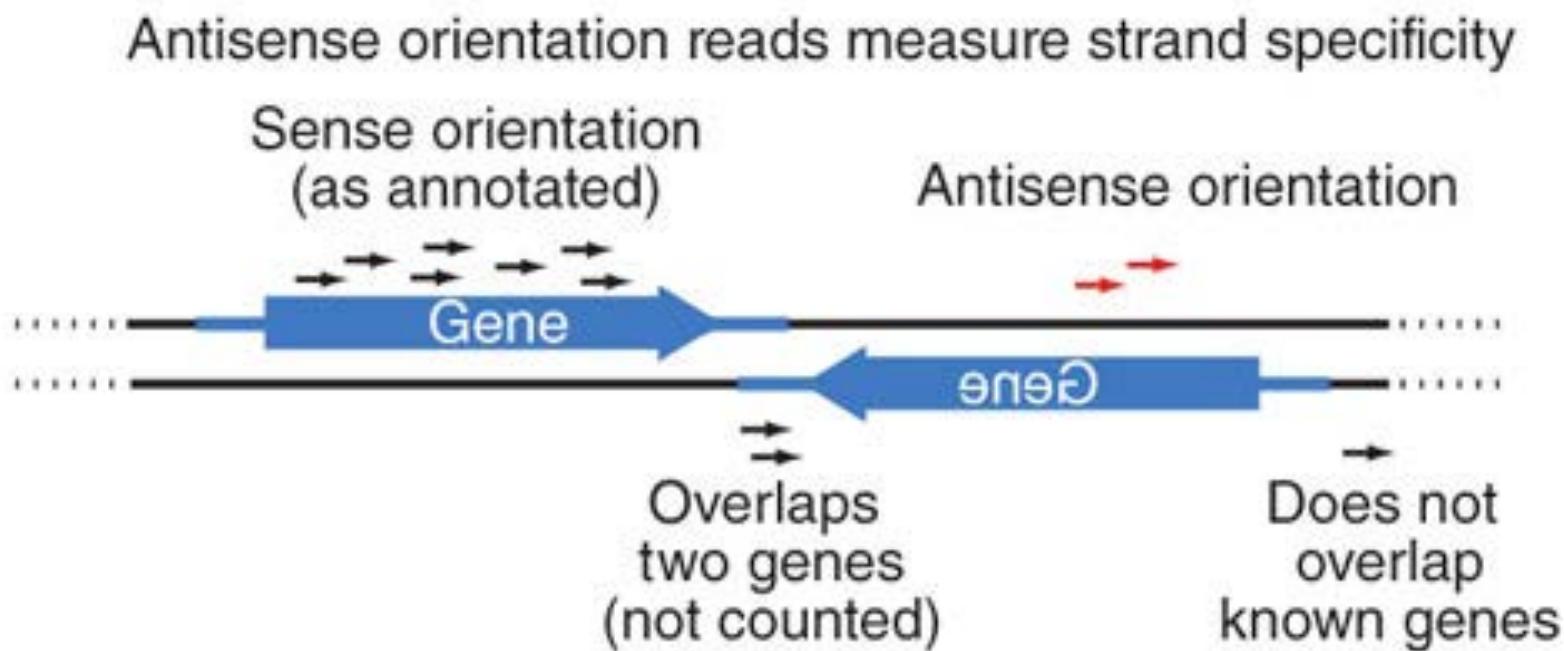


RNA-seq Library preparation process

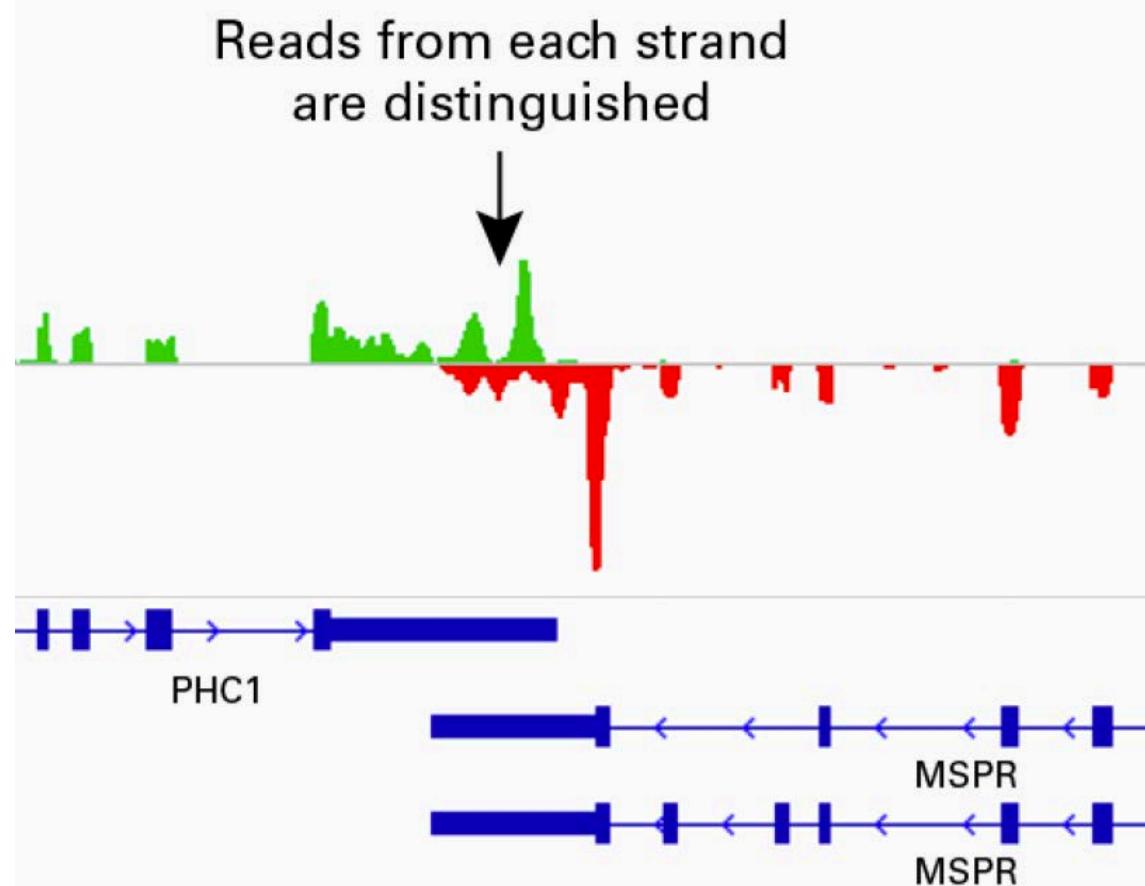
Directional (stranded)
vs non-directional



