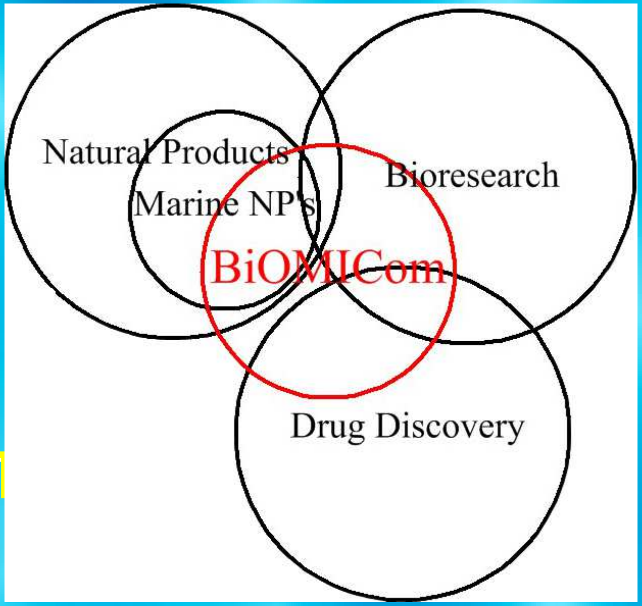


BiOMICom

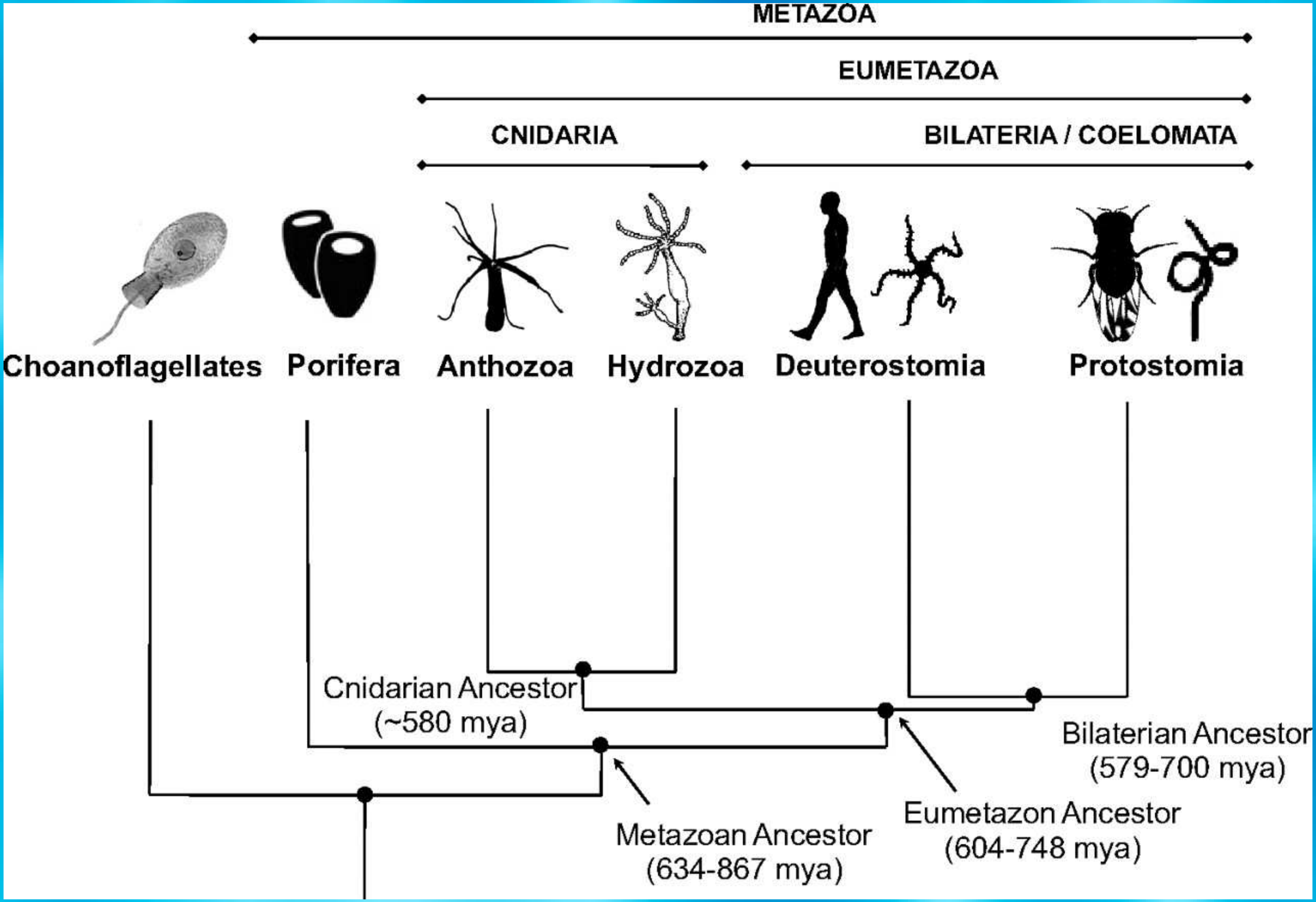
…Analyzing Life

Lineage from the Aquatic
Ecosystem as a Future
Biomedical Life-line

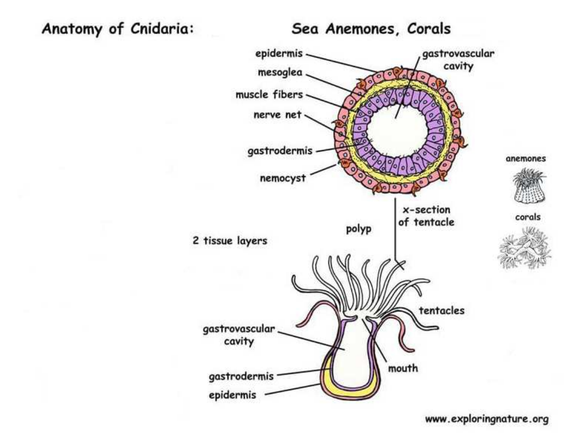
Origins of Life

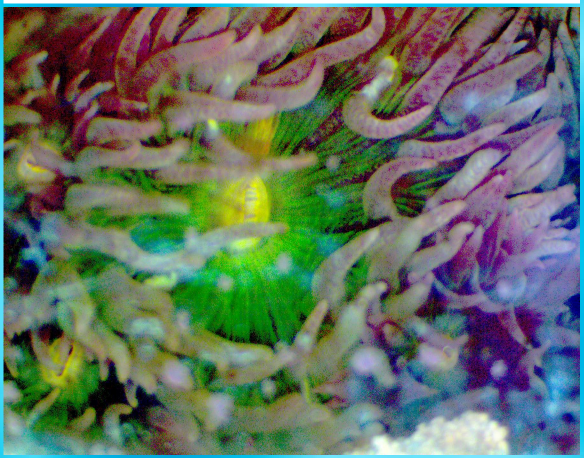
- Most theories assume life began in the seas
- Basic amino acids, building blocks of life, formed
  - Shallows where atmospheric CO_2 and NH_3 reacted on rock surfaces often with other hydrocarbons
  - Undersea hydrothermal vents where many of these same constituents could react under energetically favorable conditions
- The emergence of animals, multicellular eukaroytes of the metazoan kingdom, generally marked to sea anemones, members of the phylum cnidaria, class anthozoa
  - Appearance Ca. 600 Ma
  - Over 6,000 living species in phylum cnidaria
  - Basal metazoan phylum along with Porifera (sponges)

Metazoan Evolution (1)



Anatomy of Phylum Cnidaria
Class Anthozoa





- Common anatomical features:
  - An oral groove constituting both mouth & anus
  - A gastrovascular cavity where food is ingested, digested, and waste products excreted
  - Tentacles to move food to oral groove and to provide defense; nemocysts
  - A nerve net extending from oral groove to tentacles and between polyps in communal ecosystems
  - Two differentiated tissue types; gastrodermal inner and epidermal outer with a gastrovascular canal to share nutrients; diploblastic, bilaterian precursor

