

Long non-coding regulatory RNAs in sponges and insights into the origin of animal multicellularity

Federico Gaiti, Bernard M. Degnan & Miloš Tanurdžić


To cite this article: Federico Gaiti, Bernard M. Degnan & Miloš Tanurdžić (2018): Long non-coding regulatory RNAs in sponges and insights into the origin of animal multicellularity, RNA Biology, DOI: [10.1080/15476286.2018.1460166](https://doi.org/10.1080/15476286.2018.1460166)

To link to this article: <https://doi.org/10.1080/15476286.2018.1460166>



Accepted author version posted online: 04 Apr 2018.




[Submit your article to this journal](#) 



Article views: 21



[View related articles](#) 



[View Crossmark data](#) 

Publisher: Taylor & Francis

Journal: *RNA Biology*

DOI: <https://doi.org/10.1080/15476286.2018.1460166>

Long non-coding regulatory RNAs in sponges and insights into the origin of animal multicellularity

Federico Gaiti¹, Bernard M. Degnan, and Miloš Tanurdžić *

School of Biological Sciences, University of Queensland, Brisbane, Australia

*Corresponding author

¹Current address: Department of Medicine, Weill Cornell Medicine, and New York Genome Center, New York, New York, USA

Email addresses:

FG: feg2007@med.cornell.edu

BMD: b.degnan@uq.edu.au

MT: m.tanurdzic@uq.edu.au

Author ORCIDs:

FG: 0000-0001-5111-8816

BMD: 0000-0001-7573-8518

MT: 0000-0002-7564-0868

Corresponding author contact details:

Miloš Tanurdžić

School of Biological Sciences

University of Queensland

Brisbane QLD 4072 Australia

E-mail: m.tanurdzic@uq.edu.au

Keywords: evolution, gene regulation, animals, long non-coding RNAs, multicellularity, histone modifications, *cis*-regulation, enhancers

Abstract

How animals evolved from a single-celled ancestor over 700 million years ago is poorly understood. Recent transcriptomic and chromatin analyses in the sponge *Amphimedon queenslandica*, a morphologically-simple representative of one of the oldest animal phyletic lineages, have shed light on what innovations in the genome and its regulation underlie the emergence of animal multicellularity. Comparisons of the regulatory genome of this sponge with those of more complex bilaterian model species and even simpler unicellular relatives have revealed that fundamental changes in genome regulatory complexity accompanied the evolution of animal multicellularity. Here, we review and discuss the results of these recent investigations by specifically focusing on the contribution of long non-coding RNAs to the evolution of the animal regulatory genome.

Multicellular life has evolved independently in at least 10 different eukaryotic lineages as diverse as animals, fungi, plants, slime molds and seaweeds [1-3]. The transition from a unicellular to a multicellular lifestyle required the emergence of genomic regulatory systems to allow for dynamic spatiotemporal and cell-type specific gene expression, which allows for greater cell type functional diversity and specialization (reviewed in [4]). This process has been deemed to be orchestrated by the interplay between regulatory genes, including transcription factors (TFs) and signaling molecules, as well as with non-coding regulatory DNA and RNA sequences [5]. Although it is widely appreciated that these systems ultimately evolved from genomic regulatory mechanisms present in single-celled ancestors, the questions of *how* and *when* did multicellularity evolve, and *what* genomic innovations underpin these evolutionary steps remain a focus of evolutionary and developmental biology.

While many TFs and signaling genes evolved after the divergence of animals and their closest living relatives (*e.g.*, choanoflagellates) and, thus, are correlated with the evolution of animal multicellularity (*e.g.*, [6-8]), others have an older origin and are found in unicellular organisms [9-18]. For instance, an early burst in the diversity of LIM homeobox TFs and cell adhesion genes (*e.g.*, Type IV collagens) occurred in unicellular holozoans well before the emergence of the first animals, consistent with the later co-option of these gene families into roles in multicellular development [19]. The striking conservation of the TF family repertoire in non-bilaterian multicellular animal lineages (cnidarians, placozoans, ctenophores and sponges) – and to a lesser extent unicellular holozoan lineages – further supports the premise that the evolution of animal multicellularity must have been a result of the evolution of other regulatory features that are likely responsible to control *how* and *when* genes are employed during animal development, including *cis*-regulatory DNA, as well as chromatin modifications and non-coding RNAs (ncRNAs) [5].

Molecular reconstruction of the most recent common ancestor of animals is, therefore, crucial to determine which of these regulatory features, if not all, were

instrumental in the origin of animal multicellularity. Sponges are one of the earliest-branching animal lineages, diverging from other animals around 700 Mya [20]. Hence, key insights into animal origin could be gained through the comparison of genomic traits shared between sponges and all other animals (*i.e.*, eumetazoans). Since their divergence, sponges and eumetazoans have had radically different evolutionary histories, with the ancestor of bilaterians, cnidarians and placozoans giving rise to a range of morphologically-complex body plans, and the ancestor of sponges yielding one morphologically-simple body plan. However, despite sharing a remarkably similar repertoire of developmental genes [5, 21-25], these disparate evolutionary trajectories had yet to be reconciled in terms of regulatory non-coding genome content and organization.

Here, we review recent discoveries of long non-coding RNAs and their associations with specific chromatin states in the sponge *Amphimedon queenslandica*, and discuss their implications for the evolution of gene regulation and animal multicellularity.

