Interphylum comparisons of genetic circuits in muscle tissue development

Steven Kuntz

March 25, 2005

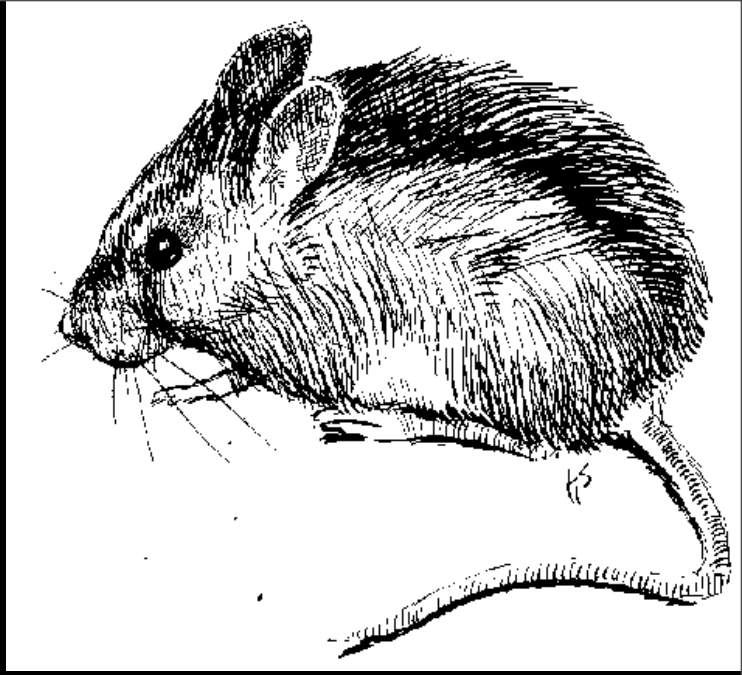
Genetic Circuits in Muscle Tissue

- Specialized tissues originated from a single ancestral tissue type
- Striated and pulsating muscles are an accessible set of specialized tissues in *C. elegans* and *M. musculus*
- Tissue specialization preceded phylum divergence
- Strong conservation of robust core regulators
- Core regulators tissue specific
- Divergence of peripheral factors and elements
- Dissectible through computational analysis with use of genomic, expression, and binding data

Model Organisms:

Ideal model organisms for this study must meet 6 criteria:

1. Reliably sequenced and annotated genome
2. Closely related species have sequenced genomes

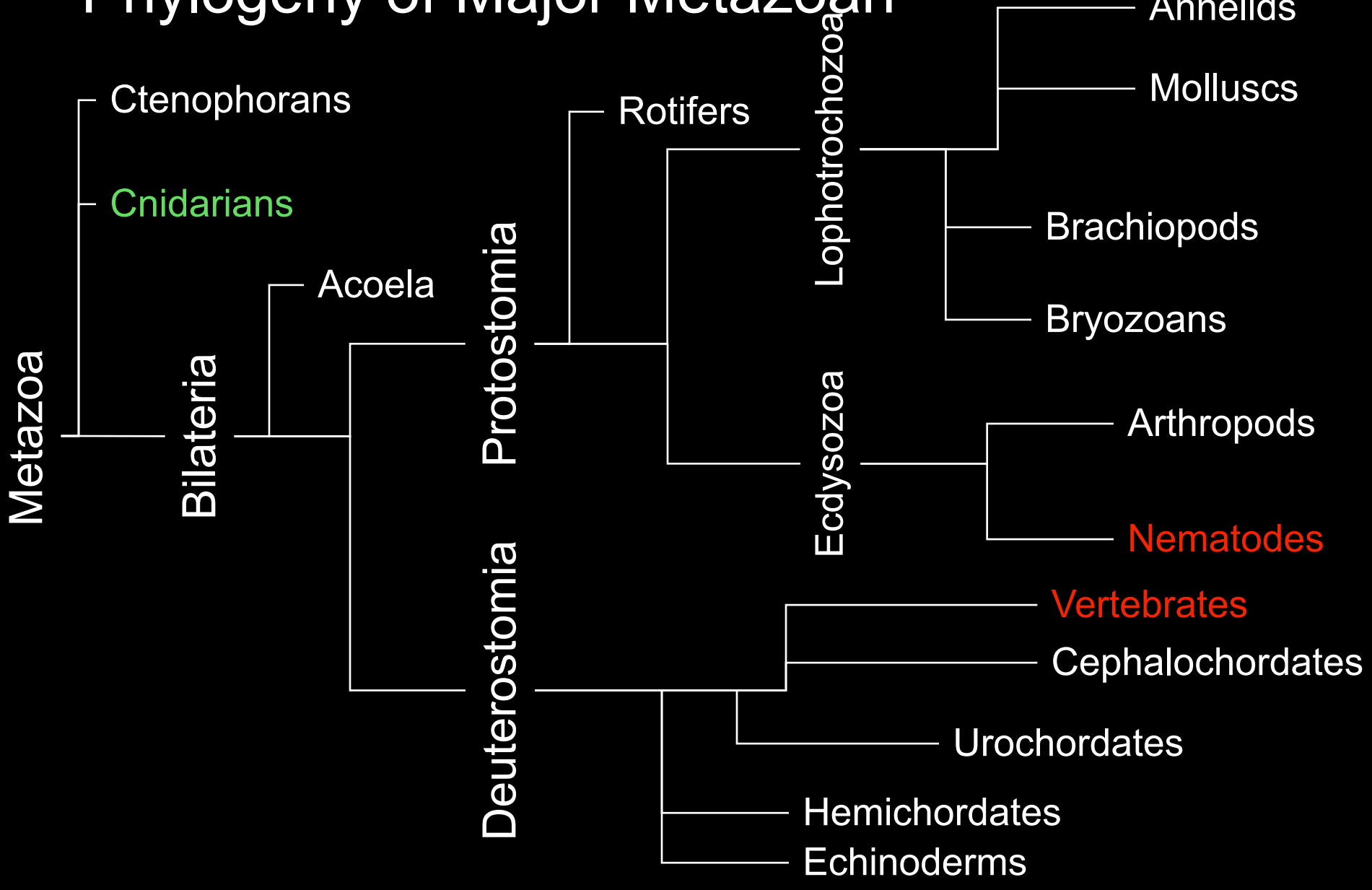


http://home.teleport.com/~hsimante/pictures/mouse.gif

http://www.desc.med.vu.nl/NL-taxi/ICE/C_elegans1.jpg

3. Tissue enrichment or culturing possible
4. Powerful epistatic models
5. Microarrays available
6. Potential for Chromatin IP

