BioEssays
www.bioessays-journal.com

Coupled Genomic Evolutionary Histories as Signatures of Organismal Innovations in Cephalopods

Co-evolutionary Signatures Across Levels of Genome Organization May Shed Light on Functional Linkage and Origin of Cephalopod Novelties

*Elena A. Ritschard, Brooke Whitelaw, Caroline B. Albertin, Ira R. Cooke, Jan M. Strugnell, and Oleg Simakov**

How genomic innovation translates into organismal organization remains largely unanswered. Possessing the largest invertebrate nervous system, in conjunction with many species-specific organs, coleoid cephalopods (octopuses, squids, cuttlefishes) provide exciting model systems to investigate how organismal novelties evolve. However, dissecting these processes requires novel approaches that enable deeper interrogation of genome evolution. Here, the existence of specific sets of genomic co-evolutionary signatures between expanded gene families, genome reorganization, and novel genes is posited. It is reasoned that their co-evolution has contributed to the complex organization of cephalopod nervous systems and the emergence of ecologically unique organs. In the course of reviewing this field, how the first cephalopod genomic studies have begun to shed light on the molecular underpinnings of morphological novelty is illustrated and their impact on directing future research is described. It is argued that the application and evolutionary profiling of evolutionary signatures from these studies will help identify and dissect the organismal principles of cephalopod innovations. By providing specific examples, the implications of this approach both within and beyond cephalopod biology are discussed.

1. Introduction

Evolutionary novelty is associated with a wide range of genetic mechanisms. Recent studies have begun to uncover the role of various levels of genome organization in the evolution of phenotypic novelties.[1–3] These include, but are not limited to, novel gene emergence (i.e., taxonomically restricted genes), gene duplications, and genome rearrangements.[4–6] Each play a role in the modification and extension of the existing regulatory landscapes that define physiological systems, organs, tissues, and cell types [7–9] (**Box 1**).

Often, however, it remains difficult to associate specific genetic or genomic changes with their morphological and functional outcomes. This problem becomes increasingly more difficult for more complex systems, such as the genetic basis of vertebrate brain development and organization[10] or that of other complex nervous systems, such as those found in cephalopod molluscs.[11] While further functional data (e.g., gene knock-down/out) is absolutely necessary in dissecting such systems, novel approaches to comparative genomics are also required to quantify and reveal the key contributors among the various genomic characters to individual morphological or developmental innovations.

Through ongoing technological progress and the availability of many genomic resources, such an approach is now becoming

E. A. Ritschard, Prof. O. Simakov
Department for Molecular Evolution and Development
University of Vienna
Austria
E-mail: oleg.simakov@univie.ac.at

B. Whitelaw, Prof. J. M. Strugnell
Centre for Sustainable Tropical Fisheries and Aquaculture
College of Science and Engineering
James Cook University
Townsville, Queensland 4811, Australia

Dr. C. B. Albertin
Marine Biological Laboratory
Woods Hole, MA 02543, USA

Dr. I. R. Cooke
Department of Molecular and Cell Biology
James Cook University
Townsville, Queensland 4811, Australia

Prof. J. M. Strugnell
Department of Ecology
Environment and Evolution
La Trobe University
Melbourne, Victoria 3086, Australia

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/bies.201900073

DOI: 10.1002/bies.201900073

Box 1
Levels of genome organization in metazoans and their regulatory context

Complex transitions in body plan organization have defined animal evolution (e.g., the origin of multicellularity, the appearance of bilateral symmetry accompanied by the development of a through-gut, the evolution of centralized nervous systems).[6] Increasing evidence from comparative genomic studies (e.g., refs. [30,32,105–109] has shifted the traditional paradigm that these and other organismal transitions were related to a "step-wise" increase in genome complexity toward the hypothesis of an already complex ur-metazoan ancestor). For instance, studies found evidence for conserved gene order (also termed synteny), protein domain architecture, and gene structure (e.g., exon–intron composition) in early branching lineages,[105,106,110] suggesting these features were present in the metazoan ancestor. Crucial cell signaling and developmental gene family complements (e.g., Wnts, TGFbeta, Hedgehog, Hox, and Sox) are highly conserved in most metazoans, including early branching lineages like sponges[105] and cnidarians.[106] Because of their absence outside Metazoa, these have been suggested to be "true" animal gene novelties.[6,109] Altogether, these ancestral genomic features have largely contributed to the evolution of modern metazoans as building blocks that were modified in different metazoan lineages. Such modifications can be expansions of pre-existing gene repertoires via gene duplication (including whole genome duplications as in the vertebrate lineage) or rearrangements of the ancestral synteny blocks (order of genes). Novel genetic material can also arise by novel/orphan gene formation (i.e., genes with no sequence homology outside of a taxon). These evolutionary events are what we here refer to as levels of genome organization, and each contributes differently to the evolution of individual metazoan lineages.