Animal long non-coding RNA-based regulation: insights from the sponge

Amphimedon queenslandica

The application of next-generation sequencing technologies over the last decade, largely in bilaterian model species, has revealed that animal genomes encode thousands of lncRNAs [26-39], which may be sense or antisense, intronic, and intergenic with respect to protein-coding genes [40]. lncRNAs range in size from ≥ 200 nucleotides to ≥ 10 kilobases in length, and are often multi-exonic and polyadenylated. Despite lacking obvious protein coding potential, lncRNAs perform a wide range of regulatory roles beyond RNA's cardinal function in the flow of genetic information (see *e.g.*, [41-44]). These regulatory roles of lncRNAs can be performed by the lncRNA transcripts themselves, or they may be mediated by lncRNAs involvement in higher-order chromatin structure, recruitment of regulatory protein complexes through RNA-protein interactions, or a combination thereof (reviewed in [45]). For instance, lncRNAs act as scaffolds to bring two or more proteins into a complex or in physical proximity [42, 46]. lncRNAs can also act as guides to recruit chromatin modifying enzymes and can be required for localization of ribonucleoprotein complexes to specific targets [41, 43]. Finally, several lncRNAs have been shown to act as decoys that titrate away microRNAs or regulatory proteins [41]. These roles suggest that evolutionary innovations

involving regulatory lncRNAs were a cornerstone for increasing complexity of genetic regulation in animals.

Unlike many protein-coding sequences, lncRNA genes are rapidly evolving and exhibit poor primary sequence similarity between species; orthologous lncRNAs are difficult to identify [47], thus precluding a detailed understanding of their evolution in terms of sequence, structure and function. In fact, while the role of lncRNAs in the regulation of developmental gene activity now appears to be widespread amongst animals [23, 48-65], only a handful of lncRNAs have thus far been shown to possess conserved function(s) in evolutionarily divergent animals; all functional studies currently are restricted to bilaterians [52, 58, 66-68] (Fig. 1A). Understanding the early evolution of these putative master orchestrators can contribute to reconstructing the origin of animal gene regulatory complexity and multicellularity.

Animal lncRNAs tend to be abundantly expressed in discrete cell types [69, 70] and exhibit more tissue [31] and developmental stage specificity [71-74] than protein-coding genes at different expression ranges, suggesting that animal development requires the fine-scale regulation of expression of specific lncRNAs [75]. Consistent with this, we recently showed that the sponge *Amphimedon queenslandica* expresses an array of lncRNAs akin to their bilaterian counterparts (*i.e.*, in a spatiotemporal and cell type-specific manner) [23, 76]. The analysis of *Amphimedon* lncRNAs expression profiles has indeed revealed that lncRNA abundances correlate with morphogenetic and developmental milestones [23, 76] – a hallmark of regulatory molecules (Fig. 1B, C). For example, while the complexity of the morphogenetic events during *Amphimedon* early embryo cleavage is reflected in the high diversity of the lncRNAs expressed at this stage, the subsequent embryonic stages have markedly fewer lncRNAs expressed at high levels [23]. This observation is similar to previous findings in bilaterians, where early embryonic stages appear to be a period of active transcription of lncRNAs [23, 31, 51, 63, 75, 77], perhaps to regulate maternal transcripts or transcription of cell-cycle genes [75].

The highly dynamic and tightly regulated expression of lncRNAs in sponges and bilaterians suggests these features were present in their last common ancestor. However, the origin and evolution of this class of non-coding RNAs remain unclear.

Evolutionary conservation of animal lncRNAs: homology or co-option?

The scarcity of lncRNA annotations (especially in non-bilaterian animals) and their rapid sequence divergence has posed challenges to understand lncRNAs evolution [47]. Only a handful of functionally homologous bilaterian lncRNAs have been identified and analyzed [52, 58, 66-68]. These studies suggest that sequence conservation is not an essential requirement for lncRNA functionality [78]. Nonetheless, highly conserved elements within lncRNA sequences (micro-homologies), interspersed with longer and less conserved stretches of nucleotide sequences, have been reported [79, 80], and include the *miR-7* binding site in the lncRNA *Cyrano* [52], the PRC2-binding elements in the lncRNA *Xist* [81], and short sequences derived from Alu repeat elements [82]. Other lncRNA features that appear to be conserved include syntenic relationships to neighboring genes [58, 79, 83], conservation of secondary structure [84-87], and specific expression patterns [79, 88, 89].

As in bilaterians where lncRNAs have been shown to be co-expressed with multiple protein-coding genes [50], *Amphimedon* lncRNAs also appear to belong to co-expressed developmental gene modules [23]. By comparing gene co-expression networks between *Amphimedon* and *Sycon ciliatum* [63] – a distantly-related calcisponge – we have recently identified several putative evolutionarily conserved developmental modules of co-expressed homologous genes and lncRNAs in sponges [76]. This is despite the lack of sequence similarity between the network-embedded sponge lncRNAs. One such example is comprised of two lncRNAs found in *Amphimedon* and *Sycon* sponges, differing in sequence but sharing in their co-expression, and thus presumably co-regulation, with the G protein-coupled receptor *Frizzled B* (a key component of the Wnt signaling pathway in animal development) and other regulatory genes (*e.g.*, *TGF-beta*) [23, 63, 76]. As gene regulatory networks and modules are central for the control and timing of animal development [90-92], the finding of similar sets of homologous protein-coding genes co-expressed with lncRNAs between evolutionarily divergent sponge species points to lncRNAs being developmental regulators that might operate in conserved gene regulatory networks. Given the lack of sequence identity of these lncRNAs and of functional data in sponges, the independent co-option of non-homologous lncRNAs into these modules cannot be discounted at this time.

Recent discoveries of lncRNAs in unicellular lineages closely related to animals generate further uncertainty about the evolutionary origin of animal lncRNAs. In fact, while

lncRNAs appear to be greatly expanded in multicellular animals [23, 63, 80, 93], several hundred lncRNAs have now been annotated in two of animal closest unicellular relatives, *Capsaspora owczarzaki* and *Creolimax fragrantissima* [94, 95]. This is consistent with this class of non-coding RNAs antedating the origin of animal multicellularity and development.