Possible insights from a directional RNAseq experiment

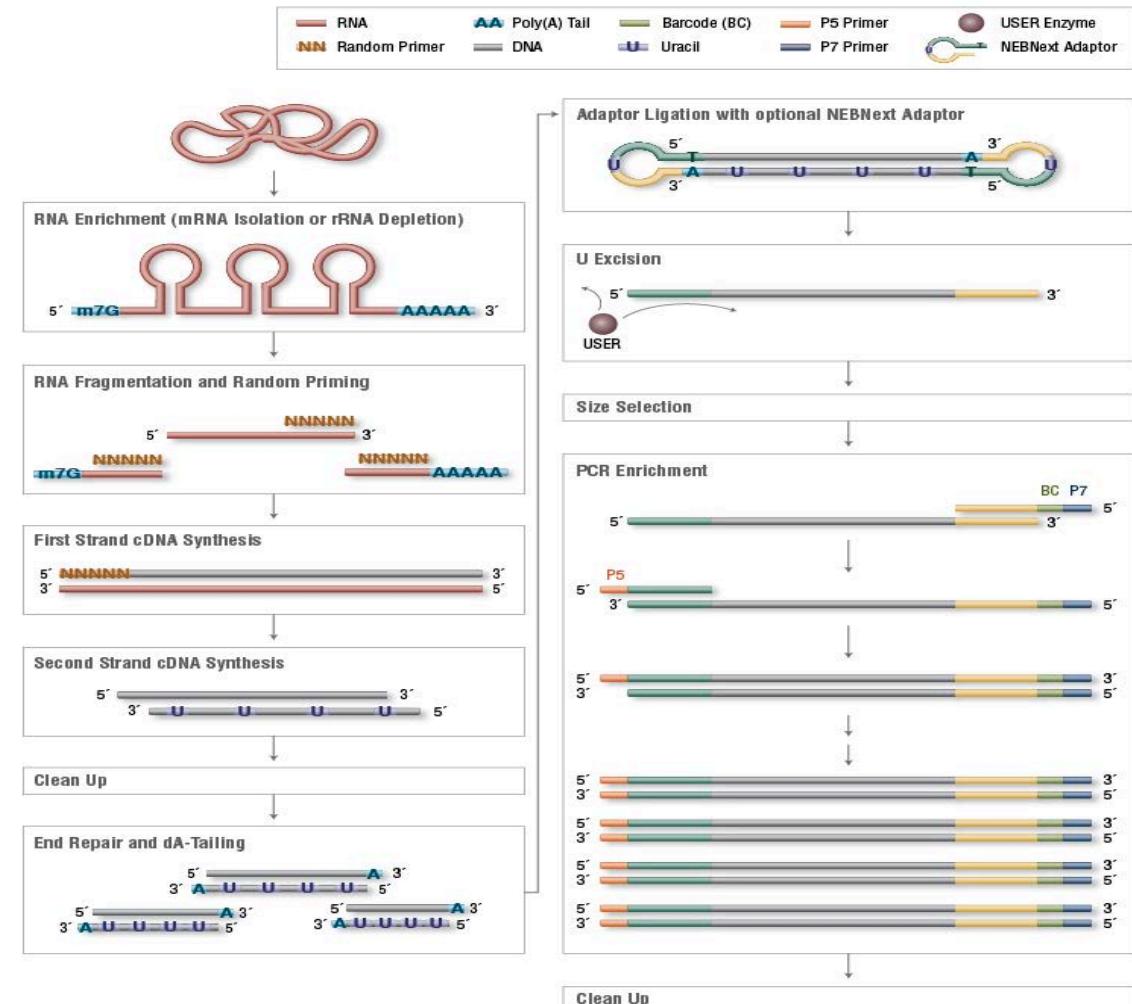


Possible insights from a directional RNAseq experiment

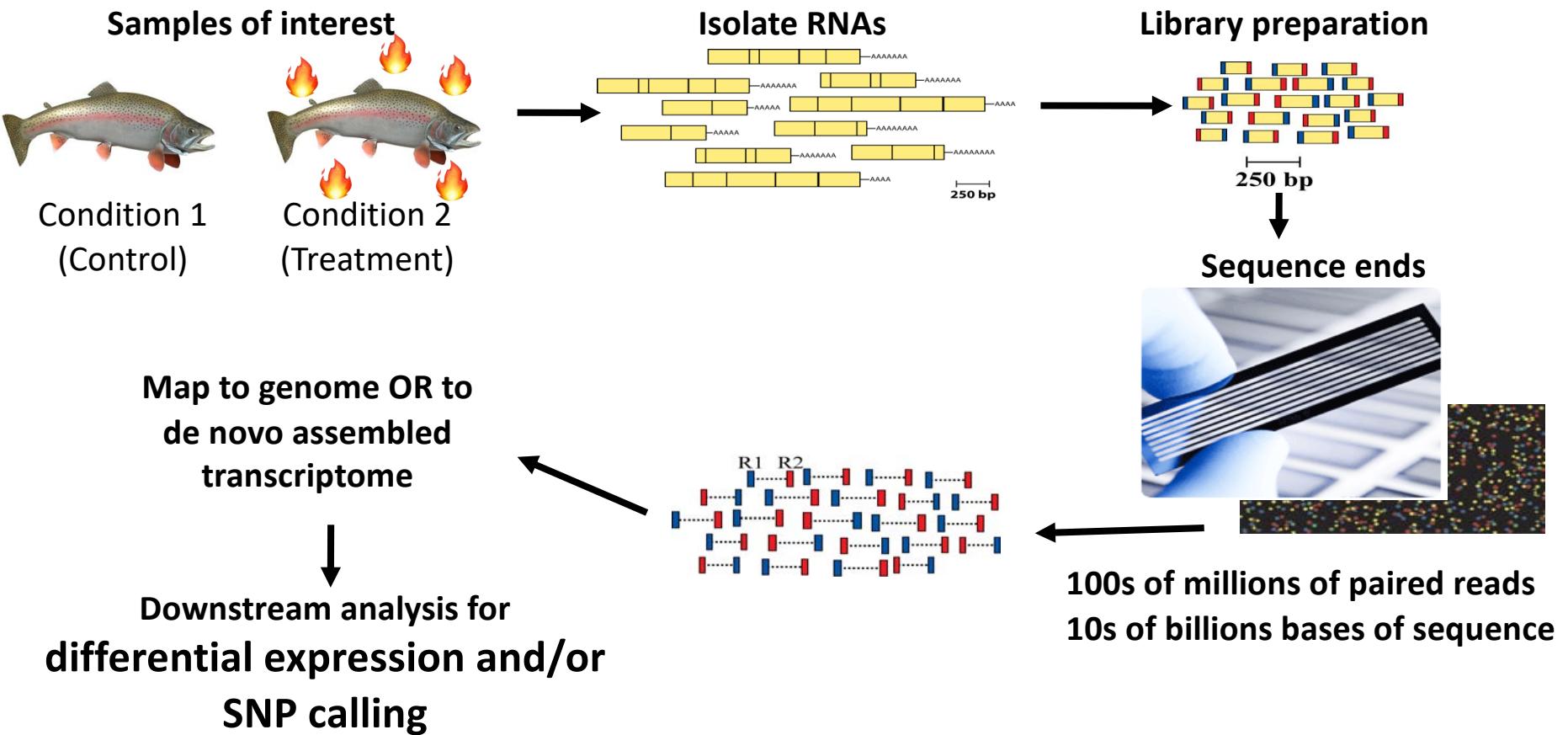


RNA-seq Library preparation process

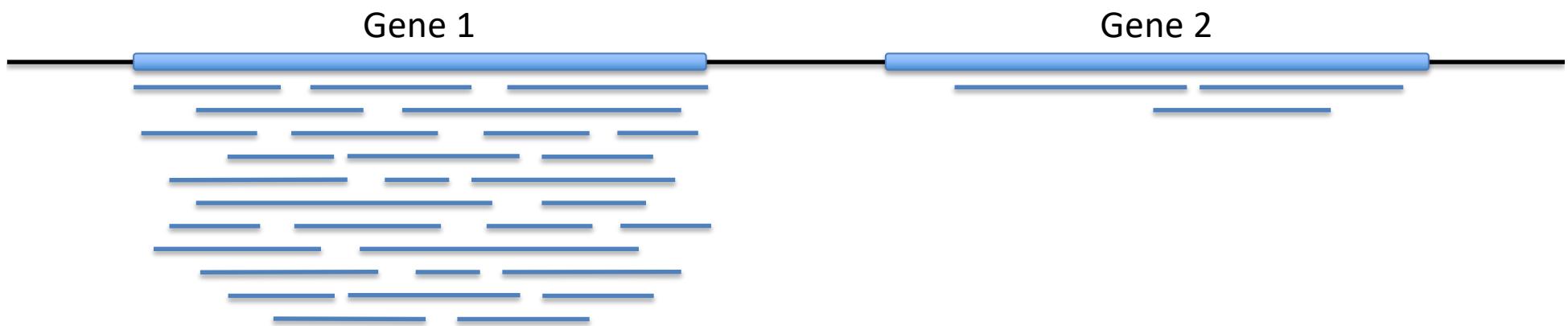
- Adapter ligation
- Barcoding
- Amplification



RNA sequencing experiment



Quantifying expression relies on mapping to a 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

