

Genomics on the Half Shell

Environmental Epigenetics, Open Science, and the Oyster

Steven Roberts
Associate Professor
University of Washington
School of Aquatic and Fishery Sciences
robertslab.info
@sr320

OSU IB Seminar - May 4, 2015

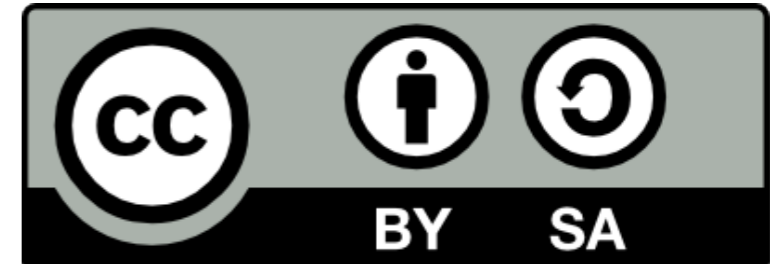


Open Science

- You are free to Share!
- Our lab practices open notebook science



IP[y]: IPython
Interactive Computing



OpenNotebookScience

@ONScience FOLLOWS YOU

Updates from #openscience notebooks,
brought to you by @openscience. Start
yours at onsnetwork.org/joinons/

onsnetwork.org

- Data, Preprints, Proposals, Lab Meetings, Web Cams, Slidedecks

These slides plus links @
github.com/sr320/talk-osu-2015

Background

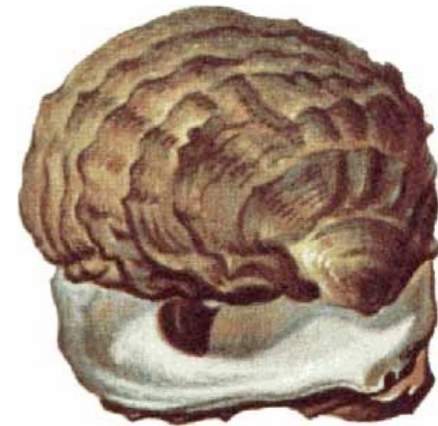
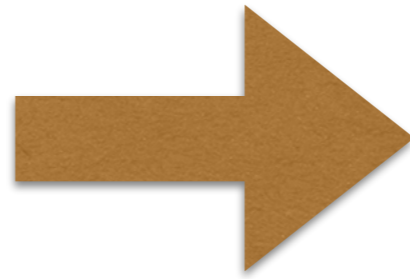
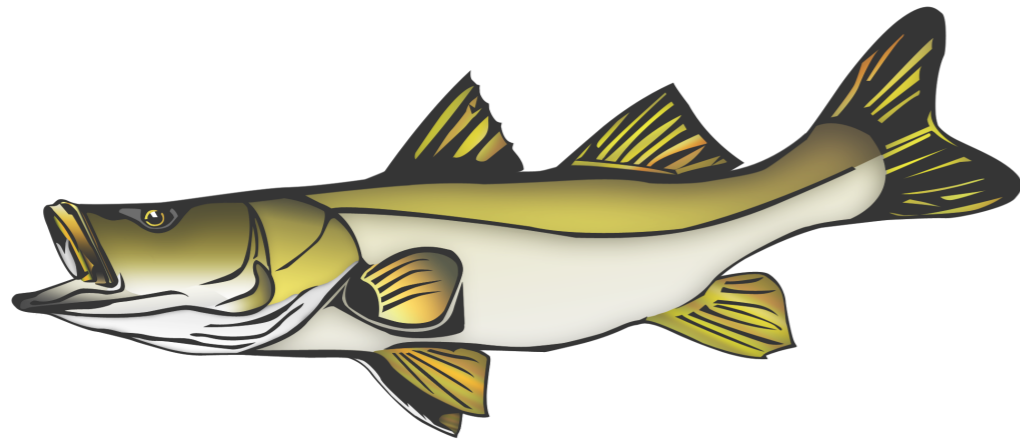
Physiology

How fundamental processes work in aquatic species

Background

Physiology

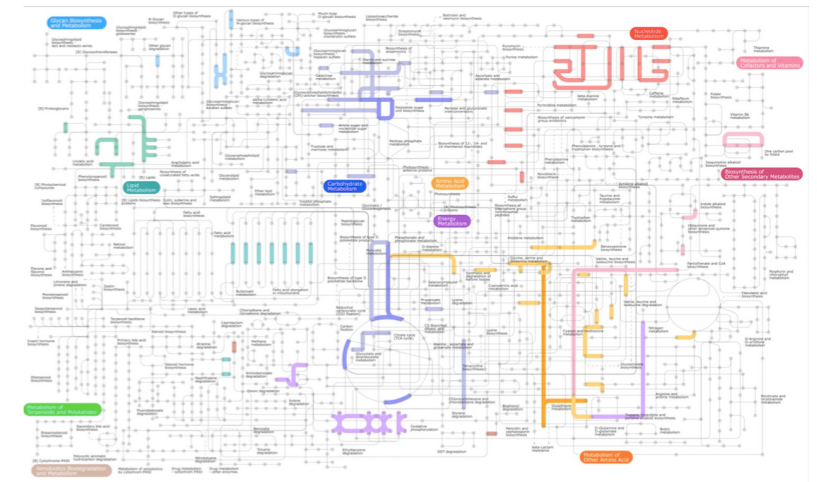
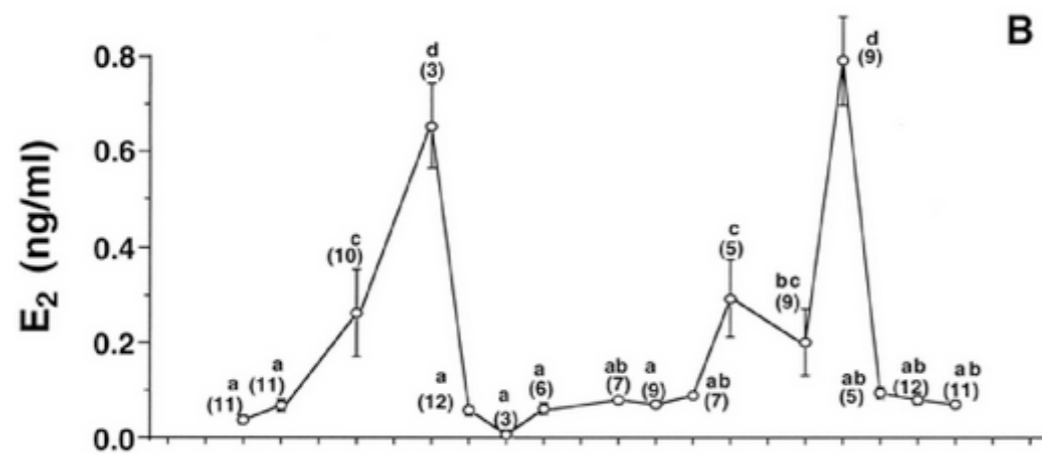
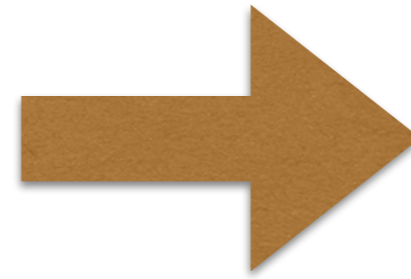
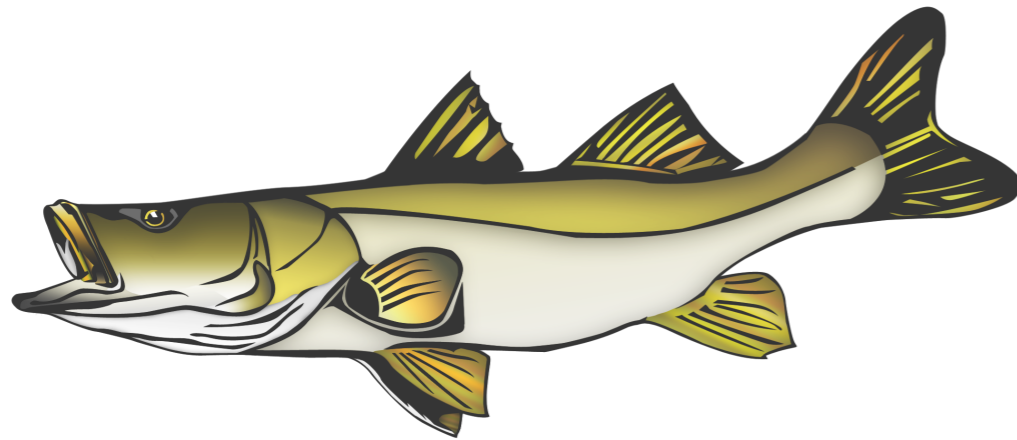
How fundamental processes work in aquatic species



Background

Physiology

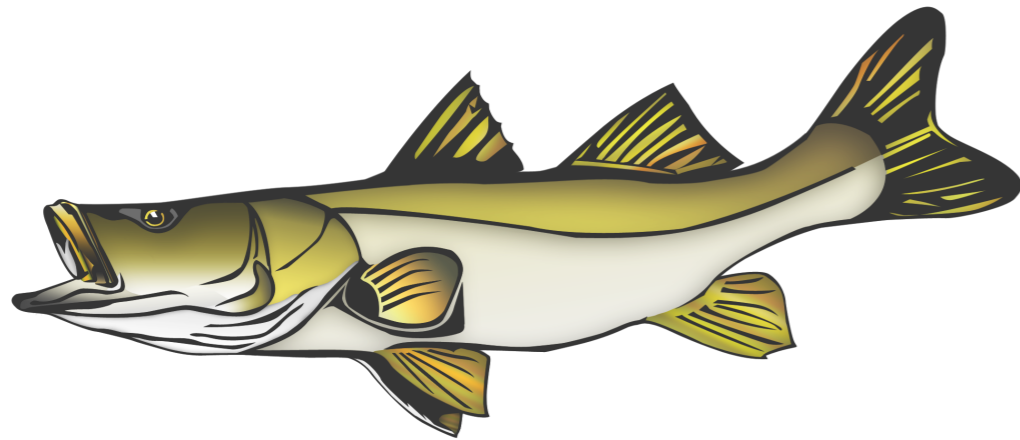
How fundamental processes work in aquatic species



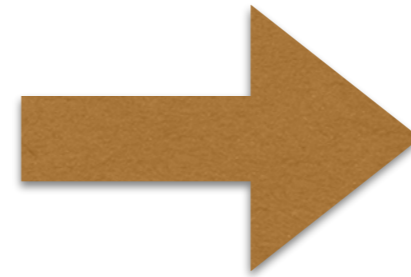
Background

Physiology

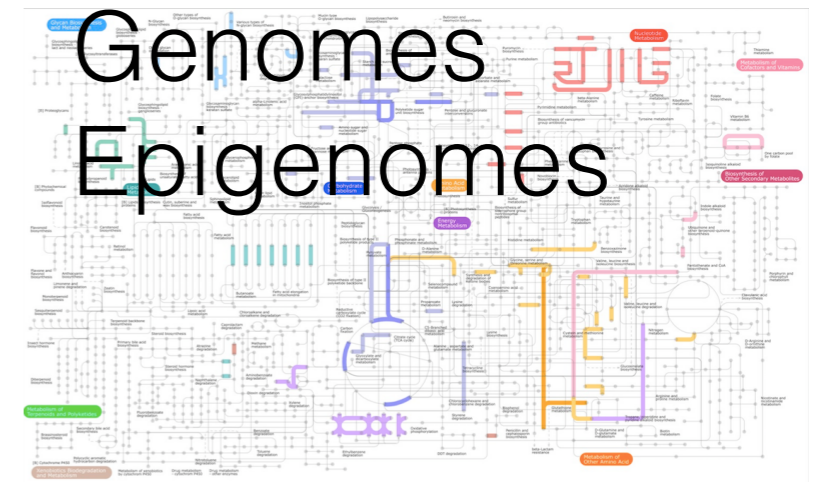
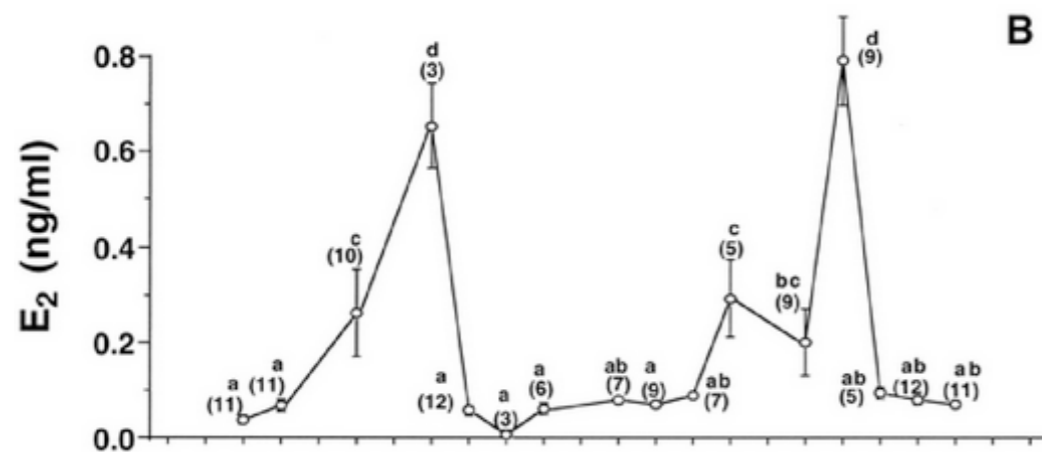
How fundamental processes work in aquatic species



Hormones
Proteins



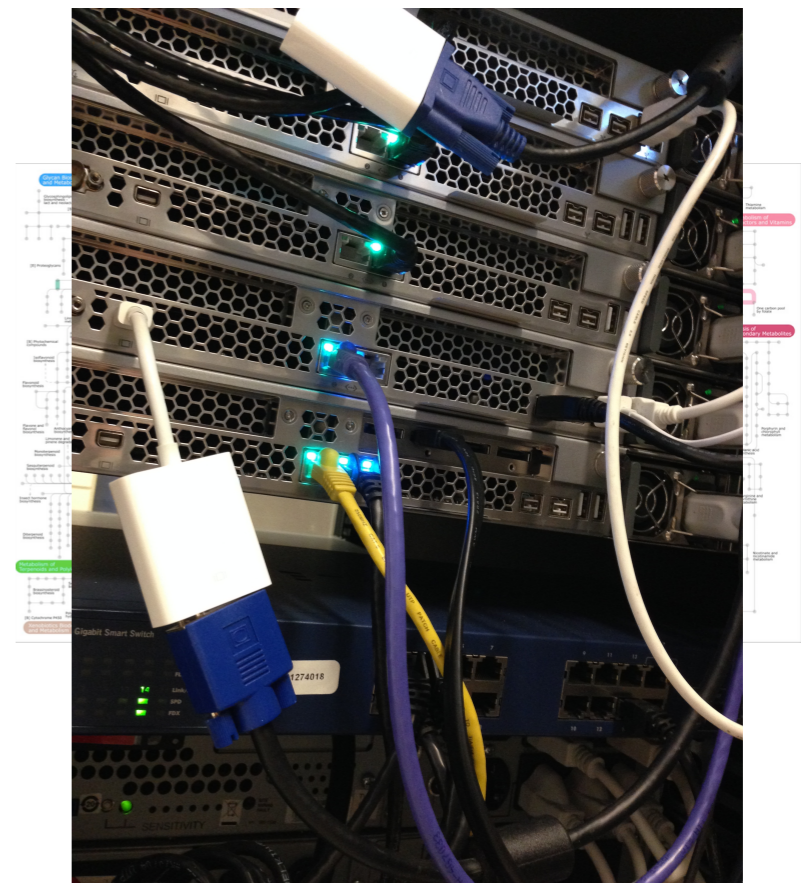
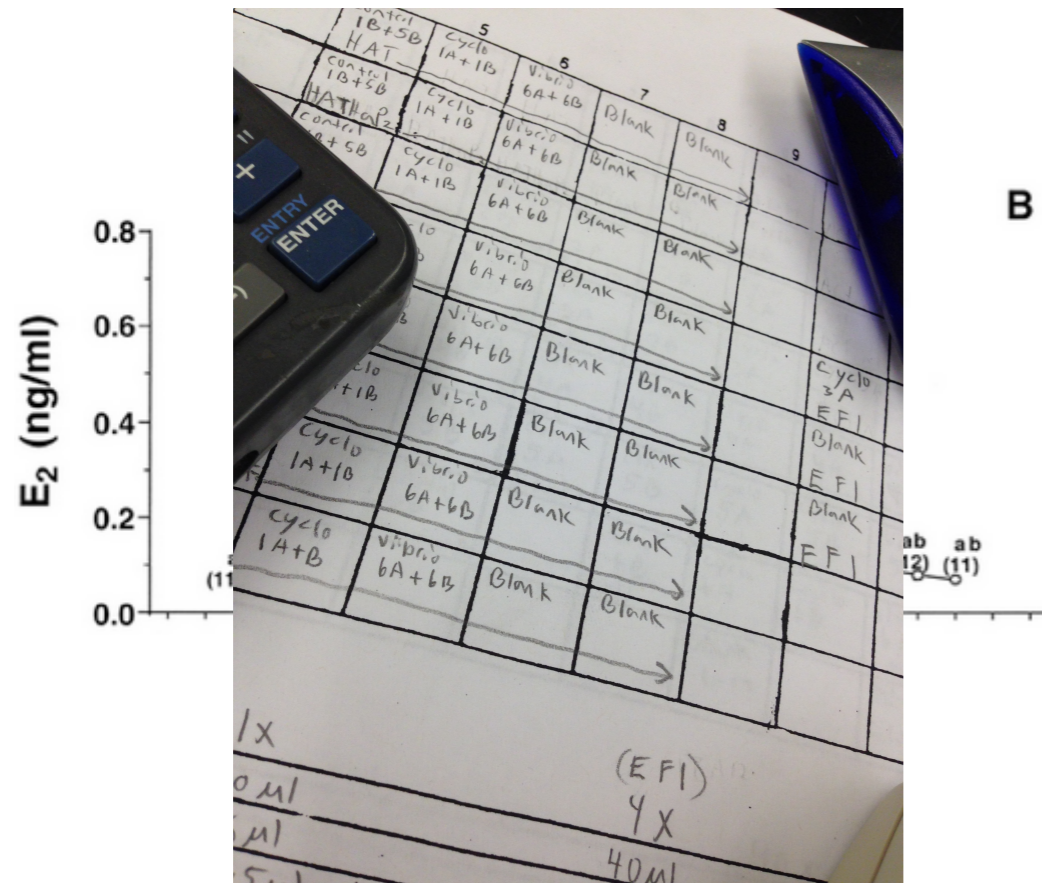
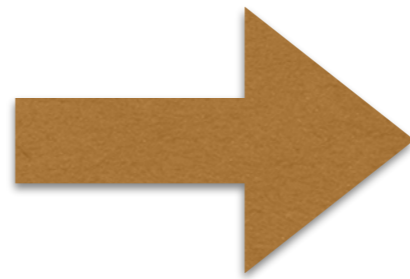
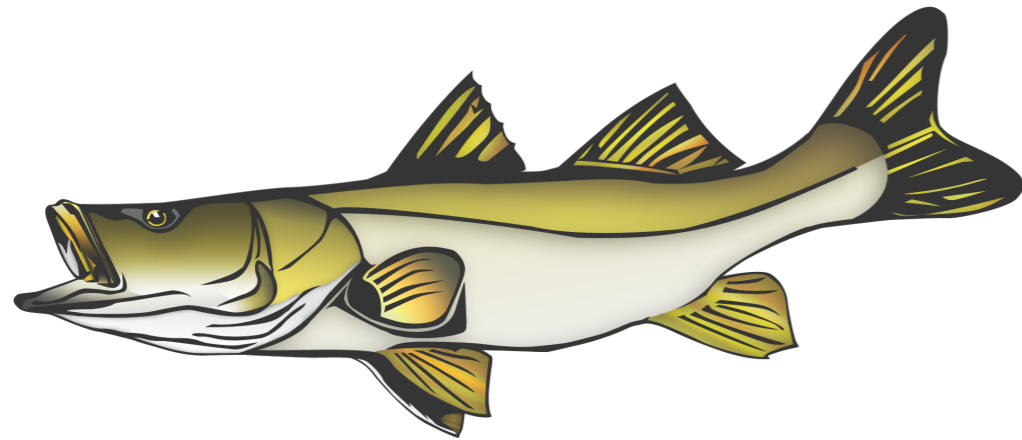
Transcriptomes
Proteomes

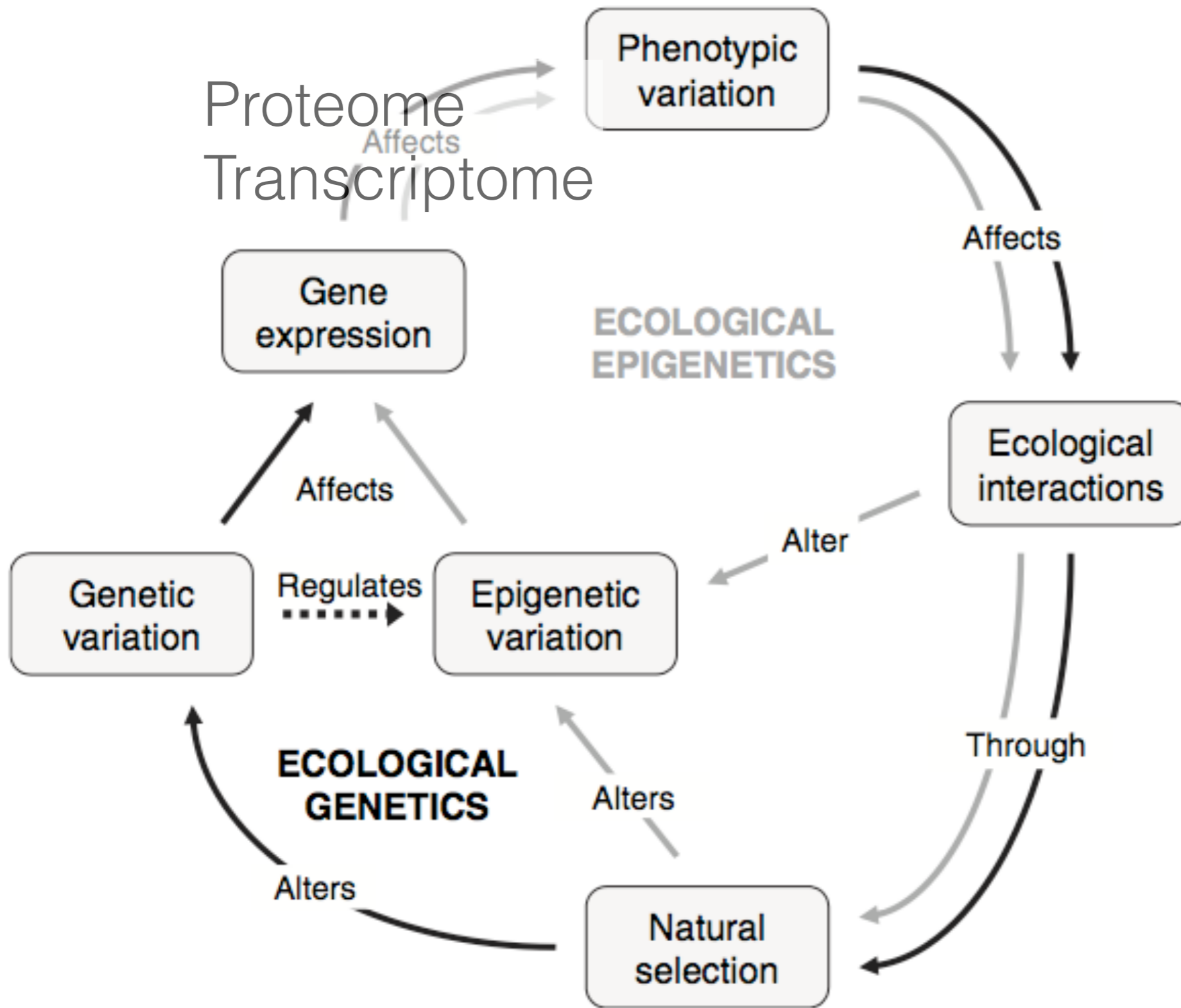


Background

Physiology

How fundamental processes work in aquatic species



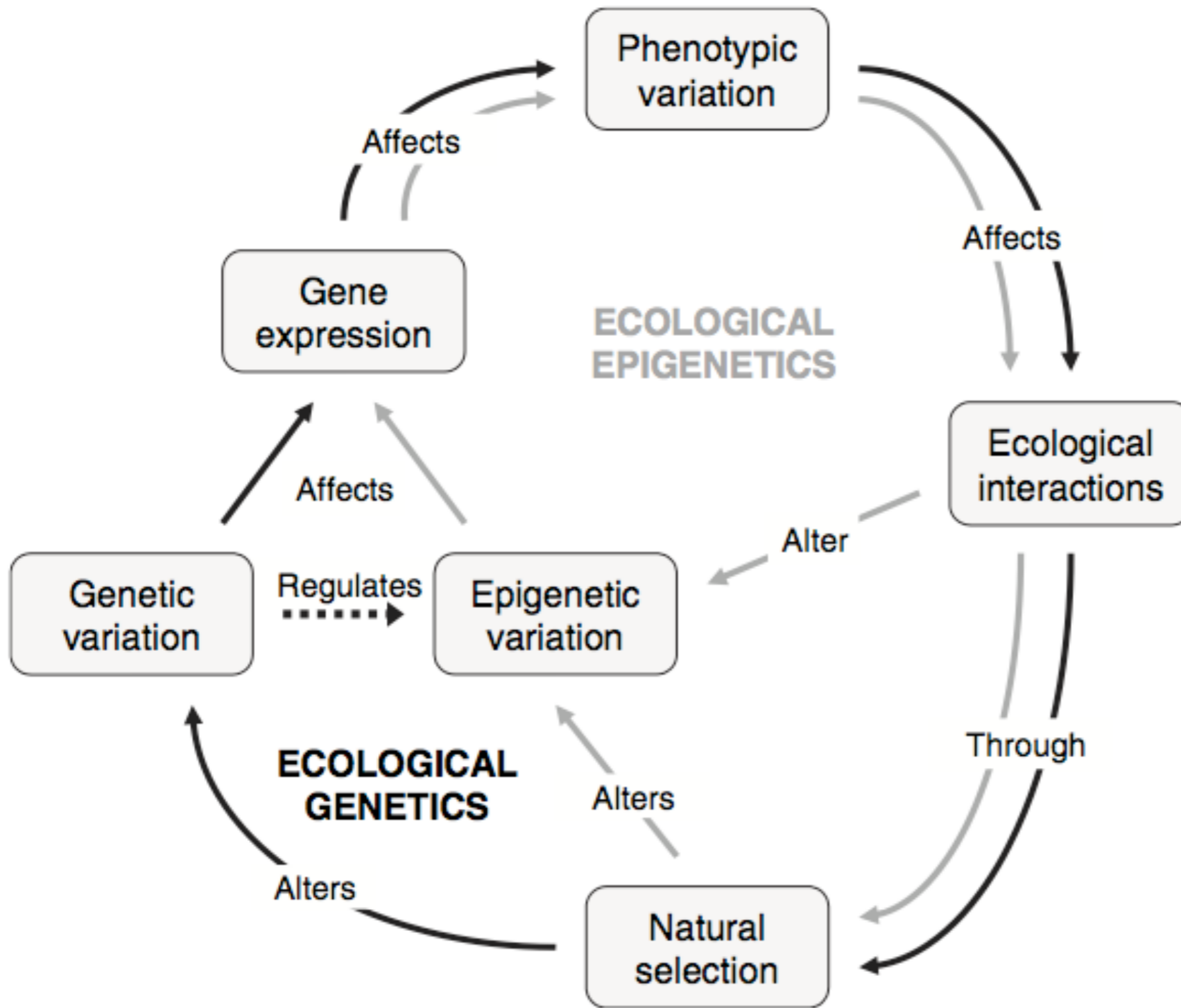


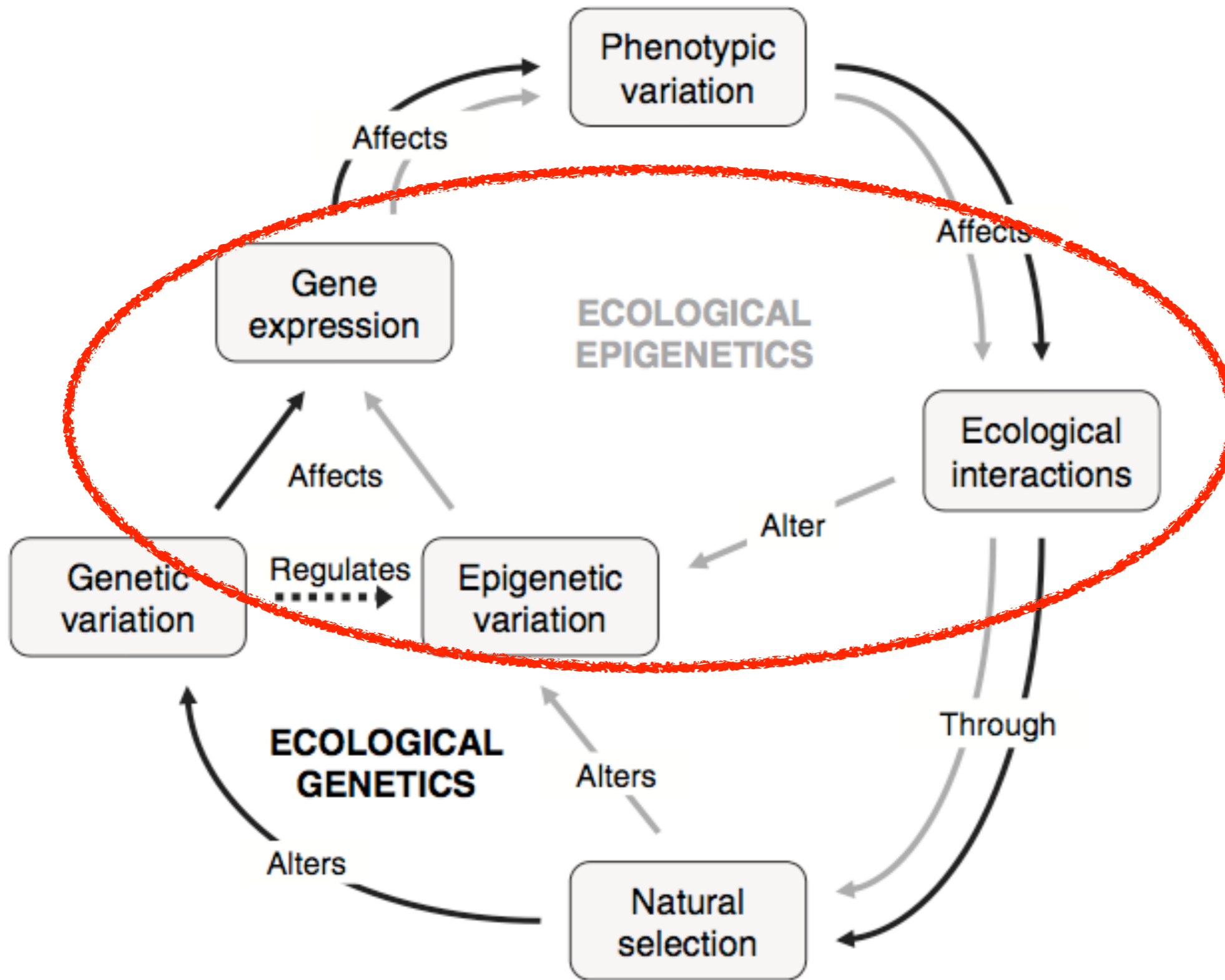
modified from

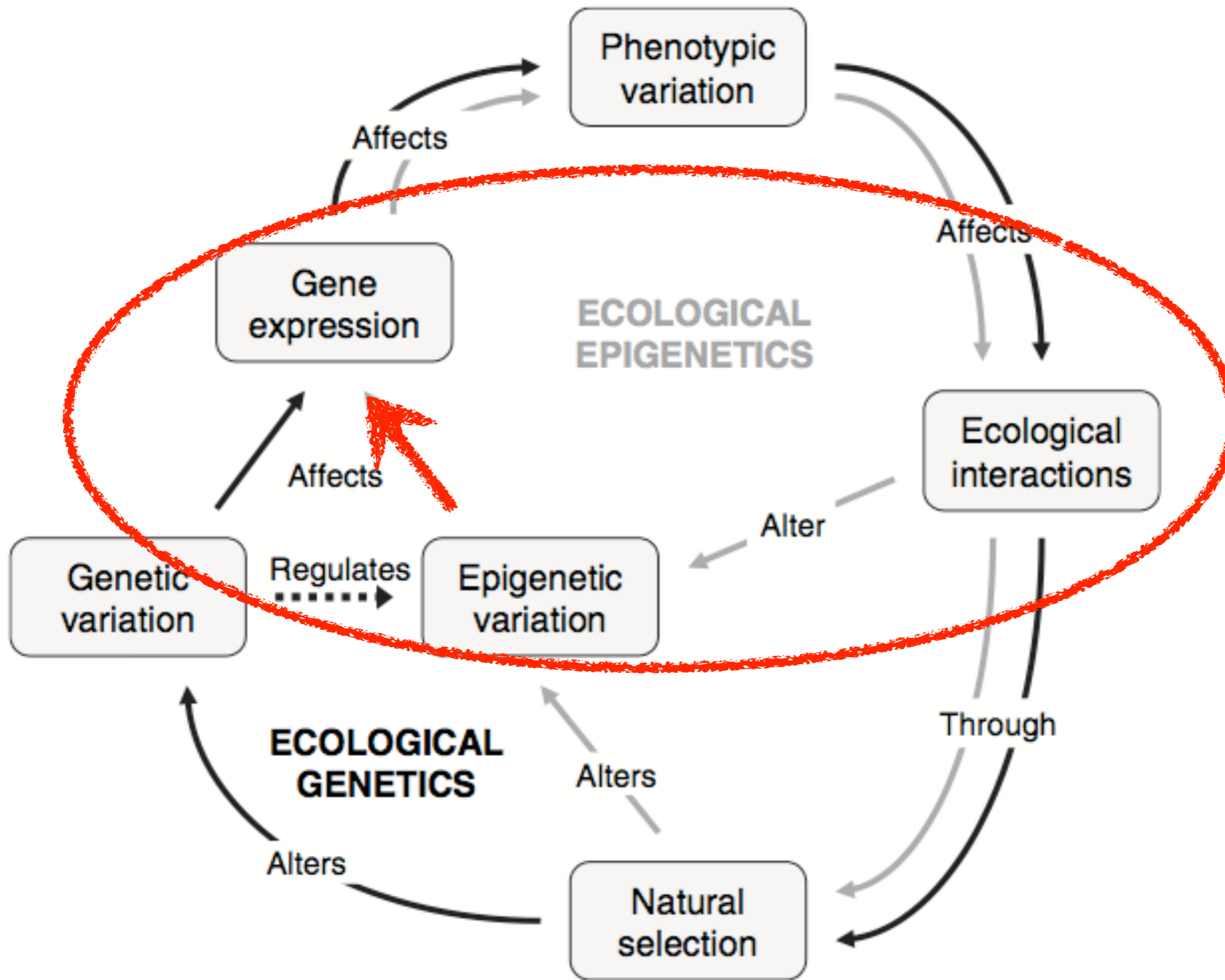
IDEA AND PERSPECTIVE

Epigenetics for ecologists

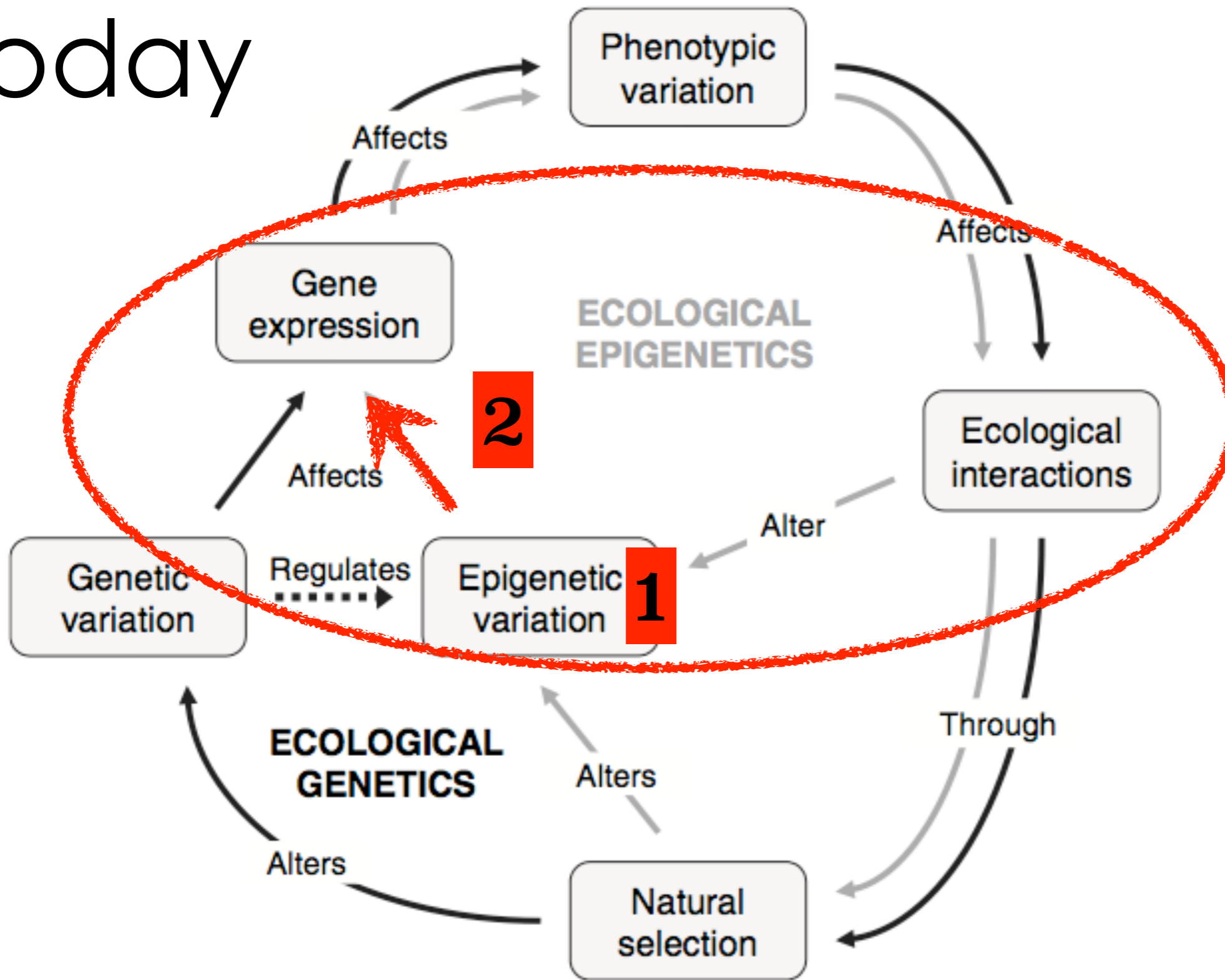
Oliver Bosdorp,^{1*} Christina L. Richards² and Massimo Pigliucci³







