Implications of duplicated *cis*-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty

Senda Jiménez-Delgado[1], Juan Pascual-Anaya, Jordi Garcia-Fernàndez*

Departament de Genètica and Insitut de Biomedicina (IBUB) of Universitat de Barcelona, Barcelona, Spain

**Corresponding author.*

[1]Present address: Center of Regenerative Medicine in Barcelona (CMRB-PRBB)

Words in abstract: 213
Words in manuscript: 4050
Figures: 2
References: 44

Corrrespondence to:
Jordi Garcia-Fernàndez
Departament de Genètica
Facultat de Biologia
Unviersitat de Barcelona
Av. Diagonal 645
08028 Barcelona, Spain
Tel: +34 934034437
Fax: +34 934034430
email: jordigarcia@ub.edu

Abstract

The discovery that most regulatory genes were conserved among animals from distant phyla challenged the ideas that gene duplication and divergence of homologous coding sequences were the basis for major morphological changes in metazoan evolution. In recent years, however, the interest for the roles, conservation and changes of non-coding sequences grew-up in parallel with genome sequencing projects. Presently, many independent studies are highlighting the importance that subtle changes in *cis*-regulatory regions had in the evolution of morphology trough the Animal Kingdom. Here we will show and discuss some of these studies, and underscore the future of *cis*-Evo/Devo research. Nevertheless, we would also explore how gene duplication, which includes duplication of regulatory regions, may have been critical for spatial or temporal co-option of new regulatory networks, causing the deployment of new transcriptome scenarios, and how these induced morphological changes were critical for the evolution of new forms. Forty years after Susumu Ohno famous sentence "natural selection merely modifies, while redundancy creates", we suggest the alternative: "natural selection modifies, while redundancy of *cis*-regulatory elements innovates", and propose the DDI model (from Duplication-Degeneration-Innovation) to explain the increased evolvability of duplicated *cis*-regulatory regions. Paradoxically, making regulation simpler by subfunctionalisation paved the path for future complexity or, in other words, “to make it simple to make it complex”.

Keywords: *cis*-regulatory evolution, gene duplication, Evo-Devo, metazoan evolution, DDI model

More than a century after his death, Darwin’s rationale can be revisited with our current knowledge of genetics and development. It is as simple as to notice that, if animals and their fitness are the result of embryonic development, and embryonic development is the result of the deployment of complex genetic networks, then evolution of the wide diversity of metazoans has had to occur through changes and bricolage of these genetic systems and networks. In the particular case of bilaterian metazoans, these changes heavily concentrated in a relatively short period of time, during the Cambrian period, around 450 Myears ago, time in which almost all the current animal morphologies were already established. Understanding the changes that allowed such impressive explosion of animal forms is the central ambition that currently occupies many evolutionary scientists’ brains and hands.

A paradox that is not a paradox

Since the identification and characterization of the anterior-posterior patterning Hox genes in

Dosophila [1], hordes of other master developmental genes have been characterized, and BMP's, Wnt's, FGF's, TGF's and many other gene family abbreviations are incorporated in any literature regarding embryonic development in almost any animal model. [2]. The major generalisation of decades of research of developmental genetics at the end of the XX century was clear: major regulatory genes are conserved in almost all metazoans. This genetic toolkit patterns major body parts and structures in different animals and was probably set up in the first metazoans, and surely was present by the time of the Cambrian Explosion. The exciting result of commonality of developmental gene networks, however, raises a most puzzling conundrum: if all metazoans use the same genetic networks, how can they be so different?

A first proposal to solve the paradox is thinking it is not a paradox: maybe these generalisations are biased to particular species (fly, mouse, nematode) and to particular genes, and a deep analyses will show up other different genes and gene networks waiting to be discovered controlling the development of different groups.

However, if current generalisations match certainties, the proposal to solve the paradox is rather obvious: slight changes in proteins and protein functions, that is, *trans*-changes (as opposed to cis-regulatory changes for a given gene, see below) (Fig. 1). In this manner, a developmental pathway could change and a new or modified trait be positively selected by evolution. Nonetheless, most developmental genes are involved in multiple pathways, given raise to several, often unrelated, traits. Thus structural changes affecting functional parts of these proteins would collectively affect several traits, and it seems unlikely that all these new traits would be co-selected simultaneously [3], hence that mutations affecting functionally the coding sequence are likely to be selected negatively. This argumentation is enhanced by two lines of evidence; first, functional domains within developmental proteins (e.g., homeodomain, b-HLH) have slower evolutionary rate than the remaining of the coding sequence [4], and second, although there are not many interspecies protein equivalence studies reported, in most of these cases, mutations of different developmental genes can be rescued by the expression of the orthologous proteins of a close or distant species. For example, human *HoxD4* in *Drosophila* substitutes the normal functions of its homologue *Deformed* [5] or the Hydra *achaete-scute* protein forms heterodimers with the endogenous *Daughterless* gene product in *Drosophila*, allowing the correct formation of sensory organs [6].

