Deep homology in the age of next-generation sequencing

Patrick Tschopp[1†], Clifford J. Tabin[1,*]

1: Department of Genetics, Harvard Medical School, 77 Ave Louis Pasteur, Boston, MA 02115, USA

†Present address: Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland.

**: tabin@genetics.med.harvard.edu*

Orcid ID, Patrick Tschopp: 0000-0001-7339-3990

Keywords: morphological novelty, regulatory evolution, transcriptomics, gene regulatory networks, enhancers, genome structure

Summary

The principle of homology is central to conceptualizing the comparative aspects of morphological evolution. The distinctions between homologous or non-homologous structures have become blurred, however, as modern evolutionary developmental biology (evo-devo) has shown that novel features often result from modification of pre-existing developmental modules, rather than arising completely *de novo*. With this realization in mind, the term “deep homology” was coined, in recognition of the remarkably conserved gene expression during the development of certain animal structures that would not be considered homologous by previous strict definitions. At its core, it can help to formulate an understanding of deeper layers of ontogenetic conservation for anatomical features that lack any clear phylogenetic continuity. Here, we review deep homology and related concepts in the context of a gene expression-based homology discussion. We then focus on how these conceptual frameworks have profited from the recent rise of high-throughput next-generation sequencing. These techniques have greatly expanded the range of organisms amenable to such studies. Moreover, they helped to elevate the traditional gene-by-gene comparison to a transcriptome-wide level. We will end with an outlook on the next challenges in the field and how technological advances might provide exciting new strategies to tackle these questions.

1. Introduction

Understanding the origins of the vast array of morphological features displayed by creatures in the natural world is a core problem in evolutionary biology. Much of the variety we can observe in living species arises through changes to structures present in ancestral species, a process neatly explained by Darwins' concept of descent with modification. It has been more problematic to explain the appearance of seemingly novel anatomic structures. The principle of homology was developed to distinguish between these situations, referring to the existence of shared ancestry between a pair of structures. However, how to best define homology has been contentious [1,2]. As others have previously pointed out, this has to some extent been a problem of the hierarchical level at which homology is considered [3-5]. For example, while all vertebrate forelimbs can be considered homologous as structural units, this structural entity has been modified evolutionarily to various different 'character states', facilitating distinct, yet in some cases analogous modes of locomotion, such as e.g. wings used for flying. The skeletal elements inside the wings of birds, bats or pterodactyls clearly imply a common basic pattern. This can be traced back to a common forelimb ground state that has subsequently been modified in each of the three lineages. Their inferred historical continuity therefore allows us to identify them as homologous as forelimbs. However, although all three are used for flying, they have to be considered functionally analogous as wings, since flight has evolved independently in these clades. The structures are thus homologous at the level of forelimbs but not at the level of wings. I.e., whether traits are classified as homologous or not becomes a hierarchy issue, dependent on the level at which homology is being discussed.

Conflicting semantics aside, at the core of the homology concept is the notion of "sameness" and some sort of "historical continuity." This has also led to the inclusion of the level of genes and proteins within the general realm of homology,

followed by the rise of modern molecular biology [6]. A similar case can be made at yet another level of organizational hierarchy, when considering homology amongst different cell types, rather than entire organs. Again, phylogenetic as well as functional or structural criteria have historically been used to assess homology of these entities [7].

In this context, it is perhaps best to frame a discussion of homology around its original definition. As first defined by Sir Richard Owen, homology refers to “the same organ in different animals under every variety of form and function” [8]. His pre-Darwinian definition of homology was thus based very much on underlying structural similarities that could not simply be explained by functional constraints. With the advent of Darwinism, Owen’s concept of homology became linked to the historical continuity of morphological structures (e.g. E.R. Lankester [9]), thus implying descent with modification from an “archetype” of a common ancestor. This type of homology is therefore intricately linked to phylogenetics and systematics, and how certain morphological characters are distributed over the evolutionary tree. It is often referred to as “historical homology” [5].

With the advent of comparative “evo-devo” biology, and a better understanding of the molecular mechanisms driving embryogenesis in classical model organisms, some have argued that considering ontogeny would sometimes be more informative when evaluating homology [10-12]. In particular, rather than following strict genealogical criteria, mechanistic constraints on the development of morphological features should be taken into account. This was largely inspired by the realization that distantly related species utilize a remarkably conserved gene toolkit during embryogenesis, for example for patterning their main body axes [13,14]. For the evo-devo field, this represented a unique opportunity to reframe the homology concept with a particular focus on developmental constraints. One of the prevailing criticisms of historical homology to this point was that anatomical structures, unlike genes, are not directly copied and inherited, but rather are

generated *de novo* during the embryonic development of each generation [15,16]. As individualized parts of a species' phenotype, these structures can change independently due to adaptive processes. Their development, however, is usually internally constrained by the underlying genetic blueprint as well as morphogenetic processes that can be inherently self-regulatory. There, the "biological homology" concept, as it has been referred to, is defined for anatomical structures that have a shared set of developmental constraints for their individualization [15,17]. As such, this form of homology mainly concerns phenotypes that result from complex regulatory interactions, rather than single-gene traits, such as color variants [18,19]. Moreover, it also insinuates a certain degree of modularity, within which self-contained developmental units can undergo evolutionary change, for example at the level of gene regulation.

2. Homology and gene expression - kernels, character identity networks and deep homology

Consideration of hierarchy in the regulation of genes led to the formulation of the "kernel" concept [20]. Fundamental to this approach, the genome is treated as a regulatory blueprint for embryogenesis, layered in both its functional impact on developmental patterning as well as its evolutionary age (with newer modules superimposed upon older ones). At the top of this regulatory hierarchy lie the so-called kernels, sub-units of gene regulatory networks (GRN) that are central to bodyplan patterning, exhibit deep evolutionary conservation and are refractory to regulatory rewiring. Their static behavior, and importance in defining fundamental embryonic patterns, has been argued to underlie the stability exhibited by different animal bodyplans since the Cambrian explosion [21]. At the base of this GRN hierarchy, so-called gene differentiation batteries direct terminal cell or organ differentiation. These are assemblies of effector genes that underlie functional specification, but lack regulatory information. Their deployment is controlled by so-called intermediate plug-ins, or I/O-switches, that transmit

kernel-contained patterning information down to its final differentiation output. From an evolutionary perspective, regulatory modifications are most likely to occur at this “plug-in” level, to ultimately result in structural novelty. Therefore, a hierarchy of regulatory homology can, to some extent, be inferred from the re-deployment of these switches. There are a number of examples of extraordinarily deep evolutionary GRN conservation that display “kernel”-like behavior: these include endomesoderm specification in echinoderms, hindbrain regionalization in chordates or, most spectacularly, the specification of heart development in clades as distant as arthopods and chordates (see references in [20]). Albeit structurally very distinct, a core set of regulatory interactions is equally important in directing heart development in these two distantly related phyla. Sub-circuit formations as well as downstream effector genes are remarkably conserved, implying a common regulatory blueprint that traces back to a primitive circulatory organ at the base of the Bilateria [20,22,23] (Fig. 1a). However, one of the potential shortcomings of the kernel concept in assessing homology, is its focus on transcription factors and their associated *cis*-regulatory sequences, with much less attention given to signaling pathways. Moreover, central to its definition is the deep evolutionary conservation that GRN parts have to display, in order to be considered a kernel.

