Genomics of *Ciona intestinalis*

- David N. Keys
  - dnkeys@lbl.gov
- Joint Genome Institute
  - DOE • LBNL • LLNL
  - Walnut Creek, CA
  - www.jgi.doe.gov

Genomics of *Ciona intestinalis*

Joint Genome Institute
LBNL - LLNL - DOE

Paramvir Dehal
Jarrod Chapman
Chris Detter

Dan Rokhsar
Paul Richardson
Trevor Hawkins

University of California Berkeley

Anna Di Gregorio
Mike Levine

Kyoto University

Yutaka Satou
Nori Satoh

A Functional Genomics Approach to Developmental Genetics

Joint Genome Institute
LBNL - LLNL - DOE

University of California Berkeley

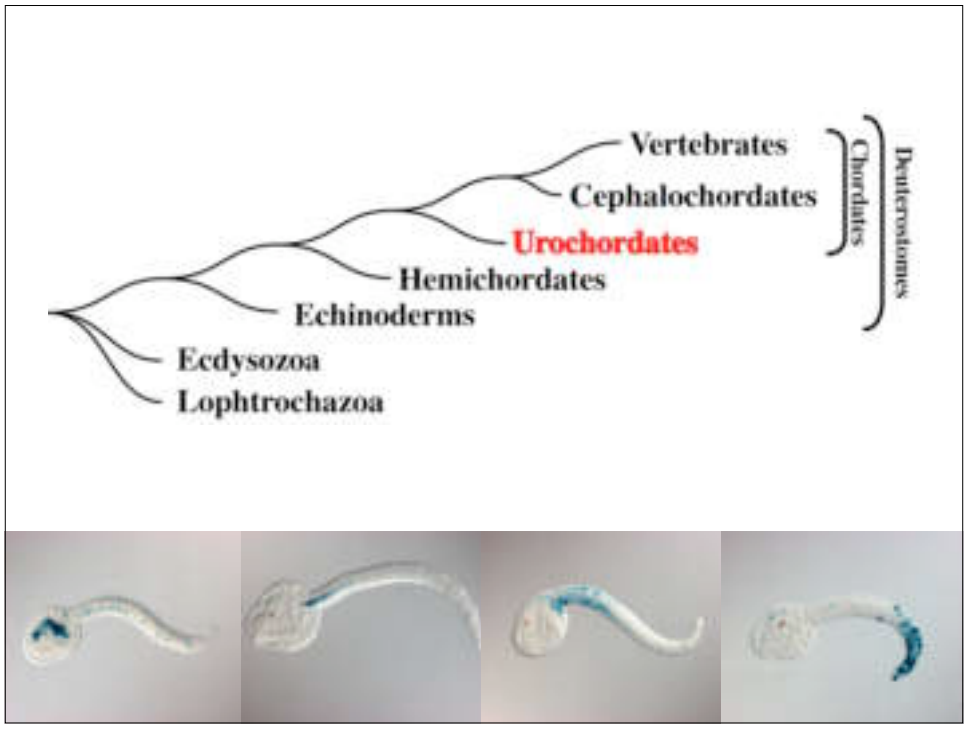
Byung-in Lee
Chris Detter
Stephan Trong

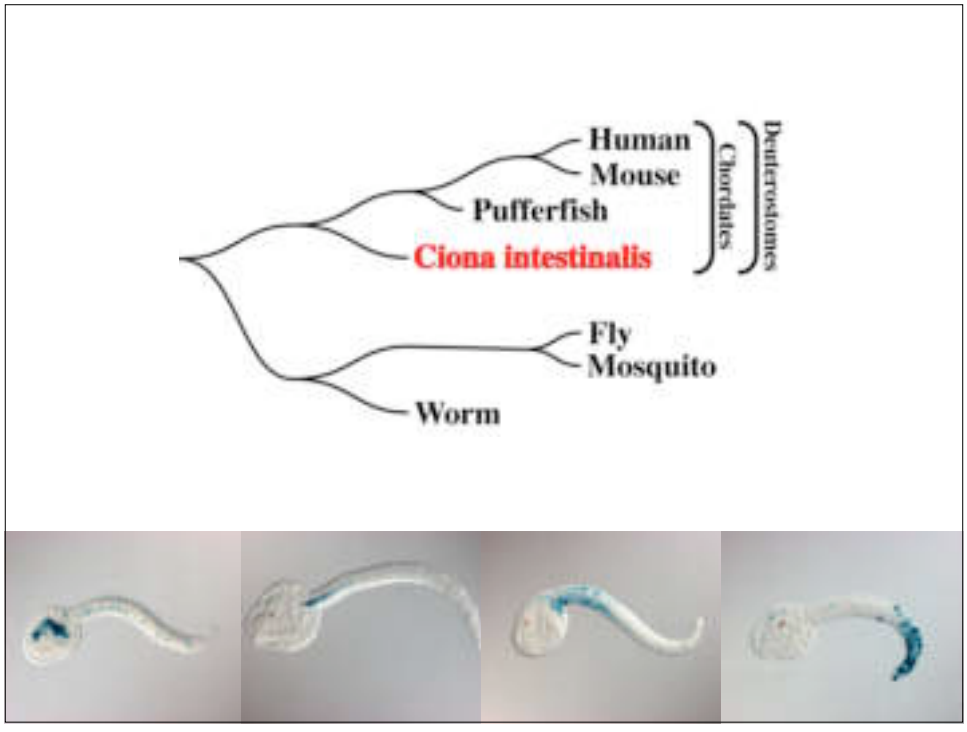
Dan Rokhsar
Paul Richardson
Trevor Hawkins

Syvia Ahn
Dave Engle
Orsalem Kahsai
Maria Shin
Joann Wang
Mei Wang

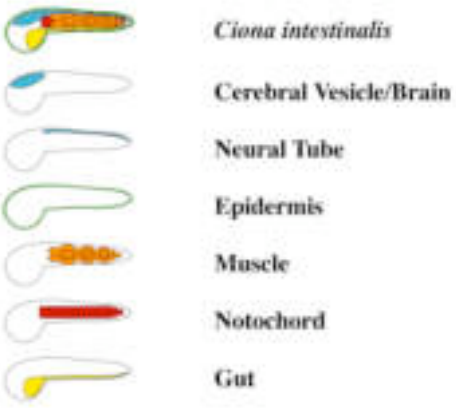
Naoe Harafuji
Anna Di Gregorio
Mike Levine

Genomics of *Ciona intestinalis*


Vertebrates
Cephalochordates
Urochordates
Hemichordates
Echinoderms
Ecdysozoa
Lophtrochazoa
Chordates
Deuterostomes


Human
Mouse
Pufferfish
Ciona intestinalis
Fly
Mosquito
Worm
Chordates
Deuterostomes

Body plan simplicity
Represents the ancestral chordate?



Genomic simplicity
Represents the ancestral chordate?

- ~16,000 genes
  - (~half of typical vertebrate)
- ~155 Mbp
  - (116.7Mbp nonrepetative in current assembly)
- Therefore ~1 gene every 10kb
- Small gene families
  - (pre-vertebrate duplication)
- *Drosophila* like genome considerations

Experimental Tractability

Disadvantage

No true genetics

Advantages

Easy transgenics

Scorable phenotypes

Availability

Functional Genomics

- Studying large sets of genes in parallel rather than single genes
- Experimental, not observational or modeled
- Invent new hypothesis testing experiments
- Scale traditional hypothesis testing experiments to the entire genome

Large Scale *Cis*-Reg Hunts

•**Primary Goal**

•Screen genomic libraries for *cis*-regulatory activity

•Catalog a large number of functionally defined *cis*-regulatory elements

•**Secondary Goal**

•Do some targeted developmental genetics along the way

Results

Catalog a large number of functionally defined *cis*-regulatory elements

- Design, implementation and results of a small scale pilot screen of random genomic DNA - 11
- Design, implementation and results of an exhaustive screen of a medium size (250kb) genomic domain
- Design, implementation of an on going large scale screen of random genomic DNA

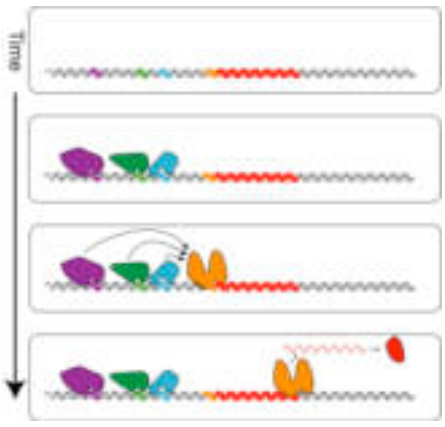
Themes to Keep in Mind

- Trade Offs
  - Number of characterized elements
  - Resolution of the characterizations
- Biases
  - Experimental biases
  - Experimenter's biases
    - Nature of enhancers vs detection methods