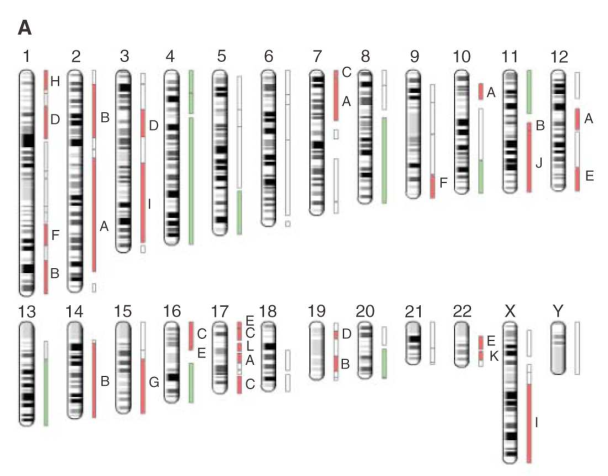
Nematostella vectensis (Nv) Model Organism

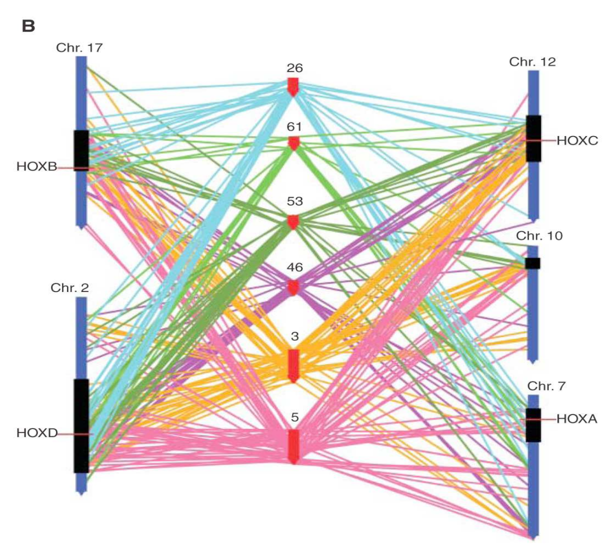
- Starlet sea anemone, estuarine, native to the atlantic coast of north america (1)
  - Native stresses include:
    - Temperature
    - Salinity
    - UV light
    - Oxygen
    - Sulfides
    - Reactive oxygen species
- Physiological tolerance is a function of ability to upregulate stress response genes
- Extensively used for research in developmental biology, evolution, and genomic studies
    - Regenerates and cells grow in culture

Genomics

- Human (homo sapiens; Hs) DNA consists of around 23,000 genes, the same approximate number as *Nematostella vectensis*
- There is remarkable overlap of lower species DNA with Hs DNA
  - Chimpanze differs only 1.2% in sequence (98.8% overlap)
  - Nv DNA has around 50% overlap of sequence
- It must therefore be concluded that the phenotypic diversity of species is in large measure driven by gene expression; transcription & translation

Genomics (2)

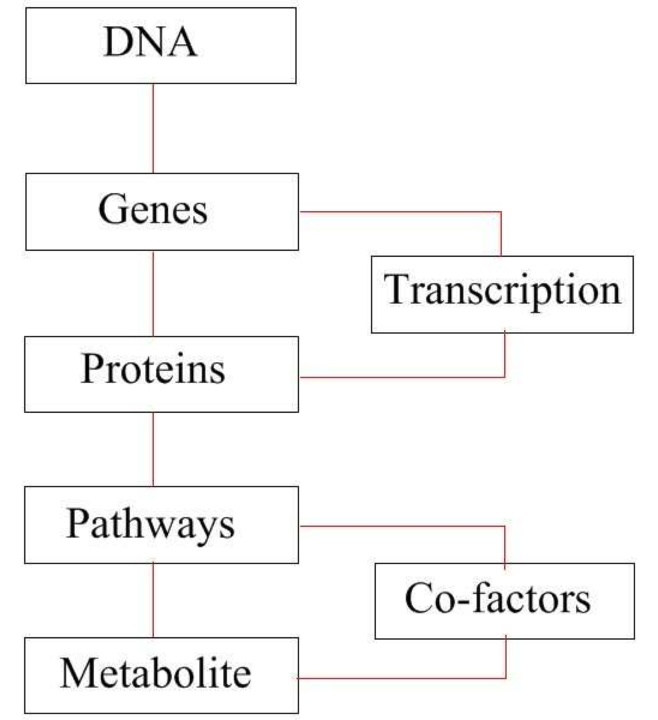




- **Chart A** = human DNA chromosomes (24) segmented into 98 regions; colored segments represent significant conservation of sequence between human and *Nematostella*, white segments do not show significant conservation of linkage