A slightly more flexible approach is provided by the character identity network (ChIN) concept [24]. Again, historical continuity of character-defining gene regulatory networks is key to its definition. However, and unlike kernels, these do not have to be evolutionary ancient (i.e. phylum or sub-phylum level for kernels, down to species level for ChINs). Central to the applicability of ChINs in discussing homology is the inherent modularity of developmental systems. Different body parts and organs develop, and are patterned, in a semi-autonomous fashion, a fact known since the early days of experimental embryology [25]. The division into discrete developmental sub-systems allows for their individualized evolution, yet shared ChINs underlying their formation in

different organisms helps us to identify the resulting anatomical features as homologous. By introducing ChINs, a historical continuity is inferred by means of their repetitive re-deployment during the embryogenesis of each following generation, while modifying their output results in varying character states across species [24]. Such reasoning can help to disentangle conflicting results coming from various lines of research, such as embryology *versus* paleontology, when trying to establish homologies between morphological characters. This has been demonstrated in the assessment of digit identity in extant avian wings, where a pentadacyl ground state has been reduced to a three-digit formula. While the most anterior wing digit develops from an embryonic position usually associated with digit II, the paleontological record of theropod digit loss suggests the remaining most anterior digit to be digit I [26,27; see also Cooper & Towers, this issue]. Using comparative RNA-sequencing revealed a strong transcriptional signature uniting the most anterior digits (MAD) of the forelimbs and hindlimbs (Fig. 1b). This implies, at the ChIN level, that the most anterior digit of the avian wing shares a common developmental blueprint with its hindlimb counterpart, and hence the forelimb digit formula should be considered I, II, III, regardless of the anatomical position from which they develop [28,29]. These findings were recently corroborated by studies using gene expression patterns to identify homeotic identities of digit primordia in species that actually have lost digit I [30]. In the meantime, ChIN-based approaches of homology have also been expanded to address the evolution of novel cell types [31].

Both kernel and ChIN arguments for homology are continuous, at least in part, with the concept of “deep homology” [32], while also mechanistically refining it. The term deep homology was originally coined to describe the repeated use of highly conserved genetic circuits in the development of anatomical features that do not share homology in a strict historical or developmental sense. For example, although evolutionary separated since the Cambrian, and morphologically and developmentally highly divergent, the development of insect and vertebrate

appendages share striking similarities in specifying their embryonic axes [32,33] (Fig. 1c). Such degree of conservation in developmental patterning, it was argued, would be most parsimoniously explained by a common ancestor that possessed a primitive body-wall outgrowth program [33,34]. The genetic blueprint for this outgrowth program would then have been co-opted and re-deployed for the independent evolution of body appendages in different animal phyla, as well as being reactivated in different anatomical locations within a developing organism to give rise to serial homologs (e.g. tetrapod fore- and hindlimbs). Modification of this deeply conserved genetic program would thus represent the molecular framework within which the morphological diversifications of animal appendages would have to be considered. Their development thus shares historical continuity at the level of the gene regulatory circuits, and is said to display deep homology (Fig. 1c). Likewise, the use of similar cellular building blocks, such as the deployment of an ancestral photosensitive cell in the generation of animal eyes, led to the dependence on homologous gene regulatory interactions and thus resulted in morphological structures that display deep homology [33]. As such, the structural entity itself (in this example, the eye) is termed ‘deeply homologous’, based on the re-deployment of genetic circuitries and/or developmental mechanism that themselves display true homology, i.e. a common historical origin. Ultimately, determining whether the expression of similar genetic cassettes underlying the development of two historically non-homologous structures represents deep homology or convergent, and potentially coincidental, deployment of related genes in two independent settings, depends on assessing the number of genes used in common and, more importantly, their epistatic relationships. Accordingly, the ability to confidently identify deep homology is already enhanced through the use of more comprehensive approaches to determine gene expression similarities.

3. Deep homology goes global – from transcripts to transcriptomes

The obvious advantage of assessing deep homology at a genome-wide level is strength in numbers, each probed gene adding robustness to formulate meaningful predictions. The emergence of hybridization-based array-techniques first opened the possibility of measuring the expression of a large set of genes in a single experiment. Moreover, compared to *in situ* hybridization techniques, array-based experiments also yield information about quantitative differences in gene expression (although often at the expense of spatial information). Indeed, several pioneering evo-devo studies exploited the potential of microarrays for comparative gene expression studies across multiple species. Variation in hybridization efficiencies, due to species-inherent sequence differences in the targeted parts of mRNAs, and other technical issues made direct interpretations difficult. A number of normalization and analytics procedures, helped to bypass these problems [35,36]. However, most of these shortcomings can now easily be avoided, thanks to the development of high-throughput next-generation sequencing (NGS) techniques [37]. The advantages of NGS over array-based methods of gene expression measurements are manifold. Massively paralleled sequencing of cDNAs, known as RNA-seq, now allows for genome-wide interrogation of expression status, as well as the global description of transcript structures [38]. RNA-seq experiments also yield a better quantitative assessment of gene expression differences, thanks to a higher dynamic range as compared to hybridization-based methods [39]. Moreover, RNA-seq opens the possibility to compare a broad range of species that traditionally would have been considered “non-model organism”, including those previously excluded by a lack of an available genome sequence [40,41]. As a consequence, comparative large-scale transcriptome studies, covering different species and tissue types, have emerged in recent years as a powerful approach to address questions of morphological evolution and homology [42]. These approaches have been utilized across several taxonomic levels, from comparing transcriptomes of closely related species in a given genus, to spanning the entire metazoan kingdom [43-45]. Such studies are a powerful proof of how one can now go far beyond the

standard realm of model organisms, eliminating the need to focus on just a select few taxa [46,47]. Moreover, NGS enables for the global assessment of quantitative gene expression differences, a parameter known to influence a variety of phenotypes [48-50]. There are, however, also analytical challenges, in normalizing and comparing these data-sets, especially when working with evolutionary distant organisms [50-53]. While many of the early trans-species RNA-seq studies have mainly focused on adult tissue samples, increasingly this approach is being expanded to developmental time-series, across species and embryonic stages. Heterochrony, i.e. species-specific differences in the relative order certain morphological structures appear during embryogenesis, may potentially confound such transcriptome comparisons [53,54]. However, focusing on the temporal dynamics of transcriptome evolution holds the greatest potential to inform us about putative developmental homologies of different anatomical features [28,51]. NGS-based global and quantitative assessment of transcriptome dynamics, across a range of species and developmental time-points, is therefore likely to reveal more cases of deep homology in the near future.