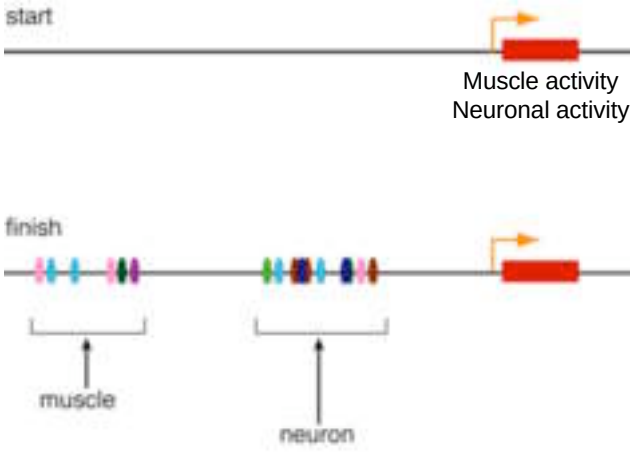
Technology to take *cis*-regulatory screening to the genomic level

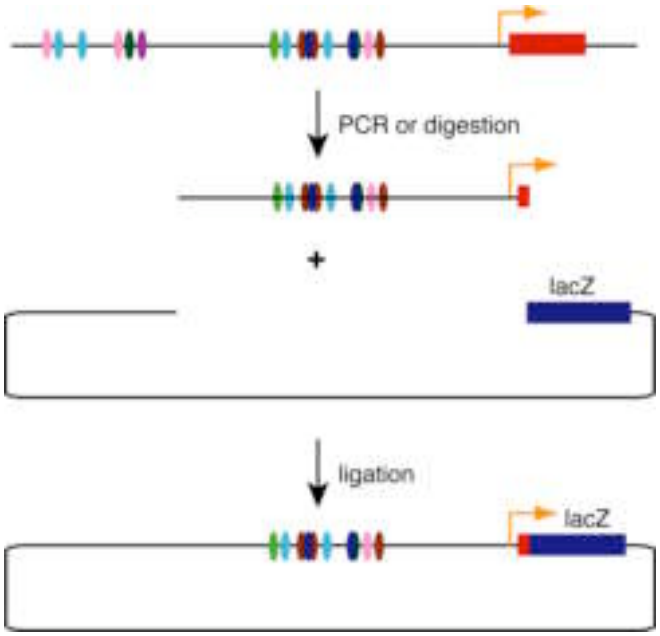
Scale traditional hypothesis testing experiments to the entire genome