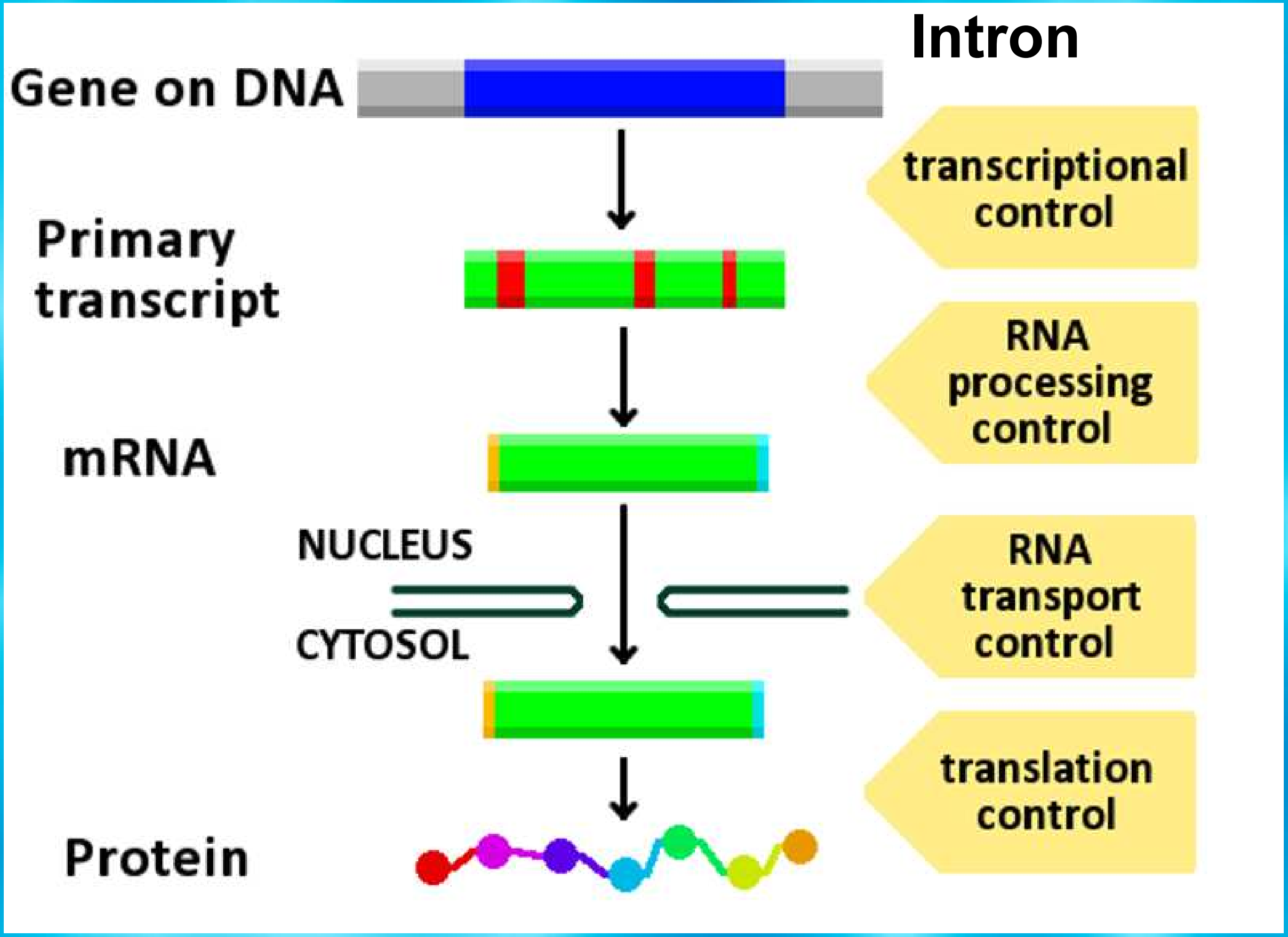
- **Chart B** = conserved linkages between human chromosomes 17, 12, 10, 7 & 2 and *Nematostella* scaffolds 26, 61, 53, 46, 3 & 5 with length proportional to the number of genes descended from the inferred ancestral set color coded from the *Nematostella* scaffold. Red lines indicate the positions of the four human Hox clusters

Molecular Biology



- Human DNA has around 3.2 billion base pairs; 800 mB of data divided into around 23,000 genes
- This information codes for around 20,000-30,000 proteins and RNA transcription nucleotides
- The cnidarian genome is around the same complexity