Today



Ecology Letters, (2008) 11: 106–115

doi: 10.1111/j.1461-0248.2007.01130.x

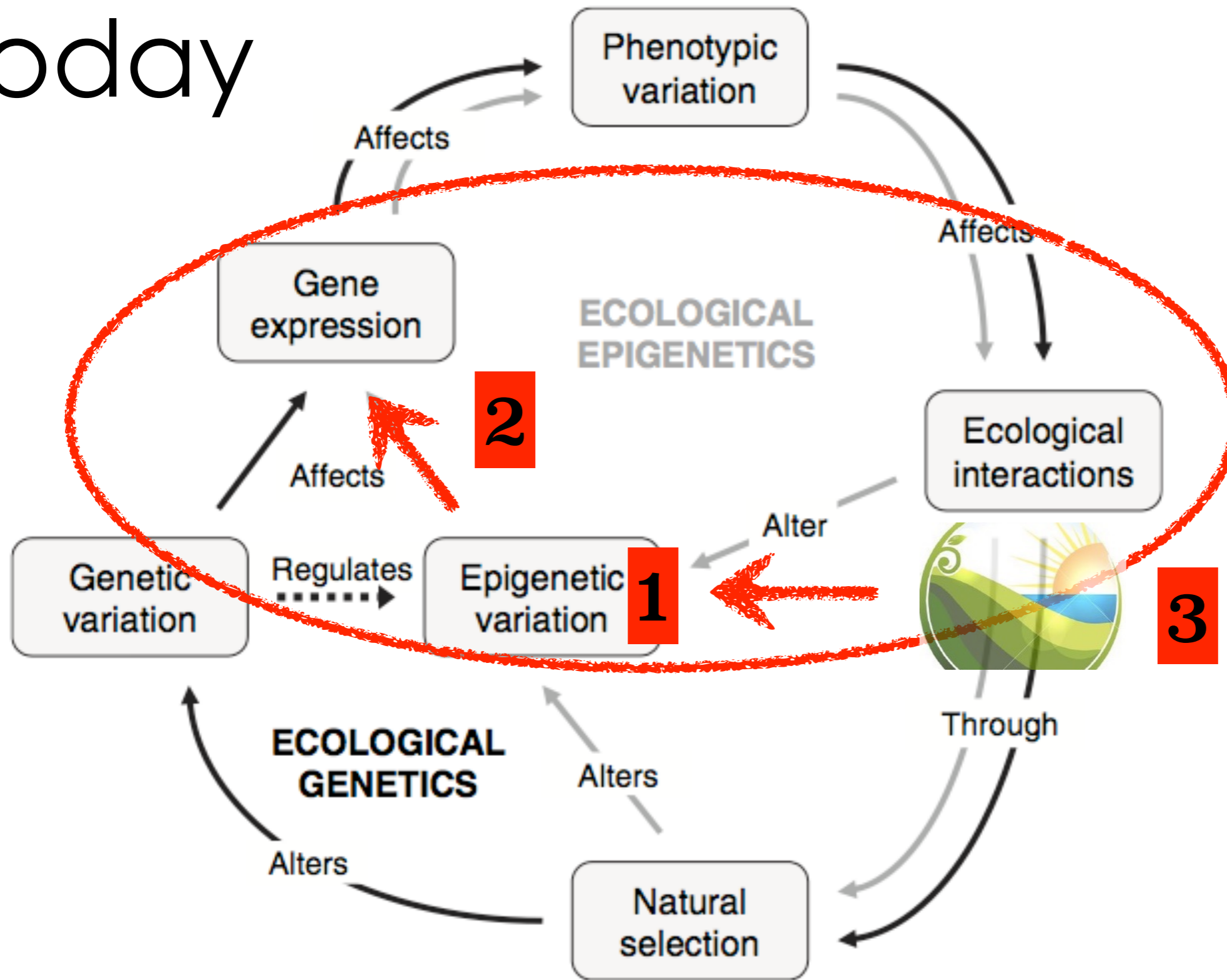
modified from

**IDEA AND
PERSPECTIVE**

Epigenetics for ecologists

Oliver Bosdorf,^{1*} Christina L. Richards² and Massimo Pigliucci³

Today



Ecology Letters, (2008) 11: 106–115

doi: 10.1111/j.1461-0248.2007.01130.x

modified from

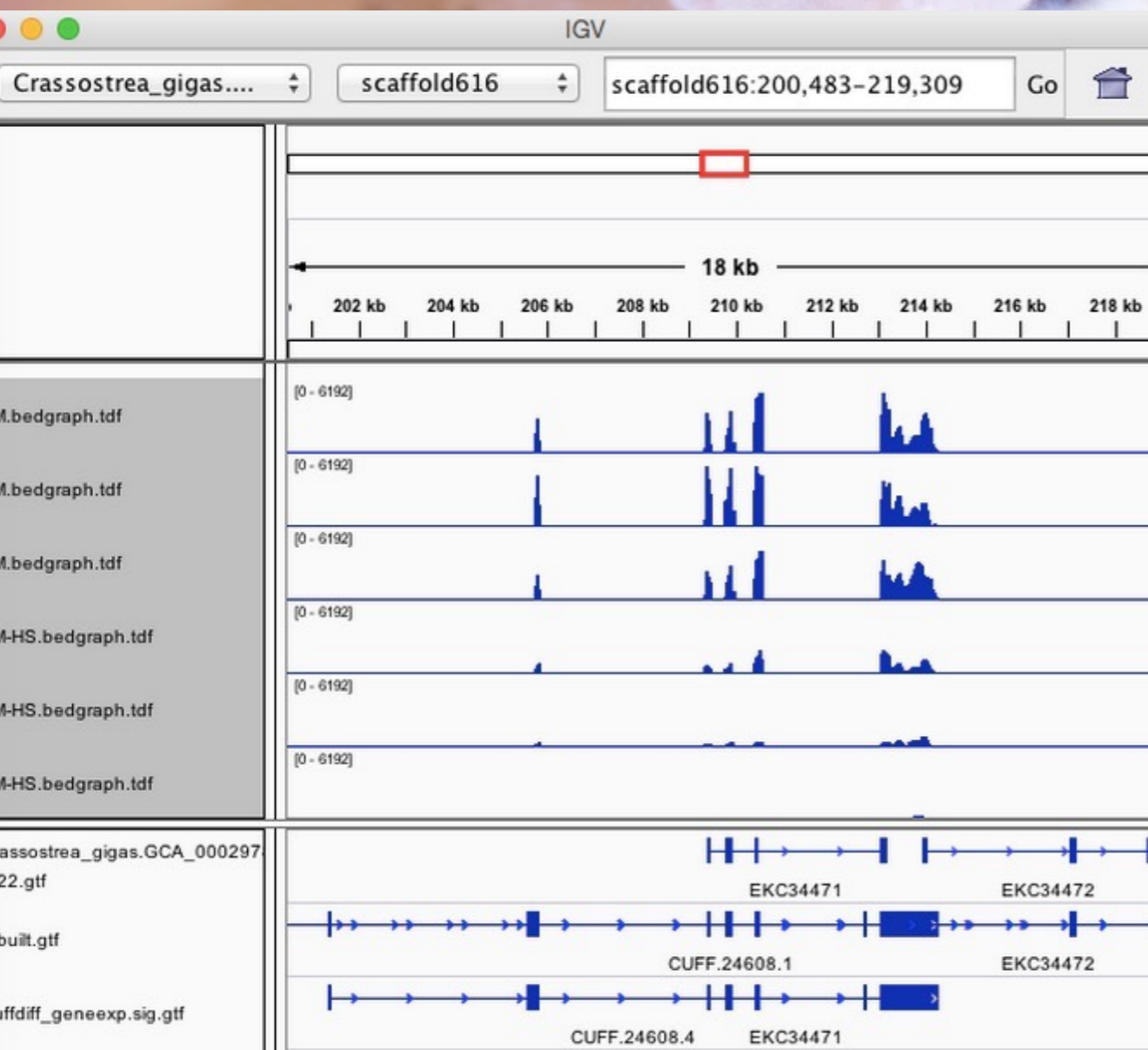
**IDEA AND
PERSPECTIVE**

Epigenetics for ecologists

Oliver Bosdorf,^{1*} Christina L. Richards² and Massimo Pigliucci³



Genome Resources



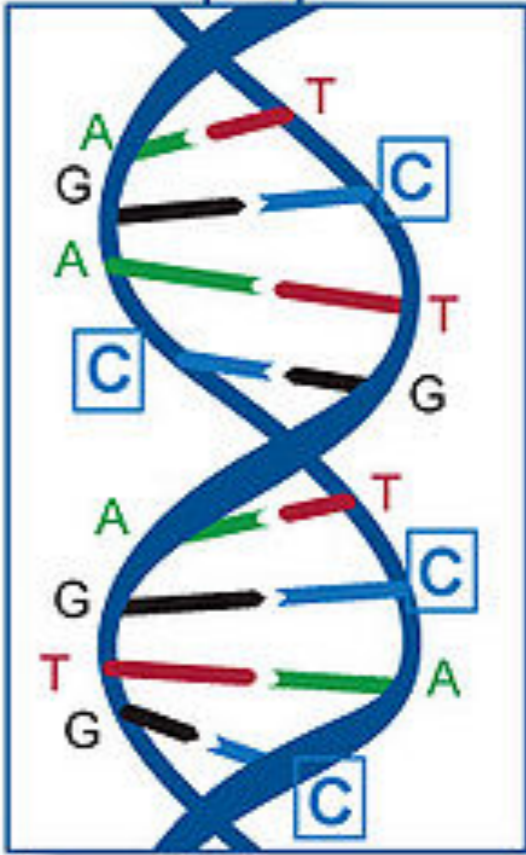
Epigenetics



Histone Modification

short RNAs

DNA Methylation

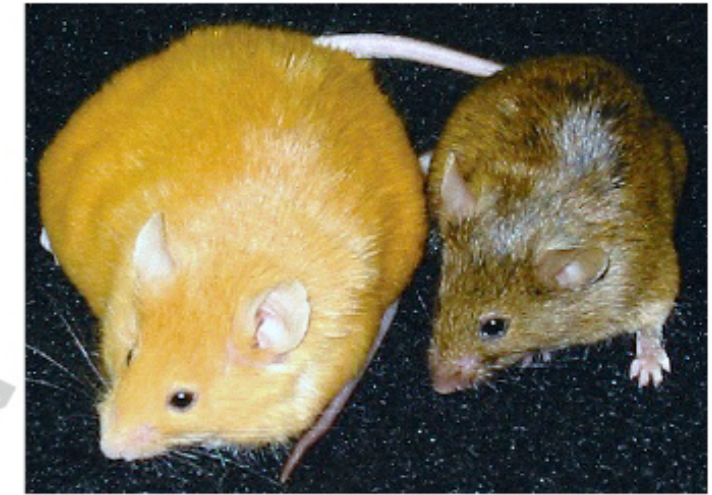




Epigenetics



These Two Mice are Genetically Identical and the Same Age



While pregnant, both of their mothers were fed Bisphenol A (BPA) but DIFFERENT DIETS:

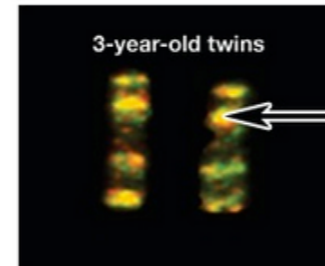
The mother of this mouse received a **normal mouse diet**

The mother of this mouse received a diet **supplemented** with choline, folic acid, betaine and vitamin B12

Nature AND Nurture

Chromosome 3 Pairs

3-year old twins vs. 50-year-old twins



3-year-old twins

Yellow shows where the twins have epigenetic tags in the same place.

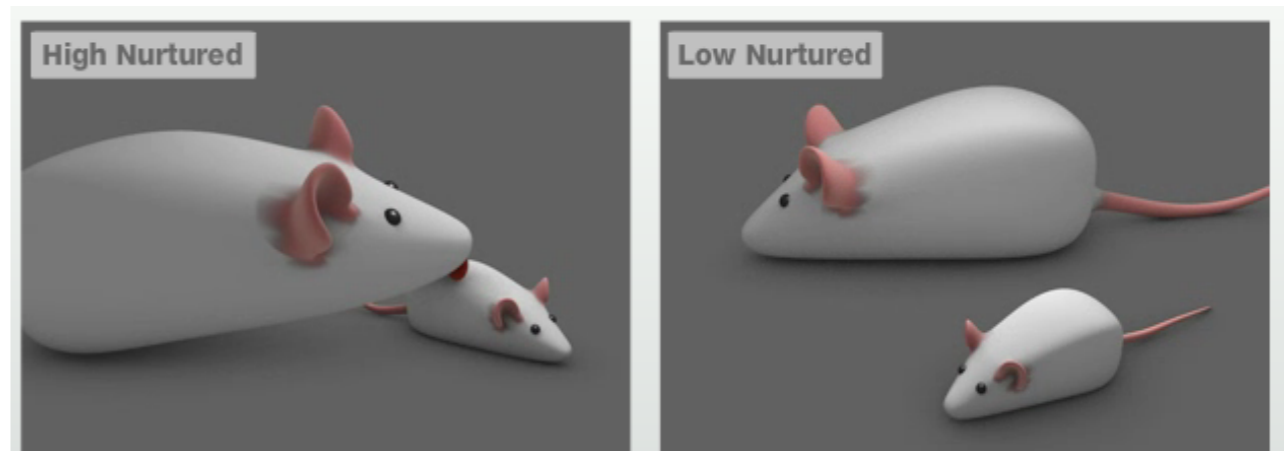
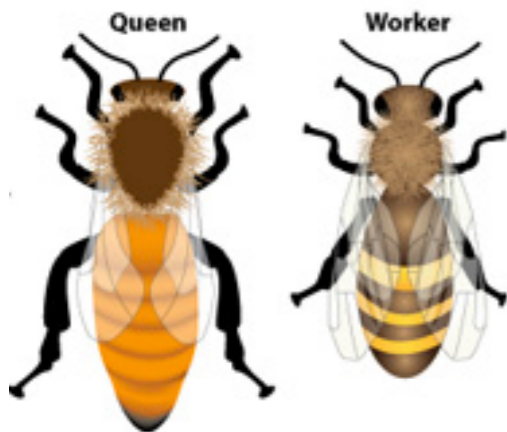


50-year-old twins

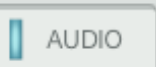
Red and green show where the twins have epigenetic tags in different places.

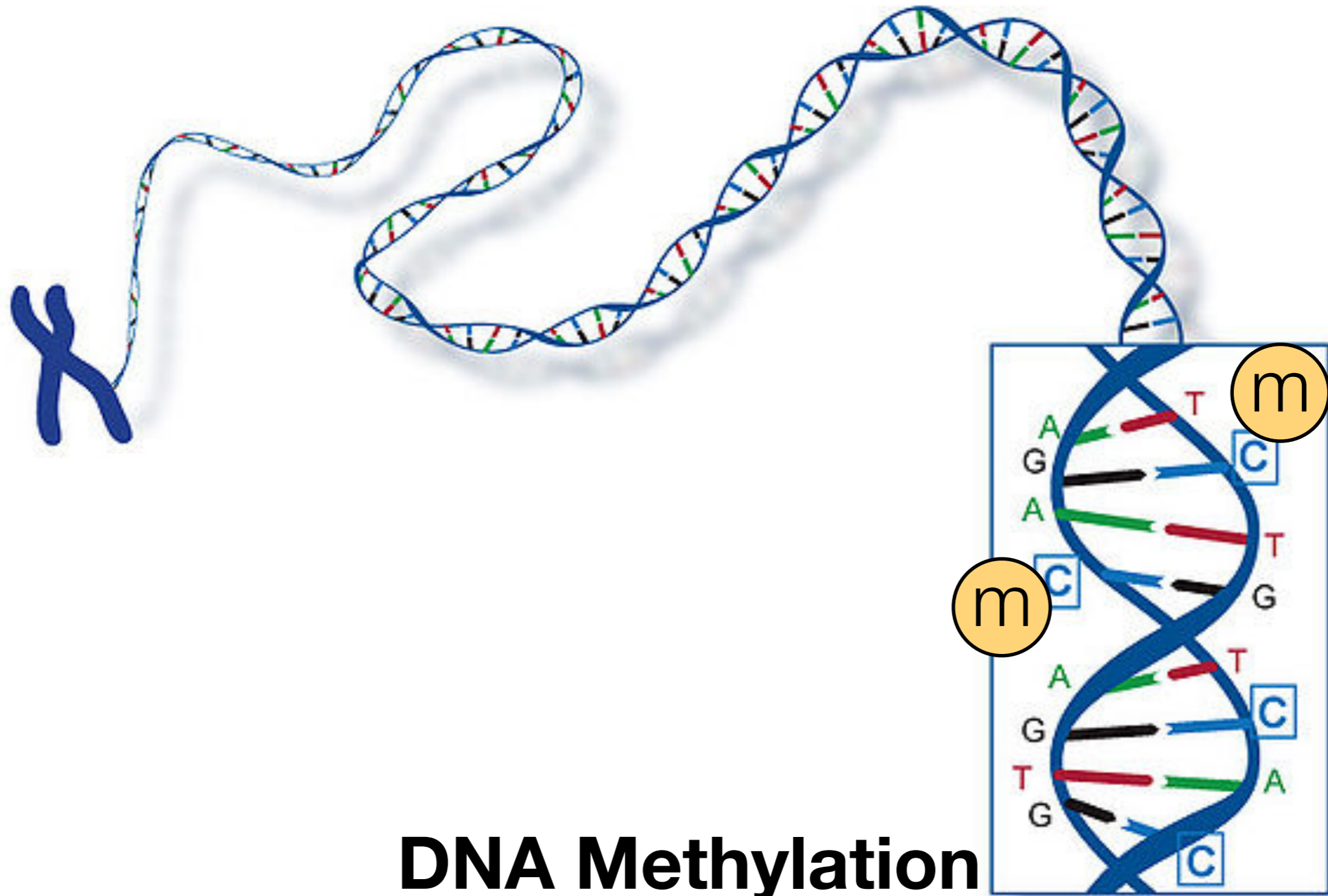


Queen Bee Larvae: Queens are raised in specially constructed cells called "queen cups," which are filled with royal jelly.



These mothers come from a long line of inbred rats, so their genomes are highly similar. But they care for their pups very differently.





DNA Methylation

The background features a complex genomic visualization. At the top, a horizontal axis is marked with values from 1,000 to 10,000. Below this, several tracks are visible: a track with orange and green bars, a track with a red bar and a green bar, a track with orange arrows, a track with purple and green arrows, and a track with blue and white text. A vertical line is drawn at the 3,754 position. Two semi-transparent images of animals are overlaid: a seal at the top and a lizard at the bottom. The text 'Non-Vertebrates?' is centered over the image.

Non-Vertebrates?

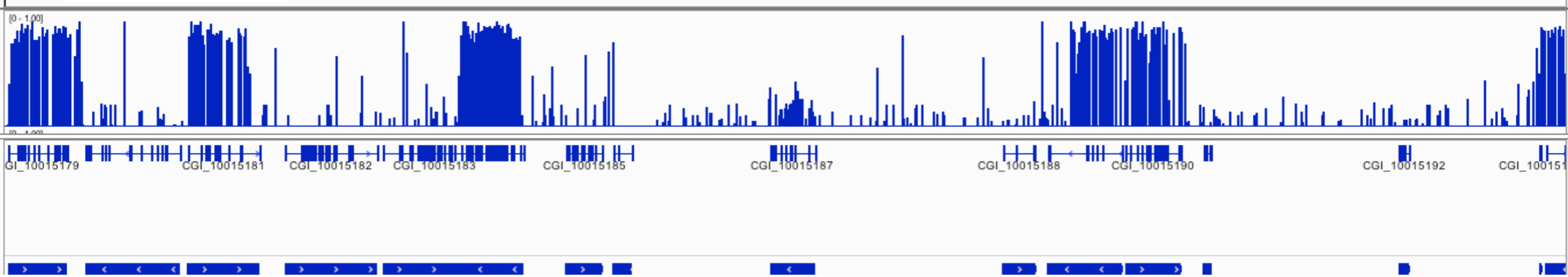
landscape and function is very different than what is observed in vertebrates

Absent in
several
model
organisms

Oysters?



Epigenetic variation **1**

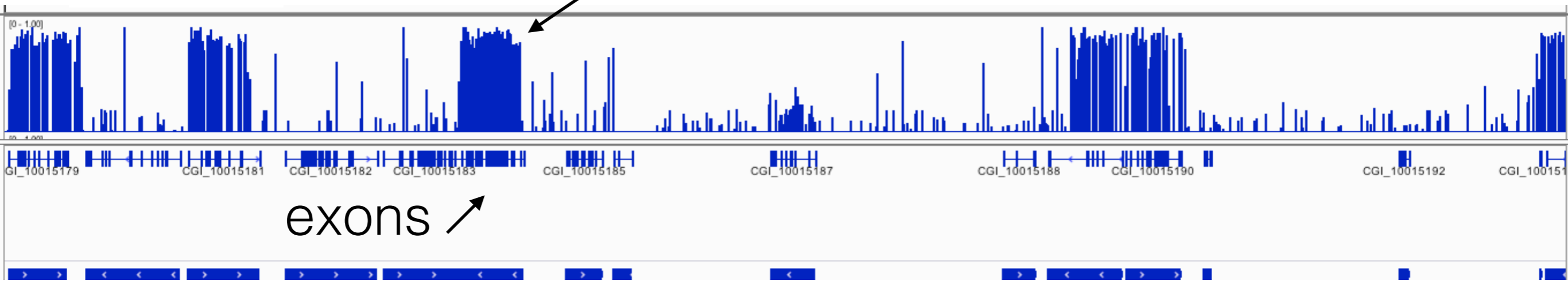


mosaic

associated with gene bodies

Epigenetic variation **1**

DNA methylation level (0-100%) @ cytosines



exons ↗

gene ↗

mosaic

associated with gene bodies

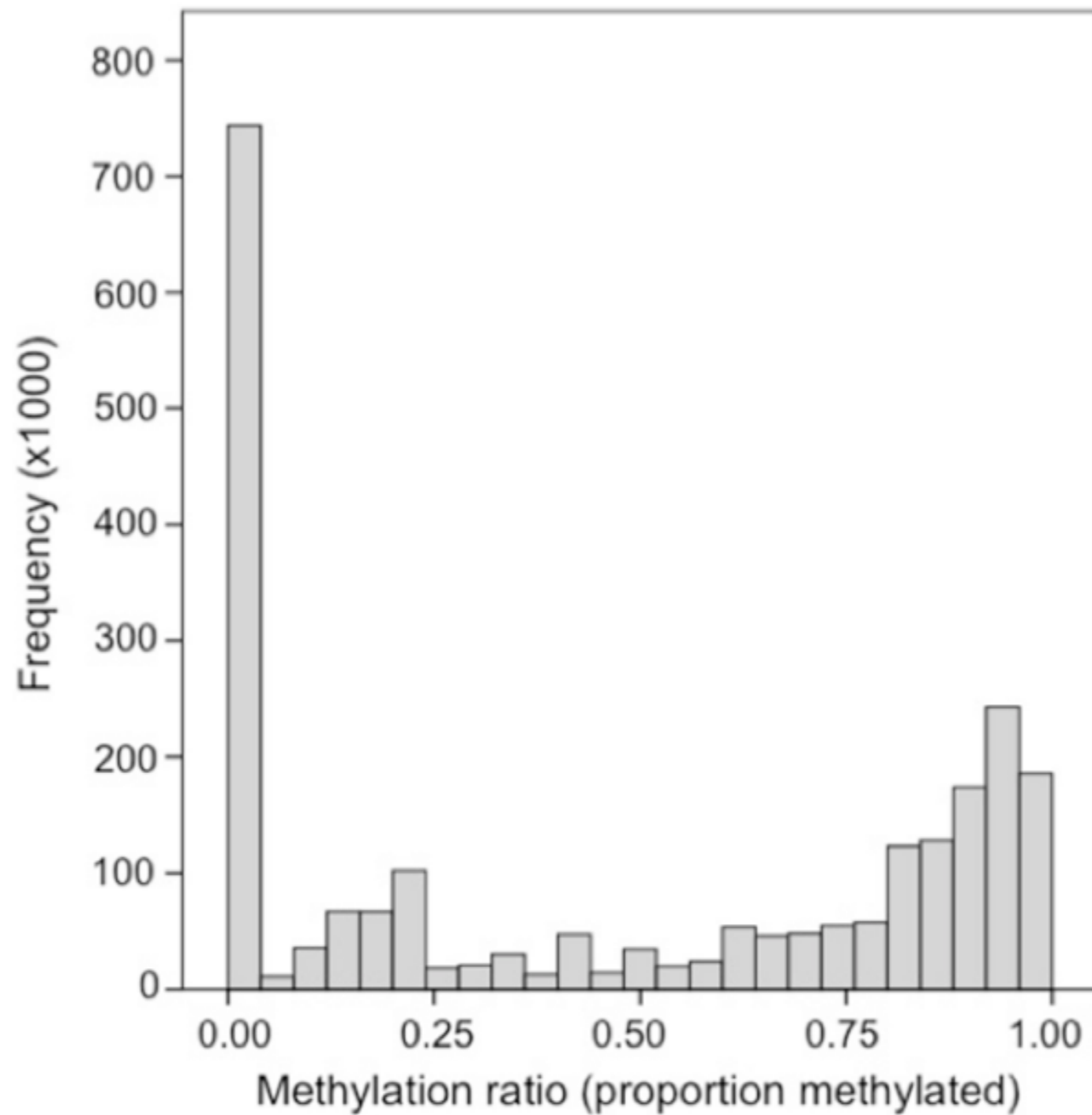
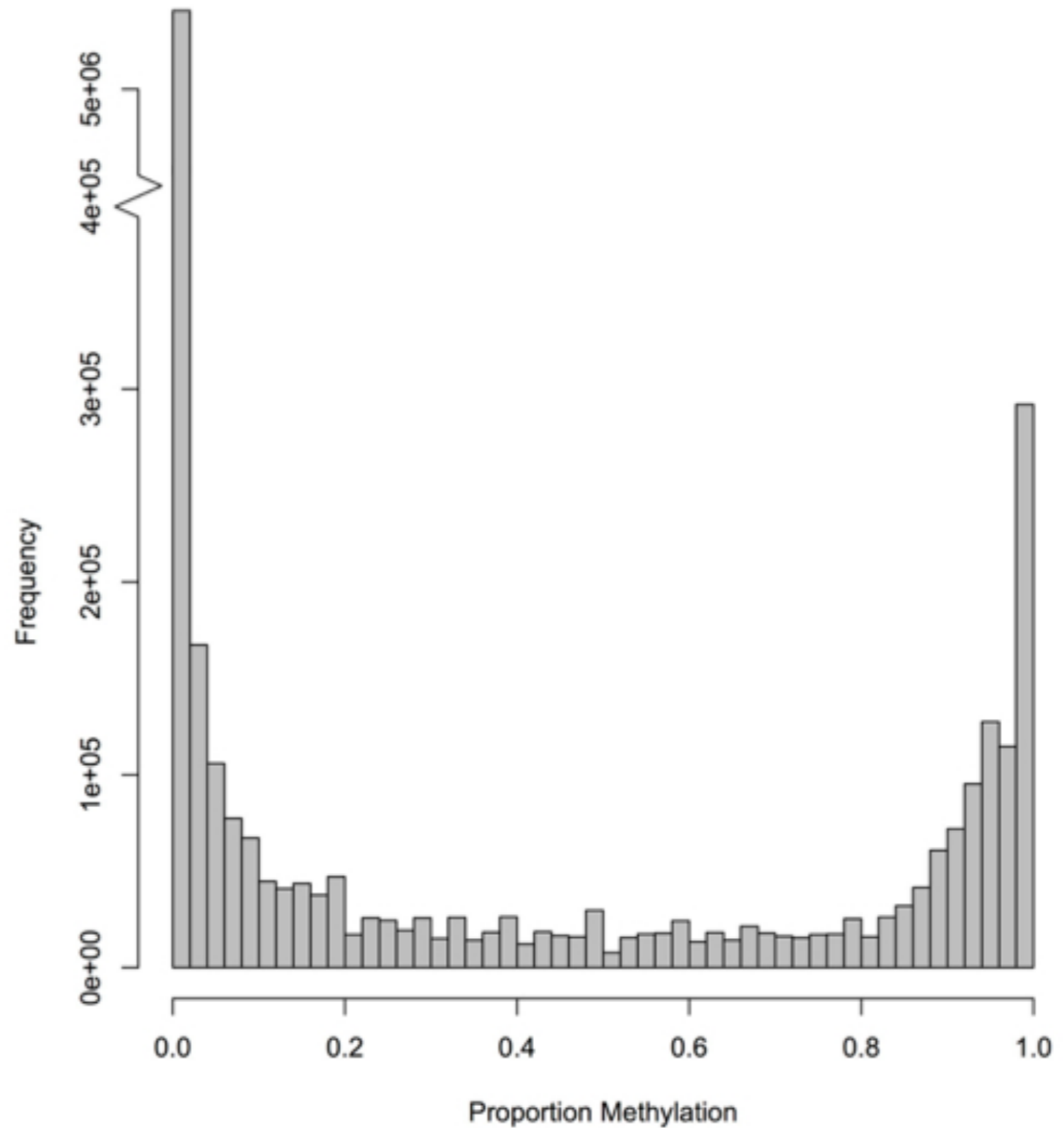


Figure 1 Frequency distribution of methylation ratios for CpG dinucleotides in oyster gill tissue. A total of 2,625,745 CpG dinucleotides with $\geq 5\times$ coverage are represented.