The origin of animal *cis*-regulatory complexity: Insights from the sponge histone modifications landscape

If differences in information content (*e.g.*, proteome or lncRNAome size) between animals and their unicellular relatives cannot fully account for their phenotypic differences, evolutionary innovations in gene regulatory mechanisms, rather than the introduction of novel regulatory genes (*i.e.*, TFs and lncRNAs), could have been a crucial step in the emergence of animal multicellularity. Recent chromatin profiling of histone H3 post-translational modifications (PTMs) in *Amphimedon* has shown that that sponges use the same chromatin-based regulatory system found in more complex animals [96]. Based on the presence and genomic location of specific histone PTMs patterns (*e.g.*, H3K27me₃, H3K4me₁ and H3K27ac) associated with regulatory features that had previously only been identified in more complex animals, these results identified novelties that may underlie the emergence of multicellular animals, including epigenetic memory mediated by chromatin repressors, and distal *cis*-regulatory elements (*i.e.*, enhancers) [96, 97].

Enhancers are regulatory elements critical for accurate spatiotemporal and cell type-specific expression of the genes that regulate development in animals and are distinguished by a unique chromatin signature of histone H3 lysine 4 and 27 monomethylation/acetylation (H3K4me₁ and H3K27ac) [98-100]. Similar to eumetazoans, analysis of predicted enhancer elements in *Amphimedon* revealed that these regulatory elements also contain putative binding sites for TFs that are important for animal development (*e.g.*, SOX) [96, 101-105]. Their identification in *Amphimedon* is, therefore, consistent with this regulatory feature evolving along the metazoan stem at the transition to multicellularity, as signatures of animal enhancers have not been detected at regulatory sites of closely related unicellular holozoan sister taxa [94] (Fig. 2A).

Enhancer elements are known to be associated with the transcription of non-coding RNA transcripts, termed enhancer RNAs (eRNAs) [106-108]. The majority of eRNAs have

been shown to be short, unspliced and non-polyadenylated transcripts. However, a second subset of long intergenic non-coding RNAs displaying enhancer-like activity (elincRNAs) has recently been described. These polyadenylated transcripts arise from genomic locations with typical enhancer properties such as enrichment of histone H3 monomethylated at lysine 4 (H3K4me1) and depletion of histone H3 trimethylated at lysine 4 (H3K4me3) chromatin marks. This is in contrast to yet another type of lincRNA, the more canonical promoter-associated lincRNAs (plincRNAs; low H3K4me1-to-H3K4me3 ratio) [94, 109-112]. These two distinct populations of chromatin-enriched lincRNAs were also recently discovered in *Capsaspora* [94] and *Amphimedon* [96] (Fig. 2B). Similar to bilaterians, these two lincRNA populations showed only minor differences in length, expression level, and expression variation [94]. While their functional significance is yet to be determined, they may represent distinct classes of polyadenylated lincRNA transcripts with putative diverse biological functions, consistent with elaborate genome regulation by lincRNAs being already present in unicellular holozoans [94, 109-112].

What can we learn about the evolution of multicellularity by investigating other lineages?

Animals are not the only multicellular organisms, and thus not the only system suitable to the study of multicellularity's origins. It has evolved independently multiple times across the tree of life, including at least 10 times across eukaryotes [3, 113]. For instance, studies on the evolution of multicellular plants showed that multicellularity evolved several times across two main plant lineages: streptophytes (charophyte algae and all land plants) and chlorophytes (sister clade which includes green algae) [114-116]. Comparative genomic analyses between two representatives of the chlorophytes – the single cell green alga *Chlamydomonas* and its multicellular relative *Volvox* [117] – show patterns of diversification in gene content very similar to what has been observed in animals, with very few differences between *Chlamydomonas* and *Volvox* genomes that could explain the drastic differences in their morphologies.

While land plants employ the same tool kit of chromatin-based gene regulatory mechanisms [118], the roles of non-coding RNAs and histone PTMs in the evolution of plant multicellularity remain unknown. In fact, while recent reports show that gene repression by

PRC2-mediated histone methylation is present in unicellular algae [119, 120], its role(s) in determining the spatiotemporal and cell-specific gene expression patterns in early multicellular plant lineages, together with the role(s) and evolution of lncRNAs, remain unknown at this time.

Concluding remarks

Recent in-depth analysis of gene regulation in the sponge *Amphimedon queenslandica* shows that fundamental changes in the non-coding regulatory architecture of the genome occurred along the metazoan stem, in concert with the evolution of the multicellular condition. It now appears that most of the genes and long non-coding regulatory mechanisms underlying the formation of complex animals, like ourselves, had an unexpected early origin – probably as early as the first steps in the evolution of multicellular animals from single-celled organisms, at least 700 Mya. Thus, the first animals likely evolved from a unicellular ancestor through the co-option of multiple ancestral gene modules, as well as the evolution of new coding gene families, distal enhancer elements and other classes of non-coding RNAs, such as microRNAs and Piwi-interacting RNAs [122]. These changes are likely to have contributed an increase in the capacity to regulate spatial and temporal gene expression, which is necessary for complex multicellularity. With a complex gene regulatory landscape already in place at the dawn of animals, the further differential expansion of genomic regulatory repertoires in bilaterian animals likely account for their increased regulatory and morphological complexity relative to non-bilaterian animals.