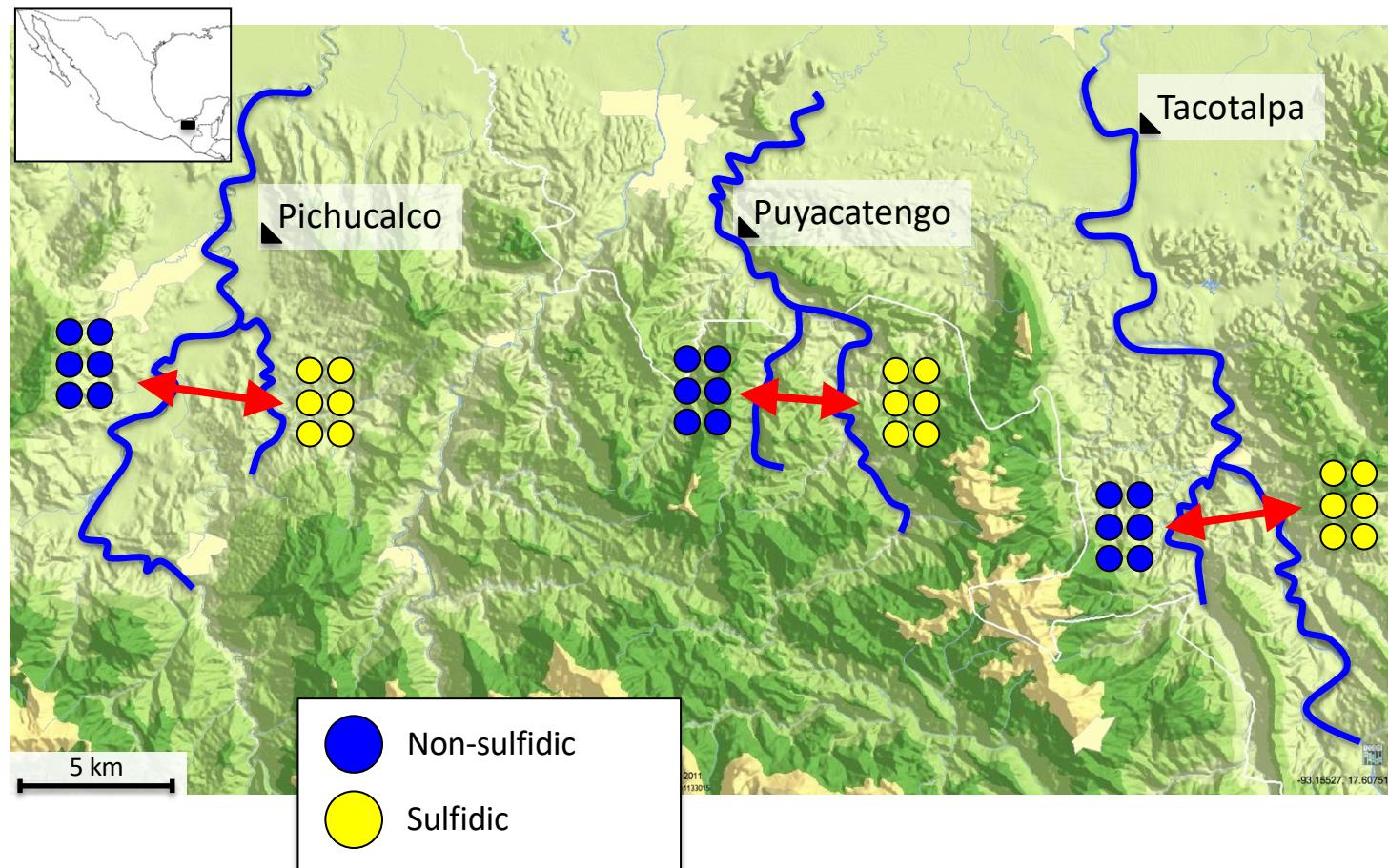
Advantages – annotations already done

No reference? No problem!

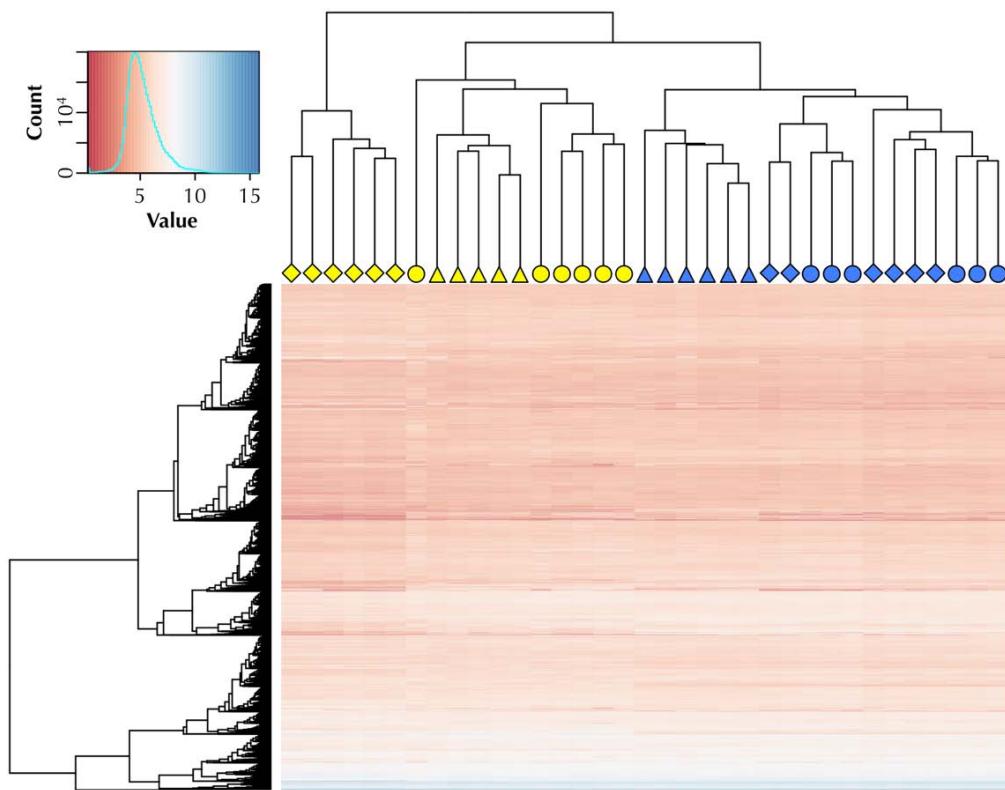
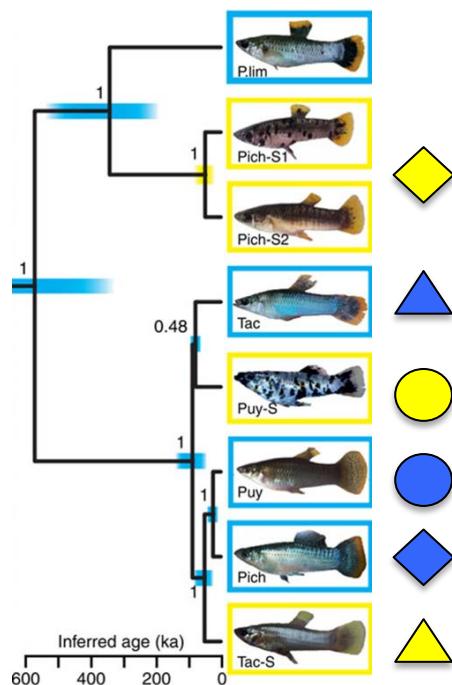
De novo reference transcriptome assembly!

- Trinity *de novo*
- Other programs