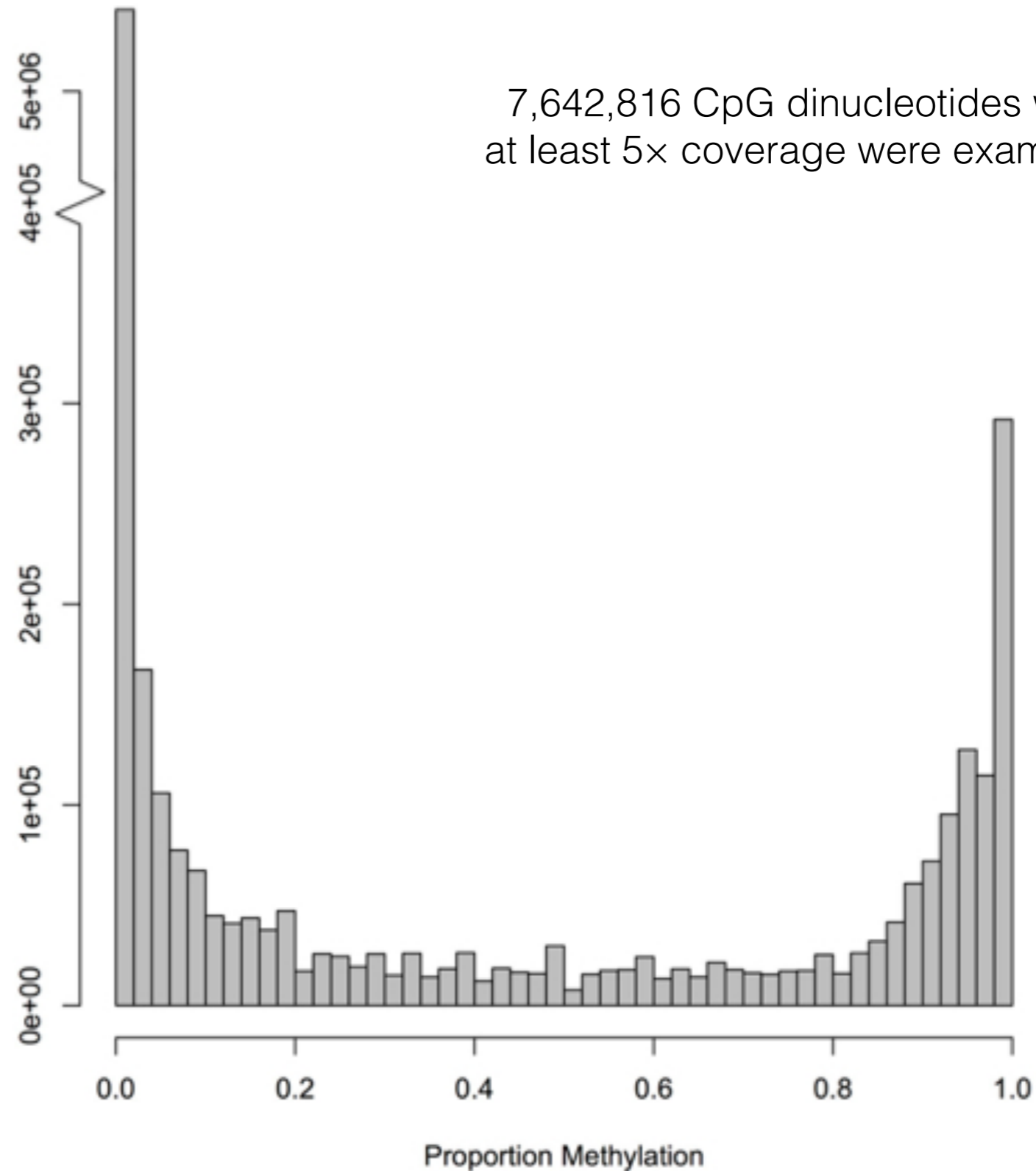
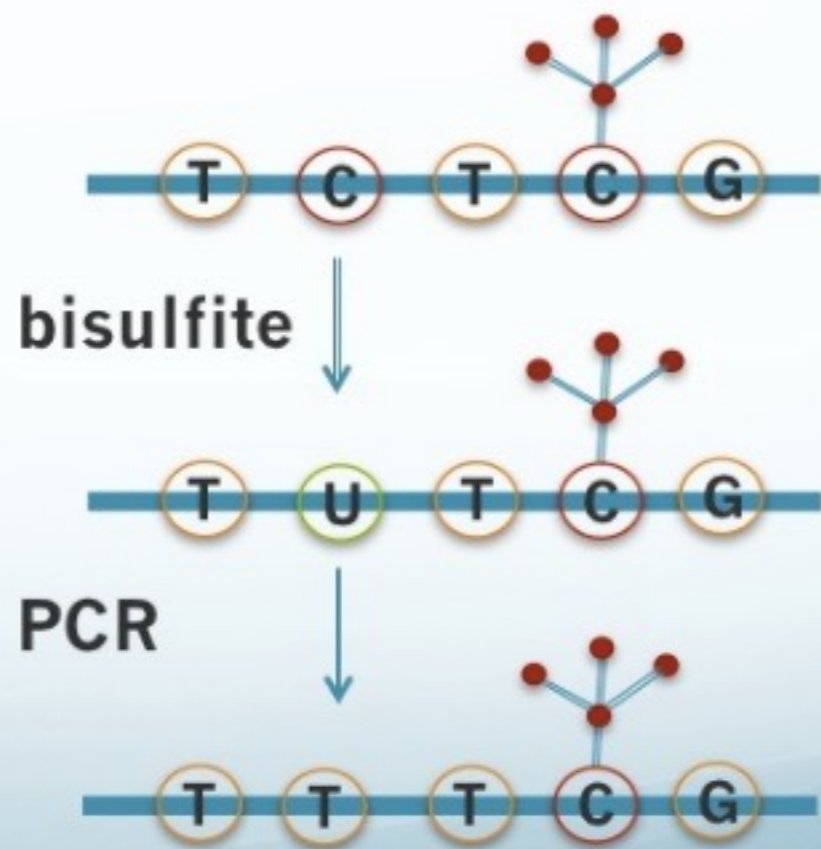
Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes

Claire E. Olson and Steven B. Roberts*

School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA

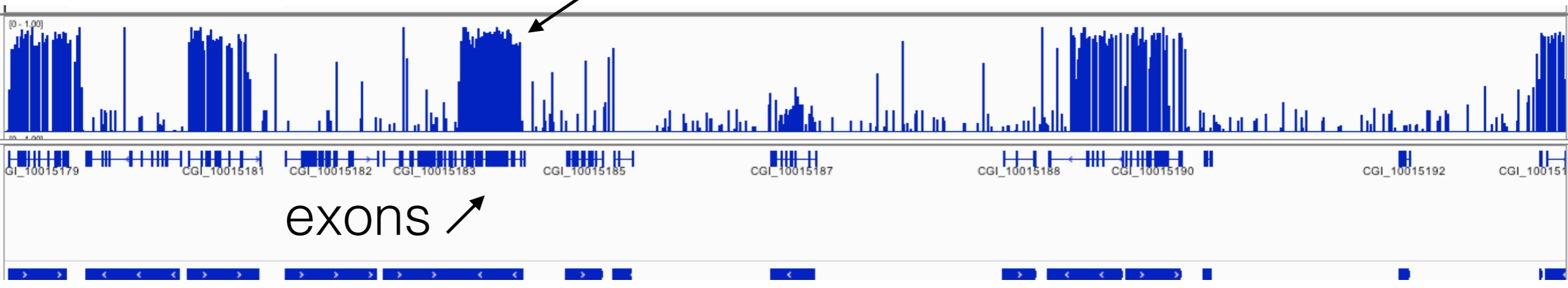
Epigenetic variation **1**

- Bisulfite conversion



Epigenetic variation **1**

DNA methylation level (0-100%) @ cytosines



exons ↗

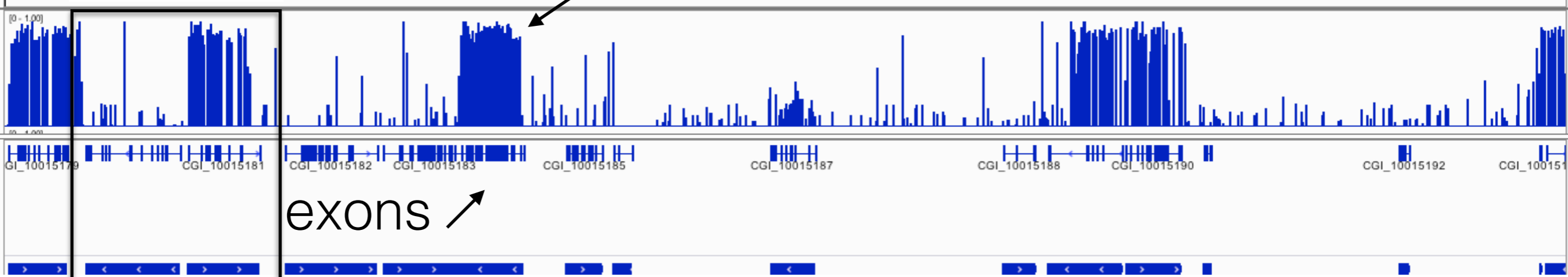
gene ↗

mosaic

associated with gene bodies

DNA methylation level (0-100%) @ cytosines

Epigenetic variation **1**



gene

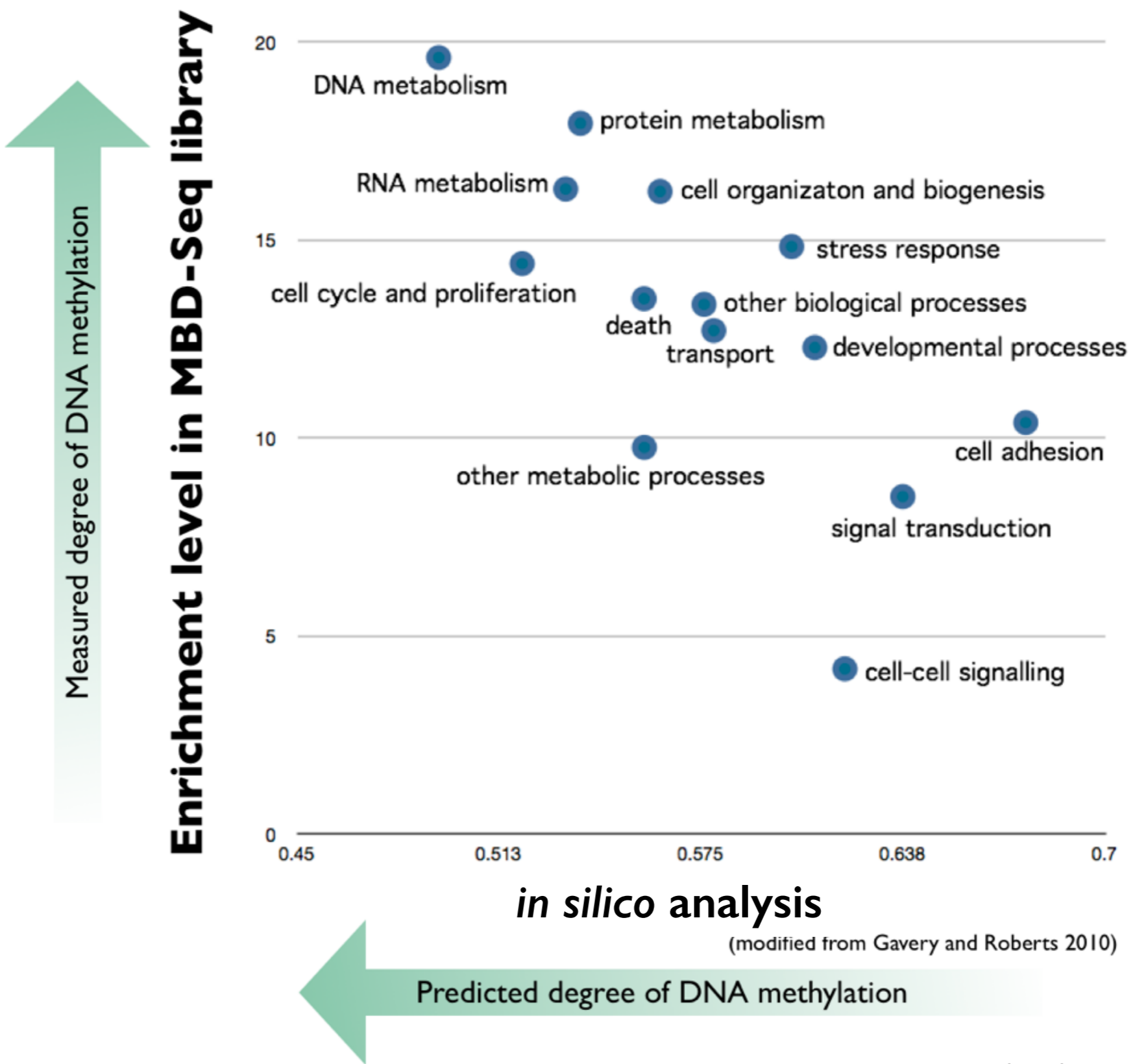
exons

mosaic

Why are only a subset of genes methylated?

associated with gene bodies

Epigenetic variation **1**



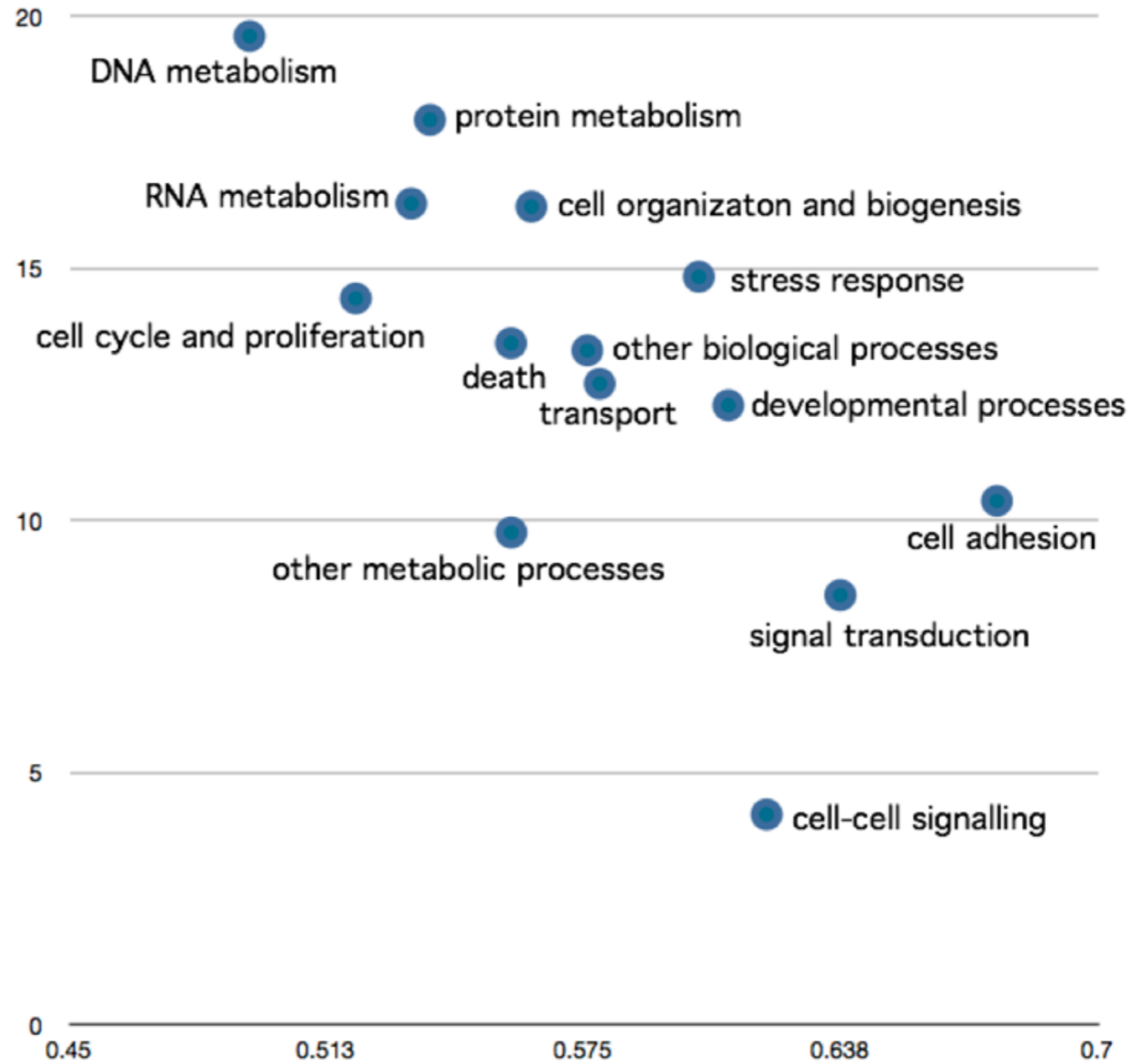
in silico analysis

(modified from Gavery and Roberts 2010)

Epigenetic variation **1**

Measured degree of DNA methylation

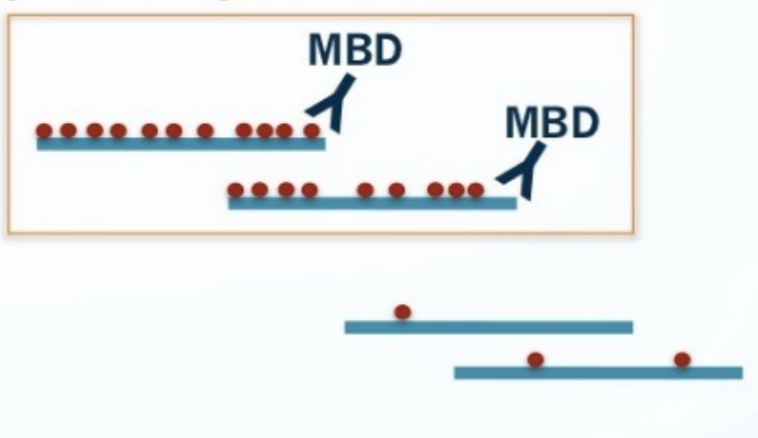
Enrichment level in MBD-Seq library



in silico analysis

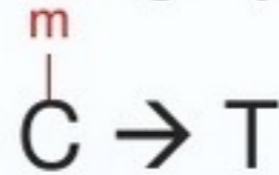
(modified from Gavery and Roberts 2010)

Predicted degree of DNA methylation



in silico approach

- Principle:
 - Methylated cytosines are highly mutable



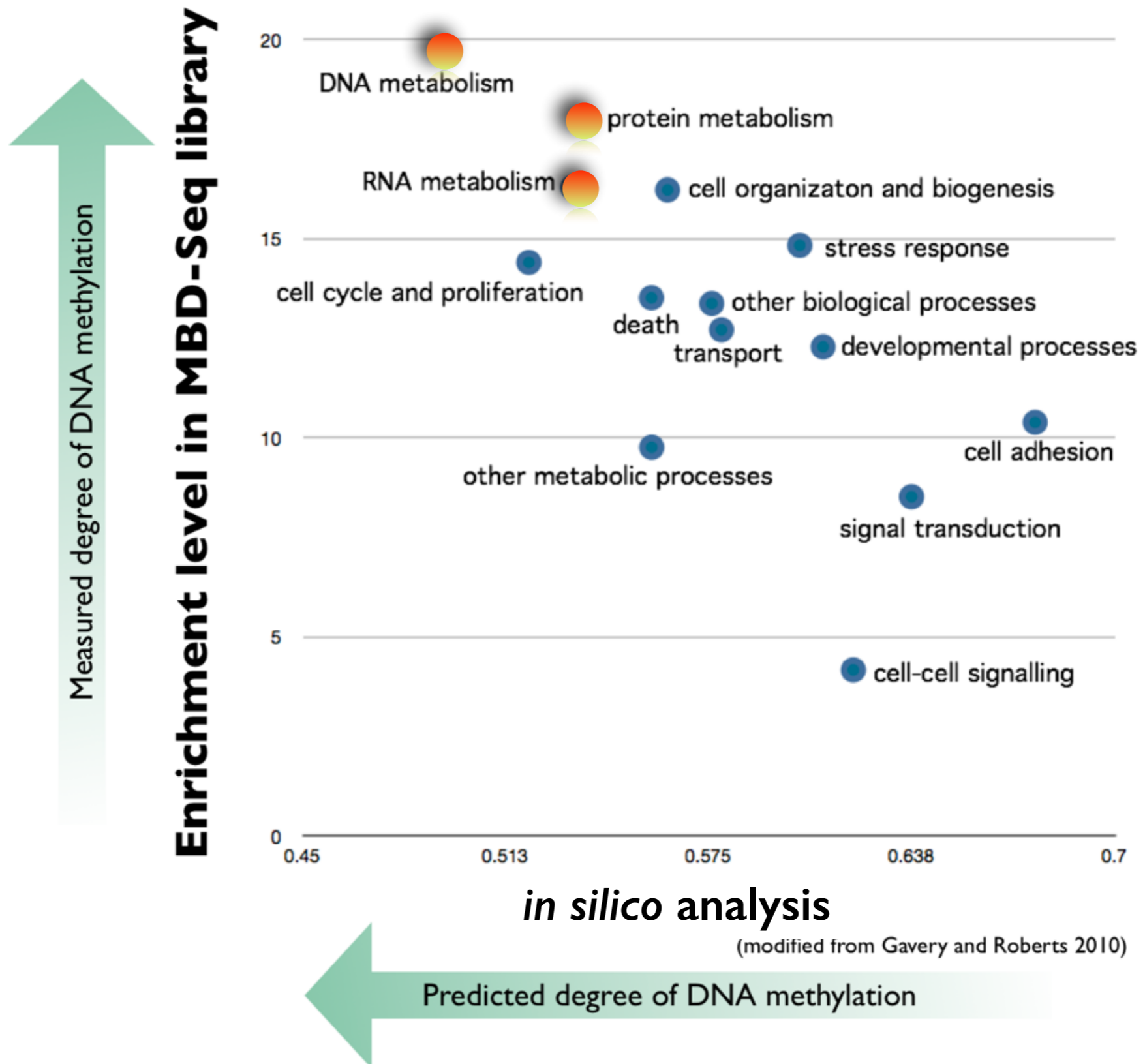
- Methylated regions of DNA are depleted of CpG dinucleotides over evolutionary time (CpG to TpG)

$$\text{CpG O/E} = \frac{\text{CpG observed}}{\text{CpG expected}}$$

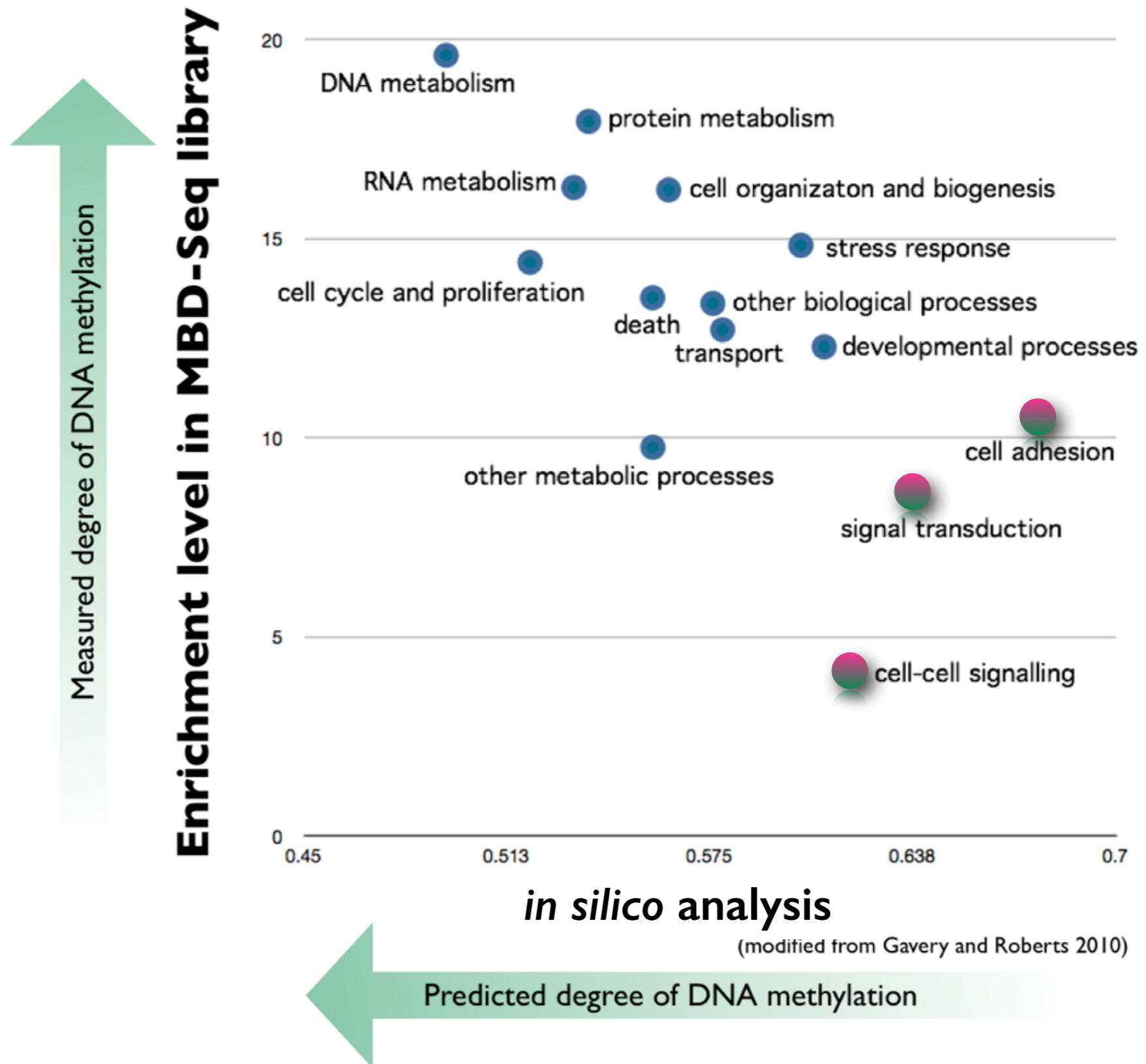


high = unmethylated

Epigenetic variation **1**



(modified from Gavery and Roberts 2010)



Jay Dimond

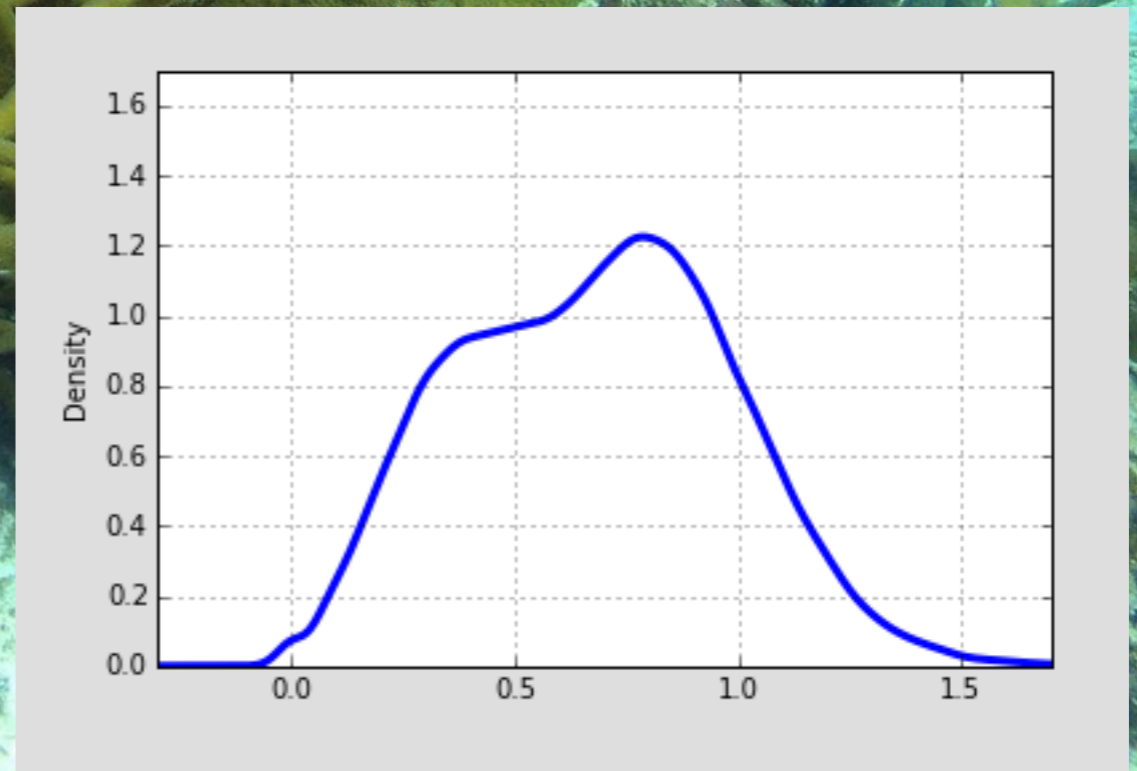
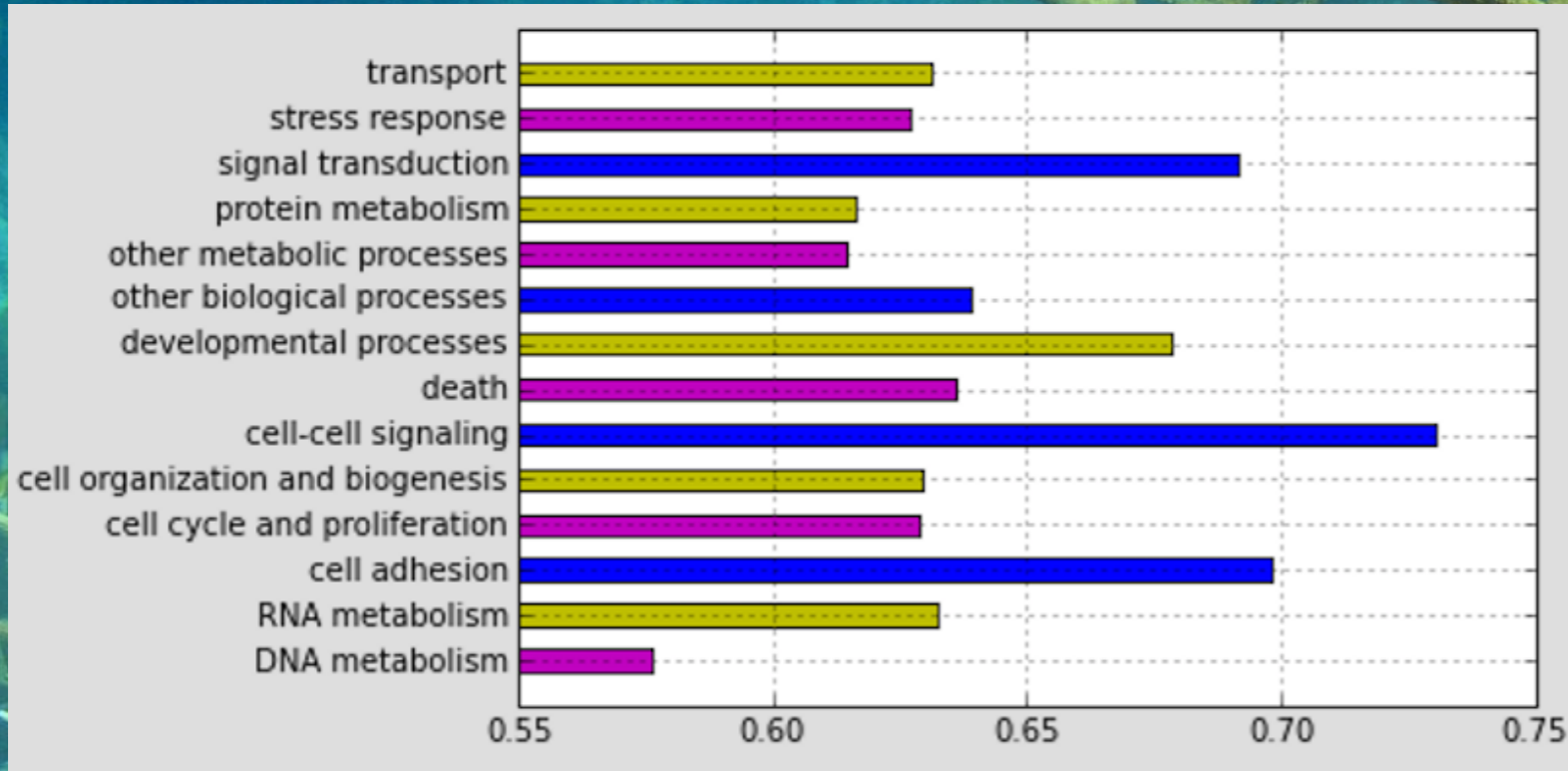
Acropora palmata

Acropora cervicornis



Jay Dimond

Acropora palmata

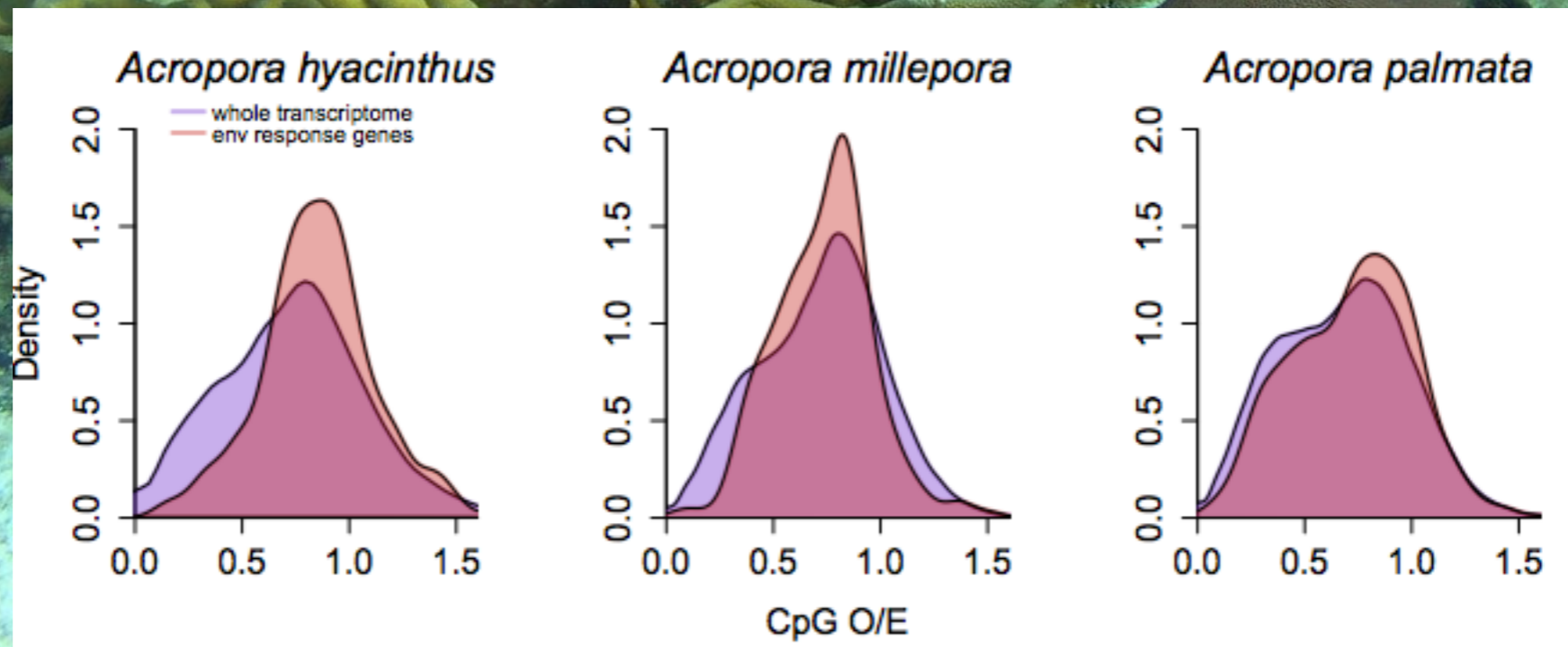


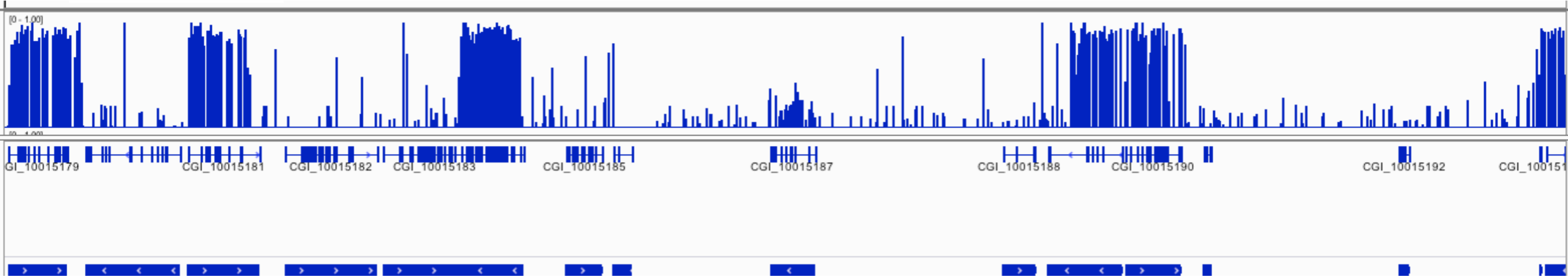
Jay Dimond



Acropora palmata

Germline DNA methylation in five coral transcriptomes



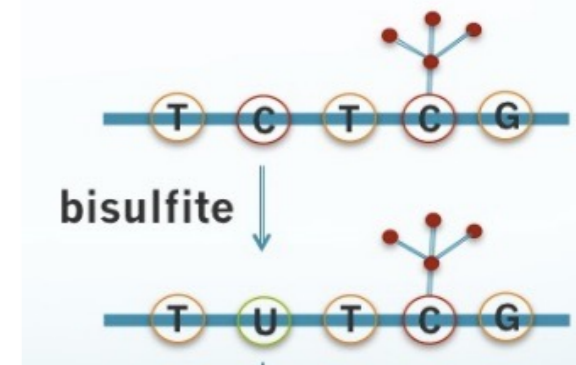


mosaic

associated with gene bodies
based on gene function

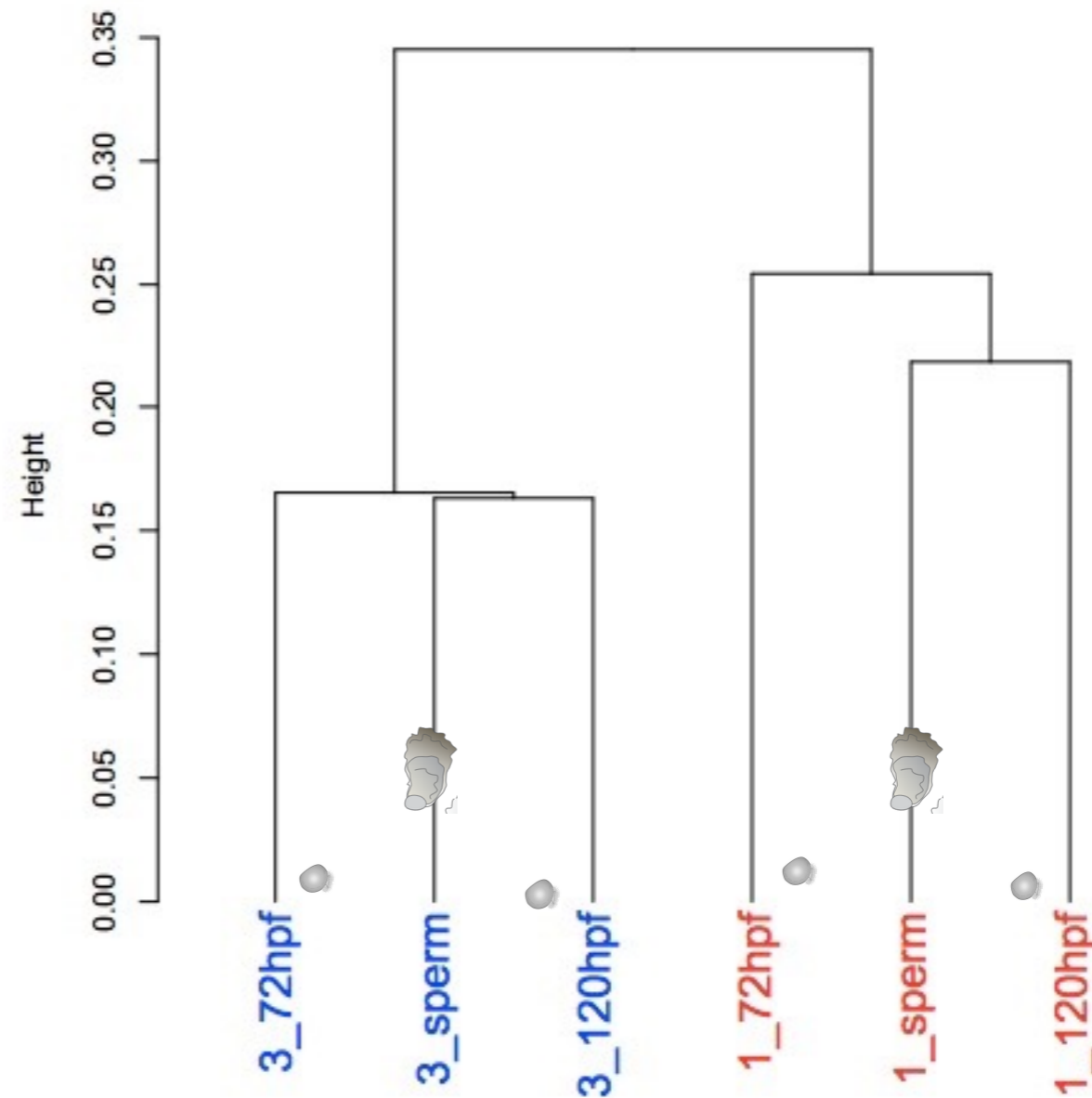
explanation?