The former does not imply that there is not evolution of protein sequences at all, and that trans-evolution has never been at the base of morphological evolution. For instance, in insects the protein engrailed has gained a new groucho-interaction motif, not present in the rest of Arthropods, that represses *wingless in vivo* [7]. In addition, a now classic demonstration of the relevance of protein modifications is the gain in the insect lineage of a repression domain in the homeotic protein *Ubx*, which then represses *Distall-less* expression and leg formation in the abdomen, explaining why fruit flies only have six legs [8]. Moreover, amino acidic sequence or structural conservation between species is not purely static: different modifications that affect

protein function are widespread, like the presence of distinct isoforms generated by alternative splicing, or post-translational modifications. For example, depending on differential splicing, the transcription factor AML1 may act as an activator or as a repressor [9]. Recent advances on the relevance of alternative splicing in morphological evolution can be found elsewhere [10]. Even though these and other examples, the vast majority of data shows that structural and functional conservation of proteins between species is extremely high, hence something else has to account for morphological evolution. Evolution of morphology has to be caused by changes in levels, other than changes in developmental proteins.

Therefore, if changes in the coding regions of DNA cannot account for most the morphological diversification, and the genes are the same, changes in temporal and spatial expression (*cis* changes) of such developmental genes, that is, the regulation of their expression, has to correlate with evolutionary change. This idea was firstly noted by King and Wilson [11], who concluded, after analyzing a large set of proteins from human and chimpanzee, that the behavioural and anatomical differences between us and them couldn't be explained by the small molecular differences they exhibited. Thirty years after King and Wilson seminal work, the sequencing of genomes from a great diversity of animal taxa have made possible a large number of comparative genomic and development studies, which make regulatory or cis-evolution the current hypothesis to explain morphological evolution [12]. In a recent report, Wilson and collaborators [13] approached regulatory evolution using a genomic approach. They take advantage of transgenic mice that harboured an extra human chromosome 21. Then they analysed wide expression profiles and showed that the expression of the genes in that human chromosome mimicked that of the human chromosome 21 and not that of the mouse orthologous regions. Given that the transcriptional machinery and cellular components were murine, the conclusion was that the *cis*-regulatory sequences were determinant for the species-specific expression profiles.

Main features of regulatory evolution:* cis*-regulatory elements.

The *cis*-regulatory region of a particular gene is composed by a number of *cis*-regulatory elements (CREs) with binding sites for one or multiple transcription factors, organised in a modular manner. The spatial and temporal window of expression of every gene is controlled by its CREs [14]. The modular organization of the regulatory regions has a strong implication on how developmental genes can evolve, since it feasibly allows changes that affect different pathways independently. Therefore, CRE evolution gives a great flexibility for morphological evolution, by changes affecting the patterning and building of one part of the body, without affecting others [15].

CREs can evolve in different ways. One of them is by sequence mutations, which affect the binding site for a transcription factor by preventing, modifying the affinity or allowing *de novo* the binding of a transcription factor and consequently altering the spatial and/or temporal expression of the gene without implying changes in coding sequences. For instance, the sexual pigmentation of different fruit fly species is due to changes in the *cis*-regulatory region of the *yellow* locus, through gaining binding sites for very conserved transcription factors of wing development, like *Engrailed*, which controls the generation of a spot in male *D.biarmipes* wings [16]. Interestingly indeed, gaining of such cis-regulatory blocks has happened independently in distinct regions of the yellow locus, in different drosophila lineages [17]. This is one of the first reported cases of proper convergence studied at deep molecular level. Furthermore, those CRE have been independently lost in distinct *Drosophila* lineages. Another reported case is the binding site of Abd-B in the same *yellow* locus, which involves pigmentation of the posterior part of the male abdomen of *D. melanogaster.* The secondary lost of these new Abd-B binding sites in other drosophila species results in the lost of such pigmentation [18]. Moreover, a striking case of regulatory evolution is found in the network involved in endoderm determination in echinoderms. Both sea stars and sea urchins use the same core gene network, but upstream regulators and downstream targets of the core are radically distinct in both clades, which indicates that *cis*-regulatory changes happened at least at two levels in these gene networks [19].