Combining next-generation transcriptomic analyses with comparative embryological and functional experiments can inform us about the developmental mechanisms that lead to the appearance of deep homology. For example, such was the motivation in studying the evolution of genital bud transcriptomes in amniotes, an embryonic structure that was known to share substantial gene expression similarities with developing limbs [55,56]. Comparative lineage tracing experiments uncovered considerable variation in the embryonic origins of amniote external genitalia [51]. In the case of squamates (lizards and snakes), early limb and genital buds share a common cellular source that results in high transcriptome similarity between the two tissues (Fig. 2a,c). In its most radical interpretation, squamate external genitalia development could therefore be considered serially homologous to hindlimbs. In contrast, in mammals, the genital

bud originates from the tailbud mesenchyme. As the genitalia of squamates and amniotes form from distinct embryonic tissues, they are not historically homologous. How then did these critical, functionally analogous, structures evolve? The answer comes from the realization that this shift in tissue of origin in the two clades is accompanied by a relative repositioning of the genitalia-inducing signaling center, the embryonic cloaca. Once positioned next to a different tissue source, the cloaca elicited similar downstream transcriptional responses in a different mesenchymal cell population. This indicates that the mammalian tail bud is competent to respond in a manner reminiscent of the squamate lateral plate tissue. Given the lack of a well-documented fossil record for developing amniote external genitalia, it is difficult to argue for either squamate or mammalian situation to be the ancestral condition. However, the still recognizable similarities in gene regulatory programs underlying both amniote limb and genital development suggest an ancestral limb-like embryonic origin for amniote external genitalia as the more likely scenario [51]. In this manner, mammalian genitalia development maintained a limb-like regulatory blueprint similar to the squamates, despite its now distinct embryonic origin (Fig. 2b,c). Thanks to this regulatory co-option, due to the common inductive signals originating from the cloaca, and a shared transcriptional response to it, squamate and mammalian external genitalia display a deep homology in their development [51]. These transcriptional similarities are likely to be mechanistically linked at the *cis*-regulatory level, as suggested by comparative enhancer studies [57]. Intriguingly, a similar co-option of an ancestral gene regulatory network seems to have paved the way for the emergence of morphological novelties in the external genitalia of several members of the *Drosophila melanogaster* clade [58]. Although shared developmental trajectories seem unlikely in this case of regulatory co-option, the *Drosophila* example further underscores the power of re-utilizing pre-existing genetic cassettes, especially in rapidly evolving structures such as external genitalia [59,60].

4. Homology assessment - gene expression and regulatory strategies

A pressing question when evaluating homology based on gene expression similarities is whether the observed common patterns of gene activity are caused by the same underlying regulatory strategies, or are rather the result of convergent evolution. Even though the evolution of gene expression seems much more constrained than originally assumed [61], it still displays a considerable amount of plasticity and can rapidly diverge in response to altered selective pressures [62]. Consequently, it has been suggested that studying underlying regulatory strategies, rather than gene expression patterns alone, is more informative when evaluating potential synapomorphies of anatomical structures [63]. Likewise, a much stronger argument for deep homology can be made if the *cis*-regulatory circuitries causing the observed gene expression similarities are conserved as well. In the case of deep homology, however, pre-existing gene regulatory modules that can easily be co-opted by the gain of expression of single “gatekeeper” transcription factors also need to be assessed. Probing gene expression at a transcriptome-wide level, thanks to NGS approaches, has opened new experimental avenues to address these questions. By restricting the analysis of whole-transcriptome data to functional sub-classes of genes, GRN inputs can be approximated from the expression status of signaling molecules and/or transcription factors [28,51]. Given the propensity of the latter to bind to DNA, some similarity in GRN regulatory input can be expected in tissue types that show a high degree of correlation for their transcription factor expression profiles. Eventually, though, dedicated experimental data is required to arrive at a more molecular understanding of the regulatory mechanism causing any observed gene expression similarities. Again, several NGS-based approaches pave the way for such epigenomic annotation of regulatory elements, in a variety of tissue types and species. These include chromatin immunoprecipitation followed by sequencing (ChIP-seq) of histone modifications associated with enhancer function [64], as well as methods to assess local and/or global

chromatin structure [65]. Of particular interest are transcriptional enhancers, *cis*-regulatory elements that can activate target genes over hundreds of kilobases away. These elements can function in a highly tissue-specific and temporally resolved fashion, making them potent drivers of morphological evolution. At the same time, by changing only the regulation of a gene, rather than its coding sequence, pleiotropic effects can be avoided [66]. Hence, potential evolutionary modifications of enhancer activities have been the subject of intense investigation in the field of regulatory evolution. For example, using ChIP-seq for histone H3 Lysine 27 acetylation (H3K27Ac), an active enhancer mark, the evolutionary and developmental dynamics of enhancer activities have been mapped in various different organs and across several taxonomic ranks [67-69]. Ideally, such enhancer activity maps will be complemented by the binding profiles of transcription factors known to be important for the development or function of the tissues in question [70,71]. At the level of chromatin structure, local DNA accessibility as well as higher-order folding can inform us of potential enhancer sequences and regulatory strategies at a given locus. DNA accessibility, as defined by nucleosome-sparse regions, can be probed with a variety of NGS-based assays, such as DnaseI-, FAIRE- or ATAC-seq [72-74]. As a result, potential transcription factor binding sites can be defined bioinformatically, using motif search algorithms. Advantages of these techniques include the considerably smaller cell numbers that they require as input, compared to transcription factor ChIP-seq experiments. These methods have been successfully employed to compare global DNA accessibility maps for different species and organs, in adult tissues or across developmental time-points [75,76].