Family and Developmental Variation

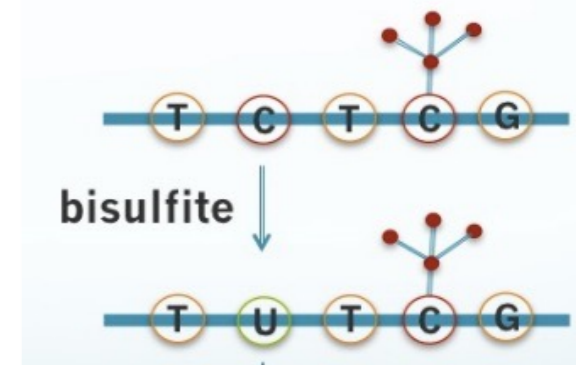


Sperm & Larvae (72h & 120h)

CpG methylation clustering

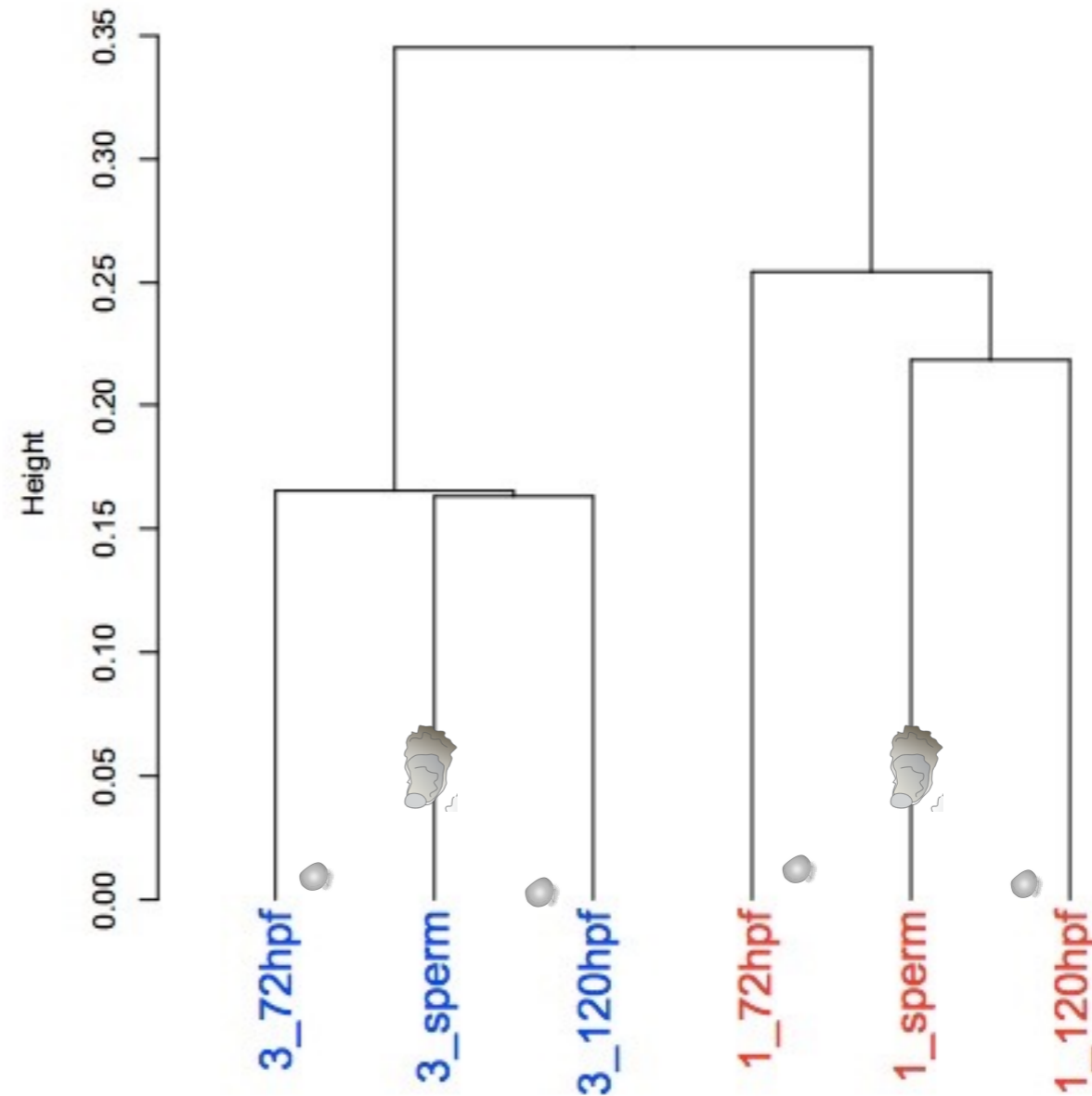


Family and Developmental Variation



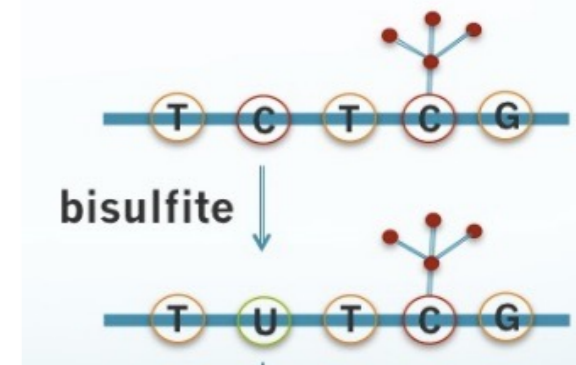
Inheritance

CpG methylation clustering



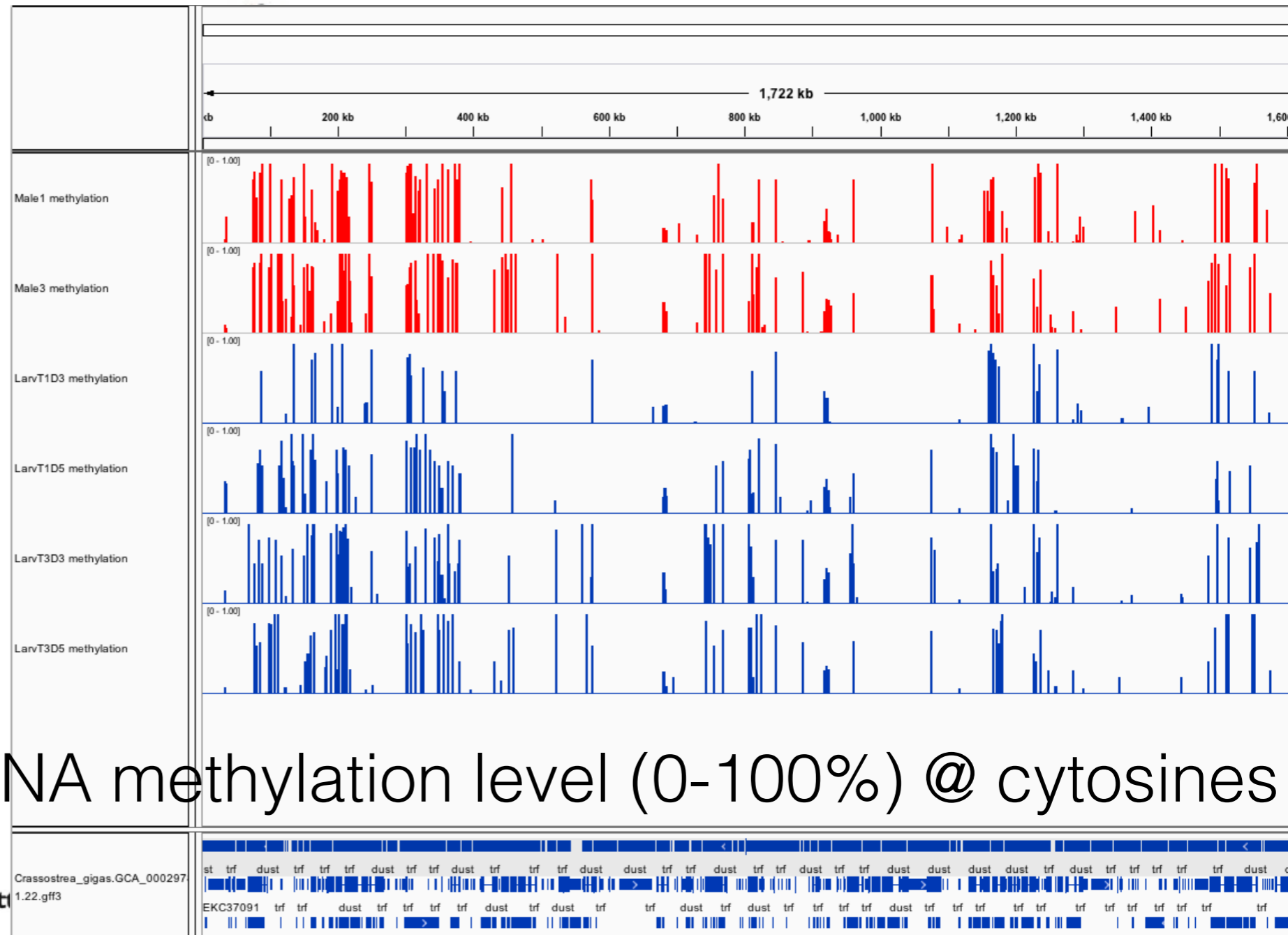
Epigenetic variation **1**

Family and Developmental Variation



CpG methylation clustering

Sperm & Larvae (72h & 120h)



DNA methylation level (0-100%) @ cytosines



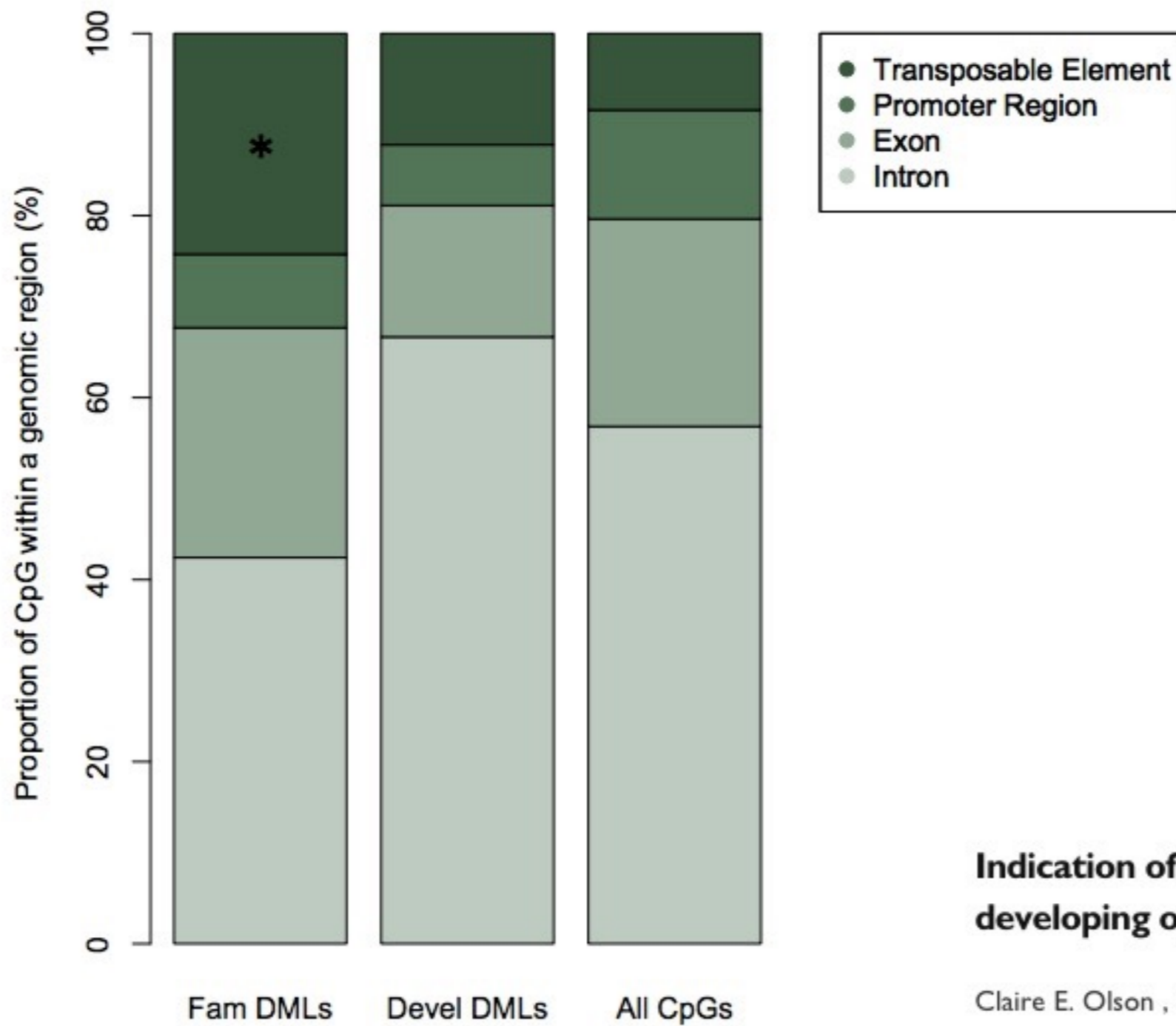
bioRxiv beta
THE PREPRINT SERVER FOR BIOLOGY

New Results

Indication of family-specific DNA methylation patterns in developing oysters

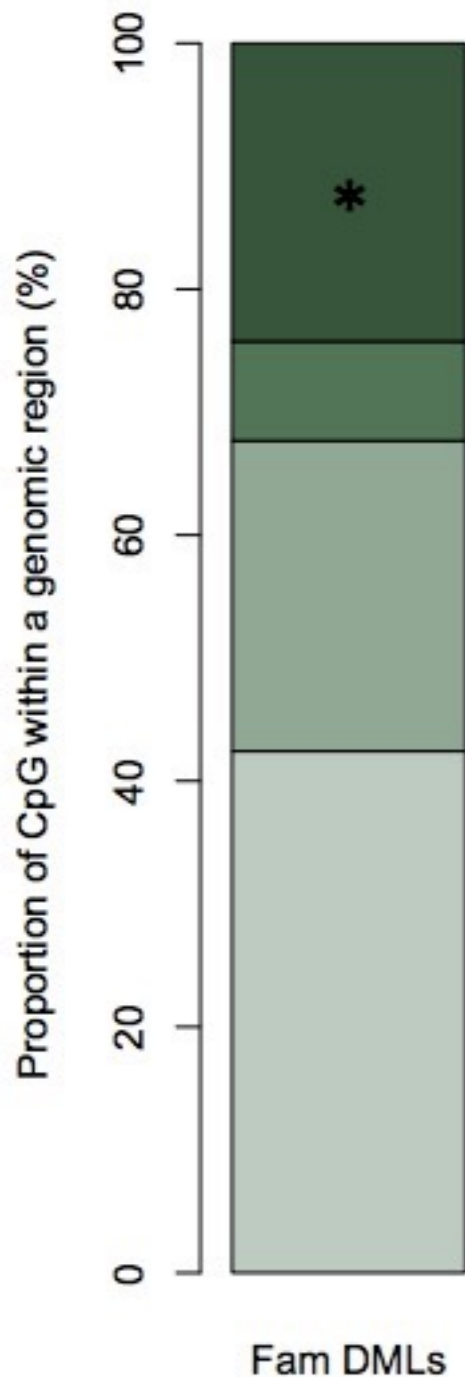
Claire E. Olson, Steven B. Roberts

doi: <http://dx.doi.org/10.1101/012831>



Indication of family-specific DNA methylation patterns in developing oysters

Claire E. Olson , Steven B. Roberts
doi: <http://dx.doi.org/10.1101/012831>



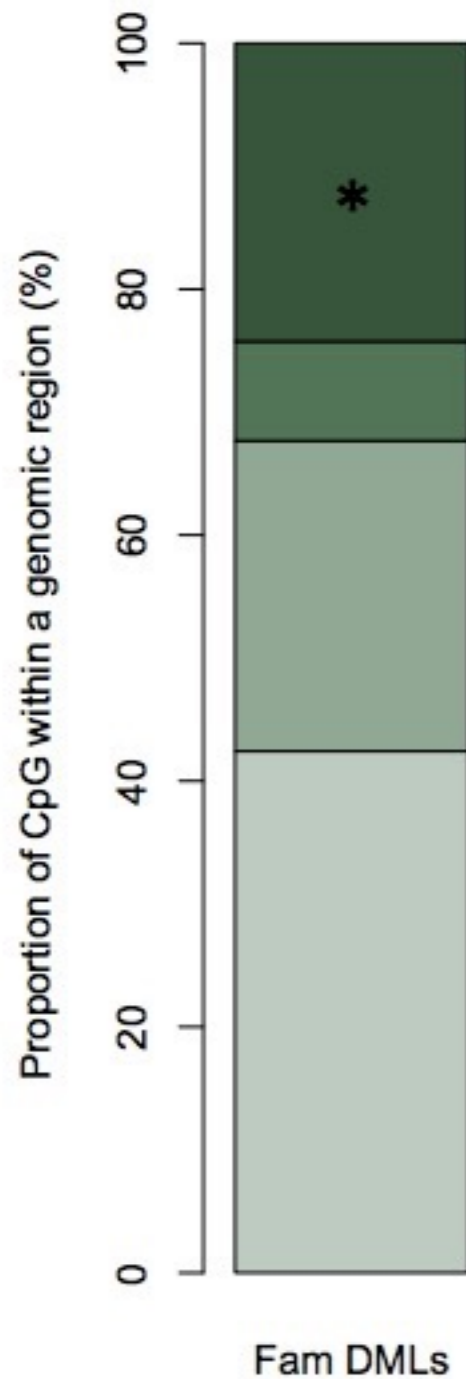
Differentially methylated loci in gene bodies could be lethal or deleterious, as they would alter gene expression.

Another possibility is that differentially methylated loci may provide advantageous phenotypic variation by increasing transposable element mobility.

Indication of family-specific DNA methylation patterns in developing oysters

Claire E. Olson , Steven B. Roberts

doi: <http://dx.doi.org/10.1101/012831>



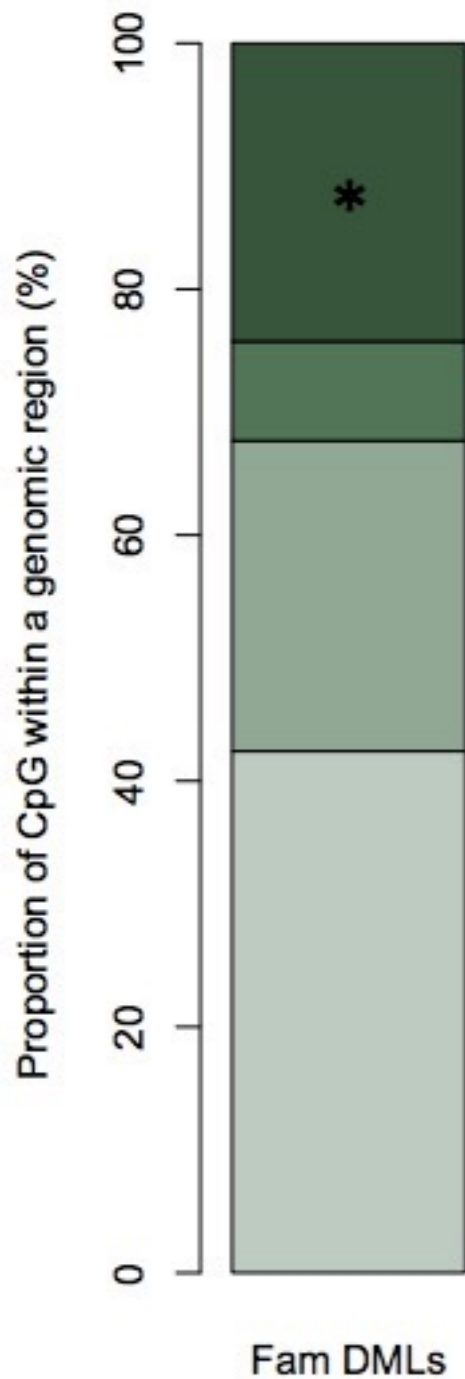
***Assumes some form
of random process
occurring during
gametogenesis?***

**Indication of family-specific DNA methylation patterns in
developing oysters**

Claire E. Olson , Steven B. Roberts

doi: <http://dx.doi.org/10.1101/012831>

Epigenetic variation **1**



This repository Search



che625 / olson-ms-nb



IPython Notebook for downloading and analyzing data the manuscript: "Indication of family-specific DNA methylation patterns in developing oysters"

bioRxiv preprint - <http://dx.doi.org/10.1101/012831>

To execute the IPython Notebook in its entirety you will need:

- IPython - [install instructions](#)
- BSMAP - [install instructions](#)
- bedtools - [install instructions](#)
- R - [install instructions](#)
- rpy2 (interface to R from Python) - [install instructions](#)

1 Summary

- Sparsely (~16 %), mosaic methylated genome
- Gene body methylation correlated with function
- DNA methylation patterns are inherited
- DMRs are predominant in transposable elements

Gene
expression



2

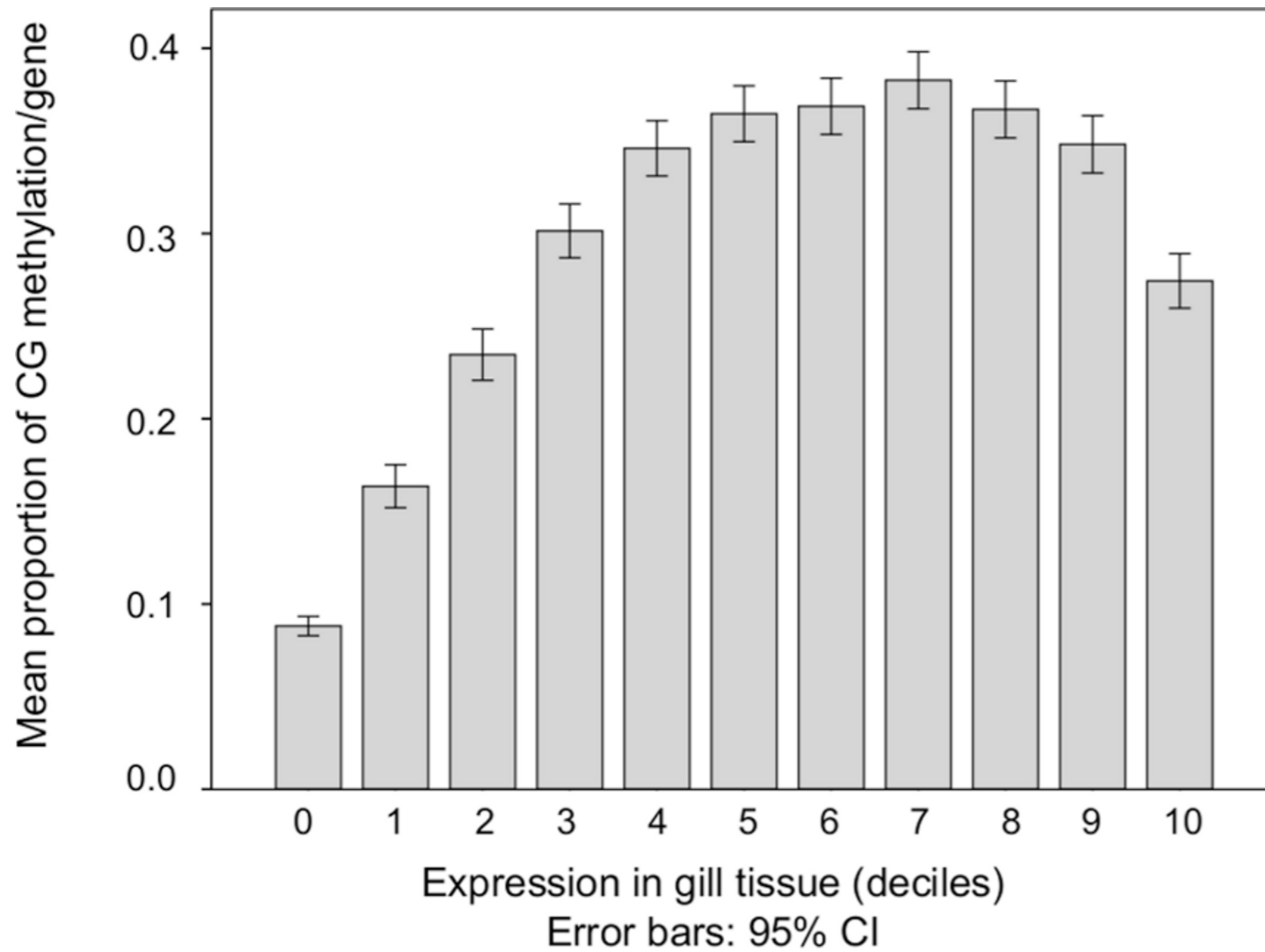
Epigenetic
variation

Function?

Gene expression

2

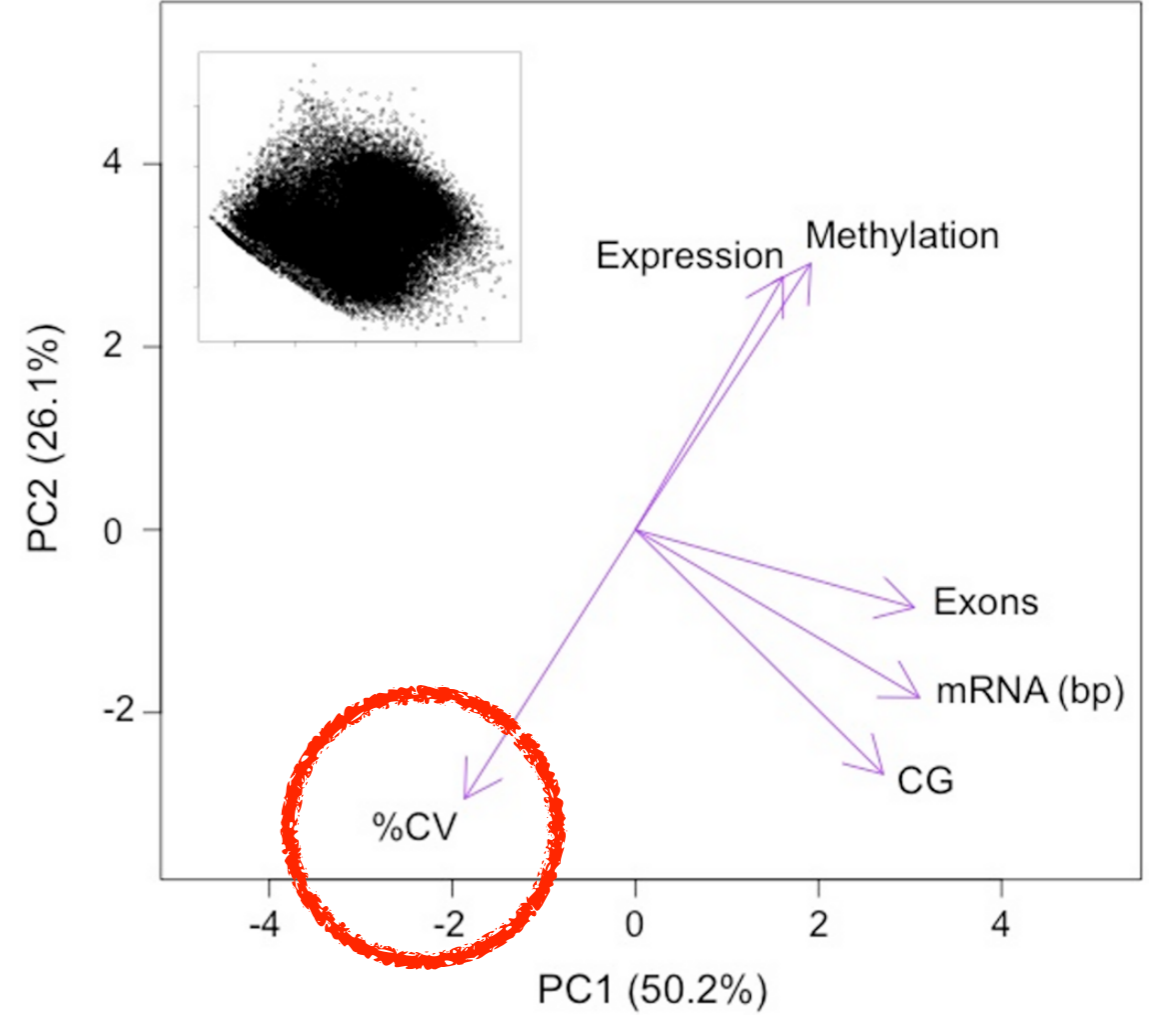
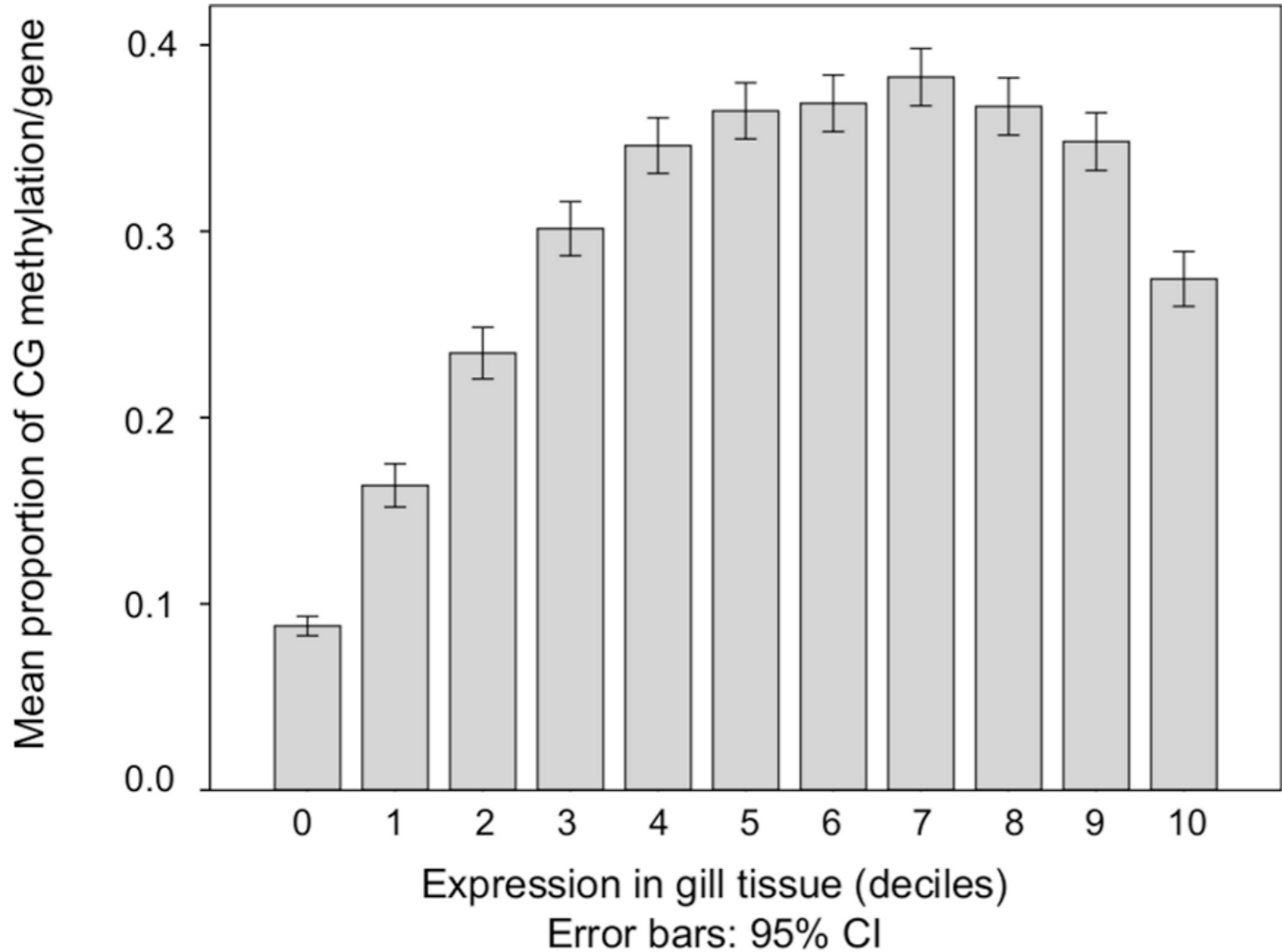
Epigenetic variation



Gene expression

2

Epigenetic variation



Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc

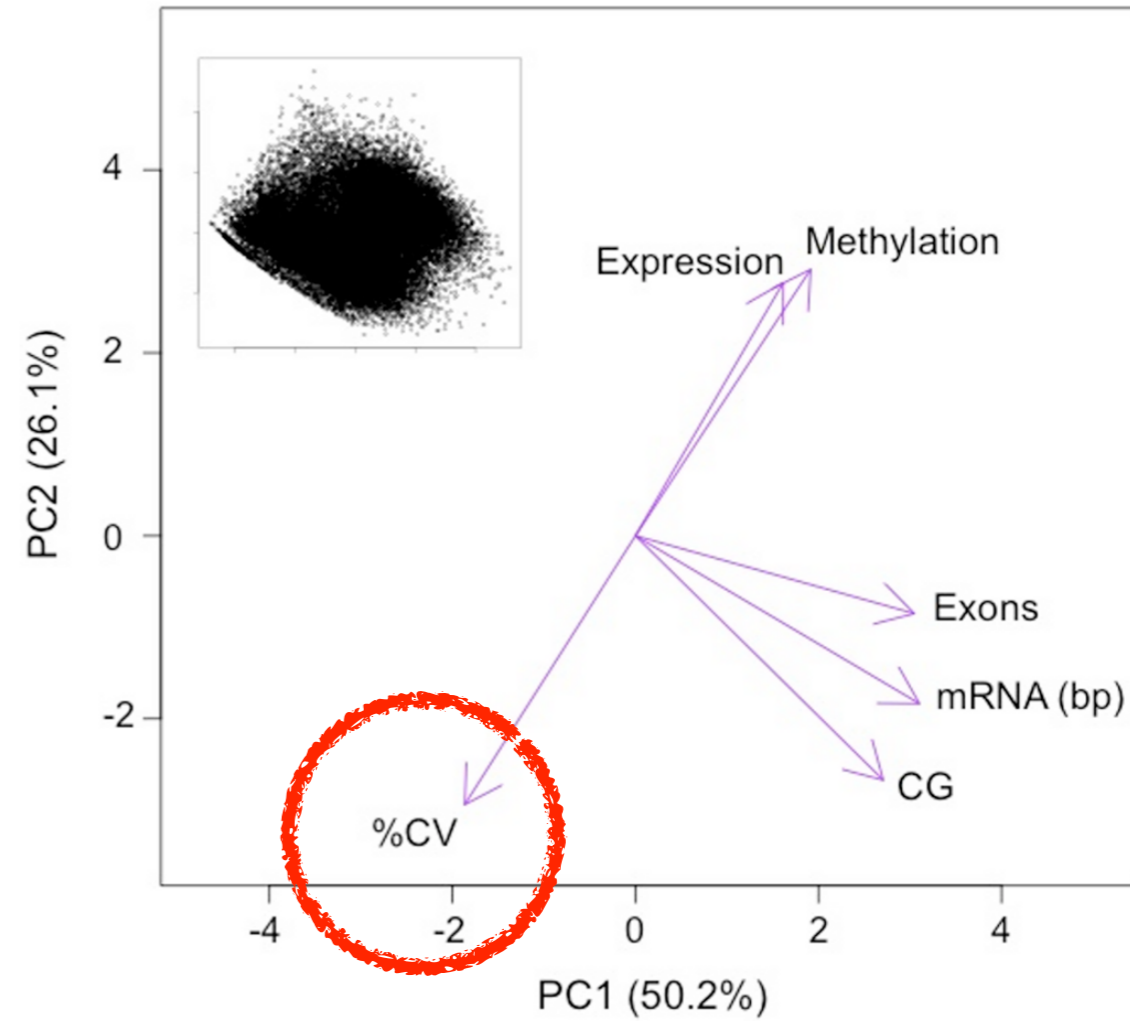
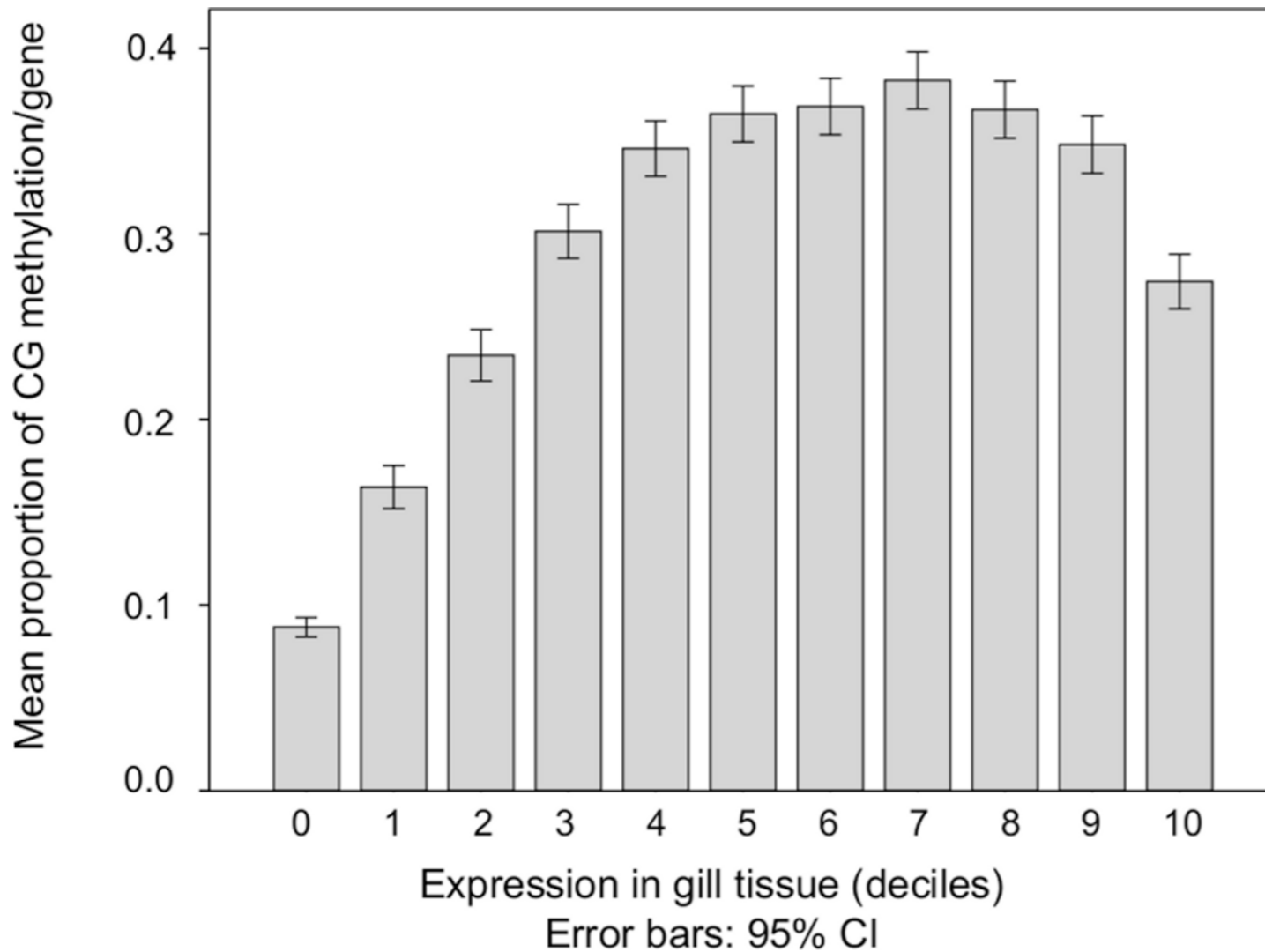
Mackenzie R. Gavary and Steven B. Roberts
School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA

Gene expression

2

Epigenetic variation

Theory: **Does not** influence expression level but rather alternative splicing.