An important aspect of the evolution of these levels of genome organization is gene regulation. Changes in protein coding regions in the genome, from single nucleotide substitutions to gene duplications, were thought to be major drivers of evolution and thus the focus of most studies. While this is true for the evolution of many physiological processes, increasing importance is now given to regulation of gene expression as an evolutionary force.[1] However, we do not consider it here as an independent force but rather a "groundfloor" in which the different levels of genome organization change, interact with each other, and consequently contribute to evolutionary transitions.

- **Gene duplication**: duplication of regulatory sequences and/or their target genes can generate new interactions in a regulatory network.[111] Following duplication, copies may share at the beginning same regulatory properties, and this redundancy will permit their subsequent divergence without altering the overall ancestral function. Such divergence can happen either by dividing the ancestral repertoire of regulatory interactions between the two copies (i.e., sub-functionalization) or by the acquirement of novel regulatory pathways by one of the paralogues (i.e., neo-functionalization). Moreover, the extent to which paralogues contribute to regulatory network evolution also depends on the duplication dynamics. For example, tandem duplication may contribute to the maintenance of the copies in close proximity to each other and thus ancestral interactions, favoring sub- over neo-functionalization. Additional genome architecture modifications may result in translocation of paralogues (e.g., by genome rearrangements), establishing new regulatory interactions that were not present in the ancestral gene because of new "territory colonization." This can favor neo-functionalization.
- **Novel gene emergence**: as in the case of duplicated genes evolving via neo-functionalization, novel genes may contribute to the generation of new regulatory interactions. Their contribution to regulatory networks will highly depend on the architectural landscape (i.e., composition of regulatory loops and promotors/enhancers found therein) of the chromosomal region where these genes originate.
- **Genome rearrangements**: maintenance of gene order is important in the context of architectural and functional looping of the genome. These loops, also termed topologically associated domains (TADs), bring regulatory elements and their targets into close proximity.[9] Therefore, genome rearrangements impose a high impact on generating novel regulatory interactions as they may break these existing regulatory loops.

feasible. Sampling gaps are starting to be filled due to advances in high throughput sequencing methods and development and application of functional tools for many non-model organisms (e.g., ref. [12]). The existing genomic and functional data for a few key metazoan species have revealed a complex picture of both coding and noncoding region evolution.[6] We are now beginning to understand how different scales of genome organization (Box 1) have contributed to the evolution of various animal groups.[6] We refer to those combinations of genomic characters as modes of genome evolution and suggest that they are a result of complex selective pressures on the regulatory architecture of the genome as a whole.

Given such improved understanding of genome evolution across metazoans, it is becoming necessary to (re)consider the organismal context of genomic novelties. In particular, genomic novelties affecting a single organ or cell type are likely to play a role in the evolution of functionally linked tissues and/or structures,[13,14] a concept generally referred to as co-evolution. In the original definition, co-evolution is implicated when, for example, the phylogenies of host and symbiont taxa are congruent, a phenomenon known as co-divergence (e.g., refs. [15,16]). However, this term has been also applied to study protein interaction evolution.[17–19] At the simplest level this involves, and can be measured as, co-evolution between genes of different gene

