mod/ENCODE Integrative Comparison

Worm, Fly, and Human

Chromatin

Regulation

Transcription

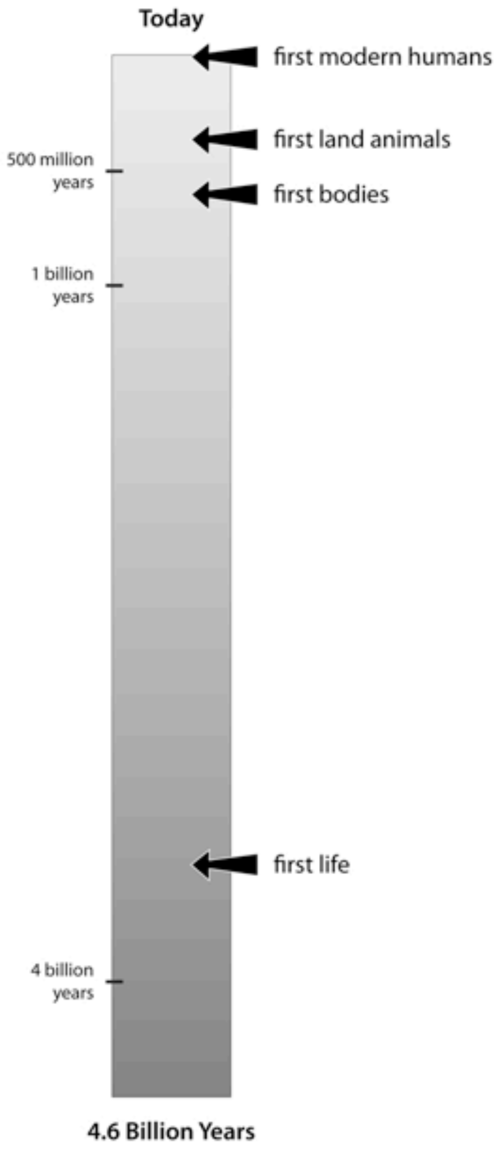
Roger Alexander
Gerstein lab group meeting

23 August 2012

from Neil Shubin's

Your Inner Fish

A Journey into the 3.5-Billion-Year History of the Human Body

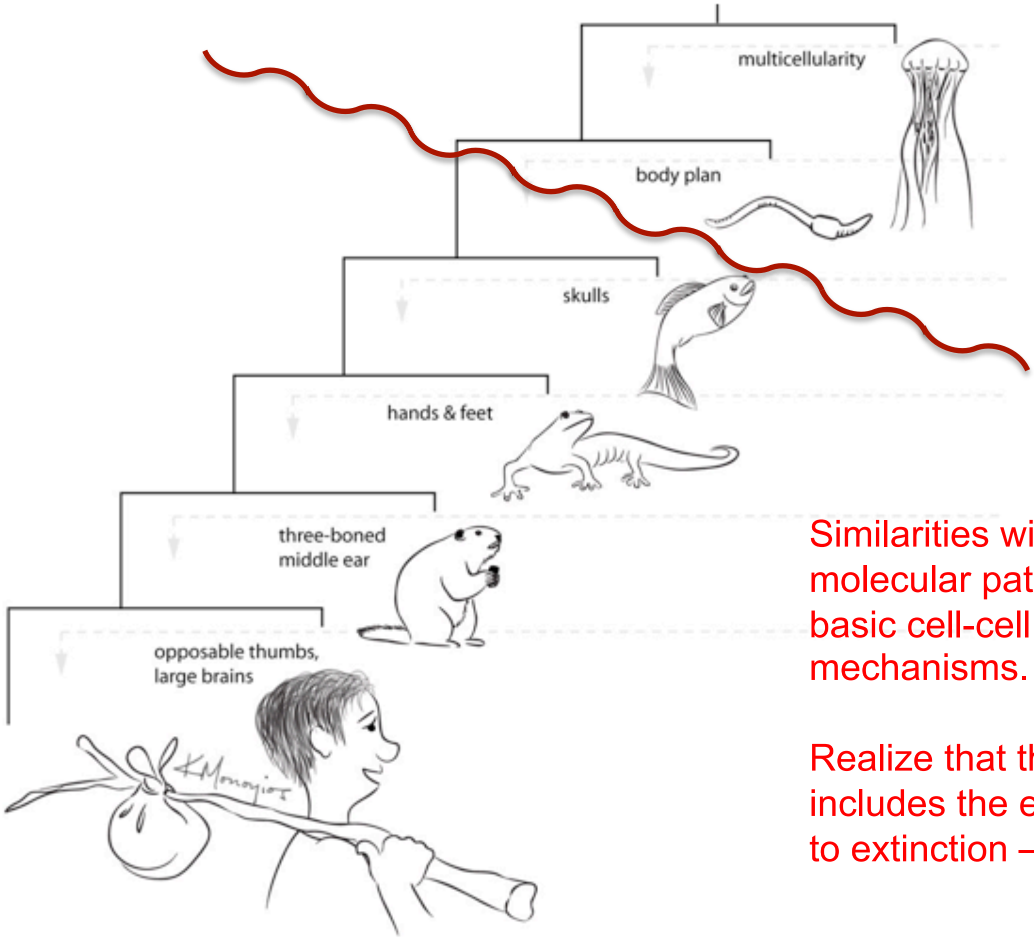


What do worms, flies, and humans have in common?

from Neil Shubin's

Your Inner Fish

A Journey into the 3.5-Billion-Year History of the Human Body



Similarities will be very basal – molecular pathways, germ layers, basic cell-cell interaction mechanisms.

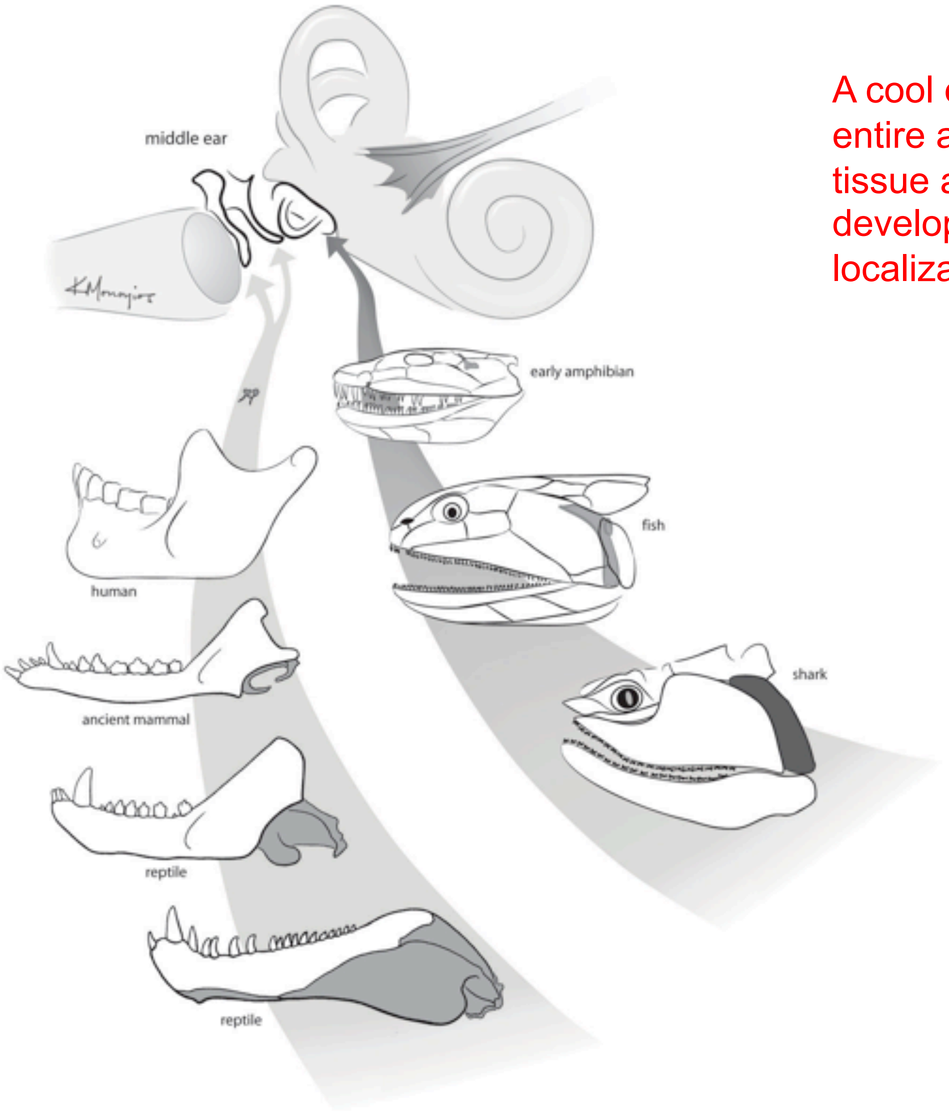
Realize that the human branch includes the entire history – birth to extinction – of the dinosaurs.

A human family tree, all the way back to jellyfish. It has the same structure as the one for the bozos.

from Neil Shubin's

Your Inner Fish

A Journey into the 3.5-Billion-Year History of the Human Body



A cool example of heterochrony – entire anatomical structures / tissue assemblages can shift in developmental timing and spatial localization.

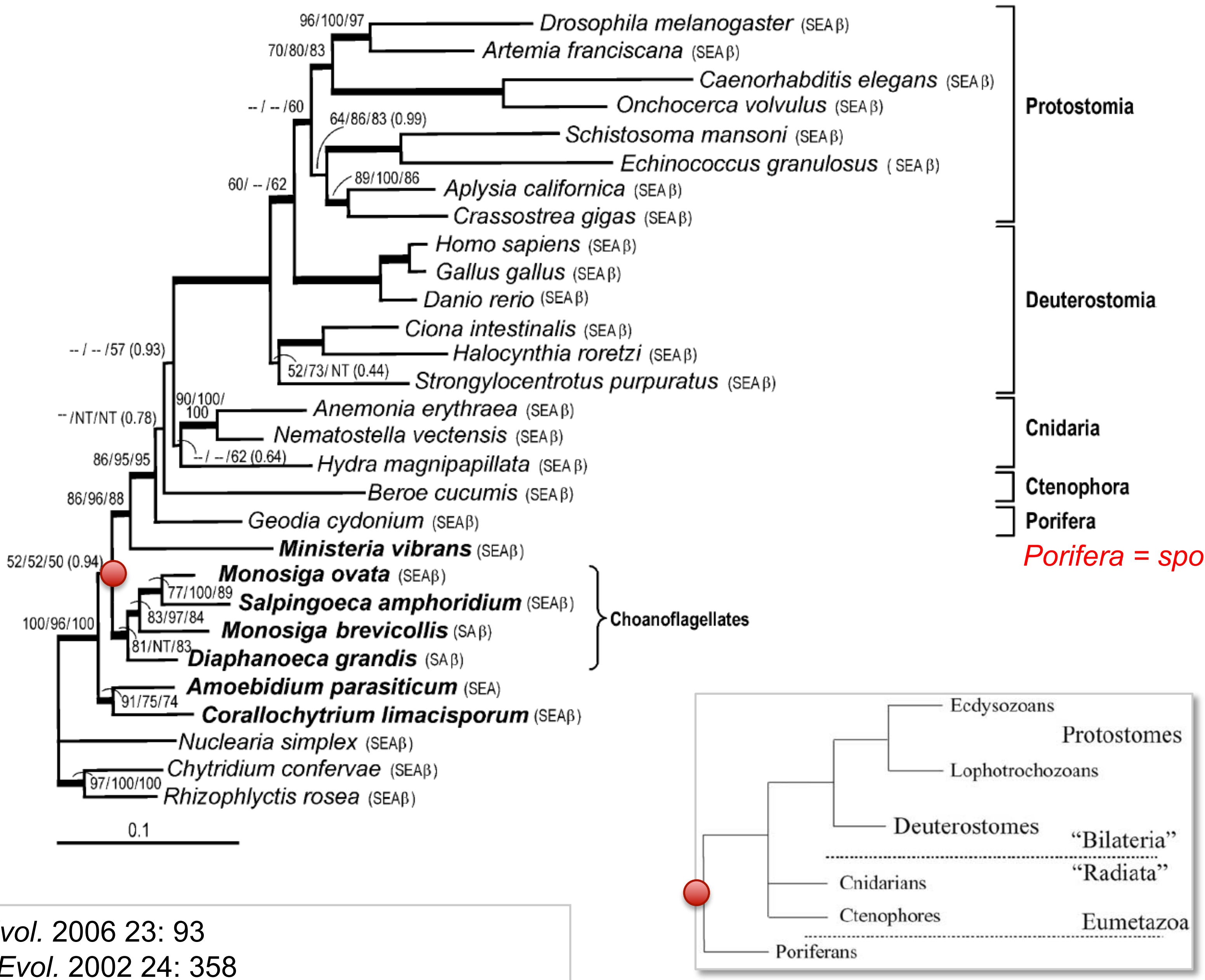
We can trace bones from gill arches to our ears, first during the transition from fish to amphibian (right), and later during the shift from reptile to mammal (left).

What do worms, flies, and humans have in common?

- They are all animals (metazoans).
- They are multicellular.
- They are triploblast – i.e. they have three germ layers
  - endoderm
  - mesoderm
  - ectoderm
- They are bilaterian – i.e. bilaterally symmetric.

- BUT humans are deuterostomes, while worms and flies are protostomes.

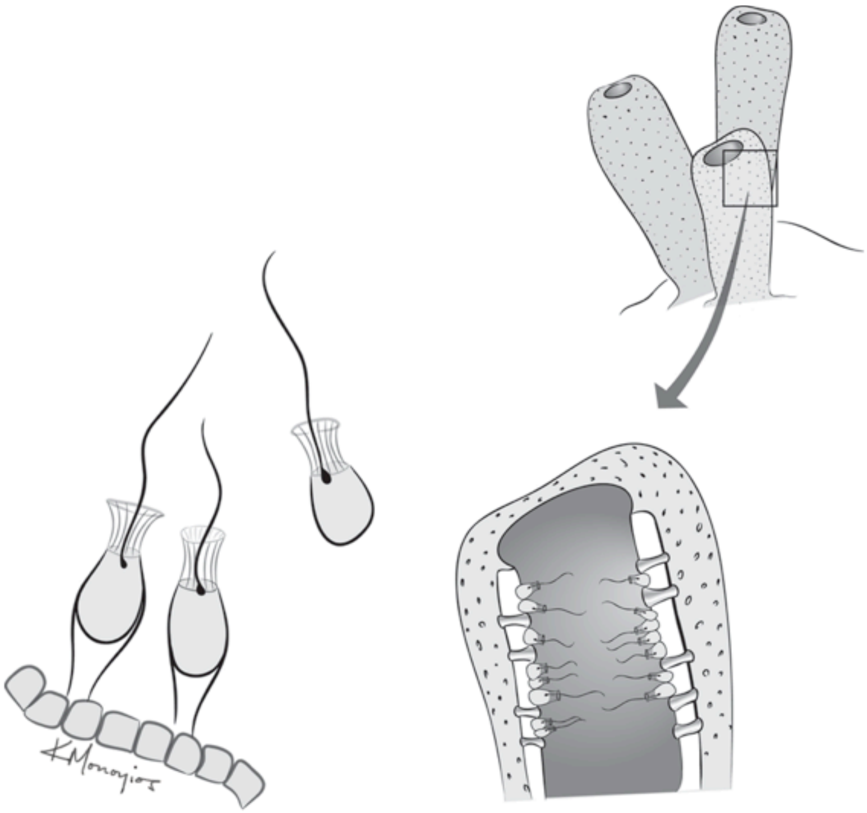
Animal Phylogeny – Origin of Multicellularity ●


96/100/97
70/80/83
-- / -- /60
64/86/83 (0.99)
89/100/86
60/ -- /62
-- / -- /57 (0.93)
52/73/ NT (0.44)
90/100/ 100
-- /NT/NT (0.78)
-- / -- /62 (0.64)
86/95/95
86/96/88
52/52/50 (0.94)
77/100/89
83/97/84
81/NT/83
100/96/100
91/75/74
97/100/100
Drosophila melanogaster (SEA β)
Artemia franciscana (SEA β)
Caenorhabditis elegans (SEA β)
Onchocerca volvulus (SEA β)
Schistosoma mansoni (SEA β)
Echinococcus granulosus (SEA β)
Aplysia californica (SEA β)
Crassostrea gigas (SEA β)
Homo sapiens (SEA β)
Gallus gallus (SEA β)
Danio rerio (SEA β)
Ciona intestinalis (SEA β)
Halocynthia roretzi (SEA β)
Strongylocentrotus purpuratus (SEA β)
Anemonia erythraea (SEA β)
Nematostella vectensis (SEA β)
Hydra magnipapillata (SEA β)
Beroe cucumis (SEA β)
Geodia cydonium (SEAβ)
Ministeria vibrans (SEAβ)
Monosiga ovata (SEAβ)
Salpingoeca amphoridium (SEAβ)
Monosiga brevicollis (SA β)
Diaphanoeca grandis (SA β)
Amoebidium parasiticum (SEA)
Corallochytrium limacisporum (SEAβ)
Nuclearia simplex (SEAβ)
Chytridium confervae (SEAβ)
Rhizophlyctis rosea (SEAβ)
0.1
Protostomia
Deuterostomia
Cnidaria
Ctenophora
Porifera
Choanoflagellates
Ecdysozoans
Protostomes
Lophotrochozoans
Deuterostomes
"Bilateria"
"Radiata"
Cnidarians
Ctenophores
Eumetazoa
Poriferans

Porifera = sponges

Mol. Biol. Evol. 2006 23: 93
Mol. Phylo Evol. 2002 24: 358

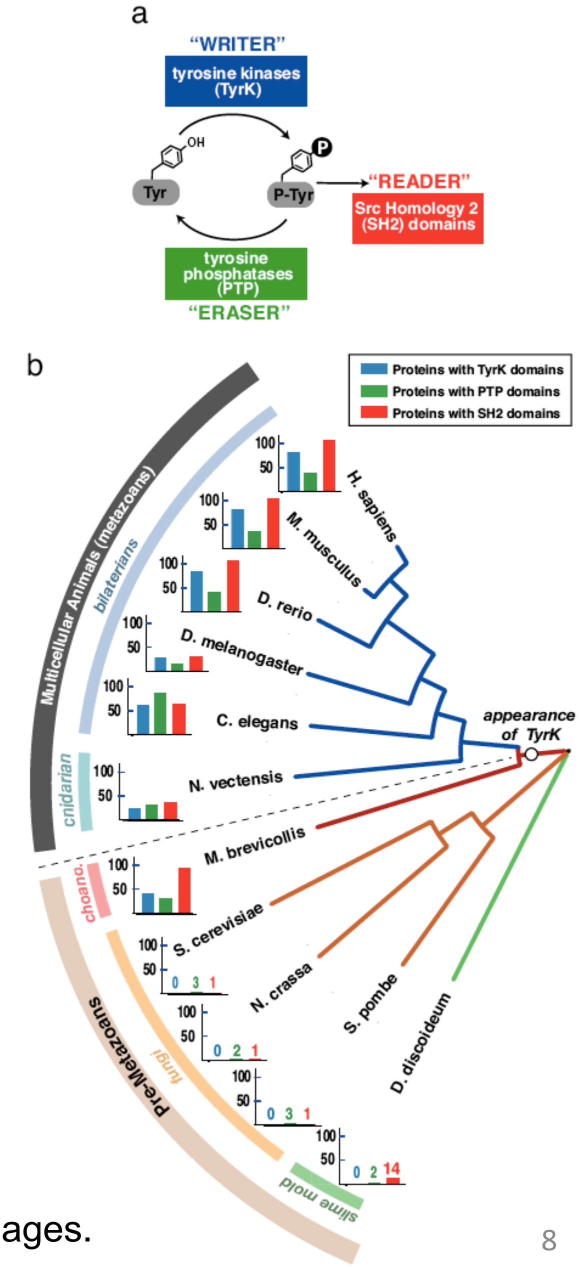
Animal Phylogeny – Origin of Multicellularity



Choanoglagellate
(e.g. *M. brevicolis)*

Sponge
(e.g. *N. vectensis)*

Illustration © Kalliopi Monoyios



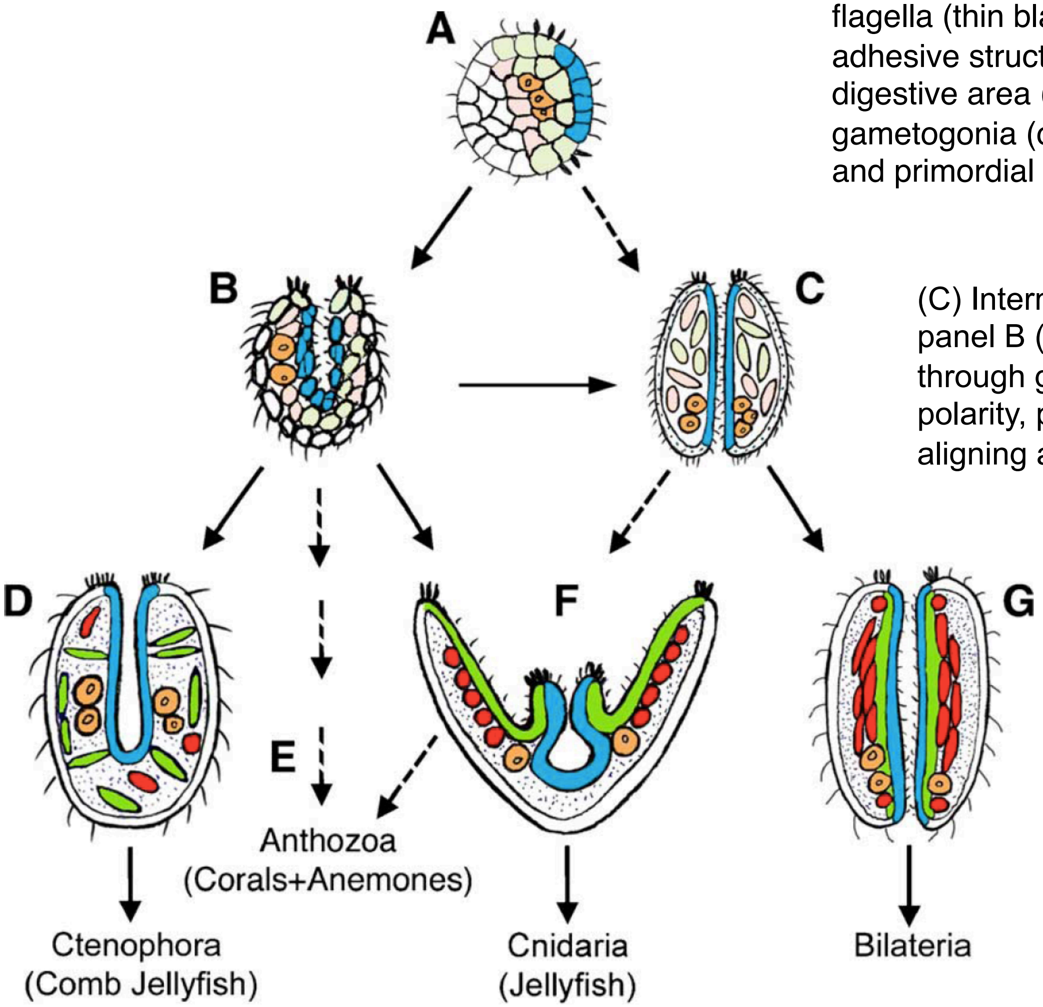
Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages.
PNAS 2008 105: 9680

What do worms, flies, and humans have in common?