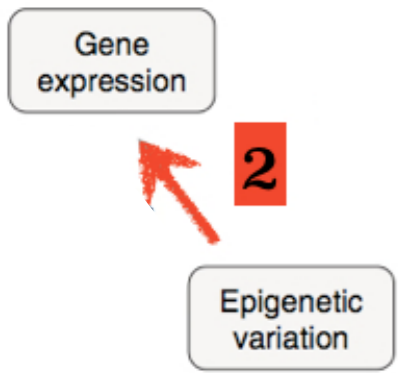


A context dependent role for DNA methylation in bivalves

Mackenzie R. Gavery and Steven B. Roberts

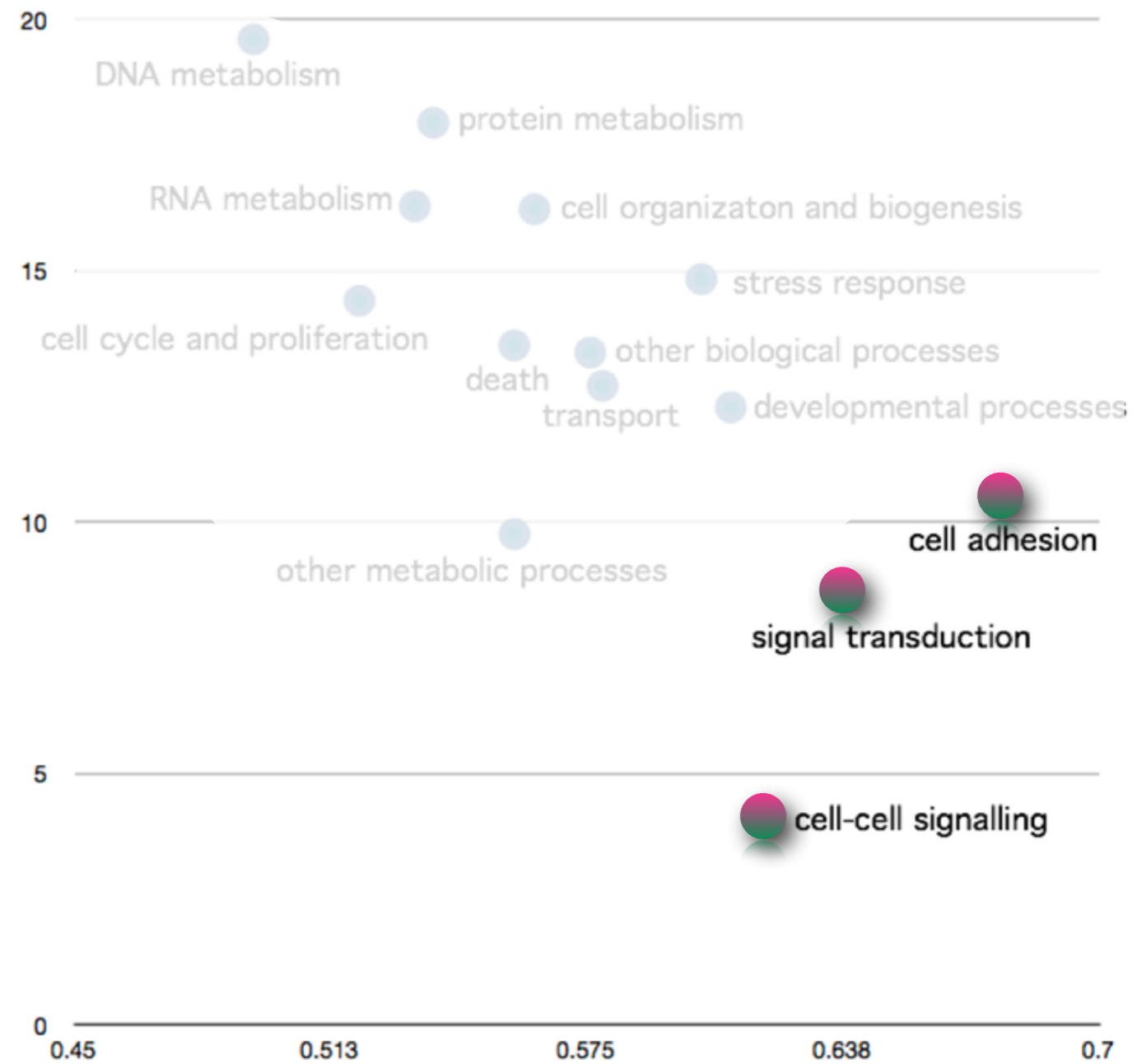
Advance Access publication date 7 January 2014

In species that experience a diverse range of environmental conditions, processes have evolved to increase the number of potential phenotypes in a population in order to improve the chances for an individual's survival.



Measured degree of DNA methylation

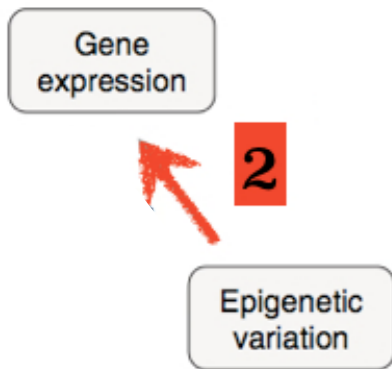
Enrichment level in MBD-Seq library



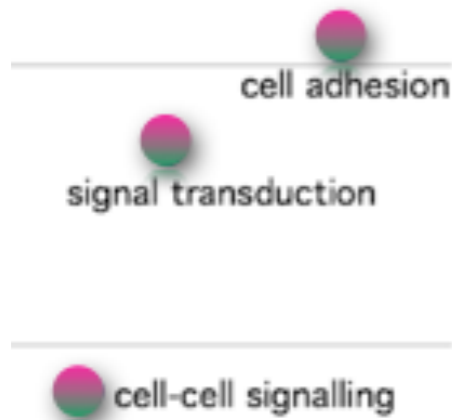
in silico analysis

(modified from Gavery and Roberts 2010)

Predicted degree of DNA methylation

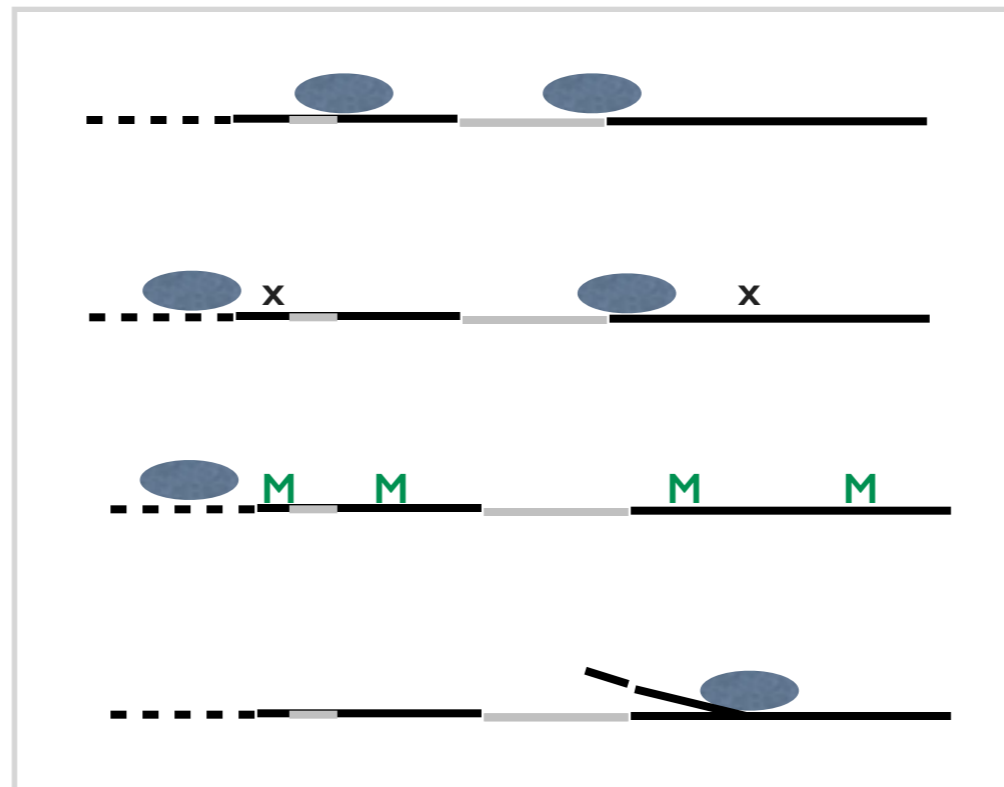


Transcriptional opportunities



sparse methylation

tissue / temporal specific and inducible genes

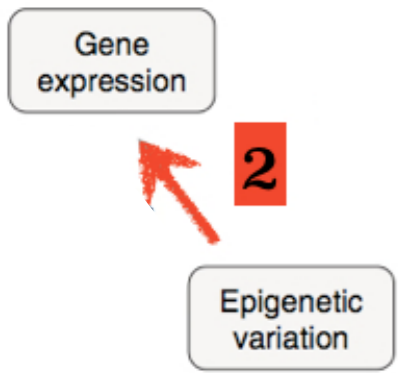


alternative start sites

sequence mutation
change AA, premature stop codon

conventional transcription
transient methylation

alternate transcript
exon skipping



Gene
expression

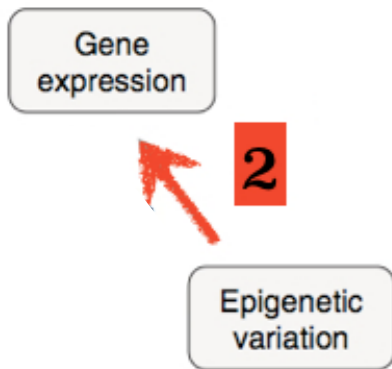


2

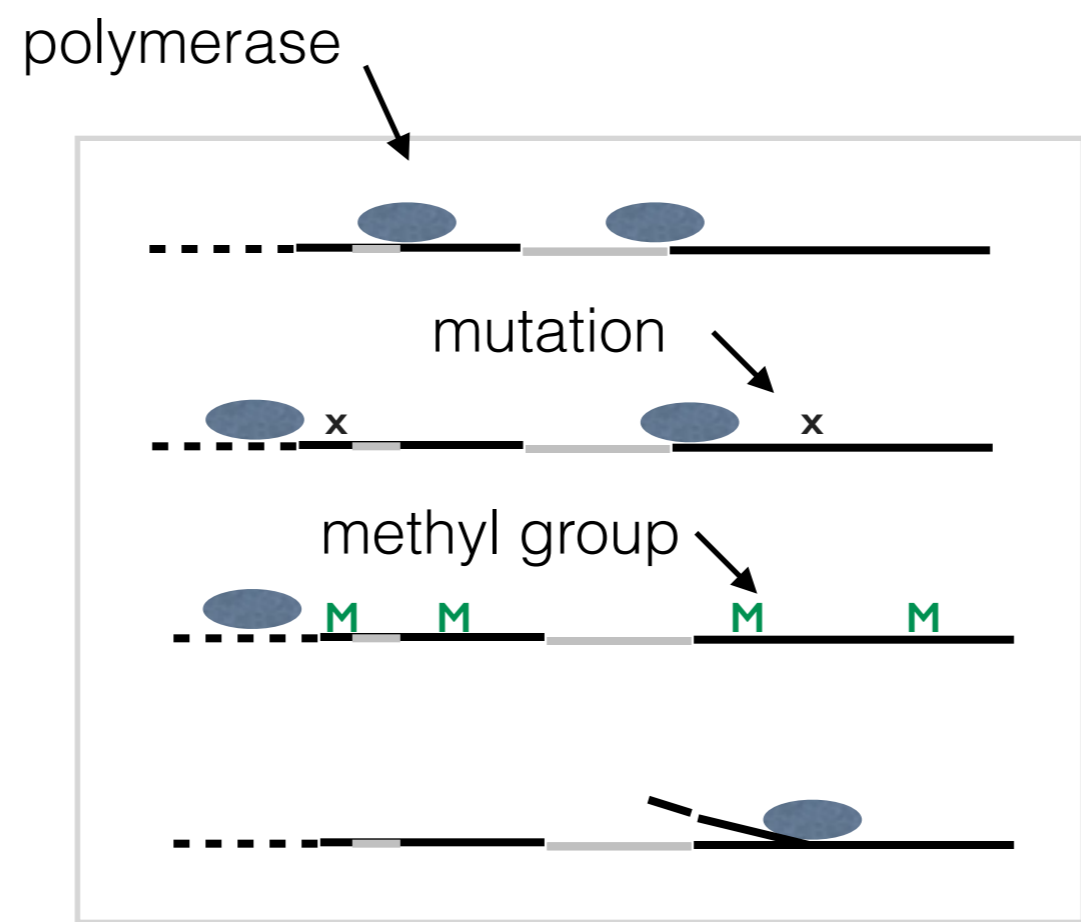
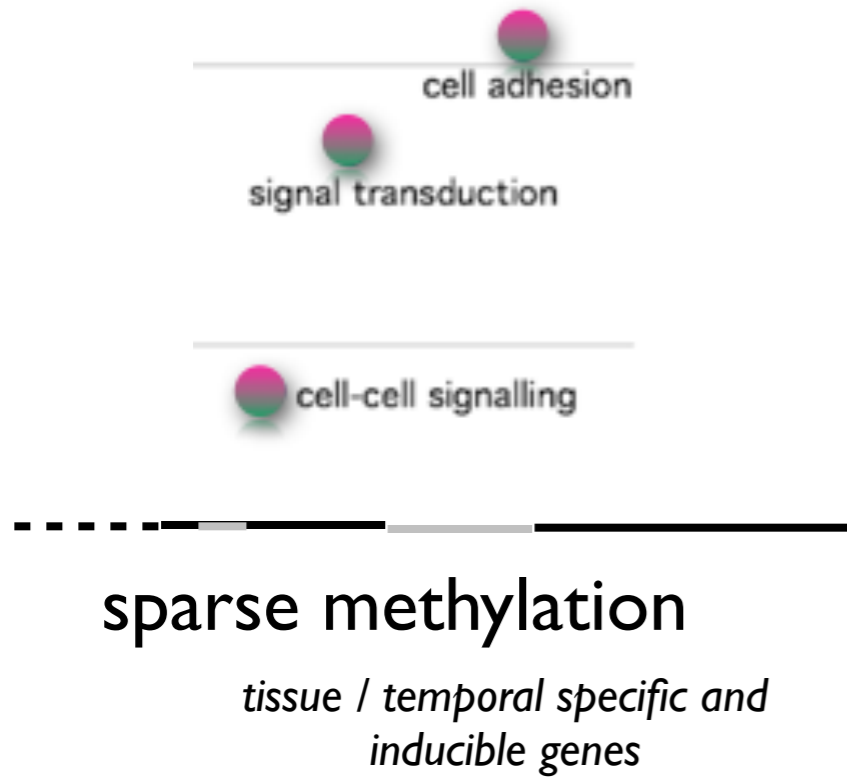
Epigenetic
variation

transcript

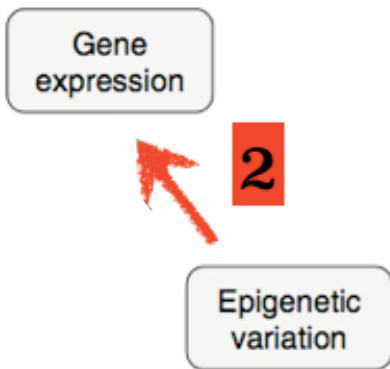




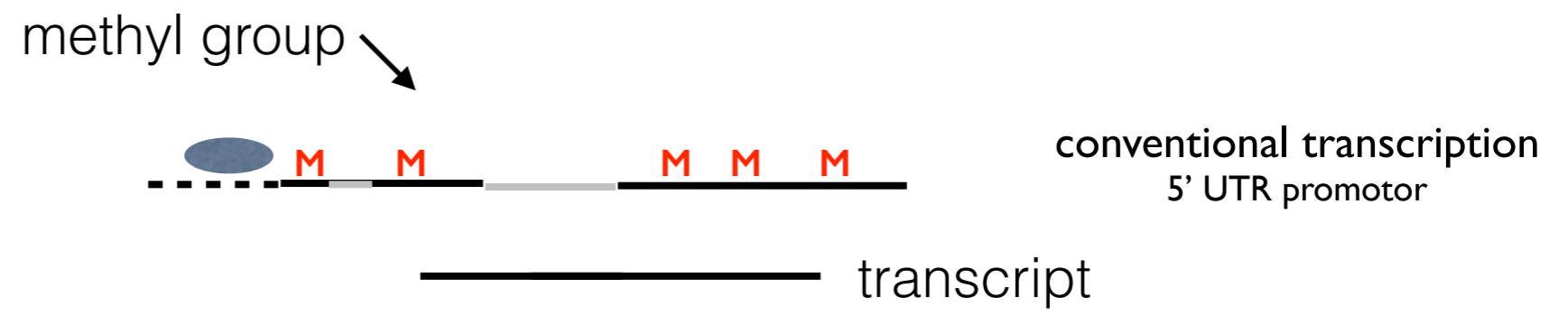
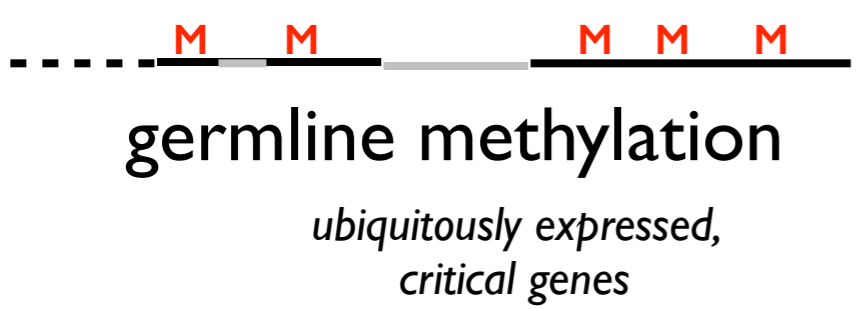
Transcriptional opportunities



- alternative start sites
- sequence mutation
change AA, premature stop codon
- conventional transcription
transient methylation
- alternate transcript
exon skipping



Transcriptional opportunities



A context dependent role for DNA methylation in bivalves

Mackenzie R. Gavery and Steven B. Roberts

Advance Access publication date 7 January 2014

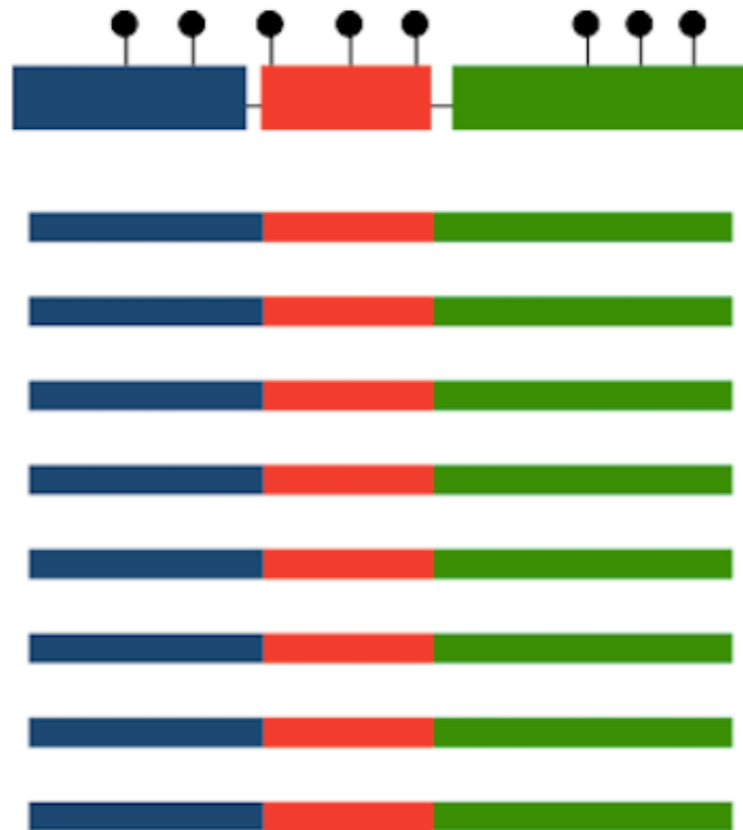
Stochastic Variation

Gene expression

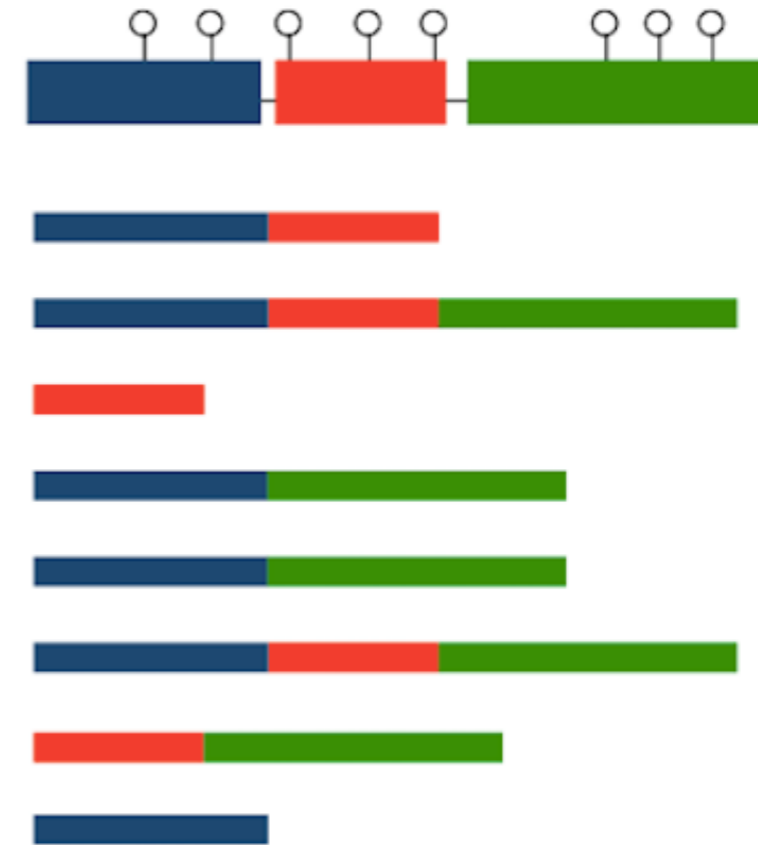
2



Epigenetic variation



housekeeping



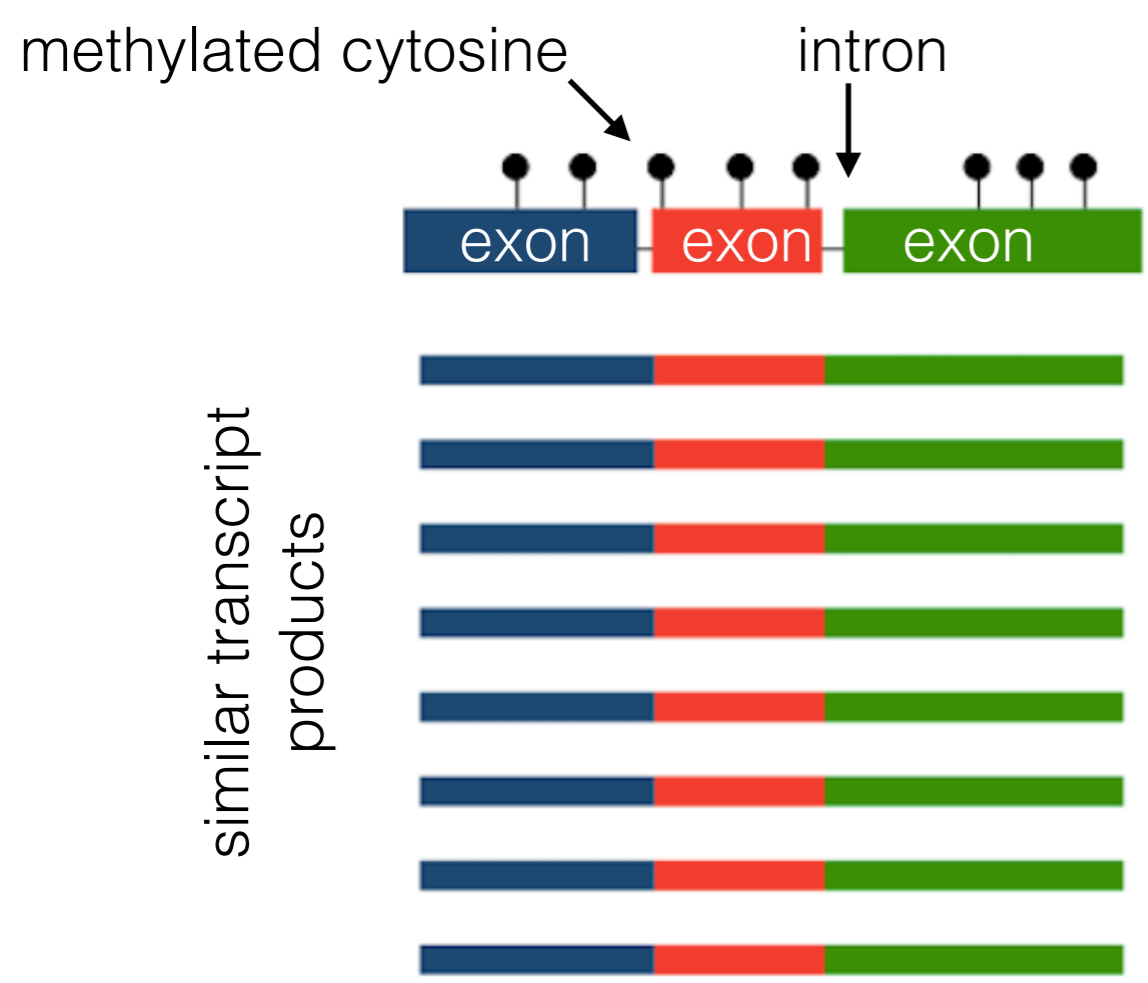
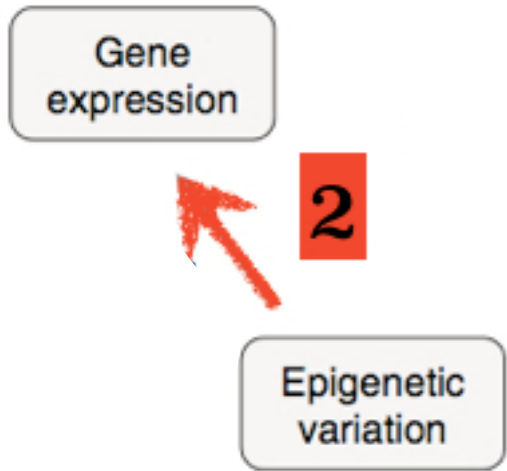
response to change

A context dependent role for DNA methylation in bivalves

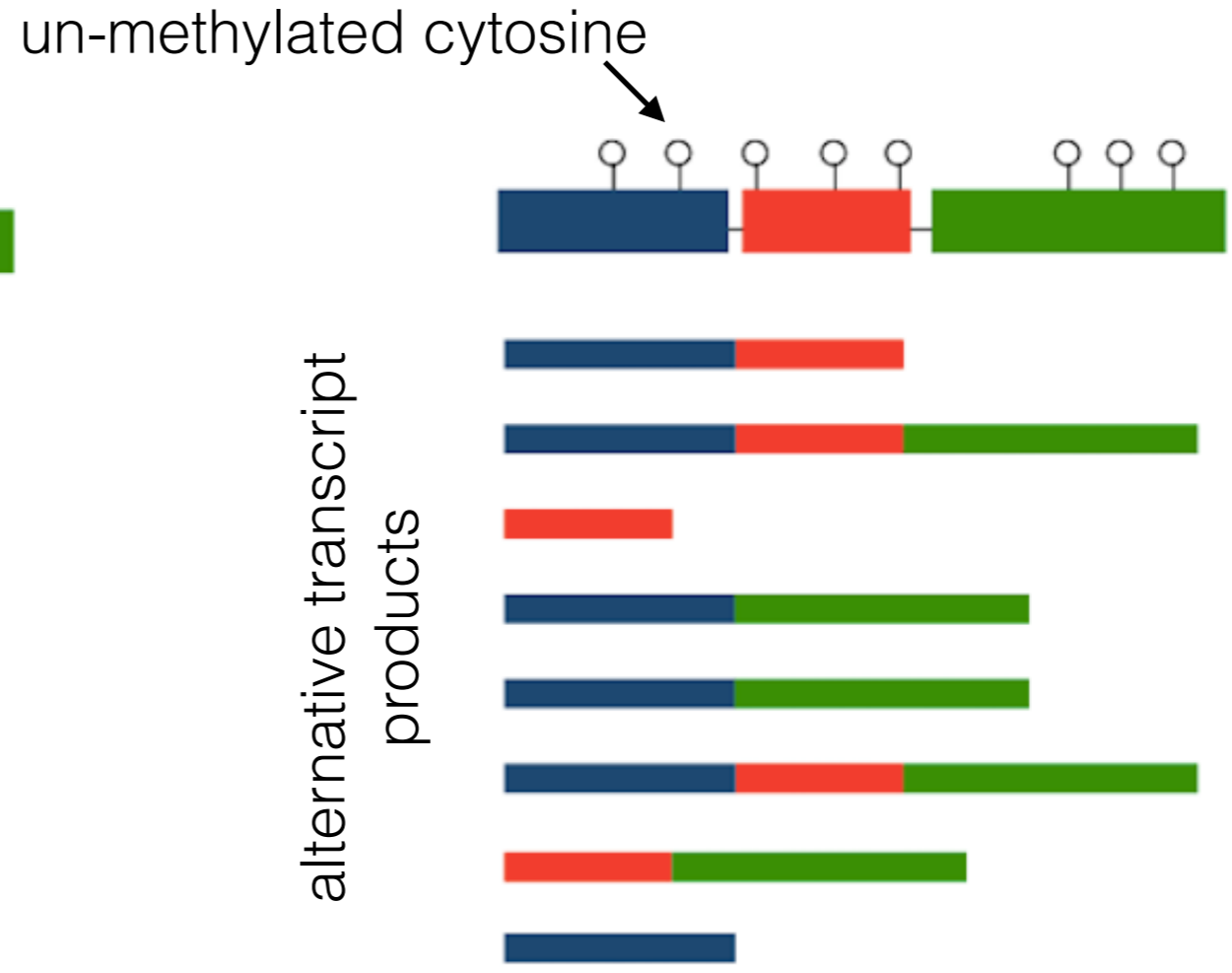
Mackenzie R. Gavery and Steven B. Roberts