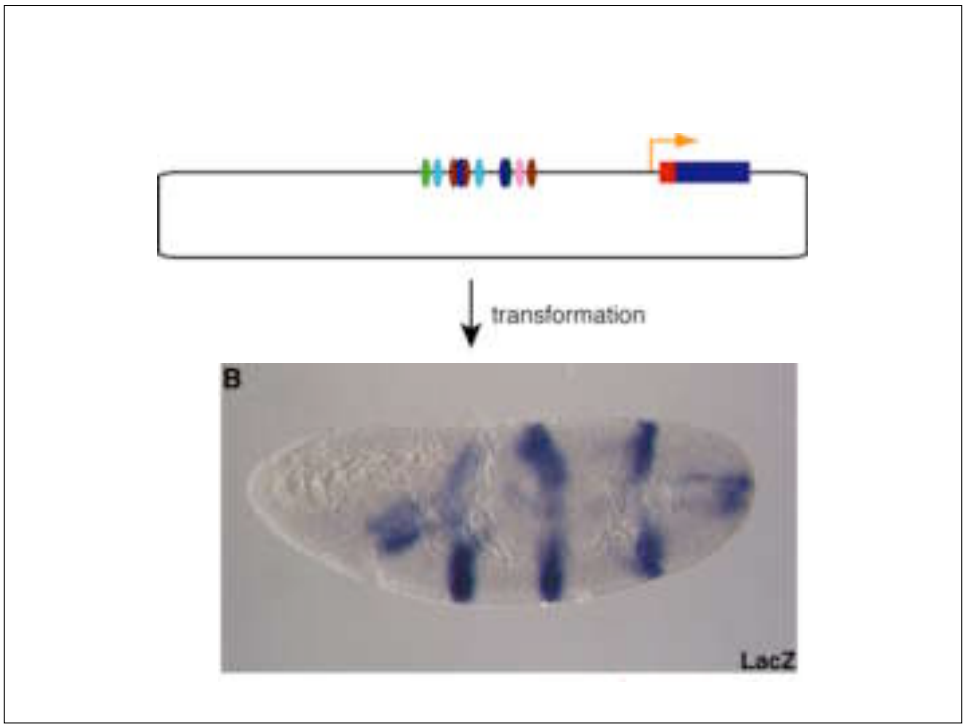
A genetic switch

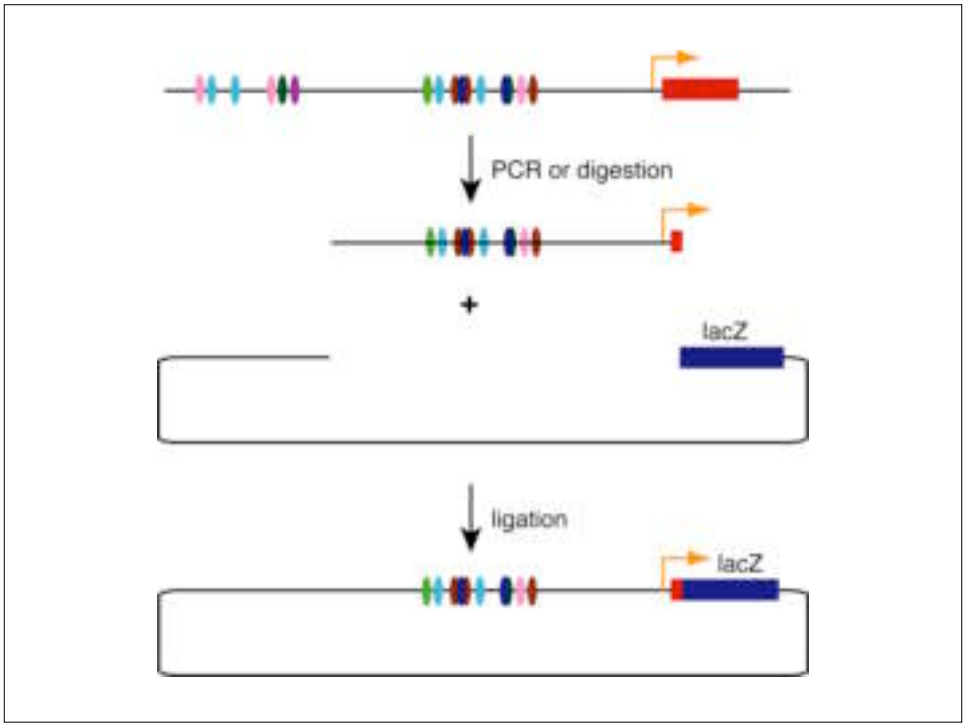


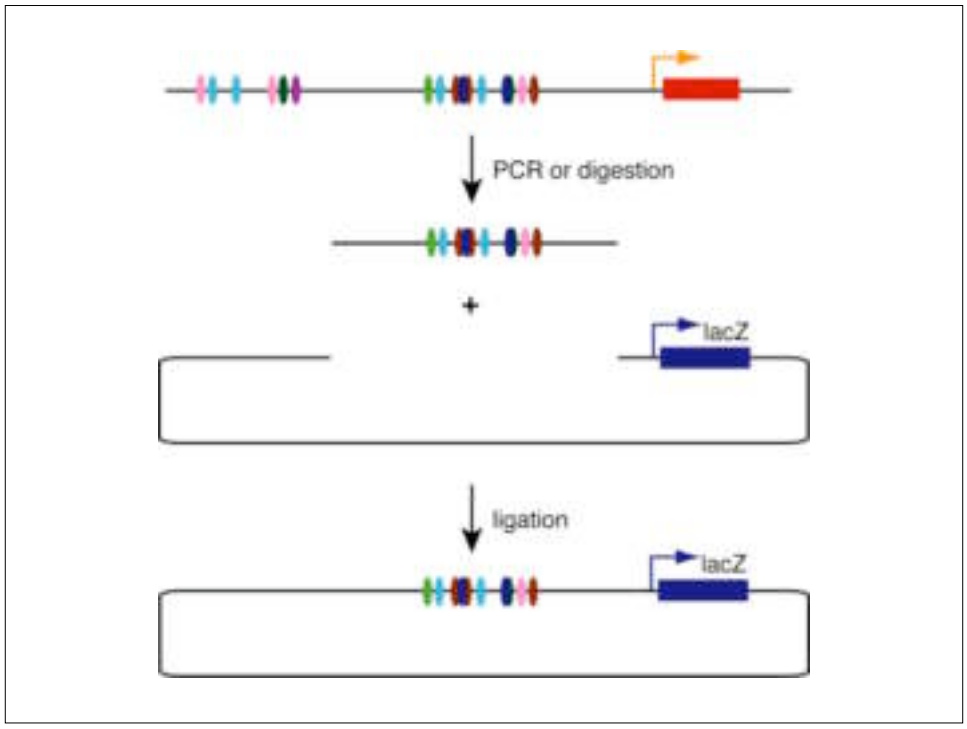
Enhancer Characterization

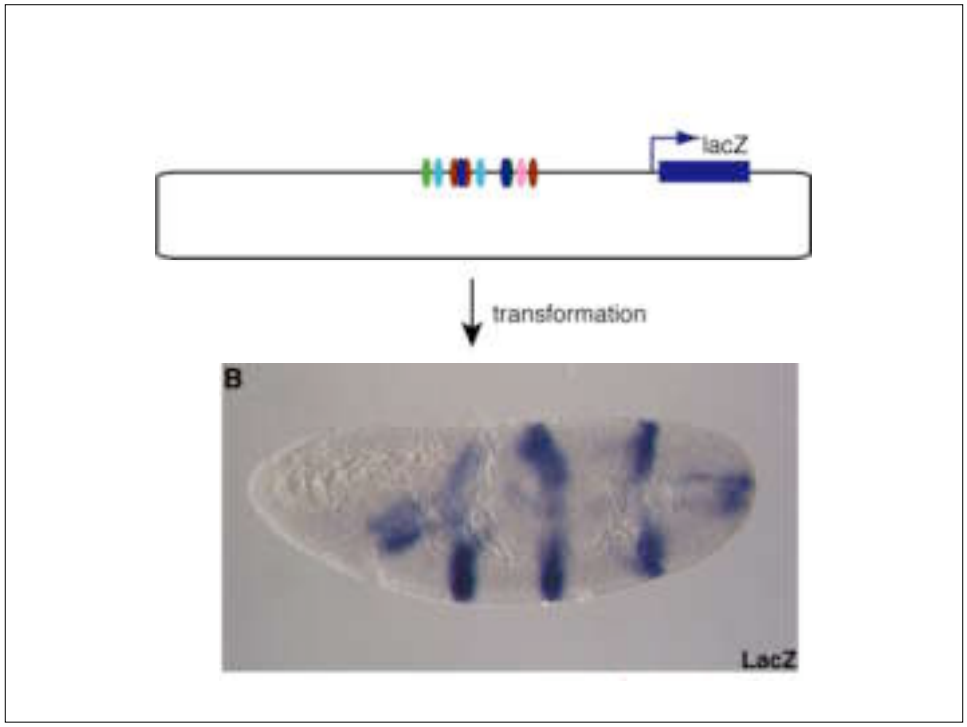





transformation
B
LacZ


PCR or digestion
+
lacZ
ligation
lacZ


PCR or digestion
+
lacZ
ligation
lacZ


lacZ
transformation
B
LacZ

Drosophila transformation

- Collect naturally laid eggs
- Dechorionation
- Transform by single embryo microinjection
- Individually rear to 2nd generation
- Screen
- **Total Time: month(s)**

Traditional Enhancer Characterization

- Targeted
- Slow/Labor intensive
  - Building specific DNA constructs
  - Transforming into animals
  - Maintaining/screening animals

Ciona intestinalis
Sea Squirts

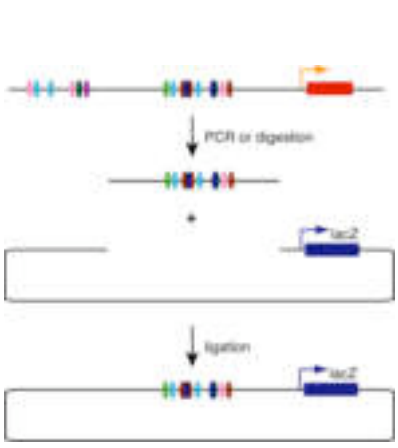
Ciona intestinalis transformation

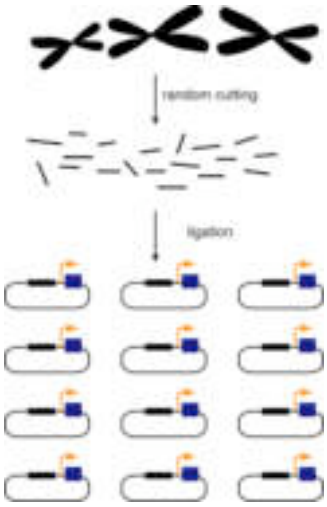
- Collect eggs & sperm via dissection
- Dechorinonation
- Mix embryos with plasmid DNA (100ug)
- Transform by batch electroporation of single cell embryos
- Incubate 1-24 hours
- Stain (GFP/lacZ/in situ) & visually screen
- **Total time: 24hours**

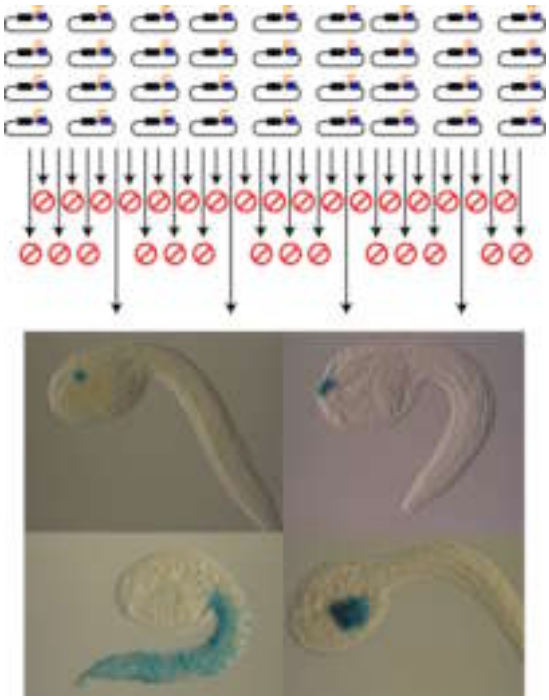
Standard *Ciona* Enhancer Characterization

- Targeted
- Slow/Labor intensive
  - Building specific DNA constructs
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals

Build enhancer screening libraries instead of specific constructs





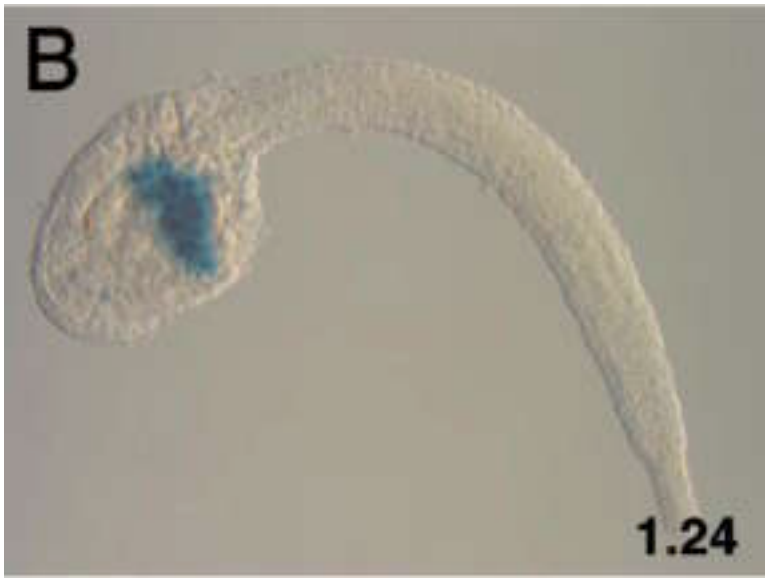


Ciona Enhancer Screening

- Non-Targeted
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals
  - Building random DNA constructs
- Limiting factors
  - DNA preps (50-100ug)
  - Transformation window (single cell embryos)
  - Imaging