Identify differentially expressed genes between ecotypes



Unique patterns of gene expression in sulfide spring populations



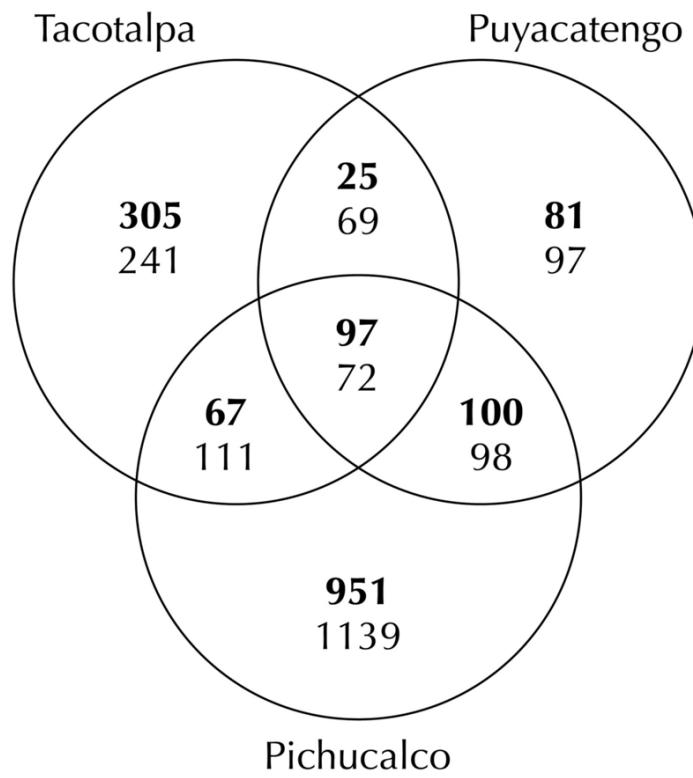
Kelley et al. 2016, MBE

Shared and unique variation in gene expression between sulfidic and non-sulfidic populations