- They are all animals (metazoans).
- They are multicellular.
- They are triploblast – i.e. they have three germ layers
  - endoderm
  - mesoderm
  - ectoderm
  - Diploblasts lack mesoderm.
- They are bilaterian – i.e. bilaterally symmetric.

- BUT humans are deuterostomes, while worms and flies are protostomes.

Triploblasty



(A) Ancestral metazoan with
flagella (thin black lines),
adhesive structures (thick black spikes),
digestive area (blue),
gametogonia (orange),
and primordial myocytes (light green and
light red)

(C) Intermediate stage formed from panel B (or from panel A). It has a through gut and anterior– posterior polarity, primordial myocytes start aligning along the digestive tube.

(D) massive extracellular matrix (ECM) has evolved; most myocytes differentiated into smooth muscle type (green)

(F) Radial animal with central gut and striated muscle (red).

(G) Zootype ancestor with digestive tube.

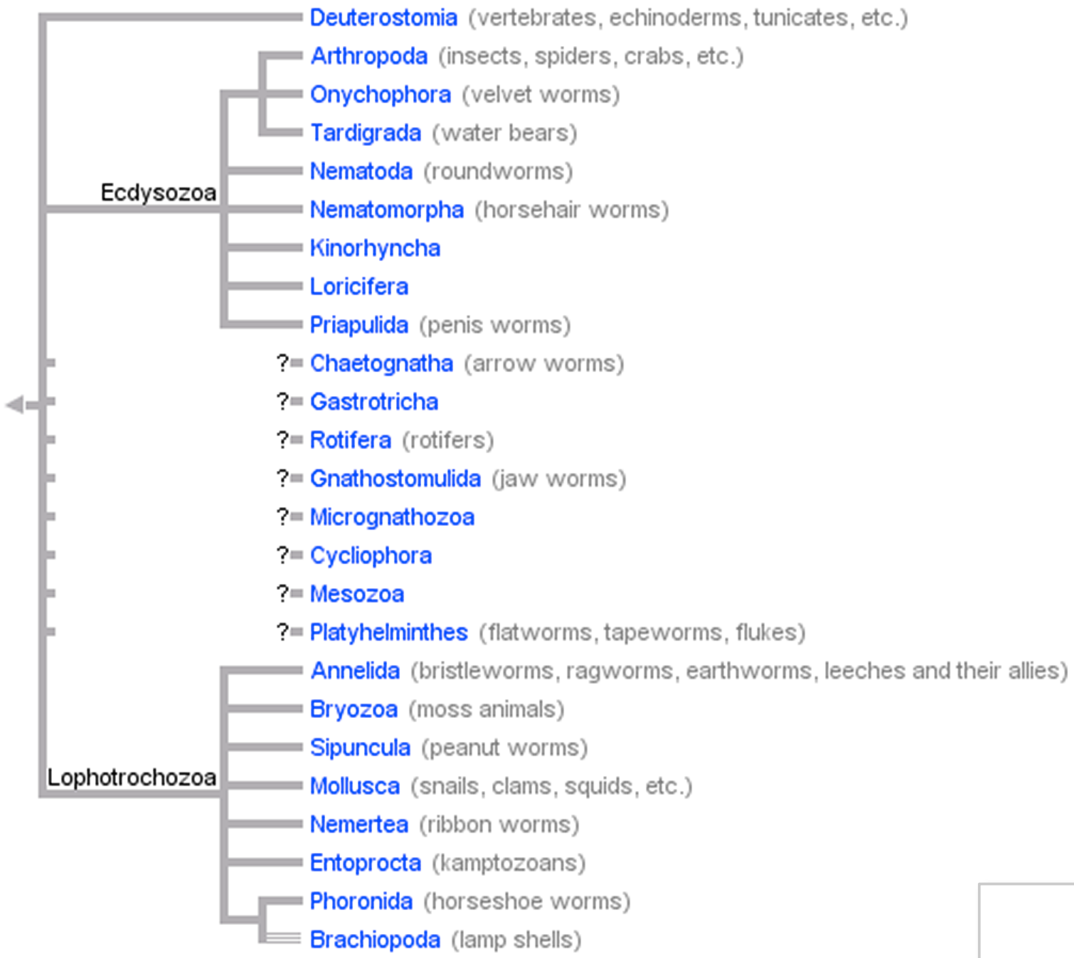
Evolution of striated muscle: Jellyfish and the origin of triploblasty
Devel. Biol. 2005 282: 14

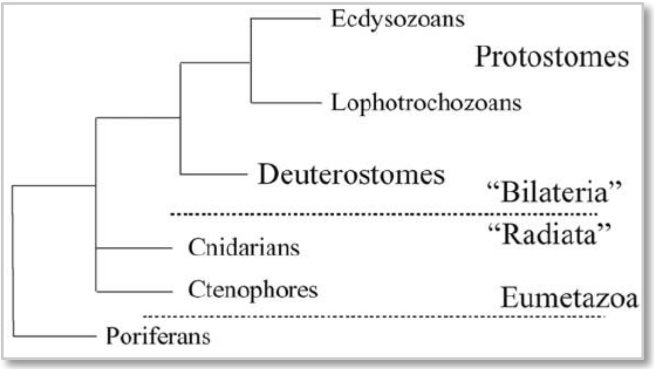
What do worms, flies, and humans have in common?

- They are all animals (metazoans).
- They are multicellular.
- They are triploblast – i.e. they have three germ layers
  - endoderm
  - mesoderm
  - ectoderm
- They are bilaterian – i.e. bilaterally symmetric.
- They have a zootypic stage, i.e. a body plan built from HOX gene regulatory networks
- and they have a phylotypic stage

- BUT humans are deuterostomes, while worms and flies are protostomes.

Animal Phylogeny – Bilateria


Deuterostomia (vertebrates, echinoderms, tunicates, etc.)
Arthropoda (insects, spiders, crabs, etc.)
Onychophora (velvet worms)
Tardigrada (water bears)
Ecdysozoa
Nematoda (roundworms)
Nematomorpha (horsehair worms)
Kinorhyncha
Loricifera
Priapulida (penis worms)
?
Chaetognatha (arrow worms)
?
Gastrotricha
?
Rotifera (rotifers)
?
Gnathostomulida (jaw worms)
?
Micrognathozoa
?
Cycliophora
?
Mesozoa
?
Platyhelminthes (flatworms, tapeworms, flukes)
Annelida (bristleworms, ragworms, earthworms, leeches and their allies)
Bryozoa (moss animals)
Sipuncula (peanut worms)
Lophotrochozoa
Mollusca (snails, clams, squids, etc.)
Nemertea (ribbon worms)
Entoprocta (kamptozoans)
Phoronida (horseshoe worms)
Brachiopoda (lamp shells)

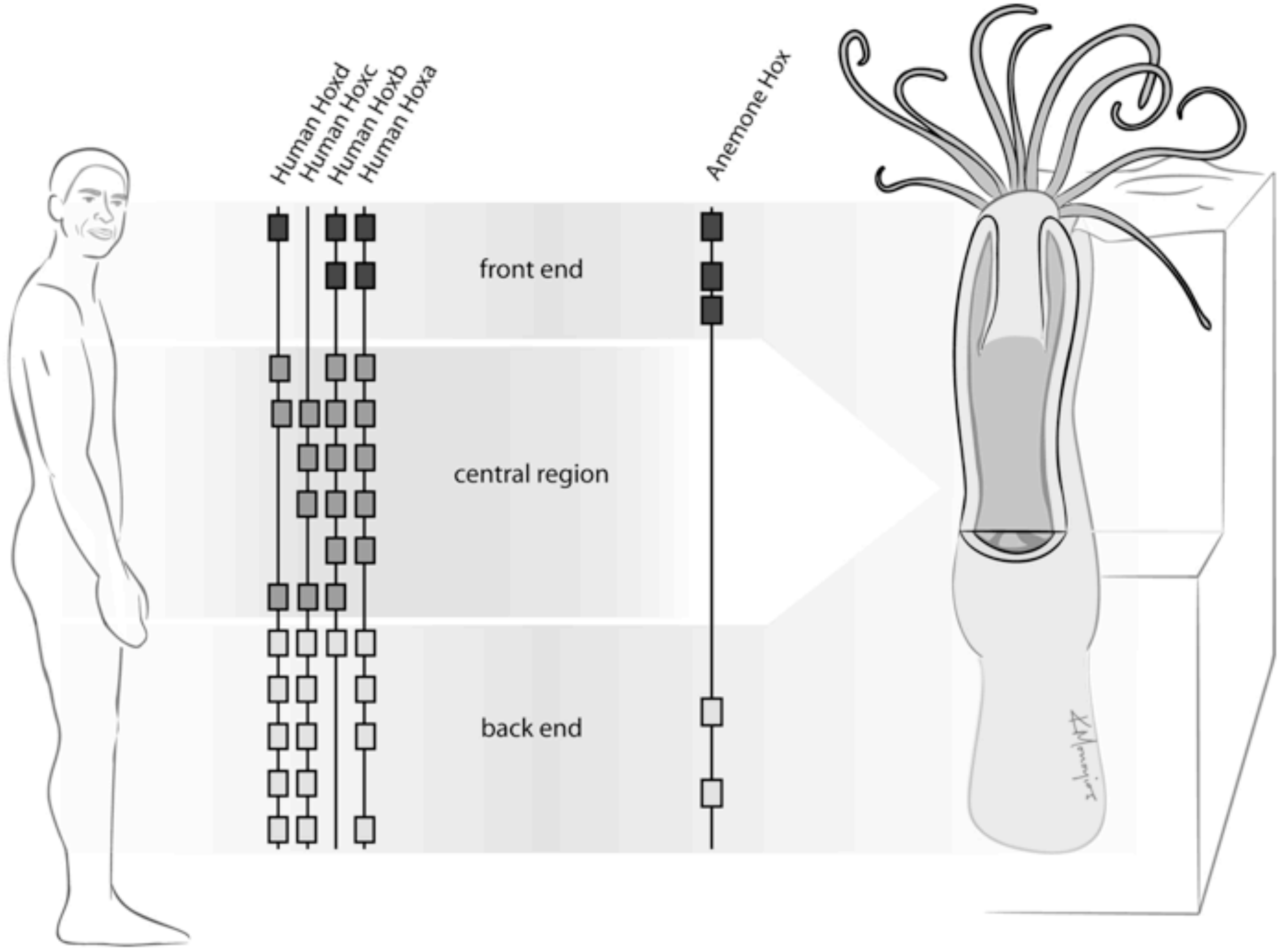


Tree of Life web (http://tolweb.org/Bilateria)
Mol. Phylo Evol. 2002 24: 358

from Neil Shubin's

Your Inner Fish

A Journey into the 3.5-Billion-Year History of the Human Body



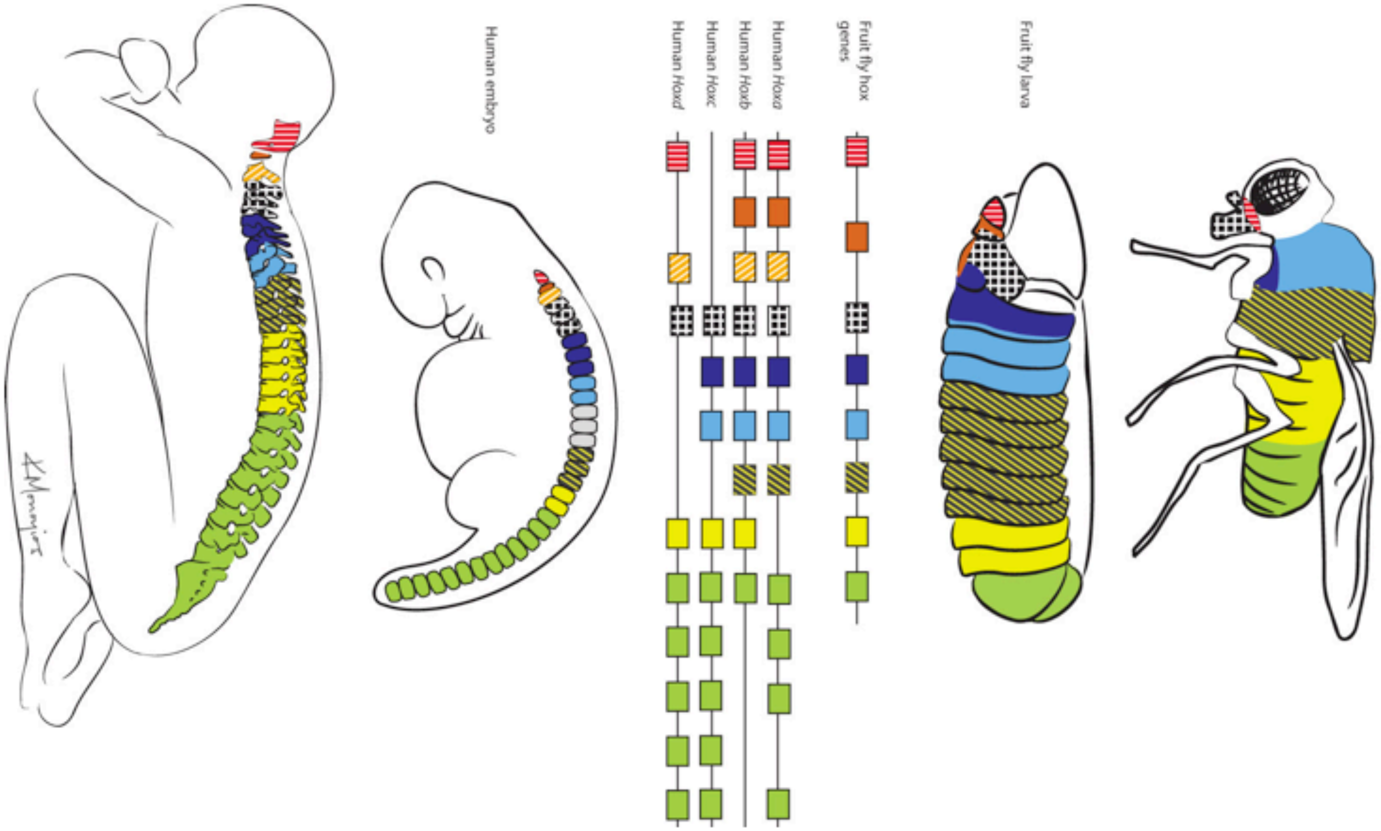
Jellyfish relatives, such as sea anemones, have a front and a back as we do, a body plan set up by versions of the same genes.

Text © Neil Shubin Illustration © Kalliopi Monoyios

from Neil Shubin's

Your Inner Fish

A Journey into the 3.5-Billion-Year History of the Human Body



Hox genes in flies and people. The head-to-tail organization of the body is under the control of different *Hox* genes. Flies have one set of eight hox genes, each represented as a little box in the diagram. Humans have four set of these genes. In flies and people, the activity of a gene mtches its position on th eDNA: genes active in the head lie at one end, those in the ail at anoher, with genes affecting the middle of the body lying in between.

What do worms, flies, and humans have in common?

- They are all animals (metazoans).
- They are multicellular.
- They are triploblast – i.e. they have three germ layers
  - endoderm
  - mesoderm
  - ectoderm
- They are bilaterian – i.e. bilaterally symmetric.
- They have a zootypic stage, i.e. a body plan built from HOX gene regulatory networks and they have a phylotypic stage

- BUT humans are deuterostomes, while worms and flies are protostomes.

Phylotypic Stage

- stage of development at which all major body parts are represented in their final positions as undifferentiated cell condensations
- OR the stage after the completion of the principal morphogenetic tissue movements
- OR the stage at which all members of the phylum show the maximum degree of similarity

- vertebrates: tailbud stage
- insects: fully segmented germband stage
- leeches: fully segmented, ventrally closed stage
- nematode after the completion of most embryonic cell divisions

- The phylotypic stage is NOT the earliest stage – variability of early stages may result from adaptation to particular types of reproductive strategy or to the demands of embryonic nutrition.

Nature 1993 361: 490

Universality of HOX genes

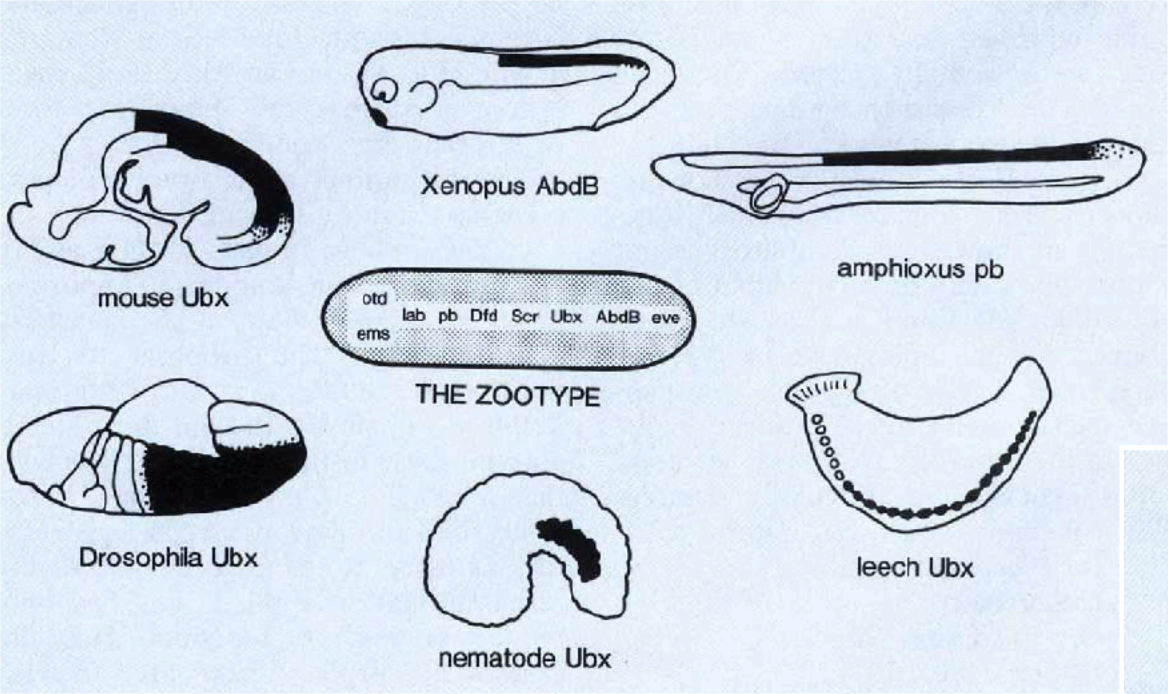
“The amphioxus-vertebrate comparison suggests that the vertebrate head is homologous to the anterior, but not cephalized, segments of the lower chordate.”

“HOX cluster genes really do seem to encode relative position within the organism rather than any specific structure, and the patterns are conserved despite major shifts in other developmental mechanisms.”

“HOX cluster genes are also present in Hydra (phylum Cnidaria).”

Nat. Comm. 2011 2: 248

HOX genes and the phylotypic stage



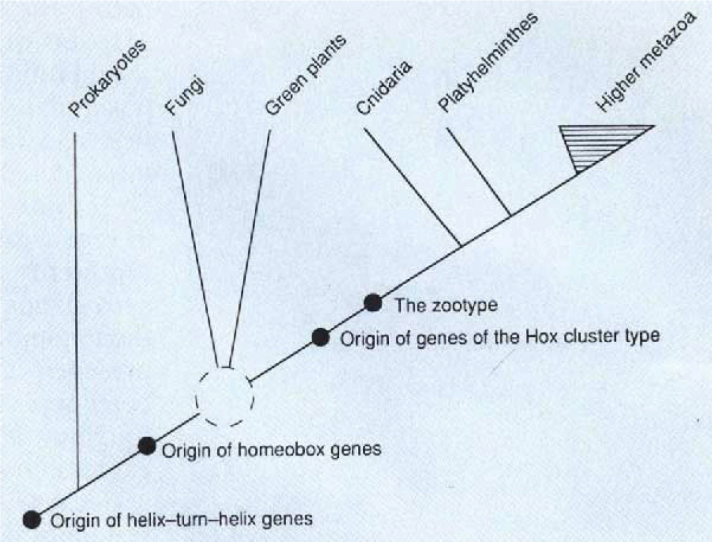


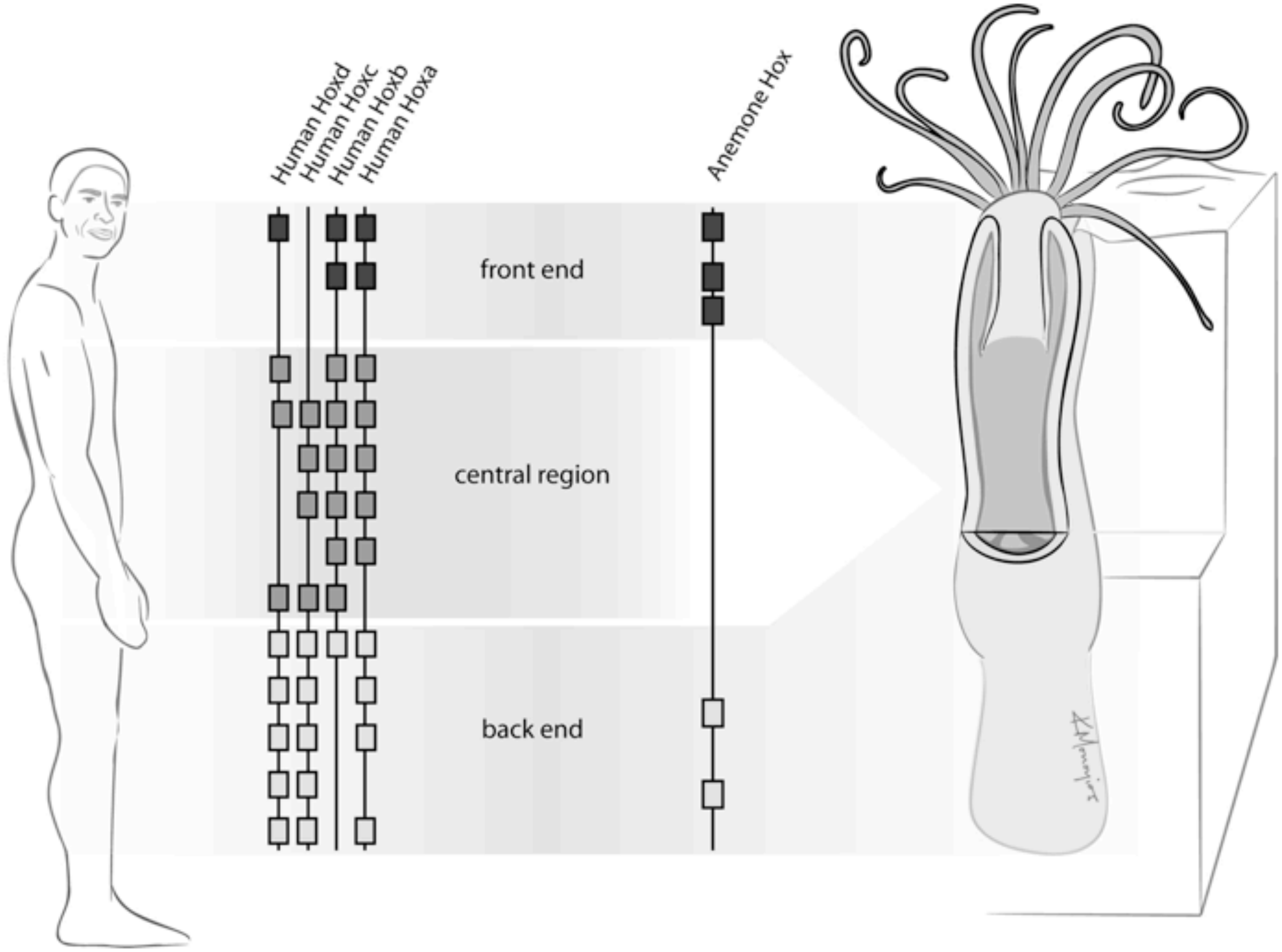
FIG. 4 Origin of the zootype on the evolutionary tree. The Hox cluster genes are a subset of the homeobox genes, which are in turn a subset of genes encoding DNA-binding proteins of the helix–turn–helix class.