Advance Access publication date 7 January 2014

Stochastic Variation



housekeeping



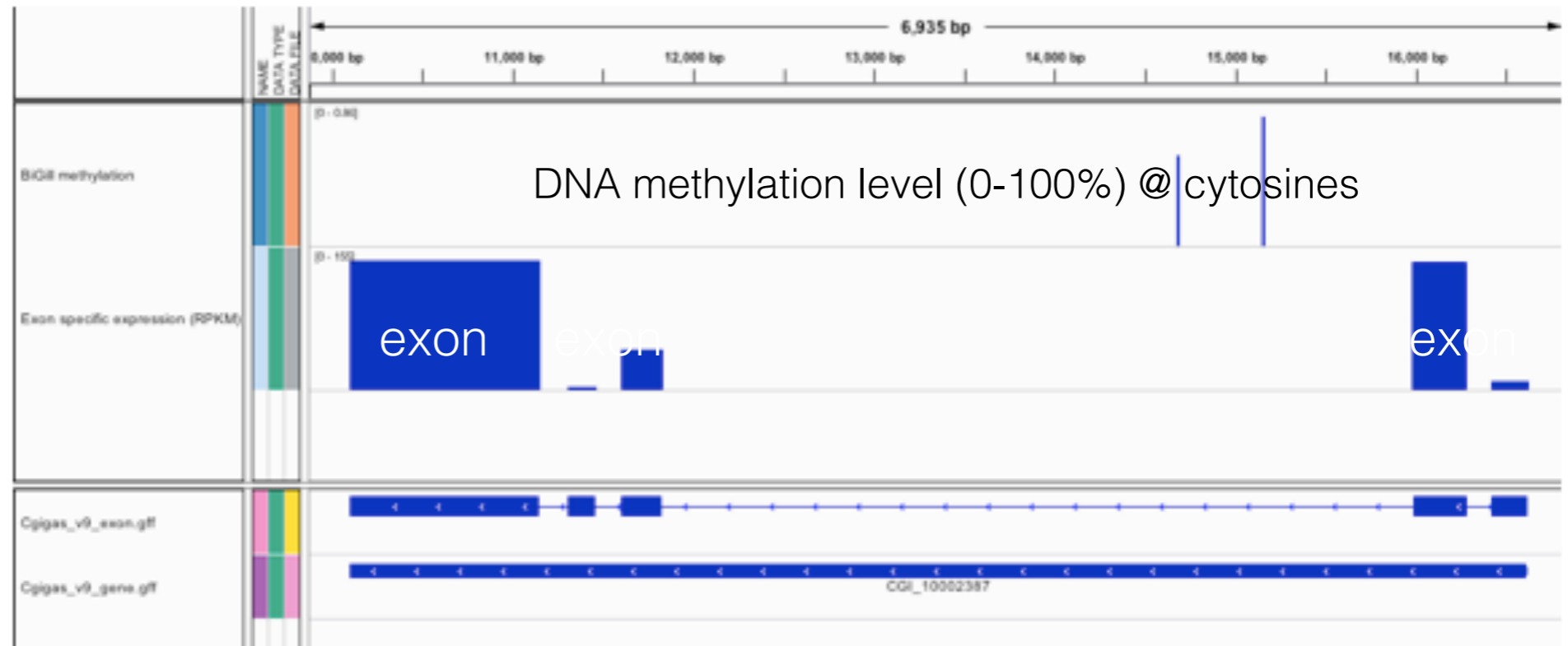
response to change

response to
change

Heat shock 70 kDa protein 12A

avg exp
199

%meth
2%

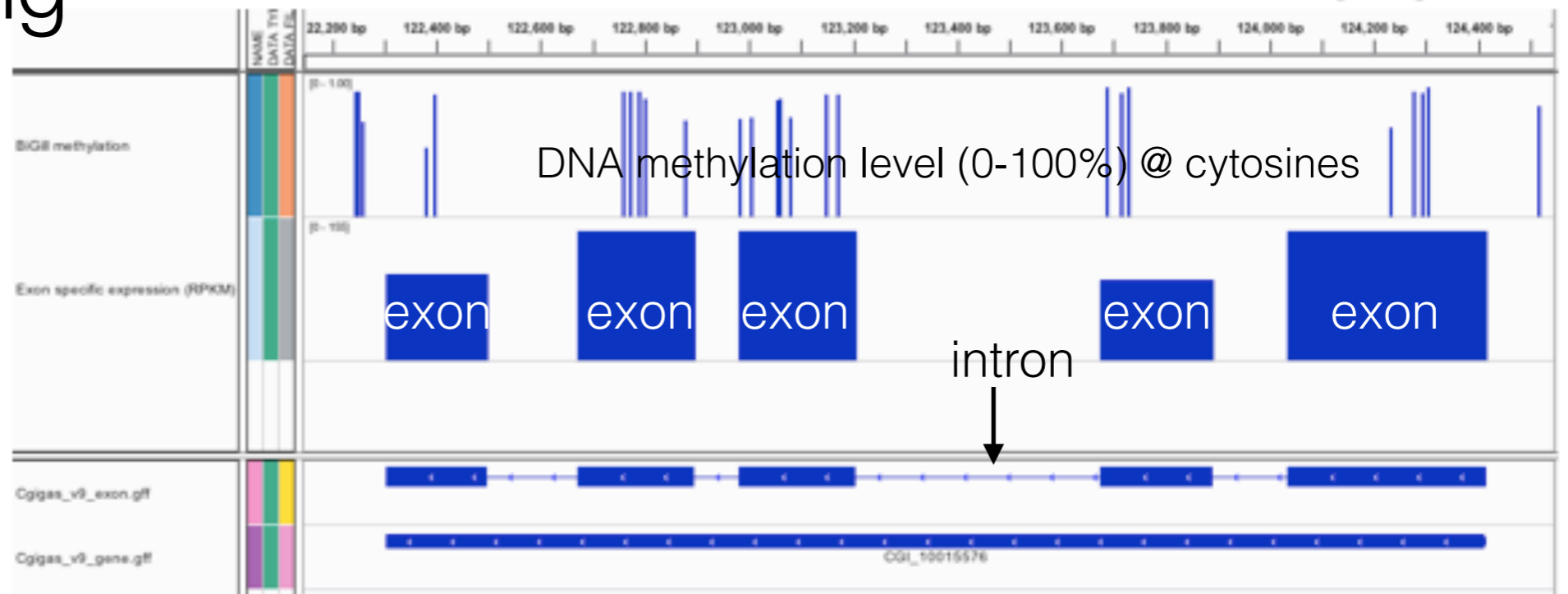


housekeeping

Tektin-2

avg exp
197

%meth
91%

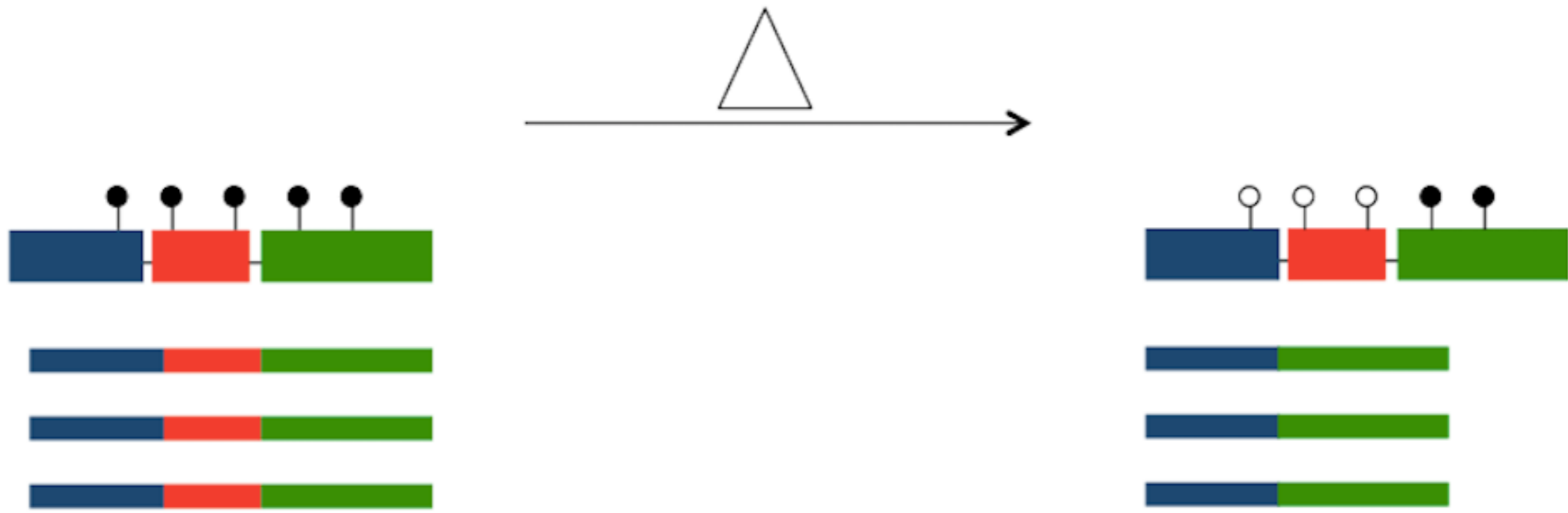
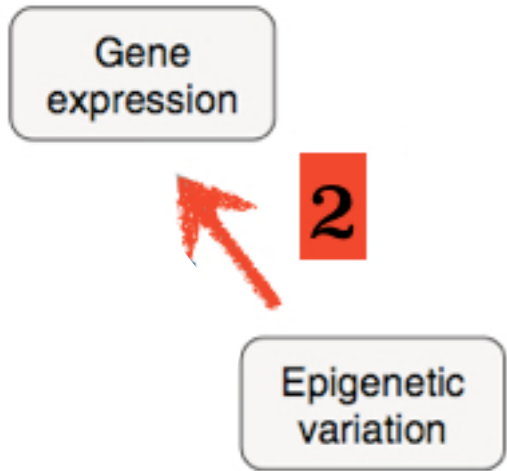


A context dependent role for DNA methylation in bivalves

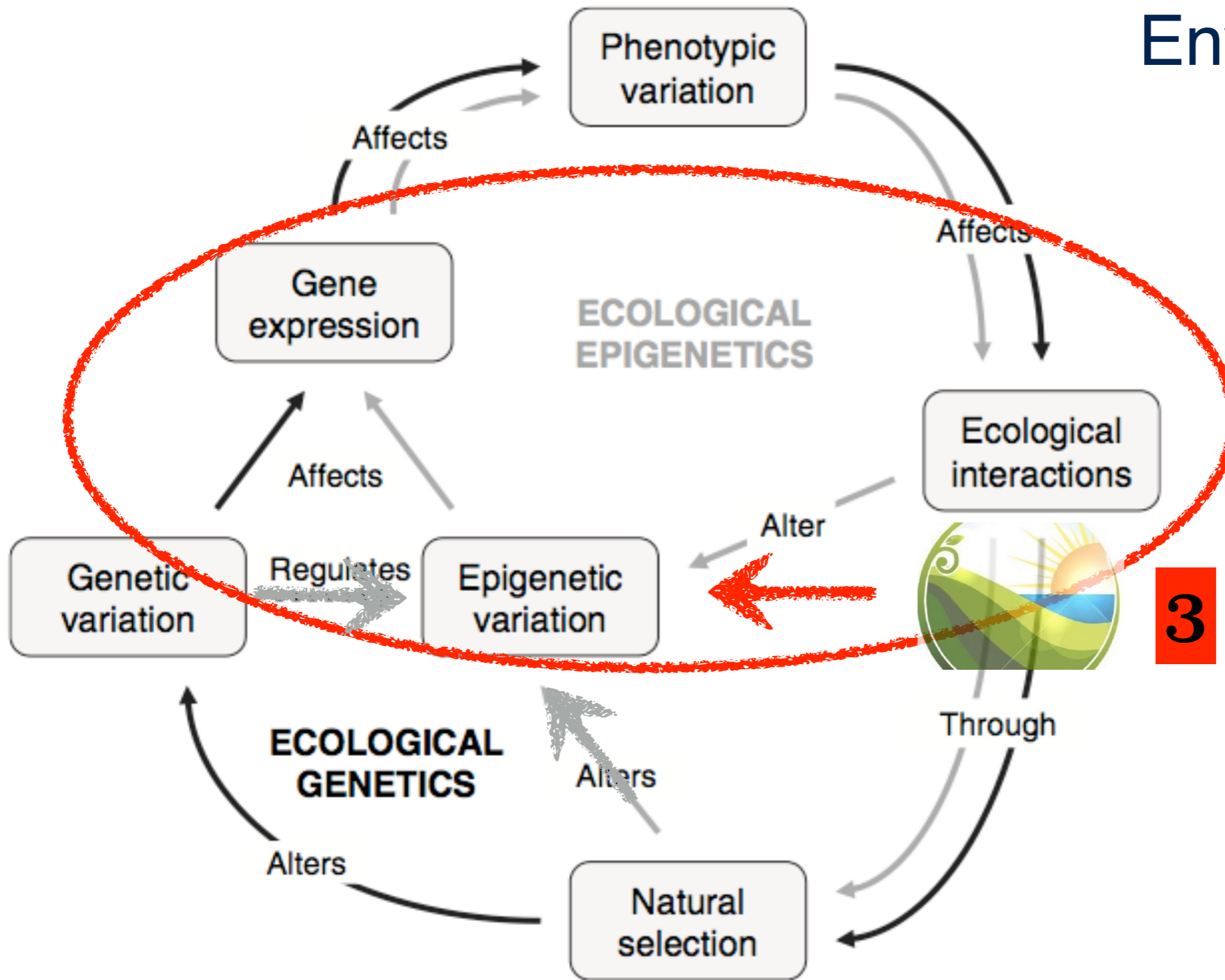
Mackenzie R. Gavery and Steven B. Roberts

Advance Access publication date 7 January 2014

Targeted Regulation



Environmental Influence



Ecology Letters, (2008) 11: 106–115

doi: 10.1111/j.1461-0248.2007.01130.x

modified from

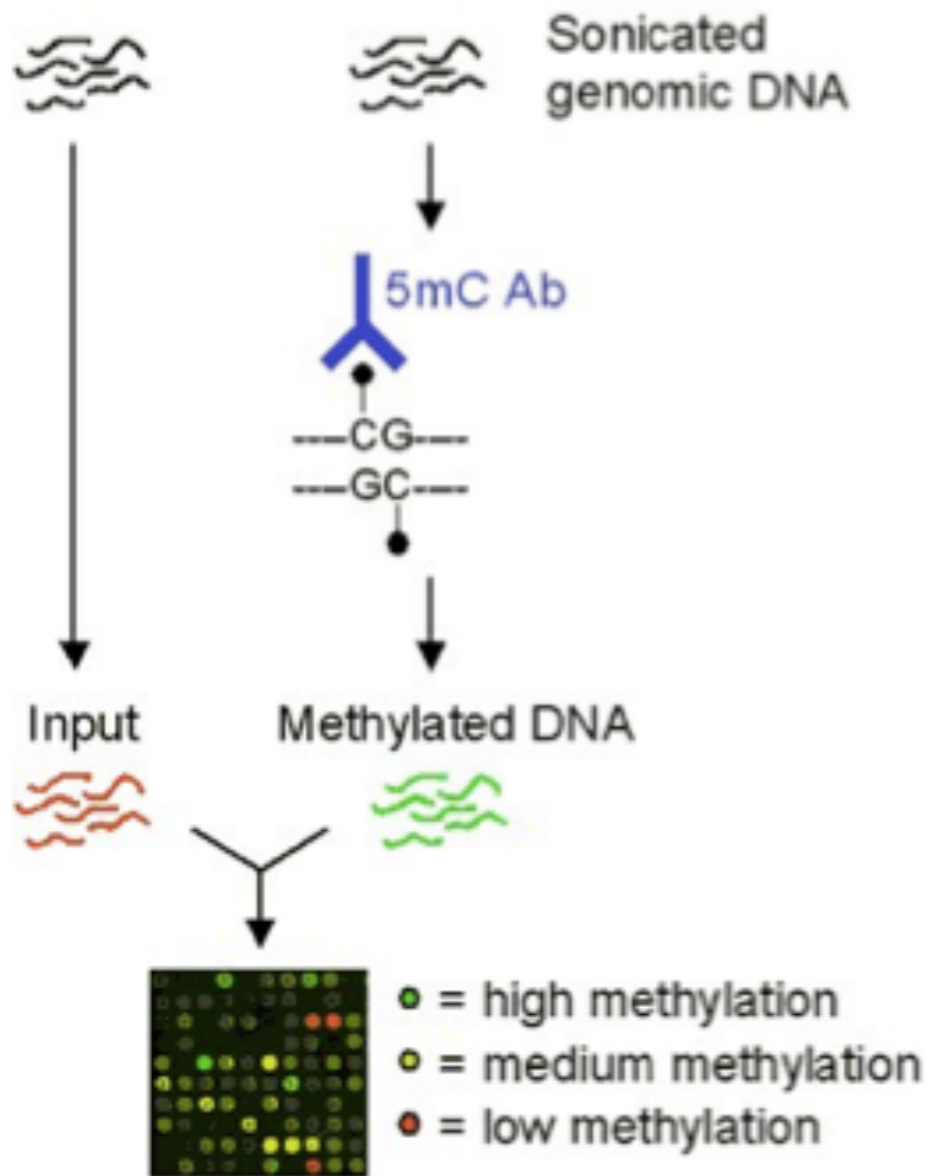
**IDEA AND
PERSPECTIVE**

Epigenetics for ecologists

Oliver Bosdorp,^{1*} Christina L. Richards² and Massimo Pigliucci³

Very new data

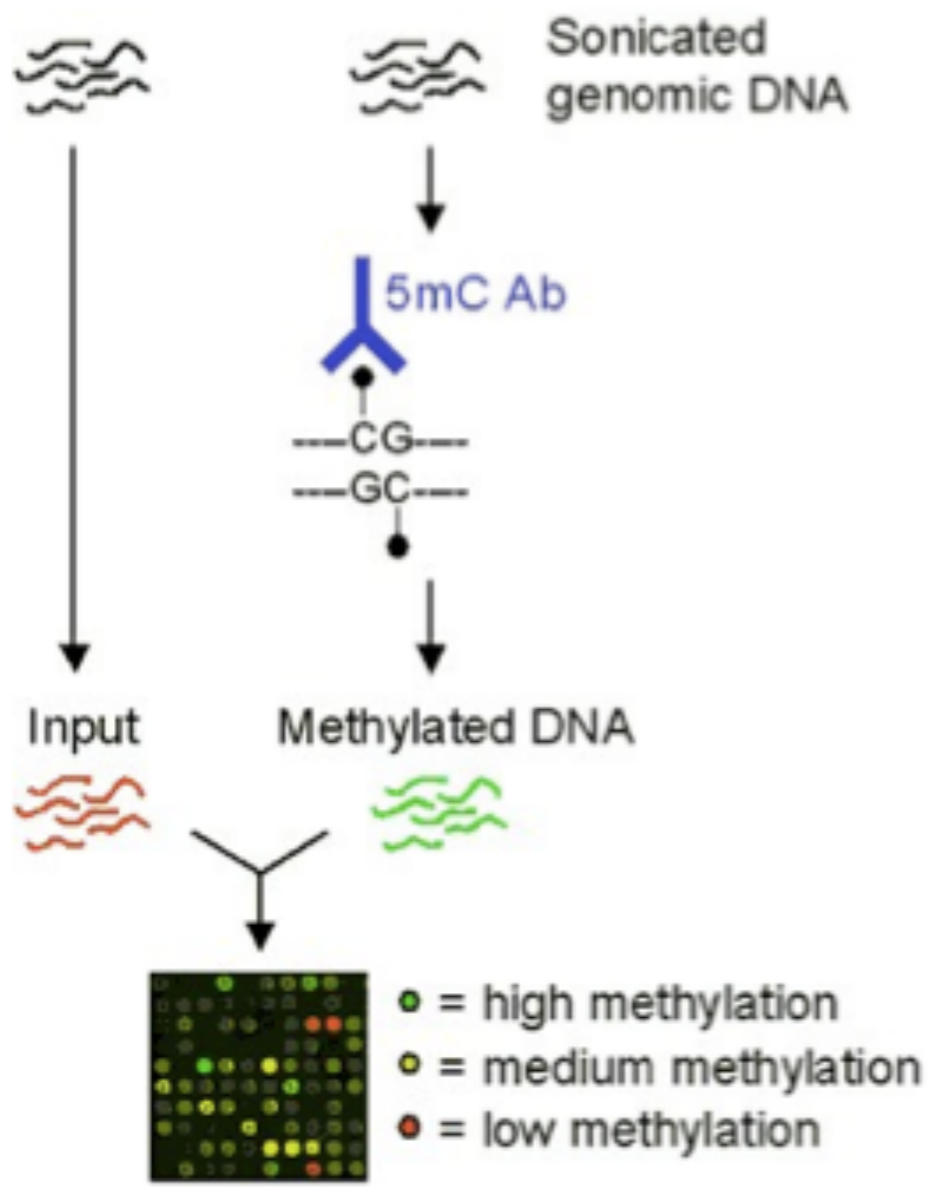
Environmental impact (Estrogens)



- 96 regions that were hypermethylated in EE2
 - 90 are in genes
 - 52 of these cross exon/intron boundary another
 - 32 are in introns
 - 6 are just in exons
- 287 regions that were hypomethylated in EE2
 - 256 are in genes
 - 138 cross exon/intron boundary
 - 114 are in introns only and
 - 4 are just in exons

Very new data

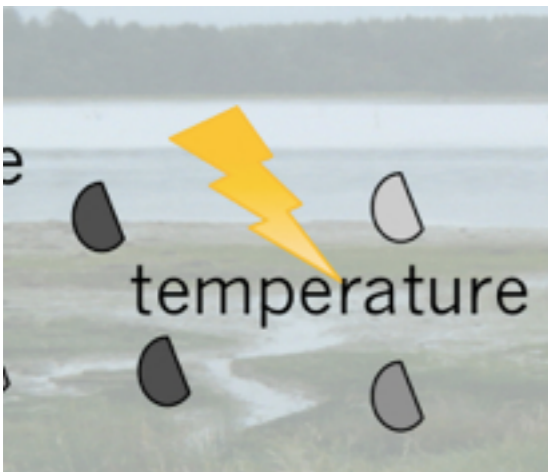
Environmental impact (Estrogens)



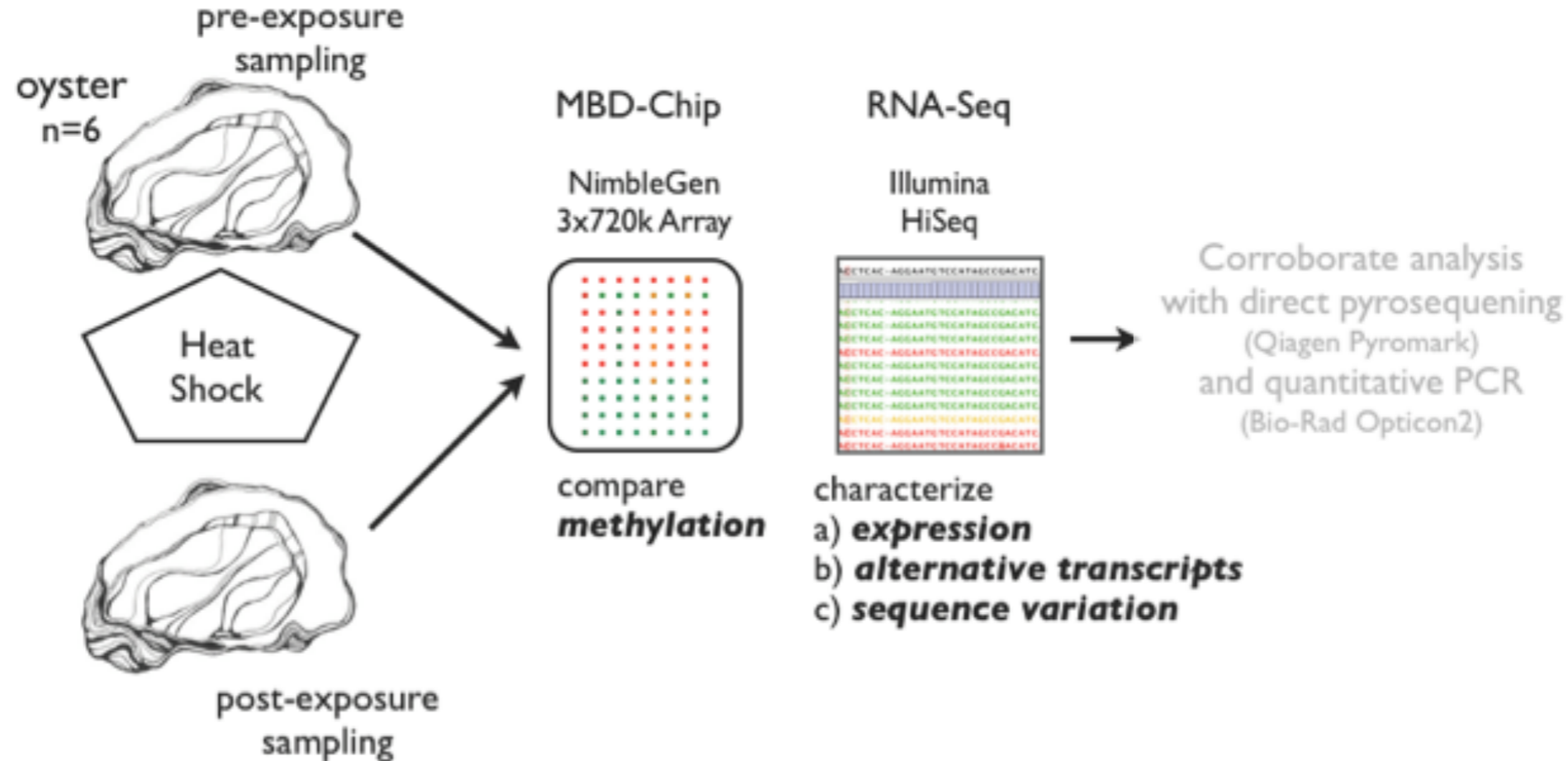
Exposure results in significant hypomethylation

Very new data

Environment and gene expression

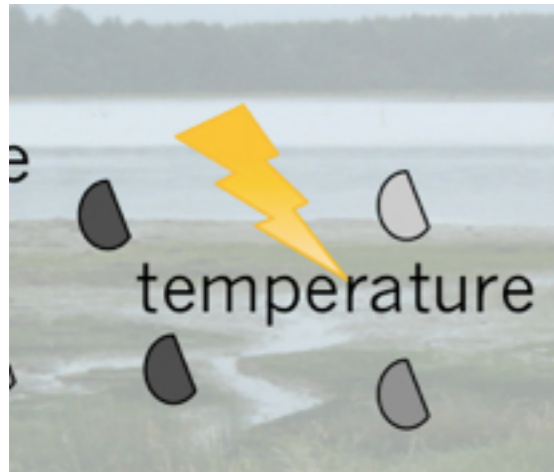


stochastic or targeted?



Very new data

Environment and gene expression

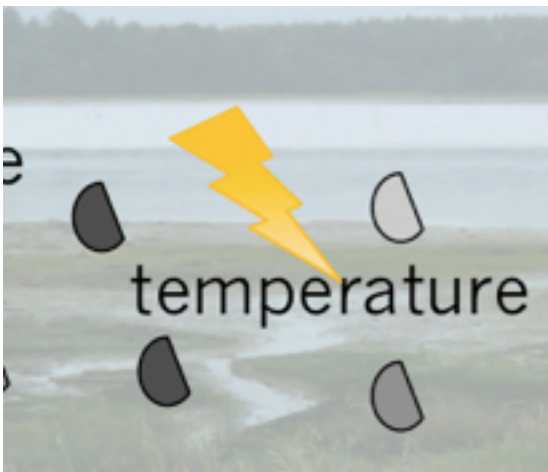


stochastic or targeted?

Oyster	Hypo-methylated	Hyper-methylated
2	7224	2803
4	6560	3587
6	7645	4044

Very new data

Environment and gene expression



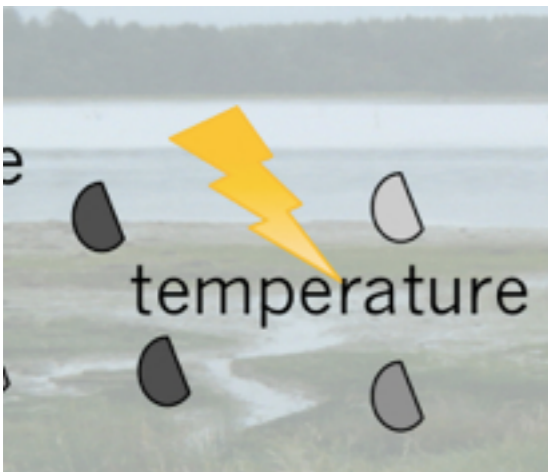
stochastic or targeted?

Oyster	Hypo-methylated	Hyper-methylated
2	7224	2803
4	6560	3587
6	7645	4044

No obvious association with genome feature including *differentially expressed genes*

Very new data

Environment and gene expression

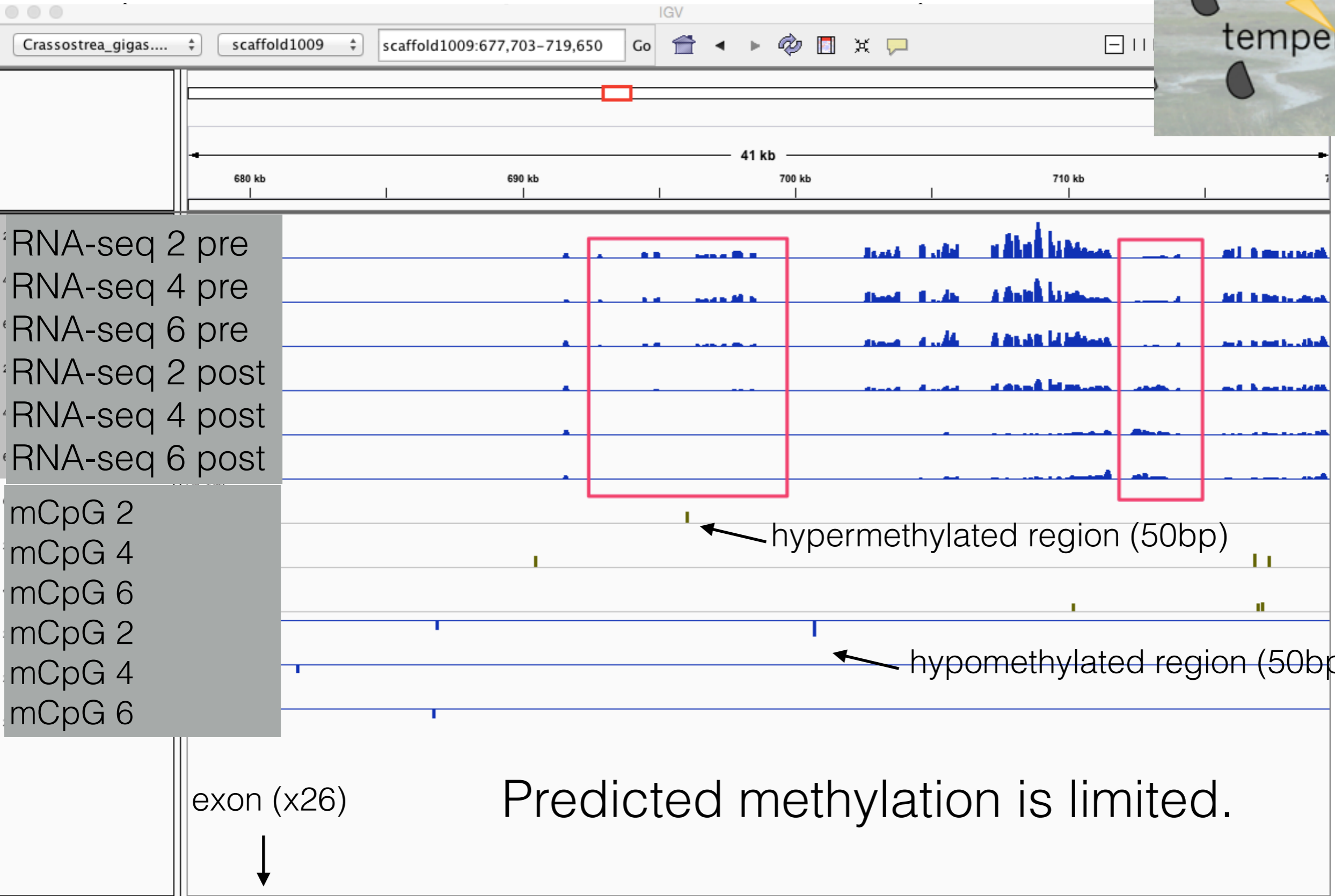
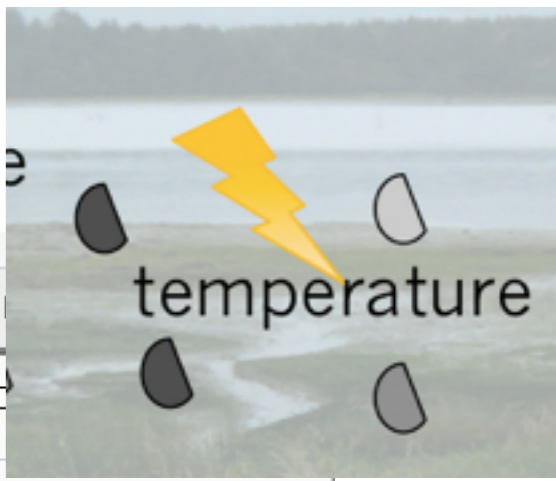


*stochastic or targeted
or ..?*

Oyster	Hypo-methylated	Hyper-methylated
2	7224	2803
4	6560	3587
6	7645	4044

Changes in methylation (either direction) are more prevalent in introns, repeats, and transposable elements.

Very new data



exon (x26)
↓

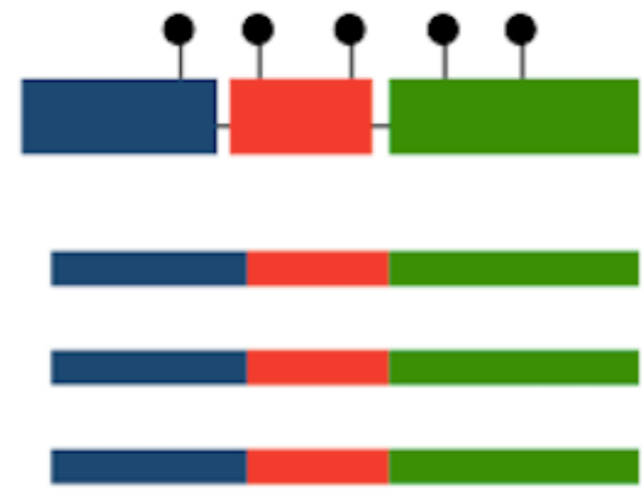
Predicted methylation is limited.



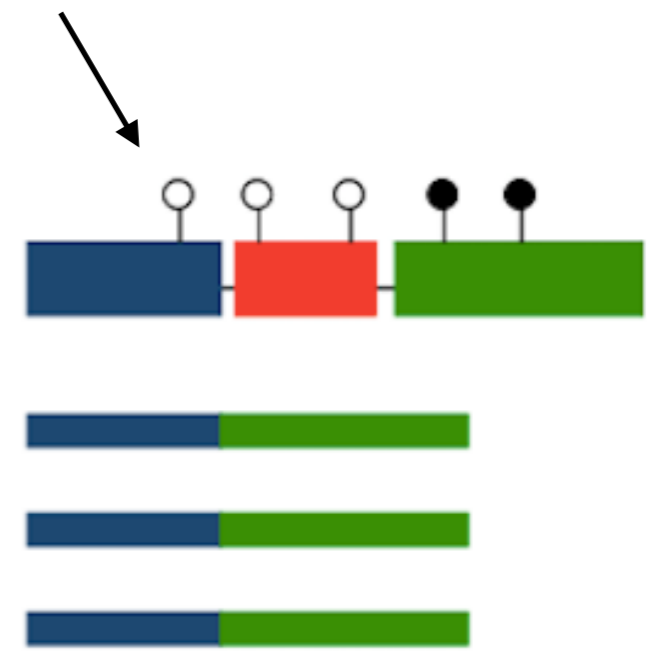
Gene expression

2

Epigenetic variation



hypomethylated region (50bp)



not in this experiment

not even consistent methylation changes at loci level

Gene expression

2

Epigenetic variation



more questions....
interesting but what is controlling?

Take Home

Oyster genome has a fantastic degree of diversity contributing to phenotypic plasticity & adaptation potential.

Take Home

Oyster genome has a fantastic degree of diversity contributing to phenotypic plasticity & adaptation potential.