Acknowledgments

This work was supported by an Australian Research Council grant (FL110100044) to BMD.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Knoll AH. The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences*. 2011;39(1):217-239. doi: 10.1146/annurev.earth.031208.100209.
2. Nguyen TA, Cisse OH, Yun Wong J, et al. Innovation and constraint leading to complex multicellularity in the Ascomycota. *Nature Communications*. 2017;8:14444. doi: 10.1038/ncomms14444.
3. Niklas KJ, Newman SA. The origins of multicellular organisms. *Evolution & Development*. 2013;15(1):41-52. doi: 10.1111/ede.12013.
4. Brunet T, King N. The Origin of Animal Multicellularity and Cell Differentiation. *Developmental Cell*. 2017;43(2):124-140. doi: <https://doi.org/10.1016/j.devcel.2017.09.016>.
5. Davidson EH, Peter IS. *Genomic control process*. Oxford: Academic Press; 2015.
6. Larroux C, Luke GN, Koopman P, et al. Genesis and expansion of metazoan transcription factor gene classes. *Molecular Biology and Evolution*. 2008;25(5):980-996. doi: 10.1093/molbev/msn047.
7. Degnan BM, Vervoort M, Larroux C, et al. Early evolution of metazoan transcription factors. *Current Opinion in Genetics & Development*. 2009;19(6):591-599. doi: <http://dx.doi.org/10.1016/j.gde.2009.09.008>.
8. Richards GS, Degnan BM. The dawn of developmental signaling in the Metazoa. *Cold Spring Harbor Symposia on Quantitative Biology*. 2009;74:81-90. doi: 10.1101/sqb.2009.74.028.
9. Seb e-Pedros A, de Mendoza A. Transcription factors and the origin of animal multicellularity. *Evolutionary Transitions to Multicellular Life: Principles and mechanisms*. Dordrecht: Springer Netherlands; 2015. p. 379-394.
10. King N. The unicellular ancestry of animal development. *Developmental Cell*. 2004;7(3):313-25. doi: 10.1016/j.devcel.2004.08.010.
11. King N, Hittinger CT, Carroll SB. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science*. 2003;301:361-3. doi: 10.1126/science.1083853.
12. Richter DJ, King N. The genomic and cellular foundations of animal origins. *Annual Review of Genetics*. 2013;47(1):509-537. doi: 10.1146/annurev-genet-111212-133456.

13. de Mendoza A, Sebé-Pedrós A, Šestak MS, et al. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proceedings of the National Academy of Sciences*. 2013;110(50):E4858-E4866. doi: 10.1073/pnas.1311818110.
14. Sebé-Pedrós A, Roger AJ, Lang FB, et al. Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proceedings of the National Academy of Sciences*. 2010;107(22):10142-10147. doi: 10.1073/pnas.1002257107.
15. Sebé-Pedrós A, de Mendoza A, Lang BF, et al. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzaki*. *Molecular Biology and Evolution*. 2011;28(3):1241-1254. doi: 10.1093/molbev/msq309.
16. Sebé-Pedrós A, Zheng Y, Ruiz-Trillo I, et al. Premetazoan origin of the Hippo signaling pathway. *Cell Reports*. 2012 1/26/;1(1):13-20. doi: <http://dx.doi.org/10.1016/j.celrep.2011.11.004>.
17. Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. The origin of Metazoa: a unicellular perspective. *Nature Reviews Genetics*. 2017;18:498. doi: 10.1038/nrg.2017.21.
18. Richter D, Fozouni P, Eisen M, et al. The ancestral animal genetic toolkit revealed by diverse choanoflagellate transcriptomes. *bioRxiv*. 2017. doi: <https://doi.org/10.1101/211789>
19. Grau-Bové X, Torruella G, Donachie S, et al. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife*. 2017;6:e26036. doi: 10.7554/eLife.26036.
20. Erwin DH, Laflamme M, Tweedt SM, et al. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*. 2011;334(6059):1091-7. doi: 10.1126/science.1206375.
21. Srivastava M, Simakov O, Chapman J, et al. The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature*;466(7307):720-6. doi: 10.1038/nature09201.
22. Richards GS, Simionato E, Perron M, et al. Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit. *Current Biology*. 2008;18(15):1156-1161. doi: <http://dx.doi.org/10.1016/j.cub.2008.06.074>.
23. Gaiti F, Fernandez-Valverde SL, Nakanishi N, et al. Dynamic and widespread lncRNA expression in a sponge and the origin of animal complexity. *Molecular biology and evolution*. 2015;32(9):2367-2382. doi: 10.1093/molbev/msv117.
24. Fernandez-Valverde SL, Degnan BM. Bilaterian-like promoters in the highly compact *Amphimedon queenslandica* genome. *Scientific Reports*. 2016;6:22496. doi: 10.1038/srep22496.
25. Fernandez-Valverde SL, Calcino AD, Degnan BM. Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene annotation in the

- sponge *Amphimedon queenslandica*. BMC Genomics. 2015;16(1):387. doi: 10.1186/s12864-015-1588-z.
26. Khalil AM, Guttman M, Huarte M, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(28):11667-72. doi: 10.1073/pnas.0904715106.
 27. Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. Science. 2005;309(5740):1559-63. doi: 10.1126/science.1112014.
 28. Ravasi T, Suzuki H, Pang KC, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome research. 2006;16(1):11-9. doi: 10.1101/gr.4200206.
 29. Bertone P, Stolc V, Royce TE, et al. Global identification of human transcribed sequences with genome tiling arrays. Science. 2004;306(5705):2242-6.
 30. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136(4):629-41. doi: 10.1016/j.cell.2009.02.006.
 31. Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes & Development. 2011;25(18):1915-27. doi: 10.1101/gad.17446611.
 32. Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. Nature. 2012;489(7414):101-8. doi: 10.1038/nature11233.
 33. Kapranov P, Cheng J, Dike S, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007;316(5830):1484-8. doi: 10.1126/science.1138341.
 34. Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome research. 2012;22(9):1775-89. doi: 10.1101/gr.132159.111.
 35. Okazaki Y, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature. 2002;420(6915):563-573.
 36. Lagarde J, Uszczyńska-Ratajczak B, Carbonell S, et al. High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. Nature Genetics. 2017;49:1731-1740. doi: 10.1038/ng.3988.
 37. Ingolia Nicholas T, Brar Gloria A, Stern-Ginossar N, et al. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. Cell Reports. 2014;8(5):1365-1379. doi: 10.1016/j.celrep.2014.07.045.