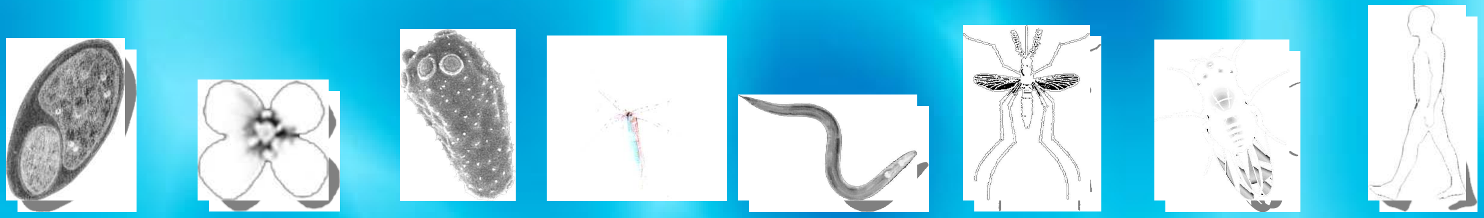
Gene Transcription (3)



Evolutionary Conservation of Introns (4)

- Introns are the non-coding regions between genes and are an important component of eukaryotic genomes
  - Represent 25% of the human genome
  - Represent 95% of the human transcriptome (Venter et al., 2001.)
- Responsible for much of the inherent cis-regulatory elements for DNA transcription
- Alternative transcripts greatly magnify transcriptome complexity
- Introns foster genetic recombination

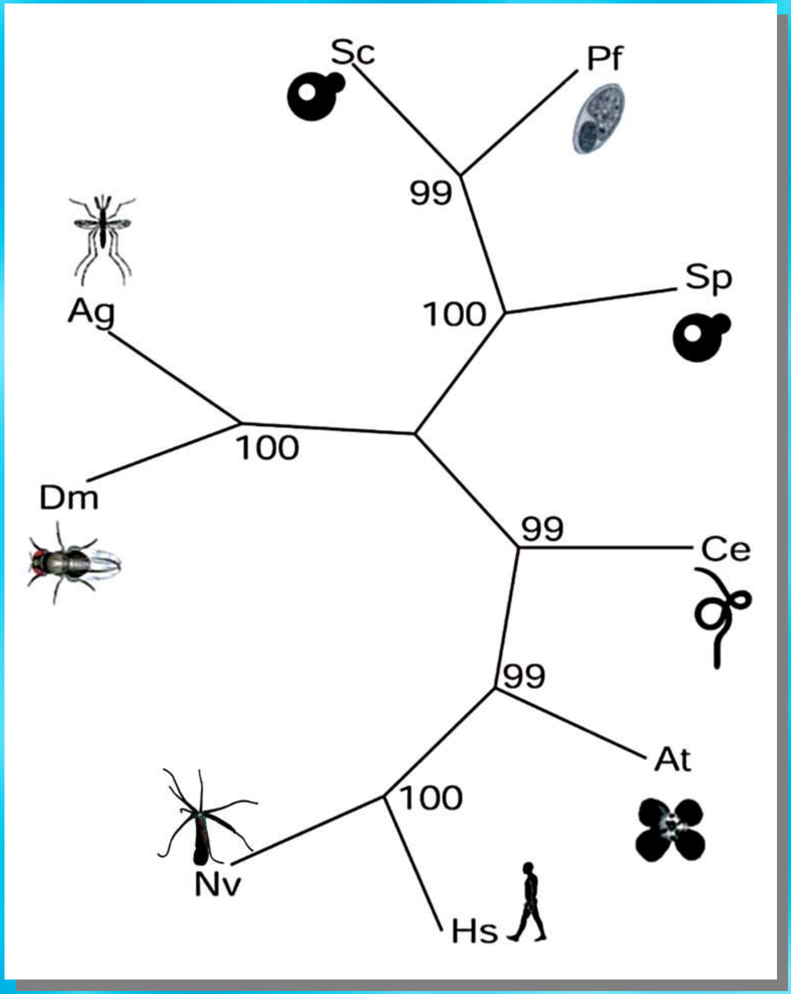
Percentage of Shared Introns (4)