Phylogeny of Major Metazoan



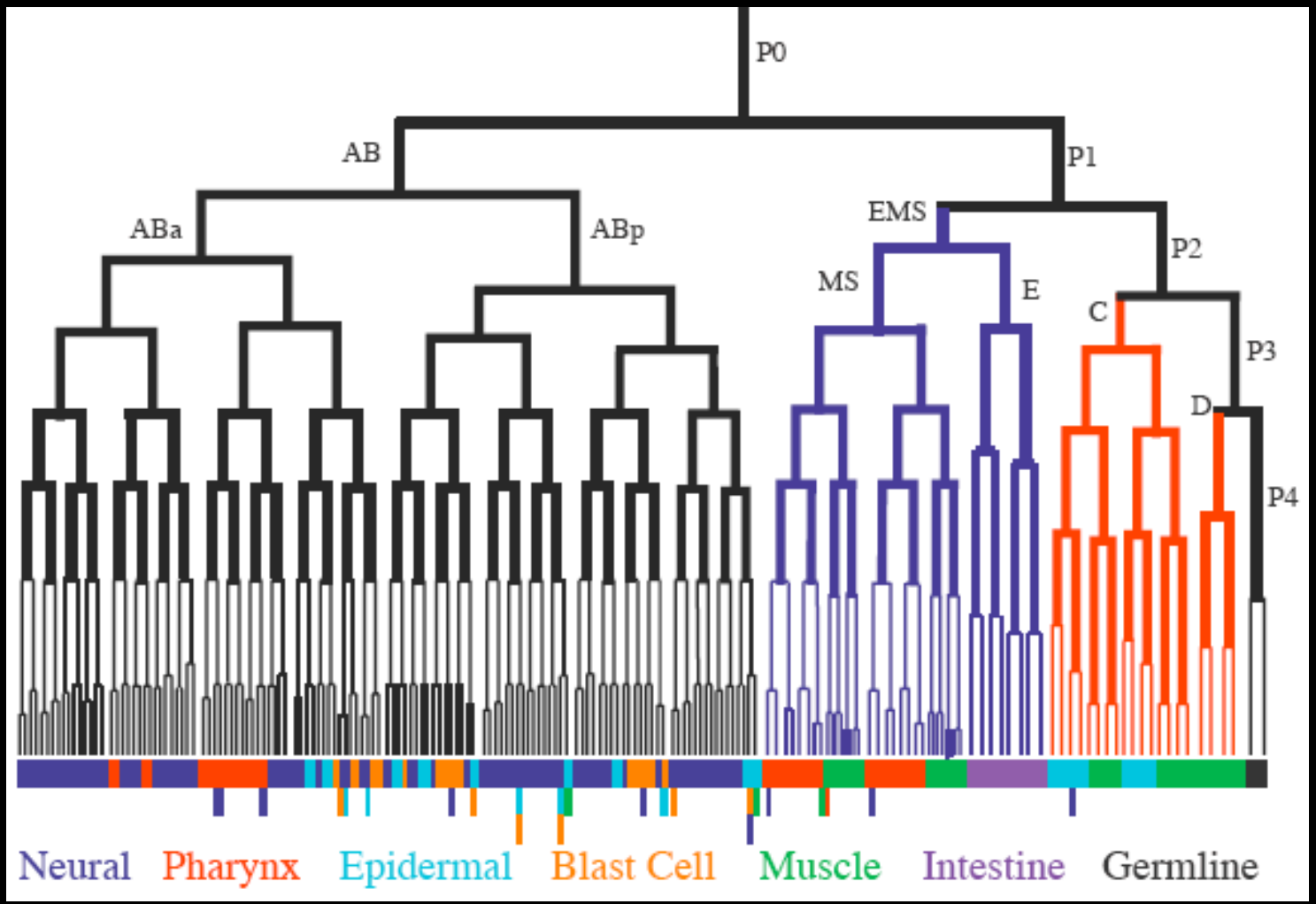
- Source: UCMP Berkeley Phylogeny Map

Background: Physiological Similarities of Muscle Tissue Types

- Cardiac vs. Pharyngeal
  - Rhythmic pulsation
  - Simple striations
  - Integration with exclusive nerve system
  - Homeodomain
- Skeletal vs. Body Wall
  - Striations
  - Linear patterning
  - bHLH, MADS-box

Activation of gene circuits

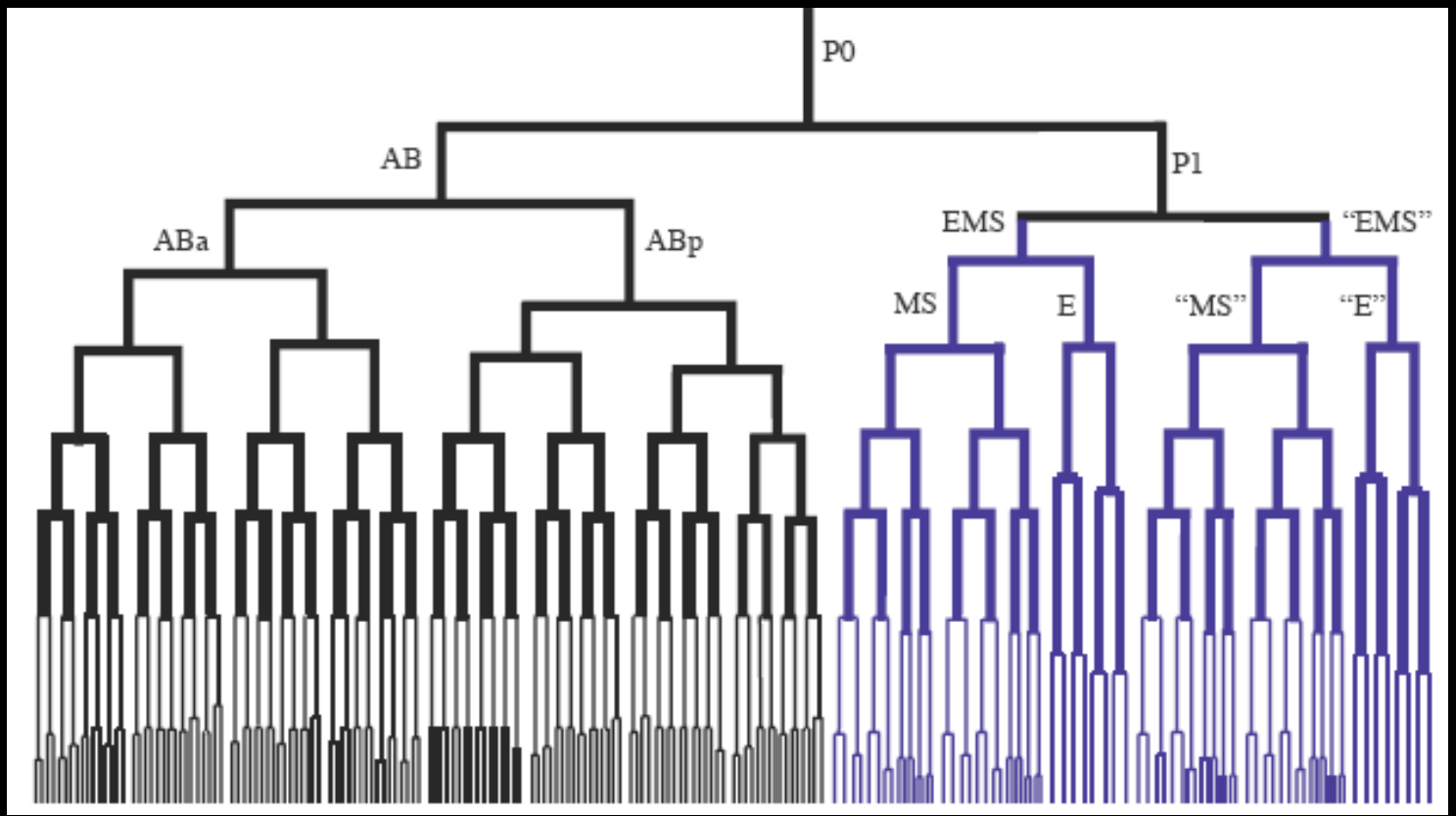
- Different lineages sometimes have the same phenotypic end point



Baugh et al. *Development* 2003.

Activation of gene circuits

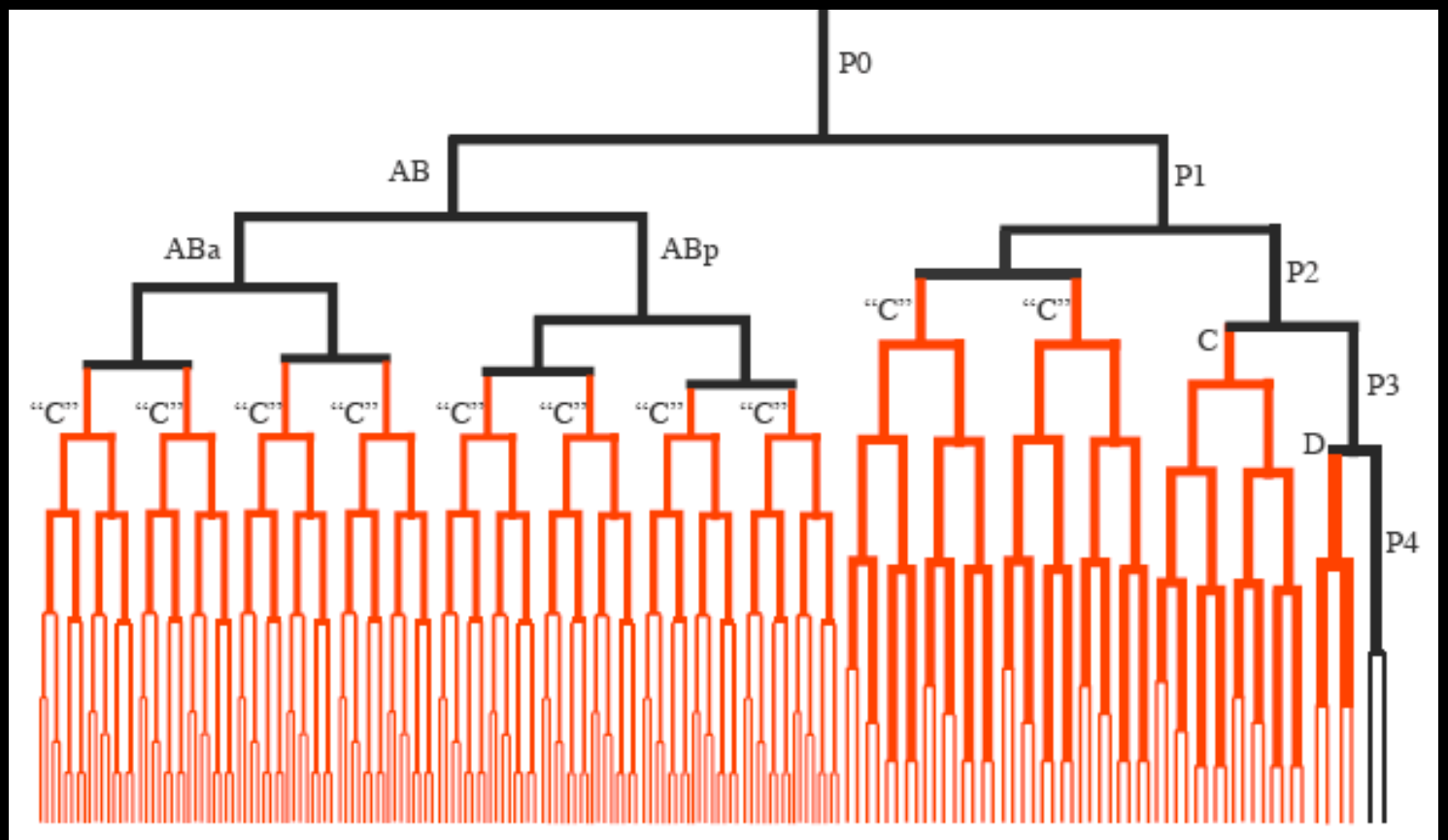
- Different lineages sometimes have the same phenotypic end point