38. Guttman M, Russell P, Ingolia NT, et al. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell*. 2013;154(1):240-51. doi: 10.1016/j.cell.2013.06.009.
39. Housman G, Ulitsky I. Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2016;1859(1):31-40. doi: <http://dx.doi.org/10.1016/j.bbagr.2015.07.017>.
40. Morris KV, Mattick JS. The rise of regulatory RNA. *Nature Reviews Genetics*. 2014;15(6):423-437. doi: 10.1038/nrg3722.
41. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nature Review Genetics*. 2016;17(1):47-62. doi: 10.1038/nrg.2015.10.
42. Engreitz JM, Ollikainen N, Guttman M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nature Review Molecular Cell Biology*. 2016 17:756–770. doi: 10.1038/nrm.2016.126.
43. Rutenberg-Schoenberg M, Sexton AN, Simon MD. The properties of long noncoding RNAs that regulate chromatin. *Annual Review of Genomics and Human Genetics*. 2016;17:69-94. doi: 10.1146/annurev-genom-090314-024939.
44. Melé M, Rinn John L. “Cat’s Cradling” the 3D Genome by the Act of LncRNA Transcription. *Molecular cell*. 2016;62(5):657-664. doi: <https://doi.org/10.1016/j.molcel.2016.05.011>.
45. Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. *Cell*. 2018;172(3):393-407. doi: 10.1016/j.cell.2018.01.011.
46. Hacisuleyman E, Goff LA, Trapnell C, et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nature Structural & Molecular Biology*. 2014 0;21(2):198-206. doi: 10.1038/nsmb.2764.
47. Ulitsky I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nature Review Genetics*. 2016;17(10):601-14. doi: 10.1038/nrg.2016.85.
48. Forouzmand E, Owens NDL, Blitz IL, et al. Developmentally regulated long non-coding RNAs in *Xenopus tropicalis*. *Developmental Biology*. 2016;pii: S0012-1606(16):30120-8. doi: <http://dx.doi.org/10.1016/j.ydbio.2016.06.016>.
49. Tan MH, Au KF, Yablonovitch AL, et al. RNA sequencing reveals a diverse and dynamic repertoire of the *Xenopus tropicalis* transcriptome over development. *Genome research*. 2013;23(1):201-216. doi: 10.1101/gr.141424.112.
50. Necșulea A, Soumillon M, Warnefors M, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature*. 2014;505(7485):635-40. doi: 10.1038/nature12943.

51. Pauli A, Valen E, Lin MF, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Research*. 2012;22(3):577-91. doi: 10.1101/gr.133009.111.
52. Ulitsky I, Shkumatava A, Jan CH, et al. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*;147(7):1537-50. doi: 10.1016/j.cell.2011.11.055.
53. Sauvageau M, Goff LA, Lodato S, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *eLife*. 2013;2:e01749. doi: 10.7554/eLife.01749.
54. Nam JW, Bartel DP. Long noncoding RNAs in *C. elegans*. *Genome Research*. 2012 Dec;22(12):2529-40. doi: 10.1101/gr.140475.112.
55. Brown JB, Boley N, Eisman R, et al. Diversity and dynamics of the *Drosophila* transcriptome. *Nature*. 2014;512(7515):393-9. doi: 10.1038/nature12962.
56. Chen B, Zhang Y, Zhang X, et al. Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila* metamorphosis. *Scientific Reports*. 2016;6:23330. doi: 10.1038/srep23330.
57. Jenkins AM, Waterhouse RM, Muskavitch MAT. Long non-coding RNA discovery across the genus anopheles reveals conserved secondary structures within and beyond the Gambiae complex. *BMC Genomics*. 2015;16(1):1-14. doi: 10.1186/s12864-015-1507-3.
58. Quinn JJ, Zhang QC, Georgiev P, et al. Rapid evolutionary turnover underlies conserved lincRNA–genome interactions. *Genes & Development*. 2016;30(2):191-207. doi: 10.1101/gad.272187.115.
59. Jayakodi M, Jung JW, Park D, et al. Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*. *BMC Genomics*. 2015;16(1):1-12. doi: 10.1186/s12864-015-1868-7.
60. Young RS, Marques AC, Tibbit C, et al. Identification and properties of 1,119 candidate lincRNA loci in the *Drosophila melanogaster* genome. *Genome Biology and Evolution*. 2012;4(4):427-42. doi: 10.1093/gbe/evs020.
61. Wu Y, Cheng T, Liu C, et al. Systematic identification and characterization of long non-coding RNAs in the silkworm, *Bombyx mori*. *PLoS one*. 2016;11(1):e0147147. doi: 10.1371/journal.pone.0147147.
62. Mu C, Wang R, Li T, et al. Long non-coding RNAs (lincRNAs) of sea cucumber: Large-scale prediction, expression profiling, non-coding network construction, and lincRNA-microRNA-gene interaction analysis of lincRNAs in *Apostichopus japonicus* and *Holothuria glaberrima* during LPS challenge and radial organ complex regeneration. *Marine Biotechnology*. 2016 2016;18(4):485-499. doi: 10.1007/s10126-016-9711-y.

