Published in final edited form as:

Bioessays. 2020 February ; 42(2): e1900173. doi:10.1002/bies.201900173.

The Developmental Gene Hypothesis for Punctuated Equilibrium: Combined Roles of Developmental Regulatory Genes and Transposable Elements

Emily L. Casanova¹, Miriam K. Konkel²

¹Department of Biomedical Sciences, University of South Carolina School of Medicine at Greenville, Greenville, South Carolina, USA

²Department of Genetics and Biochemistry, Clemson Center for Human Genetics, Biomedical Data Science and Informatics Program, Clemson University, Clemson, South Carolina, USA

Abstract

Theories of the genetics underlying Punctuated Equilibrium (PE) have been vague to date. Here we propose the Developmental Gene Hypothesis, which states that: 1) developmental regulatory (DevReg) genes are responsible for the orchestration of metazoan morphogenesis and their extreme conservation and mutation intolerance generates the equilibrium or stasis present throughout much of the fossil record; and 2) the accumulation of regulatory elements and recombination within these same genes—often derived from transposable elements—drives punctuated bursts of morphological divergence and speciation across metazoa. This two-part hypothesis helps to explain the features that characterize PE, providing a theoretical genetic basis for the once-controversial theory.

Layman's Summary

The Theory of Punctuated Equilibrium states that physical changes across species tend not to occur gradually but instead present themselves in brief punctuated events bookended by long periods of evolutionary stability or stasis. Though this theory has previously lacked a genetic explanation for the trends seen in the fossil record, the Developmental Gene Hypothesis for Punctuated Equilibrium helps account for these patterns by proposing that developmental regulatory genes, which are largely responsible for animal form, help maintain stasis due to their extreme conservation across species and overall mutation sensitivity. Similarly, when change does occur in these genes, such as through transposable element-related mechanisms, they may be co-opted (e.g., gene regulation), resulting in physical changes between species.

Keywords

developmental regulatory genes; punctuated equilibrium; macroevolution; conserved noncoding elements; transposable elements; regulome; Burgess Shale

*Address Correspondence to: Emily L. Casanova, Patewood Medical Campus, 200A Patewood Dr., Greenville, SC 29615, Tel: +1 (864) 454-1082, casanove@greenvillemed.sc.edu; emily.casanova@prisma.org; scienceoveracuppa@gmail.com.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest pertaining to this work, financial or otherwise.

1 A GENETIC HYPOTHESIS TO EXPLAIN THE THEORY OF PUNCTUATED EQUILIBRIUM

Despite his support for phyletic gradualism, Darwin himself acknowledged the abrupt appearance of new species within the fossil record. ^[1] And although this observation was partly due to the infancy of 19th century paleontology, more than a century and a half later we recognize that not all aspects of cladogenesis can be accounted for by a lack of detailed stratigraphical study and instead adhere to the theory of punctuated equilibrium.

Detractors of punctuated equilibrium have previously chided the theory as a mechanism with “no scientific use” as they claimed it could not be tested at the genetic level. ^[2] However, with the advent of large-scale genomic sequencing and extensive molecular and computational study of numerous genomes, our wealth of available data has grown substantially since the early battles over Gould and Eldredge’s theory. ^[3] Since that time we have been able to study not only transposable elements (TE) (i.e., “selfish DNA”) and their roles in molecular evolution, but also the subset of genes responsible for the regulation of morphogenesis reflected in the fossil record.

Here, we propose that there is a native genetic complement to TE insertions leading to features of punctuated equilibria in both the vertebrate and invertebrate fossil records. Specifically, this complement lies within the developmental regulatory (DevReg) genes responsible for morphogenesis and their unique mutational patterns, as well as the elements that regulate their expression. The genes’ relative mutation intolerance suggests a means by which morphology is actively conserved even in the face of exaggerated TE activity. ^[4] Yet they also exhibit a clear history of TE insertion that is strongly correlated to the presence of conserved noncoding elements (CNE) and changes to gene regulation and phenotype by acting as promoters, enhancers, repressors, terminators, insulators, and post-transcriptional effectors. ^[4–7] In addition, TE-derived RNA may act as direct regulators of these important developmental genes, strongly reminiscent of Barbara McClintock’s “controlling elements”. ^[7,8] Alongside the TE-Thrust Hypothesis ^[9] and the TE-epigenomic arms race proposed by Zeh et al. ^[10], as well as the large body of work spearheaded by Eric Davidson ^[11,12] concerning the roles of the regulome in speciation, we believe the *Developmental Gene Hypothesis* may help explain the long periods of active morphological stasis within the fossil record bookended by rare TE-associated mutational events that have long lasting effects on phenotype and have been tightly conserved over evolutionary time ^[4,13] (Fig. 1).