Beside these kind of mutations, with involved de novo appearance or lost of CREs, there are examples of CRE reshuffling. The gene *sim,* which regulates the patterning of a narrow ventral midline in arthropods, is a classical example. The regulatory region of this gene has been studied in different insects with a curious result. In the honeybee *Apis mellifera sim* is regulated only by Twist. In the case of *Anopheles gambiae sim* is regulated by the Notch pathway. Curiously, *D.melanogaster sim* is regulated by both Twist and Notch, while *sim* regulation in *D.pseudoobscura* and *D.virilis* resembled *A. gambiae* regulation. *sim* enhancers in these insects are in a dynamic turnover: whereas *A.mellifera* contains several Twist binding sites, it lacks binding sites for Notch effectors. And though in *D.pseudoobscura* and *D.virilis sim* is regulated only by Notch, there is no difference in quality or number of binding sites for Twist and Notch between *D.melanogaster* and the other two drosophilids [20]. This example clearly shows that restructuring in CREs is very dynamic and takes advantage of its modular nature. Furthermore, the reorganization and modular architecture of CREs make it possible that sequences of highly divergent regulatory regions promote equivalent downstream responses at gene expression level. For example, the regulatory region of the *even-skipped* gene from scavenger flies (*Sepsidae*), although show undetectable sequence conservation with its homologue in fruit flies, drives expression pattern in transgenic Drosophila identical to that of the endogenous gene of Drosophila [21].

An additional manner of CRE evolution is the recruitment of pre-existing CREs in other places within the genome, making possible the regulation of a new gene. Transposable elements (TEs) can be the cause of such movement. Bejerano *et al.* [22] described a SINE transposable element in the coelacanth *Latimeria menadoensis* 0.5 Mb upstream of the *ISL1* gene. This element was found to recapitulate *in vivo* part of the expression of the gene, showing enhancer properties. The element has also been reported in mammalian genomes [22]. Furthermore, *in silico* searches of putative transcription factors binding sites suggest multiple target sites, including biding sites for ESR1, TP53, POU5F1, SOX2 and CTFC within distinct families of transposable elements [23]. A further analysis in mammalian genomes reveals that almost 25% of regulatory regions under study contain TE-derived sequences [24], highlighting the importance of mobile DNA for the evolution and dynamics of CREs.

Nonetheless, CRE mutations and reorganizations are under a certain selective pressure, as it happens with protein coding sequences: a gene miss expressed in space or time during development will probably be negatively selected, as it will, in principle, alter the body plan or body homeostasis of the organism. However, as we will suggest below, a way to gain some flexibility in mutations of cis-regulatory sequences is Ohno's duplication potential idea, but applied to cis-regulatory sequences.

Duplication of CREs as a source of variability and evolvability:

When a gene is duplicated, not only the coding region is affected, but also the surrounding regions (Fig. 1A). Susumu Ohno work predated most studies of gene regulation, thus in his book *Evolution by Gene Duplication*, he focussed on the high variation potential of new proteins [25]. The duplicated genes and regulatory sequences reduce the constraints to change, since the probability for lethal mutations in all copies at once is very low. Some copies are lost because these mutations can accumulate in a duplicated copy, resulting in a pseudogene, or else, a copy might suffer mutations in its regulatory sequences that turn off its expression. Nonetheless, besides this degeneration process, duplicated genes might suffer other interesting changes.

Neofunctionalisation:

The neofunctionalisation concept is controversial. It can be viewed as the process by which a protein changes its functional properties, and in this way is common in genes studied by biochemists (e.g., steroid receptors [26]) but developmental biologists also includes here the acquisition of new functions through the co-option of the gene in a new territory or in a distinct time of development, as this co-option might eventually generate a new embryological or physiological function. The first case is not very common on major developmental genes, because the structural characteristics, usually based in the protein domains, determine its biochemical function and those changes generally have a deep impact in highly critical

developmental processes. Nonetheless, proteins might gain new domains by exon-shuffling, as it happened with the bilaterian gene *Hedgehog*, which originated by a combination of a hedge-domain and a more ancient hog/intein-domain [27]. The occurrence of the second case is much more common and is based on changes of the regulatory regions, e.g. the gain of a new CRE motif in a given promoter region alters (often, by recruitment) the spatial and/or temporal expression of the gene. Indeed, when a duplication event occurs, neofunctionalisation by amino acid substitutions is facilitated, because duplicated copies are not under selection constraints. Through accumulation of mutations, duplicated copies might acquire new functional properties (Fig.1 C). However, bibliography is scarce for such sort of mutations. We here suggest that i) duplication also facilitates neofunctionalisation by means of changes in regulatory elements, and ii) this phenomenon was widespread, particularly at the origin and early evolution of Vertebrates. For instance, the involvement of *Hox* genes in fin/limb development [28] or the recruitment of Sonic Hedgehog for determination of digits in mammals [29] are nice examples of the gain of a new expression territory, after gene duplication. Neofunctionalisation of regulatory regions does not need duplication, but we propose that the extra regulatory material generated after duplication might increase the chances of such neofunctionalisation events.