The zootype and the phylotypic stage.
Nature (1993) 361: 490

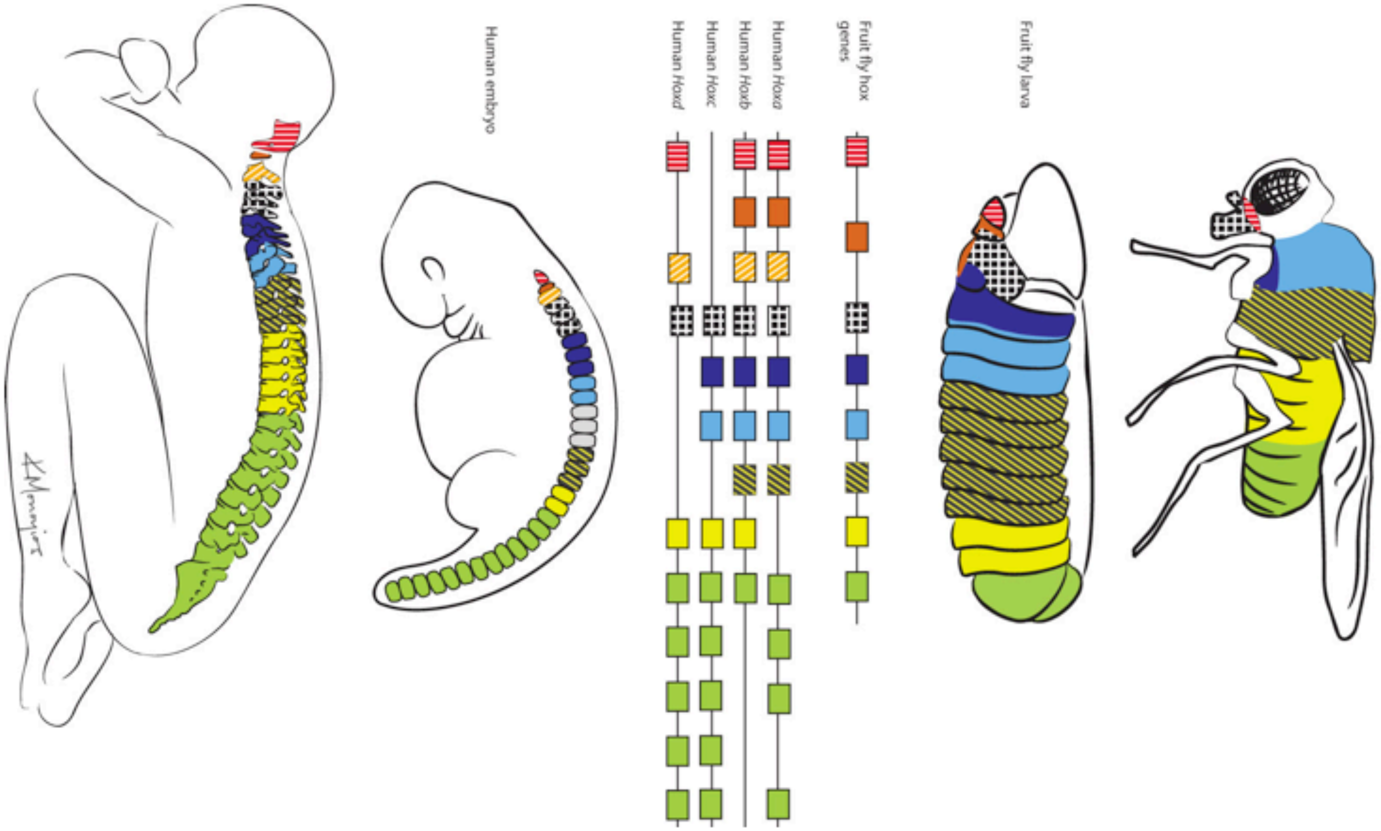
from Neil Shubin's

Your Inner Fish

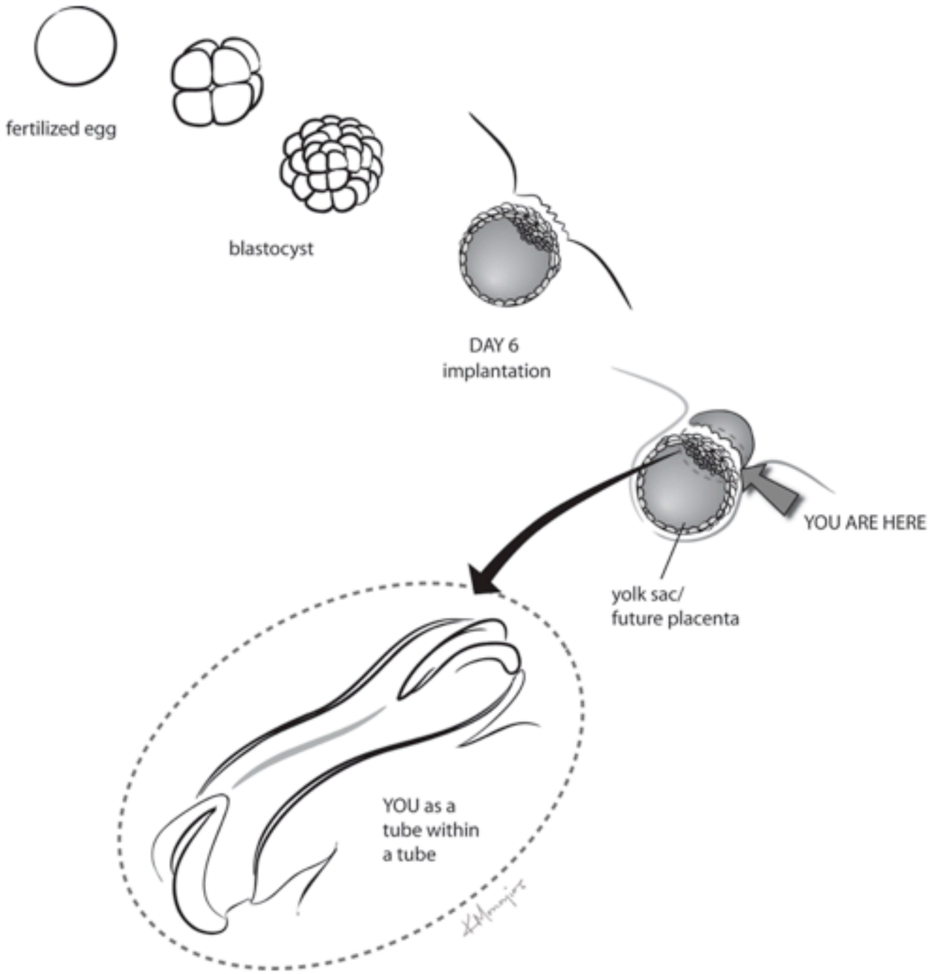
A Journey into the 3.5-Billion-Year History of the Human Body



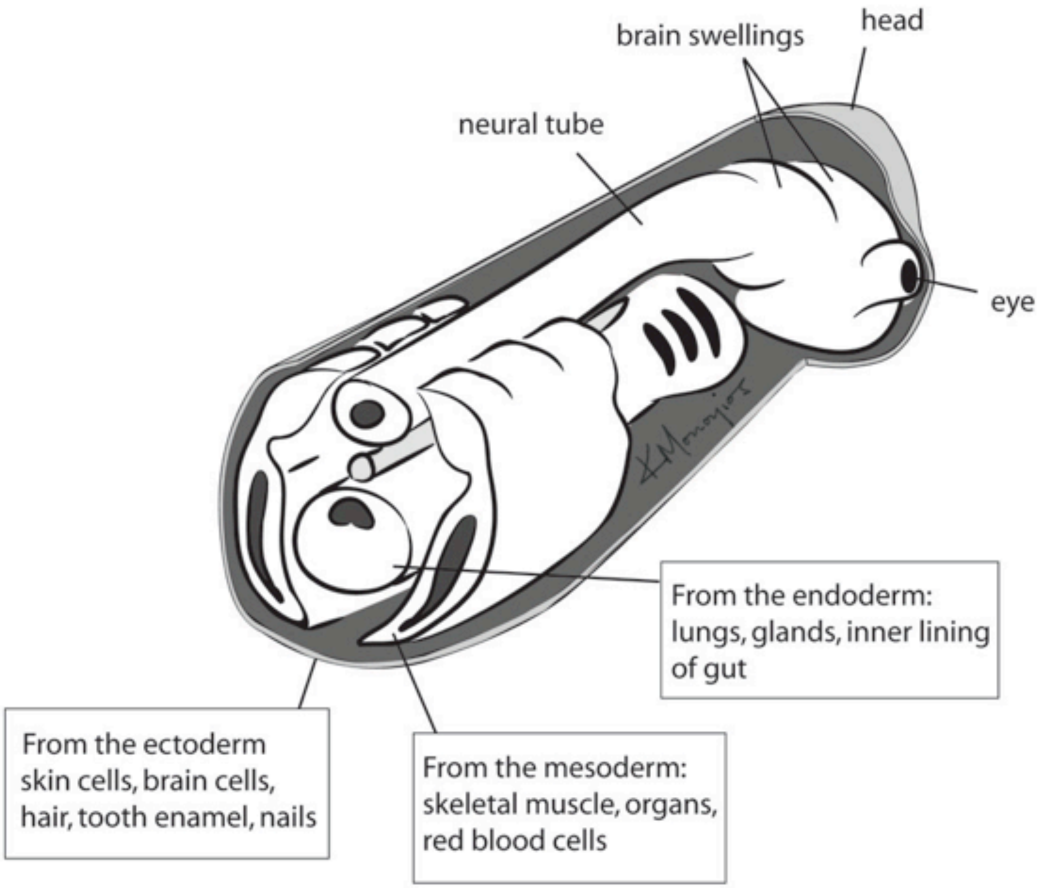
Jellyfish relatives, such as sea anemones, have a front and a back as we do, a body plan set up by versions of the same genes.



Hox genes in flies and people. The head-to-tail organization of the body is under the control of different *Hox* genes. Flies have one set of eight hox genes, each represented as a little box in the diagram. Humans have four set of these genes. In flies and people, the activity of a gene mtches its position on th eDNA: genes active in the head lie at one end, those in the ail at anoher, with genes affecting the middle of the body lying in between.



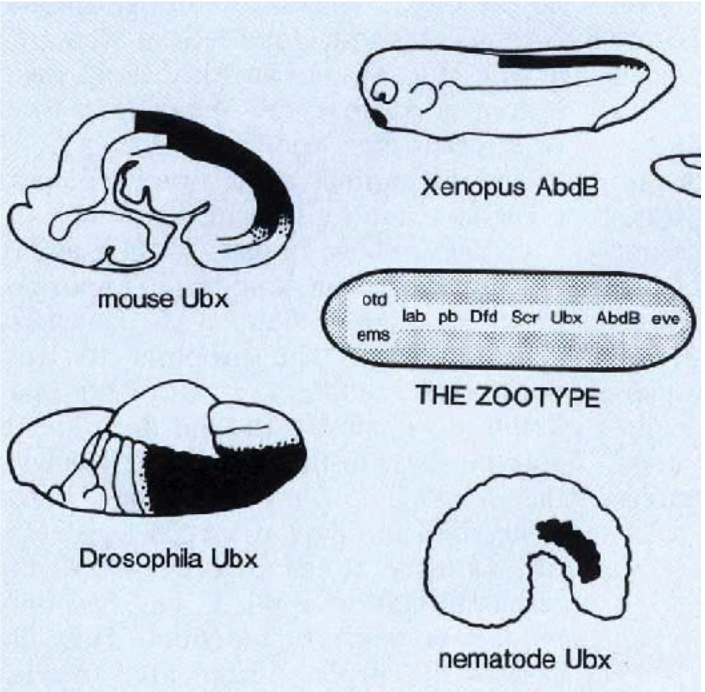
Our early days, the first three weeks after conception. We go from being a single cell to a ball of cells and end up as a tube.



From the endoderm: lungs, glands, inner lining of gut

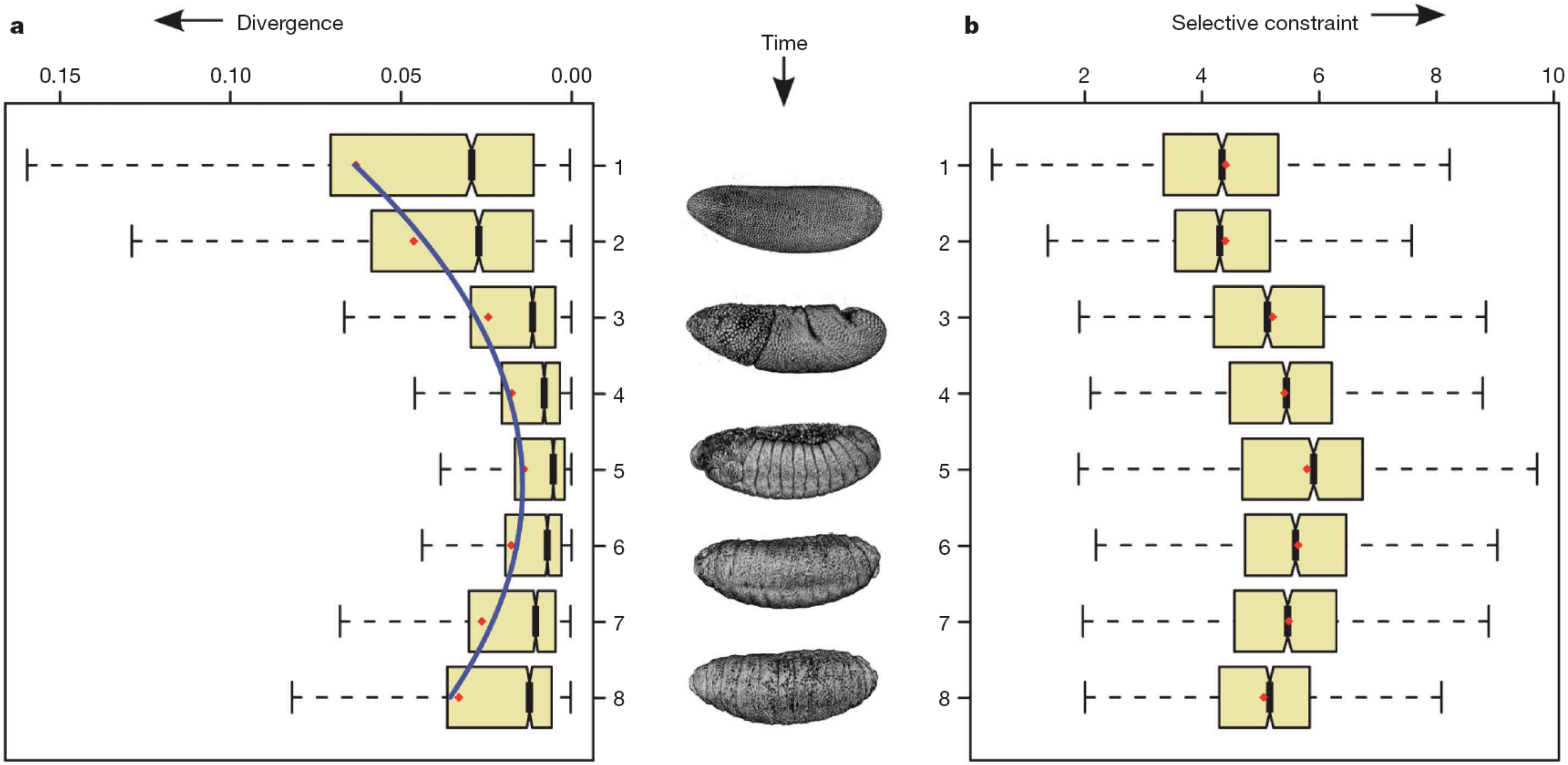
From the ectoderm skin cells, brain cells, hair, tooth enamel, nails

From the mesoderm: skeletal muscle, organs, red blood cells



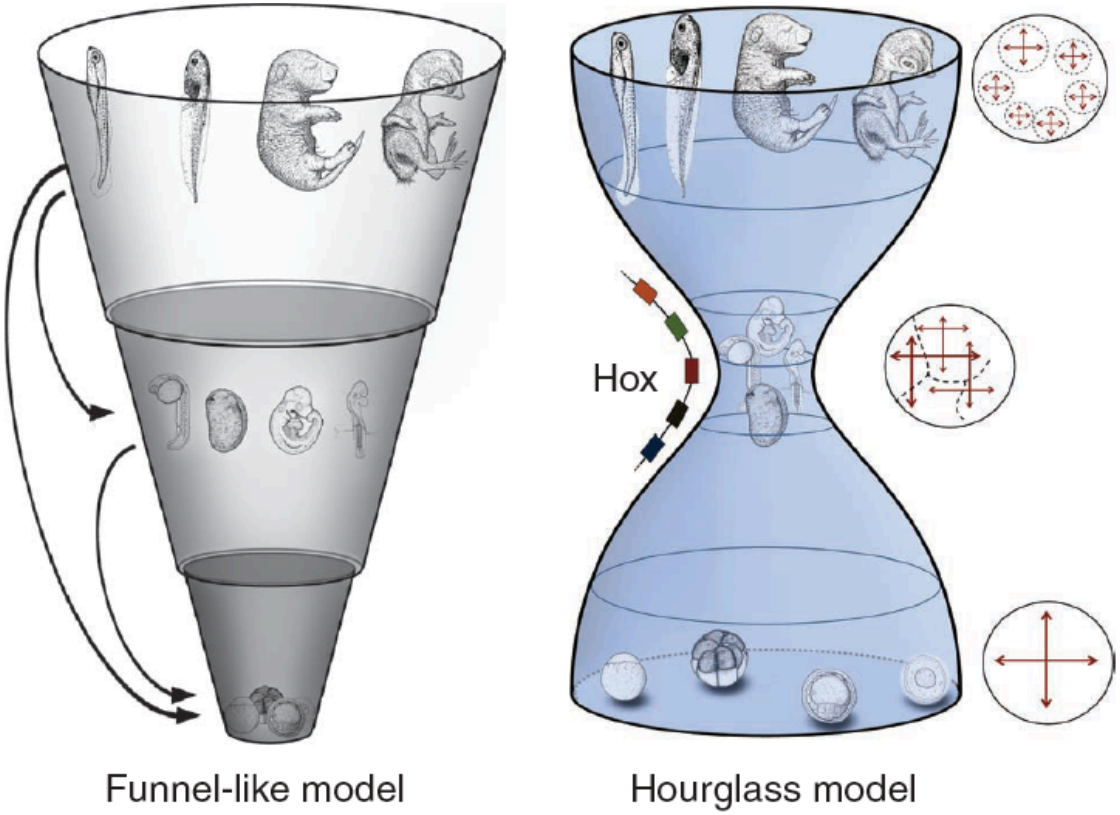
At four weeks after conception, we are a tube within a tube and have the three germ layers that give rise to all our organs.

Phylotypic Stage



Nature (2010) 468: 811

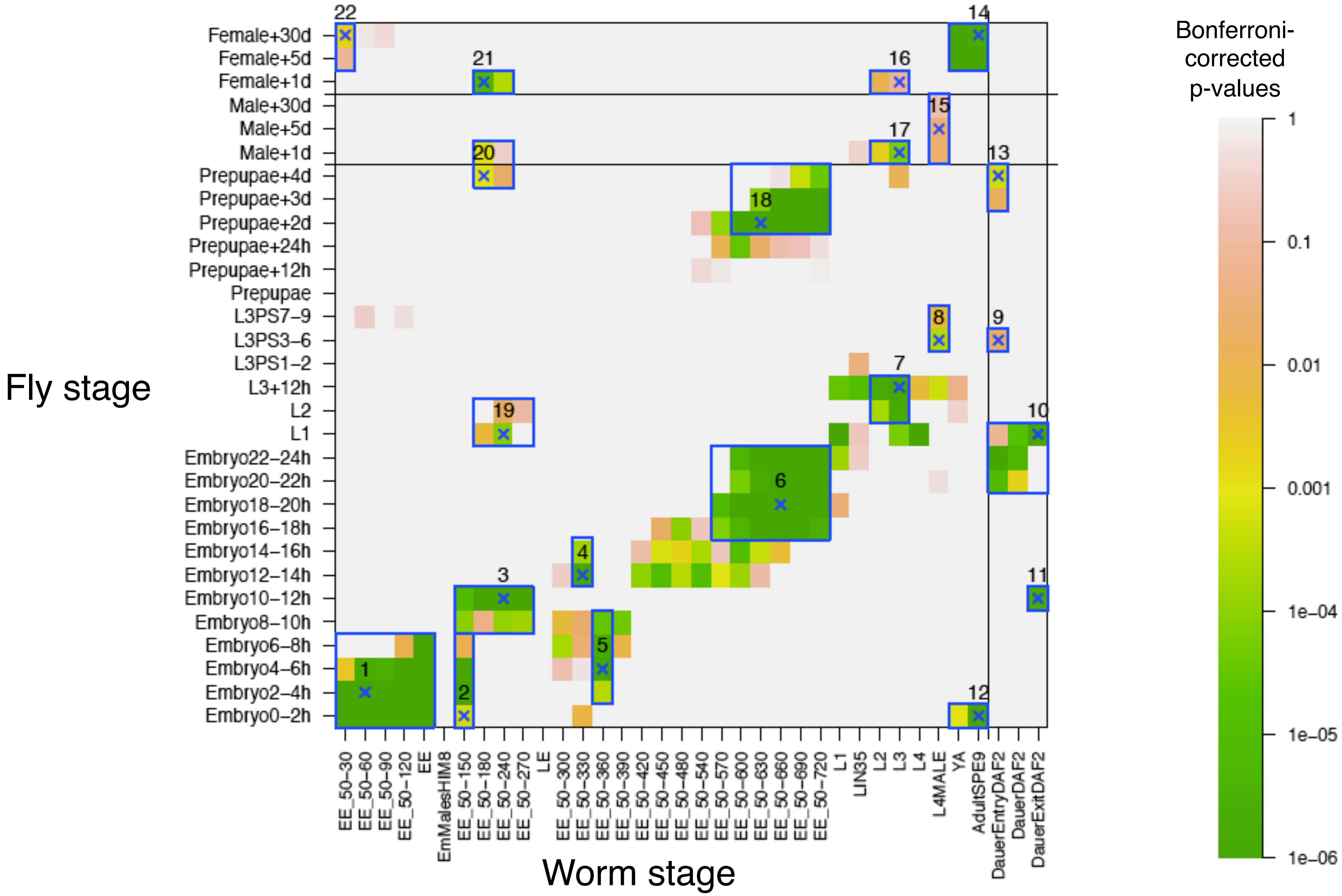
Phylotypic Stage



There is a problem with using early embryo stages for comparison across wide swaths of the phylogenetic tree. Embryonic stages have diverged further than the zootypic / phylotypic stage.