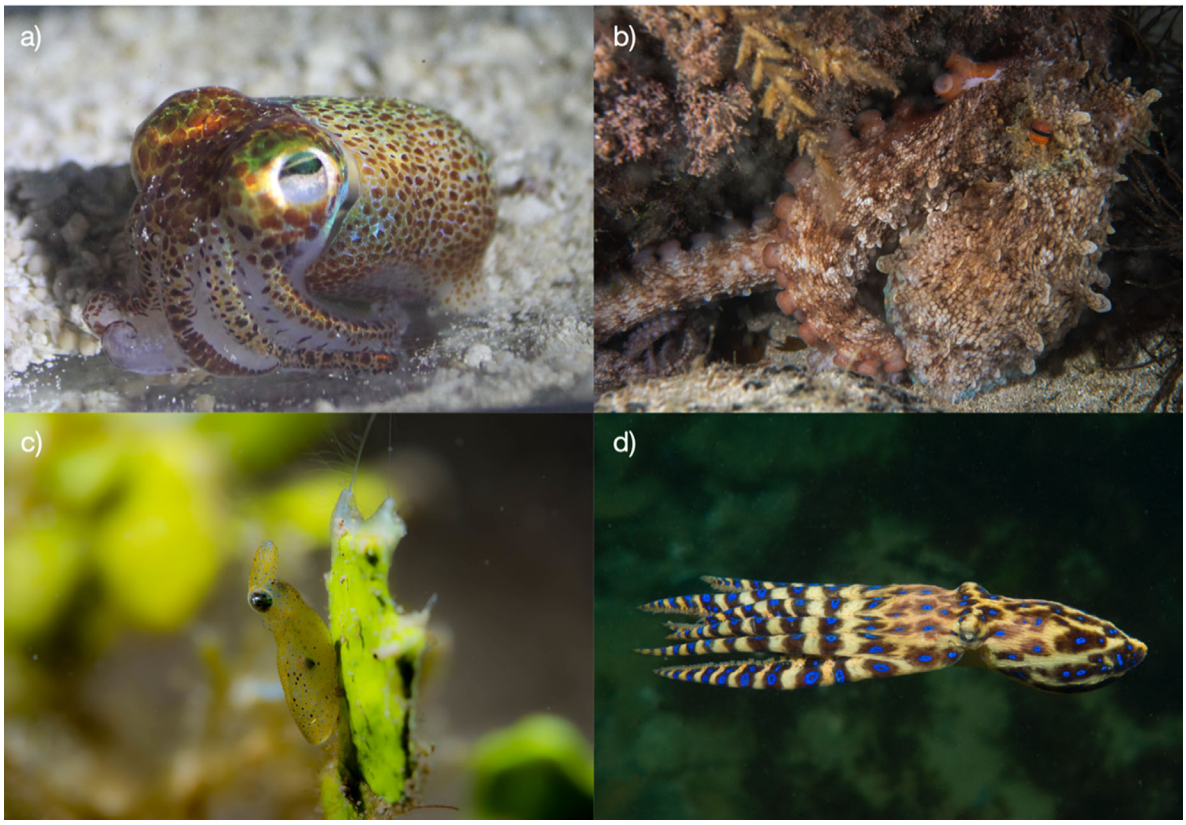

Figure 1. Cephalopod diversity and examples of unique adaptations. a) Symbiotic light organ in bobtail squid (*Euprymna scolopes*), image contribution: Sarah McAnulty, b) dynamic skin patterning for camouflaging in octopus (*Octopus tetricus*), image contribution: Julian Finn, c) adhesive glands in pygmy squid (Family *Idiosepiidae*), image contribution: Rickard Zerpe, and d) toxin-producing salivary glands in blue-ringed octopus (Genus *Hapalochlaena*), image contribution: Julian Finn.

families by, for example, phylogenetic profiling[20] or the mirrortree approach.[21] How this ultimately links to co-evolution at the morphological level, remains unclear. Therefore, a more complete approach across several levels of genomic organization is required. Such an approach should take into account the contribution of all genomic characters (mode of genome evolution), such as the origin of novel coding genes, gene duplication, noncoding elements, genome reorganization, etc. Here we explore how those different modes of genome evolution have contributed to the origin of novelties in the organismal context. We ask whether, and how, the individual genomic evolutionary signatures across and within tissues and organs can show signs of co-evolution, reflecting functional linkage.

Coleoid cephalopods (cuttlefishes, squids, and octopuses) are an ideal model taxon for this endeavor. Cephalopods are considered to possess the most elaborate invertebrate nervous system, a notion that is likely linked to their complex behavioral repertoire, their active predatory mode of life, and their ability to rapidly manipulate skin texture and color for camouflage and display. Additionally, their adaptation to a wide range of marine environments is associated with a suite of morphological novelties that have evolved across the subclass such as light and adhesive organs, accessory nidamental glands, sucker ring teeth, toxin producing salivary glands, among others (**Figure 1**). Cephalopods have thus been the target of several research areas[22,23] such as behavioral biology,[24] phylogeography,[25] development,[26–28] neurobiology,[11,28,29] and, more recently, genomics.[28] They provide a fruitful, yet almost entirely unexplored model system to study the role of genomic innovations shaping morphology, development, and behavior. Investigating genomic signatures behind those adaptations may help reveal fundamental insights into the co-evolution of novel features in the organismal context. This will open up an exciting possibility to explore functional linkage in the evolutionary histories of distinct genetic components. Additionally, the discovery of such signatures could be used to trace back the functionally linked regions within individual tissues or organs, thus helping understand their functionality by identifying key linked genomic components.

2. Which Genomic Changes Drive Organismal Novelty? A Multi-Scale Approach

Many studies have identified and linked genomic and morphological changes in metazoans (see Box 1). However, genome sequencing has failed to provide a pattern or a comprehensive answer with regards to which genetic components drive morphological novelty. So far, most studies have focused on case-by-case studies to identify individual factors of the genetic framework required for a given morphological change. Those studies have identified genetic contributions behind both monogenic and polygenic trait types well suited model systems (e.g., ref. [31]).

The origins of evolutionary novelty in other less studied systems are difficult to reveal due to the often-large genome sizes, lack of understanding of the underlying gene regulation, but also the absence of functional methods to probe it. Moreover, even studies on well-established model systems rarely consider the evolutionary balance between different genomic characters contributing to the organismal novelty. Research has been largely constrained by the methodological approaches and tools available, mostly focusing on individual gene families and/or their dynamics. No study, to our knowledge, considers the combined evolutionary history and the co-evolutionary trajectories of various genomic characters in metazoans.