2 DEVELOPMENTAL GENES AND THEIR REGULOME ARE INTEGRAL TO THE PHYLOTYPIC PERIOD OF DEVELOPMENT

Since the morphological features of an animal are the product of its developmental process, and since the developmental process in each animal is encoded in its species-specific regulatory genome, then change in animal form during evolution is the consequence of change in genomic regulatory programs for development. ^[11]

Although evolution of the eukaryotic exome has been relatively conserved across time, evolution of the regulome seems to account for considerable interspecies variation.^[11,13] The regulome is both a seat of rapid evolutionary development, as in the case of species-specific TE regulatory exaptation, and extreme conservation in the form of conserved (CNE) and ultraconserved noncoding elements (UCNE).^[14–16] UCNE networks are likewise strongly conserved across species in the form of gene regulatory blocks surrounding important developmental genes.^[17]

2.1 The Evolutionary Rate of Conserved Noncoding Elements Has Slowed

The rate of substitutions within protein-coding regions has been relatively constant across vertebrates, yet CNE have experienced significantly lower mutation rates since the dawn of amniotes.^[15] Interestingly, prior to this mutational retardation, tetrapod ancestors experienced a massive expansion of sequences leading to CNE, presumably related to the evolution of terrestrial life.^[15] Many of the UCNE are located in or near transcriptional and developmental genes, as are many of the younger less conserved noncoding elements unique to later phylogenetic branches, suggesting a strong directional selection for the retention and exaptation of potential noncoding elements in these regions^[15,18] While many ancient coding and noncoding ultraconserved elements present in bony fish and their descendants occurred in roughly equal proportions targeting splice sites, intronic, and intergenic regions (and to a lesser degree coding sequences and untranslated regions), more recent elements in amniotes, theria, or eutheria are derived primarily from introns and gene-poor sites.^[15]

Invertebrate genomes likewise exhibit conservation of noncoding elements surrounding key developmental genes, a feature they share with vertebrates. However, these CNE appear to be largely unique to each taxon and likely play an integral role in defining the general body plans particular to the major branches.^[19–21]

In contrast to the invertebrates we see today, however, it is well recognized that during the Cambrian period there was a much wider array of body plans, i.e., the so-called “Cambrian explosion”, a concept which Gould made famous in his 1989 book, *Wonderful Life: The Burgess Shale and the Nature of History*.^[22] It is now believed that rather than a single explosion, the Ediacaran-Cambrian interval experienced a series of significant radiations more akin to successive fireworks.^[23] Nevertheless, the variety in bauplan present in the Burgess Shale Lagerstätte, along with features of “interchangeable parts” of some of these strange creatures as suggested by Gould, indicates that CNE may have been experiencing a significant evolutionary flux during these early periods in metazoan evolution prior to entrenchment. Likewise, fluctuation in regulatory elements provides a theoretical means for the extreme variety present during those early periods, as well as the few but highly conserved body plans and extreme CNE mutation intolerance we see in extant organisms today.

2.2 Developmental Regulatory Genes Share Unique Features

We have recently found that DevReg genes in humans, for example, maintain a particularly high relative density of intronic CNE compared to most other genes.^[4] These genes generally produce larger proteins but the genes themselves also tend to be quite long, which

can be primarily attributed to intron size. Their sensitivity to variation is also significant and is evidenced by reduced numbers of common/rare single nucleotide variants (SNV) and copy number variants (CNV), suggesting they may be strongly conserved. [4,24] Interestingly, these genes also maintain more extensive protein-protein interaction (PPI) networks, suggesting many of them may be foundational and function as molecular hubs. [4,25] Data indicates that many of these genes are also targets of various forms of regulatory RNA derived from TE, adding yet another level of complexity to this broad developmental network. [8,26,27]

2.3 Developmental Regulatory Genes Play a Critical Role during the Phylotypic Period of Development

Many DevReg genes are expressed during the phylotypic period of development, suggesting a reason for their conservation. [28] This conservation is modeled in what Arthur [29] calls “the Fisher principle,” referring to the inverse relationship between the extent to which a mutation affects phenotype and the likelihood it will be selectively advantageous:

It seems reasonable to suppose that genes controlling early developmental decisions, such as which end of the embryo is anterior, will be subject, on average, to mutations with more major phenotypic effects than genes controlling later developmental processes, such as the production of mammalian hair ... If so, then under the Fisher principle early genes will evolve more slowly than their later counterparts ... and, under certain assumptions ... this will mean that the earlier stages themselves will be more evolutionarily conserved. [29]

Although Arthur [29] was correct about weak genic conservation in late phenotypic stages, genes expressed in very early embryogenesis are in fact poorly conserved and small in size, an observation that explains significant variability amongst the zygotes of different species. [30] Instead, genes expressed during the *middle* stages of embryogenesis (i.e., the phylotypic period) are in fact more heavily conserved and fit the profiles of the long complex genes described here. This pattern of conservation is referred to as the “hourglass model,” in which genes involved in organogenesis share greatest overlap across diverse species. [28] Various DevReg genes show the highest transcriptional expression during the phylotypic period in both mouse and human. They are typically expressed across multiple tissue types, potentially leading to shared changes across many organ systems when germline mutations do occur. [31] This can be seen, for example, in many human-specific mutation events in DevReg genes that often lead not only to intellectual disability but multiple congenital anomalies. [32] In addition, many genes associated with rare human genetic diseases share similar structural, functional, and conservation profiles as those reported by Casanova et al. [4], reinforcing the concept of their importance in development and human health. [24] Although these are all extreme pathological examples, it is nevertheless easy to envision how subtler less detrimental changes to developmental gene function could lead to adaptive changes across organ systems over larger timescales, allowing evolution to act on specific aspects of morphology, not in isolation, but in concert. [33]

Although there are considerable differences in physical form during the phylotypic stage across higher order taxa (e.g., insects, crustaceans, and vertebrates), within these taxa there

is significant conservation during early organogenesis, which, as mentioned earlier, appears to be reflected in conservation of unique regulatory elements in and around developmental genes. [19,20] Many CNE are *cis*-regulatory elements (e.g., enhancers) and are believed to reflect parallel evolution of alternative elements influencing expression of these genes and ultimately variations in body plan that differentiate the major taxa. Therefore, although most CNE are not conserved across vertebrates and invertebrates, networks that regulate these vital developmental genes are common to the metazoa. [20,37]