Pilot Genomic Screen

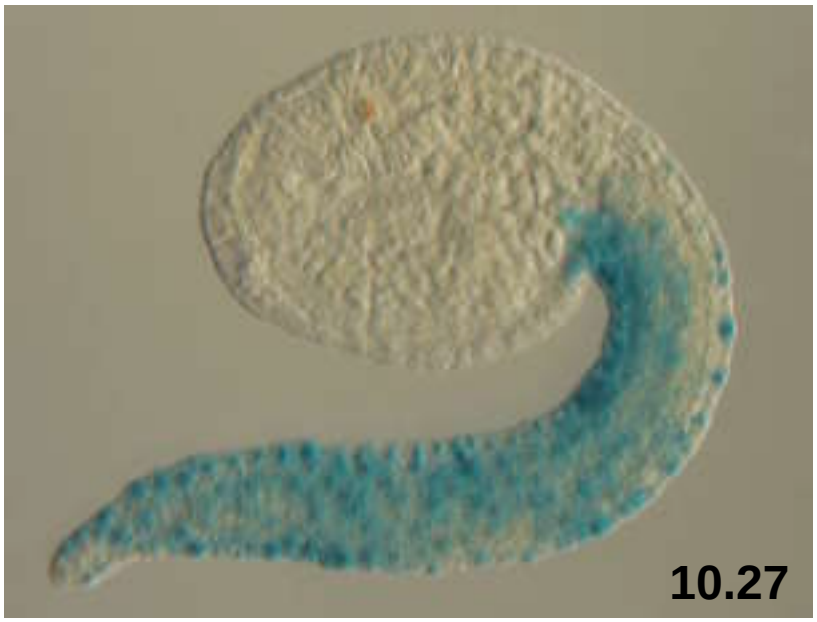
- Construct:
  - *Ciona Forkhead* basal promoter
  - *lacZ* marker detected by *beta*-Gal activity
  - Random genomic Sau3AI frags, 1.7kb average
- Prediction:
  - Will find cis-regulatory DNA
  - Gene density = 1 gene per 10kb. Therefore could find 1 enhancer every 10kb


B
1.24


2.20


2.24

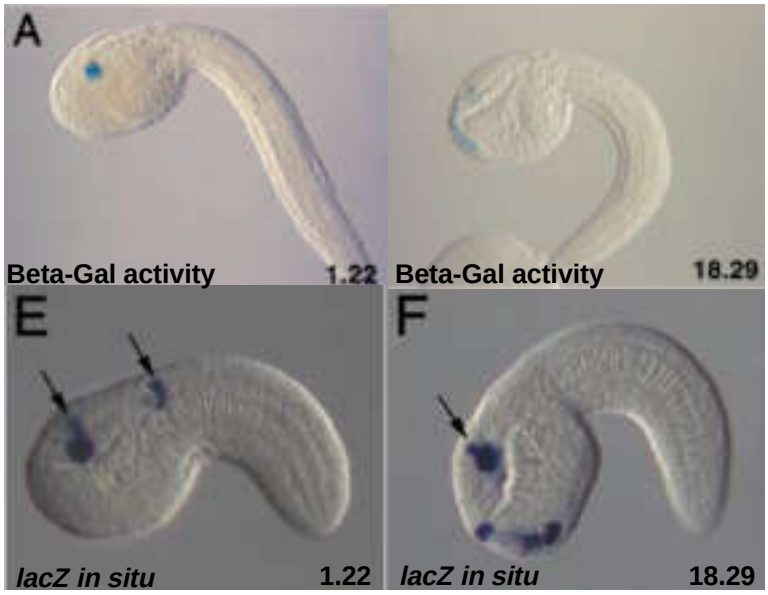

D
6.17



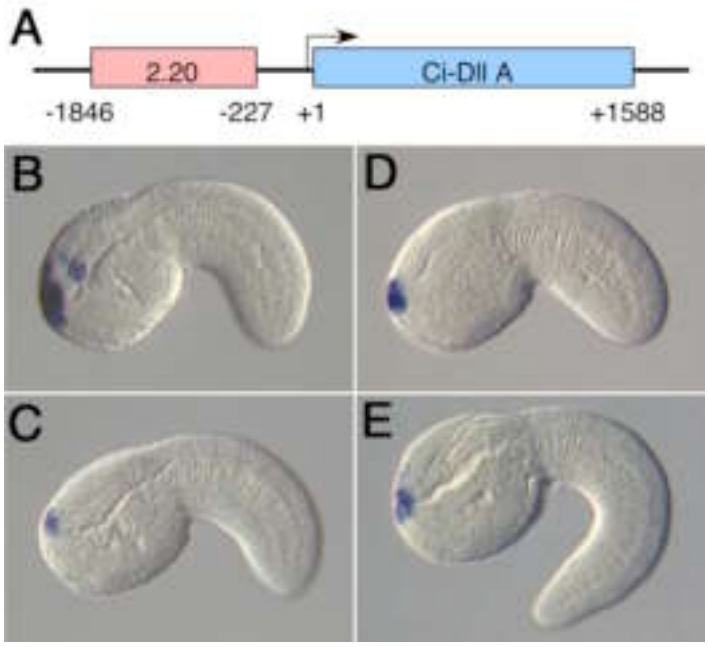
First genomic screen

- 138 constructs
- 250kb screened
- 0.15% of the genome

- **Results:**
- 11 strong cis-regulatory elements
- At least 8 appear to be "real" enhancer elements
- One confirmed enhancer
- 1 detectable element every 23-31 kb
- 1 detectable every 2-3 genes.


A
Beta-Gal activity
1.22
Beta-Gal activity
18.29
E
F
lacZ in situ
1.22
lacZ in situ
18.29


2.20



- **Prediction:**
  - Gene density = 1 gene per 10kb
  - Therefore could find 1 enhancer every 10kb
- **Results:**
  - 1 detectable element every 23-31 kb
  - 1 detectable element every 2-3 genes

Potential Issues

- Promoter specificity
- Insulators & repressors
- Enhancer Polarity
- Promoter competition
- Enhancers fragmented during cloning

- Timing
- Insufficient detection strength

Ciona Enhancer Screening

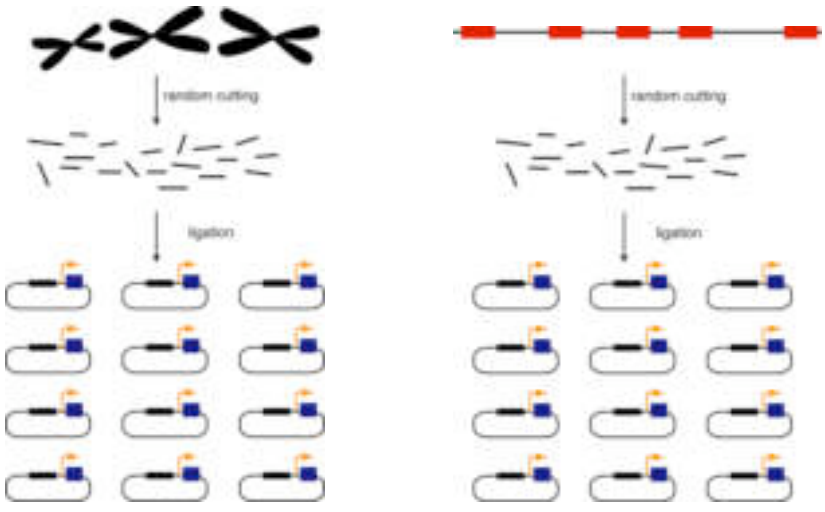
- Non-Targeted
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals
  - Building random DNA constructs
- Limiting factors
  - DNA preps (50-100ug plasmid)
  - Transformation window (single cell embryos)
  - Imaging

Limiting Factors

- DNA preps (50-100ug plasmid)
  - Qiagen Midipreps - up to 48 constructs per day
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
- Imaging
  - Quality trade offs - Tough decisions

Semi-targeted *Ciona* enhancer screen

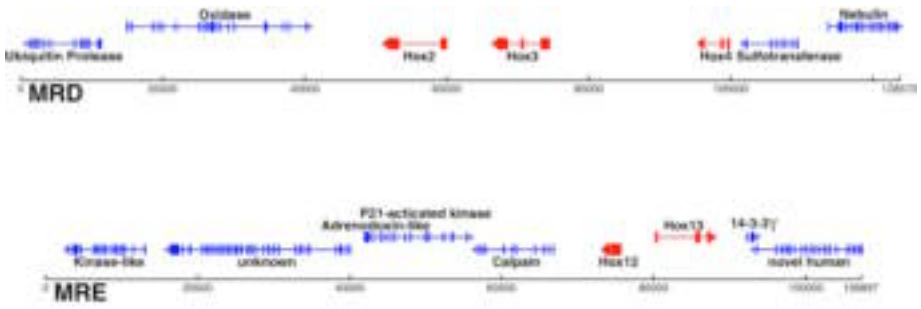
Build random libraries from limited regions