There is an increasing interest to explore this multi-scale approach as any change occurs in the global context of genome organization, encompassing several levels, from single nucleotide substitutions to gene order changes on chromosomal scales (Box 1). We previously have argued for the existence of several "modes" of genome evolution[6] where certain genomic characters are more likely to change than others depending on species or clade. This pattern has been found for several model species, and similar findings for other non-model metazoan groups can be identified.[32] One of the key emerging observations are the distinct pathways along which metazoan genome architecture can be utilized. For example, global gene loss is common in the *Drosophila* lineage,[33,34] whereas gene redundancy through whole genome duplication has shaped other species groups, most notably vertebrates.[35,36] Such genome expansions or contractions have an ultimate effect on intergenic distances, regulatory element, and gene density and their topological structure. Furthermore, different levels of genome organization may contribute preferentially to certain organs or tissues, that is, some organs may be more prone to evolve in hand with certain types of genomic novelty. For example, taxon-specific traits have been found to be enriched in novel genes (also termed orphan genes or genes with de novo origin, i.e., genes with no sequence similarity outside a species/taxon).[37,38] Because of the functional linkage between organs, tissues, and cells, it is increasingly important to understand how those different genomic changes synergistically contribute to organismal novelty.

Altogether, to begin to link genomic changes to morphology, there is a need to account for different modes of genome evolution (i.e., a multi-scale approach). This will permit us to identify the co-evolutionary trajectories of major genomic changes reflecting functional linkages in the organismal context.

3. Cephalopods: Model Systems for the Study of Genome Evolution Behind Organismal Novelty

Cephalopods (nautiloids, cuttlefish, squid, and octopus) are molluscs with a unique body plan and nervous system. The major extant coleoid lineages, the Octopodiformes and the Decapodiformes, diverged around 270 million years ago (mya).[39] Since then, there have been subsequent radiations of squids and cuttlefishes in the Jurassic and incirrate octopuses in the Cretaceous that have been linked to the radiation of bony fishes.[39] Morphological novelties in cephalopods include camera-type eyes that evolved convergently with those found in vertebrates[40] and flexible arms and tentacles that are thought to be derived from the molluscan foot.[41,42] Together with the partial or complete loss of the molluscan shell, the arm crown has been proposed as a major morphological innovation for their diversification as it likely enabled these animals to become agile predators.[43] Moreover, cephalopods have the largest nervous systems among invertebrates,[11] rivaling vertebrates in terms of size and complexity.[44] In the octopus brain, around a third of the neurons are contained in central lobes surrounding the esophagus and two optic lobes; the remaining two thirds are distributed within the arms, mainly in the axial nerve cord (ANC).[45] Despite this general configuration, there are large differences in the structure of the central nervous systems within coleoid cephalopods, likely a result of species-specific adaptations.[46] Coupled with this neuronal sophistication, cephalopods have also evolved a complex behavioral repertoire,[24] which includes their camouflaging abilities. This is mediated by chromatophore cells in the skin that contain pigment granules. Their expansion and contraction is achieved through radial muscles that, under neural control, allow the animals to rapidly adjust body coloration.[47]

Only recently could genomic studies of cephalopods be accomplished.[30,48–50] Large and repetitive genomes were hindering sequencing and assembly.[30,48] Despite the scarcity of sampling, those studies revealed various levels of genome organization such as the emergence of novel genes, gene family expansions, RNA editing, and extensive genome rearrangements, as possible drivers for the evolution of cephalopod morphological novelties (**Figure 2**). For example, the independently expanded C2H2 zinc fingers and protocadherin gene families in both vertebrates and cephalopods have been argued to have facilitated their neuronal diversity.[30,51,52] Moreover, several cephalopod- and octopus-specific expansions have been observed for G-protein coupled receptors (GPCRs); some of them related to the evolution of the ANC.[53] A recent study of the Hawaiian bobtail squid (*Euprymna scolopes*), has shown that its light organ (LO), a specialized symbiotic structure that houses luminescent bacteria, expresses tandemly duplicated genes (i.e., reflectins).[54] Another symbiotic organ present in many squid species, the accessory nidamental gland (ANG), was found to be enriched in novel genes.[49] Same scenario was found for other tissues, such as suckers of *Octopus bimaculoides*.[30] Furthermore, cephalopods employ epigenetic mechanisms such as extensive RNA editing to generate diversity of their transcripts. RNA editing was found to underlie temperature adaptation in potassium channels in polar octopuses[55] and to be enriched in transcripts expressed in nervous tissues, similar to vertebrates.[56] Finally, genome rearrangements at the coleoid ancestor split conserved gene clusters found across different metazoan levels (e.g., molluscs, lophotrochozoan, bilaterians) possibly providing novel cephalopod-specific gene regulatory domains.[49]

Overall, we are starting to gain insights into the diverse genomic mechanisms underlying morphological novelties in these animals. This now provides the opportunity to study the evolutionary interactions among those various genomic characters and unravel the diversity of cephalopod modes of genome evolution. In the following sections, we will outline hypotheses on how evolution of genomes across those different levels may be functionally linked and reflect organizational principles of cephalopod organs or tissues.