The importance of three-dimensional folding of the DNA strand itself for correct execution of gene regulatory programs, is also becoming increasingly appreciated. Such DNA looping can range from enhancer-promoter interactions over several hundred kilobases, all the way to supra-structural territories on the mega-base scale called topological associated domains, or TADs [77]. Inside of

these TADs regulatory “promiscuity” of enhancer-promoter contacts can occur to some degree. However, regulatory interactions across TAD boundaries seem inhibited, underlining their importance in maintaining proper genome organization and gene regulation [78,79]. Cross-species comparisons of TADs have revealed booth deeply conserved, as well as step-wise evolutionary dynamic assembly of such regulatory domains [80,81]. TAD conservation could thus reveal deep homology of entire regulatory landscapes, just as deep conservation of enhancer activities might inform us about evolutionary relationships of different morphological structures that they help to pattern [75]. Clearly, however, the most instructive insights at the gene regulatory level would come from the integration of different NGS technologies to first describe the regulatory architecture at loci of interest, and then test them in reciprocal, cross-species enhancer reporter experiments [82-84].

5. Concluding remarks and outlook

NGS has already significantly advanced our ability to address questions of homology, both experimentally as well as conceptually, and will likely continue to do so. Overall, it’s fair to assume that the trend of incorporating technological advances from different fields of biology, and indeed the natural sciences in general, will continue in the field of evolutionary and developmental biology, to make it a more inclusive science [85,86]. Already, the next revolution in transcriptome sequencing is on its way, with the recent ability to use single cells as input material in high-throughput experiments [87]. For the field of homology assessment, this holds special importance at several levels. Firstly, it will allow us to study the molecular mechanisms driving cell type differentiation at the relevant resolution, and learn about the evolution and putative homologous counterparts in different species, using comparative approaches [7]. Moreover, at the organ-level, it enables for the cellular deconstruction of morphological character development, which will help to resolve confounding organ cell

heterogeneity that might differ from species to species [53]. Spatial transcriptome information lost during organ dissociation can then be recovered *in silico*, down to cellular resolution, by re-mapping single-cell RNA-seq data onto grids of known marker gene *in situ* hybridization patterns [88,89]. Such high-throughput *in vivo* approaches will benefit from complementary cell culture experiments, where the controlled parameters of an *in vitro* environment can be exploited. Various cell and organ development pathways can already be re-capitulated *in vitro*, thereby helping to define their minimal differentiation requirements [90,91]. Expanding such rationale to a comparative level, between cell lines and organoids from different organisms, will allow for a reductionist approach to character identity development. Moreover, cell culture-based assays for large-scale comparative studies of epigenomic states and enhancer activities can function as invaluable proxies to delineate regulatory logic across species boundaries [92,93].

A more comprehensive inclusion of published data-sets, for example through large-scale consortia, has the potential to increase the predictive power of more targeted comparative studies on gene regulatory strategies, across, species, organs and developmental time-points [94-96]. The creation of public repositories will certainly help this endeavor (see resources in [97]), especially when created with an explicit evo-devo approach in mind [98]. Undoubtedly, though, this will present new challenges in terms of data analysis and integration. Dedicated bioinformatics efforts will thus be required, to tackle the problems inherent to cross-species comparisons, especially when considering large evolutionary distances [50,52,99]. Ideally, the future of evo-devo will gravitate toward such an integrative approach, where comparative embryology, transcriptomics and epigenomics, bioinformatics, as well as functional *in vivo* and *in vitro* work will be incorporated to study these questions at the systems level [85,86,100]. Including comparative embryological and gene regulatory data at the micro-evolutionary level, in addition to the macro-evolutionary perspective of classical evo-devo, will

allow for the integration of ecological constraints, as well as population genetics data [101].

Ultimately, a definite answer to address questions of homology amongst morphological features will only result from a highly integrative approach, taking into account a well-curated paleontological record, studies on the developmental dynamics underlying the ontogeny of these structures in extant species, as well as a molecular and genetic understanding of their underlying mechanism. At the very least, exploiting the power of NGS-technologies to investigate molecular mechanisms during development might help to disentangle conflicting results from the two former fields, i.e. paleontology and embryology (see e.g. digit identities in bird wings [29]) Moreover, building on such qualitative and quantitative molecular data during embryogenesis, modeling approaches could then move the field forward and help to consider deep homology beyond a simple comparative description of gene expression. In particular, a combination of experimental data and modeling approaches might help to delineate a 'configuration space', in which different individualized developmental and gene regulatory systems were free to evolve [86,102]. Following such synthesis, evo-devo research in general might indeed gain certain predictive powers about possible evolutionary trajectories, by defining the range of possible ontogenetic roadmaps [46,47,86]. The concept of deep homology, with its emphasis on how gene transcription can be co-opted in an evolutionary novel context, is likely to prove particularly powerful in delineating the gene regulatory dimension of such configuration spaces.

Acknowledgments

The authors wish to apologize to colleagues whose work could not be cited due to space limitations.

Authors' contribution

P.Tschopp and C.J. Tabin wrote the paper.

Competing Interests

The authors declare no competing interests

Funding

P. Tschopp is supported by an Advanced Postdoc.Mobility fellowship (P300P3_158525) from the Swiss National Science Foundation. Work in the Tabin lab is funded by grant HD032443 from the NIH.