Target: *Ciona Hox Complex*

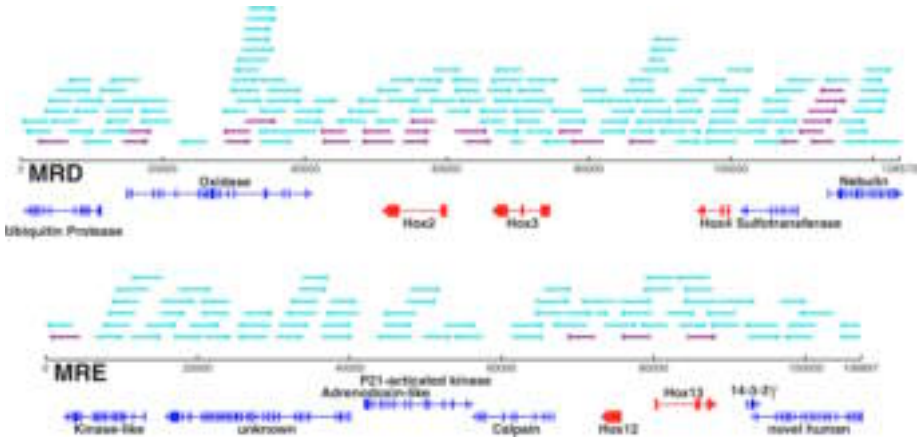
Predictions:

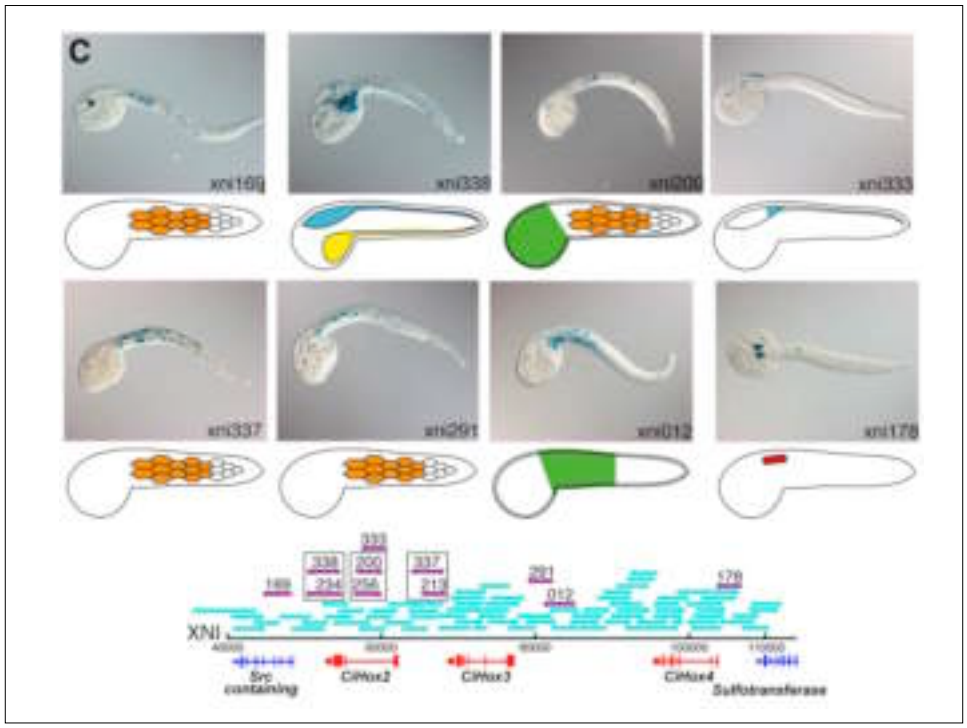
- Should be a single *Hox Complex*
- Should be a single domain
- Predictable expression patterns
  - *Hox3* & *Hox5* described by in situ
    (Branno & Di Lauro, Stazione Zoologica, Naples)

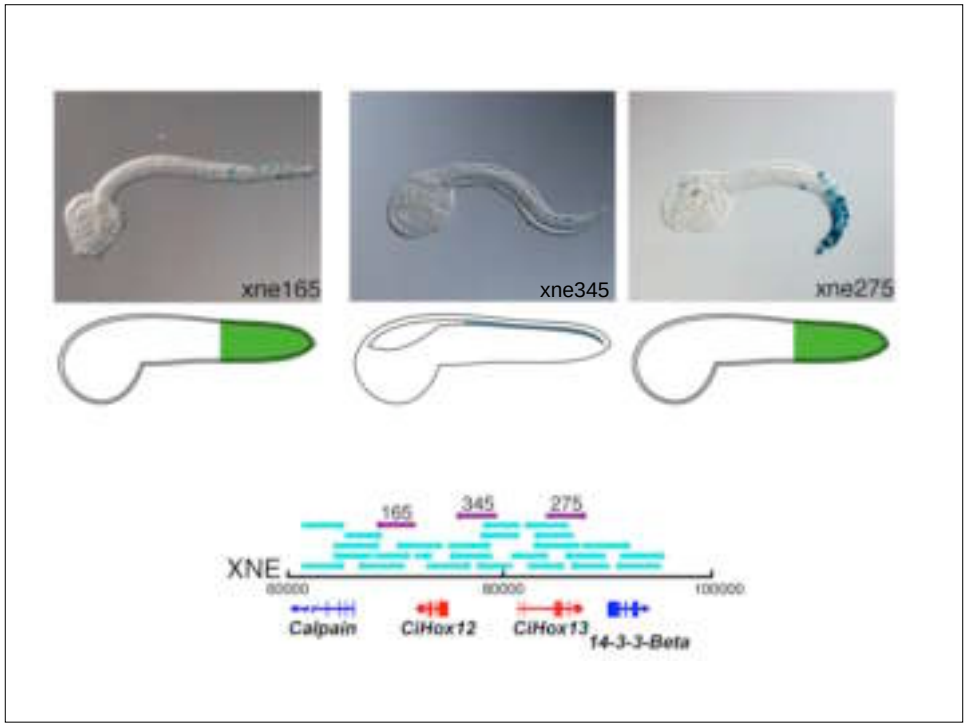


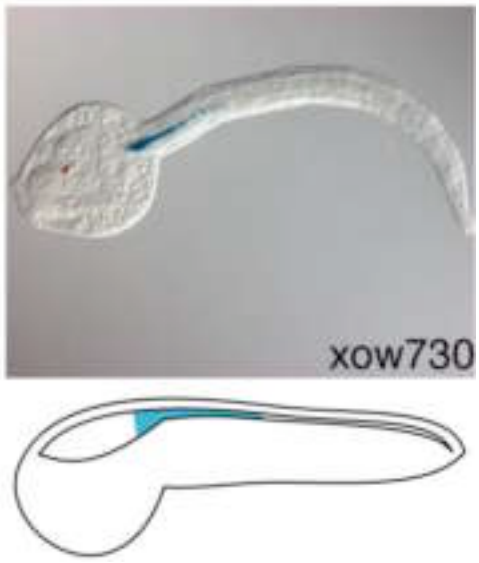
Ciona Hox Complex

- Should be a single *Hox Complex*
  - Correct
- Should be a single domain
  - Wrong, at least 4 separable domains
- Predictable expression patterns




C
xni169
xni338
xni200
xni333
xni337
xni291
xni012
xni178
333
338
200
337
169
234
256
213
261
012
178
XNI
Src containing
CiHox2
CiHox3
CiHox4
Sulfotransferase


xne165
xne345
xne275
165
345
275
XNE
Calpain
CiHox12
CiHox13
14-3-3-Beta



Primary results

- 221 clones electroporated & passed
- 39 clones showed positive signal
- Range for enhancers actually found:
  - Likely Maximum 30
  - Likely real 21
  - Minimum 17
  - Likely *Hox* 08