In general, the phylotypic period is relatively conserved, yet it is during this stage in which minor progressive changes to the regulation of DevReg genes orchestrating it may lead to the inception of heterochronic tissue development and changes in morphology that ultimately differentiate species across the fossil record. [11,38] In addition, while there is indeed notable developmental variability during this period between human and zebrafish embryos, for instance, similarities in overall morphology are also visibly apparent suggesting a conserved bauplan that is likely reflected in coding and noncoding ultraconserved elements within the genome. [15,20,39]

2.4 Is Stasis an Active Phenomenon?

Gould and Eldredge [40] once wrote:

As the most important change in research practice provoked by punctuated equilibrium, stasis has now exited from its closet of non-definition to become a subject of quantitative investigation in all major fossil groups ... Moreover, because species often maintain stability through such intense climatic changes as glacial cycling, stasis must be viewed as an active phenomenon, not a passive response to unaltered environments.

In this instance, the variation intolerance of DevReg genes provides an active safeguard against mutational events. Surprisingly, CNE within long genes such as DevReg genes are not mutation cold spots despite their functional sensitivity. Instead, such mutations are typically incompatible with life and often lead to miscarriage or serious functional impairment preventing further inheritance. [41,42]

Nevertheless, developmental genes may be key sites for species-specific evolution. As an example from human and chimpanzee: genes with longer introns (e.g., DevReg genes) are sequentially more divergent than those with shorter introns (K_i = substitutions per intronic site), suggesting they are sites for accelerated evolution between our two species despite their overall mutation sensitivity. [43,44]

While genetic isolation and ultimately hybrid dysgenesis ensure speciation, the evolution of DevReg genes plays a fundamental role in the type of morphological divergence that originally inspired the Linnaean classification system. [45] Of course, not all species can be divided according to morphological divergence (i.e., cryptic biodiversity) and, similarly, DevReg genes are not completely impervious to drift. [46,47] However, following the recognition of this disparity, it has subsequently been demonstrated that morphological evolution correlates better with regulatory gene divergence rather than with overall rates of molecular evolution. [48–50] Thus, the variation intolerance we see in DevReg genes appears

to be responsible for the significant periods of morphological stasis present in the fossil record. Meanwhile, other gene groups evolve at a more rapid, and perhaps more constant, rate.

3 TRANSPOSONS AND THE EVOLVING REGULOME DRIVE PUNCTUATED CHANGES IN ANIMAL FORM

Since the publication of the theory of punctuated equilibrium, detractors have sometimes criticized its dependence on an imperfect fossil record and the lack of an underlying genetic theory. ^[1] Oliver and Greene ^[9] and Zeh et al. ^[10] made significant strides by proposing genetic and epigenetic mechanisms that help explain the long periods of stasis bookended by periods of comparatively rapid change. Likewise, the body of work by Eric Davidson clearly outlines numerous regulatory mechanisms leading to morphological divergence across species. ^[11,12] These works, however, do not recognize the active role DevReg genes play in maintaining morphological stasis, their propensity to collect TE-derived noncoding elements over time, or their tendency to be regulated by TE-derived RNA. ^[8] All of these presumably underlie changes in gene function and morphology relevant to the patterns within the fossil record first recognized by Gould and Eldredge. ^[3] The Developmental Gene Hypothesis helps to fill in these important gaps.

3.1 Conserved Noncoding Elements Are Often Derived from Transposable Elements

CNE are key players in the evolution of DevReg genes and organismal morphology. What is more, DevReg genes in humans not only contain an increased density of CNE, they also have a higher density of TEs. ^[4] In fact, intronic CNE and TE content across the entire human genome correlates strongly with one another ($r = 0.856$, $p < 0.0001$) (Fig. 3). ^[4] This relationship is perhaps unsurprising since some CNE are derived from TE insertion and exaptation events. ^[51] For instance, Lowe et al. ^[52] reported that thousands of TE fragments in the human genome have been coopted and have been under strong purifying selection over the last 100 million years. Indeed, at least 20% of regulatory sequences in humans are derived from mobile elements and because many ancient transposition events are no longer recognizable, an even wider breadth of TE-derived CNE may exist. ^[51,53] Lowe and Haussler ^[51] went on to report that, across twenty-nine mammalian genomes, TE fragments are consistently exapted as gene regulatory sequences, especially transcription factor binding sites. Meanwhile, Villar et al. ^[54] reported that, across 20 mammalian species, species-specific enhancers are often derived from ancient transposon-rich DNA, suggesting that in addition to TE insertion, genetic drift may play an important role in the development of new enhancer elements. A similarly slow phenomenon of “optimization” is noted in TE-derived exonization. ^[55]

3.2 Transposon-derived Regulatory Elements Are Key Drivers of Speciation

Within mammals, there are strong links between TE activity and speciation. ^[56,57] As mentioned, TEs provide a prime source of potential regulatory material to the host genome and the primate lineage is an excellent example of such a relationship. As Zeh et al. ^[10] review, the primate-specific retrotransposons known as *Alu* elements have been exceptionally successful in this lineage. ^[58,59] For instance, the *AluJ* family arose prior to

the divergence of haplorrhines and strepsirrhines approximately 65 million years ago. More recently, about 40–35 million years ago, the *AluS* family arose, followed by *AluY*. *AluS* was actively propagating during the radiation of tarsiers from platyrrhines and before the radiation of New World monkeys (platyrrhines) from catarrhines. [60–62] In fact, the reconstruction of the *Alu* genetic “fossil record”¹ strongly supports that their rate of propagation was highest during this time period. [64,65] *AluY* eventually became the leading family in catarrhines prior to the radiation of Old World monkeys and apes. [66] Finally, *AluYa5* and *AluYb8* – the two main *AluY* subfamilies currently active in humans – arose a few million years ago during the emergence of our own lineage, the hominins. [58,63,67–69] Given the timing of events and the fact that *Alu* elements are associated with regulatory functions, such as transcription factor binding sites, adenosine-to-inosine (A-to-I) editing, and circular RNA, it is likely this family of retrotransposons has played an integral role in the evolution of primate morphology, including our own. [70–74]