Subfunctionalisation:

An interesting process that might occur when a gene is duplicated is the subsequent copy subfunctionalisation, a process described by the DDC model (Duplication-Degeneration-Complementation) [30]. The model described the differential loss of CREs in two duplicated copies, which divided their function, so that only the expression of the two genes, as a group, recapitulated the expression of the original ancestral gene. Several cases of DDC have been clearly identified. For example, the zebra fish *Hoxb1a* and *Hoxb1b* subfunctionalized the expression domains of the ancestral gene *Hoxb1* without altering the protein functionality [31]. Another illustrative example is the expansion up to eight copies of the Hairy genes in the cephalochordate amphioxus, whereas mice bear a single hairy gene. The combined expression pattern of four out of those eight genes mimic the expression of the single mouse Hairy gene [32]. Computing analyses of the 5' region of these duplicates suggested complex events of differential degeneration between the amphioxus paralogues, [33]. Finally, the *Xenopus laevis* genome has retained approximately 40% of the duplicated genes produced in a recent whole genome duplication (WGD) event when compared to its unduplicated relative *X. tropicalis*, Sixty eight out of 1300 pairs of duplicated genes suffered a reduction of gene expression in some tissues, and one third of them had followed a process of subfunctionalization [34].

The DDI model (Duplicacion-Degeneration-Innovation)

Here we propose that subfunctionalisation and neofunctionalisation are not only not mutually exclusive phenomena, but rather they are intimately related (fig 1B). Subfunctionalisation of

highly regulated genes is a very common event. Subsequently, enhancers and other regulatory modules that were lost in the duplicated copies might well become raw material retaining structural enhancer properties, which by subtle mutations –supposedly neutral, as these regions became not functional after the DDC– might thus acquire new spatial-temporal expression profiles, and this on its own might mean neofunctionalisation, or even result in morphological innovation.

Gene duplication *per se* is not a real force to induce morphologic changes, but rather, in our view, may facilitate them, although not precisely in the way that Susumu Ohno suggested. We now know that gene networks are essentially the same in most metazoans, and that they were present, more or less interconnected, at the origin of metazoans. What we suggest is that gene duplication by means of subfunctionalisation diminishes gene pleiotropy, and that this reduced level of pleiotropy increases the evolvability of gene regulation. In the original DCC model [30], neofunctionalisation by acquisition of new enhancers was included as a possible route after gene duplication. We go further and propose, more explicitly, that after duplication the duplicate genes became more specialized, so that they are expressed in fewer territories. Therefore, their regulatory regions became simpler, and as such, less constrained to change. We surmise that, as a consequence, subtle changes in these "fresher" regulatory regions would allow co-option of these duplicated genes in new territories and times. These novel co-options might allow new interconnections amongst distinct gene networks, increasing their complexity and connectivity, hence providing the grounds for further increase in organism complexity, that is, paving the path for evolution (controversially assuming here that evolution may be linked to an increase in complexity). What we think is extremely exciting, is the apparent paradox within a paradox, namely, that the initial reduction of pleiotropy after gene duplication might eventually result in a huge leap in the complexity of gene networks interaction, that is, in an increase in organism complexity. We propose the term DDI (from Duplication-Degeneration-Innovation) to explain this increased evolvability of duplicated *cis*-regulatory regions. The apparent solution to this conundrum: "making it simple to make it complex".

The ambition to the future: testing* cis*-regulatory evolution experimentally:

Testing any scientific hypothesis in the lab by means of *in vivo* or *in vitro* experiments is a crucial part of any science project, and Evolution should become a discipline not out of this logical process. However, most evolutionary studies so far come from inferences on comparisons (anatomical, genetic, genomic). We are entering the so-called third wave of Evo-Devo [35] that is, after finding the surprising similarities (first wave) and subtle differences (second wave) between animal clades, it is time to test in the lab those genetic changes identified, and induce, if they are the true causal events, those morphological changes predicted.

If *cis*-regulation is the main driving force of evolutionary change, those experimental Evo-Devo studies should concentrate in analysing the function and the change in function of CRE. Two main types of experiments can be envisaged. First, the concrete function of a given identified CRE can be identified using transgenesis in which the putative CRE controls a reporter gene in the endogenous context. By this, we can elucidate which component of the temporal and/or spatial expression is due to a given CRE [36-38]. Of course, bioinformatics has a very important role in identifying *in silico* those good CRE candidates for transgenesis analysis and excellent bibliography to do so is available [13, 34] and Ovcharenko, I., in this issue. Then, when a proper function has been identified, the extent of evolutionary conservation of function must be tested. For this purpose, heterologous transgenesis - CRE from species A in species B- will inform if the function of this *cis*-regulatory region is conserved or not. These sorts of experiments get highly exciting when working with key evolutionary novelties, for instance, the origin of tetrapod limb. Posterior Hox genes of mammalian HoxA and HoxD clusters are expressed in, and control the patterning of developing limbs [39]. Posterior HoxD genes are controlled by the Global Control Region (GCR), a region placed far upstream the cluster [40], which also drives expression of these genes in the genital bud and the central nervous system. The GCR is made up by a set of *cis*-regulatory regions that control all the posterior genes of the cluster as a whole. This meta-*cis*-regulation was probably responsible for the consolidation of vertebrate clusters [41]. Spitz et al. [40] injected a β-lac reporter gene in mice under the control of a 7.8 Kb GCR region from the fish *Tetraodon* and recover reporter expression in the central nervous system (where the expression in the forebrain was enlarged) and the genital bud, but not in limb buds. Consequently, a plausible explanation for this observation was that the tetrapod GCR had gained CREs for the regulation of expression of posterior HoxD genes concomitantly with the origin of the tetrapod limbs. Heterologous injection of dissected tetrapod GCR in fishes and dissected teleost GCR in mouse will help to elucidate the proper syntaxis of CREs that was concomitant with the origin of the tetrapod limb.