63. Bråte J, Adamski M, Neumann RS, et al. Regulatory RNA at the root of animals: dynamic expression of developmental lincRNAs in the calcisponge *Sycon ciliatum* [10.1098/rspb.2015.1746]. Proceedings of the Royal Society of London B: Biological Sciences. 2015;282(1821):20151746.
64. Perry RB-T, Ulitsky I. The functions of long noncoding RNAs in development and stem cells. Development. 2016;143(21):3882-3894.
65. Huang C, Morlighem J-ÉRL, Cai J, et al. Identification of long non-coding RNAs in two anthozoan species and their possible implications for coral bleaching. Scientific Reports. 2017;7(1):5333. doi: 10.1038/s41598-017-02561-y.
66. Heard E, Mongelard F, Arnaud D, et al. Human XIST yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(12):6841-6846.
67. Migeon BR, Kazi E, Haisley-Royster C, et al. Human X inactivation center induces random X chromosome inactivation in male transgenic mice. Genomics. 1999;59(2):113-121. doi: <http://dx.doi.org/10.1006/geno.1999.5861>.
68. Grant J, Mahadevaiah SK, Khil P, et al. Rxs is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature. 2012;487(7406):254-258.
69. Liu SJ, Nowakowski TJ, Pollen AA, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. Genome Biology. 2016;17(1):1-17. doi: 10.1186/s13059-016-0932-1.
70. Cabili MN, Dunagin MC, McClanahan PD, et al. Localization and abundance analysis of human lincRNAs at single-cell and single-molecule resolution. Genome Biology. 2015;16(1):1-16. doi: 10.1186/s13059-015-0586-4.
71. Ponjavic J, Oliver PL, Lunter G, et al. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. PLoS Genetics. 2009;5(8):e1000617. doi: 10.1371/journal.pgen.1000617.
72. Mercer TR, Dinger ME, Sunken SM, et al. Specific expression of long noncoding RNAs in the mouse brain. Proceedings of the National Academy of Sciences. 2008;105(2):716-721. doi: 10.1073/pnas.0706729105.
73. Zappulo A, van den Bruck D, Ciolli Mattioli C, et al. RNA localization is a key determinant of neurite-enriched proteome. Nature Communications. 2017;8(1):583. doi: 10.1038/s41467-017-00690-6.
74. Goff LA, Groff AF, Sauvageau M, et al. Spatiotemporal expression and transcriptional perturbations by long noncoding RNAs in the mouse brain. Proceedings of the National Academy of Sciences. 2015;112(22):6855-6862. doi: 10.1073/pnas.1411263112.

75. Pauli A, Rinn JL, Schier AF. Non-coding RNAs as regulators of embryogenesis. *Nature Review Genetics*. 2011 Feb;12(2):136-49. doi: 10.1038/nrg2904.
76. Gaiti F, Hatleberg WL, Tanurdzic M, Degan BM. Sponge long non-coding RNAs are expressed in specific cell types and conserved networks. *Non-coding RNA*. 2018;4(1), 6. doi:10.3390/ncrna4010006.
77. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nature Review Genetics*. 2014;15(1):7-21. doi: 10.1038/nrg3606.
78. Pang KC, Frith MC, Mattick JS. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends in genetics*. 2006 Jan;22(1):1-5. doi: 10.1016/j.tig.2005.10.003.
79. Hezroni H, Koppstein D, Schwartz Matthew G, et al. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Reports*. 2015;11(7):1110-1122. doi: <http://dx.doi.org/10.1016/j.celrep.2015.04.023>.
80. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell*. 2013;154(1):26-46. doi: 10.1016/j.cell.2013.06.020.
81. Maenner S, Blaud M, Fouillen L, et al. 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS biology*. 2010;8(1):e1000276. doi: 10.1371/journal.pbio.1000276.
82. Lubelsky Y, Ulitsky I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature*. 2018 01/24/online. doi: 10.1038/nature25757
83. Amaral PP, Leonardi T, Han N, et al. Genomic positional conservation identifies topological anchor point (tap)RNAs linked to developmental loci. *bioRxiv*. 2016. doi: <https://doi.org/10.1101/051052>.
84. Hawkes Emily J, Hennelly Scott P, Novikova Irina V, et al. *COOLAIR* antisense RNAs form evolutionarily conserved elaborate secondary structures. *Cell Reports*. 2016;16(12):3087-3096. doi: 10.1016/j.celrep.2016.08.045.
85. Novikova IV, Hennelly SP, Sanbonmatsu KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Research*. 2012;40(11):5034-5051. doi: 10.1093/nar/gks071.
86. Sanbonmatsu KY. Towards structural classification of long non-coding RNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2016;1859(1):41-45. doi: <http://dx.doi.org/10.1016/j.bbagr.2015.09.011>.
87. Tichon A, Gil N, Lubelsky Y, et al. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nature Communications*. 2016;7:12209. doi: 10.1038/ncomms12209.