- large gene families
- very high mutation rate (snp/50bp)
- numerous exons per gene (potential for alternatives)
- genome full of repeats region
- high number of transposable elements
- lack of methylation of transposable suggest mobility
- family variation of methylation
- limited methylation environmental response genes is associated with spurious transcription
- inheritance of epigenetic marks as mechanism of improved adaptation



Next Steps

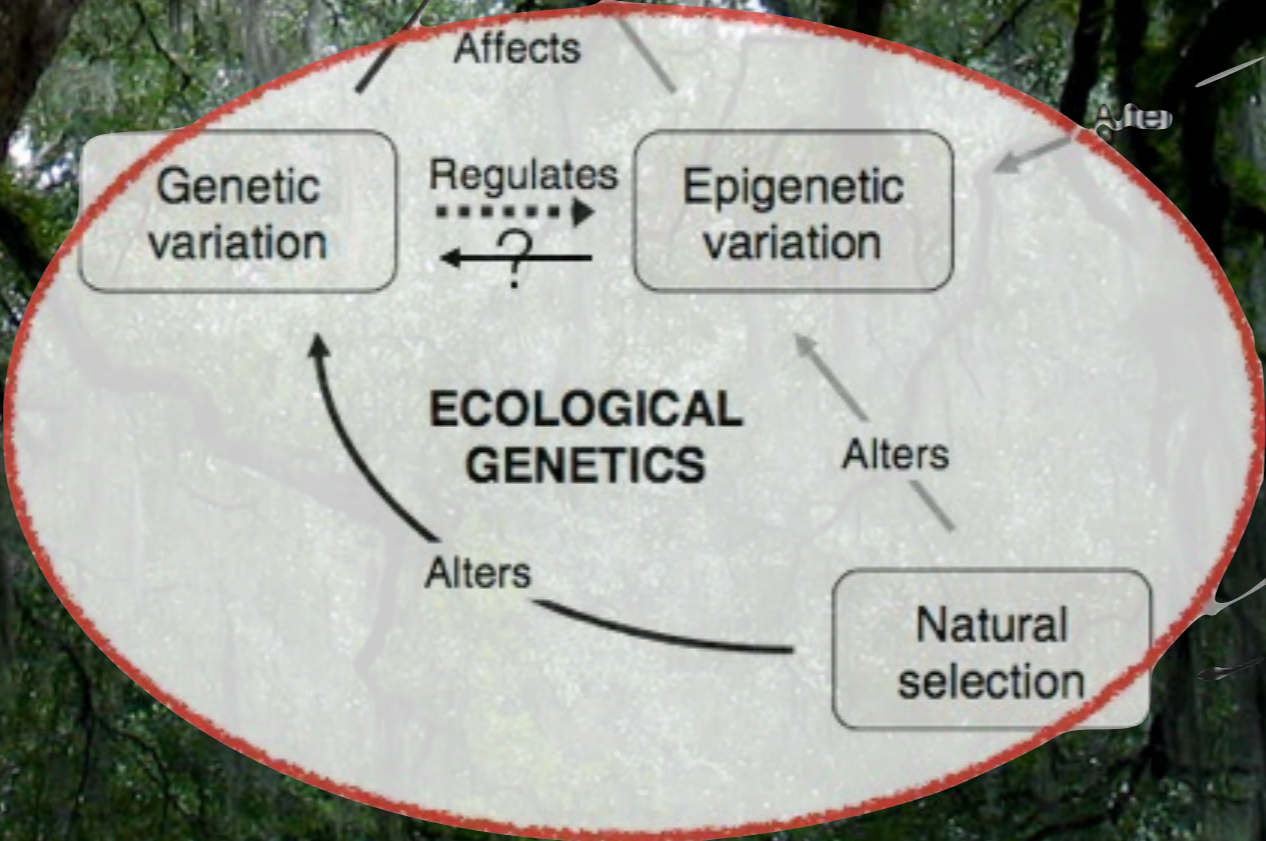
Dig into the
heat-shock data

Transposable
Elements

Consider other
epigenetic process

Photo credit: Flickr, Creative Commons, csessums

Next Steps



~~Very new data~~
Heritability
Plasticity
Local Adaptation

*Genetics
versus
Epigenetics*

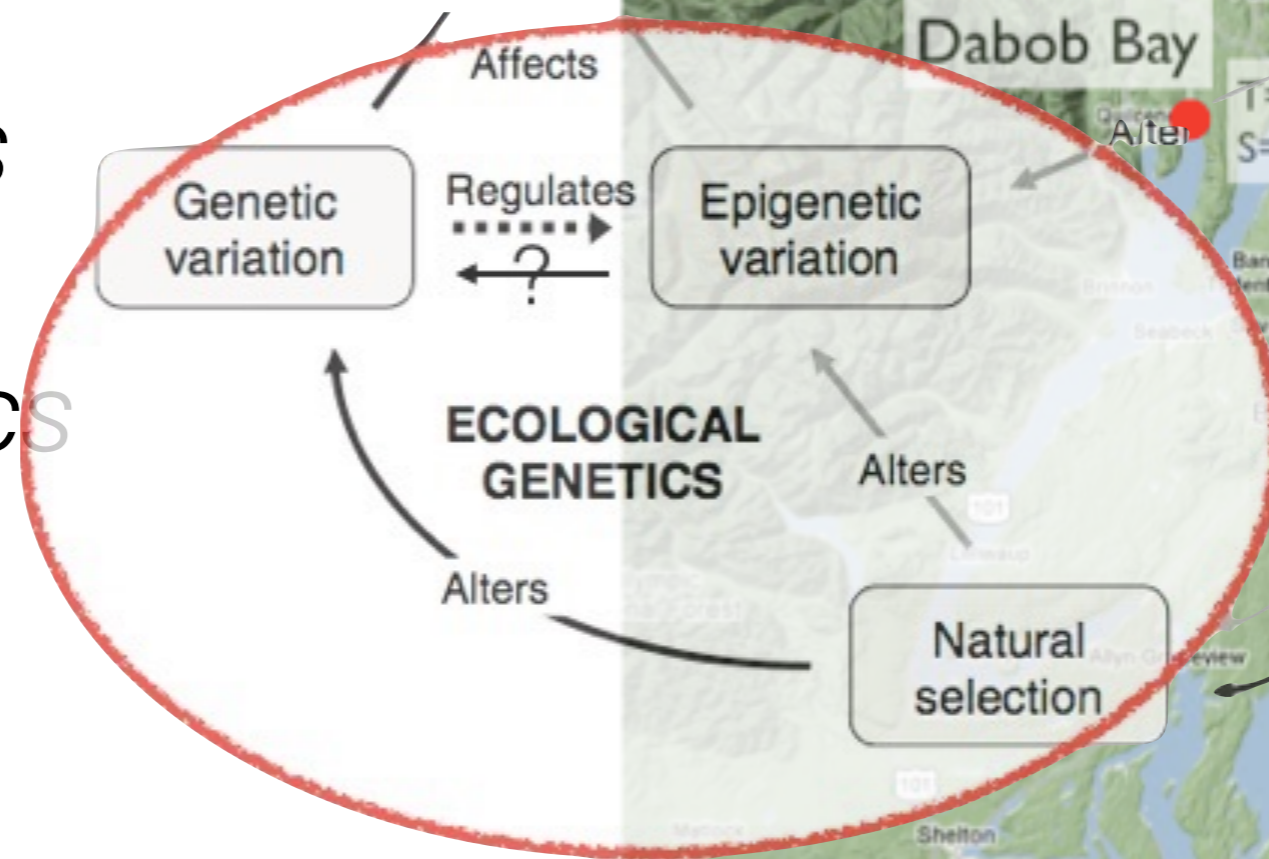
Common Garden Experiment



~~Very new data~~
Heritability
Plasticity
Local Adaptation

*Genetics
versus
Epigenetics*

Common Garden Experiment



Open Science

Open Science

web-native scholarship

Sharing



Photo credit: Flickr, Creative Commons, speechless

Open Data

Open Methods

Open Data

Sharing raw data

Example



Based on Ensembl release 40 - Aug 2006

Use Ensembl to...

- Search Contig Browser
- Data mining [BioMart]
- Export data
- Download data
- Digital Differential Display
- Venn Diagrams

Docs and downloads

- About Contig Browser
- Contig Browser data

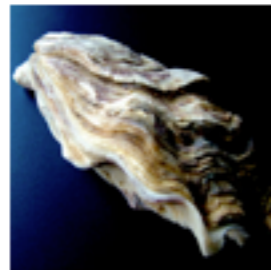
Other links

- Home
- Sigenae
- Agenae



Browse the *Crassostrea gigas* contigs database

About the *Crassostrea gigas* contigs database



Assembly

The Oyster EST contig browser aims to produce and maintain an automatic annotation of Oyster EST libraries. This database [GigasDatabase](#) was initiated within the frame of the [AquaFirst](#) European project, it now gathers EST sequences produced by a [Marine Genomics Europe](#) project (GOCE-CT-2004-505403) and a [Genoscope](#) project. [GigasDatabase](#) is regularly updated in the context of the ANR project "Gametogenes" (ANR-08-GENM-041).

Annotation

Contigs were annotated searching sequence homologies against following databases:

- UniProtKB/Swiss-Prot Release 2011_03 of 08-Mar-2011
- RefSeq Protein Index Blast of 09-Jan-2011
- Pfam Release 24.0 of Jul-2009
- Sigenae Oyster Contigs V8
- RefSeq RNA Index Blast of 09-Jan-2011
- TIGR Fugu FGI 3.0

Statistics

Assembly:	pcg8, Mar 2011
Number of base pairs (Contigs):	67,875,621
Number of sequences (EST/mRNA):	1,013,570
Total number of contigs:	82,312
Number of singlets:	20,148
Polymorphism:	Mar 2011
Number of putative SNPs:	29,493
Number of contigs including SNP:	10,392

Example Data Points

Jump directly to the example entry points:

Contig: [BQ426804.p.cg.8](#)

SNP: [snp_EW777925_683](#)

Example



Based on Ensembl release 40 - Aug 2006

Use Ensembl to...

Browse the *Crassostrea gigas* contigs database

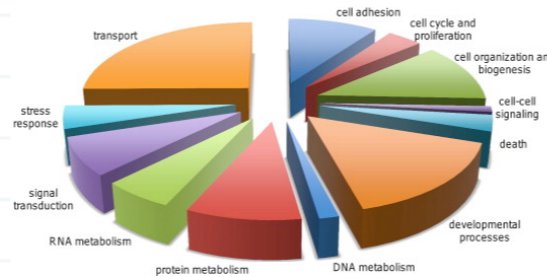
Search Contig Browser

About the *Crassostrea gigas* contigs

Statistics

Table 1. Summary of assembly and RNA-Seq statistics for de novo assembly and reference mapping (GigasDatabase v8).

		De novo assembly	Reference mapping
Assembly	Mapped reads	8,407,963	29,107,760
	Unmapped reads	36,944,698	16,244,901
	Contigs	18,510	77,433
	Average contig length	276	554
	Average contig coverage	62	16
RNA-Seq	Differentially expressed genes	2991	427
	Enriched GO biological process	15	3
	Contigs annotated to GO Slim	3931	7296



Assembly:	pcg8, Mar 2011
Number of base pairs (Contigs):	67,875,621
Number of sequences (EST/mRNA):	1,013,570
Total number of contigs:	82,312
Number of singlets:	20,148
Polymorphism:	Mar 2011
Number of putative SNPs:	29,493
Number of contigs including SNP:	10,392

Sigenae

Agene

Characterizing short read sequencing for gene discovery and RNA-Seq analysis in *Crassostrea gigas*

Mackenzie R. Gavary, Steven B. Roberts*



- ▶ Sigenae Oyster Contigs V8
- ▶ RefSeq RNA Index Blast of 09-Jan-2011
- ▶ TIGR Fugu FGI 3.0

Example Data Points

Jump directly to the example entry points:

Contig: [BQ426804.p.cg.8](#)

SNP: [snp_EW777925_683](#)

Raw Data

1) As sequencing facility provides data, files are downloaded to our local NAS (owl), in the root `nightingales` directory.
<http://owl.fish.washington.edu/nightingales/>

2) The Nightingales Google Spreadsheet is updated.

3) Update the Nightingales Google Fusion Table with new information from the Nightingales Google Spreadsheet. This is accomplished by:

- deleting all rows in the Nightingales Google Fusion Table (Edit > Delete all rows)
- Importing data from the Nightingales Google Spreadsheet (File > Import more rows...)

4) Once metadata is included in the Google Fusion Table [Nightingales](#), respective data files are moved to subdirectory labelled with first letter of genus followed by species ie `c_gigas` . Check url in Nightingales Fusion table to ensure it is accurate.

Nightingales

Imported at Wed Jun 25 07:25:38 PDT 2014 from Nightir

[Add Attribution](#) - Edited on December 15, 2014

File Edit Tools Help

Rows 1

Filter

No filters applied. Sorted by SeqDat

1-100 of 153

SeqID



This repository Search



sr320 / LabDocs

github
SOCIAL CODING

branch: master

LabDocs / DMPseq.md



sr320 on Jan 12 added nightingales backup info

2 contributors



Querying Disparate Datasets



Querying Disparate Oyster Datasets | qDOD

The goal of this project is to produce a web-based interface for querying and visualizing *Crassostrea gigas* genomic datasets. This site serves as a portal for documenting our efforts, providing user access, as well as a means to gather feedback.

Preliminary Phase: Aggregating Datasets

Using [SQLShare](#) as a platform we have already begun to aggregate and format data. Anyone can view (and contribute) using the tag "qDOD". Below is a table describing some of the relevant datasets. "Snapshot" provides you with a screenshot of the data in SQLShare and "Direct Link" brings you directly to the data in SQLShare. You can also [open the table in a new webpage](#).

qDOD online			
qDOD_Cgigas_gene_fasta	sequence fasta file. Exon only.	http://goo.gl/ogCxl	https://sqlshare.esc.edu/
qDOD_Zhang_Gil_gene_RNA-seq	Gill RNA-seq data (gene based)	http://goo.gl/8oISR	https://sqlshare.esc.edu/
qDOD_Zhang_Mgo_gene_RNA-seq	Male Gonad RNA-seq data (gene based)	http://goo.gl/6buVz	https://sqlshare.esc.edu/

DATA



Preliminary Phase: Aggregating Datasets

Using [SQLShare](#) as a platform we have already begun to aggregate and format data. Anyone can view (and contribute) using the tag "qodod". Below is a table describing some of the relevant datasets. "Snapshot" provides you with a screenshot of the data in SQLShare and "Direct Link" brings you directly to the data in SQLShare. You can also [open the table in a new webpage](#).

qDOD online			
qDOD_Cgigas_gene_fasta	sequence fasta file. Exon only.	http://goo.gl/ogCxl	https://sqlshare.esc.edu
qDOD_Zhang_Gil_gene_RNA-seq	Gill RNA-seq data (gene based)	http://goo.gl/8oISR	https://sqlshare.esc.edu
qDOD_Zhang_Mgo_gene_RNA-seq	Male Gonad RNA-seq data (gene based)	http://goo.gl/6buVz	https://sqlshare.esc.edu

DATA

- Your datasets
- All datasets
- Shared datasets
- Recent activity... **293**
- Recently viewed »

- Upload dataset
- New query

YOUR TOP VIEWED

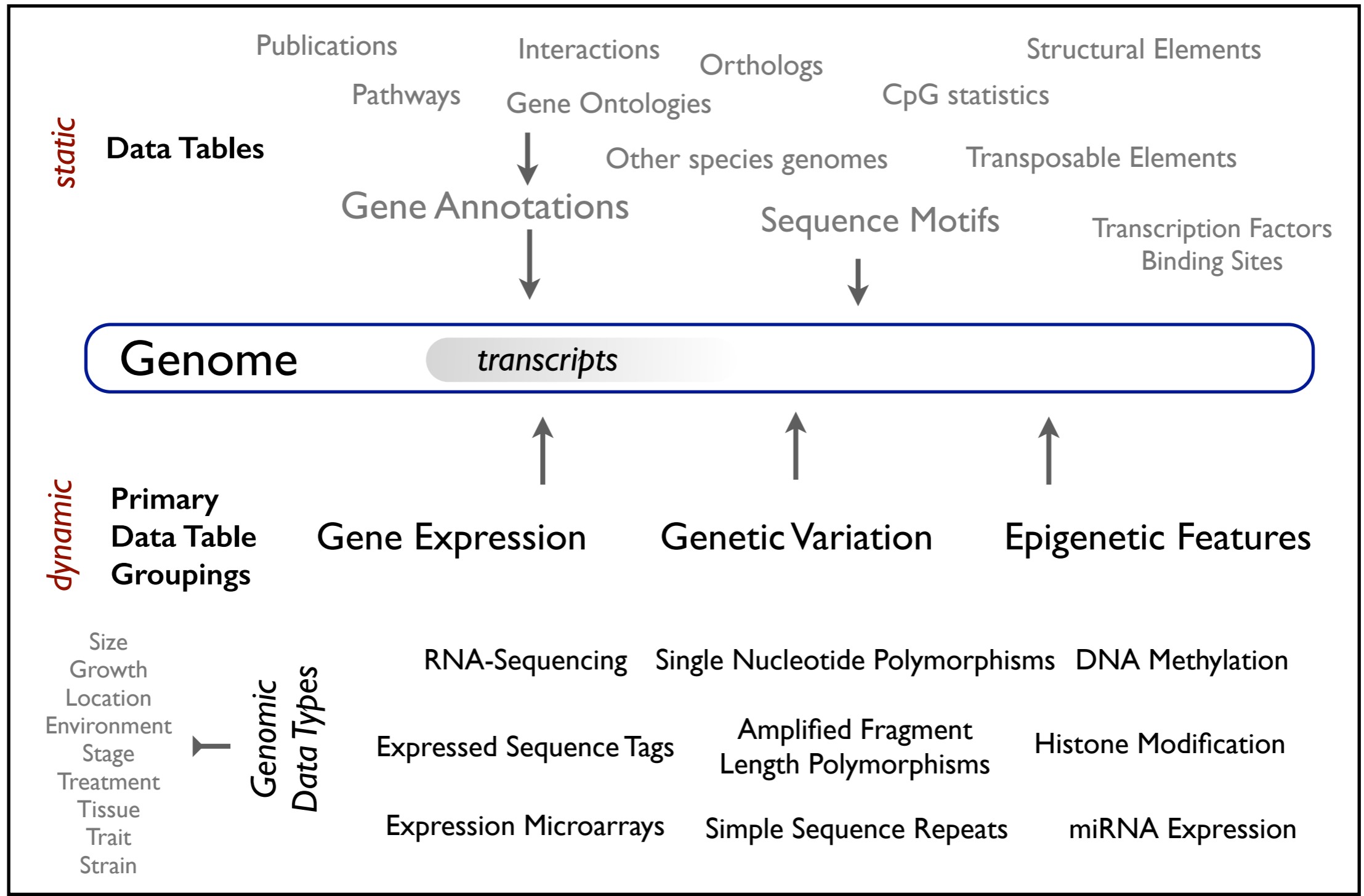
- qDOD_scaffol... 35
- oyster_v9_mR... 28
- oyster_v9_CD... 23
- BiGO_betty_pl... 16
- qDOD_Cgigas... 16

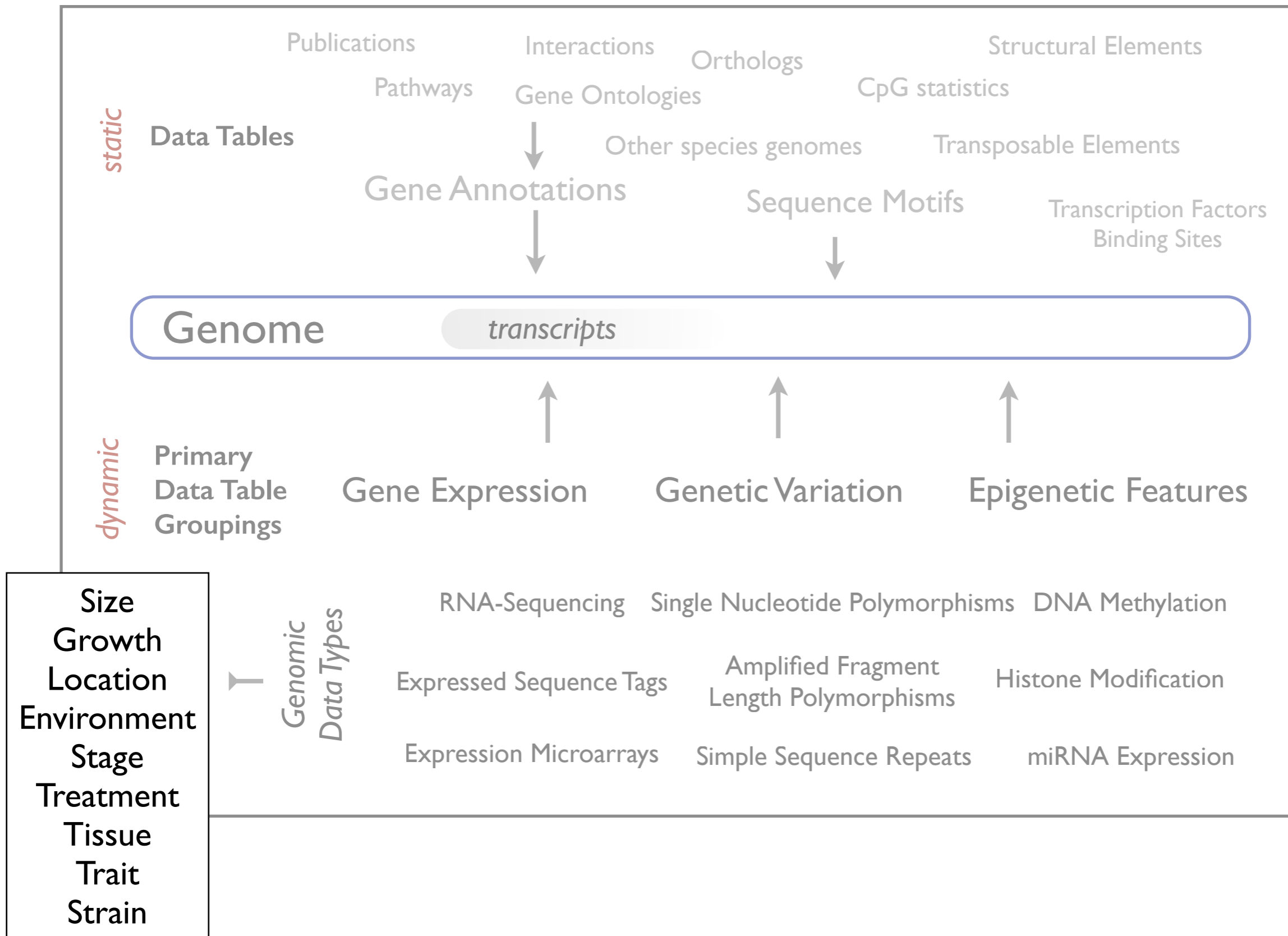
POPULAR TAGS

- protein 47
- oyster 45
- proteomics 21
- bioinformatics 21
- oa 20

Tagged Datasets | qdod

Name	Sharing /
BiGO_Methylation_oysterv9_GFF GFF format file with percent methylation (score) information for bsmmap gff gonad methylation oyster qdod sperm	sr320
qDOD_gene_length Sequence lengths of all genes in the oyster genome. This only includes trans fasta gene length oyster qdod	sr320
BiGill_methratio_v9_A.txt Methylation ratio information (MBD-Seq) from oyster gill tissue. Combir bsmmap epigenetic gill mbd methylation oyster qdod	sr320
qDOD_Protein_Sequences Amino acid sequence for all proteins (v9) fasta oyster protein qdod	sr320
qDOD_oyster_gene_exon_number Number of exons for all genes in oyster genome (v9) exon oyster qdod	sr320
Zhang_et_al_SuppTable14 S14. Transcriptomic representation of genes (RPKM) at different develc oyster qdod zhang	sr320
qDOD_Cgigas_gene_fasta Tabular format of Cgigas gene sequence fasta file Derived using Data: fasta oyster qdod	sr320
qDOD_Zhang_Gil_gene_RNA-seq Paired end Gill RNA-Seq data from Zhang et al 2012. Exported gill oyster qdod rna-seq zhang	sr320
Mgo_RNAseq_transcript Paired end Male Gonad RNA-Seq data from Zhang et al 2012 Exported f oyster qdod rna-seq zhang	sr320
qDOD_Zhang_Mgo_gene_RNA-seq Paired end Male Gonad RNA-Seq data from Zhang et al 2012 gonad oyster qdod rna-seq sperm zhang	sr320





```
SELECT cgslim.CGI_ID,Description,evaluate,SPID,GOID,term,GOSlim_bin,sequence
FROM [sr320@washington.edu].[qDOD_Cgigas_GO_GOslim] cgslim

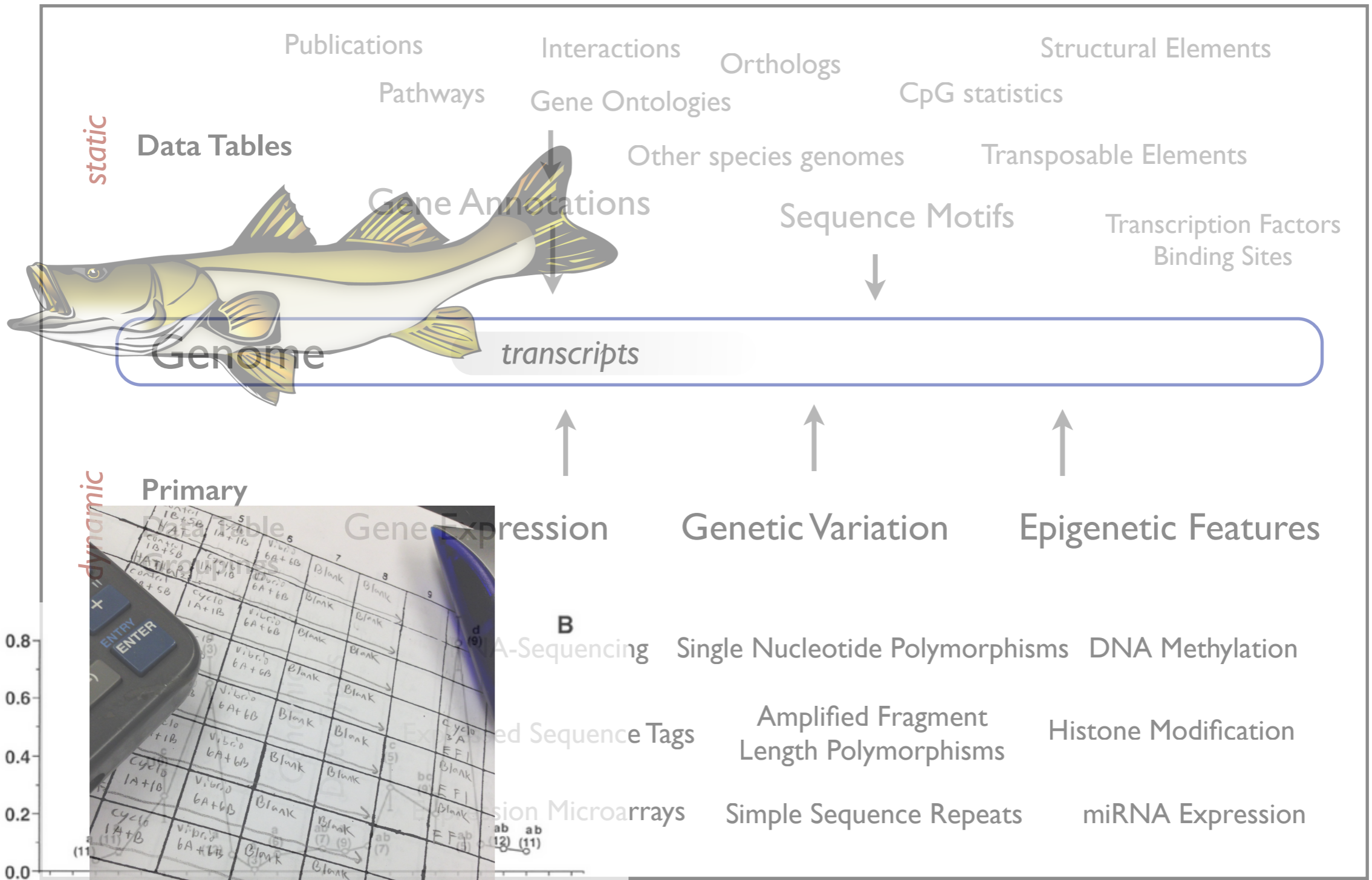
LEFT JOIN [sr320@washington.edu].[qDOD_Cgigas_gene_fasta] cgf
on cgslim.CGI_ID = cgf.CGI_ID

Where term LIKE '%methyl%'
OR
term LIKE '%histone%'
```

What genes are associated with epigenetics?

What genes are associated with immune response?

Which genetic markers are associated with
fast growth?
resilience?
disease tolerance?



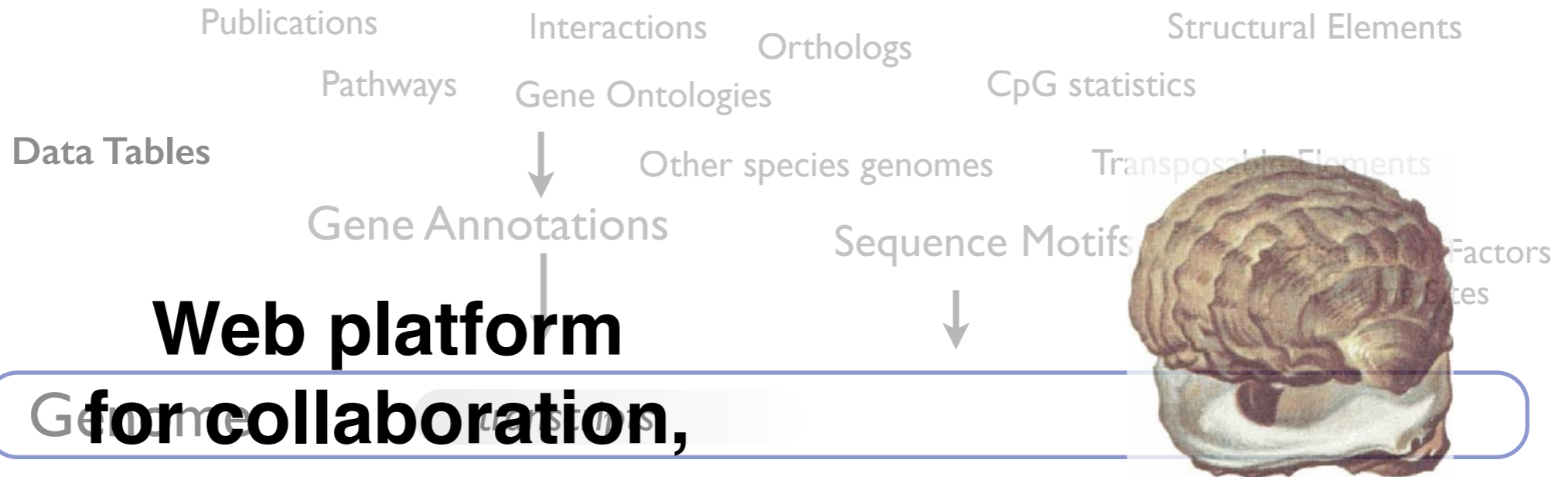
E₂ (ng/ml)

control	10+5B	5	6	7	8	9
control	10+5B	1A+1B	V. vibr.	6A+6B	Blank	Blank
control	6A+6B	Cyclo	6A+6B	Blank	Blank	Blank
control	1A+1B	V. vibr.	6A+6B	Blank	Blank	Blank
control	6A+6B	V. vibr.	6A+6B	Blank	Blank	Blank
control	1A+1B	V. vibr.	6A+6B	Blank	Blank	Blank
control	6A+6B	V. vibr.	6A+6B	Blank	Blank	Blank
control	1A+1B	V. vibr.	6A+6B	Blank	Blank	Blank
control	6A+6B	V. vibr.	6A+6B	Blank	Blank	Blank
control	1A+1B	V. vibr.	6A+6B	Blank	Blank	Blank
control	6A+6B	V. vibr.	6A+6B	Blank	Blank	Blank

1x
0.1M
0.1M
5.0

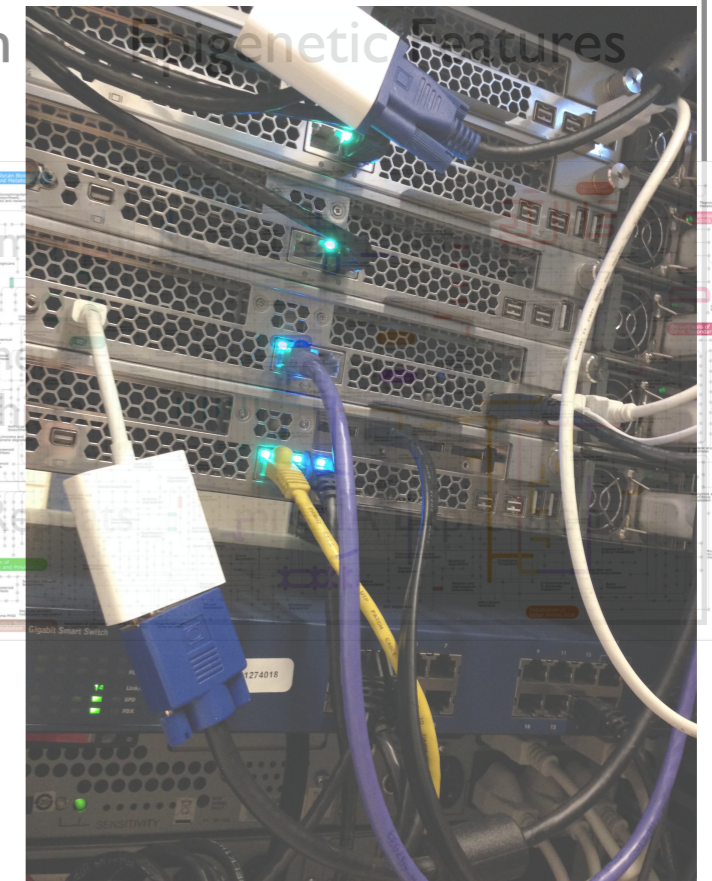
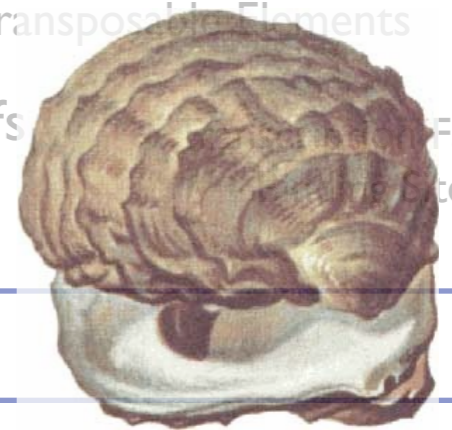
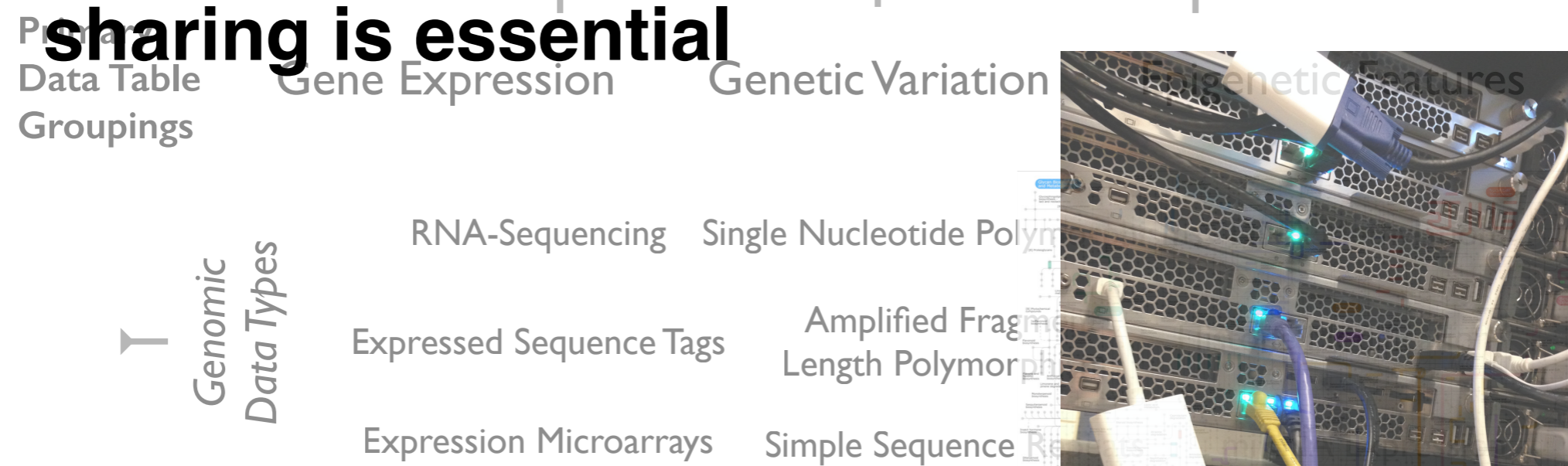
(E F I)
4x
40M

static



Web platform for collaboration, analyses, and sharing is essential

dynamic



Open Methods

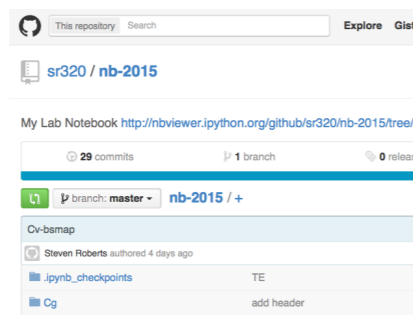
Open Methods

reproducibility

Data Acquisition and Analysis



Open Notebook Science



<http://genefish.wikispaces.com/>

Open Notebook Science

... there is a URL to a laboratory notebook that is freely available and indexed on common search engines. It does not necessarily have to look like a paper notebook but it is essential that all of the information available to the researchers to make their conclusions is equally available to the rest of the world.

—Jean-Claude Bradley

How

open notebooks

-wiki

-evernote

-IPython

How
open notebooks
-wiki
-evernote
-IPython

The screenshot shows a web browser window displaying the homepage of the Roberts Lab Wiki. The browser's address bar shows the URL genefish.wikispaces.com. The page features a green header with the Roberts Lab Wiki logo. A navigation menu on the left includes links for Wiki Home, Recent Changes, Pages and Files, and Members, along with a search box. The main content area is titled "home" and contains a paragraph explaining the wiki's purpose as a resource for lab personnel and students. Below this, there are three main sections: "Laboratory Reference Material" with a list of links to protocols, manuals, and inventories; "Lab Activity and Communication" with links to meetings, project ideas, and a calendar; and "Data Repositories" with a link to "The Eagle". On the right side, there is a smaller version of the Roberts Lab website, which includes a banner about lab notebooks and data online, and sections for Research, Outreach, and News and Notes.