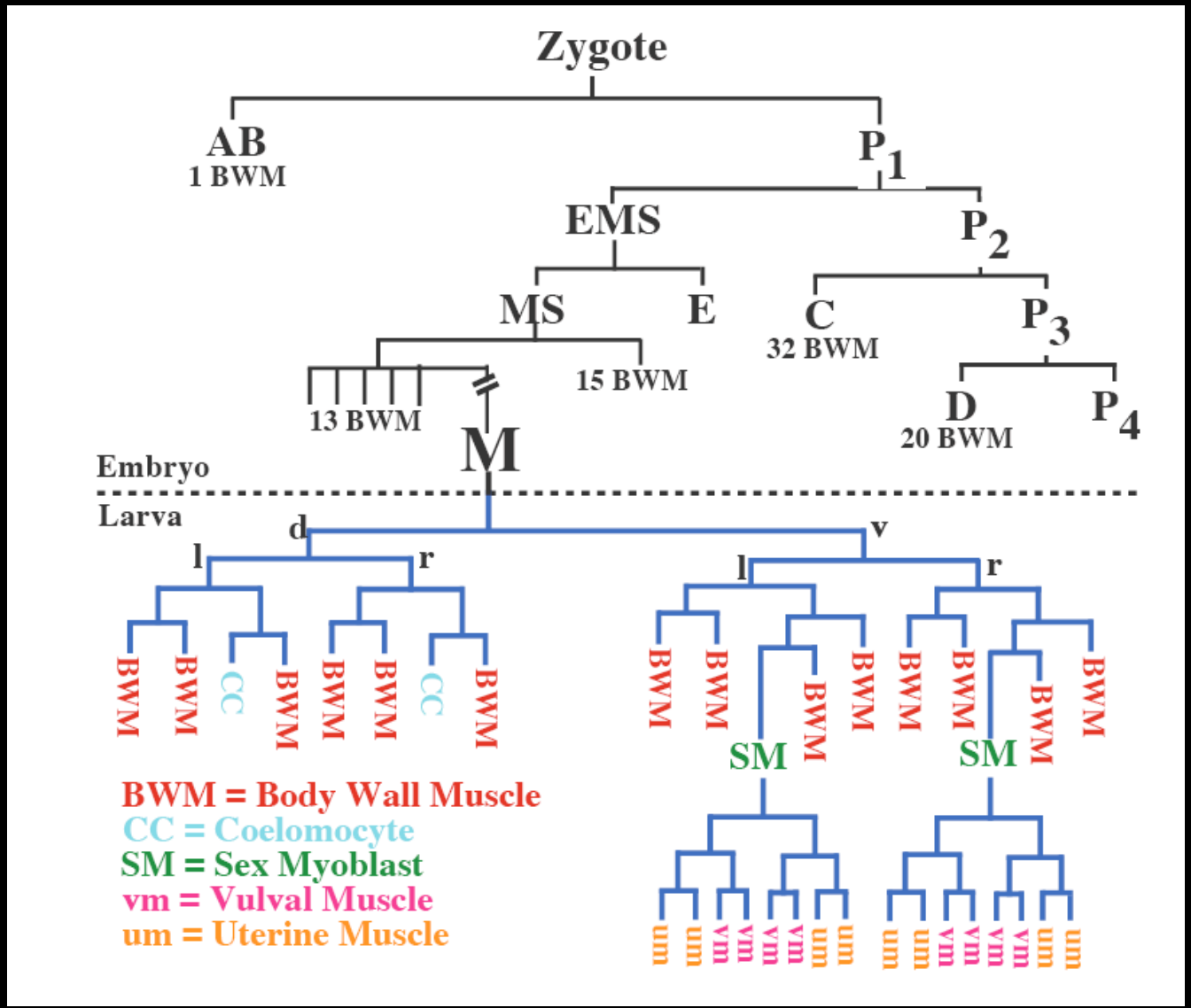
Baugh et al. *Development* 2003.

Activation of gene circuits

- Different lineages sometimes have the same phenotypic end point

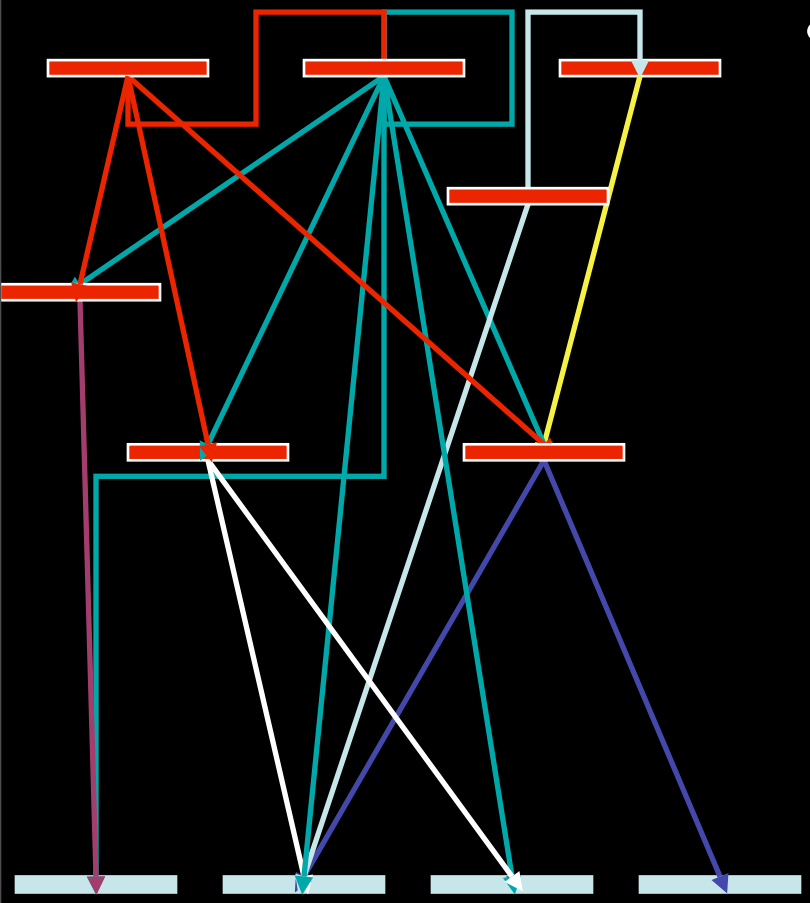


Baugh et al. *Development* 2003.



Harfe et al. *Development*, 2479-2488 (1998)

Diagram of Gene Network



- Trans-factors
  - Interacts with other factors in a network exclusive combination
  - May be involved with other networks