88. Chodroff RA, Goodstadt L, Sirey TM, et al. Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes. *Genome Biology*. 2010;11(7):R72. doi: 10.1186/gb-2010-11-7-r72.
89. Washietl S, Kellis M, Garber M. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. *Genome Research*. 2014;24(4):616-28. doi: 10.1101/gr.165035.113.
90. Peter Isabelle S, Davidson Eric H. Evolution of gene regulatory networks controlling body plan development. *Cell*. 2011;144(6):970-985. doi: <http://dx.doi.org/10.1016/j.cell.2011.02.017>.
91. Erwin DH, Davidson EH. The evolution of hierarchical gene regulatory networks. *Nature Review Genetics*. 2009;10(2):141-148.
92. Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans. *Science*. 2006;311. doi: 10.1126/science.1113832.
93. Kapusta A, Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. *Trends in Genetics*. 2014;30(10):439-452. doi: 10.1016/j.tig.2014.08.004.
94. Seb e-Pedr os A, Ballar e C, Parra-Acero H, et al. The dynamic regulatory genome of *Capsaspora* and the origin of animal multicellularity. *Cell*. 2016;165(5):1224-37. doi: 10.1016/j.cell.2016.03.034.
95. de Mendoza A, Suga H, Permanyer J, et al. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *eLife*. 2015 (4):e08904.
96. Gaiti F, Jindrich K, Fernandez-Valverde SL, et al. Landscape of histone modifications in a sponge reveals the origin of animal cis-regulatory complexity. *eLife*. 2017 2017/04/11;6:e22194. doi: 10.7554/eLife.22194.
97. Hinman V, Cary G. The evolution of gene regulation. *eLife*. 2017 2017/05/12;6:e27291. doi: 10.7554/eLife.27291.
98. Levine M, Cattoglio C, Tjian R. Looping back to leap forward: Transcription enters a new era. *Cell*. 2014;157(1):13-25. doi: 10.1016/j.cell.2014.02.009.
99. Long HK, Prescott SL, Wysocka J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell*. 2017;167(5):1170-1187. doi: 10.1016/j.cell.2016.09.018.
100. Dickel DE, Ypsilanti AR, Pla R, et al. Ultraconserved Enhancers Are Required for Normal Development. *Cell*. 2018. doi: 10.1016/j.cell.2017.12.017.
101. Schwaiger M, Schonauer A, Rendeiro AF, et al. Evolutionary conservation of the eumetazoan gene regulatory landscape. *Genome Research*. 2014 Apr;24(4):639-50. doi: 10.1101/gr.162529.113.

102. Rada-Iglesias A, Bajpai R, Swigut T, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470(7333):279-283.
103. Creighton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences*. 2010;107(50):21931-21936. doi: 10.1073/pnas.1016071107.
104. Heintzman ND, Hon GC, Hawkins RD, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459. doi: 10.1038/nature07829.
105. Daugherty AC, Yeo RW, Buenrostro JD, et al. Chromatin accessibility dynamics reveal novel functional enhancers in *C. elegans*. *Genome Research*. 2017; 27: 2096-2107.
106. Natoli G, Andrau J-C. Noncoding transcription at enhancers: general principles and functional models. *Annual Review of Genetics*. 2012;46(1):1-19. doi: 10.1146/annurev-genet-110711-155459.
107. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nature Reviews Genetics*. 2016;17(4):207-23. doi: 10.1038/nrg.2016.4.
108. Kim T-K, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465(7295):182-187.
109. Marques A, Hughes J, Graham B, et al. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. *Genome Biology*. 2013;14(11):R131. doi:10.1186/gb-2013-14-11-r131.
110. Illott NE, Heward JA, Roux B, et al. Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. *Nature Communications*. 2014;5:3979. doi: 10.1038/ncomms4979.
111. Werner MS, Sullivan MA, Shah RN, et al. Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription. *Nature Structural & Molecular Biology*. 2017;24:596. doi: 10.1038/nsmb.3424.
112. Gayen S, Kalantry S. Chromatin-enriched lncRNAs: a novel class of enhancer RNAs. *Nature Structural & Molecular Biology*. 2017;24:556. doi: 10.1038/nsmb.3430.
113. Grosberg RK, Strathmann RR. The Evolution of Multicellularity: A Minor Major Transition? *Annual Review of Ecology, Evolution, and Systematics*. 2007;38(1):621-654. doi: 10.1146/annurev.ecolsys.36.102403.114735.
114. Umen JG. Green Algae and the Origins of Multicellularity in the Plant Kingdom. *Cold Spring Harbor Perspectives in Biology*. 2014;6(11).

115. Leliaert F, Smith DR, Moreau H, et al. Phylogeny and Molecular Evolution of the Green Algae. *Critical Reviews in Plant Sciences*. 2012;31(1):1-46. doi: 10.1080/07352689.2011.615705.
116. Featherston J, Arakaki Y, Hanschen ER, et al. The 4-celled *Tetrabaena socialis* nuclear genome reveals the essential components for genetic control of cell number at the origin of multicellularity in the volvocine lineage. *Molecular Biology and Evolution*. 2017:msx332-msx332. doi: 10.1093/molbev/msx332.
117. Prochnik SE, Umen J, Nedelcu AM, et al. Genomic Analysis of Organismal Complexity in the Multicellular Green Alga *Volvox carteri*. *Science*. 2010;329(5988):223.
118. Vergara Z, Gutierrez C. Emerging roles of chromatin in the maintenance of genome organization and function in plants. *Genome Biology*. 2017;18(1):96. doi: 10.1186/s13059-017-1236-9.
119. Shaver S, Casas-Mollano JA, Cerny RL, et al. Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*. *Epigenetics*. 2010;5(4):301-312. doi: 10.4161/epi.5.4.11608.
120. Mikulski P, Komarynets O, Fachinelli F, et al. Characterization of the Polycomb-Group Mark H3K27me3 in Unicellular Algae. *Frontiers in Plant Science*. 2017;8(607). doi: 10.3389/fpls.2017.00607.
121. Leys SP, Degnan BM. Embryogenesis and metamorphosis in a haplosclerid demosponge: gastrulation and transdifferentiation of larval ciliated cells to choanocytes. *Invertebrate Biology*. 2002;121(3):171-189.
122. Gaiti F, Calcino AD, Tanurđić M, et al. Origin and evolution of the metazoan non-coding regulatory genome. *Developmental Biology*. 2017;427(2):193-202. doi: <https://doi.org/10.1016/j.ydbio.2016.11.013>.