TE insertions have also played an important role in the evolution of pregnancy, particularly in the decidualization of the connective tissue within the uterus. For instance, certain endogenous retroviruses (ERV) are highly expressed in mammalian uterus in a tissue-specific manner and, among other things, help to drive cell fusion within the trophoblast layers via action of the ERV-derived envelope glycoprotein, syncytin. [75,76] Interestingly, the *Mabuya* lizard, which evolved approximately 25 million years ago (mya), is viviparous and has an unusually mammalian-like placenta that also expresses an ERV-derived envelope glycoprotein functionally identical to mammalian syncytin. [77] Similarly, fossilized evidence of the Jurassic marine reptiles, ichthyosaurs, show they were also viviparous, indicating that live birth has evolved multiple times in the reptilian lineage and may commonly be linked with the exaptation of retroviruses (see Fig. 4).

Thousands of other transposon-derived *cis*-regulatory elements have been identified that also regulate placental function, many of which have been exapted as hormone response elements (HRE). [78] For instance, *Alu* elements have been shown to house high-affinity binding sites for the estrogen, thyroid, and retinoic acid receptors. [79] SVA elements likewise appear to house HRE half-sites and also bind the glucocorticoid receptor. [80] Therefore, evidence suggests that both TE-derived HRE and ERV-derived envelope glycoproteins have played important roles in the evolution of mammalian pregnancy.

3.3 Transposon-mediated Recombination Can Drive Speciation

Aside from insertion and exaptation, TEs are also frequent sites for recombination due to their repetitive natures, resulting in high copy numbers. [81–85] This provides additional means for escape from evolutionary stasis. Often, recombination results in deletion of genetic material but it can also lead to duplications or more complex forms of recombination such as translocation. [86–89] For instance, compared to human, the chimpanzee genome shows evidence of over 600 *Alu*-mediated deletion events alone (panTro1, hg17). [90] Meanwhile, the human genome contains close to 500 unique *Alu*-mediated deletions. [91] In addition, scientists have found that recombination rates are increased around conserved

¹Terminology used to refer to the genetic record of *Alu* insertions within the genome. [63]

noncoding regions in human as compared to chimpanzee, indicating accelerated evolution in those areas. ^[92,93] There are also numerous examples of genes associated with human diseases whose insertions and deletions are the result of *Alu*-mediated recombination. ^[82,95–100] Therefore, both the loss and addition of TE-associated genetic materials have been shown to play active roles in primate speciation and pathology. ^[66,101]

Primates, particularly humans and the great apes, show evidence of higher rates of segmental duplications. ^[102] Interestingly, evidence of segmental duplications has also been identified in other clades that have undergone rapid speciation events. One such example is the infraorder, cetacea, which first began diverging from artiodactyls over 50 mya. Neoceti, which is made up of the extant parvorders, odontocetes and mystocetes, diverged from archaeocetes at least 36 mya. ^[103] From the time of the partly aquatic whale ancestor, *Indohyus*, it took this lineage approximately 10–12 million years to become fully aquatic—an extraordinarily rapid series of speciation events. ^[104,105]

Although scientists are unable to study the genomes of extinct lineages, all modern toothed and baleen whales sequenced to date show evidence of significant numbers of segmental duplications enriched in regions housing morphological genes. ^[106] The presence of duplications in these regions suggests these variants have influenced cetacean morphology at least from the time of the basilosaurids, if not earlier, and may have played significant roles in this lineage's return to the sea.

The genomes of many invertebrates, in contrast to those of mammalian vertebrates, show comparatively minimal evidence of TE insertion into gene-rich areas, with some exceptions. ^[107,108] Given the hypothesis presented here and the relative morphological conservation apparent in many invertebrate lines, a slower evolutionary pace may influence DevReg genes within these taxa. ^[11] An intriguing and notable exception, however, appears to be the octopus, which has experienced significant TE expansion leading to large-scale genomic rearrangements and an enlargement of specific gene families influencing morphogenesis and central nervous system development in particular. ^[109] Interestingly, octopuses are known for some object-oriented play and tool use and are a model organism for the study of advanced cognitive abilities in invertebrates. ^[110]

4 Conclusions and Outlook

While punctuated equilibrium as a theory has stood the test of time, until now there has been no clear genetic explanation for the trends present within the fossil record. Work by Oliver and Greene ^[9], Zeh et al. ^[10], and Davidson ^[11] provide admirable starting points and, we think, vital pieces of the larger story.

The Developmental Gene Hypothesis, however, proposes a clear and testable mechanism for stasis via the strong purifying selection acting upon DevReg genes. Likewise, a measurable record of transposable element insertion, exaptation, and recombination within these same genes provides a primary mechanism for bursts of adaptation and evidence of accelerated evolution in these genes across select species lends further support to this notion.