	Pf	At	Sp	Nv	Ce	Dm	Ag	Hs
Pf	159	3%	3%	3%	2%	2%	3%	2%
At		1076	5%	17%	7%	5%	5%	16%
Sp			173	6%	4%	7%	6%	7%
Nv				1459	14%	9%	10%	47%
Ce					560	11%	11%	15%
Dm						260	43%	13%
Ag							269	12%
Hs								1246

$$= \frac{\text{(Introns shared by A \& B)}}{\text{(Introns shared by A \& B)} + \text{(Introns unique to A)} + \text{(Introns unique to B)}}$$

Shared Introns (4)



- Diagonal of previous slide represents total intron count of selected taxa
- Percentages on previous slide represents total intron overlap
  - *Nematostella vectensis* (Nv) intron overlap with homosapien (Hs) is 47%
  - Nv possesses 69% of all Hs introns (862/1246)
- Highest overlap for any lower metazoan comparison; Hs and Nv grouped alone and together using parsimony analysis of this data

microRNA & PiwiRNA (5)
Conservation (cis-regulation)

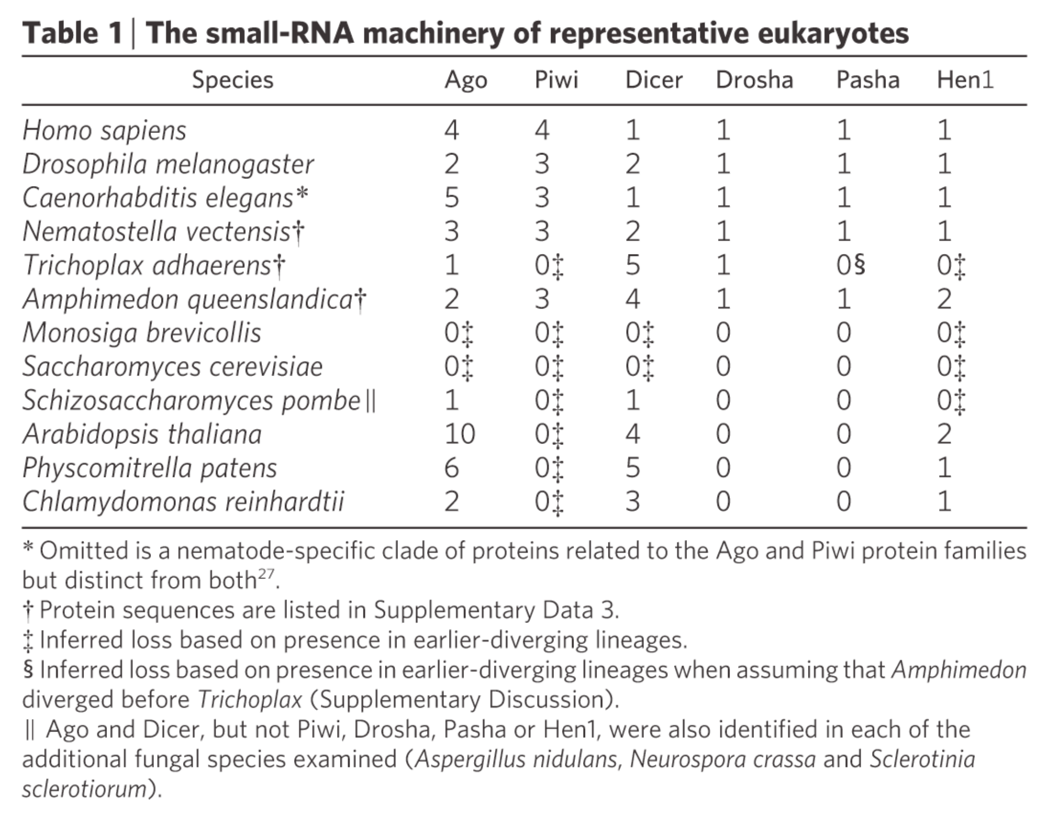


Table 1 | The small-RNA machinery of representative eukaryotes

Species	Ago	Piwi	Dicer	Drosha	Pasha	Hen1
Homo sapiens	4	4	1	1	1	1
Drosophila melanogaster	2	3	2	1	1	1
*Caenorhabditis elegans**	5	3	1	1	1	1
Nematostella vectensis†	3	3	2	1	1	1
Trichoplax adhaerens†	1	0‡	5	1	0§	0‡
Amphimedon queenslandica†	2	3	4	1	1	2
Monosiga brevicollis	0‡	0‡	0‡	0	0	0‡
Saccharomyces cerevisiae	0‡	0‡	0‡	0	0	0‡
Schizosaccharomyces pombe‖	1	0‡	1	0	0	0‡
Arabidopsis thaliana	10	0‡	4	0	0	2
Physcomitrella patens	6	0‡	5	0	0	1
Chlamydomonas reinhardtii	2	0‡	3	0	0	1

* Omitted is a nematode-specific clade of proteins related to the Ago and Piwi protein families but distinct from both[27].