Trans-factor

Terminal target

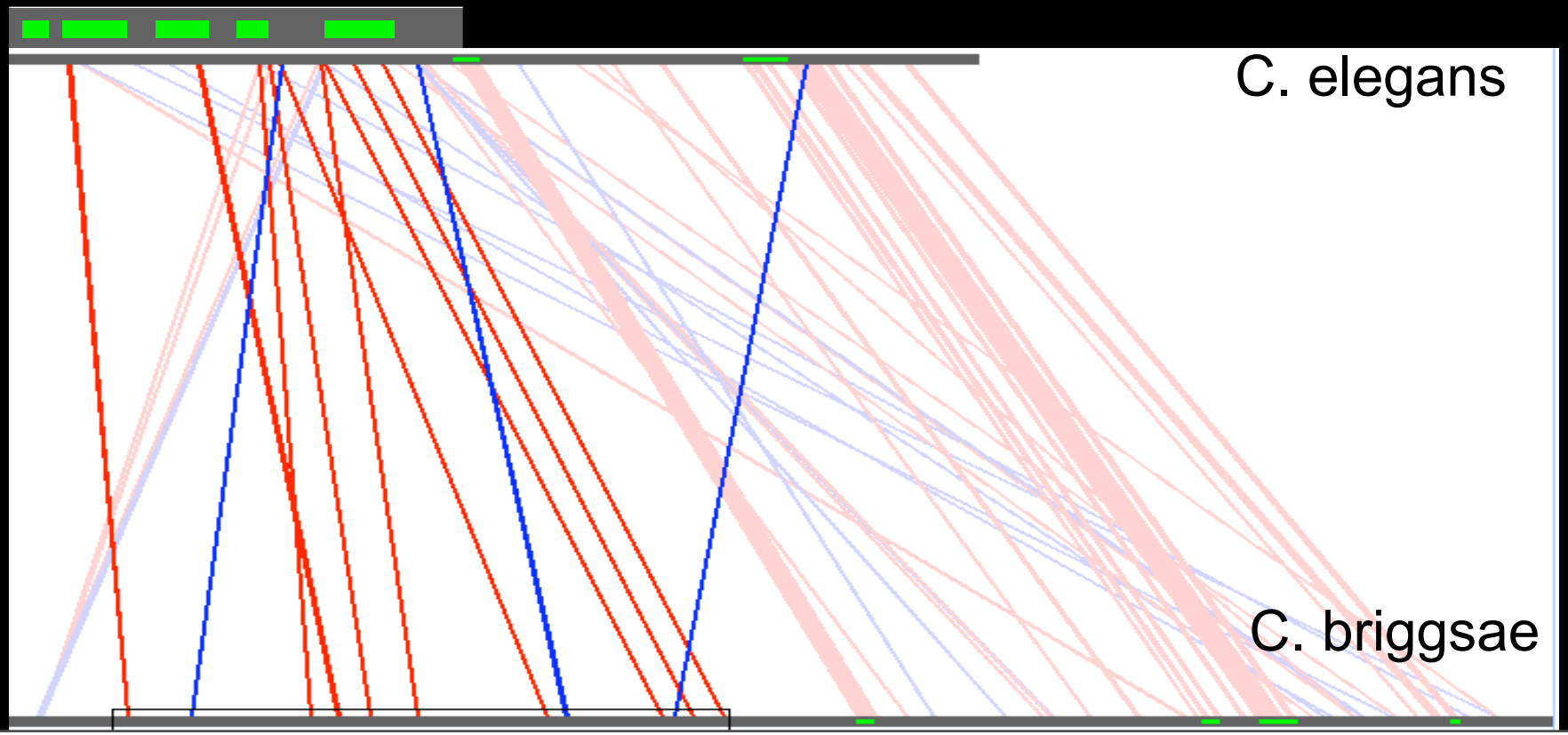
Trans/Cis connection

bHLH Shuffling

- Mouse: MyoD, Myf5, Mrf4, Myogenin
- Worm: hlh-1, hnd-1, hlh-8
- Fly: nautilus, twist

Preliminary Results: Computation

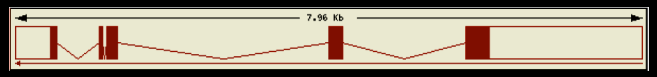
- Mussa comparison and experimentally defined elements
- Hlh-1 promoter described by Krause et al.



Preliminary Results: Computation

C. elegans vs. C. briggsae vs. CB5161: lin-39 and ceh-13

- Window: 30 Threshold: 25



Lin-39 (- strand)

Ceh-13 (+ strand)

H1 H2 H4 H H7 H9

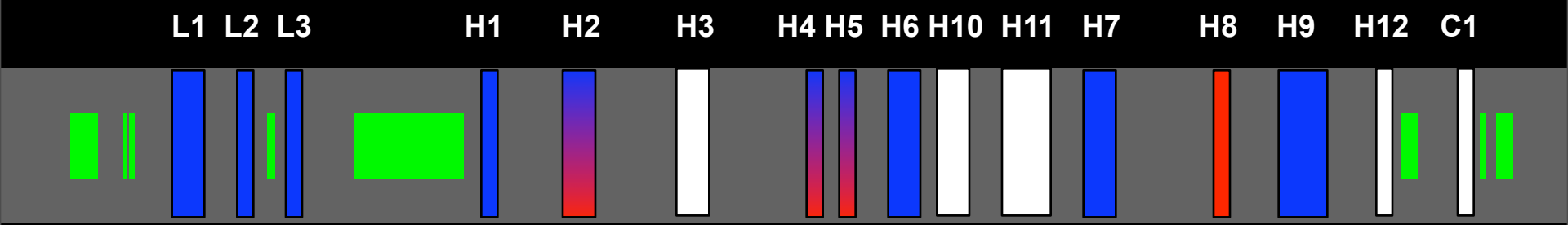
5 6

C. elegans

C. briggsae

CB5161

Enhancer Element Distribution



lin-39

ceh-13

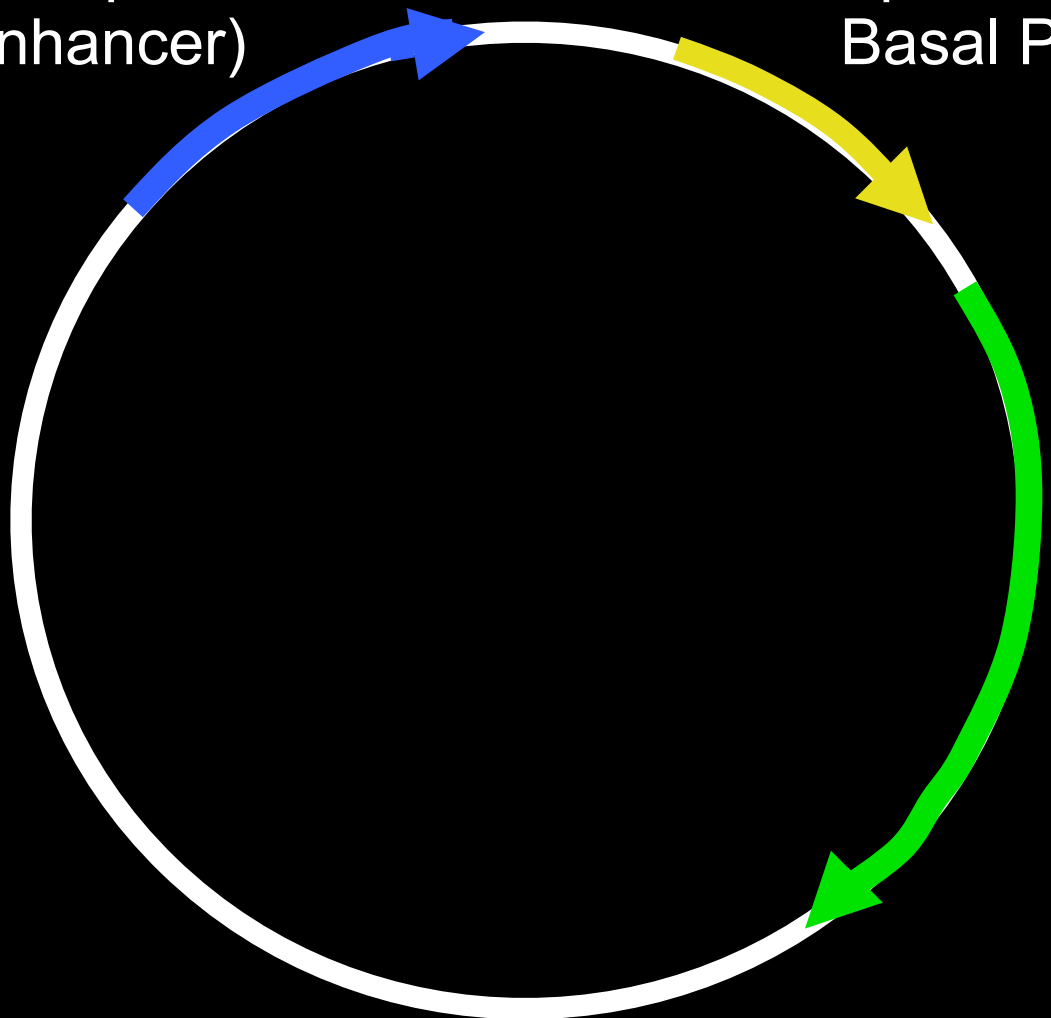
- C. elegans vs. C. briggsae
- C. elegans vs. C. briggsae vs. CB5161
- C. elegans vs. vertebrate
- Transcribed regions

Transgenic Nematode Progress

	Cloned	RC Cloned	Described	Injected	Transformed Lines	Integrated Lines
L1	X	X		X	12	
L2						
L3	X					
H1	X	X		X	?	
H2	X	X		X	?	
H3	X					
H4	X	X		X	?	
H5	X	X		X	3 (+)	12 (+)
H6	X	X	Liu et al.	X	?	
H7	X	X	Streit et al.	X	4	
H8	X	X	Streit et al.			
H9	X			X	3/12	
H10	X					
H11	X					
H12	X					
C1	X					