Number of genes

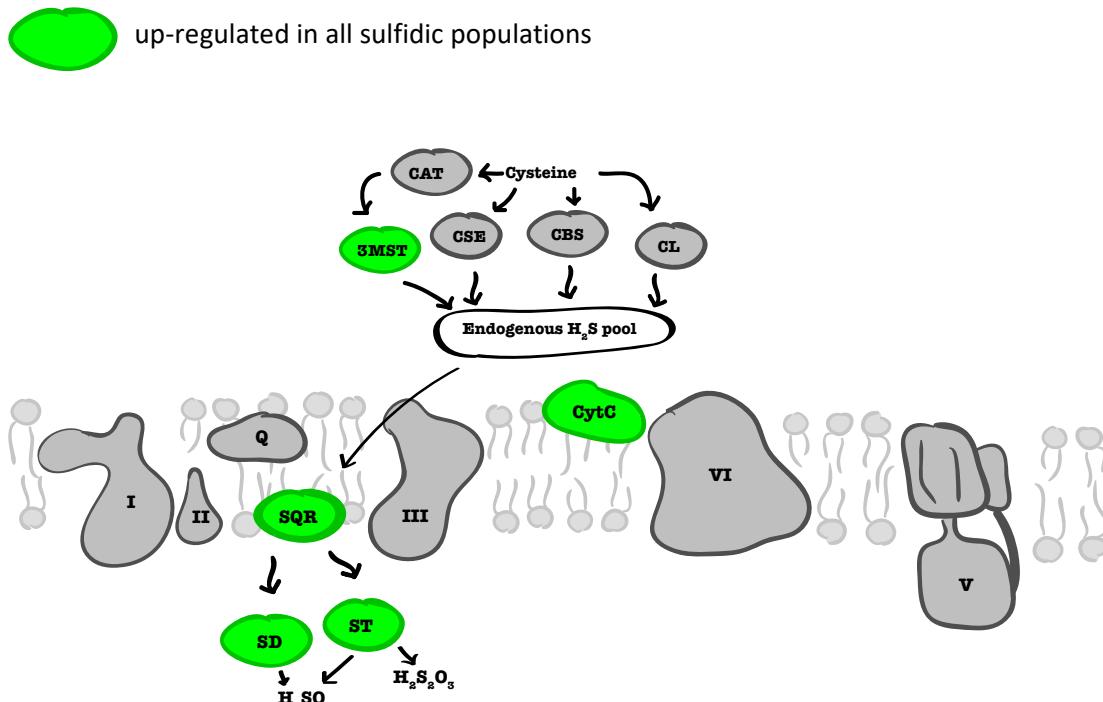
Up-regulated in H₂S

Down-regulated in H₂S



Kelley et al. 2016, MBE

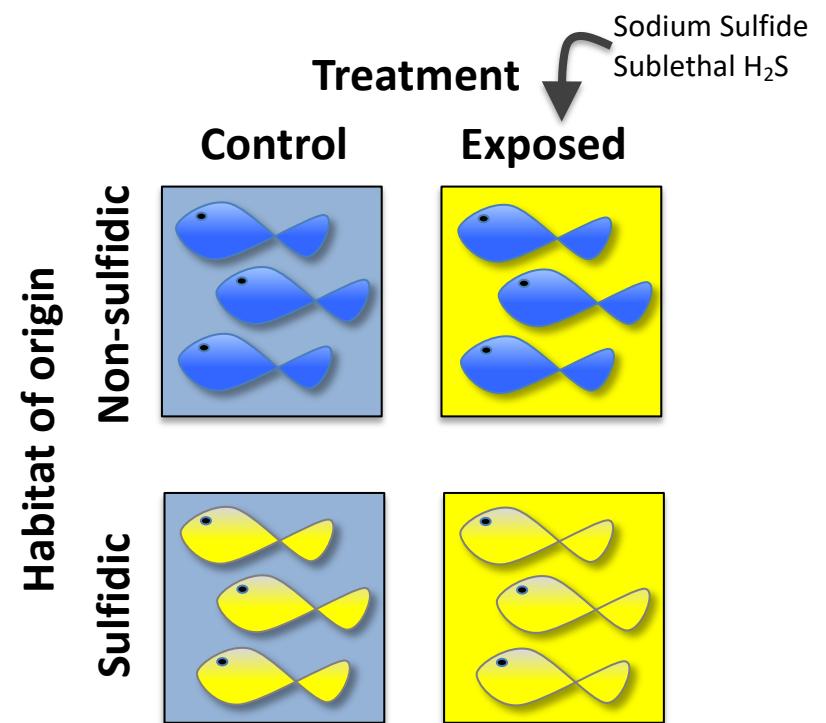
Shared expression of candidate genes



- No evidence for down-regulation of genes related to H_2S production
- Significant up-regulation of genes related to detoxification
- Consistent up-regulation of complex III and cytochrome c
 - Capturing electrons during blockage of COX