Future avenues of research may include paired approaches using cladistics and genetics, with a particular focus on evolutionary patterns of DevReg genes across extant species and their relationship to morphological divergence. In addition, the study of these genes in species that share cryptic biodiversity may provide additional contrast. With further investigation, we believe the Developmental Gene Hypothesis will provide biologists a better understanding of observable trends within the fossil record and of evolution of this unique and important group of genes responsible for animal form.

FUNDING

This work was funded by National Institutes of Health grant [R01 HD-65279].

REFERENCES

1. Gould SJ, Eldredge N, *Sys. Zool* 1986, 35, 143.
2. Penny D, *Syst. Zool* 1985, 34, 201.
3. Gould SJ, Eldredge N, in *Phyletic Gradualism: Our Old and Present Picture* 1972, pp. 82–115.
4. Casanova EL, Switala AE, Dandamudi S, Hickman AR, Vandenbrink J, Sharp JL, et al., *Autism Res.*, 2019, 12, 860. [PubMed: 31025836]
5. Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, et al., *Cell Stem Cell*, 2019, 24, 724. [PubMed: 31006620]
6. Chuong EB, Elde NC, Feschotte C, *Nat. Rev. Genet.*, 2017, 18, 71. [PubMed: 27867194]
7. Comfort NC, *J Hist Biol.*, 1999, 32, 133. [PubMed: 11623812]
8. Platt RN 2nd, Vandeweghe MW, Kern C, Schmidt CJ, Hoffmann FG, Ray DA, *Mol. Biol. Evol.*, 2014, 31, 1536. [PubMed: 24692655]
9. Oliver KR, Greene WK, *Mobile DNA*, 2011, 2, 8. [PubMed: 21627776]
10. Zeh DW, Zeh JA, Ishida Y, *Bioessays*, 2009, 31, 715. [PubMed: 19472370]
11. Davidson EH, *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*, Elsevier, Amsterdam, 2006.
12. Britten RJ, Davidson EH, *Q. Rev. Bio* 1971, 46, 111. [PubMed: 5160087]
13. Elgar G, Vavouri T, *Trend. Genet.*, 2008, 24, 344.
14. Sundaram V, Cheng Y, Ma Z, Li D, Xing X, Edge P, et al., *Genome Res.*, 2014, 24, 1963. [PubMed: 25319995]
15. Stephen S, Pheasant M, Makunin IV, Mattick JS, *Mol. Biol. Evol.*, 2008, 25, 402. [PubMed: 18056681]
16. Stefflova K, Thybert D, Wilson MD, Streeter I, Aleksic J, Karagianni P, et al., *Cell*, 2013, 154, 530. [PubMed: 23911320]
17. Dimitrieva S, Bucher P, *Bioinformatics*, 2012, 28, i395. [PubMed: 22962458]
18. Sironi M, Menozzi G, Comi GP, Cagliani R, Bresolin N, Pozzoli U, *Hum. Mol. Genet.*, 2005, 14, 2533. [PubMed: 16037065]
19. Galis F, Sinervo B, *Contribut. Zool*, 2002, 71, 101.
20. Vavouri T, Walter K, Gilks WR, Lehner B, Elgar G, *Genome Biol.*, 2007, 8, R15. [PubMed: 17274809]
21. Tabin CJ, Carroll SB, Panganiban G, *Am. Zoologist*, 1999, 39, 650.
22. Gould SJ, *Wonderful Life: The Burgess Shale and the Nature of History*, New York: W. W. Norton, 1989.
23. Wood R, Liu AG, Bowyer F, Wilby PR, S Dunn F, Kenchington CG, et al., *Nat. Ecol. Evol.*, 2019, 3, 528. [PubMed: 30858589]
24. López-Bigas N, N., Ouzounis CA, *Nucl. Acid. Res.*, 2004, 32, 3108.
25. Ekman D, Light S, Björklund A, Elofsson A, *Genome Biol.*, 2006, 7, R45. [PubMed: 16780599]

26. Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, L., Bourque G, et al., PLoS Genet., 2013, 9, e1003470. [PubMed: 23637635]
27. Piriyaopongsa J, Mariño-Ramírez L, L., Jordan IK, Genetics, 2007, 176, 1323. [PubMed: 17435244]
28. Irie N, N., Kuratani S, S. 2011. Nat. Commun, 2011, 2, 248. [PubMed: 21427719]
29. Arthur W, The W Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology, Cambridge University Press, Cambridge, UK, 1997.
30. Heyn P, Kircher M, Dahl A, Kelso J, Tomancak P, Kalinka AT, Neugebauer KM, Cell Report., 2014, 6, 285.
31. Fang H, Yang Y, Li C, Fu S, Yang Z, Jin G, et al., Dev. Cell, 2010, 19, 174. [PubMed: 20643359]
32. Casanova EL, Gerstner Z, Sharp JL, Casanova MF, Feltus FA, Front. Psychiatry, 2018, 9, e535.
33. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, et al., Nature, 2006, 444, 499. [PubMed: 17086198]
34. Richardson L, Venkataraman S, Stevenson P, Yang Y, Moss J, Graham L, et al., Nucleic Acids Res., 2014, 42, D835. [PubMed: 24265223]
35. Vulto-van Silfhout AT, Rajamanickam S, Jensik PJ, Vergult S, de Rocker N, Newhall KJ, et al., Am. J. Hum. Genet, 2014, 94, 649. [PubMed: 24726472]
36. Veraska A, Kennison J, McGinnis W, Genesis, 2002, 33, 67. [PubMed: 12112874]
37. Raff EC, Popodi EM, Kauffman JS, Sly BJ, Turner FR, Morris VB, et al., Evol. Dev, 2003, 5, 478. [PubMed: 12950627]
38. Richardson MK, Dev. Biol, 1995, 172, 412. [PubMed: 8612960]
39. Richards RJ, Biol. Phil, 2009, 24, 147.
40. Gould SJ, Eldredge N, Nature, 1993, 366, 223. [PubMed: 8232582]
41. Comeran JM, Kreitman M, Genetics, 2000, 156, 1175. [PubMed: 11063693]
42. Drake JA, Bird C, Nemesh J, Thomas DJ, Newton-Cheh C, Reymond A, et al., Nat. Genet, 2006, 38, 223. [PubMed: 16380714]
43. Gazave E, Marqués-Bonet T, Fernando O, Charlesworth B, Navarro A, Genome Biol., 2007, 8, R21. [PubMed: 17309804]
44. Hubisz MJ, Pollard KS, Curr. Opin. Genet. Dev, 2014, 29, 15. [PubMed: 25156517]
45. Presgraves DC, Nat. Rev. Genet, 2010, 11, 175. [PubMed: 20051985]
46. Pfenninger M, Schwenk K, BMC Evol. Biol, 2007, 7, 21. [PubMed: 17300732]
47. True JR, Haag ES, Evol. Dev, 2001, 3, 109. [PubMed: 11341673]
48. King MC, Wilson AC, Science, 1975, 188, 107. [PubMed: 1090005]
49. Cherty LM, Case SM, Wilson AC, Science, 1978, 200, 209. [PubMed: 635583]
50. Barrier M, Robichaux RH, Purugganan MD, PNAS, 2001, 98, 10208. [PubMed: 11517318]
51. Lowe C, Haussler D, 2012. PLoS One, 2012, 7, e43128. [PubMed: 22952639]
52. Lowe CB, Bejerano G, Haussler D, Proc. Natl. Acad. Sci. USA, 2007, 104, 8005. [PubMed: 17463089]
53. Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, et al., Nature, 2006, 441, 87. [PubMed: 16625209]
54. Villar D, Berthelot CS, Aldridge S, Rayner TF, Lukk M, Pignatelli M, et al., Cell, 2015, 160, 554. [PubMed: 25635462]
55. Schmitz J, Brosius J, Biochimie, 2011, 93, 1928. [PubMed: 21787833]
56. Ricci M, Peona V, Guichard E, Tacciolo C, Boattini A, J. Mol. Evol, 2018, 86, 303. [PubMed: 29855654]
57. Kazazian HH, Science, 2004, 303, 1626. [PubMed: 15016989]
58. Batzer MA, Deininger PL, Nat. Rev. Genet, 2002, 3, 370. [PubMed: 11988762]
59. Bailey JA, Liu G, Eichler EE, Am. J. Hum. Genet, 2003, 73, 823. [PubMed: 14505274]
60. Ray DA, Batzer MA, BMC Evol Biol., 2005, 5, 51. [PubMed: 16209711]
61. Liu GE, Alkan C, Jiang L, Zhao S, Eichler EE, Genome Res., 2009, 19, 876. [PubMed: 19411604]
62. Worley KC, Warren WC, Rogers J, Locke D, Muzny DM, Mardis ER, et al., Nat Genet., 2014, 46, 850. [PubMed: 25038751]