Hox2 xni338

Hox2 xni333

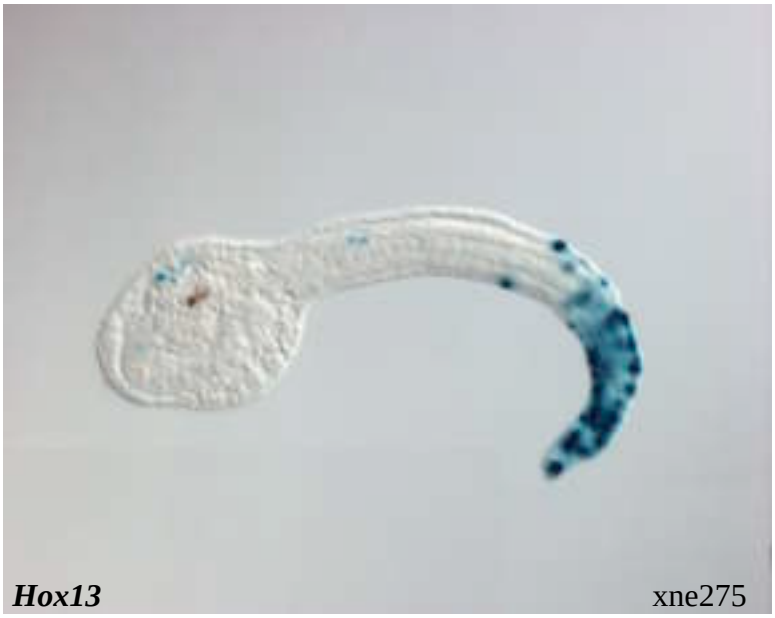
Hox5 xow730

Hox12/13 xne345

Hox2 xni200

Hox3 xni012

Hox12 xne165



Hox13 xne275

Ciona Hox Complex

- Should be a single *Hox* Complex
  - Correct
- Should be a single domain
  - Wrong, At least 4 separable domains
- Predictable expression patterns
  - Correct, Nested CNS
  - Unexpected, Nested Epidermis

Limiting Factors

- DNA preps (50-100ug plasmid)
  - Qiagen Midipreps - up to 48 plasmids per day
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
- Imaging
  - Quality trade offs - Tough decisions

Full Genome Scale Up

Limiting Factors

- DNA preps (50-100ug plasmid)
  - Rolling Circle Amplification
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
  - 480 constructs per week
- Imaging
  - Quality trade offs - Tough decisions
  - Automation??

337 Total Positives
84 Two or More Domains
59 Cerebral Vesicle/Brain
38 Neural Tube
121 Epidermis
154 Muscle
29 Notochord
12 Gut
6 Other
0 Ubiquitous

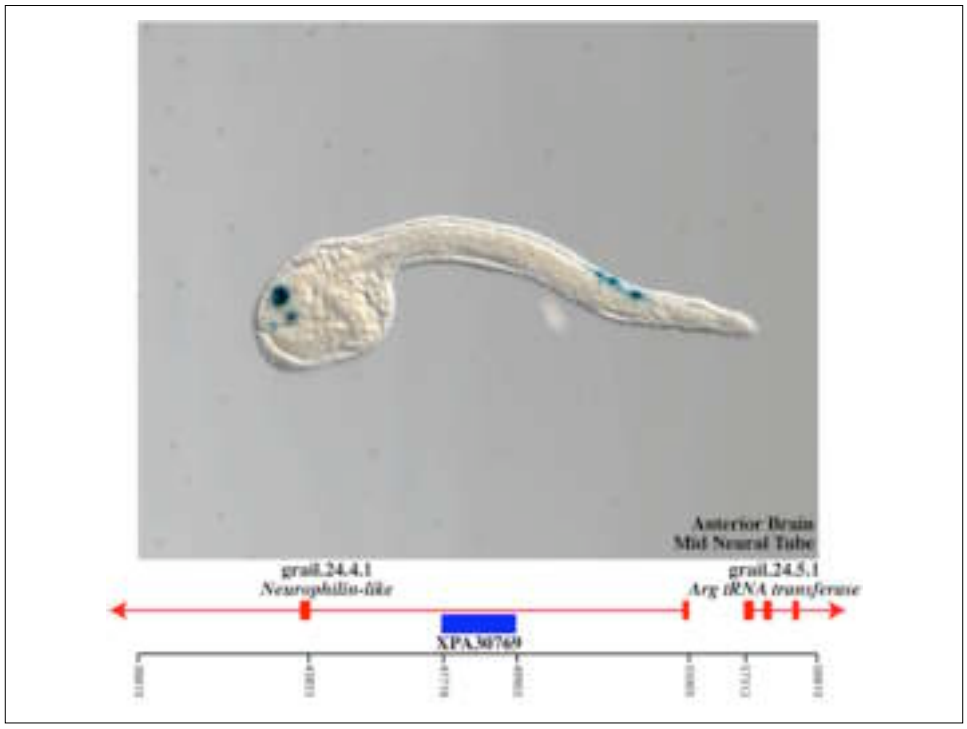
32
82

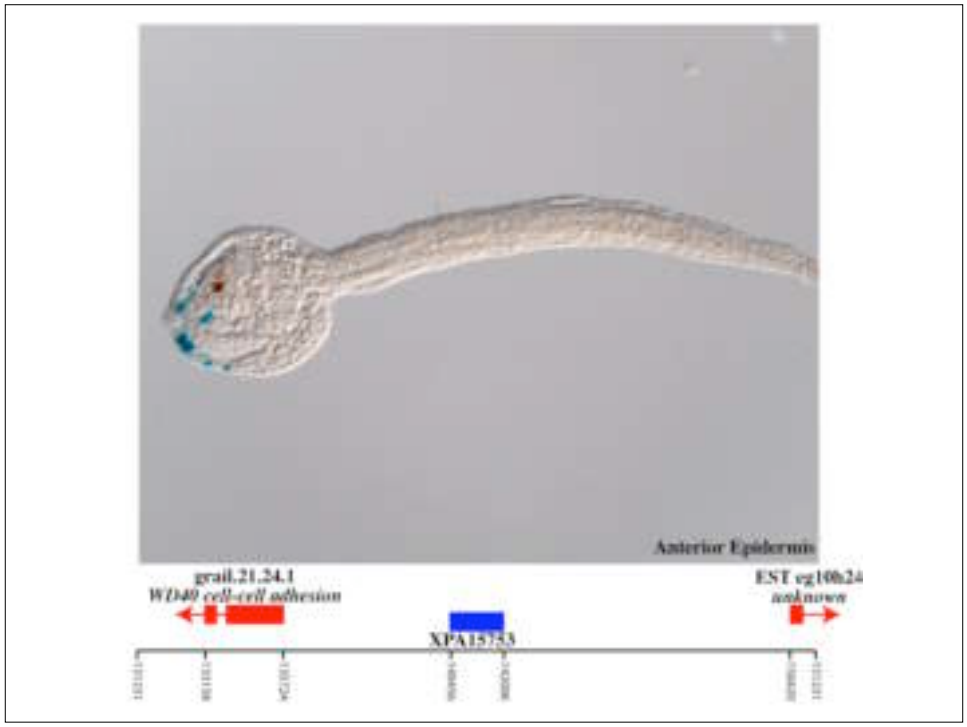
Scale Up

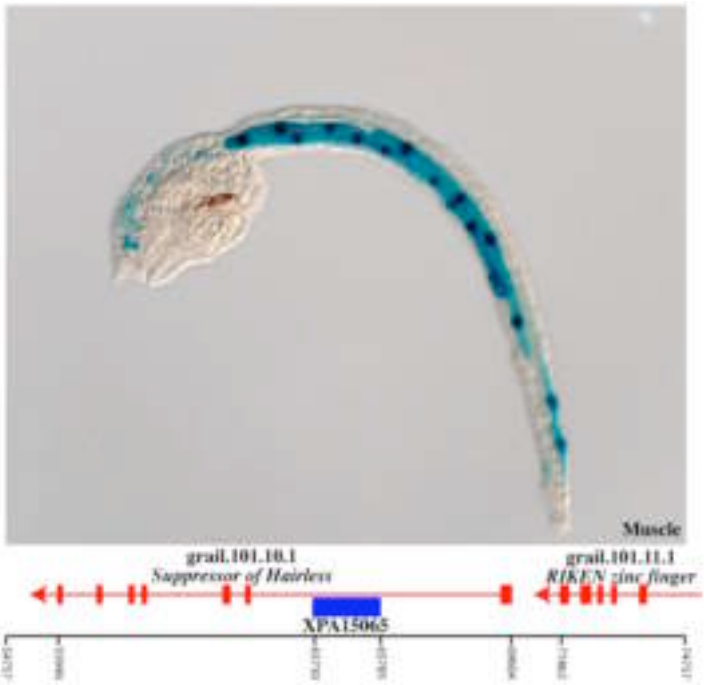
• XPA28186	All epidermis	Hypotheical 109.7 kDa protein
• XPA28213	Tail Muscle	Serine/Threonine Kinase MASK
• XPA28241	Ventral Mid Brain	Homolog to cDNA FLJ10540
• XPA28134	Notochord	Low Sequence quality
• XPA30404	Tail Muscle	RAR Related Steroid Receptor
• XPA30769	Dorsal Brain, Neural Tube	Arginine tRNA protein transferase
• XPA30770	Muscle & Notochord	Proline Oxidase 1
• XPA31107	Post Tail Epidermis	Wnt-2
• XPA28831	All CSN & Epidermis	MORN motif containing
• XPA28492	Single Cell in Brain	unknown but conserved protein
• XPA28855	Post brain & Neural Tube	Protein kinase Ck2-beta
• XPA29631	Neural Tube, All Gut	unknown but conserved protein
• XPA25239	Unknown cells in head	MEC-8 like