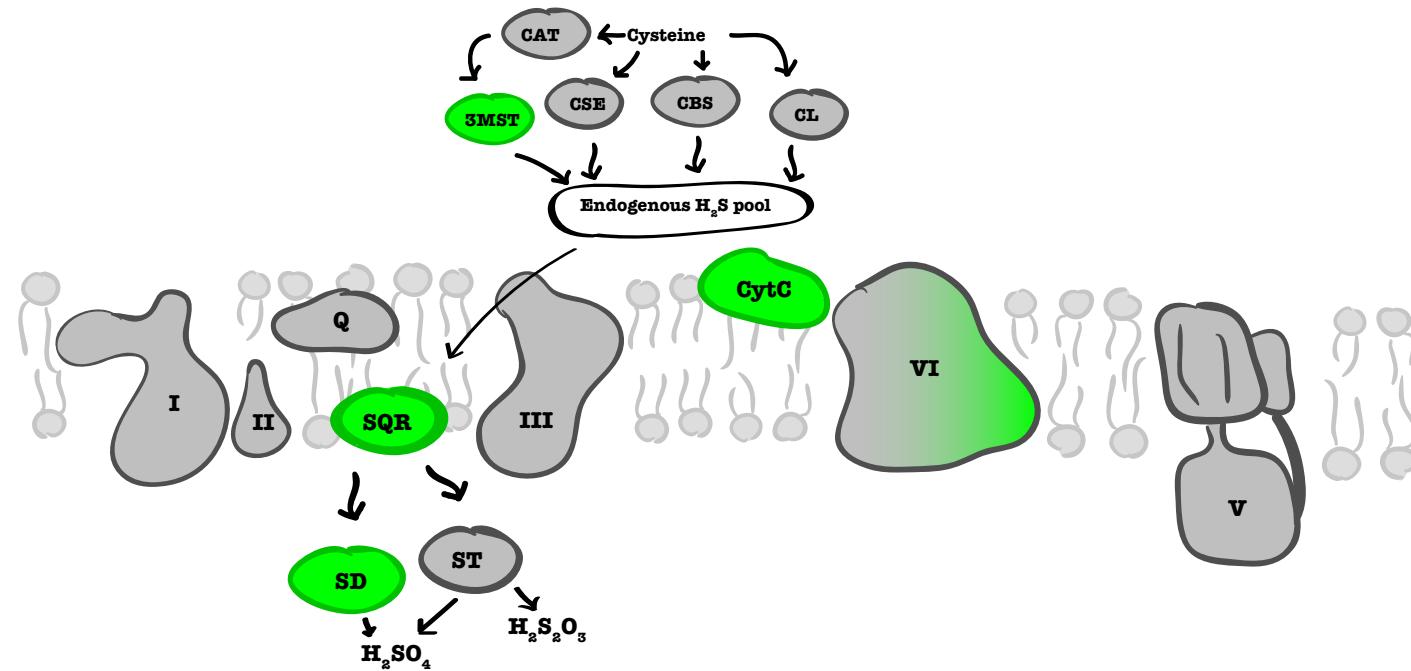
What differences in gene expression between ecotypes are adaptive versus plastic?

- Common garden rearing experiment for two of the *Poecilia mexicana* populations
- Exposure experiment of **lab-raised** fish to sulfidic and non-sulfidic conditions
- Determine the role of phenotypic plasticity vs heritable differences in gene expression?



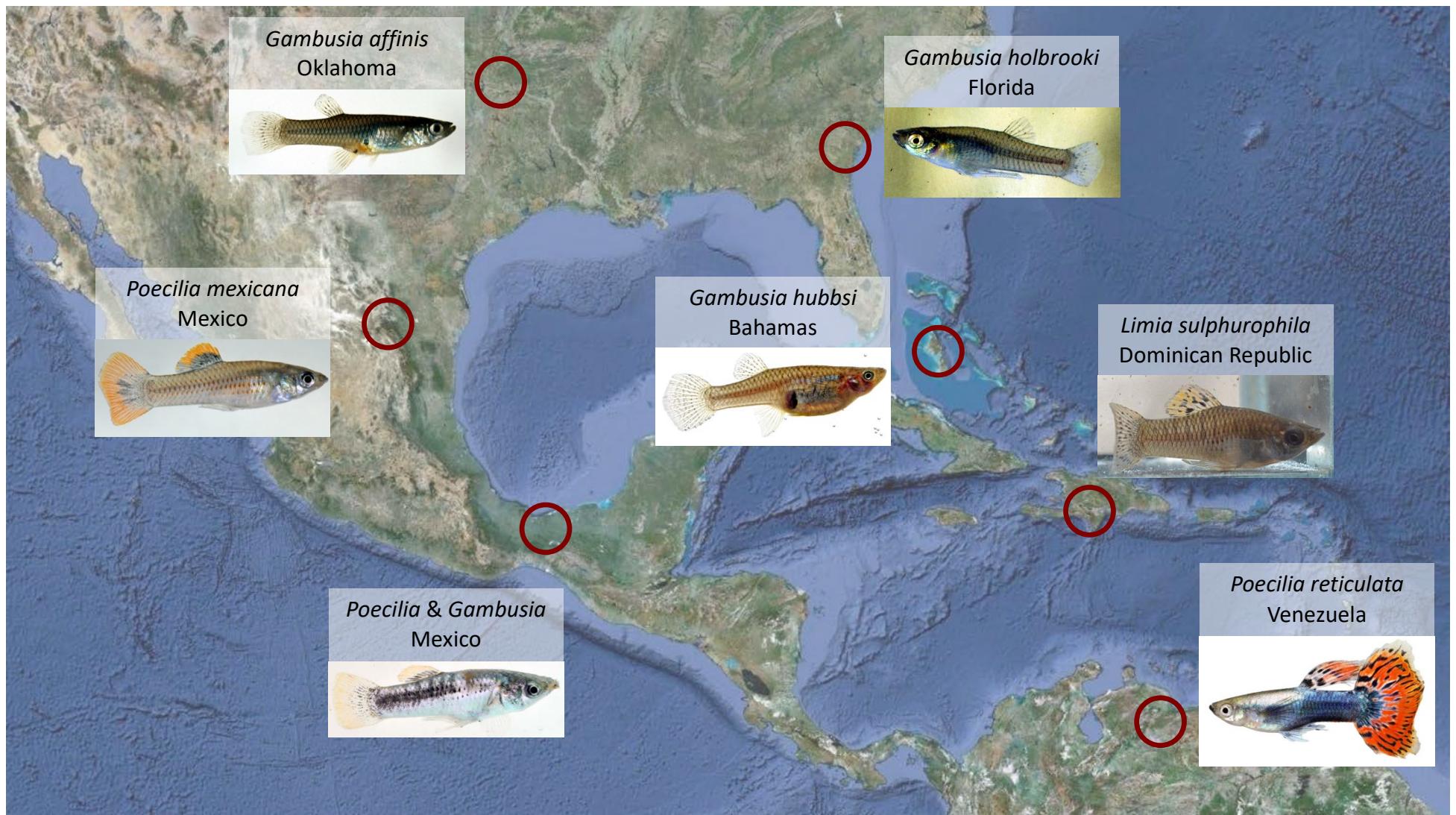
Passow et al 2017 *Molecular Ecology*

Evidence for evolved changes in gene expression in candidate genes



Mechanisms driving variation in gene expression

- Variation in differential expression is mainly driven by evolved differences, rather than ancestral plasticity.
- Although hydrogen sulfide is a strong source of selection and is driving some of the observed variation, there are other factors driving variation in differential gene expression.
- Significant progress toward understanding factors shaping gene expression variation in a natural system with strong physicochemical stressors