Nat. Comm. 2011 2: 248

Developmental stage mapping between worm and fly based on co-expression clustering of orthologs



Jingyi Jessica Li, Peter Bickel, Haiyan Huang, Steven Brenner

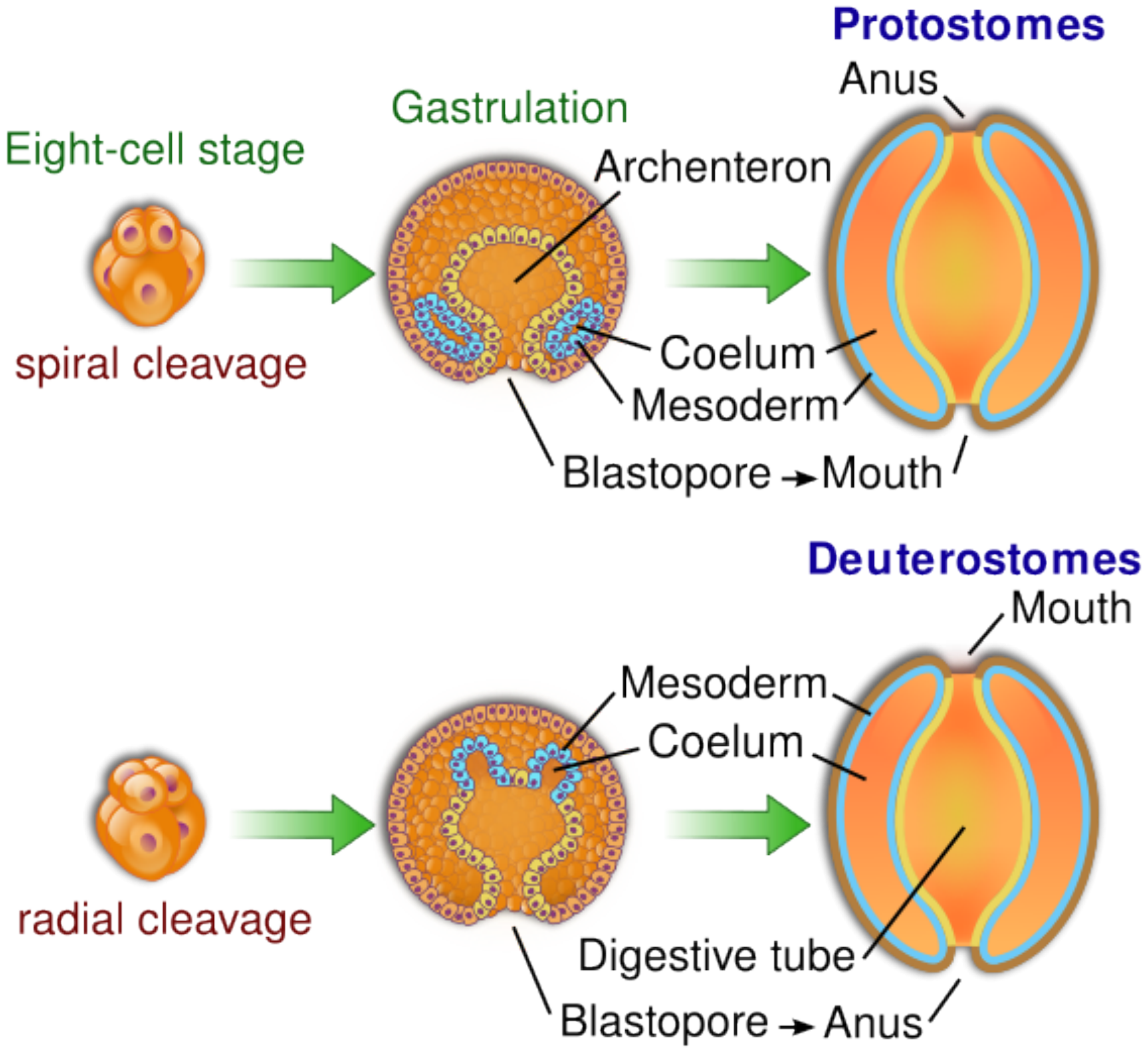
What do worms, flies, and humans have in common?

- They are all animals (metazoans).
- They are multicellular.
- They are triploblast – i.e. they have three germ layers
  - endoderm
  - mesoderm
  - ectoderm
- They are bilaterian – i.e. bilaterally symmetric.
- They have a zootypic stage, i.e. a body plan built from HOX gene regulatory networks
- and they have a phylotypic stage

BUT humans are deuterostomes, while worms and flies are protostomes.

- This difference in later development, after the phylotypic stage, appears unimportant for our analysis.

Protostomes vs Deuterostomes



.

Wikipedia user YassineMrabet

mod/ENCODE Integrative Comparison

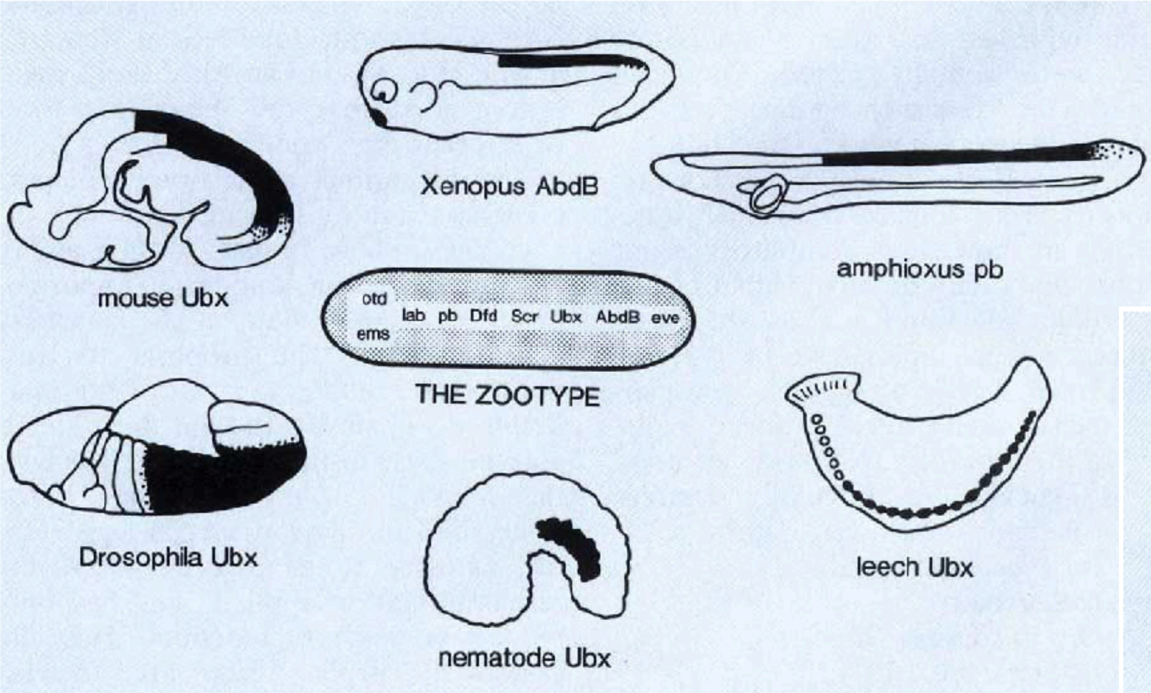
Worm, Fly, and Human

☀ Chromatin

Regulation

Transcription

Chromatin: How much of the histone code evolved at or before the origin of Bilateria?



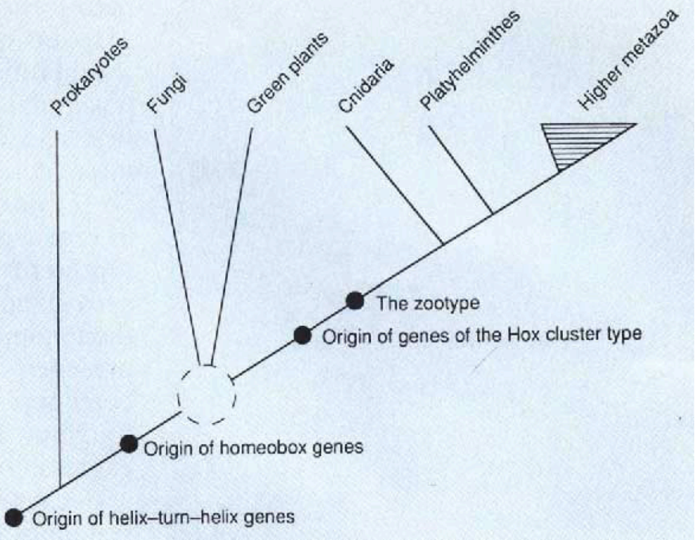
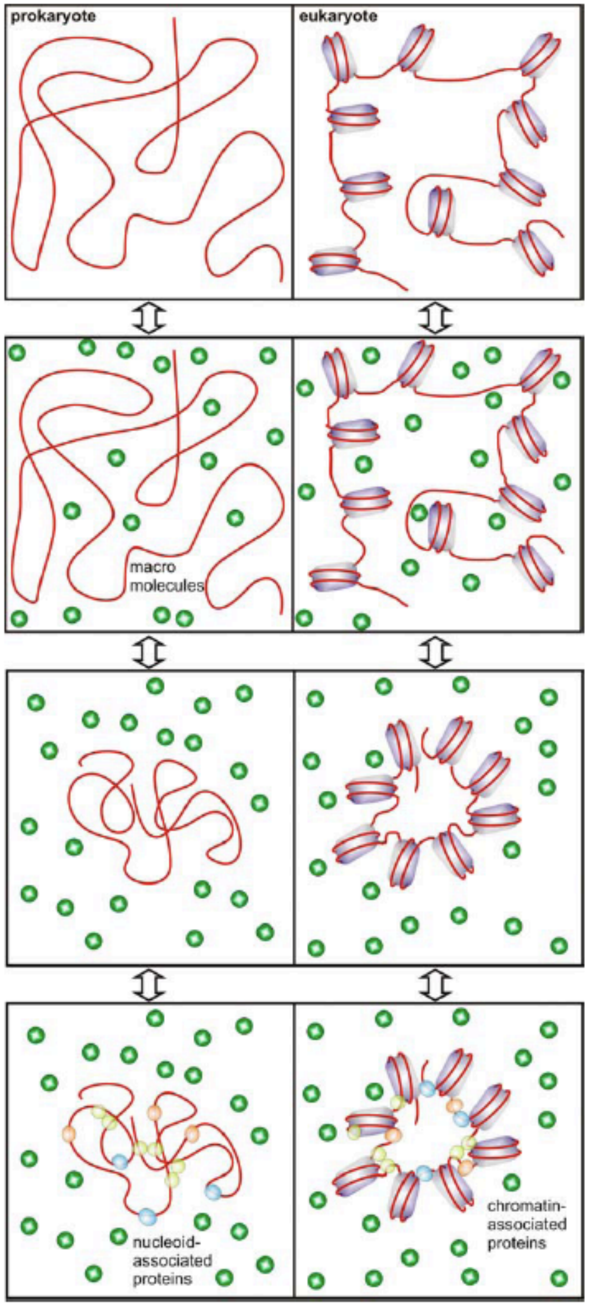


FIG. 4 Origin of the zootype on the evolutionary tree. The Hox cluster genes are a subset of the homeobox genes, which are in turn a subset of genes encoding DNA-binding proteins of the helix–turn–helix class.

The zootype and the phylotypic stage.
Nature (1993) 361: 490

Biophysics of chromatin architecture

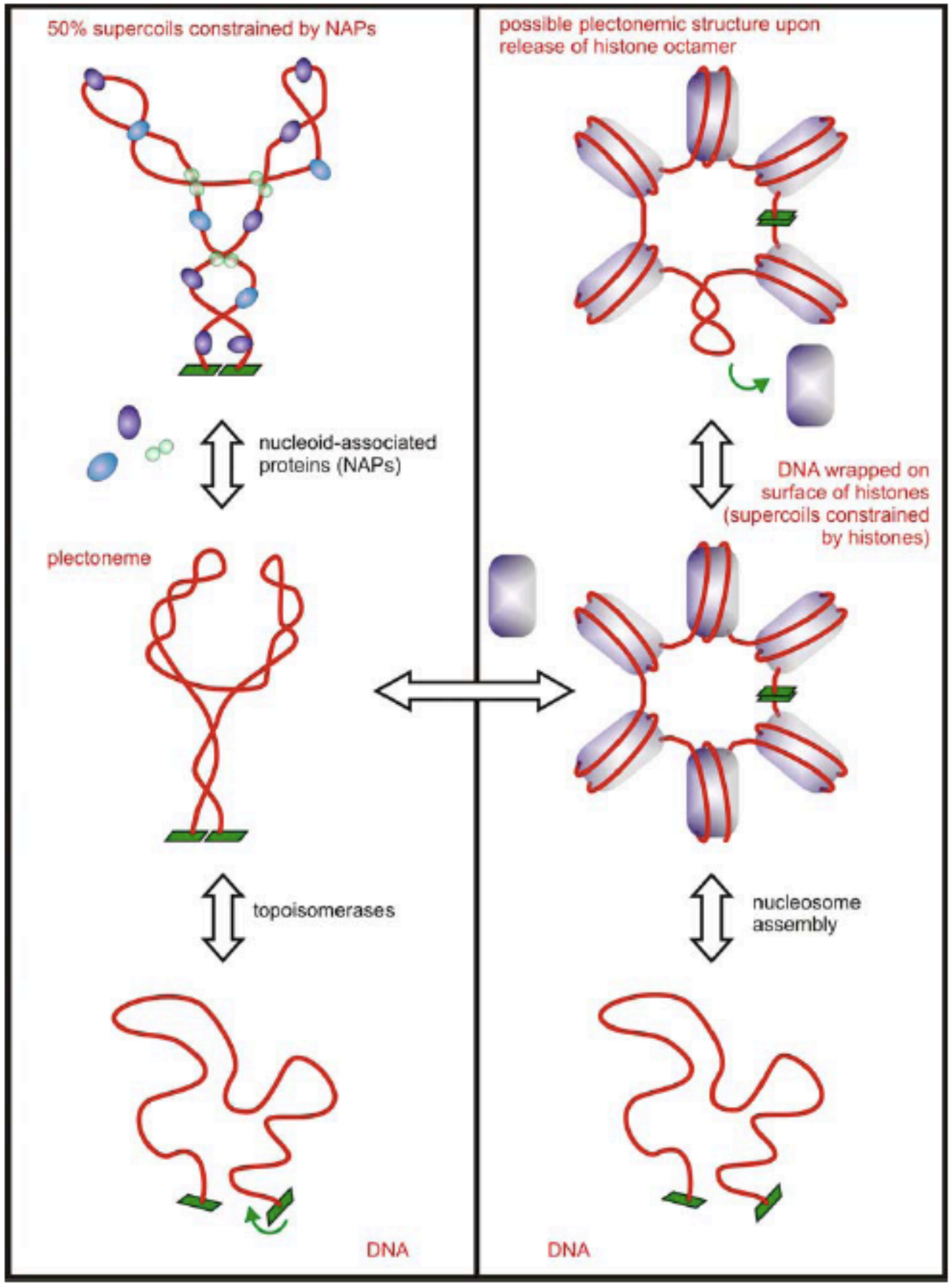
Macromolecular crowding forces chromatin condensation with or without the presence of chromatin-binding proteins.



The Major Architects of Chromatin: Architectural Proteins in Bacteria, Archaea and Eukaryotes. *Crit. Rev. Biochem. Molec. Biol.* (2008) 43: 393

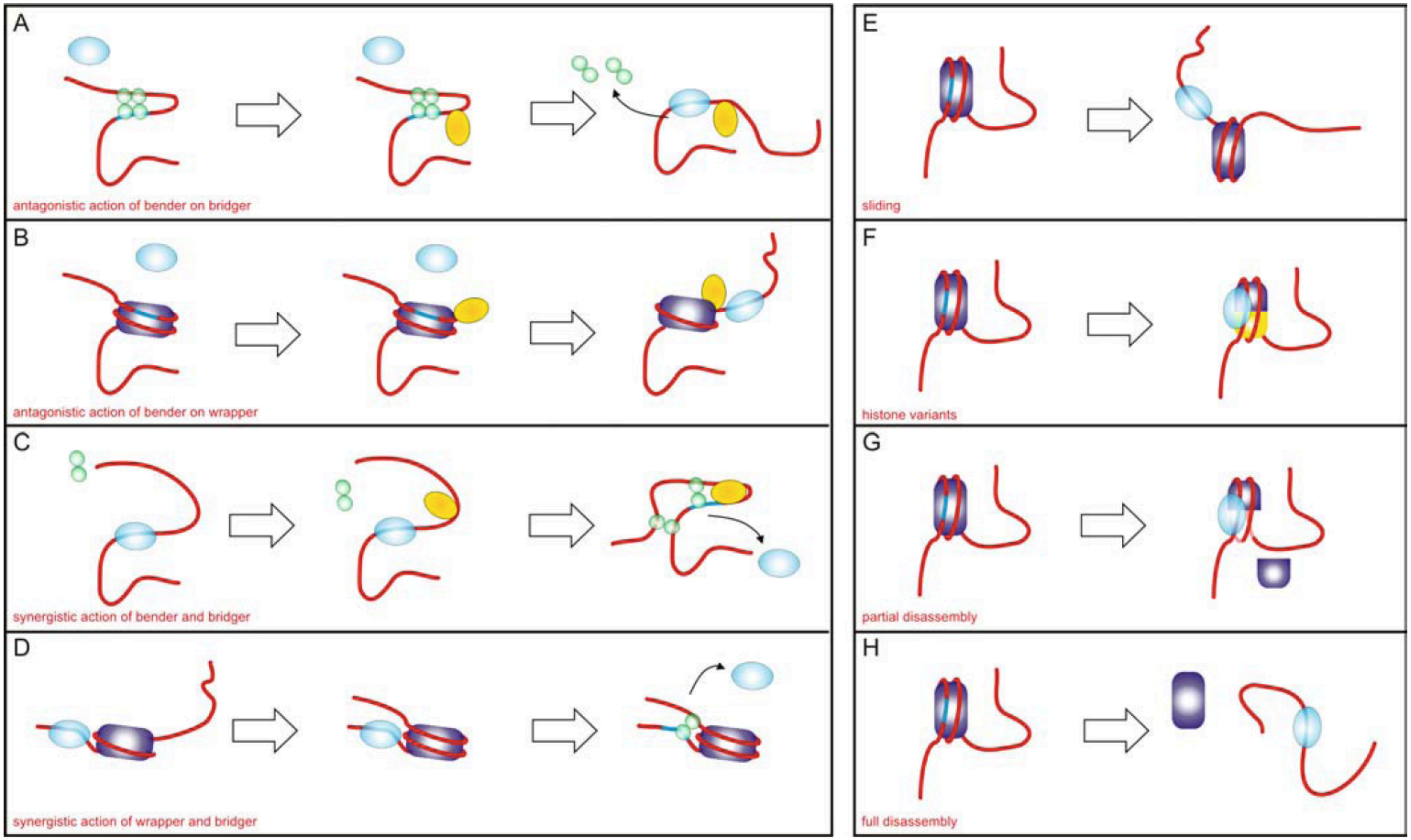
Biophysics of chromatin architecture

Supercoiling, tension, and torque are key to genome architecture in bacteria, archea, and eukaryotes.