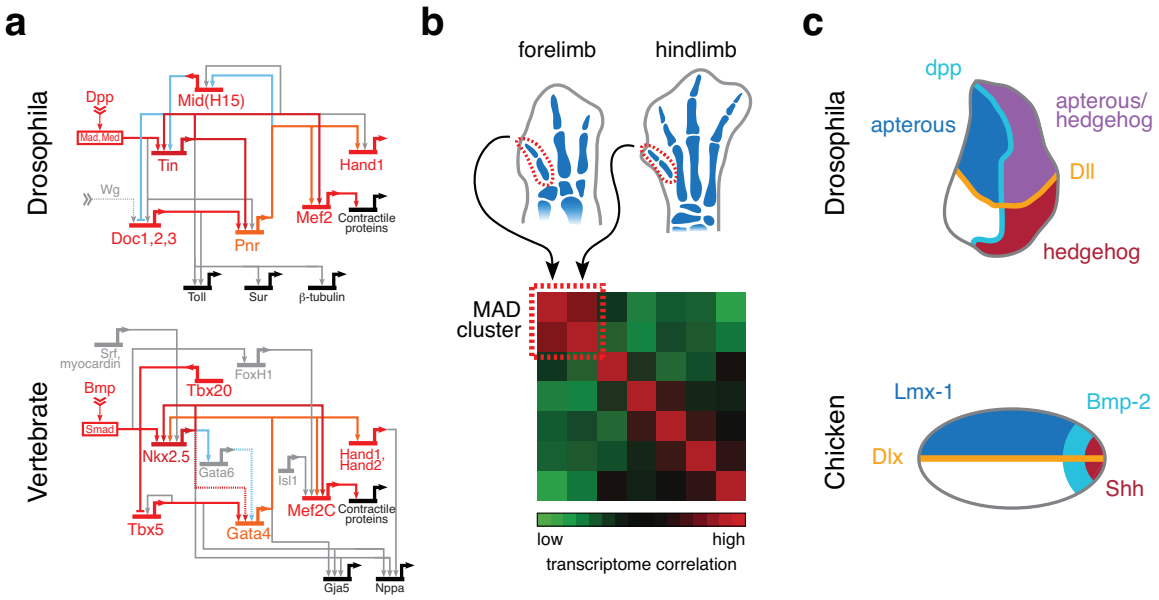
Figure captions

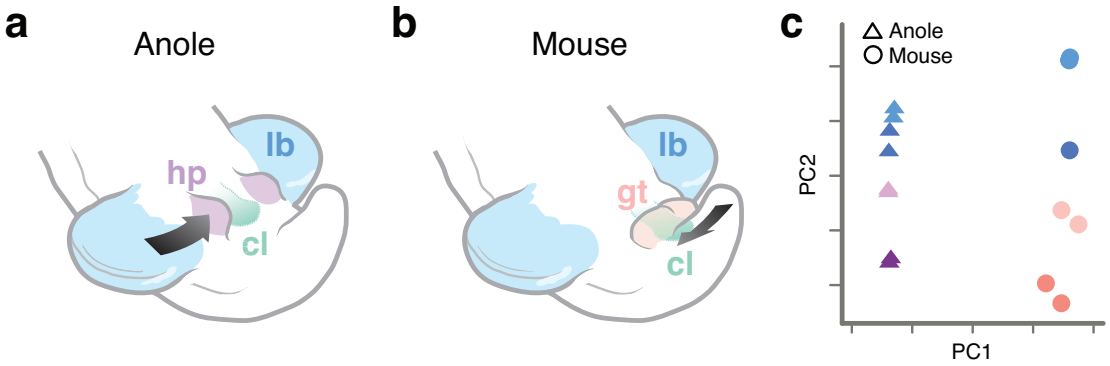
Figure 1

Gene expression-based homology assessment. a) The kernel concept. Deep evolutionary conservation of the core gene regulatory network underlying bilaterian heart development. Parts of the gene regulatory diagrams for *Drosophila* (top) and mouse (bottom) heart development are depicted. Genes and regulatory interactions conserved between the two distantly related species, i.e. the kernel, are highlighted in red. Interactions conserved *via* intermediate relays are highlighted in blue. Modified after [20,23]. b) The character identity network concept. During their development, the most anterior digits (MAD) of both chicken fore- and hindlimbs cluster together based on high transcriptome correlation, identifying them as homologous with regards to their underlying gene regulatory signature. Modified after [28]. c) The deep homology concept. Molecular patterning of developing bilaterian body appendages, in the *Drosophila* wing disc (top) and the chicken limb bud (bottom). Although these structures do not share any historical homology, they have their embryonic axes defined by remarkably conserved genetic circuitries, and hence display deep homology. Modified after [32].

Figure 2

Deep homology of amniote external genitalia. a) Stylized scheme of external genitalia development in a squamate, the anole lizard. For hemipenis bud initiation, the cloacal signaling center recruits cells with a hindlimb-like developmental origin (black arrow). b) In mammals, by contrast, the cloaca is positioned closer to the posterior end, and attracts cells from the tailbud for genital tubercle outgrowth (black arrow). c) Representative principal component analysis (PCA) depiction of limb bud and genitalia transcriptomes in anole lizard and mouse. While PC1 is dominated by a species signal, PC2 reveals tissue-type similarities between the two species, anoles (triangles) and mice (circles). Anole limb and genitalia transcriptomes show a higher degree of relatedness than in mouse, due to a shared embryonic origin. The development of both lizard and mouse genitalia, however, still shows high PC2 tissue-type similarity, and hence displays deep homology at the transcriptome level. lb, hindlimb; hp, hemipenis; gt, genital tubercle; cl, cloaca. Modified after [51].


a
Drosophila
Dpp
Mad,Med
Mid(H15)
Tin
Hand1
Wg
Doc1,2,3
Pnr
Mef2
Contractile proteins
Toll
Sur
β-tubulin
Vertebrate
Srf, myocardin
FoxH1
Bmp
Tbx20
Smad
Nkx2.5
Gata6
Isl1
Hand1, Hand2
Mef2C
Contractile proteins
Tbx5
Gata4
Gja5
Nppa
b
forelimb
hindlimb
MAD cluster
low
high
transcriptome correlation
c
Drosophila
dpp
apterous
apterous/ hedgehog
Dll
hedgehog
Chicken
Lmx-1
Bmp-2
Dlx
Shh


a
Anole
lb
hp
cl
b
Mouse
lb
gt
cl
c
Anole
Mouse
PC2
PC1

References

1. Brigandt, I. 2003 Homology in comparative, molecular, and evolutionary developmental biology: the radiation of a concept. *J. Exp. Zool.* **299**, 9–17. (doi:10.1002/jez.b.36)

2. Roux, J. & Robinson-Rechavi, M. 2010 An ontology to clarify homology-related concepts. *Trends Genet.* **26**, 99–102. (doi:10.1016/j.tig.2009.12.012)

3. Roth, V. L. 1991 Homology and hierarchies: Problems solved and unresolved. *Journal of Evolutionary Biology* **4**, 167–194. (doi:10.1046/j.1420-9101.1991.4020167.x)

4. Wake, D. B. 1994 Comparative Terminology. *Science* **265**, 268–269. (doi:10.1126/science.265.5169.268)

5. Butler, A. B. & Saidel, W. M. 2000 Defining sameness: historical, biological, and generative homology. *Bioessays* **22**, 846–853. (doi:10.1002/1521-1878(200009)22:9<846::AID-BIES10>3.0.CO;2-R)

6. Patterson, C. 1988 Homology in classical and molecular biology. *Molecular Biology and Evolution* **5**, 603–625.

7. Arendt, D. 2008 The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet* **9**, 868–882. (doi:10.1038/nrg2416)

8. Owen, R. 1843 *Lectures on Invertebrate Animals*.

9. Lankester, E. R. 1870 *On the Use of the Term Homology in Modern Zoology, and the Distinction between Homogenetic and Homoplastic Agreements*. *The Annals and Magazine of Natural History, Zoology, Botany, and Geology* **6**, 34–43.