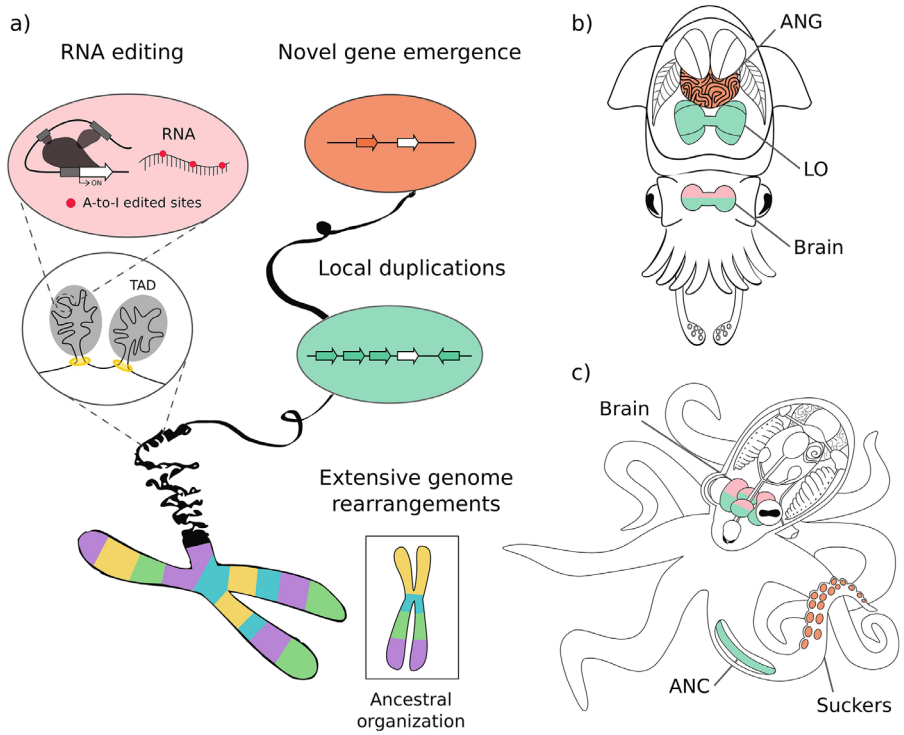

Figure 2. Different levels of genome organization and their contribution to morphological novelties in cephalopods. a) Individual levels of genome organization (i.e., genome rearrangements, local duplications, novel gene emergence, and RNA editing) and their contribution at organismal level (see text) in b) *E. scolopes* (illustration by Hannah Schmidbaur) and c) *O. bimaculoides* are color-coded. ANG: accessory nidamental gland; LO: light organ; ANC: axial nerve cord.

4. Co-Expanding Gene Families Explain the Diversification of Cephalopod Neural Circuits

4.1. Gene Duplications as a Major Driver of Neural Circuit Divergence

Gene duplication as proposed by Ohno[57] promoted the notion of novelty emergence via sub- or neofunctionalization of gene copies (Box 1). Many studies have identified the role of this powerful mechanism during nervous system evolution (e.g., refs. [58–61]). For instance, gene duplication has been proposed to be involved in whole brain pathway duplication during the evolution of complex behaviors, such as vocal learning.[59] Particularly, duplication of transcription factors (e.g., Hox genes) has been hypothesized to result in the formation of parallel circuits. These replicated circuits could ultimately diverge driven by changes in other genes involved in axon guidance[59] and neural connectivity. For instance, membrane receptors, upon duplication, facilitate diversification of synapses (and neural cell identity) via subfunctionalization.[61] One specific example is the extensive gain and losses of odorant receptors (ORs) in mammals, which have resulted in chemosensory repertoires of various sizes in different lineages.[62,63] OR genes are expressed in sensory neurons of the olfactory epithelium following a "one receptor per neuron" pattern.[64] The specific expression of these genes has been shown to determine the transcription of cell adhesive and repulsive proteins (e.g., protocadherins[65]) that control axon fasciculation during the convergence of same-type OR neurons into a glomerulus in the olfactory bulb.[64] Additionally, during early embryogenesis, axonal projections of these neurons is regulated by genes of other highly expanded families in vertebrates, the zinc finger transcription factors.[66] Therefore, changes in OR gene repertoires, together with other expanded gene families, play a role in shaping the olfactory system architecture with ultimate implications for olfactory system adaptations to the environment.[58]

Similarly, in cephalopods, the biggest gene family expansions (i.e., C2H2 zinc fingers and protocadherins) are almost exclusively expressed in nervous tissues.[30] Genes from other expansions, like the GPCRs, are mainly expressed outside of the brain, but in novel or highly derived cephalopod structures such as the ANC, suckers, and skin.[53] Therefore, it can be expected that gene family expansions in cephalopod genomes have played a role in the evolution of their complex nervous system and functionally linked structures. Here, we propose two hypothetical scenarios in which signatures of co-evolution between expanded families can help us to address neural architecture and innovation in cephalopods (**Figure 3**).