50% supercoils constrained by NAPs
possible plectonemic structure upon release of histone octamer
nucleoid-associated proteins (NAPs)
DNA wrapped on surface of histones (supercoils constrained by histones)
plectoneme
topoisomerases
nucleosome assembly
DNA
DNA

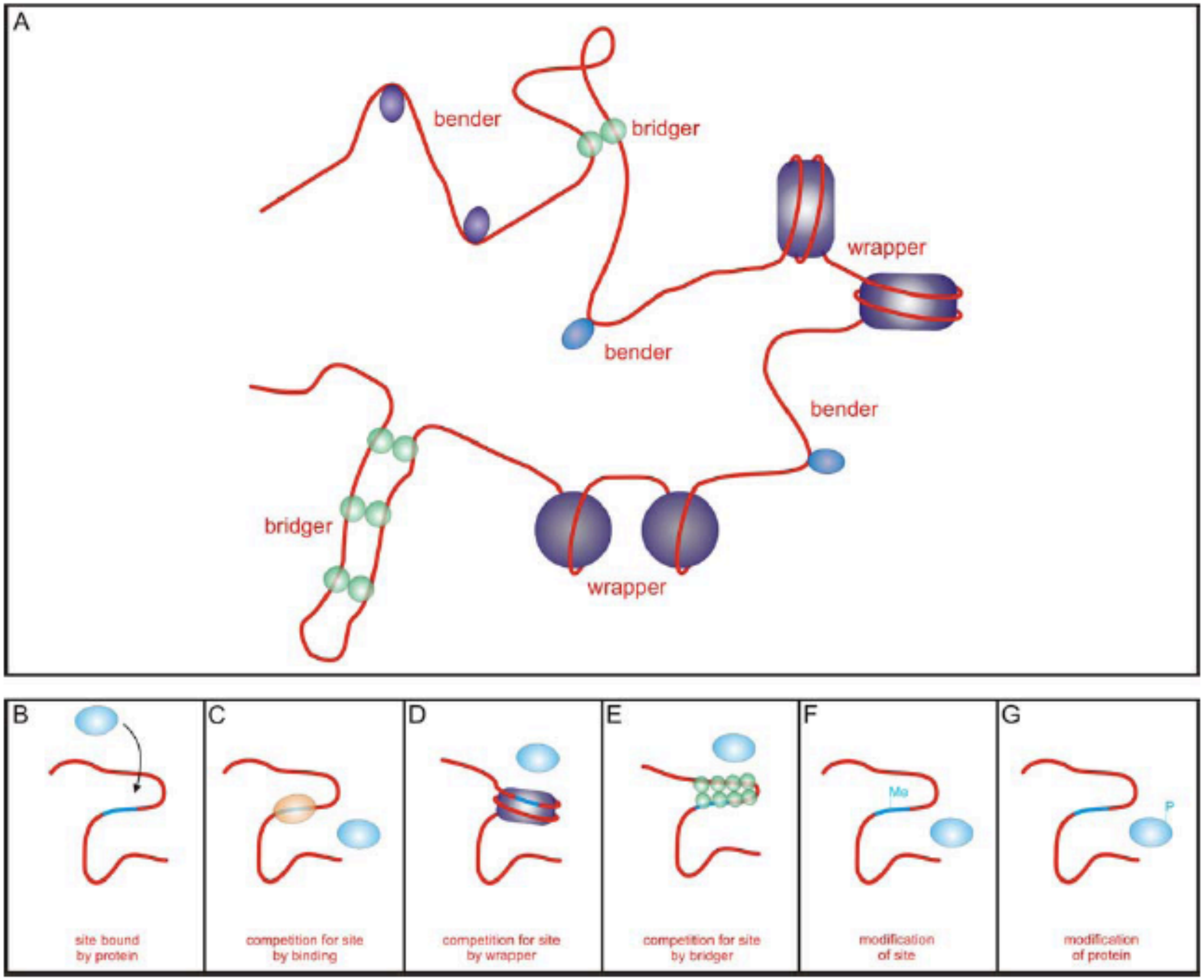
The Major Architects of Chromatin: Architectural Proteins in Bacteria, Archaea and Eukaryotes. *Crit. Rev. Biochem. Molec. Biol.* (2008) 43: 393

Chromatin binding and remodeling mechanisms


A
antagonistic action of bender on bridger
B
antagonistic action of bender on wrapper
C
synergistic action of bender and bridger
D
synergistic action of wrapper and bridger
E
sliding
F
histone variants
G
partial disassembly
H
full disassembly

The Major Architects of Chromatin: Architectural Proteins in Bacteria, Archaea and Eukaryotes. *Crit. Rev. Biochem. Molec. Biol.* (2008) 43: 393

Chromatin binding and remodeling mechanisms



The Major Architects of Chromatin: Architectural Proteins in Bacteria, Archaea and Eukaryotes. *Crit. Rev. Biochem. Molec. Biol.* (2008) 43: 393

mod/ENCODE Integrative Comparison

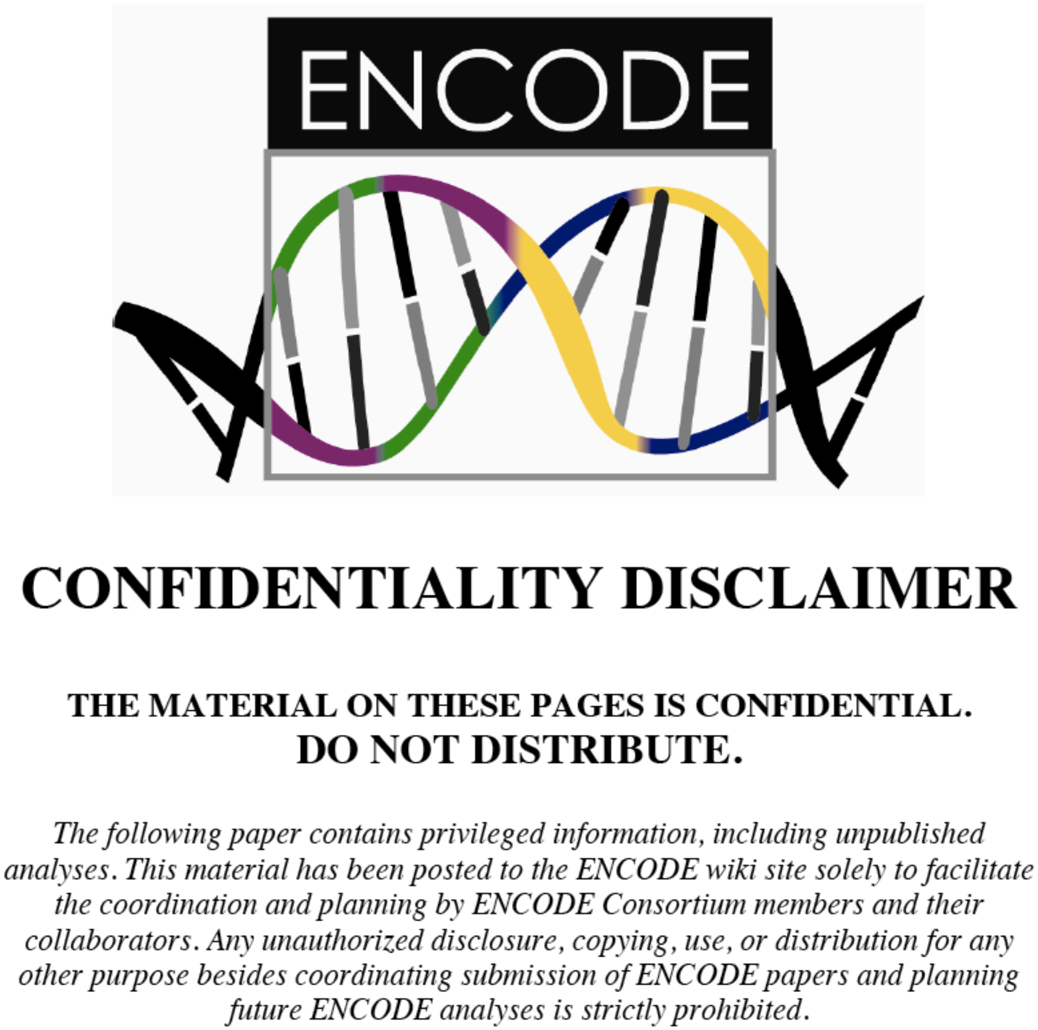
Worm, Fly, and Human

- Chromatin
- ☀ Regulation
- Transcription

Cell-type- and Tissue-specific Regulatory Networks from DNase Data

Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

Cell-type- and Tissue-specific Regulatory Networks from DNase Data



CONFIDENTIALITY DISCLAIMER

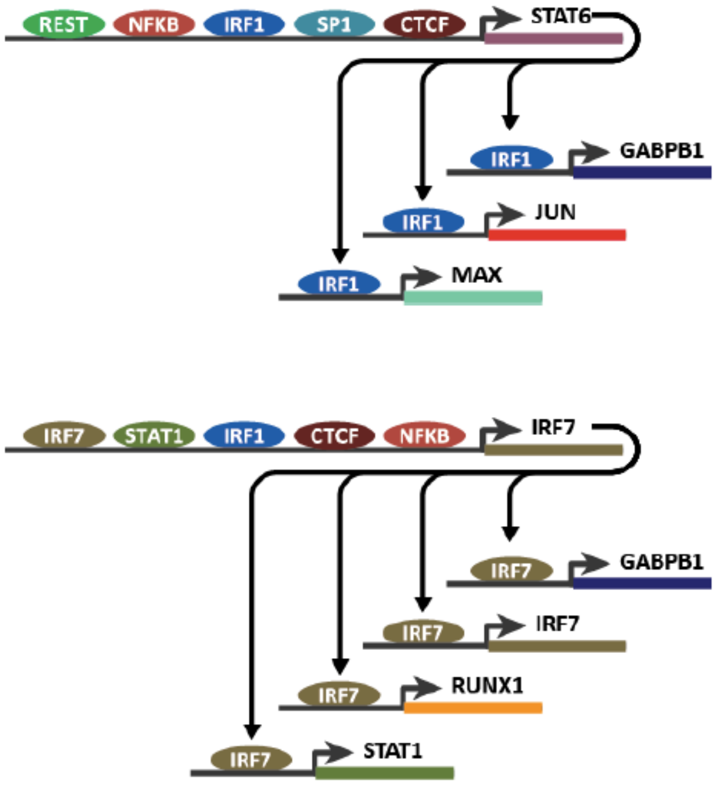
THE MATERIAL ON THESE PAGES IS CONFIDENTIAL.
DO NOT DISTRIBUTE.

The following paper contains privileged information, including unpublished analyses. This material has been posted to the ENCODE wiki site solely to facilitate the coordination and planning by ENCODE Consortium members and their collaborators. Any unauthorized disclosure, copying, use, or distribution for any other purpose besides coordinating submission of ENCODE papers and planning future ENCODE analyses is strictly prohibited.

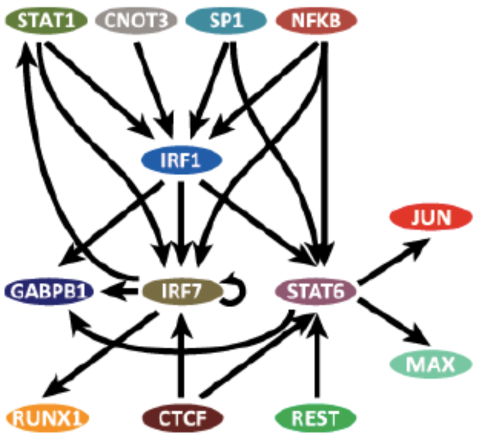
Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

Cell-type- and Tissue-specific Regulatory Networks from DNase Data

Delineating the circuitry of human TFs



. . . . etc . . . .



Repeat for all 475 TF genes with annotated recognition sequences

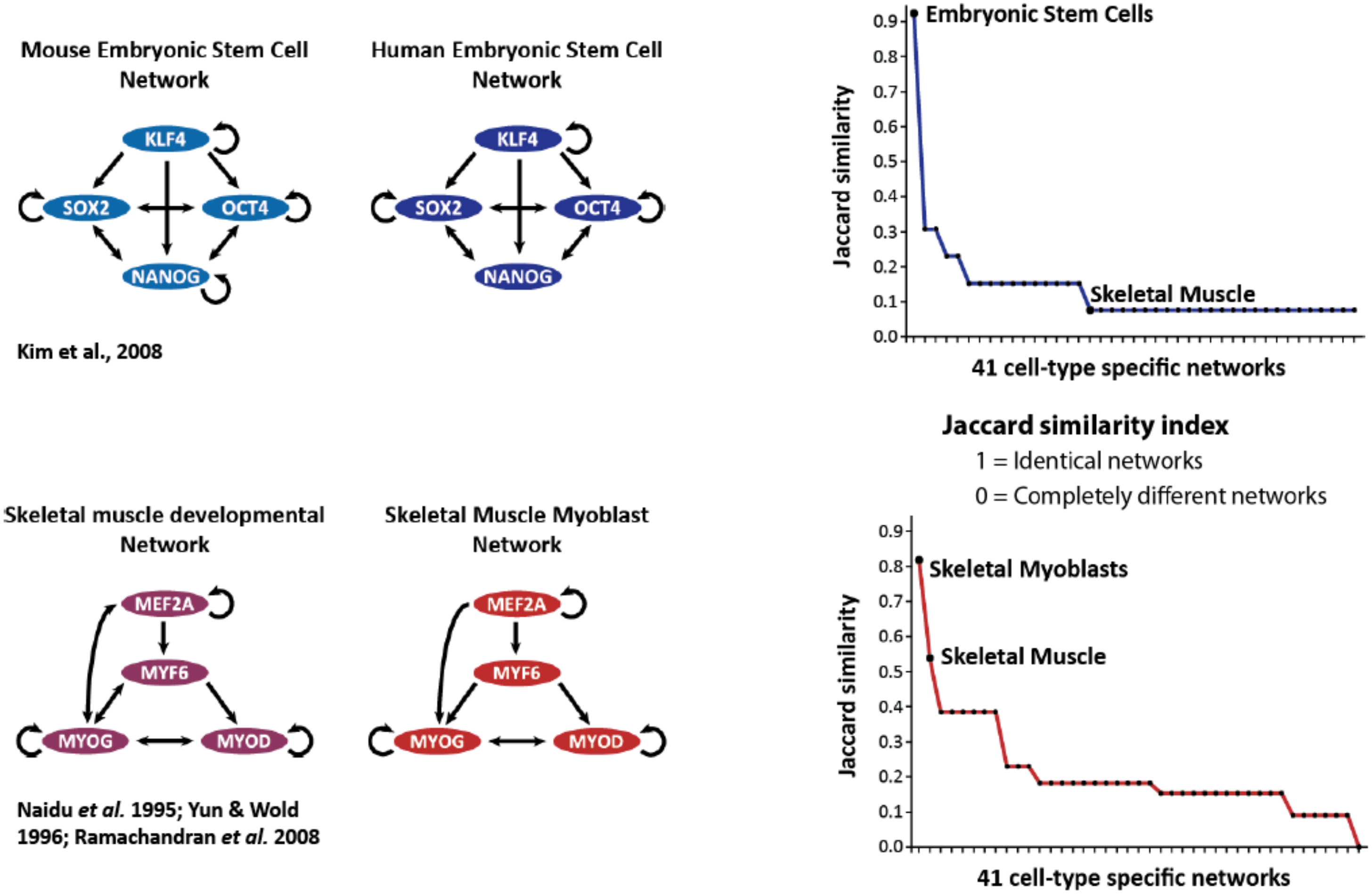
then

Repeat for 41 different cell types

Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

Cell-type- and Tissue-specific Regulatory Networks from DNase Data

De novo-derived networks accurately recapitulate known TF-to-TF network relationships



Kim et al., 2008

Naidu *et al.* 1995; Yun & Wold 1996; Ramachandran *et al.* 2008

Jaccard similarity index

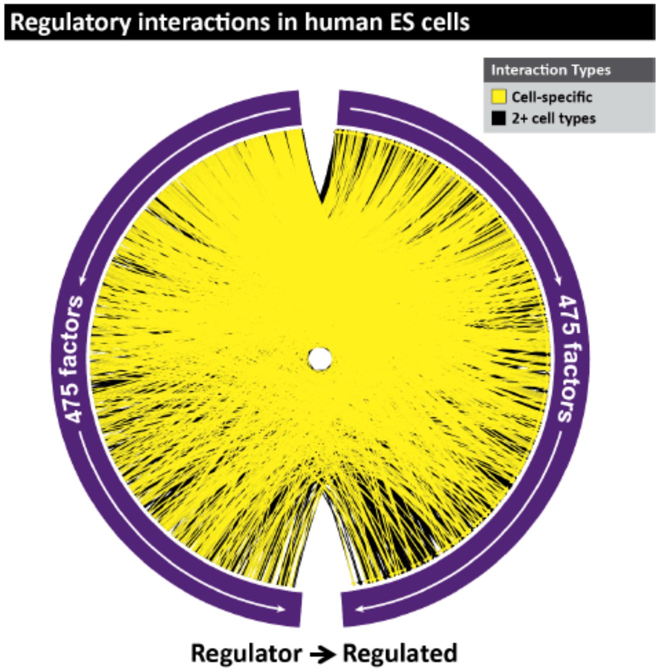
1 = Identical networks

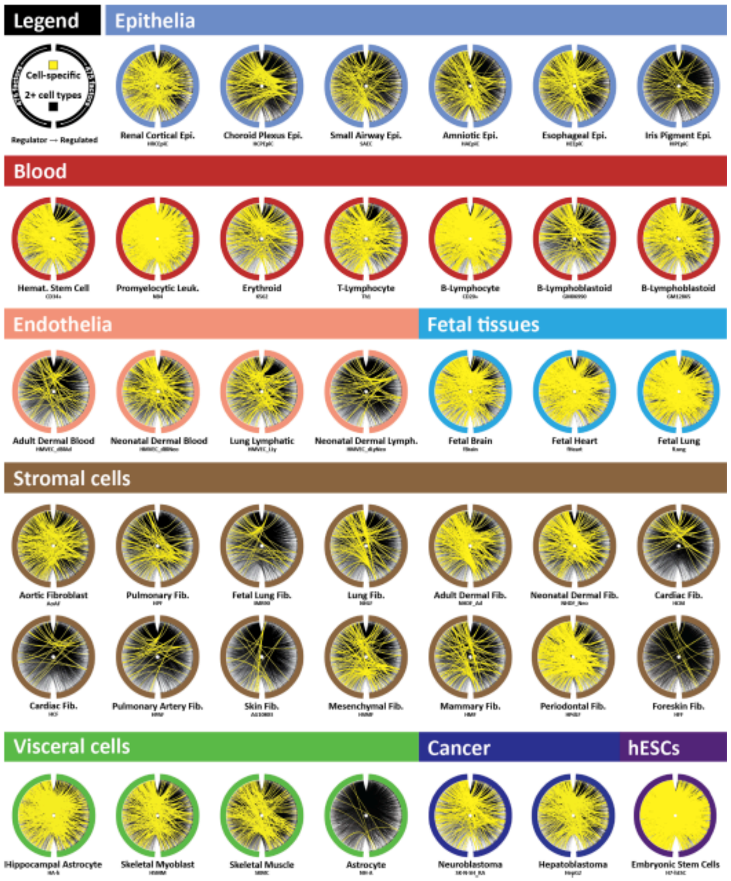
0 = Completely different networks

Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

Cell-type- and Tissue-specific Regulatory Networks from DNase Data

Transcription factor regulatory networks are highly cell-selective

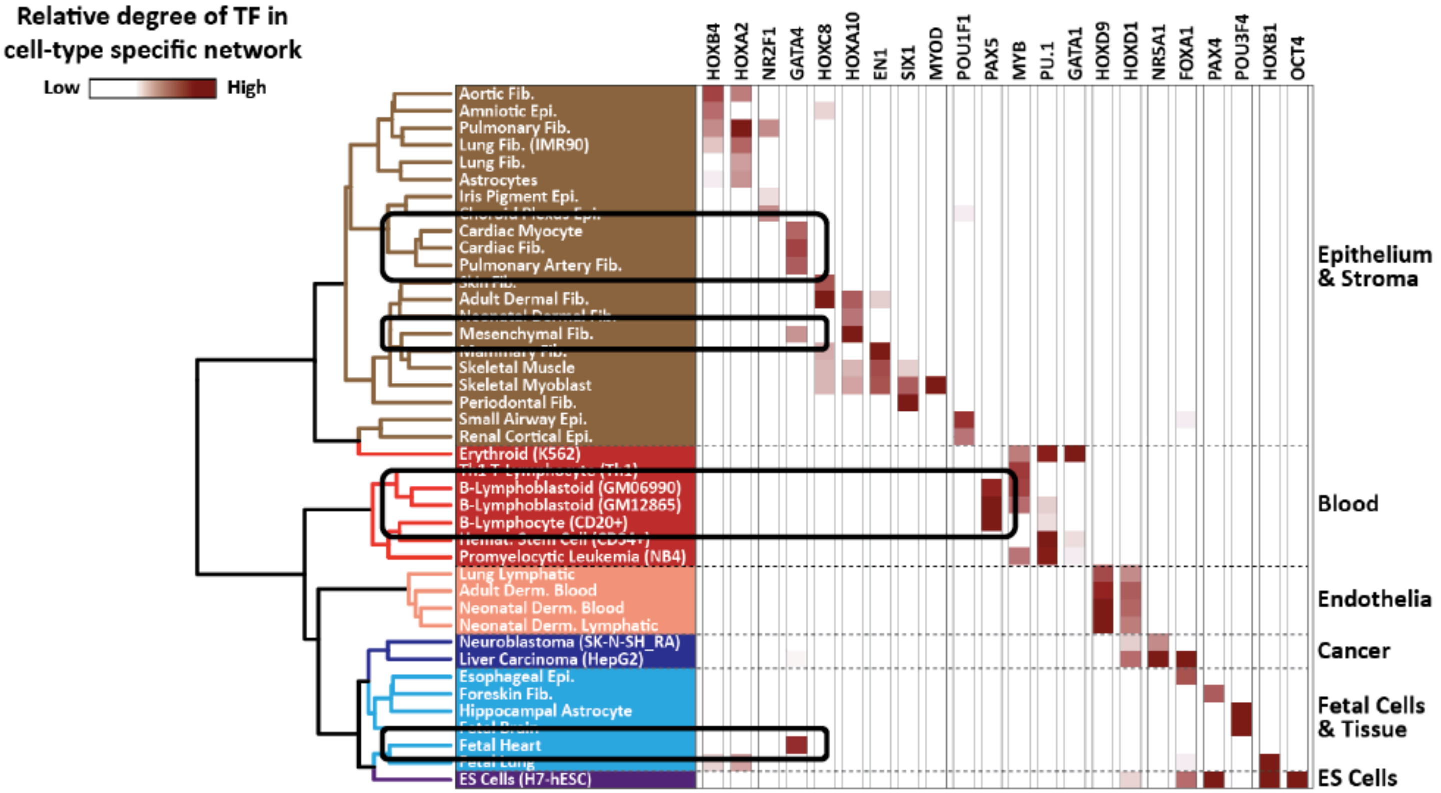




Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

Cell-type- and Tissue-specific Regulatory Networks from DNase Data

Functionally related cell types share similar core transcriptional regulatory networks