Expression Vector/Fusion

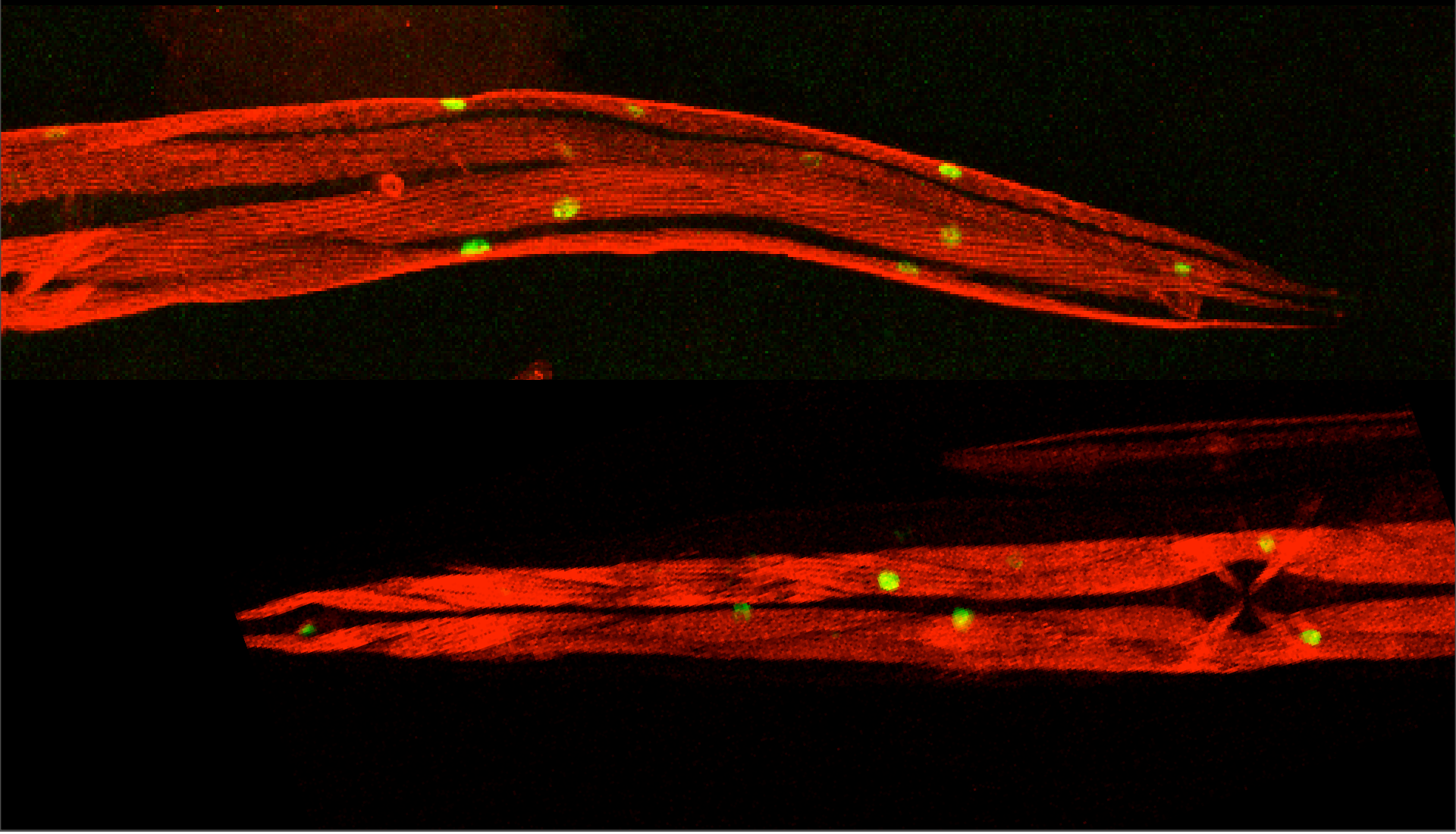
Conserved sequence
(Putative enhancer)



Δpes-10
Basal Promoter

GFP-
LacZ-
NLS
Reporter

Adult Posterior Expression Driven by H5 in Bodywall Muscle



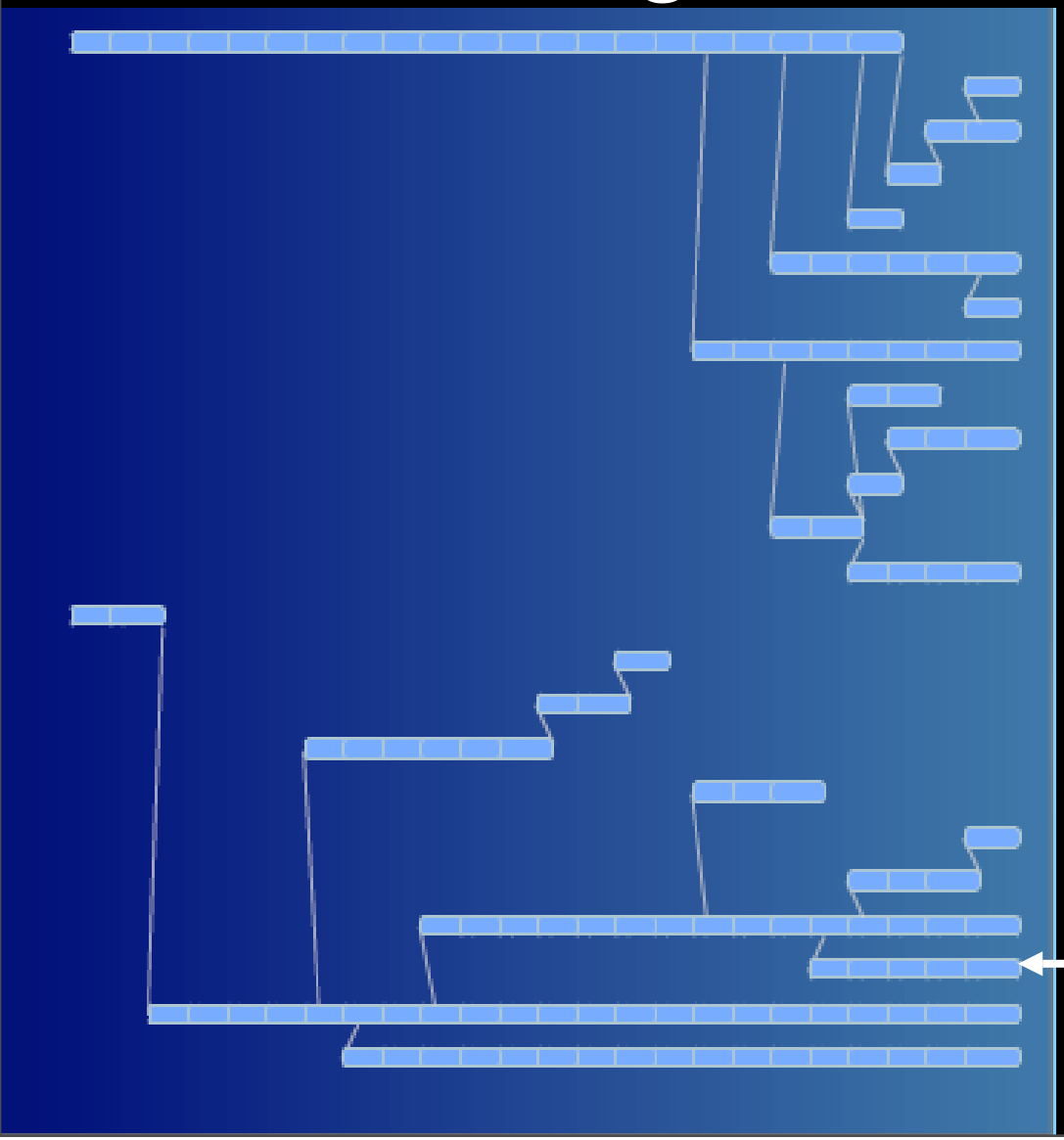
Specific Aims

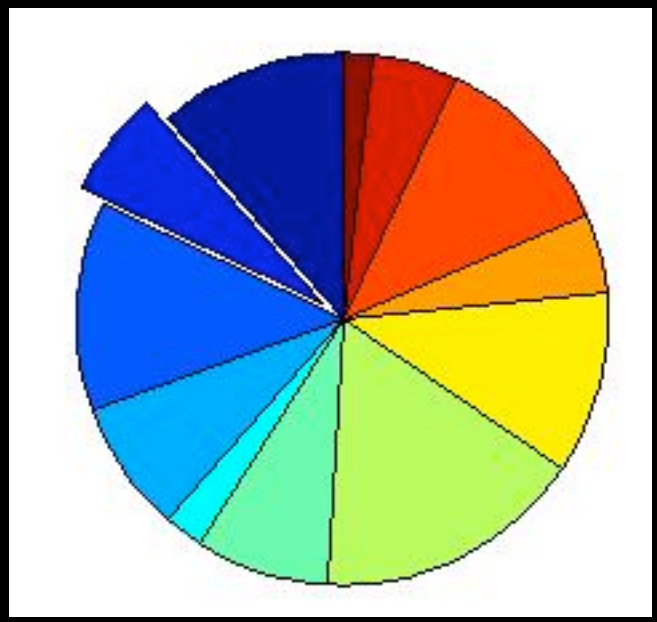
1. A muscle development regulatory network will be computationally isolated
2. The uniqueness of the regulatory core will be demonstrated
3. The interphylum relationship of the muscle development regulatory cores will be determined

Specific Aim 1: The Muscle Development Regulatory Network

- A muscle development regulatory network will be identified through computational analysis of genomic and expression data alone