4.2. Co-Evolving Adhesive Proteins Explain Neural Wiring in the Axial Nerve Cord

As in vertebrate olfactory neurons,[64,65] adhesive molecules can be involved in neural wiring, forming or avoiding links between particular neurons. In cephalopods, two groups of adhesion proteins have been found to be expanded: protocadherins[30] and

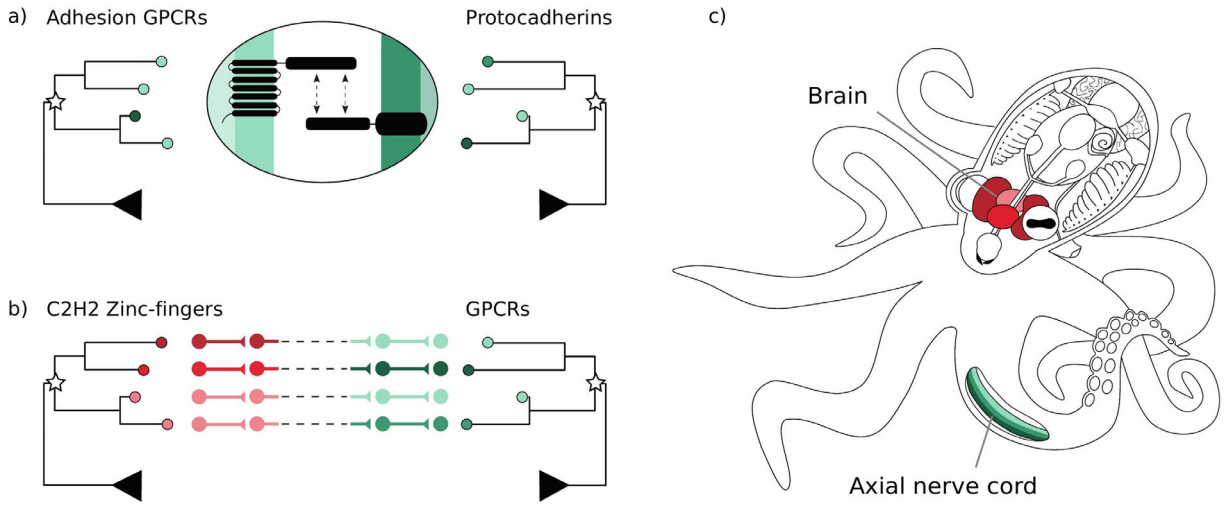

Figure 3. Dissecting cephalopod neuronal complexity through signatures of genomic co-evolution. When comparing the topology of two gene family expansions (e.g., a) Protocadherins and GPCRs and b) C2H2 zinc fingers and GPCRs), a co-evolutionary scenario can be suggested by similar branching patterns, evolutionary rates (i.e., branch lengths), and selection signatures (here represented as stars) of a particular group of duplicated genes. Co-evolution at the genome level can then be related to the co-evolution of cell types and/or structures where the expanded gene families are expressed (brain and the axial nerve cord, as shown in (c)).

adhesion GPCRs.[53] It has been hypothesized that the octopus-specific adhesion GPCR expansion complements the enlarged protocadherin repertoire in the ANC, as both have been found to be highly expressed in that neural tissue.[53] We suggest that this complementation could be either positive (e.g., if both expanded families appear to be co-expressed in same neural cells) or negative (e.g., if the use of one type of adhesion protein restricts the use of the other one and thus contributes to different cell type identities) and is related to the neural wiring in the ANC. Single cell expression data and associated reconstruction of gene regulatory networks or expression modules could help elucidate the nature of this hypothesized complementation. Moreover, comparisons between duplication events of ANC-specific adhesion GPCRs and protocadherins could give insights into the possible scenario of co-evolution of these expanded gene families and their interaction (Figure 3a). For instance, mirror phylogenies of co-evolving genes can be a result of compensatory changes between interacting proteins or mutual involvement in same cellular processes.[17] Therefore, a possible scenario of co-evolution between these two gene family expansions could be explained by compensatory changes after gene duplications because of their involvement in neural identity within the ANC. Additionally, expanded adhesion GPCRs were found to evolve under positive selection via neofunctionalization[53] and if there was co-evolution with a set of expanded protocadherins, we can expect to see the same selection signatures (Figure 3a) using synonymous to non-synonymous ratio tests. Altogether, such an event might have diversified neural wiring and contributed to new functions in the ANC.