Cluster cell types → Identify which cell types are governed by similar TFs

Stam lab, ENCODE NCP008. *Nature* (6 Sept 2012)

mod/ENCODE Integrative Comparison

Worm, Fly, and Human

Chromatin

Regulation

✹ Transcription

Classes of non-coding RNA

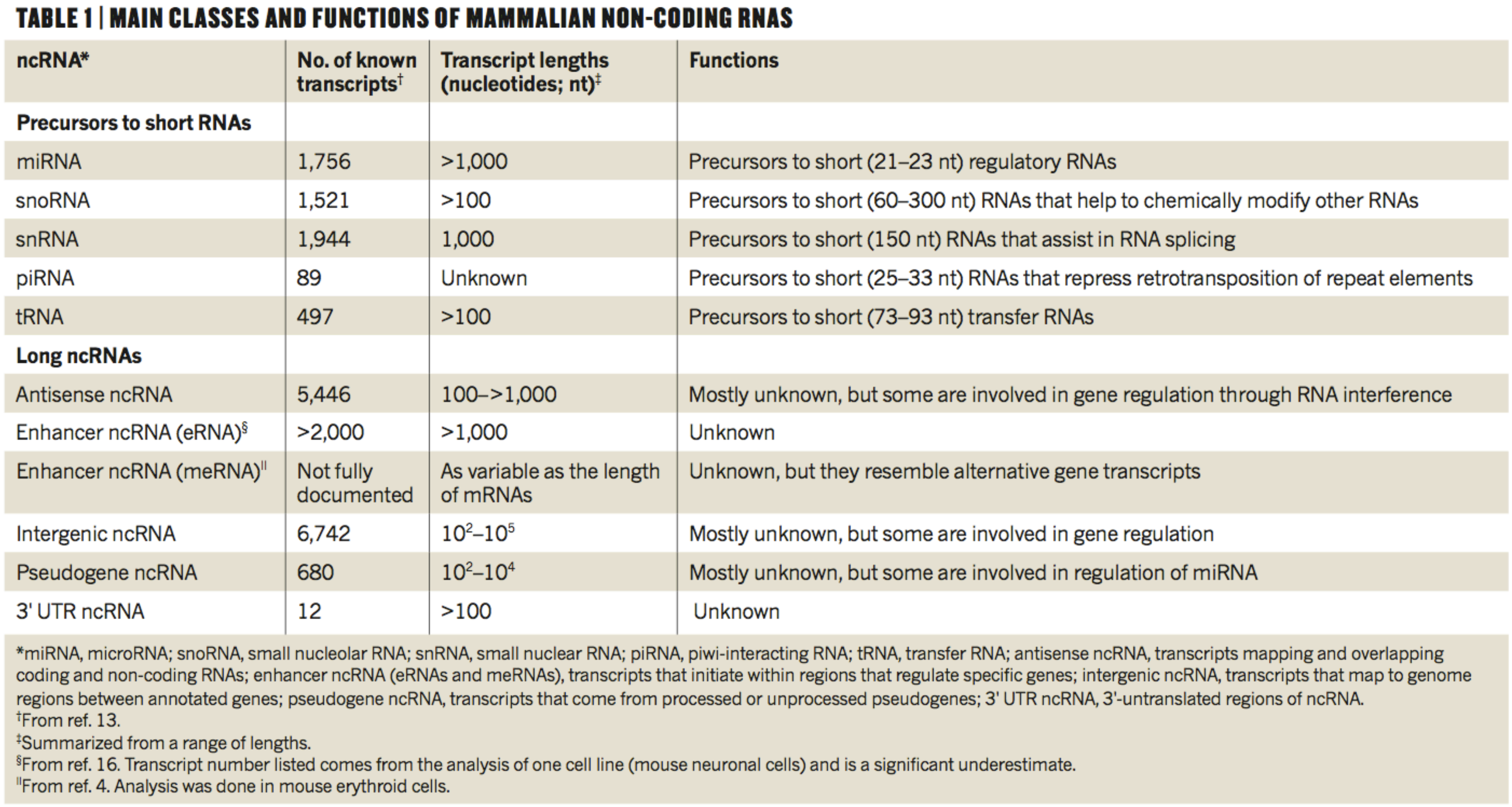


TABLE 1 | MAIN CLASSES AND FUNCTIONS OF MAMMALIAN NON-CODING RNAS

ncRNA*	No. of known transcripts[†]	Transcript lengths (nucleotides; nt)[‡]	Functions
Precursors to short RNAs			
miRNA	1,756	>1,000	Precursors to short (21–23 nt) regulatory RNAs
snoRNA	1,521	>100	Precursors to short (60–300 nt) RNAs that help to chemically modify other RNAs
snRNA	1,944	1,000	Precursors to short (150 nt) RNAs that assist in RNA splicing
piRNA	89	Unknown	Precursors to short (25–33 nt) RNAs that repress retrotransposition of repeat elements
tRNA	497	>100	Precursors to short (73–93 nt) transfer RNAs
Long ncRNAs			
Antisense ncRNA	5,446	100–>1,000	Mostly unknown, but some are involved in gene regulation through RNA interference
Enhancer ncRNA (eRNA)[§]	>2,000	>1,000	Unknown
Enhancer ncRNA (meRNA)[‖]	Not fully documented	As variable as the length of mRNAs	Unknown, but they resemble alternative gene transcripts
Intergenic ncRNA	6,742	10^2–10^5	Mostly unknown, but some are involved in gene regulation
Pseudogene ncRNA	680	10^2–10^4	Mostly unknown, but some are involved in regulation of miRNA
3' UTR ncRNA	12	>100	Unknown

*miRNA, microRNA; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; piRNA, piwi-interacting RNA; tRNA, transfer RNA; antisense ncRNA, transcripts mapping and overlapping coding and non-coding RNAs; enhancer ncRNA (eRNAs and meRNAs), transcripts that initiate within regions that regulate specific genes; intergenic ncRNA, transcripts that map to genome regions between annotated genes; pseudogene ncRNA, transcripts that come from processed or unprocessed pseudogenes; 3' UTR ncRNA, 3'-untranslated regions of ncRNA.

[†] From ref. 13.

[‡] Summarized from a range of lengths.

[§] From ref. 16. Transcript number listed comes from the analysis of one cell line (mouse neuronal cells) and is a significant underestimate.

[‖] From ref. 4. Analysis was done in mouse erythroid cells.

Nature (2012) 482: 310

ncRNA Discovery in *C. elegans* – useful resource

- The majority of the yet un-characterized is-ncRNA are expressed at low levels and/or only during specific stages of *C. elegans* development (Wang et al. 2011) and would thus only be detectable by very large sequencing depths.
- We reasoned that cleavage fragments (or other processed fragments) of mature rRNAs and mRNAs would most likely have monophosphate 59 termini and could thus largely be eliminated by treatment with Terminator 59-phosphate-dependent exonuclease (TEX).

Runsheng Chen, Institute of Biophysics, Chinese Academy of Sciences, Beijing

RNA (2012) 18: 626

ncRNA Discovery in *C. elegans*

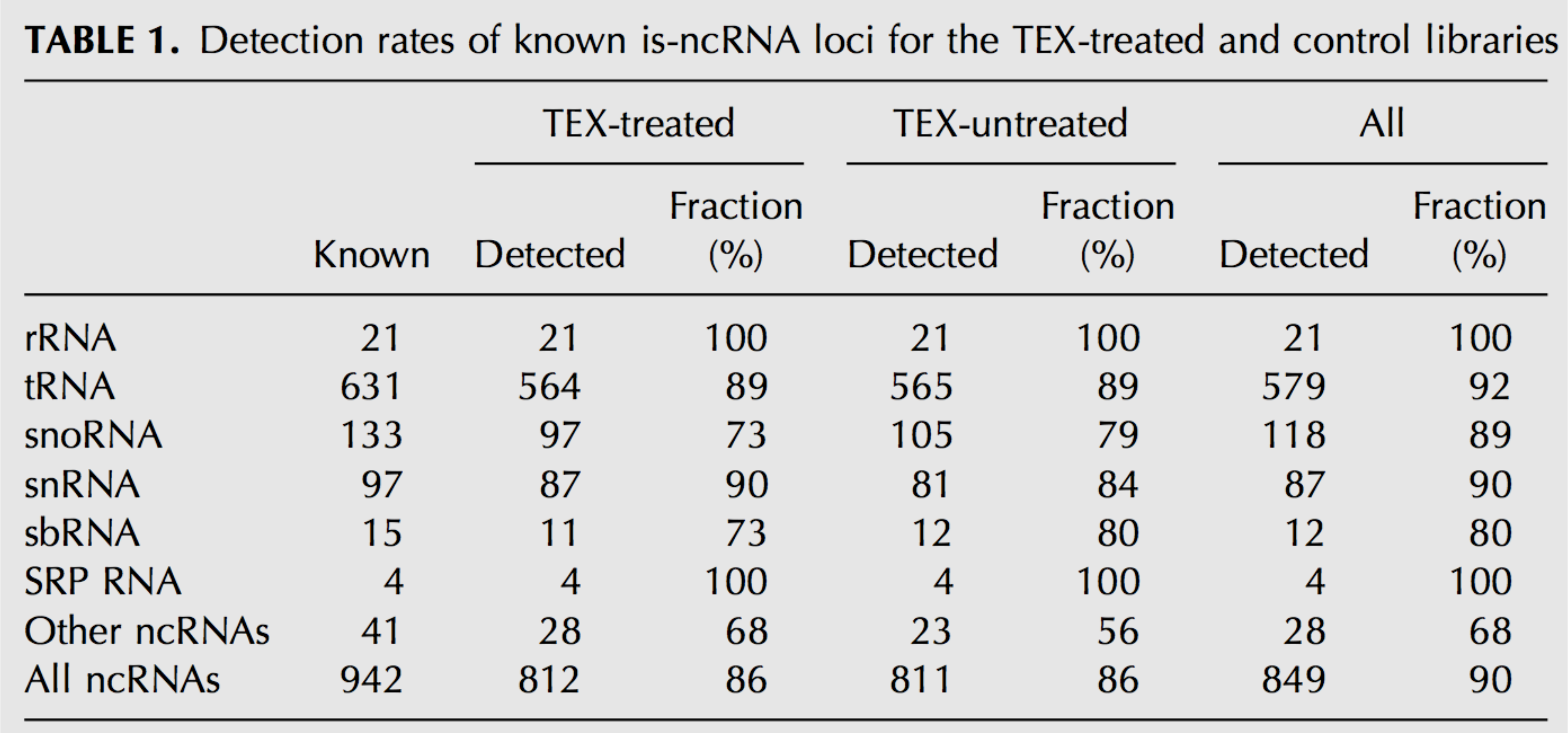
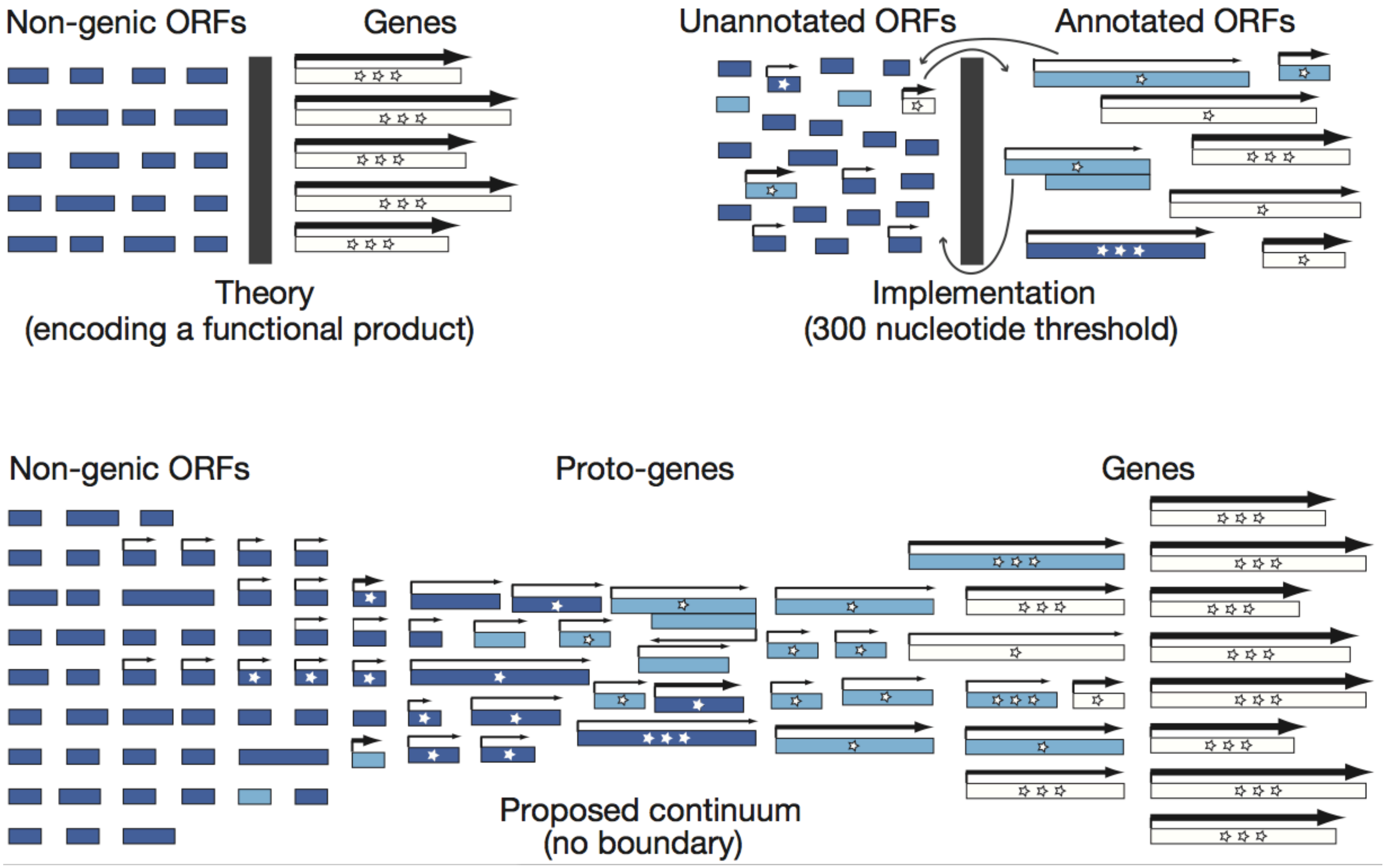


TABLE 1. Detection rates of known is-ncRNA loci for the TEX-treated and control libraries

		TEX-treated		TEX-untreated		All	
	Known	Detected	Fraction (%)	Detected	Fraction (%)	Detected	Fraction (%)
rRNA	21	21	100	21	100	21	100
tRNA	631	564	89	565	89	579	92
snoRNA	133	97	73	105	79	118	89
snRNA	97	87	90	81	84	87	90
sbRNA	15	11	73	12	80	12	80
SRP RNA	4	4	100	4	100	4	100
Other ncRNAs	41	28	68	23	56	28	68
All ncRNAs	942	812	86	811	86	849	90

RNA (2012) 18: 626

Possible function for ncRNA:
de novo gene birth via proto-genes



Nature (2012) 487: 370

Transcription Paper Outline

- Comparison of protein-coding genes
  - Comparison with existing annotations (Hillier, Davis, Brown)
  - Splicing complexity (Graveley)
  - Comparison of select orthologs (Mortazavi, Harrow, Celniker)
- Comparison on non-coding RNAs (Brown, Lai, Gerstein, Guigo, Samsonova)
- Comparison of pseudogenes (Gerstein)
- Analysis of relationship of upstream regions to transcript level (Gerstein, Weng)
- Expression clustering (Brenner, Gerstein)

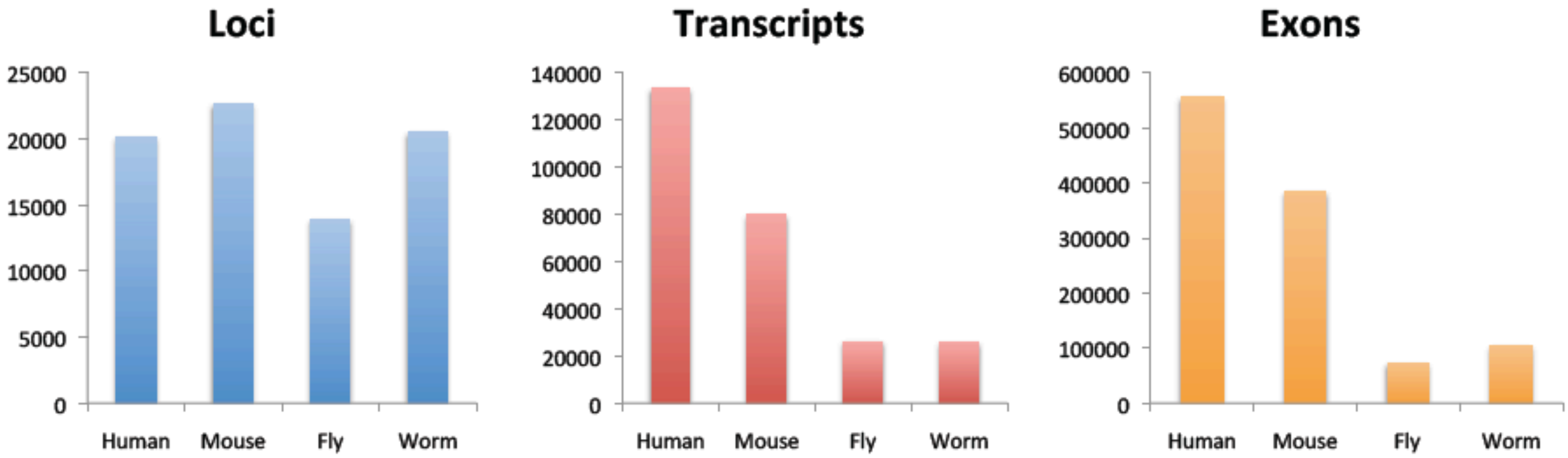
Datasets

- agreed-upon “expression compendium”
  - total RNA
  - ENCODE Tier 1
- developmental time courses (worm, fly)
- matched embryonic datasets

Comparison with existing annotations

- Because of the difficulty of assembling full transcripts with short reads and comparing their expression across species, we will focus on comparing transcript elements:
  - Transcript Start Sites (TSSs)
  - Transcript End Sites (TESs)
  - Splice Junctions (SJ)
  - de novo exons
  - de novo genes
  - de novo transcripts
  - Expression values for each above element
  - Expression values for the annotations

Number of protein-coding genes

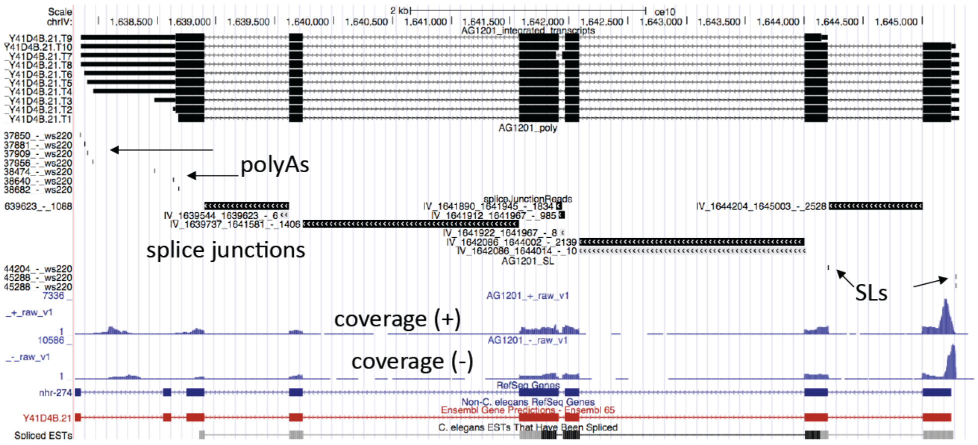


Adam Frankish

Finding all isoforms of a gene can be difficult

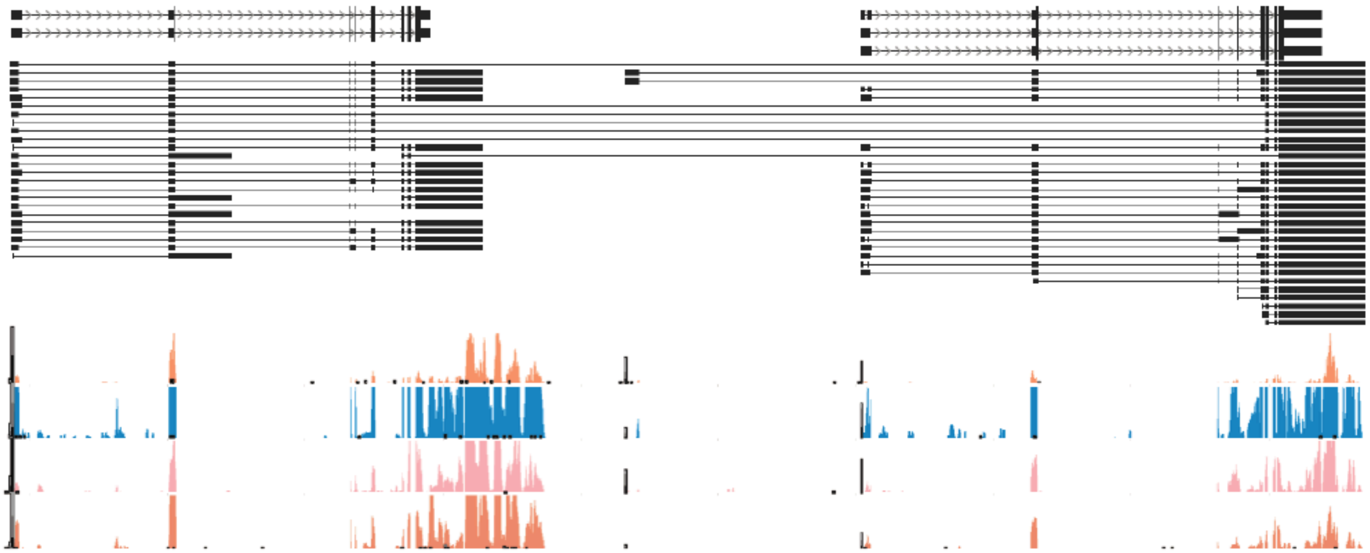
Simple Case

aggregate models



C. elegans refseq models and spliced ESTs

Hard Case



Flybase r5.32

Sparsified GRIT Transcripts

RNAseq & CAGE

Analysis of Splicing Complexity

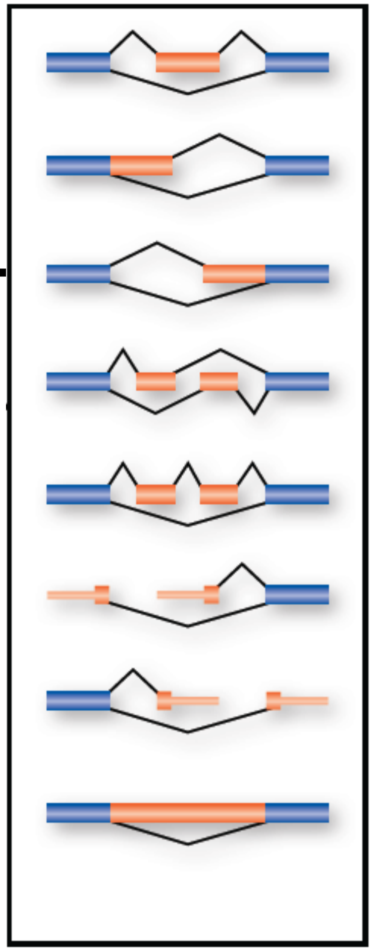
Transcript Annotations

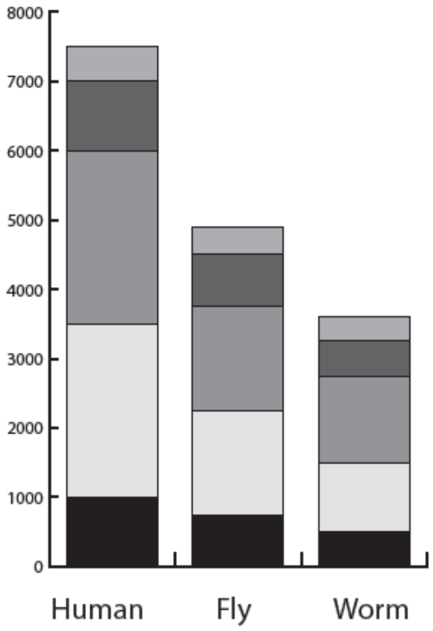
↓

Event Classifications

↓

Comparison of Event Types





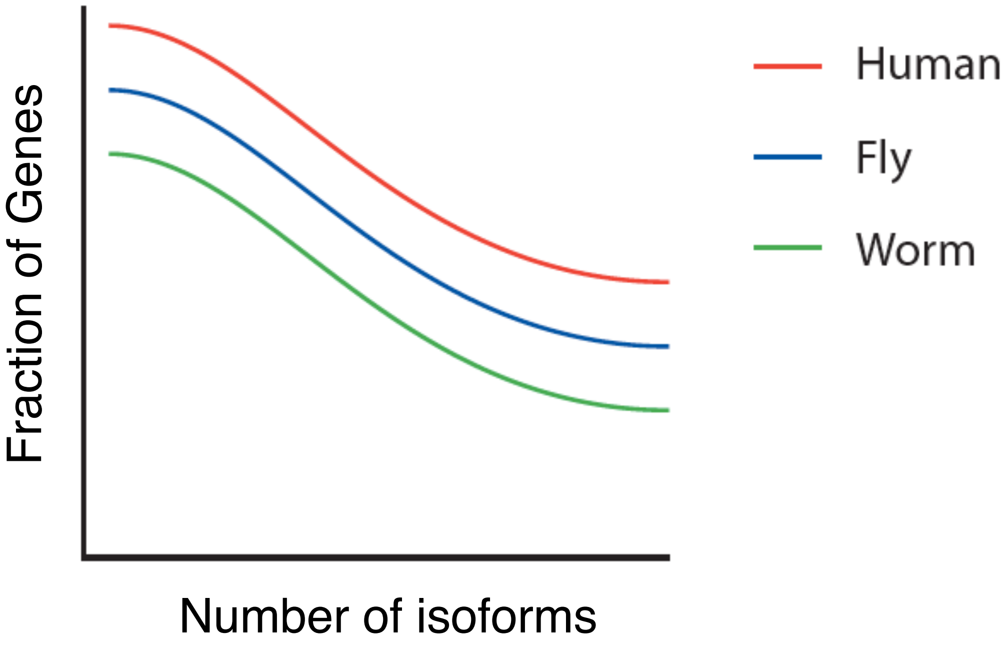
Brenton Graveley

Analysis of Splicing Complexity

- For all three species, compare motifs and conservation at splice sites for constitutive vs. alternative exons, and highly switching vs. low switching.