63. Hedges DJ, Callinan PA, Cordaux R, Xing J, Barnes E, et al., *Genome Res.*, 2004, 14, 1068. [PubMed: 15173113]
64. Konkel MK, Walker JA, Batzer MA, *Evol. Anthropol.*, 2010, 19, 236. [PubMed: 25147443]
65. Deininger P, *Genome Biol.*, 2011, 12, 236. [PubMed: 22204421]
66. Han K, Konkel MK, Xing J, Wang H, Lee J, Meyer T, et al., *Science*, 2007, 316, 238. [PubMed: 17431169]
67. Bennet EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, et al., *Res.*, 2008, 18, 1875.
68. Konkel MK, Walker JA, Hotard AB, Ranck MC, Fontenont CC, Storer J, et al., *Genome Biol Evol.*, 2015, 7, 2608. [PubMed: 26319576]
69. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al., *Nature*, 2001, 409, 860. [PubMed: 11237011]
70. Polak P, Domany E, *BMC Genomics*, 2006, 7, 133. [PubMed: 16740159]
71. Athanasiadis A, Rich A, Maas S, *PLoS Biol.*, 2004, 2, e391. [PubMed: 15534692]
72. Jeck WR, Sorrention JA, Wang K, Slevin MK, Burd CE, Liu J, et al., *RNA*, 2013, 19, 141. [PubMed: 23249747]
73. Häslér J, Strub K, *Nucleic Acids Res.*, 2006, 34, 5491. [PubMed: 17020921]
74. Paz-Yaacov N, Levanon EY, Nevo E, Kinar Y, Harmelin A, Jacob-Hirsch J, et al., *Proc Natl Acad Sci USA*, 2010, 107, 12174. [PubMed: 20566853]
75. Sugimoto J, Schust DJ, *Reprod. Sci.*, 2009, 16, 1023. [PubMed: 19474286]
76. Mi S, Lee X, Veldman GM, Finnerty H, Racie L, LaVallie E, et al., *Nature*, 2000, 403, 785. [PubMed: 10693809]
77. Cornelis G, Funk M, Vernochet C, Leal F, Tarazona OA, Meurice G, et al., *Proc. Natl. Acad. Sci. USA*, 2017, 114, E10991. [PubMed: 29162694]
78. Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, et al., *Cell Rep.*, 2015, 10, 551. [PubMed: 25640180]
79. Babich V, Aksenov N, Alexeenko V, Oei SL, Buchlow G, Tomilin N, *Gene*, 1999, 239, 341. [PubMed: 10548736]
80. Hancks DC, Kazazian HH Jr., *Semin. Cancer Biol.*, 2010, 20, 234. [PubMed: 20416380]
81. McFadden J, Knowles G, *J. Theor. Biol.*, 1997, 186, 441. [PubMed: 9278720]
82. Konkel MK, Batzer MA, *Semin. Cancer Biol.*, 2010, 20, 211. [PubMed: 20307669]
83. Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al., *Genome Res.*, 2006, 16, 1182. [PubMed: 16902084]
84. Jurka J, Kohany O, Pavlicek A, Kapitonov VV, Jurka M, *Genome Res.*, 2005, 110, 117.
85. Cardoso AR, Oliveira M, Amorim A, Azevedo L, *Hum. Genomics*, 2016, 10, 30. [PubMed: 27663310]
86. Ashizawa T, Wells RD, in *Genetic Instabilities & Neurological Diseases*, 2nd Ed. (Eds.: Wells RD, Ashizawa T), Academic Press, Amsterdam, Netherlands, 2006, pp. 3–17.
87. Cerbin S, Jiang N, *Curr. Opin. Genet. Dev.*, 2018, 49, 63. [PubMed: 29571044]
88. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al., *Nature*, 2011, 470, 59. [PubMed: 21293372]
89. Jurka J, Kohany O, Pavlicek A, Kapitonov VV, Jurka MV, *Proc Natl Acad Sci USA*, 2004, 101, 1268. [PubMed: 14736919]
90. Han K, Lee J, Meyer TJ, Wang J, Sen SK, Srikanta D, et al., *PLoS Genet.*, 2007, 3, 1939. [PubMed: 17953488]
91. Sen SK, Han K, Wang J, Lee J, Wang H, Callinan P, et al., *Am. J. Hum. Gent.*, 2006, 79, 41.
92. Freudenberg J, Fu YH, Prácek LJ, *Bionformatics*, 2007, 23, 1441.
93. Capra JA, Erwin GD, McKingsey G, Rubenstein JLR, S Pollard K, *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 2013, 368, 20130025. [PubMed: 24218637]
94. De Brakeleer S, De Grève J, Lissens W, Teugels E, *Hum. Mutat.*, 2013, 34, 785. [PubMed: 23420552]
95. Brooks EM, Branda RF, Nicklas JA, O'Neill JP, *Mutat. Res.*, 2001, 476, 43. [PubMed: 11336982]

96. O'Neil J, Tchinda J, Gutierrez A, Moreau L, Maser RS, Wong KK, et al., J. Exp. Med, 2007, 204, 3059. [PubMed: 18070937]
97. Belancio VP, Hedges DJ, Deininger P, Genome Res., 2008, 18, 343. [PubMed: 18256243]
98. Cordaux R, Batzer MA, Nat Rev Genet., 2010, 10, 691.
99. Hancks DC, Kazazian HH Jr., Mob. DNA, 2016, 7, 9. [PubMed: 27158268]
100. Kazazian HH Jr., Moran V, N Engl J Med., 2017, 377, 361. [PubMed: 28745987]
101. Lam HT, Mu XJ, Stütz AM, Tanzer A, Cayting PD, Snyder M, et al., Nat. Biotechnol, 2010, 28, 47. [PubMed: 20037582]
102. Marques-Bonet T, Girirajan S, Eichler EE, Trends Genet., 2009, 25, 443. [PubMed: 19796838]
103. Geisler JH, Curr. Biol, 2017, 27, R596. [PubMed: 28633029]
104. Thewissen JG, Cooper LN, Clementz MT, Bajpai S, Tiwari BN, Nature, 2007, 450, 1190. [PubMed: 18097400]
105. Uyeda JC, Handsen TF, Arnold SJ, Pienaar J, Proc. Natl. Acad. Sci. USA, 2011, 108, 15908. [PubMed: 21873251]
106. Tollis M, Robbins J, Webb AE, Kuderna LFK, Caulin AF, Garcia JD, et al., Mol. Biol. Evol, 2019, 36, 1746. [PubMed: 31070747]
107. Sela N, Kim E, Ast G, Genome Biol., 2010, 11, R59. [PubMed: 20525173]
108. Chisholm GE, Cooper TG, J. Bacteriol, 1992, 174, 2548. [PubMed: 1313414]
109. Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, et al., Nature, 2015, 524, 220. [PubMed: 26268193]
110. Kuba MJ, Byrne RA, Meisel DV, Mather JA, J. Comp. Psychol, 2006, 120, 184. [PubMed: 16893255]

BOX 1:**HIGHLIGHTS**

- The TE (transposable element)-Thrust Hypothesis (Oliver & Greene, 2011) and TE-epigenomic arms race by Zeh et al. (2009) do not address the active role developmental regulatory genes play as a barrier to TE insertion and other mutations due to loss-of-function sensitivity.
- Developmental regulatory genes maintain significant density of intronic TE and conserved noncoding elements (CNE) in human.
- CNE are often derived from TE insertion events.
- Young CNE occur in the same regions as older CNE, indicating selective pressure for their aggregation in these regions.
- The Developmental Gene Hypothesis proposes that patterns of resistance to and retention of TE within DevReg genes further supports the genetic basis of Punctuated Equilibrium (PE).