4.3. Signatures of Co-Evolution in Gene Duplications Help Dissect Cephalopod Neural Circuit Organization

Co-evolutionary patterns between gene family expansions may furthermore help us to understand what type of network connectivity and signal encoding cephalopods employ. Many animals represent the body in the brain through somatotopy, where regions of the body map to specific regions in the brain, as in the somatosensory cortex in vertebrates. An alternative solution could be a temporally encoded network, where input from more distal regions would accumulate as it reaches the central brain.[67] How signals are encoded in cephalopods is not well understood. While the presence of somatotopic maps has been proposed,[11] there is no evidence for their existence in higher brain centers.[68] We argue that co-evolution signatures of distinct gene families employed in region specification of the central brain (e.g., C2H2 zinc fingers) and neuronal functionality in the arms and distal structures (e.g., GPCRs in suckers, skin, and ANC) may help reveal which scenario is most likely. For example, if somatotopic representation holds, we can expect co-evolution both at the timing of duplication events (as inferred by the tree topology) as well as in the presence of signatures of positive selection at certain phylogenetic tree nodes (Figure 3b). Such scenario could also hint at whether brain pathway duplication is linked to gene duplication as proposed for vocal learning evolution.[59] Absence of such co-evolutionary signatures, on the other hand, may indicate a different, non-somatotopic, form of signal encoding and type of nervous system organization.[11,67,69]

Taken together, co-evolutionary signatures may reveal principles of cephalopod neuronal organization, both at the local cell–cell interaction and neural circuit wiring.

5. Gene Duplications Explain Lineage-Specific Adaptations in Cephalopods

Outside of the brain, co-evolutionary signatures may help reveal molecular players behind the emergence of many other species-specific organs involved in adaptation. For example, venom glands of cephalopods provide for an opportunity to study the effect of toxin evolution through gene duplication and their impact

on the whole organism. Recent studies based on extensive quantitative transcriptomics across many body tissues in venomous and nonvenomous snakes have proposed that ancestral versions of venom genes are widely expressed, including in the venom gland.[70] Based on these findings, the "restriction" model was proposed, where the key event leading to the evolution of toxicity is a mutation in a regulatory region leading to venom-gland-specific expression of one of the gene copies. A related model of "stepwise intermediate nearly neutral evolutionary recruitment" (SINNER)[71] suggests "proto-venom" genes are expressed in oral secretory glands at low levels. A switch to higher levels of gene expression in venom glands is followed by a reduction in gene expression in other tissues driven by toxicity of the tissue itself.[71] This fine scale control and specialization is a prime example of mechanisms involved in venom evolution and potentially required for the recruitment of novel gene copies.[72,73]

Likewise, gene duplications are believed to be behind the evolution of cephalopod venoms. The evolutionary origin of putative toxin proteins in these animals has been attributed to several gene families. One example are serine proteases that show dramatic expansions with almost exclusive expression within the venom gland.[74] Another example are tachykinin and tachykinin-like venoms. Tachykinins are a large family of neuropeptides found throughout the metazoans. They play a role in regulation of vasodilation and muscle contraction. In octopods, the existence of a tachykinin receptor throughout peripheral tissues and the heart[75] suggests that these ancient roles are preserved,[76] but in some species tachykinins have also been recruited as key active components of venom.[77–79] It thus seems likely that the venom tachykinins have arisen through neofunctionalization of existing endogenous tachykinins. However, it has not yet been possible to properly investigate the evolutionary history of these genes in cephalopods due to a lack of genome sequences for species known to produce tachykinin venom proteins. Additionally, there are few transcriptomes of cephalopod venom glands that can provide information on expression within the tissue.[80] Despite this, a dynamic evolutionary picture is emerging, for example, a loss of tachykinin expression in the venom gland has been observed in the southern blue-ringed octopus (*Hapalochlaena maculosa*).[81,82] Due to the inclusion of the potent non-proteinaceous neurotoxin tetrodotoxin (TTX) in *H. maculosa*, proteinaceous neurotoxins are considered to be redundant. Taken together, we expect to find co-evolutionary signatures also between genes expressed in neuronal and nonneuronal structures.

6. Emerging Key Role of Other Genomic Characters Driving Cephalopod Organismal Innovations

While evolution by gene duplication is one of the main mechanisms of genome evolution, it only comprises a single type of genomic innovation. For example, many changes observed in human and primate brain evolution rely on single amino acid changes as well as changes in the regulation of main regulators, such as FoxP2, in combination with gene duplications.[83,84] This illustrates the need for a holistic view of genome evolution in understanding genomic background behind various morphological novelties. Previous studies reveal existence of other levels of genome organization such as novel gene formation and genome rearrangements contributing to cephalopod evolution (Figure 2). Below we discuss their combined contribution to evolutionary innovations in these animals and outline the evidence for the presence of interactions across levels of genome organization that can be informative in unraveling fundamental principles of functional coupling between tissues and organs in cephalopods.