The second sort of experiments involves the functional testing of *cis*-regulatory changes. The expression of an endogenous gene under the control of an enhancer that is responsible of an evolutionary novelty might allow knowing whether this enhancer is truly the main responsible to account for such change. In a precious series of experiments, *cis*-changes that allowed the generation of wings from ancestral forelegs in bats are being identified. The enlarged length of bats' forelimbs bones with respect to others mammals is critical for the correct formation of wings, and consequently, for achieving powered flight. This evolutionary novelty is due in part to differences in *Bmp*2 expression in some digits (Fig.2A) [42] and in differences of regulation of *Prx1*, an important regulator of long bone elongation during limb development [43]. First, Sears et al. [42] proposed that the differential level of expression in digits between mice and bats is likely due to differences in *cis*-regulation, by acquisition of a strong enhancer in the bat lineage. And recently, Cretekos et al. [43] substituted the mice enhancer of the *Prx1* gene by the

corresponding region of the bat gene [44]. Strikingly, these transgenic mice exhibited a significant increase of a ~6% in forelimb bone length with respect to wild type (see Fig. 2B). Although it does not explain on its own the leg to wing transition, and surely many other genes/processes were involved, the experiment by Cretekos *et al*. is a breakthrough in the study of *cis*-regulatory evolution. In the coming years, this sort of experiments targeting major evolutionary innovations at key crossroads of evolution will, surely, shed light on the mysteries and mechanisms of evolution.

Conclusions

The answer to the initial question on the molecular mechanism that drove morphological evolution is far from being simple. We emphasized here the potential of CRE evolution, and highlighted an intriguing link between gene duplication and increasing in the evolvability of gene regulation. We introduced the DDI model, in which the differential loss and simplification of regulatory regions after gene duplication increases the chances to evolve new enhancers, increasing the interconnections within regulatory gene networks, a path that may well flow into morphological innovations. However, it does not escape our attention that other processes, like the generation of protein variants by changing alternative splicing, exon shuffling, subtle protein mutations, changes in regulatory RNAs changes in gene editing, and many other processes sure had relevance at distinct evolutionary times. Only a deeper understanding of the evolution of genome architecture and regulation, and in-depth analyses of complex gene networks in a comparative manner, together with the exciting new experimentally approaches to test those changes, will allow us to understand how these multitude of “endless forms most beautiful” appeared in Earth.

Acknowledgements

We are indebted to Manuel Irimia, Ignacio Maeso, Salvatore D’Aniello and Beatriz Albuixech for long lasting and stimulating discussions and to Hannah Katz Rotnitzsky for checking the English. JGF also thanks Gemma Marfany for a highly motivating discussion, Jaume Baguñà for sharing passion and deep knowledge, and Laura for an inspirating dialogue. JGF’s lab is funded by grant BMC2008-03776 from the Spanish *Ministerio de Educación y Ciencia*. JPA holds a FI fellowship (Generalitat de Catalunya).

Key points

- Regulatory gene networks are highly conserved in Metazoans
- Evolution was mostly driven by changes in *cis*-regulatory elements (CREs)

Deleted: h

Formatted: Font: Italic

- After gene duplication, cis-regulatory regions became simpler by subfunctionalisation, beget more evolvable, and smooth the progress to gene recruitment, facilitating innovation and novelty
- We propose the DDI model (Duplication-Degeneration-Innovation) to explain the increase of evolvability of duplicated regulatory regions
- Experimental Evo-Devo will serve to test in the laboratory the regulatory changes responsible of morphological evolution

Biographical note

Trained in Barcelona and Oxford, Jordi Garcia-Fernàndez is Professor of Genetics/Evo-Devo and his lab is excitingly immersed in the goal of experimentally testing the invertebrate–vertebrate transition, using amphioxus as a model system. Senda Jiménez-Delgado did a Ph.D. in the lab studying the origin of somitogenesis, and Juan Pascual-Anaya is currently finishing his Ph.D. centred on the evolution of Hox genes and vertebrate innovations