10. Goodwin, B. C. 1982 Development and evolution. *J. Theor. Biol.* **97**, 43–55. (doi:10.1016/0022-5193(82)90275-2)

11. Mindell, D. P. & Meyer, A. 2001 Homology evolving. *Trends in Ecology & Evolution* **16**, 434–440. (doi:10.1016/S0169-5347(01)02206-6)

12. Arthur, W. 2002 The emerging conceptual framework of evolutionary developmental biology. *Nature* **415**, 757–764. (doi:10.1038/415757a)

13. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. 1984

A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**, 403–408.

14. Krumlauf, R. 1994 Hox Genes in Vertebrate Development Review. *Cell* **78**, 191–201.

15. Wagner, G. 1989 The Biological Homology Concept. *Annual Review of Ecology and Systematics* **20**, 51–69. (doi:10.1146/annurev.ecolsys.20.1.51)

16. Müller, G. B. 2007 Evo-devo: extending the evolutionary synthesis. *Nat Rev Genet* **8**, 943–949. (doi:10.1038/nrg2219)

17. Roth, V. L. 1984 On homology. *Biological Journal of the Linnean Society* **22**, 13–29. (doi:10.1111/j.1095-8312.1984.tb00796.x)

18. Nachman, M. W., Hoekstra, H. E. & D'Agostino, S. L. 2003 The genetic basis of adaptive melanism in pocket mice. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5268–5273. (doi:10.1073/pnas.0431157100)

19. Wittkopp, P. J., Carroll, S. B. & Kopp, A. 2003 Evolution in black and white: genetic control of pigment patterns in Drosophila. *Trends Genet.* **19**, 495–504. (doi:10.1016/S0168-9525(03)00194-X)

20. Davidson, E. H. & Erwin, D. H. 2006 Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796–800. (doi:10.1126/science.1113832)

21. Davidson, E. H. et al. 2002 A genomic regulatory network for development. *Science* **295**, 1669–1678. (doi:10.1126/science.1069883)

22. Reim, I. & Frasch, M. 2005 The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in Drosophila. *Development* **132**, 4911–4925. (doi:10.1242/dev.02077)

23. Bruneau, B. G. 2013 Signaling and Transcriptional Networks in Heart Development and Regeneration. *Cold Spring Harbor Perspectives in Biology* **5**, a008292–a008292. (doi:10.1101/cshperspect.a008292)

24. Wagner, G. P. 2007 The developmental genetics of homology. *Nat Rev Genet* **8**, 473–479. (doi:10.1038/nrg2099)

25. Dassow, von, G. & Munro, E. 1999 Modularity in animal development and evolution: Elements of a conceptual framework for EvoDevo. *Journal of Experimental Zoology* **285**, 307–325. (doi:10.1002/(SICI)1097-010X(19991215)285:4<307::AID-JEZ2>3.0.CO;2-V)

26. Burke, A. C. & Feduccia, A. 1997 Developmental Patterns and the Identification of Homologies in the Avian Hand. *Science* **278**, 666–668. (doi:10.1126/science.278.5338.666)

27. Vargas, A. O. & Fallon, J. F. 2005 The digits of the wing of birds are 1, 2, and 3. a review. *J. Exp. Zool.* **304B**, 206–219. (doi:10.1002/jez.b.21051)

28. Wang, Z., Young, R. L., Xue, H. & Wagner, G. P. 2011 Transcriptomic analysis of avian digits reveals conserved and derived digit identities in birds. *Nature* **477**, 583–586. (doi:10.1038/nature10391)

29. Young, R. L. & Wagner, G. P. 2011 Why ontogenetic homology criteria can be misleading: lessons from digit identity transformations. *J. Exp. Zool.* **316B**, 165–170. (doi:10.1002/jez.b.21396)

30. Salinas-Saavedra, M., Gonzalez-Cabrera, C., Ossa-Fuentes, L., Botelho, J. F., Ruiz-Flores, M. & Vargas, A. O. 2014 New developmental evidence supports a homeotic frameshift of digit identity in the evolution of the bird wing. *Frontiers in Zoology* **11**, 33. (doi:10.1186/1742-9994-11-33)

31. Kin, K., Nnamani, M. C., Lynch, V. J., Michaelides, E. & Wagner, G. P. 2015 Cell-type Phylogenetics and the Origin of Endometrial Stromal Cells. *CellReports* **10**, 1398–1409. (doi:10.1016/j.celrep.2015.01.062)

32. Shubin, N., Tabin, C. & Carroll, S. 1997 Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648. (doi:10.1038/41710)

33. Shubin, N., Tabin, C. & Carroll, S. 2009 Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818–823. (doi:10.1038/nature07891)

34. Panganiban, G. et al. 1997 The origin and evolution of animal appendages. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5162–5166.

35. Lu, Y., Huggins, P. & Bar-Joseph, Z. 2009 Cross species analysis of microarray expression data. *Bioinformatics* **25**, 1476–1483. (doi:10.1093/bioinformatics/btp247)

36. Gilad, Y., Rifkin, S. A., Bertone, P., Gerstein, M. & White, K. P. 2005 Multi-species microarrays reveal the effect of sequence divergence on gene expression profiles. *Genome Research* **15**, 674–680. (doi:10.1101/gr.3335705)

37. Metzker, M. L. 2010 Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31–46. (doi:10.1038/nrg2626)

38. Ozsolak, F. & Milos, P. M. 2011 RNA sequencing: advances, challenges

and opportunities. *Nat Rev Genet* **12**, 87–98. (doi:10.1038/nrg2934)

39. Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. & Gilad, Y. 2008 RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research* **18**, 1509–1517. (doi:10.1101/gr.079558.108)

40. Cook, C. E., Chenevert, J., Larsson, T. A., Arendt, D., Houliston, E. & Lénárt, P. 2016 Old knowledge and new technologies allow rapid development of model organisms. *Mol. Biol. Cell* **27**, 882–887. (doi:10.1091/mbc.E15-10-0682)

41. Haas, B. J. et al. 2013 De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**, 1494–1512. (doi:10.1038/nprot.2013.084)

42. Romero, I. G., Ruvinsky, I. & Gilad, Y. 2012 Comparative studies of gene expression and the evolution of gene regulation. *Nat Rev Genet* **13**, 505–516. (doi:10.1038/nrg3229)