BOX 2:**OUTSTANDING QUESTIONS**

- How has the evolution of developmental regulatory genes varied across species, genera, class, phyla, and even kingdoms?
- When comparing two closely related genomes, will accelerated evolution be consistently apparent in and near developmental regulatory genes when morphology is significantly divergent between the two species?
- Do older classes of developmental regulatory genes display a different evolutionary pattern than younger classes and are the latter more tightly conserved?
- How have transposable elements affected the evolving regulome of developmental regulatory genes across vertebrates and invertebrates?

BOX 3:**GLOSSARY****Punctuated Equilibrium (PE):**

A theory that posits the fossil record is characterized by episodes of rapid speciation bookended by broader periods of morphological stasis.

Developmental Regulatory (DevReg) Genes:

Genes involved in regulating morphogenesis and embryonic development.

Transposable elements (TE):

Sequences of DNA that move about the genome and replicate either through transposition (cut-and-paste) or retrotransposition (copy-and-paste) mechanisms.

Ultraconserved Noncoding Elements (UCNE):

Ancient noncoding elements that are strongly conserved across vertebrates or older phyla and are presumably involved in regulation.

Conserved Noncoding Elements (CNE):

Noncoding elements involved in regulation that are strongly conserved across species (e.g., placentals or mammals) but more limited in span compared to ultraconserved noncoding elements.

Regulome:

The full complement of regulatory components in the cell.

Phylotypic Period:

The most conserved period of prenatal development during which an animal most closely resembles other species of its phylum.

Figure 1.

A schematic for the Developmental Gene Hypothesis. Evolution of developmental regulatory (DevReg) genes is strongly conserved, though is not impervious to genetic drift. Changes primarily within non-DevReg genes may result in cryptic biodiversity (i.e., two species that cannot be morphologically differentiated), which at the genetic level are deemed separate species. Eventually, flurries of transposition events, exaptation of more ancient transposable elements (TE), and/or recombination as a result of TE-influenced genomic instability in or surrounding DevReg genes—together with environmental influences—may lead to a “punctuated” series of rapid morphological change. These changes in morphology are measurable within the fossil record and have given rise to the theory of punctuated equilibrium ^[3] as well as the Developmental Gene Hypothesis presented here.

Figure 2.

Deaf1 is a developmental regulatory gene that is common to bilateria. (A) Artist illustration of embryonic *Deaf1* expression representing *in situ* hybridization in mouse and zebrafish and immunofluorescence staining in fruit fly, compared across roughly equivalent stages [embryonic day (E) 10.5, hours post fertilization (hpf) 24, and stage (S) 13 respectively]. The widespread expression across numerous developing organs is apparent in all three species despite significant differences in morphology. Drawings based on Richardson et al. [34], Vulto-van Silfhout et al. [35], and Veraska et al. [36]. (B) Amino acid sequence similarity

of Deaf1 protein across the three species. Mouse : zebrafish = 53.39% identity; mouse : fruit fly = 55.65% identity; zebrafish : fruit fly = 31.65% identity.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Correlation of total intronic transposable elements (TE) and conserved noncoding elements (CNE) per gene across the entire human genome. The strong correlation suggests a significant relationship between the two, potentially reflecting the TE derivation of many CNE. (Data published in Casanova et al. ^[4])

Figure 4.

The ichthyosaur, *Stenopterygius*, known as “Jurassic Mom,” (A) which died and was fossilized while pregnant with seven fetuses, shows clear evidence of viviparity. The skull of one of the fetuses can clearly be seen within the mother’s ribcage (B) (indicated by white arrowhead) along with at least two other vertebral columns from other fetuses located superior to the skull. (Fossil housed at the Houston Natural History Museum in Houston, Texas.)

Table 1.

Average protein length, total and relative densities of conserved noncoding elements (CNE) and transposable elements (TE), and transcript numbers per gene within the human genome broken down according to gene size. Note that protein length and transcript numbers increase consistently with gene size, although intron length is primarily responsible for overall gene size. CNE density increases with gene size but is somewhat variable and TE density peaks between the 10,000bp-200,000bp then modestly decreases with gene size. The vast majority of genes in the human genome (91%) fall under 150,000bp in size. (See Casanova et al., 2019, Supplementary Material 2 for full published data.)

Gene Length (bp)	N (genes)	Protein Length (aa)	Total CNE (CNE/gene)	CNE Density (CNE/gene normalized by gene length)	Total TE (TE/gene)	TE Density (TE/gene normalized by gene length)	Number of Transcripts (per gene)
147–10,000	5,802	318	39	6.4	4	0.7	3.2
10,001–25,000	4,311	465	183	11.0	26	1.5	4.0
25,001–50,000	3,395	603	411	11.5	60	1.7	4.4
50,001–100,000	2,631	751	796	11.4	117	1.7	4.5
100,001–150,000	1,155	964	1,349	11.1	192	1.6	4.8
150,001–200,000	562	1,078	1,935	11.2	263	1.5	4.8
200,001–400,000	773	1,146	3,259	11.7	399	1.4	5.2
400,001–600,000	228	1,212	5,661	11.8	646	1.3	5.5
600,001–800,000	70	1,297	8,062	11.7	929	1.3	5.7
800,001–1,000,000	41	1,257	11,250	12.5	1,160	1.3	5.8
1,000,001–1,200,000	19	894	11,923	10.5	1,421	1.3	6.1
1,200,001–1,400,000	11	999	16,130	12.7	1,732	1.4	5.8
1,400,001+	17	1,846	24,229	13.0	2,325	1.3	6.6