Acknowledgements



Michi Tobler, PhD
Kansas State University



Ryan Greenway
Kansas State University
(now EAWAG)



Courtney Passow, PhD
University of Minnesota



Anthony Brown
Washington State University
(now UCSF)

- Washington State University
 - Alexandra Fraik
 - Kerry McGowan
 - Corey Quackenbush
- Kansas State University
 - Nick Barts
- Univ. Juárez Autonoma de Tabasco, MX
 - Lenin Arias Rodrigues
- University of Frankfurt
 - Martin Pfenninger
 - Martin Plath



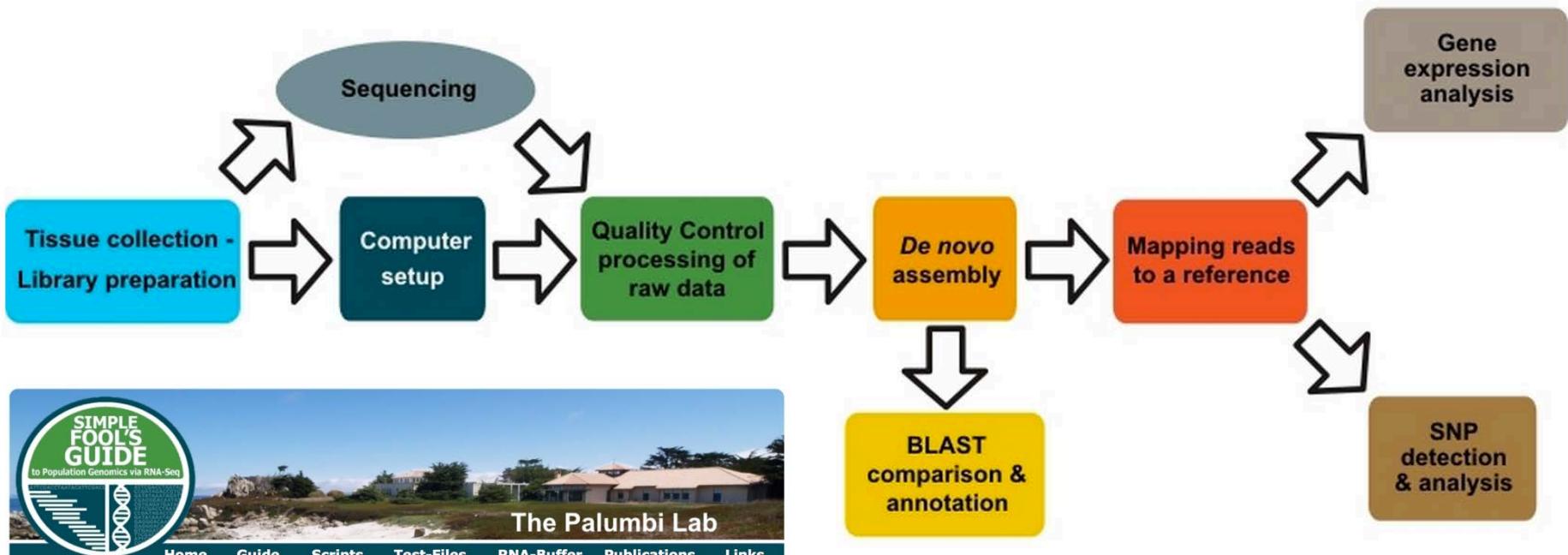
NSF: 1557795, 1931650





© Robbie Stone

Joanna Kelley, joanna.l.kelley@wsu.edu, <http://labs.wsu.edu/genomes>



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.