43. Assis, R., Zhou, Q. & Bachtrog, D. 2012 Sex-biased transcriptome evolution in Drosophila. *Genome Biol Evol* **4**, 1189–1200. (doi:10.1093/gbe/evs093)

44. Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Baker, J. C., Grützner, F. & Kaessmann, H. 2014 The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**, 635–640. (doi:10.1038/nature12943)

45. Levin, M. et al. 2016 The mid-developmental transition and the evolution of animal body plans. *Nature*, 1–18. (doi:10.1038/nature16994)

46. Duboule, D. 2010 The evo-devo comet. *EMBO Rep* **11**, 489–489. (doi:10.1038/embor.2010.94)

47. Jaeger, J., Laubichler, M. & Callebaut, W. 2015 The Comet Cometh: Evolving Developmental Systems. *Biol Theory* **10**, 36–49. (doi:10.1007/s13752-015-0203-5)

48. Montavon, T., Le Garrec, J. F., Kerszberg, M. & Duboule, D. 2008 Modeling Hox gene regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes & Development* **22**, 346–359. (doi:10.1101/gad.1631708)

49. de Navas, L. F., Reed, H., Akam, M., Barrio, R., Alonso, C. R. & Sánchez-Herrero, E. 2011 Integration of RNA processing and expression level

control modulates the function of the Drosophila Hox gene Ultrabithorax during adult development. *Development* **138**, 107–116. (doi:10.1242/dev.051409)

50. Brawand, D. et al. 2011 The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348. (doi:10.1038/nature10532)

51. Tschopp, P., Sherratt, E., Sanger, T. J., Groner, A. C., Aspiras, A. C., Hu, J. K., Pourquié, O., Gros, J. & Tabin, C. J. 2014 A relative shift in cloacal location repositions external genitalia in amniote evolution. *Nature* **516**, 391–394. (doi:10.1038/nature13819)

52. Musser, J. M. & Wagner, G. P. 2015 Character trees from transcriptome data: Origin and individuation of morphological characters and the so-called 'species signal'. *J. Exp. Zool.*, n/a–n/a. (doi:10.1002/jez.b.22636)

53. Roux, J., Rosikiewicz, M. & Robinson-Rechavi, M. 2015 What to compare and how: Comparative transcriptomics for Evo-Devo. *J. Exp. Zool.* **324**, 372–382. (doi:10.1002/jez.b.22618)

54. Anavy, L., Levin, M., Khair, S., Nakanishi, N., Fernandez-Valverde, S. L., Degnan, B. M. & Yanai, I. 2014 BLIND ordering of large-scale transcriptomic developmental timecourses. *Development* **141**, 1161–1166. (doi:10.1242/dev.105288)

55. Lin, C., Yin, Y., Bell, S. M., Veith, G. M., Chen, H., Huh, S.-H., Ornitz, D. M. & Ma, L. 2013 Delineating a conserved genetic cassette promoting outgrowth of body appendages. *PLoS Genet.* **9**, e1003231. (doi:10.1371/journal.pgen.1003231)

56. Yamada, G. et al. 2006 Molecular genetic cascades for external genitalia formation: An emerging organogenesis program. *Dev. Dyn.* **235**, 1738–1752. (doi:10.1002/dvdy.20807)

57. Infante, C. R., Mihala, A. G., Park, S., Wang, J. S., Johnson, K. K., Lauderdale, J. D. & Menke, D. B. 2015 Shared Enhancer Activity in the Limbs and Phallus and Functional Divergence of a Limb-Genital cis-Regulatory Element in Snakes. *Developmental Cell* **35**, 107–119. (doi:10.1016/j.devcel.2015.09.003)

58. Glassford, W. J., Johnson, W. C., Dall, N. R., Smith, S. J., Liu, Y., Boll, W., Noll, M. & Rebeiz, M. 2015 Co-option of an Ancestral Hox-Regulated Network Underlies a Recently Evolved Morphological Novelty. *Developmental Cell* **34**, 520–531. (doi:10.1016/j.devcel.2015.08.005)

59. Hosken, D. J. & Stockley, P. 2004 Sexual selection and genital evolution.

Trends in Ecology & Evolution **19**, 87–93. (doi:10.1016/j.tree.2003.11.012)

60. Eberhard, W. G. 2010 Evolution of genitalia: theories, evidence, and new directions. *Genetica* **138**, 5–18. (doi:10.1007/s10709-009-9358-y)

61. Necsulea, A. & Kaessmann, H. 2014 Evolutionary dynamics of coding and non-coding transcriptomes. *Nat Rev Genet* **15**, 734–748. (doi:10.1038/nrg3802)

62. Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N. & Hughes, K. A. 2015 Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* (doi:10.1038/nature15256)

63. Woltering, J. M. & Duboule, D. 2010 The Origin of Digits: Expression Patterns versus Regulatory Mechanisms. *Developmental Cell* **18**, 526–532. (doi:10.1016/j.devcel.2010.04.002)

64. Harmston, N. & Lenhard, B. 2013 Chromatin and epigenetic features of long-range gene regulation. *Nucleic Acids Research* **41**, 7185–7199. (doi:10.1093/nar/gkt499)

65. Nora, E. P., Dekker, J. & Heard, E. 2013 Segmental folding of chromosomes: A basis for structural and regulatory chromosomal neighborhoods? *Bioessays* **35**, 818–828. (doi:10.1002/bies.201300040)

66. Levine, M. 2010 Transcriptional Enhancers in Animal Development and Evolution. *Current Biology* **20**, R754–R763. (doi:10.1016/j.cub.2010.06.070)

67. Cotney, J., Leng, J., Yin, J., Reilly, S. K., DeMare, L. E., Emera, D., Ayoub, A. E., Rakic, P. & Noonan, J. P. 2013 The Evolution of Lineage-Specific Regulatory Activitiesin the Human Embryonic Limb. *Cell* **154**, 185–196. (doi:10.1016/j.cell.2013.05.056)

68. Reilly, S. K., Yin, J., Ayoub, A. E., Emera, D., Leng, J., Cotney, J., Sarro, R., Rakic, P. & Noonan, J. P. 2015 Evolutionary changes in promoter and enhancer activity during human corticogenesis. *Science* **347**, 1155–1159. (doi:10.1126/science.1260943)

69. Villar, D. et al. 2015 Enhancer evolution across 20 mammalian species. *Cell* **160**, 554–566. (doi:10.1016/j.cell.2015.01.006)

70. Spitz, F. & Furlong, E. E. M. 2012 Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613–626. (doi:10.1038/nrg3207)