- Analyze number of isoforms per gene.

  Highlight outliers (Dscam, etc.)

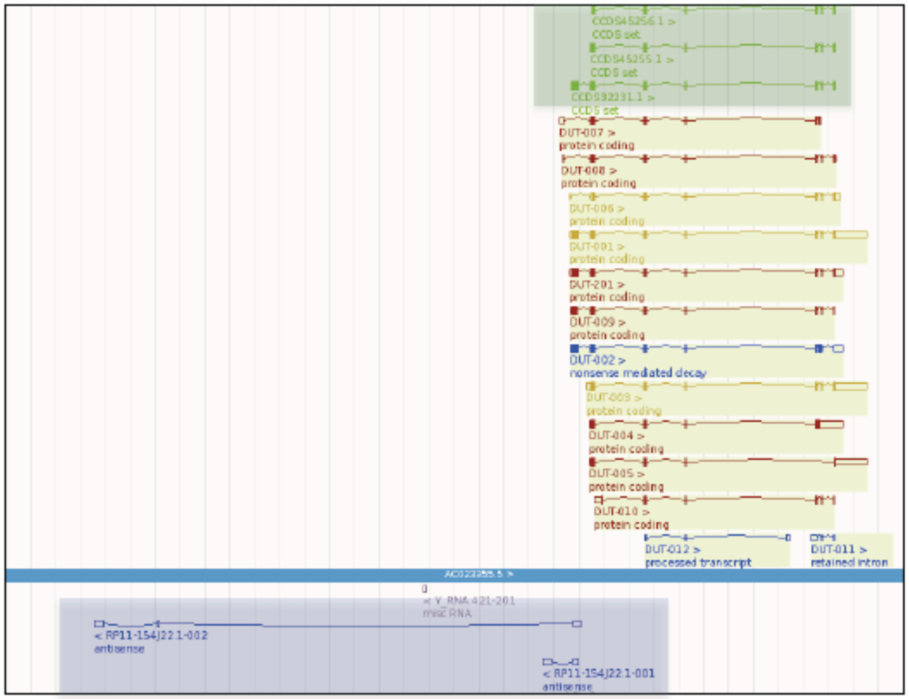


Brenton Graveley

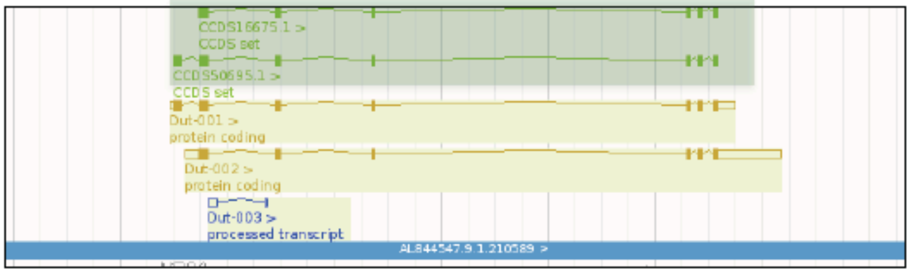
Comparison of select orthologs

Case Study: DUT / Dut / dUTPase / dut-1

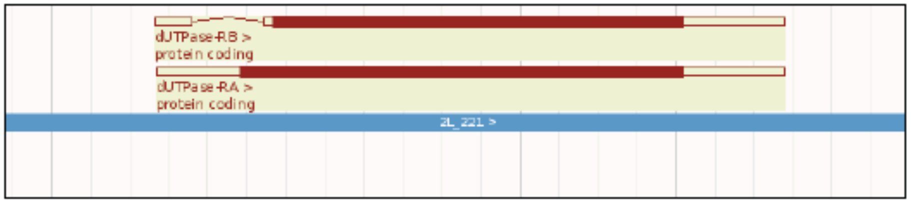
Human DUT



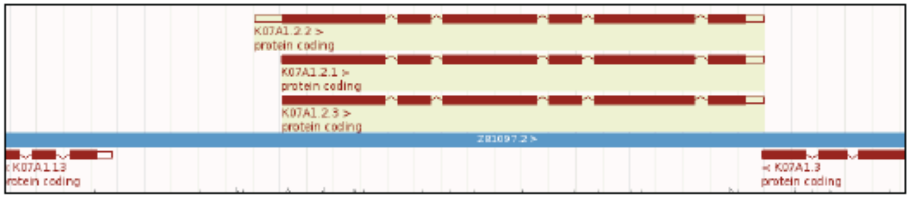
Mouse Dut



Fly dUTPase



Worm dut-1



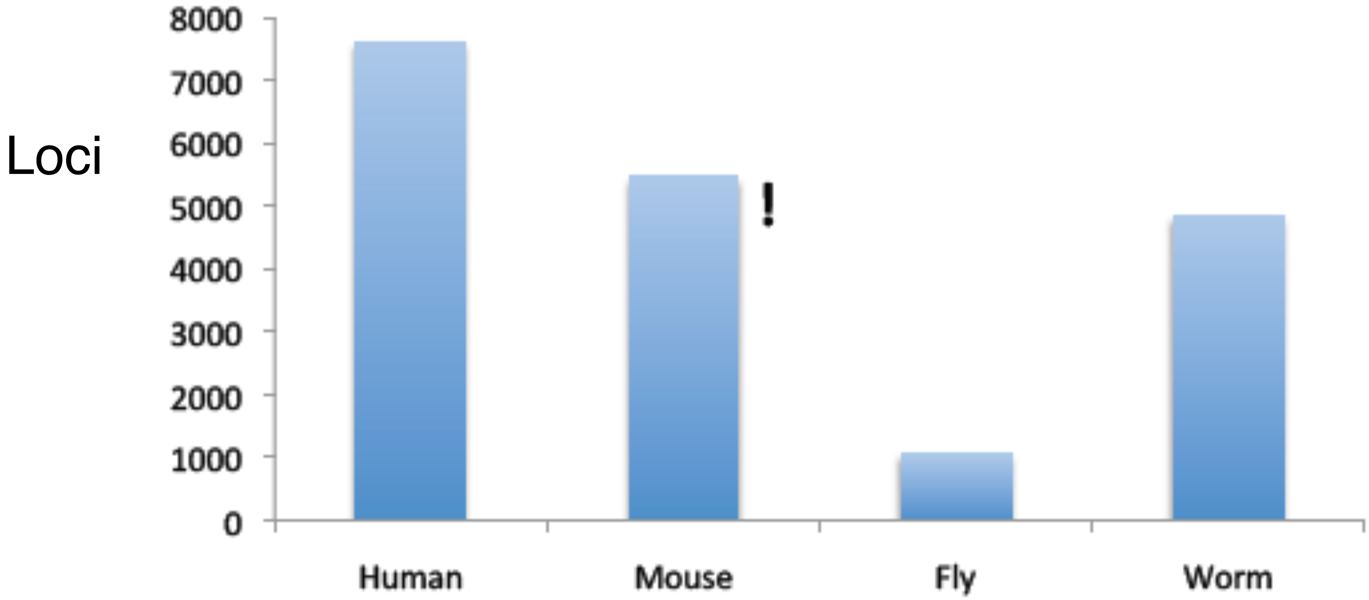
Adam Frankish

Comparison of non-coding RNAs

- How much of the nc genome is transcribed?
  - per megabase
  - across entire agreed-upon “expression compendium”
  - in ~matched embryonic stages
  - Ubiquitous vs Stage- / Cell-line specific transcription
- You cannot directly compare annotations (Gencode vs Flybase vs Wormbase)
- so, use a tiered approach; build a table or pie chart
  - first compare the existing annotations
  - incRNA algorithm
    - breakdown by RNA class
  - *de novo* mapping / TAR calling
    - issues: repeats, multi-mapped vs unique reads

Comparison of existing annotations

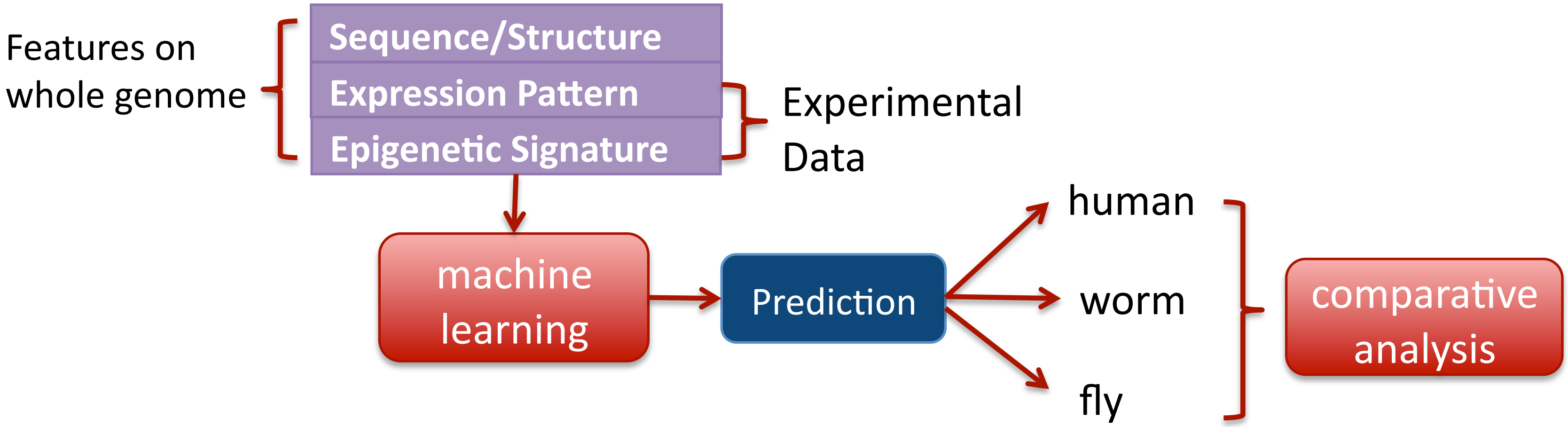
Number of short ncRNA



rRNA, tRNA, miRNA, snRNA, snoRNA (! mouse excludes tRNA)

Adam Frankish

incRNA algorithm



Features on whole genome

Sequence/Structure

Expression Pattern

Epigenetic Signature

Experimental Data

machine learning

Prediction

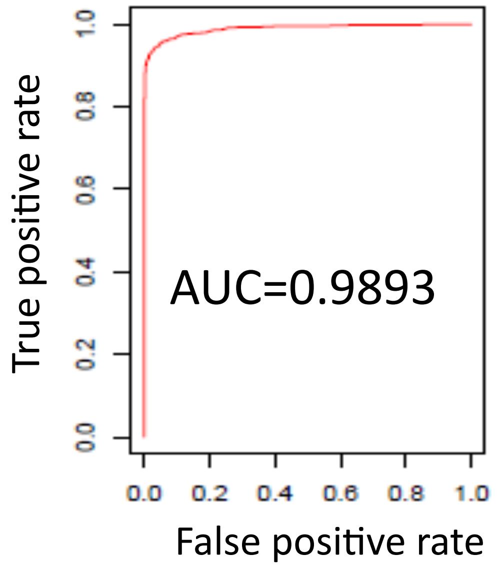
human

worm

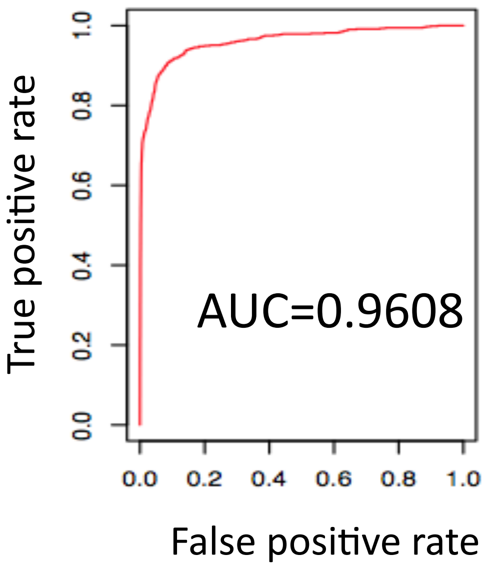
fly

comparative analysis

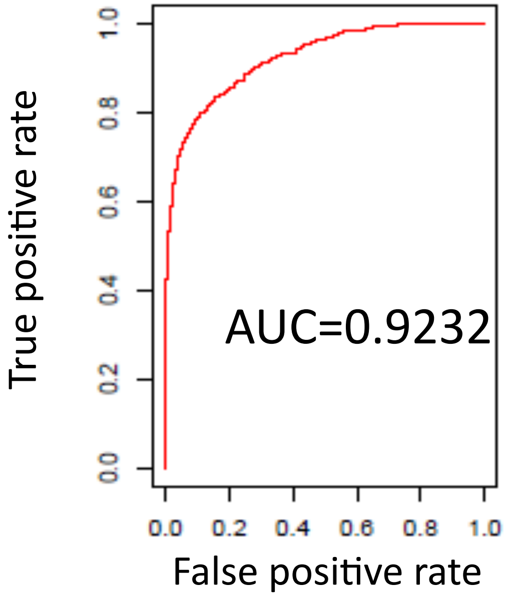
Results for known types of ncRNAs:



fly



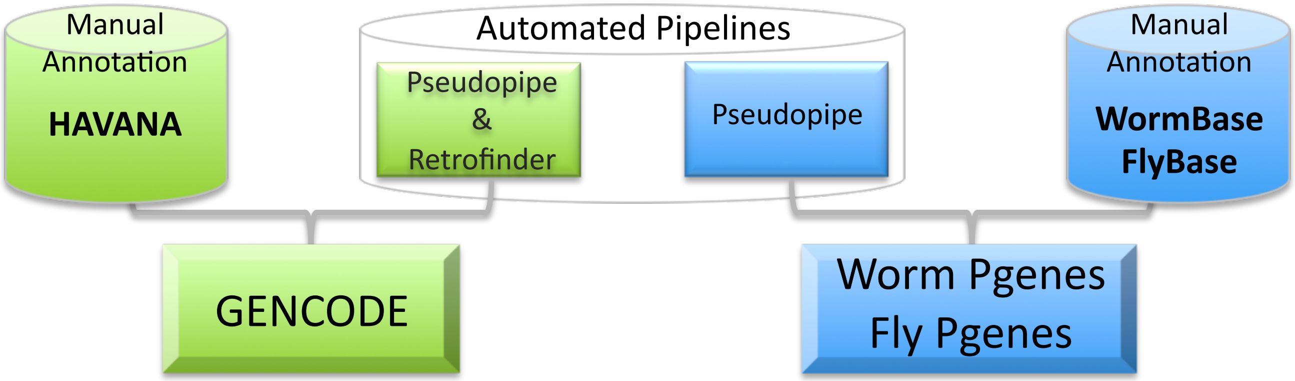
worm



human

Comparison of pseudogenes

- Pseudogenes annotated using automated pipelines intersected with manual curation



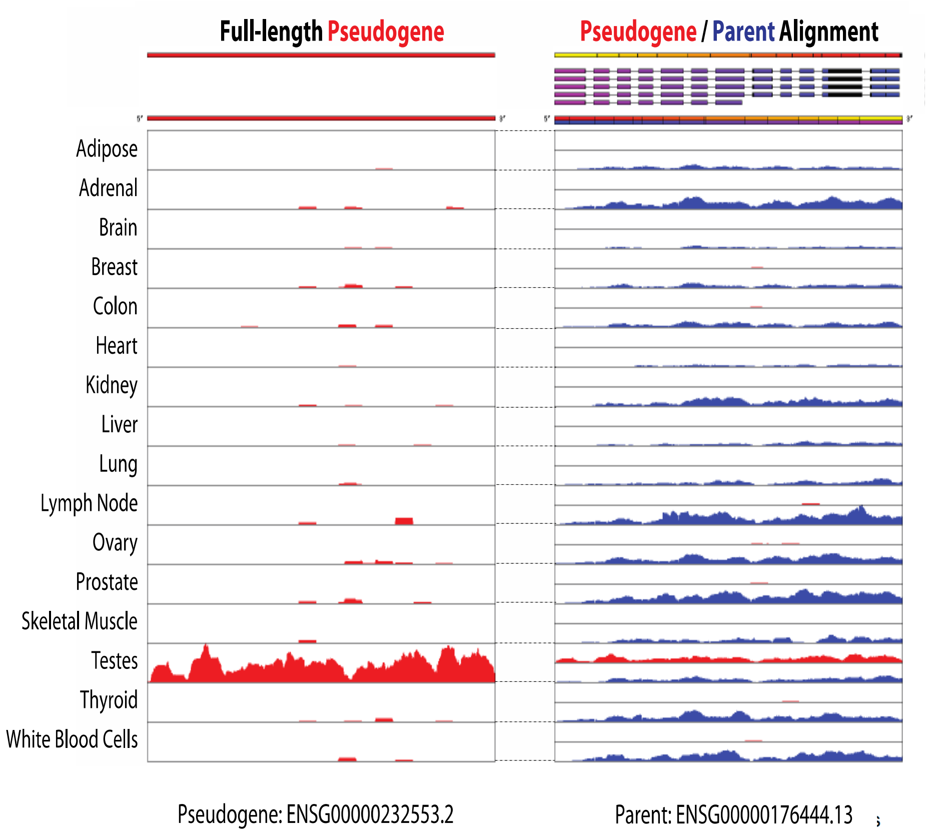
	Human – GENCODE	Worm	Fly
Total	11240 (14112*)	1198	529
Duplicated	2158	538	119
Processed	8715	255	95
Ambiguous	23	405	315
Others**	344		

* Estimated total number of pseudogenes in human genome.