1. Lewis EB. A gene complex controlling segmentation in *Drosophila*, Nature 1978;276:565-570.
2. Gilbert S. Developmental Biology. Sunderland: Sinauer Associates, 2006.
3. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution, Cell 2008;134:25-36.
4. Lynch VJ, Roth JJ, Wagner GP. Adaptive evolution of Hox-gene homeodomains after cluster duplications, BMC Evol Biol 2006;6:86.
5. McGinnis N, Kuziora MA, McGinnis W. Human *Hox*-4.2 and *Drosophila Deformed* encode similar regulatory specificities in *Drosophila* embryos and larvae, Cell 1990;63:969-976.
6. Grens A, Mason E, Marsh JL et al. Evolutionary conservation of a cell fate specification gene: the *Hydra achaete-scute* homolog has proneural activity in *Drosophila*, Development 1995;121:4027-4035.
7. Hittinger CT, Carroll SB. Evolution of an insect-specific GROUCHO-interaction motif in the ENGRAILED selector protein, Evol Dev 2008;10:537-545.
8. Galant R, Carroll SB. Evolution of a transcriptional repression domain in an insect Hox protein, Nature 2002;415:910-913.
9. Lopez AJ. Developmental role of transcription factor isoforms generated by alternative splicing, Dev Biol 1995;172:396-411.
10. Irimia M, Rukov JL, Roy SW et al. Quantitative regulation of alternative splicing in evolution and development, Bioessays 2009;31:40-50.
11. King M-C, Wilson A. Evolution at two levels in humans and chimpanzees., Science 1975;118:107-116.
12. Carroll SB. Evolution at two levels: on genes and form, PLoS Biol 2005;3:e245.
13. Wilson MD, Barbosa-Morais NL, Schmidt D et al. Species-specific transcription in mice carrying human chromosome 21, Science. 2008;322:434-438.
14. Davidson E. Genomic regulatory systems: development and evolution. San Diego: Academic, 2001.

Formatted: Font: (Default) Times New Roman, English (UK)

Formatted: Font: (Default) Times New Roman, English (UK)

15. Carroll SB. Homeotic genes and the evolution of arthropods and chordates, Nature 1995;376:479-485.
16. Gompel N, Prud'homme B, Wittkopp PJ et al. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila, Nature 2005;433:481-487.
17. Prud'homme B, Gompel N, Carroll SB. Emerging principles of regulatory evolution, Proc Natl Acad Sci U S A 2007;104 Suppl 1:8605-8612.
18. Jeong S, Rokas A, Carroll SB. Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution, Cell 2006;125:1387-1399.
19. Hinman VF, Davidson EH. Evolutionary plasticity of developmental gene regulatory network architecture, Proc Natl Acad Sci U S A 2007;104:19404-19409.
20. Zinzen RP, Cande J, Ronshaugen M et al. Evolution of the ventral midline in insect embryos, Dev Cell 2006;11:895-902.
21. Hare EE, Peterson BK, Iyer VN et al. Sepsid *even-skipped* enhancers are functionally conserved in Drosophila despite lack of sequence conservation, PLoS Genet 2008;4:e1000106.
22. Bejerano G, Lowe CB, Ahituv N et al. A distal enhancer and an ultraconserved exon are derived from a novel retroposon, Nature 2006;441:87-90.
23. Bourque G, Leong B, Vega VB et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements, Genome Res 2008;18:1752-1762.
24. Jordan IK, Rogozin IB, Glazko GV et al. Origin of a substantial fraction of human regulatory sequences from transposable elements, Trends Genet 2003;19:68-72.
25. Ohno S. Evolution by gene duplication. New York: Springer-Verlag, 1970.
26. Freilich S, Goldovsky L, Ouzounis CA, Thornton JM. Metabolic innovations towards the human lineage, BMC Evol Biol 2008; 8:247.
27. Adamska M, Matus DQ, Adamski M et al. The evolutionary origin of hedgehog proteins, Curr Biol 2007;17:R836-R837.
28. Freitas R, Zhang G, Cohn MJ. Evidence that mechanisms of fin development evolved in the midline of early vertebrates, Nature 2006;442:1033-1037.
29. Chan DC, Laufer E, Tabin C et al. Polydactylous limbs in Strong's Luxoid mice result from ectopic polarizing activity, Development 1995;121:1971-1978.
30. Force A, Lynch M, Pickett FB et al. Preservation of duplicate genes by complementary, degenerative mutations, Genetics 1999;151:1531-1545.
31. McClintock JM, Kheirbek MA, Prince VE. Knockdown of duplicated zebrafish *hoxb1* genes reveals distinct roles in hindbrain patterning and a novel mechanism of duplicate gene retention, Development 2002;129:2339-2354.
32. Minguillon C, Jiménez-Delgado S, Panopoulou G et al. The amphioxus Hairy family: differential fate after duplication, Development 2003;130:5903-5914.
33. Jiménez-Delgado S, Crespo M, Permanyer J et al. Evolutionary genomics of the recently duplicated amphioxus *Hairy* genes, Int J Biol Sci 2006;2:66-72.
34. Sémon M, Wolfe KH. Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*, Proc Natl Acad Sci USA. 2008;105:8333-8338.
35. Baguñá J, Garcia-Fernàndez J. Evo-Devo: the long and winding road, Int J Dev Biol 2003;47:705-713.
36. Hadrys T, Punnamoottil B, Pieper M et al. Conserved co-regulation and promoter sharing of *hoxb3a* and *hoxb4a* in zebrafish, Dev Biol 2006;297:26-43.