71. Villar, D., Flicek, P. & Odom, D. T. 2014 Evolution of transcription factor binding in metazoans - mechanisms and functional implications. *Nat Rev Genet* **15**, 221–233. (doi:10.1038/nrg3481)

72. John, S. et al. 2013 Genome-scale mapping of DNase I hypersensitivity. *Curr Protoc Mol Biol* **Chapter 27**, Unit 21.27–21.27.20. (doi:10.1002/0471142727.mb2127s103)

73. Giresi, P. G., Kim, J., McDaniell, R. M., Iyer, V. R. & Lieb, J. D. 2007 FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Research* **17**, 877–885. (doi:10.1101/gr.5533506)

74. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. 2013 Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Meth* **10**, 1213–1218. (doi:10.1038/nmeth.2688)

75. Gehrke, A. R. et al. 2015 Deep conservation of wrist and digit enhancers in fish. *Proceedings of the National Academy of Sciences* **112**, 803–808. (doi:10.1073/pnas.1420208112)

76. Vierstra, J. et al. 2014 Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. *Science* **346**, 1007–1012. (doi:10.1126/science.1246426)

77. Remeseiro, S., Hörnblad, A. & Spitz, F. 2016 Gene regulation during development in the light of topologically associating domains. *Wiley Interdiscip Rev Dev Biol* **5**, 169–185. (doi:10.1002/wdev.218)

78. Lupiáñez, D. G. et al. 2015 Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* (doi:10.1016/j.cell.2015.04.004)

79. Tsujimura, T., Klein, F. A., Langenfeld, K., Glaser, J., Huber, W. & Spitz, F. 2015 A discrete transition zone organizes the topological and regulatory autonomy of the adjacent tfap2c and bmp7 genes. *PLoS Genet.* **11**, e1004897. (doi:10.1371/journal.pgen.1004897)

80. Gómez-Marín, C. et al. 2015 Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proceedings of the National Academy of Sciences* **112**, 7542–7547. (doi:10.1073/pnas.1505463112)

81. Acemel, R. D. et al. 2016 A single three-dimensional chromatin

compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nature Genetics* **48**, 336–341. (doi:10.1038/ng.3497)

82. Woltering, J. M., Noordermeer, D., Leleu, M. & Duboule, D. 2014 Conservation and divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. *PLoS Biol* **12**, e1001773. (doi:10.1371/journal.pbio.1001773)

83. Parker, H. J., Bronner, M. E. & Krumlauf, R. 2014 A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature* **514**, 490–493. (doi:10.1038/nature13723)

84. Yao, Y. et al. 2016 Cis-regulatory architecture of a brain signaling center predates the origin of chordates. *Nature Genetics* (doi:10.1038/ng.3542)

85. Wagner, G. P. 2012 Next Gen Devo-Evo. *J. Exp. Zool.* **318**, 519–520. (doi:10.1002/jez.b.22463)

86. Soyer, O. S. & O'Malley, M. A. 2013 Evolutionary systems biology: What it is and why it matters. *Bioessays* **35**, 696–705. (doi:10.1002/bies.201300029)

87. Saliba, A.-E., Westermann, A. J., Gorski, S. A. & Vogel, J. 2014 Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Research* (doi:10.1093/nar/gku555)

88. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. 2015 Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, 495–502. (doi:10.1038/nbt.3192)

89. Achim, K., Pettit, J.-B., Saraiva, L. R., Gavriouchkina, D., Larsson, T., Arendt, D. & Marioni, J. C. 2015 High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol* **33**, 503–509. (doi:10.1038/nbt.3209)

90. Sánchez Alvarado, A. & Yamanaka, S. 2014 Rethinking differentiation: stem cells, regeneration, and plasticity. *Cell* **157**, 110–119. (doi:10.1016/j.cell.2014.02.041)

91. Sato, T. & Clevers, H. 2015 SnapShot: Growing Organoids from Stem Cells. *Cell* **161**, 1700–1700.e1. (doi:10.1016/j.cell.2015.06.028)

92. Arnold, C. D., Gerlach, D., Spies, D., Matts, J. A., Sytnikova, Y. A., Pagani, M., Lau, N. C. & Stark, A. 2014 Quantitative genome-wide enhancer activity maps for five Drosophila species show functional

enhancer conservation and turnover during cis-regulatory evolution. *Nature Genetics* **46**, 685–692. (doi:10.1038/ng.3009)

93. Prescott, S. L., Srinivasan, R., Marchetto, M. C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F. H., Swigut, T. & Wysocka, J. 2015 Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83. (doi:10.1016/j.cell.2015.08.036)

94. Gerstein, M. B. et al. 2014 Comparative analysis of the transcriptome across distant species. *Nature* **512**, 445–448. (doi:10.1038/nature13424)

95. Ho, J. W. K. et al. 2014 Comparative analysis of metazoan chromatin organization. *Nature* **512**, 449–452. (doi:10.1038/nature13415)

96. Stergachis, A. B. et al. 2014 Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* **515**, 365–370. (doi:10.1038/nature13972)

97. Rung, J. & Brazma, A. 2013 Reuse of public genome-wide gene expression data. *Nat Rev Genet* **14**, 89–99. (doi:10.1038/nrg3394)

98. Bastian, F., Parmentier, G., Roux, J., Moretti, S., Laudet, V. & Robinson-Rechavi, M. 2008 Bgee: Integrating and Comparing Heterogeneous Transcriptome Data Among Species. In *Data Integration in the Life Sciences*, pp. 124–131. Berlin, Heidelberg: Springer Berlin Heidelberg. (doi:10.1007/978-3-540-69828-9_12)

99. Piasecka, B., Kutalik, Z., Roux, J., Bergmann, S. & Robinson-Rechavi, M. 2012 Comparative modular analysis of gene expression in vertebrate organs. *BMC Genomics* **13**, 124. (doi:10.1186/1471-2164-13-124)

100. Thompson, D., Regev, A. & Roy, S. 2015 Comparative analysis of gene regulatory networks: from network reconstruction to evolution. *Annu. Rev. Cell Dev. Biol.* **31**, 399–428. (doi:10.1146/annurev-cellbio-100913-012908)

101. Nunes, M. D. S., Arif, S., Schlötterer, C. & McGregor, A. P. 2013 A perspective on micro-evo-devo: progress and potential. *Genetics* **195**, 625–634. (doi:10.1534/genetics.113.156463)

102. Crombach, A. & Hogeweg, P. 2008 Evolution of Evolvability in Gene Regulatory Networks. *PLoS Comput Biol* **4**, e1000112. (doi:10.1371/journal.pcbi.1000112)