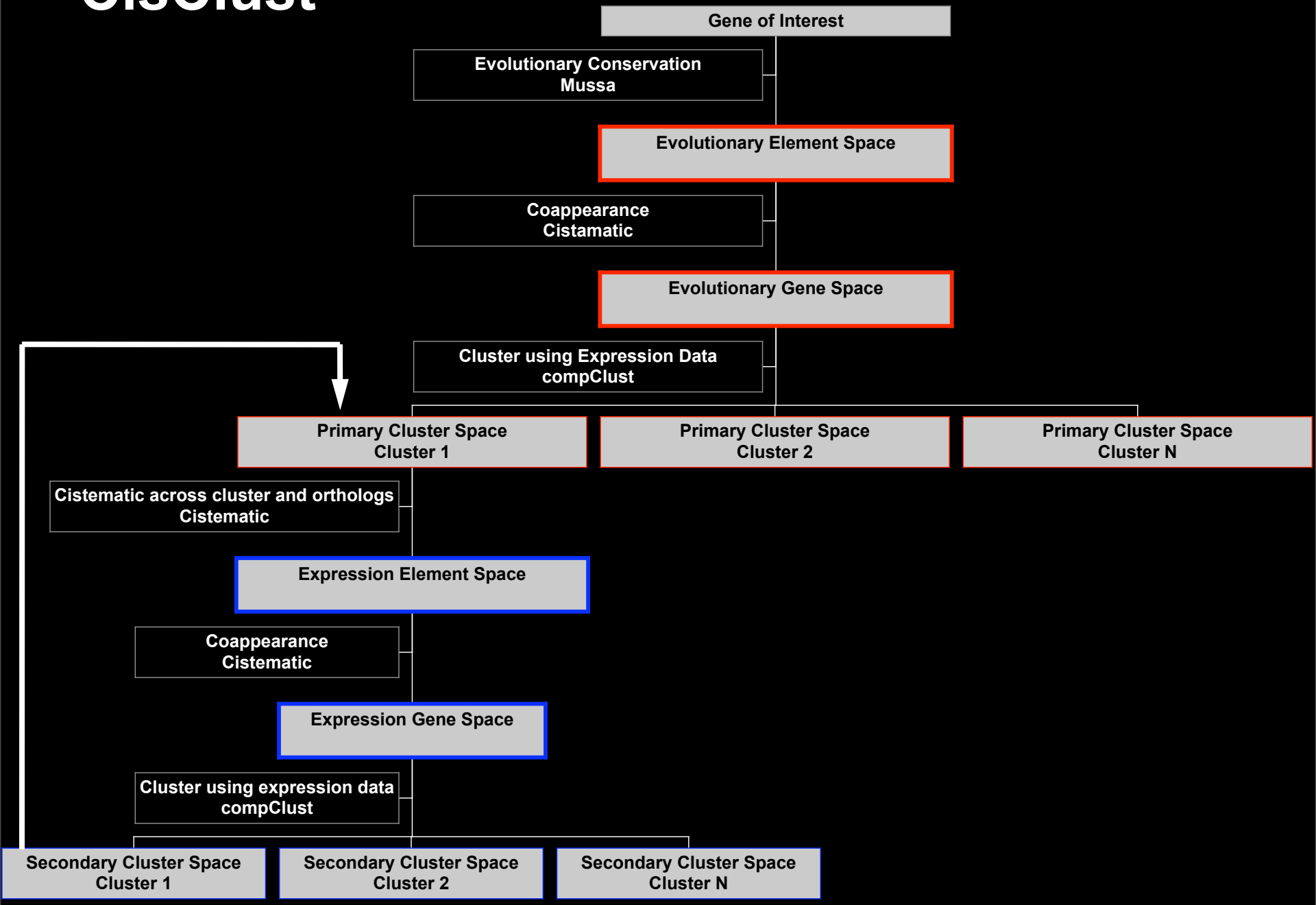
Iterative Signature Algorithm (ISA) Clustering of the GNF dataset





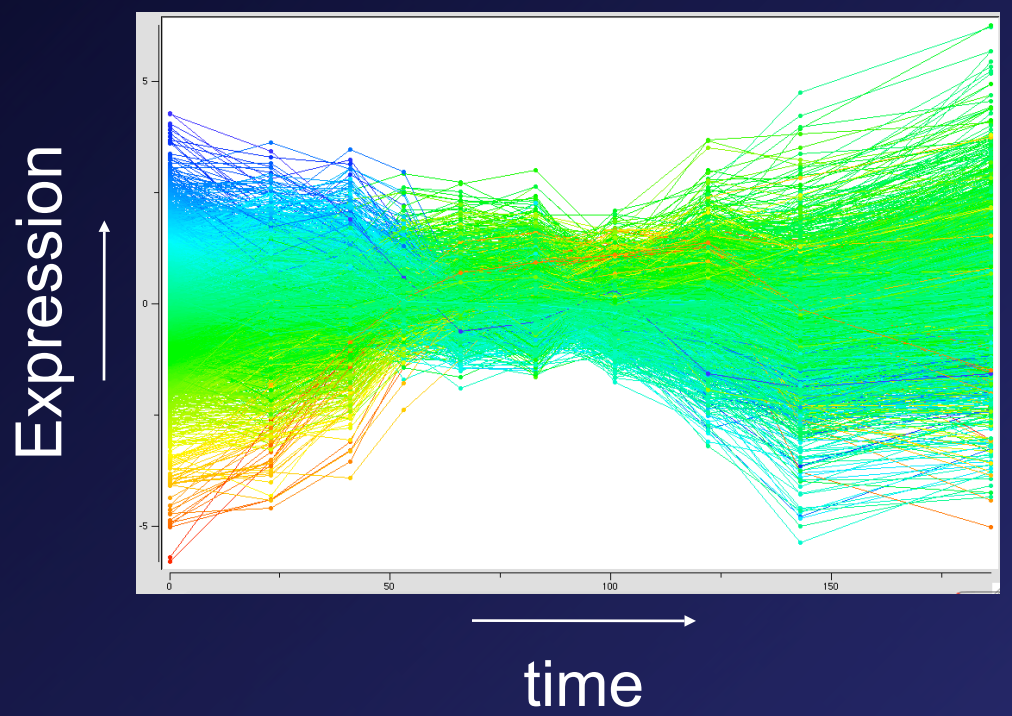
Skeletal and cardiac muscle

CisClust

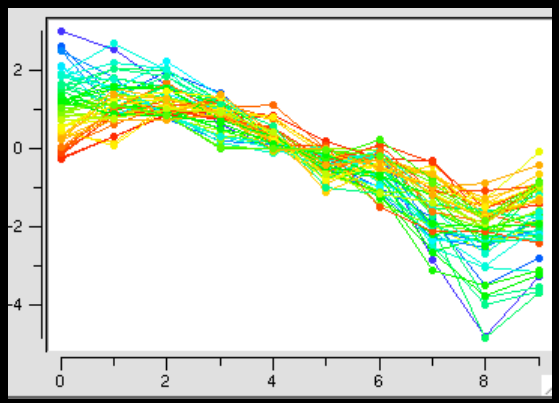

Gene of Interest
Evolutionary Conservation
Mussa
Evolutionary Element Space
Coappearance
Cistamatic
Evolutionary Gene Space
Cluster using Expression Data
compClust
Primary Cluster Space
Cluster 1
Primary Cluster Space
Cluster 2
Primary Cluster Space
Cluster N
Cistematic across cluster and orthologs
Cistematic
Expression Element Space
Coappearance
Cistematic
Expression Gene Space
Cluster using expression data
compClust
Secondary Cluster Space
Cluster 1
Secondary Cluster Space
Cluster 2
Secondary Cluster Space
Cluster N

Available Expression Data

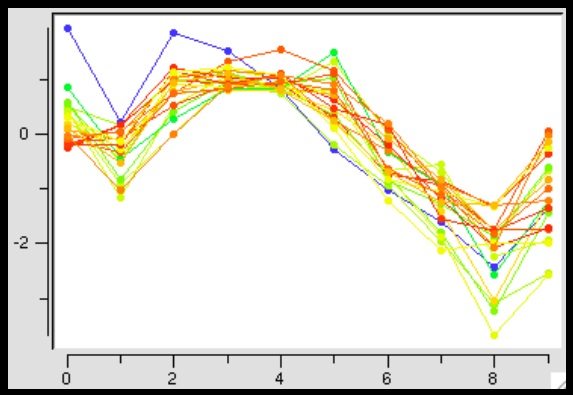
- Time courses
  - C2C12 differentiation
  - N2, mex-3;skn-1, and pie-1 early embryonic (Baugh et al. *Development* 2003)
- Tissue specific
  - GNF data
  - Mango group
  - Kim group



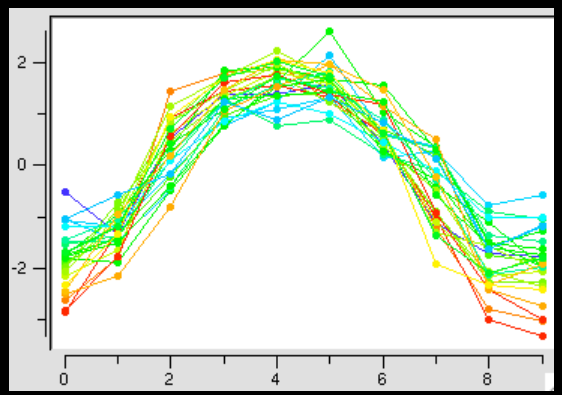
Clustered genes (mex-3)