Figure Legends

Figure 1. Early evolution of animal long non-coding RNAs: Insights from the sponge

***Amphimedon queenslandica*.** (A) Despite a growing number of lncRNAs having been identified in bilaterian animals, the systematic investigation of lncRNAs in non-bilaterian animals has been lagging behind and, thus, we lack an understanding of their origin and early evolution. Yellow background highlights the animal kingdom. (B) and (C) Identification of *Amphimedon* lncRNAs. (B) Schematic representation of the *Amphimedon queenslandica* life cycle. Larvae emerge from maternal brood chambers and then swim in the water column as precompetent larvae before they develop competence to settle and initiate metamorphosis. Upon settling, the larva adopts a flattened morphology as it metamorphoses into a juvenile, which displays the hallmarks of the adult body plan. This juvenile will grow and mature into a benthic adult [121]. Adapted from [23]. (C) Developmental expression profiles of *Amphimedon* lncRNAs. Expression profiles of the top 50 differentially expressed lncRNAs during the transition from pelagic swimming competent larva to benthic juvenile. Each row represents data for one lncRNA. Pelagic stages include precompetent (P) and competent (C) larva; benthic stages include juvenile (J) and adult (A). Red indicates high expression level, light blue low expression. Adapted from [23].

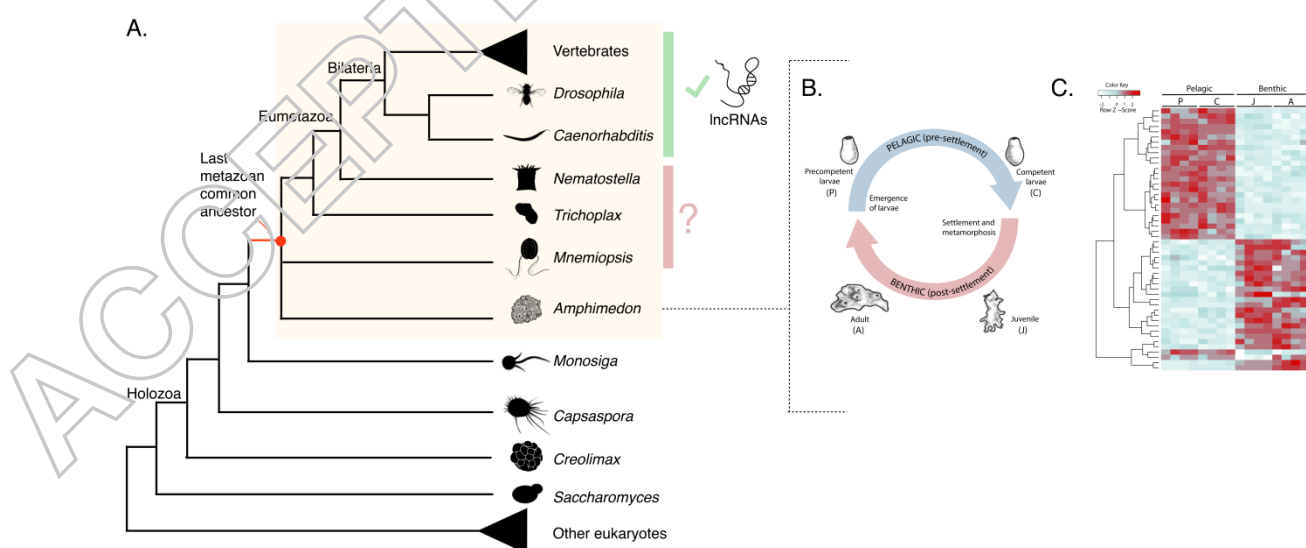
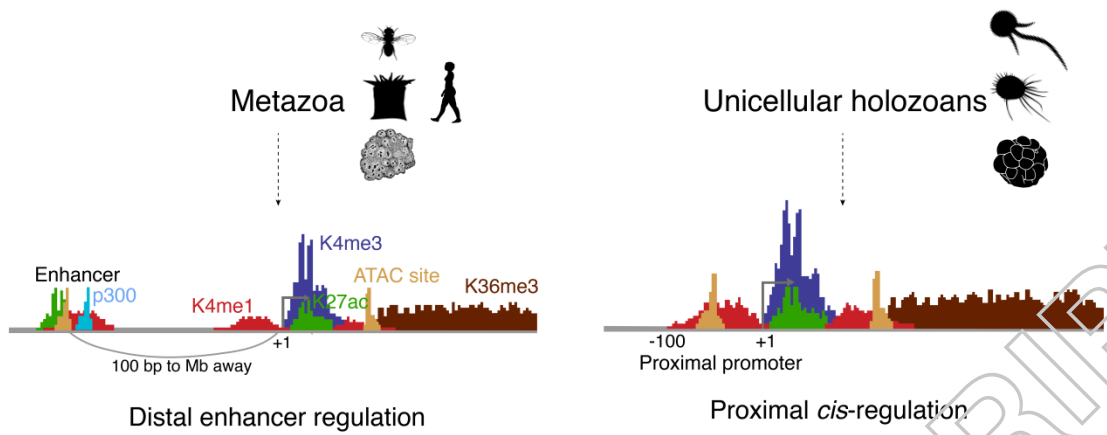


Figure 2. Long non-coding RNAs are defined by specific chromatin signatures. (A) Recent analyses [94, 96, 101] of non-coding regulatory DNA and histone marks have revealed that some *cis*-regulatory mechanisms, such as those associated with proximal promoters, are present in non-metazoan holozoans (right panel) while others appear to be metazoan innovations, most notably distal enhancer regulation (left panel). Shown is a schematic representation of the presence or absence of the typical chromatin signatures associated with animal distal enhancer elements [the transcriptional cofactor p300, histone 3 lysine 4 monomethylation (H3K4me1), histone 3 lysine 27 acetylation (H3K27ac), and ATAC site]. Adapted from [122]. (B) LincRNAs can be separated in two distinct populations of polyadenylated transcripts based on the chromatin status at their transcription-start sites. Shown is the enrichment of H3K4me1 (left) and H3K4me3 (right) (ChIP versus input) at enhancer-associated and promoter-like lincRNAs, respectively.

A.



B.