† Protein sequences are listed in Supplementary Data 3.

‡ Inferred loss based on presence in earlier-diverging lineages.

§ Inferred loss based on presence in earlier-diverging lineages when assuming that *Amphimedon* diverged before *Trichoplax* (Supplementary Discussion).

‖ Ago and Dicer, but not Piwi, Drosha, Pasha or Hen1, were also identified in each of the additional fungal species examined (*Aspergillus nidulans*, *Neurospora crassa* and *Sclerotinia sclerotiorum*).

- 30% of the Human DNA transcription uses small RNA (micro 21-24 nt; Piwi 25-30 nt) binding
- Conservation of these transcription mechanisms throughout bilaterian evolution

Metabolome

- The metabolome of an organism is comprised of all of the small molecules (less than around 2,000 Da) produced by all biochemical processes. Metabolites are generally a result of the catalysis properties of proteins within pathways
- Two groups of metabolites
  - Primary metabolites produced as a direct result of basic life sustaining processes; energy production or storage, reproduction, cell function maintenance, etc.
  - Secondary metabolites produced as a direct result of external or endogenous stresses and are considered critical to the long term survival of the organism
  - Secondary metabolites of marine invertebrates have been viewed as rich sources of bioactive chemical agents useful in the treatment of a wide variety of disease states

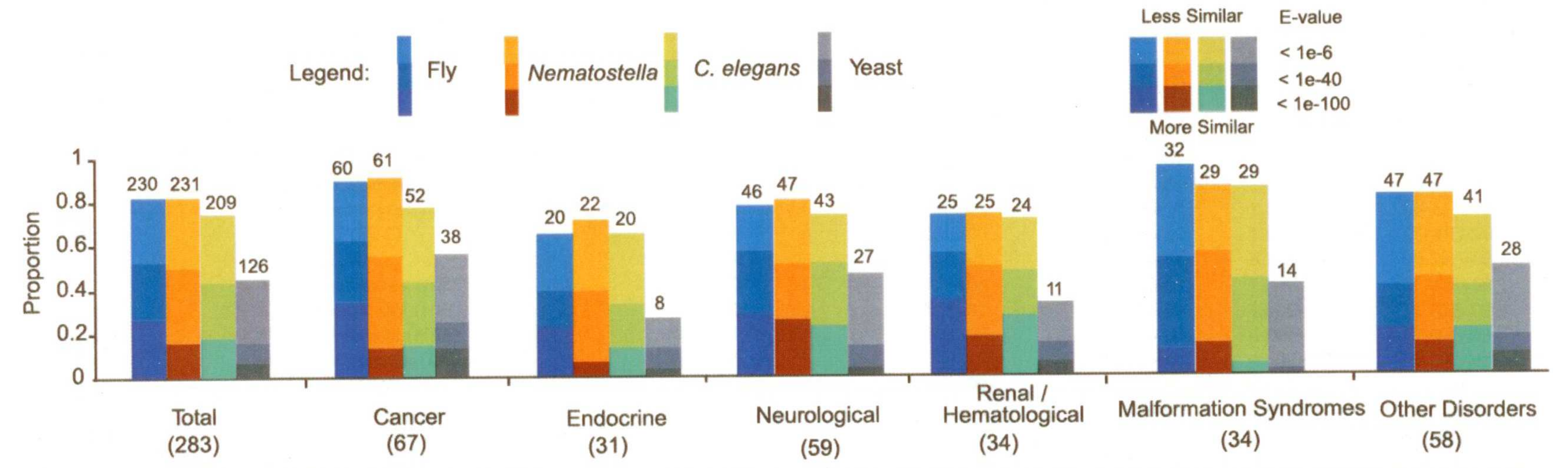
Defensome (6)