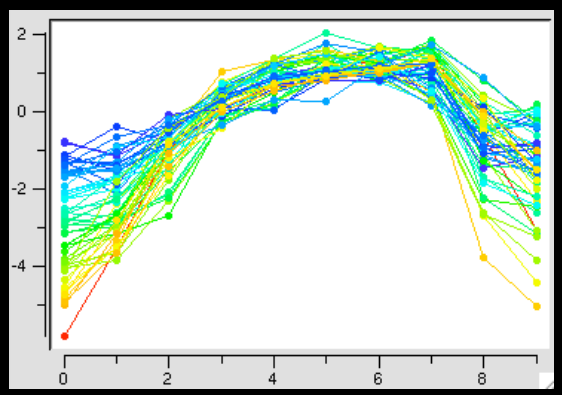
1st maternal: 54 genes



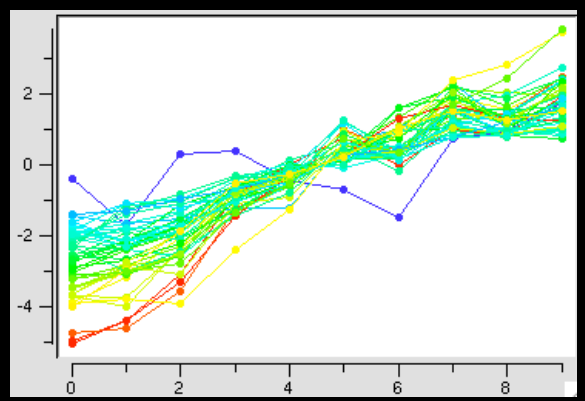
2nd maternal: 23 genes



1st embryonic: 26 genes

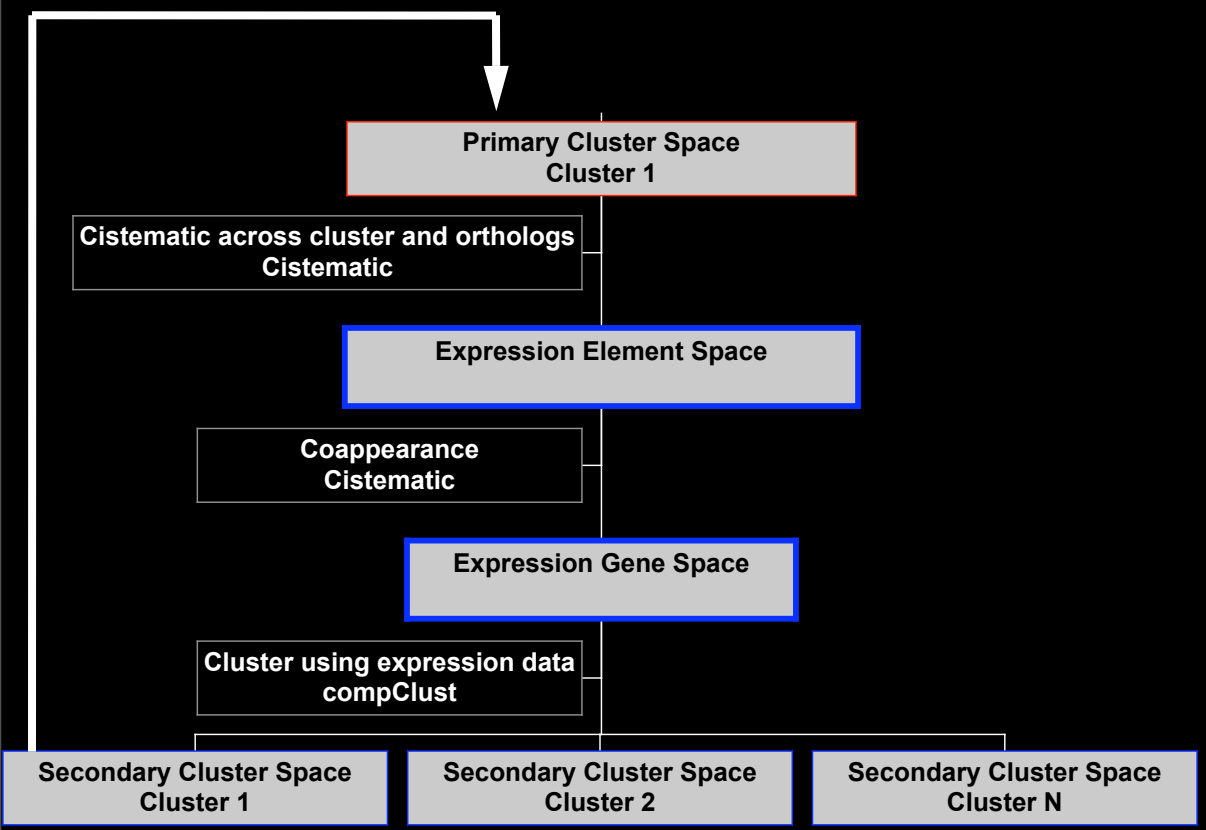


2nd embryonic: 54 genes

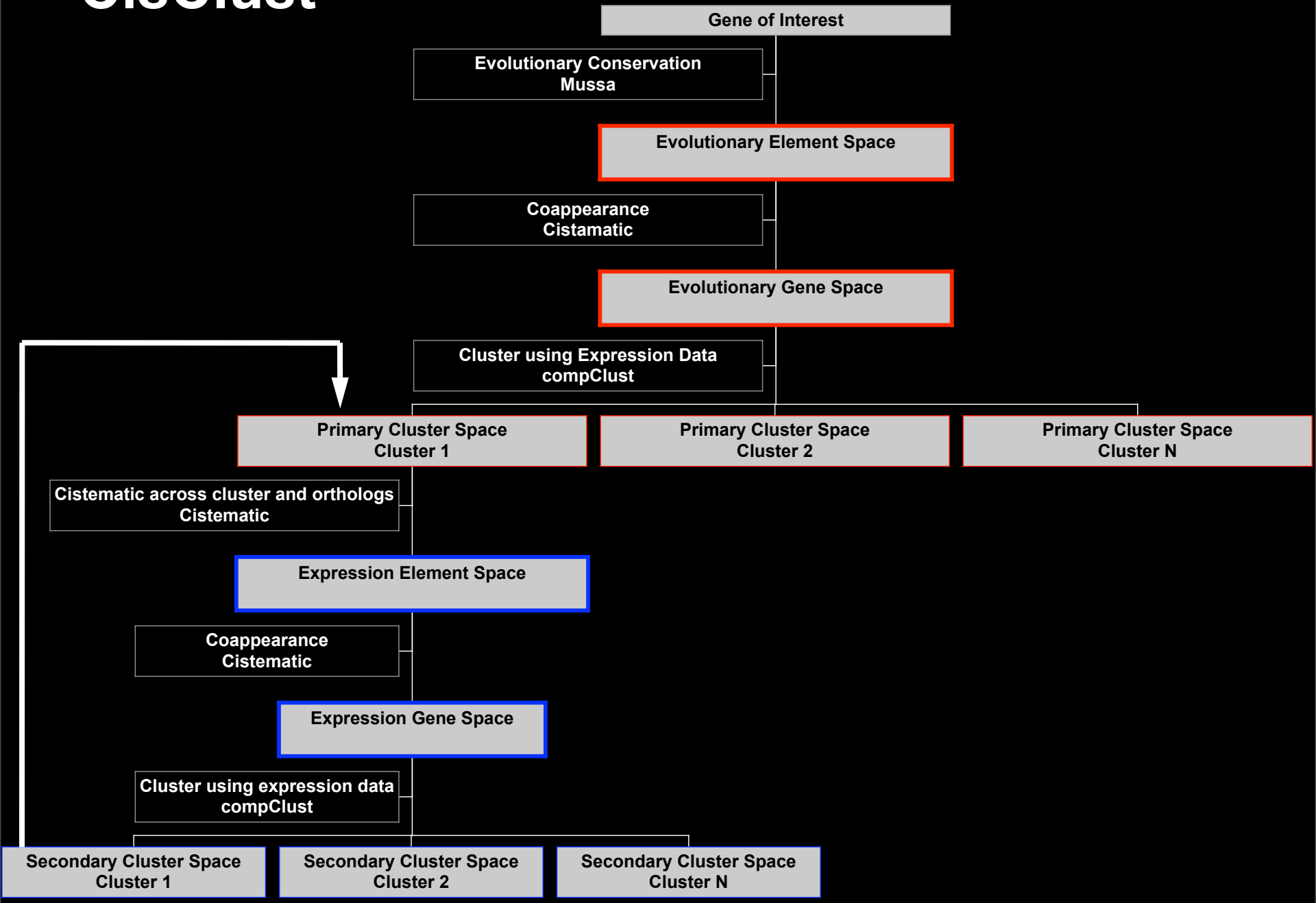


3rd embryonic: 54 genes

CisClust


Primary Cluster Space
Cluster 1
Cistematic across cluster and orthologs
Cistematic
Expression Element Space
Coappearance
Cistematic
Expression Gene Space
Cluster using expression data
compClust
Secondary Cluster Space
Cluster 1
Secondary Cluster Space
Cluster 2
Secondary Cluster Space
Cluster N

CisClust


Gene of Interest
Evolutionary Conservation
Mussa
Evolutionary Element Space
Coappearance
Cistamatic
Evolutionary Gene Space
Cluster using Expression Data
compClust
Primary Cluster Space
Cluster 1
Primary Cluster Space
Cluster 2
Primary Cluster Space
Cluster N
Cistematic across cluster and orthologs
Cistematic
Expression Element Space
Coappearance
Cistematic
Expression Gene Space
Cluster using expression data
compClust
Secondary Cluster Space
Cluster 1
Secondary Cluster Space
Cluster 2
Secondary Cluster Space
Cluster N

Towards Specific Aim1:

- Gene list output refined
  - Subject interesting targets to further expression analysis using qRT-PCR during critical time points of development
  - Screen against whole-tissue micro-array analyses (Roy et al.)