Genomic Integration

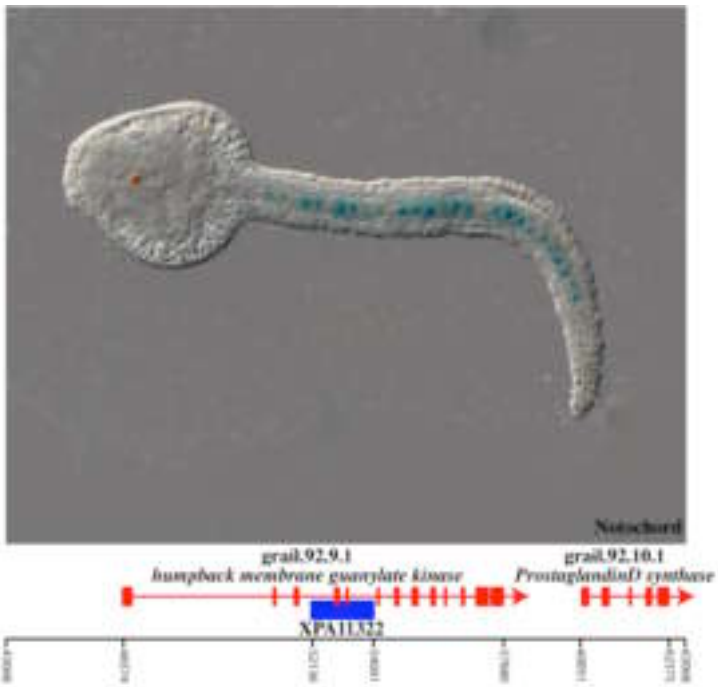
- For most random constructs, 2 end runs will
  - Identify entire subcloned sequence
  - Identify both flanking ORFs
  - Tie into EST in situ project

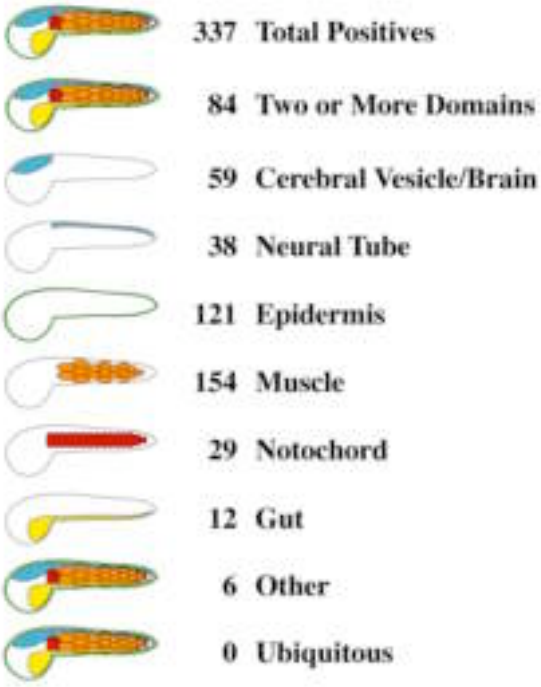

Anterior Brain
Mid Neural Tube
grail.24.4.1
Neurophilin-like
grail.24.5.1
Arg tRNA transferase
XPA.30769


Anterior Epidermis
grail.21.24.1
WD40 cell-cell adhesion
EST eg100b24
unknown
XPA.15753


Muscle
grail.101.10.1
Suppressor of Hairless
grail.101.11.1
RIKEN zinc finger
XPA15065


Epidermis
grail.82.43.1
C2H2 zinc finger
XPA14726





337 Total Positives

84 Two or More Domains

59 Cerebral Vesicle/Brain

38 Neural Tube

121 Epidermis

154 Muscle

29 Notochord

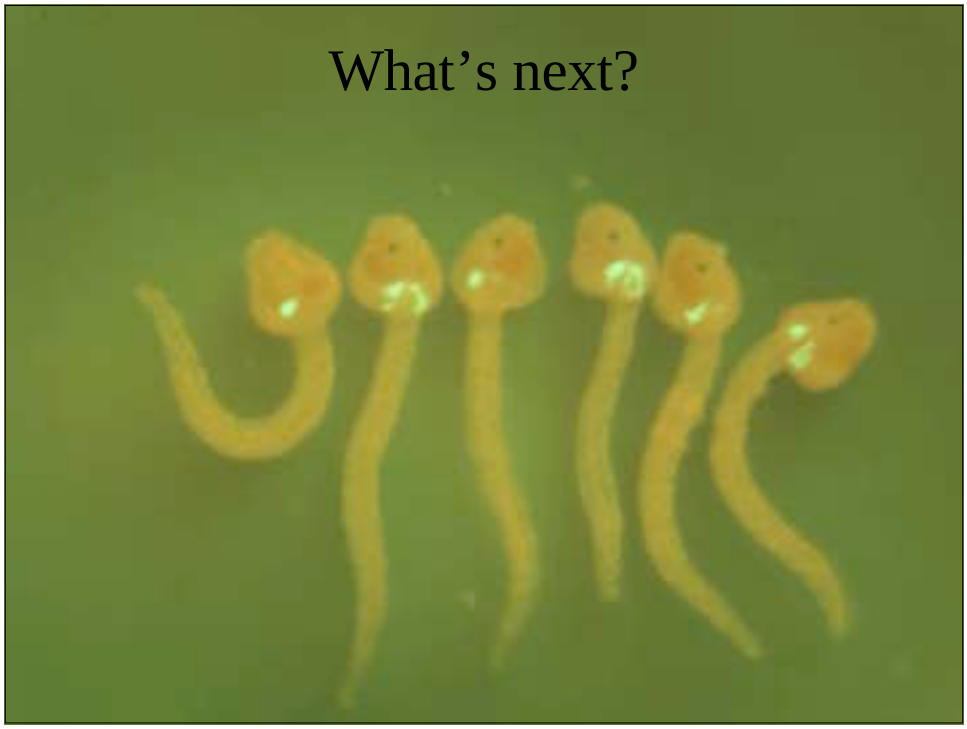
12 Gut

6 Other

0 Ubiquitous

32

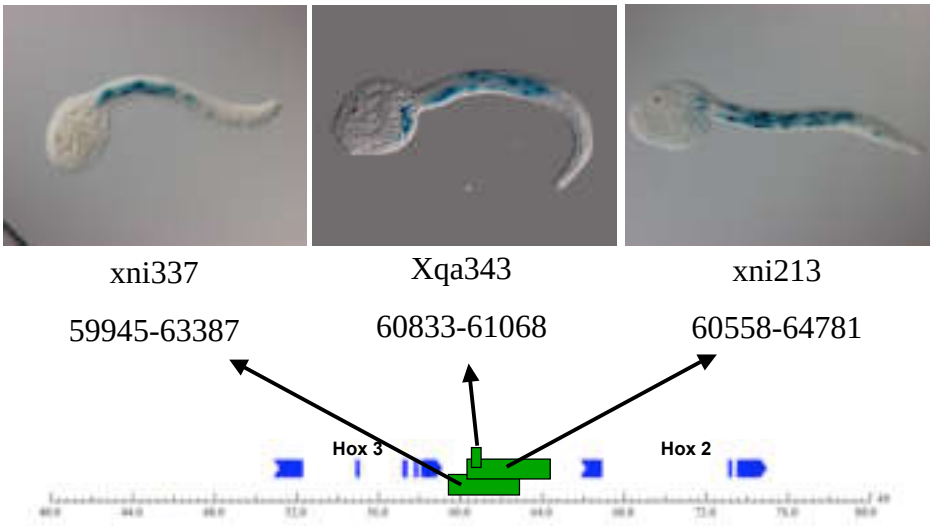
What's next?