- Gene families that defend against chemical stressors, and the transcription factors that regulate these genes, have in total been termed the "chemical defensome"[3]
- For the starlet sea anemone *Nematostella vectensis* Nv DNA this ratio is around 1% (266/27,200), about the same ratio as humans

Gene Preservation & Expression

- Remarkably high degree of preserved and expressed genes comparing the human & cnidarian genomes
  - Cell-Cell signaling: the four major bilaterian signaling pathways (7)
    - Wnt genes: 11 of 12 subfamilies (8)
    - TGF-b: cytokine protein
    - Notch: notch transmembrane receptors
    - Hedgehog
  - Defensome (6)
    - CYP; cytochrome P-450 oxidation
    - Various conjugating enzymes
      - Human: 13 SULT, 13 UGT, 21 GST
      - Sea anemone: 22 SULT, 9 UGT, 28 GST
    - ATP-dependent efflux transporters
    - Oxidative detoxification genes

Human Disease Genes (9)



- Ns has more orthologous human disease genes than other nearer neighbors to Hs
- Genbank gene reference count shown under disease state in parenthesis

Marine & Human Systems

- Both ocean and human body fluids are aqueous based
- Chemical exchange occurs between cells in the extracellular matrix (ECM) and in marine environment between organisms via both contact and diffusion
- pH
  - Ocean = 8.15-8.25
  - Human serum = 7.3-7.4
  - Pancreatic = 8.1
- Salinity
  - Ocean = 34-36 ppt
  - Human serum = 9 ppt
- The diploblastic nature of cnidarians, and the access to both tissue types, may be advantageous due to a higher chemical exchange with the surrounding ecosystem than a triploblastic organism having a mesoderm

Conclusions

- There has been substantial conservation of important gene families, and pathways, from the basal cnidarian phylum to the homo sapien species.
- There has been substantial conservation of small RNA machinery and intron regions responsible for gene transcription from the basal cnidarian phylum to the homo sapien species.
- Although many conserved & expressed genes may be used in the basal invertebrates for different functions, the pathway regulation agonists & antagonists maintain relevance for human bioactivity (assuming proper ADME and toxicity properties).
- The process of drug discovery within the marine secondary metabolome could be significantly enhanced using experimental design incorporating biotic and abiotic stress factors.
- The entire Cnidarian phylum is now receiving renewed attention for marine natural product prospecting (10)

References

1. "Genomic Survey of Candidate Stress-Response Genes in the Estuarine Anemone *Nematostella vectensis*", JR Finnerty, et al., *Biol.Bull.* **214**: 233-254 (June 2008).
2. "Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic Organization", DS Rokhsar, et al., *Science* **317**: 86-94 (6 July 2007).
3. "Regulation of Gene Expression", www.wikipedia.org.
4. "A High Percentage of Introns in Human Genes Were Present Early in Animal Evolution: Evidence from the Basal Metazoan *Nematostella vectensis*" JC Sullivan, AM Reitzel, JR Finnerty, *Genome Informatics* **17(1)**: 219-229 (2006).
5. "Early Origins and Evolution of microRNAs and Piwi-interacting RNAs in Animals", A Grimson, et al., *Nature* **455**: 1193-1198 (Oct. 2008).
6. "Environmental Sensing and Response Genes in Cnidaria: The Chemical Defensome in the Sea Anemone *Nematostella vectensis*", JV Goldstone, Cell Biol. Toxicol. **24(6)**: 483-502 (Dec. 2008).
7. "The Developmental Origins of Animal Bodyplans", DH Erwin, Topics in Geobiology, **Vol 27**, Chapter 6, *Neoproterozoic Geobiology and Paleobiology*, Springer (2006).

References

8) “Functional Conservation of Nematostella Wnts in Canonical and Non-canonical Wnt-signaling”, D Wedlich *Biology Open* **1**, 43-51 (2011).
9) “A Surprising Abundance of Human Disease Genes in a Simple “Basal” Animal, the Starlet Sea Anemone (*Nematostella vectensis*)”, JC Sullivan & JR Finnerty, *Genome* **50**: 689-692 (2007).
10) “Cnidarians as a Source of New Marine Bioactive Compounds-An Overview of the Last Decade and Future Steps for Bioprospecting”, J Rocha, et al., *Marine Drugs* **9**: 1860-1886 (2011).