← → ↻ 🏠 genefish.wikispaces.com

wiki my pinboard pinboard Compose Mail cnidarian Save to Mendeley Read Later Google URL Shortener >> Other Book

You are not a member of this wiki. [Join now](#) [Dismiss](#)

Roberts Lab Wiki

- Wiki Home
- Recent Changes
- Pages and Files
- Members

OPEN Notebook Science

Lab Notebooks

- [Sam's Notebook](#)
- [Mac's Notebook](#)
- [Emma's Notebook](#)
- [Claire's Notebook](#)
- [Steven's Notebook](#)
- [Halley's Notebook](#)
- [Katie's Notebook](#)
- [Brent's Notebook](#)
- [Doug's Notebook](#)
- [Charles' Notebook](#)
- [Jessica's Notebook](#)
- [Ahmed's Notebook](#)
- [Etillet's Notebook](#)
- [Hannah's Notebook](#)

home

[Edit](#) [0](#) [218](#) ..

This wiki has been developed as a resource for lab personnel and students to access information and publish research activities using an [open notebook science](#) based system. All lab notebooks can be accessed via the side menu. The Roberts Lab is in the [School of Aquatic and Fishery Sciences](#) within the [College of Environment](#) at the [University of Washington](#). More information can be found concerning [research](#), [personnel](#), and [outreach](#) on the [Roberts Lab Official Webpage](#).

Laboratory Reference Material

- [Laboratory Protocols](#)
- [How-to](#)
- [Emergency Contact Information](#)
- [UW Lab Safety Manual](#)
- [UW Biosafety Manual](#)
- [Chemical Inventory](#)

Lab Activity and Communication

- [Lab Meetings](#)
- [SPIAOT: Small Project Ideas and other tasks](#)
- [IPUS: Information for Prospective Undergraduate Students](#)
- [Lab Calendar](#)

Data Repositories

- [The Eagle](#)

Roberts Lab [People](#) [Research](#) [Products](#) [Outreach](#) [Courses](#) [Contact](#)

Lab Notebooks
Why do you put your lab notebook and data online?
The usual follow up to that question is "What if someone copies your research? I will not disagree this is a possibility. Others are concerned with intellectual property and I respect persons have their..."

Research
Research in our lab focuses on characterizing physiological responses of marine organisms to environmental change. Using integrative approaches we strive to examine impacts and adaptive potential from the nucleotide to organismal level. A core component of this includes investigating the functional relationship of genetics, epigenetics, and transcription.

Outreach
We are continually looking for new ways to engage and share with students, scientists, and the general public. Everyone can follow us on tumblr, facebook, youtube, twitter, wikispaces, flickr, and figure. Web-based resources we have developed include an online version of the Cotton Shellfish Collection, crabsatrome, Research Notes on Ocean Acidification and oyster genes.

News and Notes
-Congrats to Mackenzie Gavery, who won a Student Spotlight Award at Aquaculture 2013.
- "Papers" page rebranded as "Products".
-Claire Ellis launches her website: check it out (and is now on twitter!)
-Archive

Labcam **Connect** **Access** **Funding**

How open notebooks -wiki -evernote -IPython

← → ↻ 🏠 [genefish.wikispaces.com](#)

wiki my pinboard pinboard Compose Mail cnidarian Save to Mendeley Read Later Google URL Shortener >> Other Bookmarks

You are not a member of this wiki. [Join now](#) [Dismiss](#)

Roberts Lab Wiki

Wiki Home
Recent Changes
Pages and Files
Members
Search

OPEN Notebook Science

Lab Notebooks

Sam's Notebook
Mac's Notebook
Emma's Notebook
Claire's Notebook

home

Edit 0 218 ..

This wiki has been developed as a resource for lab personnel and students to access information and publish research activities using an [open notebook science](#) based system. All lab notebooks can be accessed via the side menu. The Roberts Lab is in the [School of Aquatic and Fishery Sciences](#) within the [College of Environment](#) at the [University of Washington](#). More information can be found concerning [research](#), [personnel](#), and [outreach](#) on the [Roberts Lab Official Webpage](#).

Laboratory Reference Material

- [Laboratory Protocols](#)
- [How-to](#)
- [Emergency Contact Information](#)
- [UW Lab Safety Manual](#)
- [UW Biosafety Manual](#)
- [Chemical Inventory](#)

Lab Activity and Communication

- [Lab Meetings](#)
- [SPIAOT: Small Project Ideas and other tasks](#)
- [IPUS: Information for Prospective Undergraduate Students](#)
- [Lab Calendar](#)

Data Repositories

- [The Eagle](#)

Roberts Lab

People Research Products Outreach Courses Contact

Lab Notebooks

Why do you put your lab notebook and data online?
The usual follow up to that question is "What if someone copies your research? I will not disagree this is a possibility. Others are concerned with intellectual property and I respect persons have their..."

Research: Research in our lab focuses on characterizing physiological responses of marine organisms to environmental change. Using integrative approaches we strive to examine impacts and adaptive potential from the nucleotide to organismal level. A core component of this includes investigating the functional relationship of genetics, epigenetics, and transcription.

Outreach: We are continually looking for new ways to engage and share with students, scientists, and the general public. Everyone can follow us on tumblr, facebook, youtube, twitter, wikispaces, flickr, and figure. Web-based resources we have developed include an online version of the Cotton Shellfish Collection, crassostreome, Research Notes on Ocean Acidification and oyster genes.

News and Notes: -Congrats to Mackenzie Gavery, who won a Student Spotlight Award at Aquaculture 2013. -"Papers" page rebranded as "Products". -Claire Ellis launches her website- check it out (and is now on twitter!) -Archive

Labcam Connect Access Funding

Labcam @ur320 206.685.3742

Wiki Home
Recent Changes
Pages and Files
Members
Search

OPEN Notebook Science

Lab Notebooks

Sam's Notebook
Mac's Notebook
Emma's Notebook
Claire's Notebook
Steven's Notebook
Halley's Notebook
Katie's Notebook
Brent's Notebook
Doug's Notebook
Charles' Notebook
Jessica's Notebook
Ahmed's Notebook
Etillet's Notebook
Hannah's Notebook

Featured Pages
[crassostreome](#)

More Pages

Sam's Notebook

20130729
DNA Isolation - Claire's C.gigas Development Isolated

20130409
PCR - Hexokinase Partial CDS
Performed PCR using the primers CG_HK_CDS_2132-2158 (SRID: 1521) and Cg_HK_CD: C.gigas cDNA (from DATE).

Master mix calcs and cycling params are here.

Samples were run in duplicate.

Results:

How open notebooks

- wiki
- evernote
- IPython

SR320

SHORTCUTS

- fu
- PAG
- OA

RECENT NOTES

- Mussel RNAseq
- MyFinancial.desk...
- TRAVEL SUPPOR...

Notes

Notebooks

Tags

Atlas

Trunk

All Notes mgavery's no...

AUGUST 2013 3

MethylKit w/ Final BiGO and BiGill files
8/30/13 modify Methratio outputs to format needed for methylKit in SQLShare starting...

MGAVERY MGAVERY'S NOTEBOOK

promoter methylation by...
8/27/13 generated proportion methylation for promoter (1kb upstream) using result of int...

EE2v2 sampling and stats
8/26/13 Day 60/61 sampling (3/25/3/26) 227 oysters were sampled (113 treated and 11...

JULY 2013 5

IntersectBed for genomic...
7/30/13 canonical genome files: http://nbviewer.ipython.org/urls/ra...

Generating TE canonical GFF from Re...
7/29/13 The starting file for this is the output of RepeatProteinMask performed by SR (look towards the bottom of this entry): htt...

mgavery's notebook click to add tags

Created: Aug 30, 2013 Updated: Aug 30, 2013

You are viewing a note that is shared with 2 people

MethylKit w/ Final BiGO and BiGill files

modify Methratio outputs to format needed for methylKit in SQLShare

starting file for Gill: BiGill_methratio_v9_A
starting file for Gonad: BiGO_betty_plain_methratio_v1

Used 3 steps in SQLshare - here is workflow for gill:

step1

```
SELECT
chr as chr,
pos as start,
'+' as strand,
cast (CT_count as float) as CT_count,
cast (C_count as float) as C_count
FROM [sr320@washington.edu].[BiGill_methratio_v9_A.txt]
where
context like '__CG_'
and
CT_Count >= 5
and
ratio <> 'NA'
```

step2

```
SELECT
chr as chr,
start as start,
strand as strand,
CT_count as CT_count,
C_count as C_count,
C_count/CT_count as freqC
FROM [mgavery@washington.edu].[test]
```

step3

```
SELECT
chr as chr,
start as start,
strand as strand,
CT_count as CT_count,
freqC as freqC,
1-freqC as freqT
FROM [mgavery@washington.edu].[test2]
```

name of step3 table: [mgavery@washington.edu].[BiGill_v9_A_MethylKit]

EVERNOTE

sr320 has shared a notebook with you.

Join Notebook

Lab Notebook

Transition to IPython
8/2/13 As of July 2013 I have transitioned my lab notebook to IPython. My New Lab Notebook can be found at http://sr320.github.io/ipython_nb/

BiGill - combined notebook pages
5/21/13 Running on Hummingbird d-128-95-149-219:bsmap-2.74 sr320\$./bsmap -w 1000 -a /Volumes/NGS/Drive/NGS/ Raw

Bismark: BiGill
5/17/13 ./bismark -n 1 -l50 ./genome/ /Volumes/Bay3/Software/bismark_v0.7.12/filtered_174gm_A_ACTGATA_L002_R1.fastq --path_to_bowtie

Bismark: BiGo
5/17/13 Running Genome Prep on genfish, and greenbird Genome prep complete will try robertsmac:bismark_v0.7.12 sr320\$./bismark -n 1 -l50 ./genome/ -1

Betty BiGO
5/16/13 Summary of Successful flow in bold- ./bsmap -a /Volumes/betty/filtered_174gm_A_NoIndex_L006_R1.fastq.gz -b /Volumes/betty/filtered_174gm_A_NoIndex_L006_R2.fastq.gz -d

Fixing MG Bedtools
5/15/13 {"metadata": {"name": "MG_bedtools"}, "nbformat": 3, "nbformat_minor": 0, "worksheets": [{"cells": [{"cell_type": "raw", "metadata": {}, "source": "MG: I would like to know how many"}]}]}

BiGO - Running Gill methratio through Galaxy
5/9/13 https://main.g2.bx.psu.edu/u/sroberts320/w/methratio-processing-imported-from-uploaded-file

Lifting Coordinates within genome..
5/7/13 The Ensembl API. The final example I described above (converting between coordinate systems within a single genome assembly) can be accomplished with the Ensembl core API. Many

QPX Revisions
5/3/13 ALL Files

QPX - Making a Genome Browser
5/3/13 Files for IGV
http://bio533.wikispaces.com/QPX+Genome+Annotation Min length 10,000

Created: May 21, 2013 Modified: Aug 01, 2013

BSMAP complete on Hummingbird
Total number of aligned reads: 2545683 (1.8%) Done. P...

python methratio.py -d /Volumes/web/cnidarian/TJGR_RepProMask... /Volumes/web/cnidarian/BiGill_methratio_TOnly_A.txt -s /Volumes/B... /Volumes/web/cnidarian/BiGill_BSMAP_TOnly_v2.sam

genome files:
| exon: http://sh.washington.edu/bivalvia/enomefiles_MBDbsSeq_gill/gffs/oy...

Link to data: http://oogole.com/url?q=http2F%2Feagle.fish.was...

Having look at raw output:

C10013	77	+	GTCC	0.000	1.00	0	1
C10013	78	+	TCCC	0.000	1.00	0	1
C10013	79	+	CCCT	0.000	1.00	0	1
C10013	82	-	TTGA	0.000	2.00	0	2
C10013	85	+	ATCC	0.000	1.00	0	1
C10013	86	+	TCCC	0.000	1.00	0	1
C10013	87	+	CCCT	0.000	1.00	0	1
C10013	93	+	AACT	0.000	1.00	0	1
C10013	97	+	GGGT	0.000	1.00	0	1
C10013	100	+	TTCA	0.000	1.00	0	1
C10013	103	-	ATCA	NA	0.00	0	1
C10013	111	-	ATGT	0.000	1.00	0	1
C10013	113	-	GTGA	0.000	1.00	0	1
C10013	116	-	TAGT	0.000	2.00	0	2
C10013	120	+	TTCT	0.000	1.00	0	1

How
open notebooks
-wiki
-evernote
-IPython*

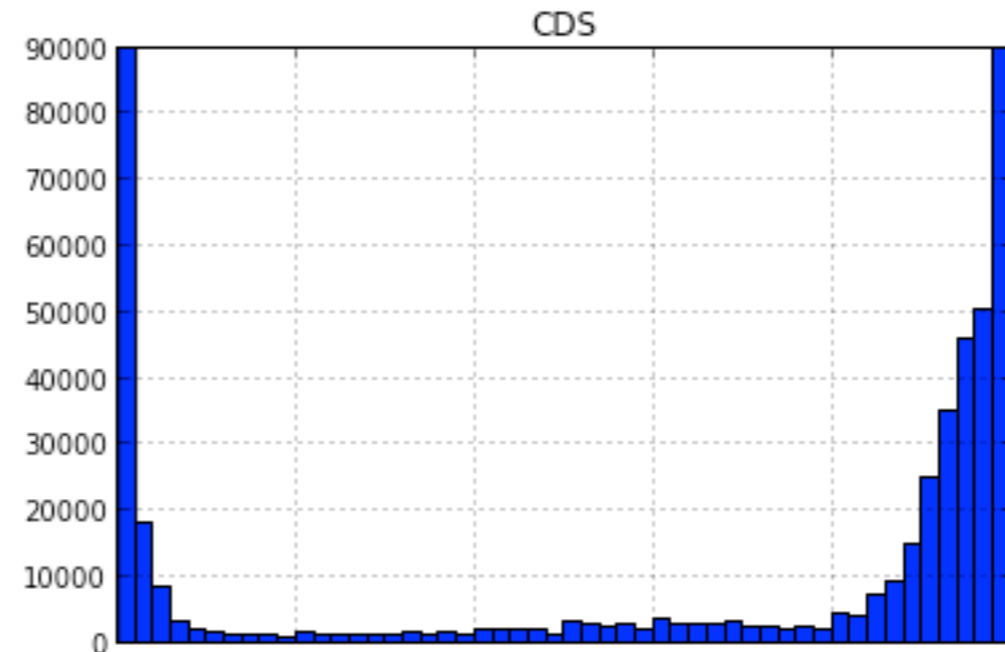
DNA methylation of Oyster Sperm based on

methratio file in SQLShare <https://sqlshare.escience.washington.edu/sqlshare#s=query/sr320>

In [38]: `from pandas import *`

```
# read data from data file into a pandas DataFrame
CDSmr = read_table("/Volumes/web/cnidarian/BiGo_methratio_boop_...
                  #sep="\t", # what character separates each column?
                  #na_values=["", " "], # what values should be consi
                  header=None)
```

```
CDSmr[5].hist(bins=50);
#Axis limits are changed using the axis([xmin, xmax, ymin, ymax]
plt.axis([0, 1, 0, 90000]);
plt.title('CDS');
```



In [44]: `!sed 's/mRNA/promoter/g' </Volumes/web/cnidarian/TJGR_Promoter_1k5p.gff> /Volumes/web`

In [45]: `!head /Volumes/web/cnidarian/TJGR_Promoter_1k5p_b.gff`

C16582	GLEAN	promoter	386	395	0.555898	-	.	ID=C
C17212	GLEAN	promoter	1	30	0.999572	+	.	ID=C
C17316	GLEAN	promoter	1	29	0.555898	+	.	ID=C
C17476	GLEAN	promoter	258	491	0.998947	-	.	ID=C
C17998	GLEAN	promoter	388	559	1	-	.	ID=CGI_10000
C18346	GLEAN	promoter	1	173	1	+	.	ID=CGI_10000
C18428	GLEAN	promoter	547	611	0.555898	-	.	ID=C
C18964	GLEAN	promoter	659	714	0.999572	-	.	ID=C

In []: `#join in SQLShare`

In [45]: `!python /Users/sr320/sqlshare-pythonclient/tools/singleupload.py -d OA_enrich2 /`
 processing chunk line 0 to 51 (0.00206899642944 s elapsed)
 pushing /Volumes/web/cnidarian/mod_chart_B1049AF0BD891379525818063.txt...
 parsing C8BE2D77...
 finished OA_enrich2

In [46]: `!python /Users/sr320/sqlshare-pythonclient/tools/fetchdata.py -s "SELECT * FROM`

In [47]: `!head /Volumes/web/cnidarian/OA_enrich2_join_SPID.csv`

```
Category,Term,Count,%,PValue,Genes,List Total,Pop Hits,Pop Total,Fold Enrichment
,value,Description
GOTERM_BP_FAT,GO:0045449-regulation of transcription,8,8.98876404494382,0.092168
7272727,1,0.998517222392313,76.9925811614,CGI_10000058,Q2KJJ0,6E-16,Pre-mRNA-pro
GOTERM_BP_FAT,GO:0045216-cell-cell junction organization,4,4.49438202247191,0.00
```



Fasta2Slim

This IPython notebook is intended to serve as a structured means to annotate sequences using UniProt/SwissProt database. The notebook can be easily modified to personal preferences. As developed, the notebook requires the user has the following software installed ...

- IPython
- NCBI Blast
- SQLShare Python Client

Instructions for use.

In a working directory of your choosing place query fasta file, naming as `query.fa`. Edit the cell below, providing the path to said working directory.

Identify the location of the blast database you would like to use and indicate path in the cell below.

Identify the location of your `sqlshare-pythonclient/tools` and indicate path in the cell below.

Change the input to the `usr` variable to reflect your SQLShare user account.

```
In [2]: #Location Variables
wd="~/Desktop/test/"

db="/Volumes/Bay3/Software/ncbi-blast-2.2.29\+/db/uniprot_sprot_r2013_12"

sqls="~/sqlshare-pythonclient/tools/"

usr="sr320@washington.edu"
```

```
In [254]: !head {wd}query.fa

>PiuraChilensis_v1_contig_1
ATTTACAATACGAAGTAAAATAGATAACGTGAAAATAATCTTGGTGCTGGATGATCGATC
AAGTTCACCAATATTTTATTGTAATAAATCATTCTAAACAGCATGAAATCGTGTACAATG
TATAAACAAGCAAATATATAACACTAAAGCAAGAGGGCGTAAGTGGGGGGGTGGGTGAGA
GTAAAAAATTCAAACATGTCAAATACCCCGGCGTTAGCCTTAAAAGCACCATGGACTTCT
CGCTTCATTAAGCATAAATTAAGAGCGTAAATAGAGCAATTCATTAAGCATTAAGAGCA
```

File Edit View Insert Cell Kernel Help



Cell Toolbar: None

Fasta2Slim

This IPython notebook is structured means to annotate sequences using UniProt/SwissProt. As per personal preferences. As

- IPython
- NCBI Blast
- SQLShare Python Client

structured means to annotate sequences using UniProt/SwissProt. As per personal preferences. As requires the user has the following software installed ...

Run
Run and Select Below
Run and Insert Below
Run All
Run All Above
Run All Below
Cell Type ▶
Current Output ▶
All Output ▶

Instructions for use.

In a working directory of your choosing place query fasta file, naming as `query.fa`. Edit the cell below, providing the

Identify the location of the blast database you would like to use and indicate path in the cell below.

Identify the location of your `sqlshare-pythonclient/tools` and indicate path in the cell below.

Change the input to the `usr` variable to reflect your SQLShare user account.

```
In [2]: #Location Variables
wd=~/Desktop/test/

db="/Volumes/Bay3/Software/ncbi-blast-2.2.29\+/db/uniprot_sprot_r2013_12"

sqls=~/sqlshare-pythonclient/tools/

usr="sr320@washington.edu"
```

```
In [254]: !head {wd}query.fa
```



```
In [2]: !head {wd}query.fa
```

```
>PiuraChilensis_v1_contig_1
ATTTACAATACGAAGTAAAATAGATAACGTGAAAATAATCTTGGTGCTGGATGATCGATC
AAGTTCACCAATATTTTATTGTAAAAAATCATTCTAAACAGCATGAAATCGTGTACAATG
TATAACAAGCAAATATATAACACTAAAGCAAGAGGGCGTAAGTGGGGGGGTGGGTGAGA
GTAAAAAATTCAAACATGTCAAATACCCCGGCGTTAGCCTTAAAAGCACCATGGACTTCT
GCCTTCAATAAGCATAAAATTAAAACACCTAATACACAATGAATATACAGATAAAACAGA
TTTATGAATAGTTGGTGTTACATCTTTTACAGCCATAAGCCTTCATTTTGCTTCCAAACG
TATAAAATCTGACTTGGACAATATACAGCCATGAGATATGACACAGCGAGCACTACAAT
ATATATTTATCTTGTACTATACAGCCTGTACAAGAAAATTCTGGAATTGTCTTCACAAGA
GACAGAAAATAGTTGCAATGTGAATGCTAGTCTACTATTTGATCACAATTGGATAGAAA
```

```
In [3]: #number of sequences
!fgrep -c ">" {wd}query.fa
```

```
282
```

Blast

```
In [4]: !blastx \
-query {wd}query.fa \
-db {db} \
-max_target_seqs 1 \
-max_hsps 1 \
-outfmt 6 \
-evalue 1E-05 \
-num_threads 2 \
-out {wd}blast_sprot.tab
```

Number of matched sequences:

```
In [5]: !wc -l {wd}blast_sprot.tab
```

```
211 /Users/sr320/Desktop/test/blast_sprot.tab
```

```
In [6]: !tr '|' '\t' <{wd}blast_sprot.tab> {wd}blast_sprot_sql.tab
!head -1 {wd}blast_sprot.tab
!echo SQLShare ready version has Pipes converted to Tabs ....
!head -1 {wd}blast_sprot_sql.tab
```

```
PiuraChilensis_v1_contig_3      sp|Q6P9A1|ZN530_HUMAN  33.33  105  61
3          825      1118      414      516      1e-07  57.4
SQLShare ready version has Pipes converted to Tabs ....
PiuraChilensis_v1_contig_3      sp          Q6P9A1  ZN530_HUMAN      33.33  105
61          3          825      1118      414      516      1e-07  57.4
```

Joining in SQL Share

```
In [7]: !python {sqls}singleupload.py \
-d _blast_sprot \
{wd}blast_sprot_sql.tab
```

```
processing chunk line 0 to 211 (0.000264167785645 s elapsed)
pushing /Users/sr320/Desktop/test/blast_sprot_sql.tab...
parsing 983DD315...
finished _blast_sprot
```

```
In [8]: !python {sqls}fetchdata.py \
-s "SELECT Column1, term, GOSlim_bin, aspect, ProteinName FROM [{usr}].[_b
last_sprot]md left join [samwhite@washington.edu].[UniprotProtNamesReviewe
d_yes20130610]sp on md.Column3=sp.SPID left join [sr320@washington.edu].[S
PID and GO Numbers]go on md.Column3=go.SPID left join [sr320@washington.ed
u].[GO_to_GOslim]slim on go.GOID=slim.GO_id where aspect like 'P'" \
-f tsv \
-o {wd}Godescriptions.txt
```

Plot GoSlim terms

```
In [10]: pylab inline
```

Populating the interactive namespace from numpy and matplotlib

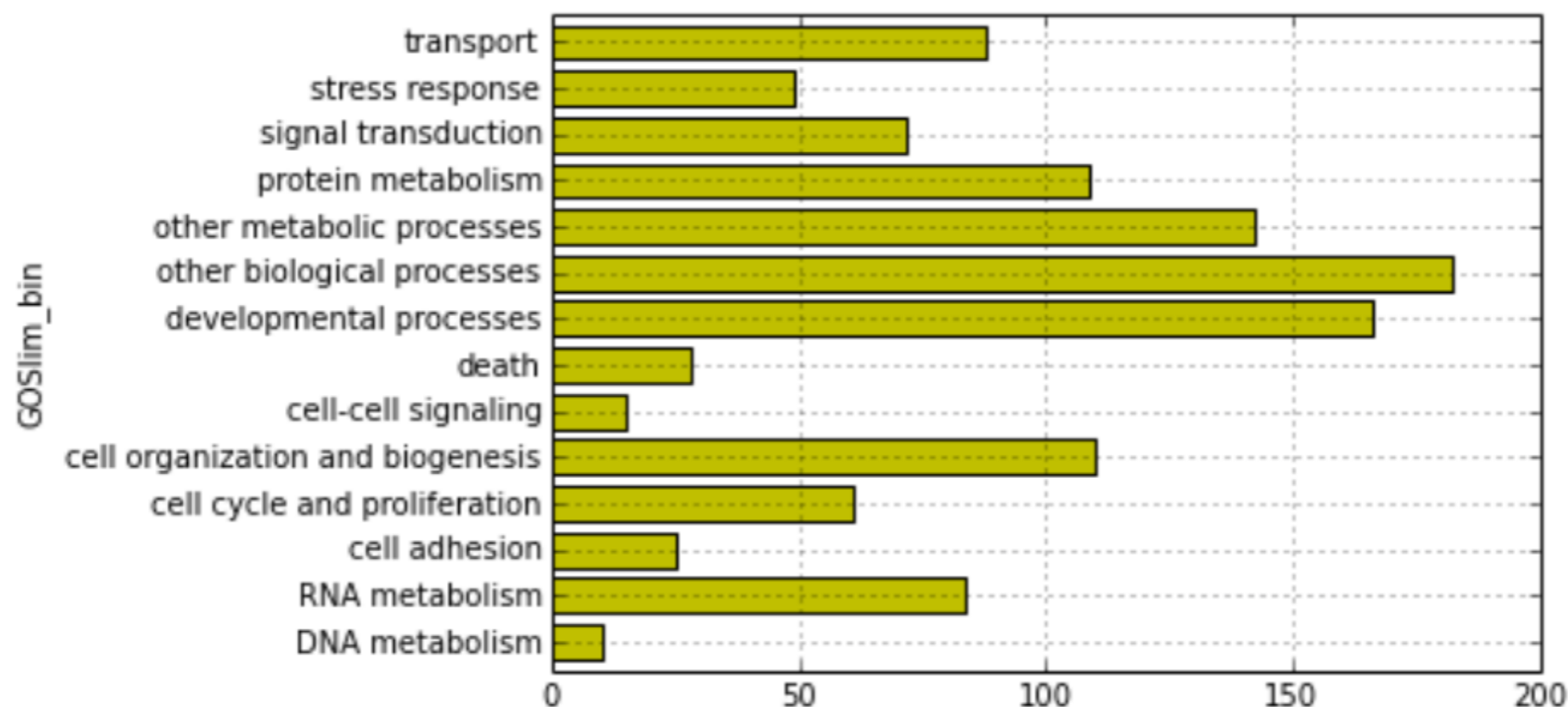
```
In [11]: cd {wd}
```

```
/Users/sr320/Desktop/test
```

```
In [12]: from pandas import *
```

```
gs = read_table('G0descriptions.txt')
```

```
In [13]: gs.groupby('GOSlim_bin').Column1.count().plot(kind='barh', color=list('y'))  
savefig('GOSlim.png', bbox_inches='tight')
```



Plot GoSlim terms

```
In [10]: pylab inline
```

Populating the interactive namespace from numpy and matplotlib

```
In [11]: cd {wd}
```

```
/Users/sr320/Desktop/test
```



This repository Search



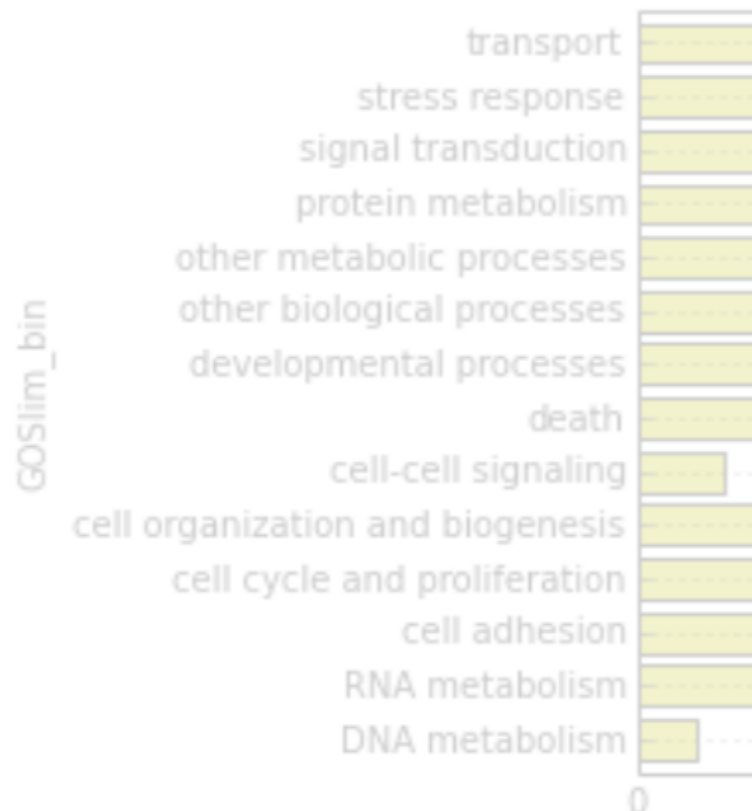
che625 / olson-ms-nb

```
In [12]: from pandas import *
```

```
gs = read_table('GOdescriptions.txt')
```

```
In [13]: gs.groupby('GOSlim_bin').Column1.count().plot(kind='barh',  
)  
savefig('GOSlim.png', bbox_inches='tight')
```

github
SOCIAL CODING



**IPython Notebook for downloading and analyzing data
the manuscript: "Indication of family-specific DNA
methylation patterns in developing oysters"**

bioRxiv preprint - <http://dx.doi.org/10.1101/012831>

To execute the IPython Notebook in its entirety you will need:

- IPython - [install instructions](#)
- BSMAP - [install instructions](#)
- bedtools - [install instructions](#)
- R - [install instructions](#)
- rpy2 (interface to R from Python) - [install instructions](#)

Plot GoSlim terms

```
In [10]: pylab inline
```

```
Populating the interactive namespace from numpy and matplotlib
```

```
In [11]: cd {wd}
```

```
/Users/sr320/Desktop/test
```

README.md

IPython Notebooks and data supplemental to the manuscript: "Up in arms: Immune and nervous system response to sea star wasting disease"

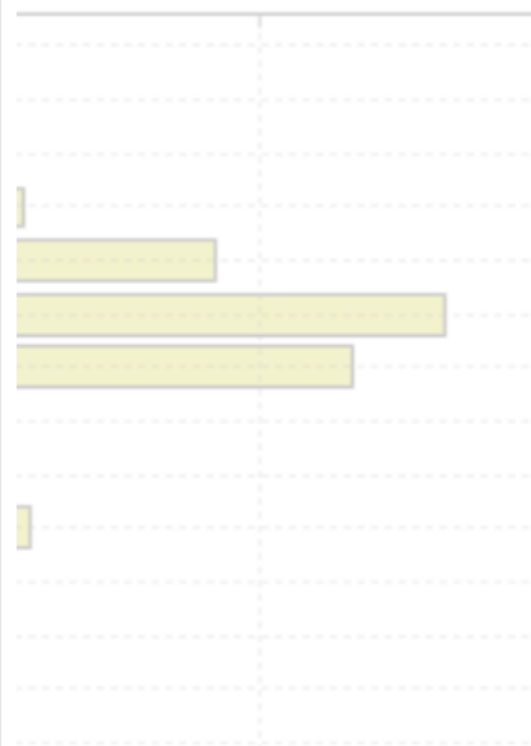
The repository includes IPython notebooks (.ipynb file) that can be downloaded locally and interactively executed. The code in the IPython notebook `eimd_analysis.ipynb` will process data such that figures in the manuscript are reproduced (in theory).

Description of Files

- `eimd_analysis.ipynb` - IPython notebook that can be interactively executed locally or viewed online designed so that user can **replicate all analysis**. Requires several dependencies (see below). [nbviewer version](#)
- `eimd_data-only.ipynb` - IPython notebook that can be interactively executed locally or viewed online designed so that user can **simply explore** data files. *Only requires IPython*. [nbviewer version](#)
- `data/Phel_transcriptome.fasta` - P hel coelocytes transcriptome. Contains xxxx contigs from de novo assembly.
- `data/Phel_countdata.txt` - Tab-delimited text file with read count data from 6 P hel RNA-seq libraries, 3 treated and 3 control libraries.

```
t(kind='barh',
```

github
SOCIAL CODING



github.com/sr320/eimd-sswd

Open Science Philosophy

Transparency with limited effort

will try just about anything

Biology

Environment

Molecular

Data Analysis

eScience

iPlant Galaxy

Notebooks

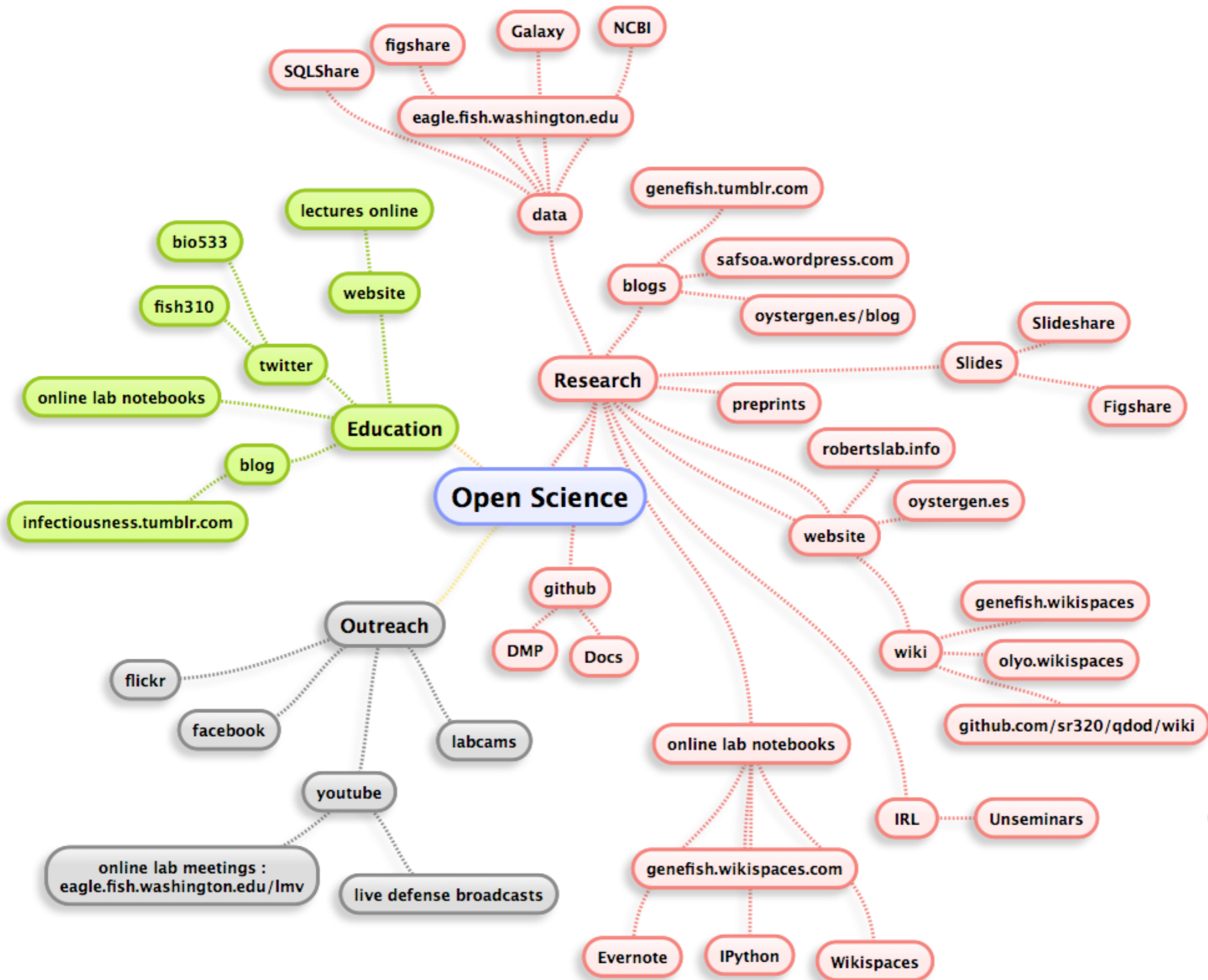
Rationale

Platforms

Open Science

Data

everything else...



Biology

Environment

Molecular

Data Analysis

eScience

iPlant Galaxy

Notebooks

Rationale

Platforms

Open Science

Data

everything else...



Male
20.026
Female

start them young



Acknowledgements

Mackenzie Gavery

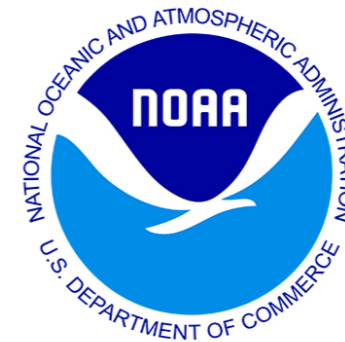
Claire Olson

Sam White

Brent Vadopalas

Jake Heare

Jay Dimond



Bill Howe

Dan Halperin



slides, data & more @
github.com/sr320/talk-osu-2015