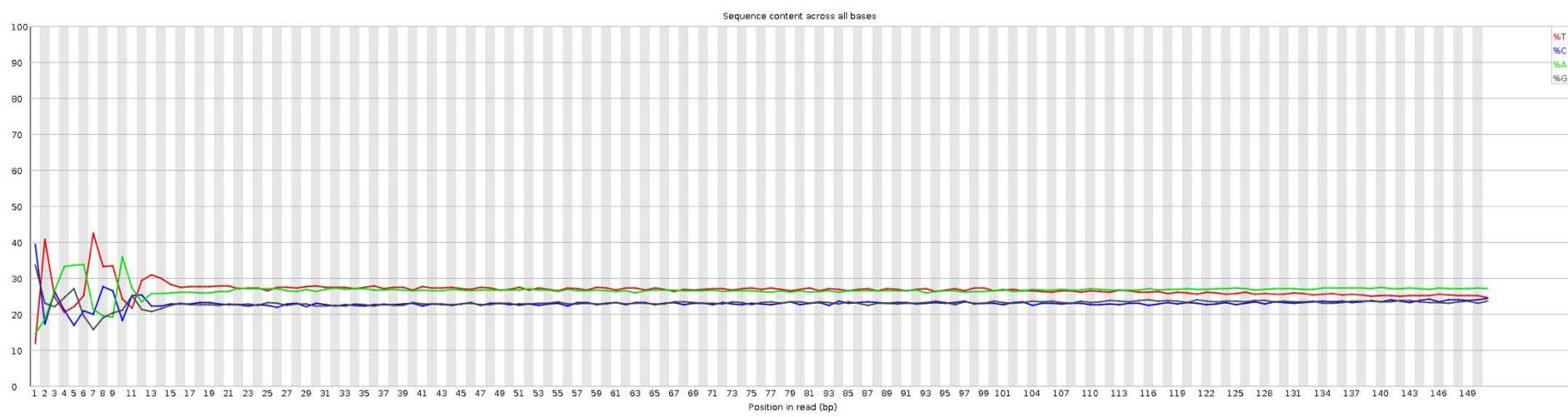


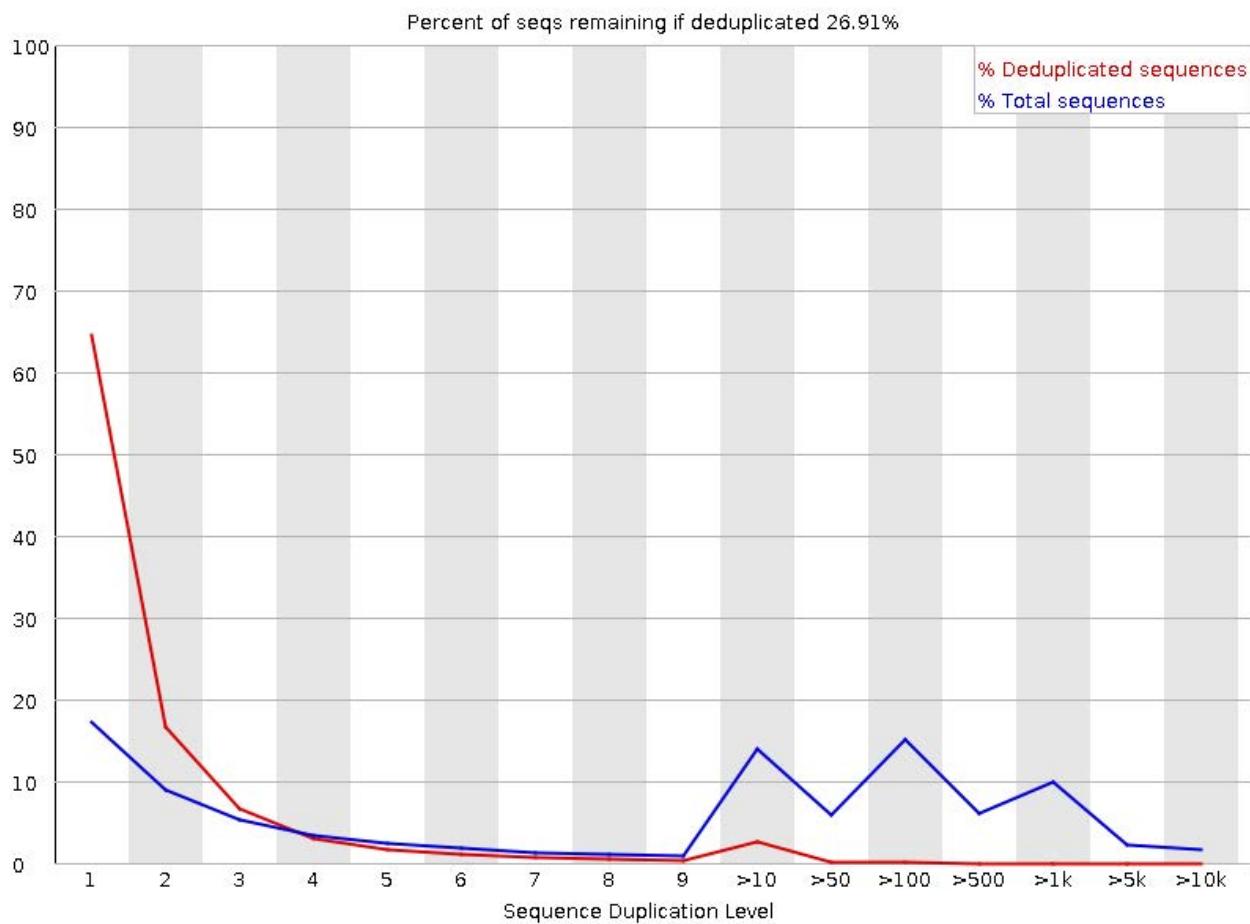
You may cite this work as follows:

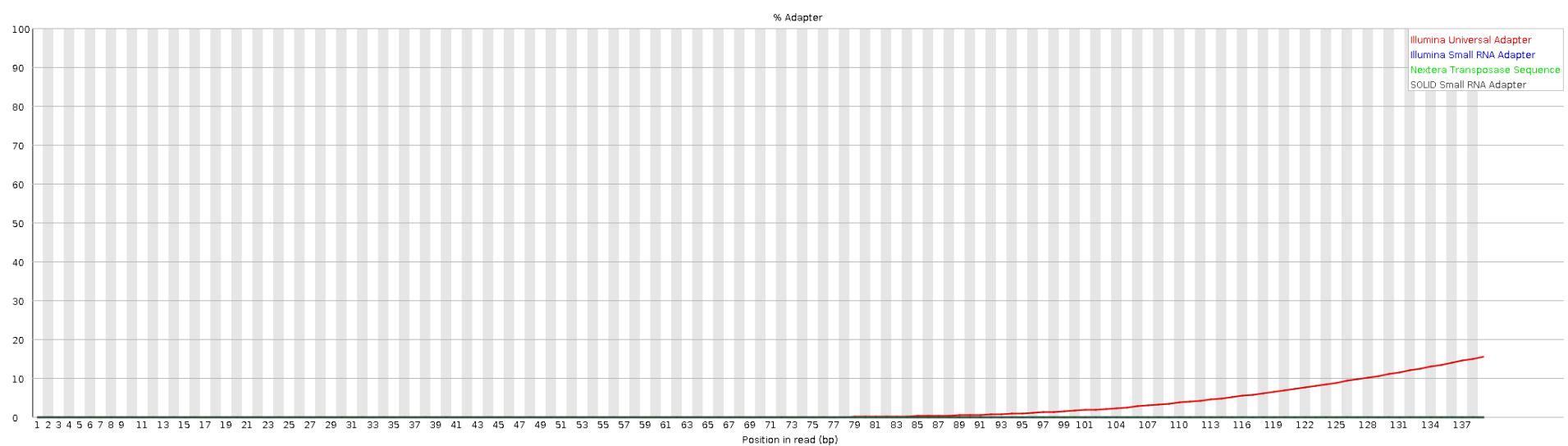
De Wit P, Pespeli MH, Ladner JT, Barshis DJ, Seneca F, Jaris H, Overgaard Therkildsen N, Morikawa M and Palumbi SR (2012) The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Molecular Ecology Resources* **12**, 1058-1067.

Data analysis (there are no pipelines)

- Demultiplex (should be done by the sequencing core)
- Look at the data (for example, fastQC)







Data analysis (there are no pipelines)

- Trim adapters
- Trim low quality reads and/or bases
- Look at the data again

Data analysis (there are no pipelines)

- Mapping
- Look at the data again (IGV)
- Determining differentially expressed genes

- Continue with **worksheet!**