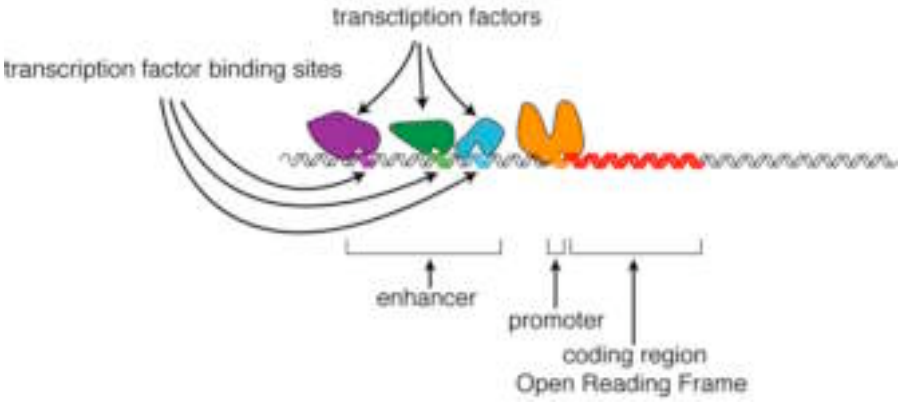
Potential Issues

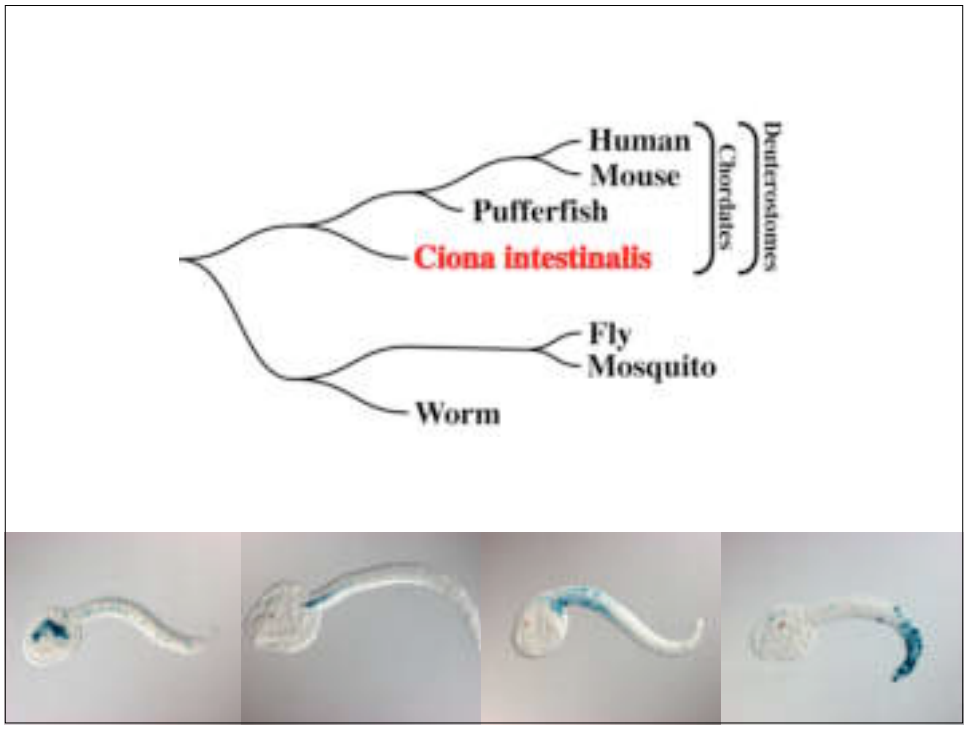
- Promoter specificity
- Insulators & repressors
- Enhancer Polarity
- Promoter competition
- Enhancers fragmented during cloning

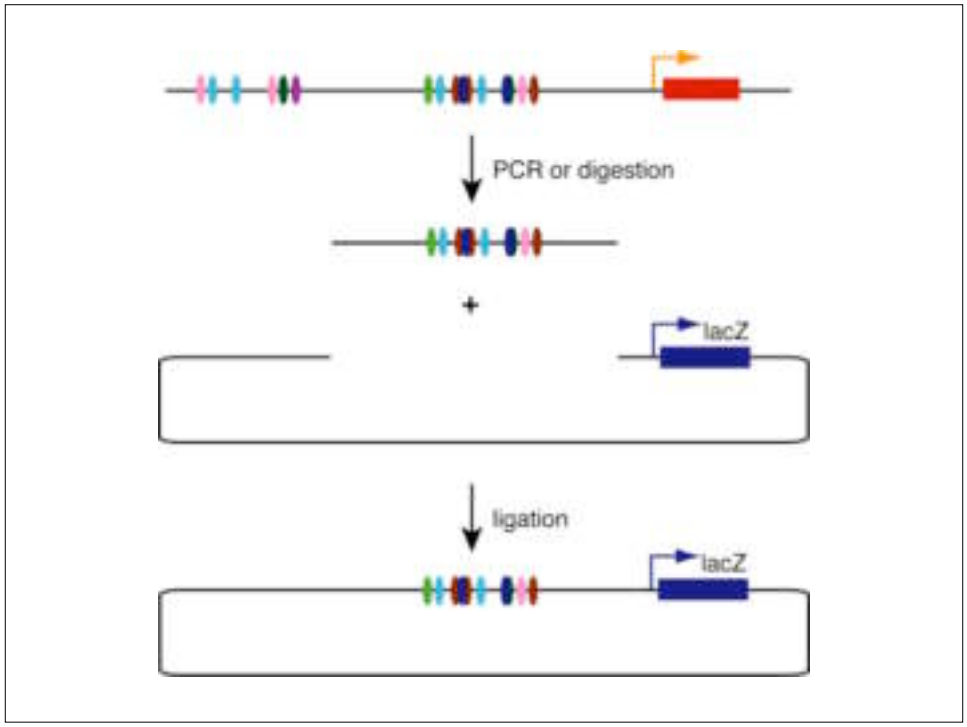
- Timing
- Insufficient detection strength

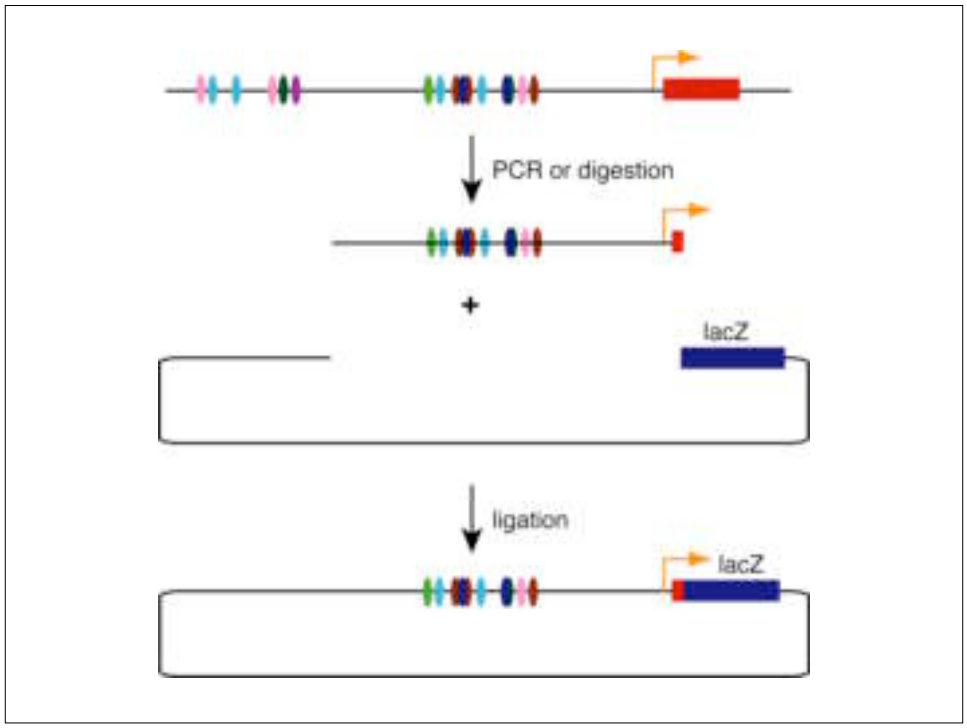


Characterize functional transcription factor binding sites




Human
Mouse
Pufferfish
Ciona intestinalis
Fly
Mosquito
Worm
Chordates
Deuterostomes


PCR or digestion
+
lacZ
ligation
lacZ



A Functional Genomics Approach to Developmental Genetics

Joint Genome Institute

LBNL - LLNL - DOE

David Keys
Byung-in Lee
Chris Detter
Stephan Trong

Dan Rokhsar
Paul Richardson

Sylvia Ahn
Dave Engle
Orsalem kahsai
Maria Shin
Joann Wang
Mei Wang

University of California Berkeley

Naoe Harafuji
Anna Di Gregorio
Mike Levine