37. Russo MT, Donizetti A, Locascio A et al. Regulatory elements controlling *Ci-msxb* tissue-specific expression during *Ciona intestinalis* embryonic development, Dev Biol 2004;267:517-528.
38. Yu JK, Holland ND, Holland LZ. Tissue-specific expression of *FoxD* reporter constructs in amphioxus embryos, Dev Biol 2004;274:452-461.
39. Zakany J, Duboule D. The role of Hox genes during vertebrate limb development, Curr Opin Genet Dev 2007;17:359-366.
40. Spitz F, Gonzalez F, Duboule D. A Global Control Region Defines a Chromosomal Regulatory Landscape Containing the HoxD Cluster, Cell 2003;113:405-417.
41. Duboule D. The rise and fall of Hox gene clusters, Development 2007;134:2549-2560.
42. Sears KE, Behringer RR, Rasweiler JJt et al. Development of bat flight: morphologic and molecular evolution of bat wing digits, Proc Natl Acad Sci U S A 2006;103:6581-6586.
43. Cretekos CJ, Wang Y, Green ED et al. Regulatory divergence modifies limb length between mammals, Genes Dev 2008;22:141-151.
44. Martin JF, Olson EN. Identification of a *prx1* limb enhancer, Genesis 2000;4:225-229.

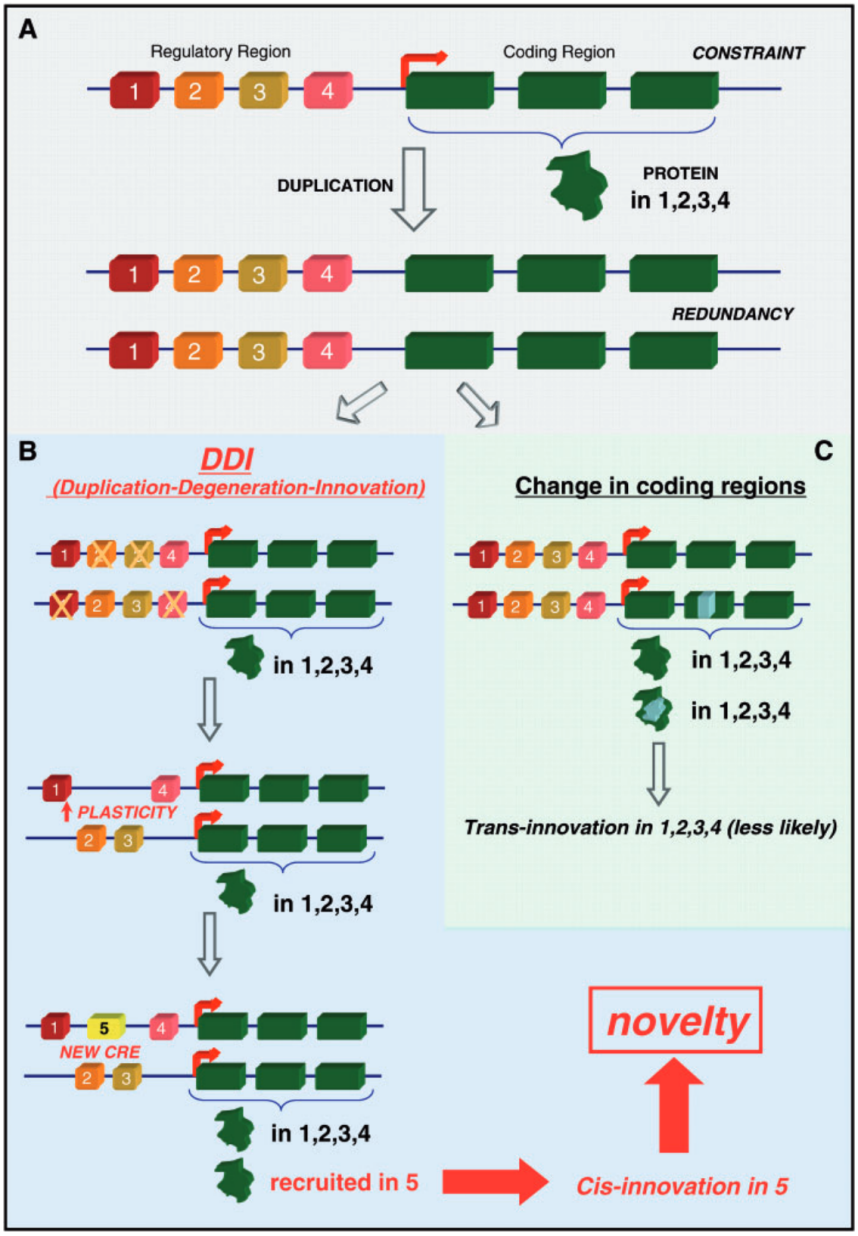

Figure 1. **The DDI model (Duplication-Degeneration-Innovation) for regulatory evolution.** Or how simplicity after duplication and degeneration turns into plasticity and innovation. The scheme shows a pleiotropic gene (**A**) regulated by four enhancers that drive expression in distinct temporal-spatial territories. The gene is evolutionary constrained, as modifications of the coding region or the complex regulatory region may affect the functionality of the gene in