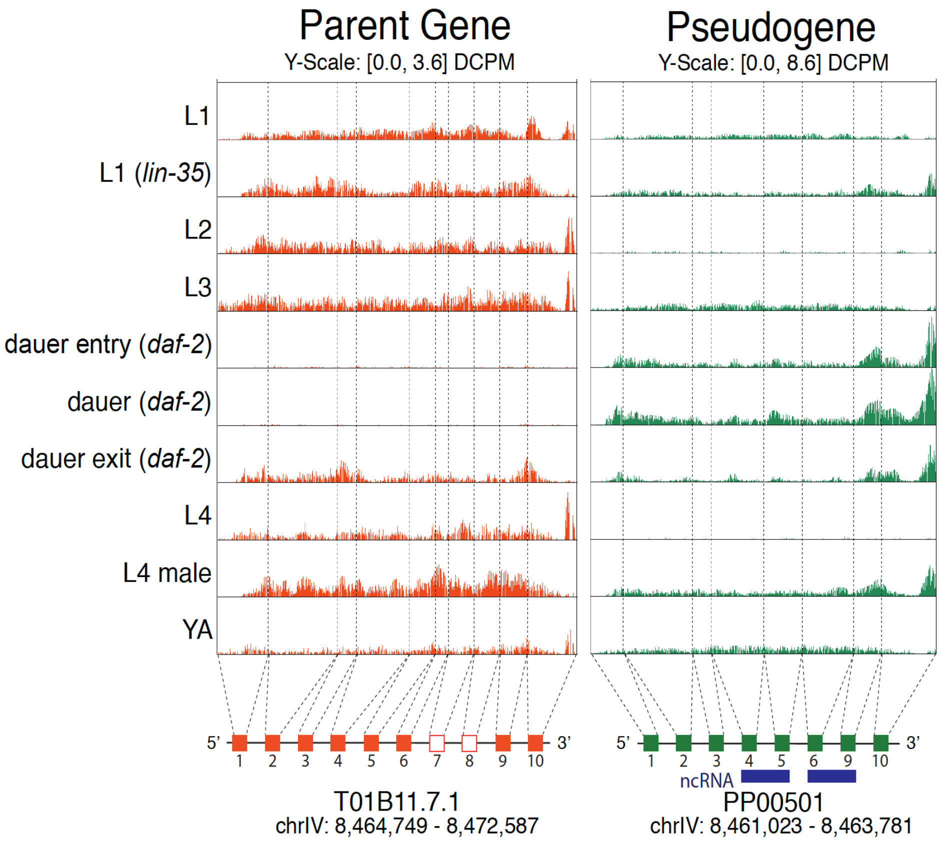
** Including Unitary (138), IG (161) TR V (21) and polymorphic (24) pseudogenes

*Transcribed Pseudogenes

Human

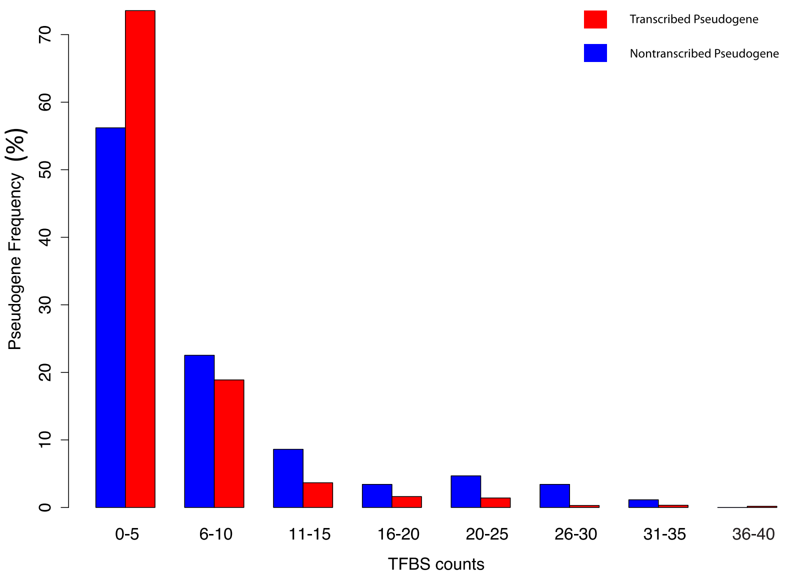


Worm

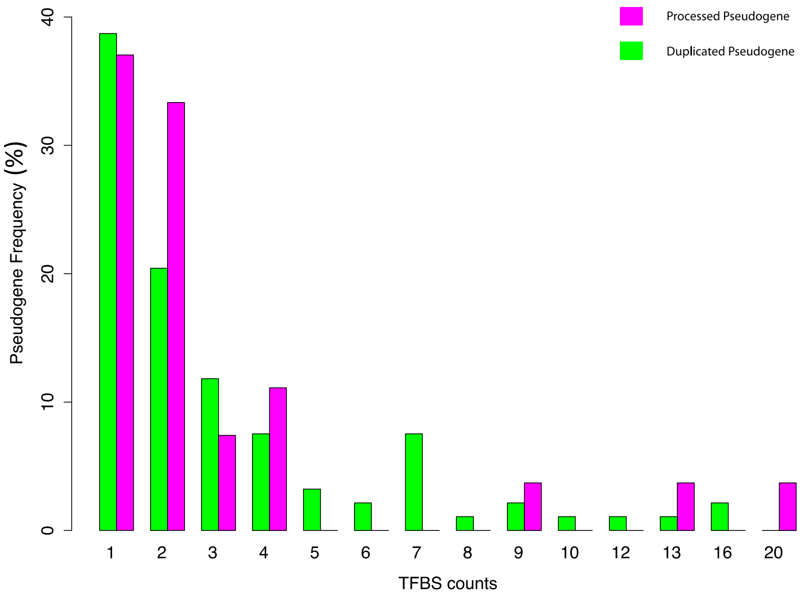


*Transcription Factor Binding Sites

Human



Worm



- TFBS were selected within 2kb upstream of the pseudogene start site
- 95 (58) duplicated and 29 (20) processed pseudogenes had TFBS in the upstream region

Analysis of relationship of upstream regions to transcript level

Analysis of relationship of upstream regions to transcript level

Analysis of relationship of upstream regions to transcript level

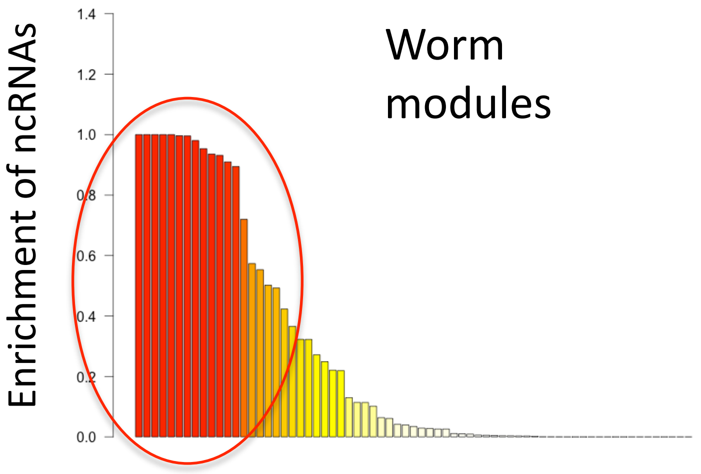
Expression clustering of protein-coding and ncRNA genes in embryo development

Species	Developmental stages	Protein-coding genes*	Non-coding RNAs*	Co-expression modules**
Worm (C. elegans)	111	9114	855	69
Fly (D. mel.)	50	8340	357	46

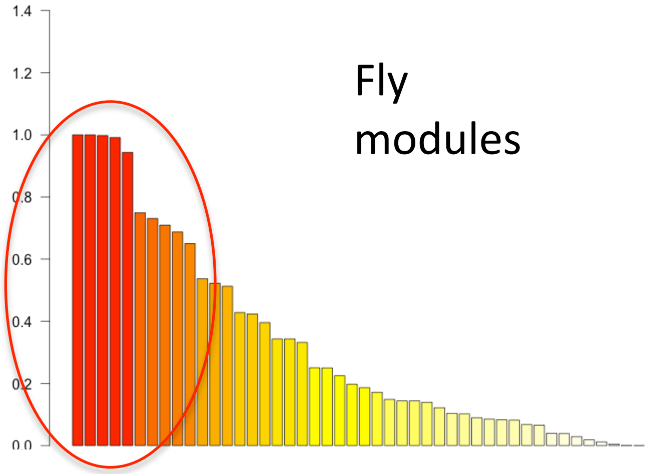
* >80% valid samples, coeff. of variance > 1 in the modENCODE finalized datasets in June 2012

** clustering via weighted gene co-expression network analysis (WGCNA)

Many co-expression modules are enriched with ncRNAs (red circles).



Worm modules

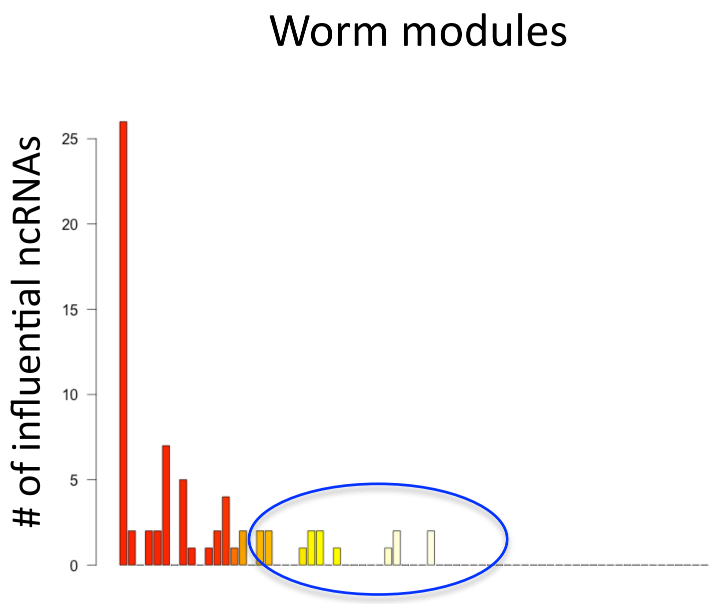


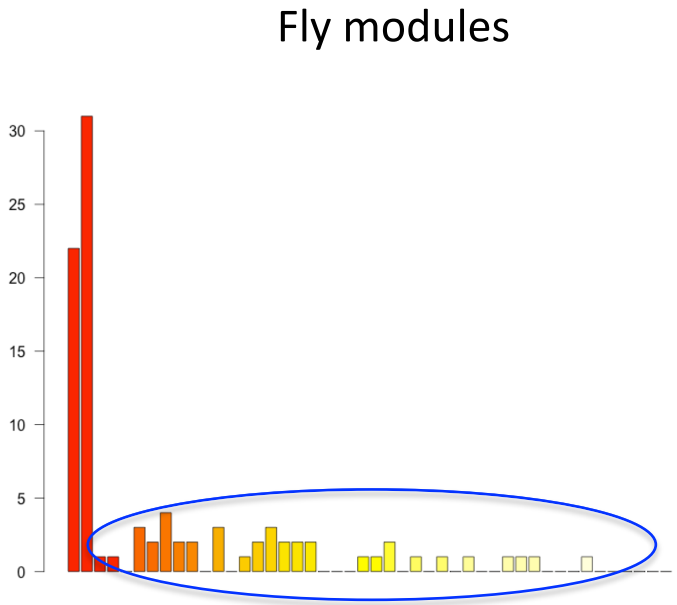
Fly modules

Daifeng Wang

Influence of ncRNA hubs on protein-coding co-expression modules

Influential ncRNAs (high network centrality) exist in modules NOT enriched with ncRNAs (blue circles).



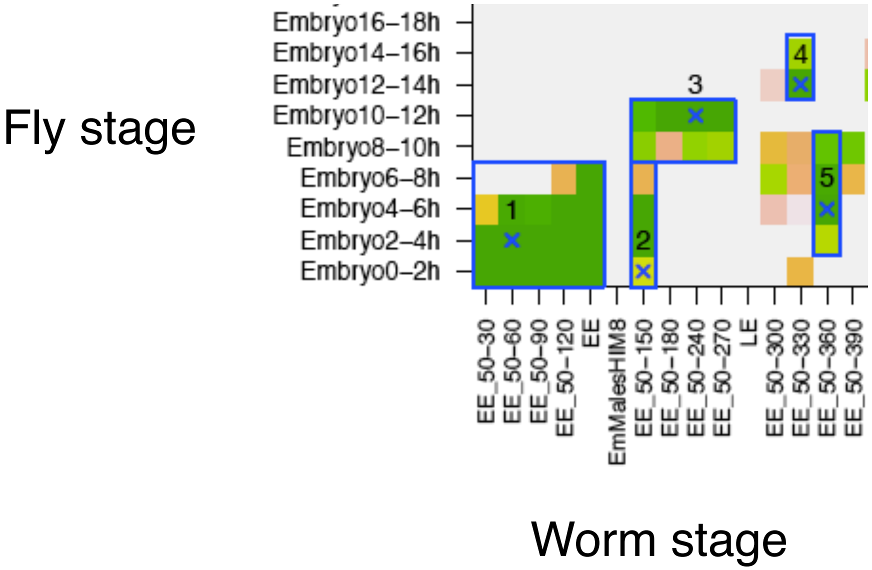


Daifeng Wang

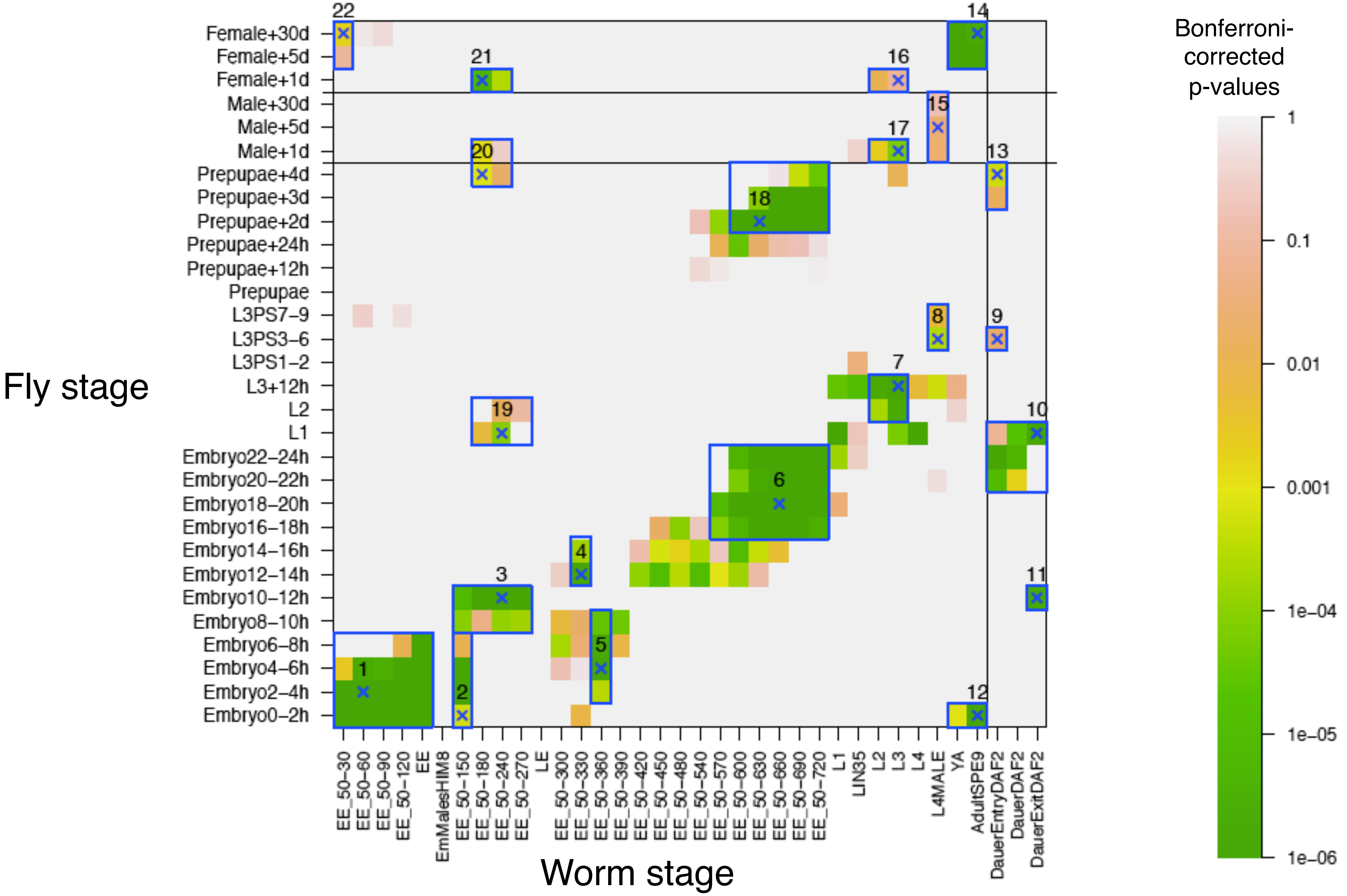
Developmental stage mapping between worm and fly based on co-expression clustering of orthologs

- Gene expression threshold: FPKM >=1 and z >= 1.5
- Significance calculated from fraction of orthologs co-expressed between pairs of stages compared to hypergeometric expectation

- Cluster numbering facilitates follow-on analysis:



Developmental stage mapping between worm and fly based on co-expression clustering of orthologs



Jingyi Jessica Li, Peter Bickel, Haiyan Huang, Steven Brenner

END

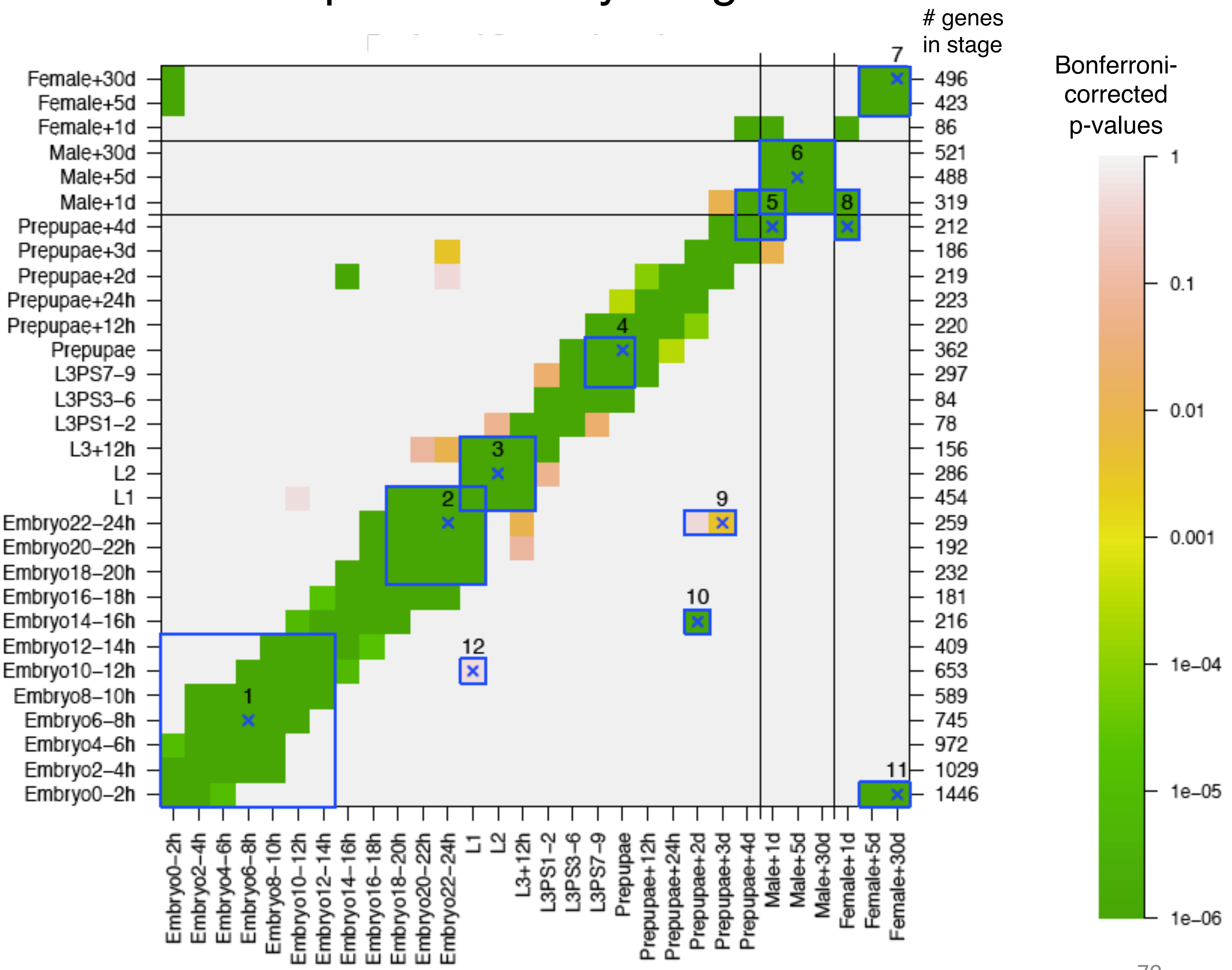
Production Stats - Worm

	Samples	**Total** Reads	**Total Unique** Reads
Embryonic Time Course	106	1,633,419,670	1,031,557,649
Life Stages	70	2,401,311,389	1,420,342,487
Other Species	54	1,779,775,463	946,431,824
Pathogens	11	702,645,329	489,536,643
Tissues	183	3,560,398,393	1,322,552,917
Totals		10,077,550,244	5,210,421,520

Production Stats - Fly

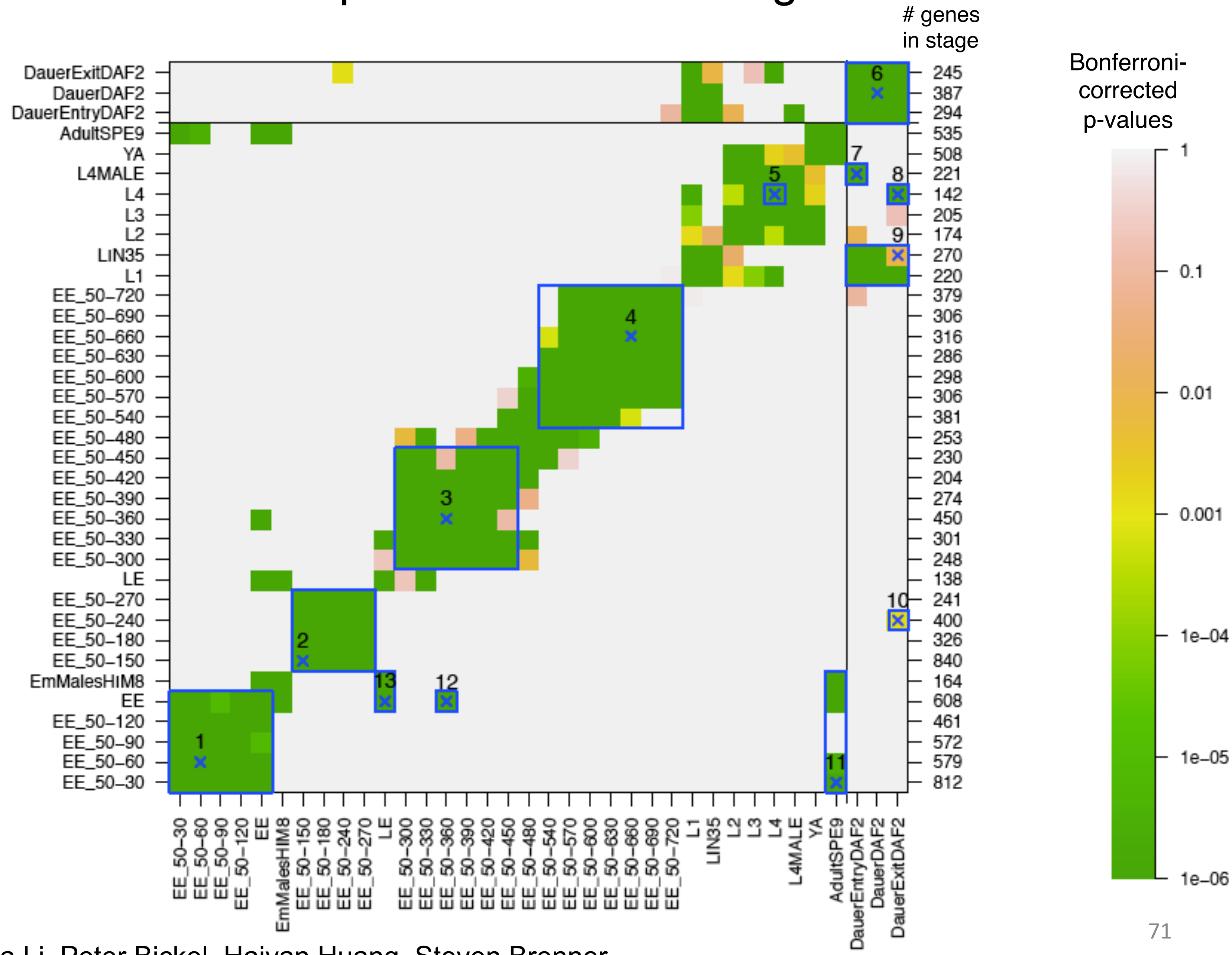
Experiment	Samples	Total Reads	Total Unique Reads	Total Unique bp
Cell Lines	25	1,677,980,920	1,272,452,612	96,706,398,512
Tissues	29	4,265,585,752	3,667,365,400	278,719,770,400
Treatment	21	6,495,812,560	4,949,215,447	376,140,373,972
Poly(A) Tail Enrichment	29	845,610,153	638,882,610	48,555,078,360
Developmental Time Course*	30	3,538,880,404	2,282,408,273	171,180,620,475
Genome Resequencing	25	943,927,826	N/A	71,738,514,776
Total	247	17,767,797,615	12,810,324,342	1,043,040,756,495

Comparison of Fly Stages



Jingyi Jessica Li, Peter Bickel, Haiyan Huang, Steven Brenner

Comparison of Worm Stages



Jingyi Jessica Li, Peter Bickel, Haiyan Huang, Steven Brenner