Specific Aim 2: The Regulatory Core

- Muscle development regulatory networks are structured around a regulatory core
- The core will be dissected out with pattern recognition across the available networks
- The necessity and sufficiency of core components will be demonstrated
- Within an organism, the regulatory core is unique to a tissue in terms of the trans-factor and cis-element interactions and combinations

Identification of Muscle Trans-Factors

- Reconstruct fundamentals of the core
  - Use literature studies
  - Use pattern recognition of cis-elements and expression clustering
- Use protein annotations/homology screening to isolate transcription factors
  - Screen battery components for known transcription factors or transcription factor domains
- Tag ‘important’ candidates (~20)
  - Identify by importance from literature or patterns
  - Create GFP-fusions, His-tags, or other tags
  - Rescue mutant background

Validating the Regulatory Core

- Synthetic target of cis-elements driving GFP reporter
  - Sufficiency of cis-elements tested
  - Check expression for specific set of cis-elements
- Synthetic target with critical mutations in cis-elements
  - Necessity of cis-elements tested
  - Check expression for specific set of mutated cis-elements
- Native target driving GFP reporter in null background
  - Necessity of trans-factors tested
  - Expression in mutant background and/or RNAi
  - Sufficiency of tagged trans-factors to rescue mutant
  - ChIP/EMSA to test rescue

The Regulatory Core is Tissue Specific

- Test substituted proteins for rescue
- To enhance phenotype, add RNAi to mimic synthetic lethal when necessary
  - Example: For unc-120 null add hlh-1 RNAi

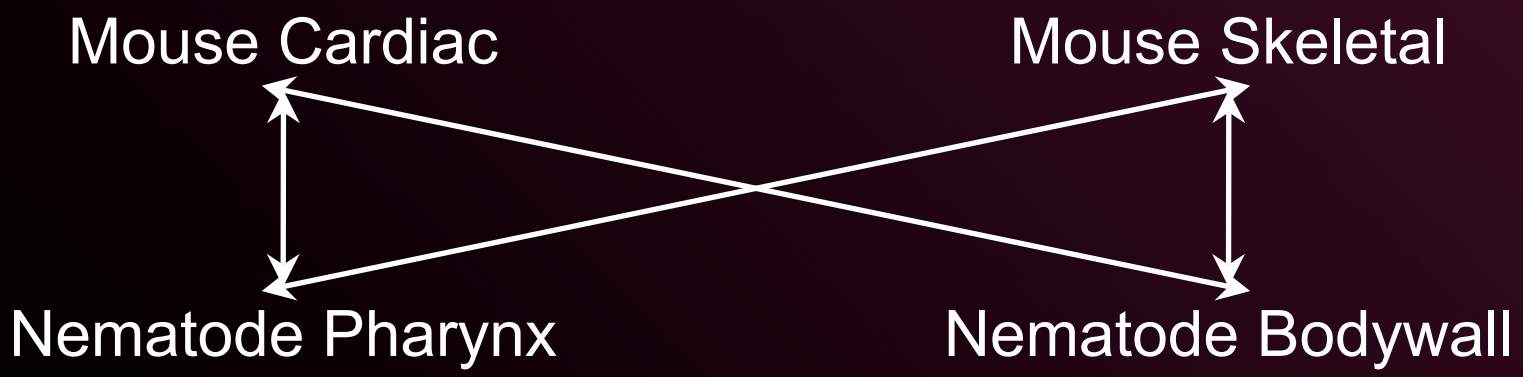
	unc-120:mef-2	mef-2:unc-120
mef-2 null	Pharynx rescue?	Pharynx rescue?
unc-120 null	Bodywall rescue?	Bodywall rescue?

Specific Aim 3: Core conservation

- Tissue-specific regulatory cores are conserved between phyla
- Regulatory cores share greater homology with orthologous cores across phyla than with other muscle cores in the same organism

Cross-Phyla Analysis

- Hybrid protein expression studies (FIGURE)
  - Changing the active site to that of another phylum
  - Screen with RNAi to enhance phenotype
- Target substitution expression studies
  - Introducing mammalian upstream sequence to a nematode factor (ceh-22, ceh-24 vs. Nkx2-4, Nkx2-5)
- Cross-reactivity between pulsating and striated



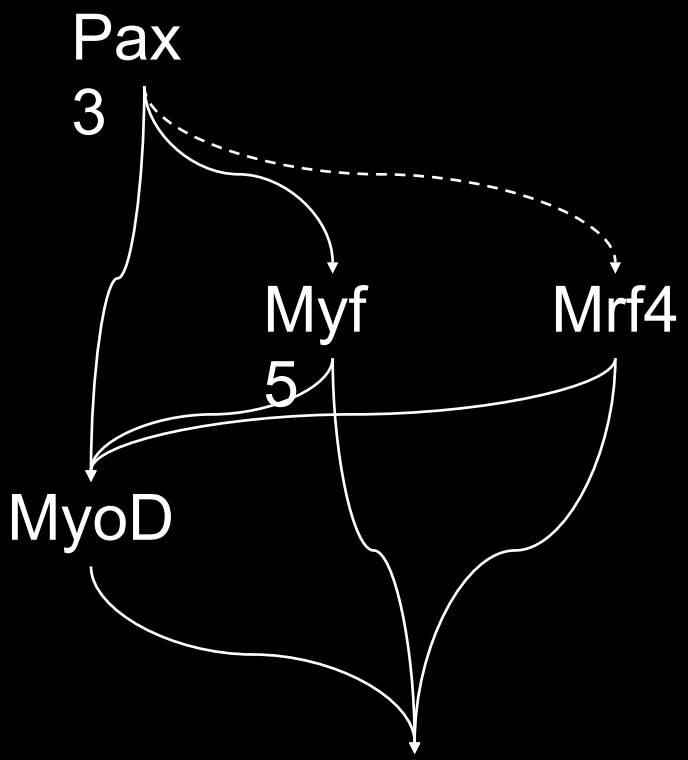
Acknowledgments

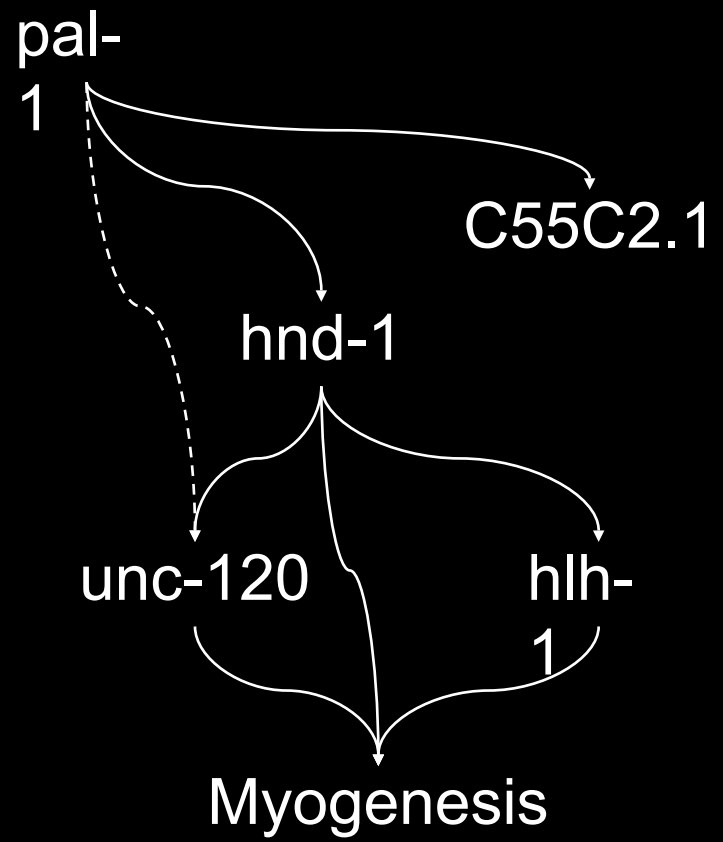
- Wold Lab
  - Ali Mortazavi
  - Tristan DeBuysscher
  - Chris Hart
  - Diane Trout
  - Brandon King
  - Leslie Dunipace
  - Gilberto Hernandez
  - Brian Williams
- Davidson Lab
  - C. Titus Brown
- Sternberg Lab
  - Erich Schwarz
  - Mihoko Kato
  - David Sherwood
  - Igor Antoshechkin
  - Takao Inoue
  - Ryan Baugh
  - Jennifer Sanders
  - Adeline Seah
  - John DeModena
  - Jolene Fernandes

Key Transcription Factors are

- Eye formation (Stierwald et al. *Developmental Biology* 274 (2004) 70-81.)
  - Pax6
  - Six, Dach, Eya
- Neuron development (Thor and Thomas. *Curr Op Genetics and Dev,* 12:558-564 2002.)
  - Islet, Lhx, Hb9
  - Eve/vab-7
  - Evx
  - Atonal?
- Muscle development
  - MEF (MADS-box)
  - MyoD (bHLH)

Some known transcription





Candidate CREs

- bHLH CANNTG
- MADS-box AT-rich
- Zn-finger CAXXTG
- GuhaThakarta
- Mango
- Baugh CCCNTCCAANNGCCGC, GGNNGNGAAANAAC, TGTCACACCNNNC
- Other