519 multiple territories. After gene duplication, changes in coding regions (**C**), as suggested by
520 Sosumo Ohno, would allow the expression of a new protein in all the territories in which the
521 original gene was expressed. For high pleiotropic genes, this will likely be negatively selected,
522 as it will affect too many processes. (**B**) DDI: After the differential degeneration of regulatory
523 enhancers [30] in the duplicated copies (subfunctionalisation), the resulting genes will had its
524 pleiotropy reduced, and the regulation of a particular copy will be simpler than originally. This
525 will turn into higher plasticity or evolvability of these regions, thus facilitating the appearance of
526 new regulatory regions which will allow the recruitment of the gene and its associated gene
527 network in a new territory (5), paving the path to morphological innovations.
528

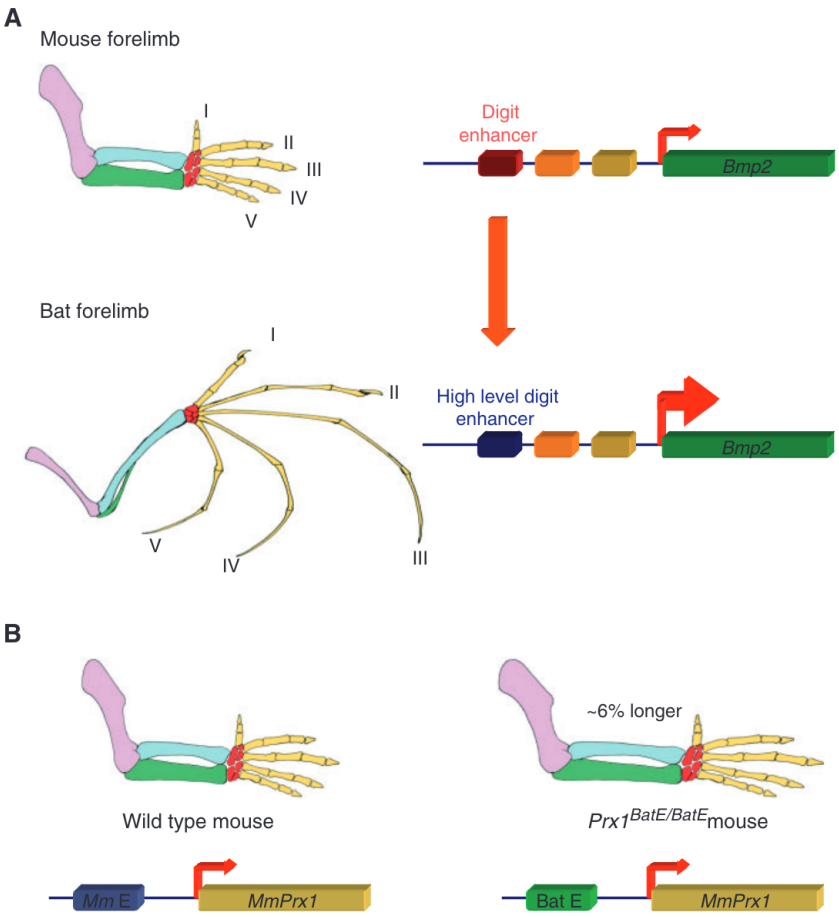

Figure 2. **Regulatory evolution and experimental Evo-Devo.** (**A**) Bats wing evolved from an ancestor with forelimbs similar to actual mice. The BMP2 gene is expressed at high levels in the digits of developing bat forelimb, probably due to a regulatory change in the regulatory region of the gene. The high level of BMP2 expression explains partly the increase in the size of the skeletal components of the bat forelimb. (**B**). Experimental Evo-Devo in vertebrates. Transgenic mice in which the endogenous enhancer of the limb-expressed gene *Prx1* was replaced by the *BatE* enhancer of the bat *Prx1* gene showed an increase of ~6% in the length of the forelimb, mimicking partly the increase that facilitated the origin of flights in the bat lineage. A based in reference [42], B based in reference [43].

Note added in proof

In a recent article, Kassahn and collaborators analysed changes in protein domains and regulatory control in the teleost lineage, and concluded that the most significant consequence of whole genome duplication was to enable more specialized regulatory control of development via de acquisition of novel spatio-temporal domains in duplicated gene copies, in a sharp example of the DDI model. (Evolution of gene function and regulatory control after whole-genome duplication: comparative analyses in vertebrates. Genome Research e-pub 12/09/2009)