6.1. Genome Reorganization in the Coleoid Ancestor Shifted the Regulatory Context of Cephalopod Innovations

While many studies identified various types of gene novelties, their genomic location was usually not considered due to the fragmented nature of many available genomes. With improving sequencing technologies, we can now study how genes are organized in chromosomes and how their location changes across animal genomes. Recent studies highlight highly local aspects of gene regulation, primarily due to the 3D organization of the genome, which subdivides the genome into regions of loops with high contact density.[85] Together, such loops form topologically associating domains (TADs),[86] which coordinate transcription by bringing the transcription start site and often large enhancer regions with bound transcription factors into closer proximity with one another.[87] Thus, the genomic location and the regulatory context of any novelty is important. For instance, many genes duplicate tandemly, presumably preserving their regulatory regions. This results in higher co-expression, as can be observed for recent GPCR expansions in cephalopods.[53] In contrast, when gene duplicates are inserted through retro-transposition in a more distant region, this will likely lead to a completely different level of expression and potential function (see Box 1).

In this respect, genome reorganization can break the existing regulatory landscape and generate novel domains (**Figure 4**).[88] Using gene order conservation analysis (synteny), recent studies revealed the presence of a large-scale genome reorganization in the ancestral coleoid cephalopod genome.[49] Novel cephalopod gene syntenies seem to express in different tissues without a clear enrichment.[49] Thus, it is yet to be revealed how much the expression of the novel and expanded genes has been affected by this reorganization. Furthermore, the role of genome reorganization after the diversification of the coleoid ancestor remains unknown as we currently lack genomes for many of the more recently diverged representatives. Further comparative genomic and transcriptomic studies are required to decipher the role of genome reorganization in cephalopods.

Another driver behind genomic context formation is genome expansion through transposable elements (TEs). Many studies have shown that transposons can generate additional material for the evolution of enhancer and other regulatory elements, comprising complex regulatory code.[89,90] Cephalopod genomes have been shown to possess an abundance of repeats and the repeat content is strikingly different between the two major lineages (i.e., Octopodiformes and Decapodiformes). While short interspersed nuclear elements (SINEs) seem to dominate octopus genomes, long interspersed nuclear elements (LINEs) contributed to the large expansion of squid genomes.[49] Together, these expansions have likely contributed to the large intergenic distances observed in cephalopod lineages. The most outstanding case is the partial Hox gene cluster in *E. scolopes* with

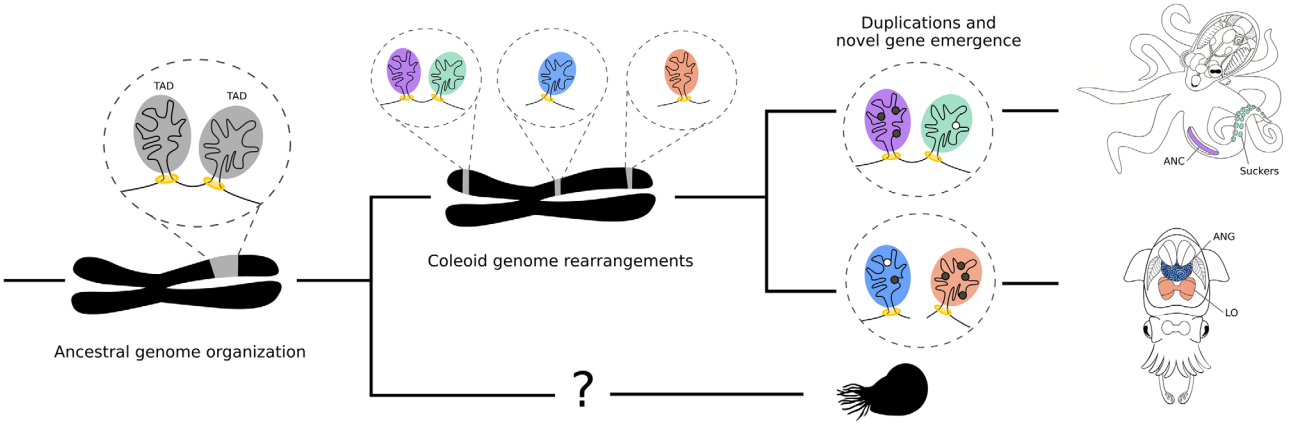

Figure 4. Regulatory landscape transition in cephalopods after extensive genome reorganizations and its impact on other genomic novelties. Genome rearrangements are proposed to dominate early coleoid evolution, where ancestral regulatory architecture (grey) broke and re-shuffled, possibly contributing to the formation of unique cephalopod regulatory landscapes (represented here as purple, green, blue, and orange). Clade-specific evolution was largely driven by gene family and repeat expansions (black circles) and novel gene emergence (white circles) in both squids and octopuses. This rearranged genomic context of novelty likely contributed to the emergence of novel expression domains and/or changes of pre-existing ones (hypothetical expression patterns are shown in squid and octopus, right), providing a "base-line" for other levels of genome organization to interact with each other (e.g., novel genes and gene duplications). ANC: axial nerve cord; ANG: accessory nidamental gland; LO: light organ.

intergenic distances between neighboring Hox genes of around 2Mb,[49] the largest so far reported in the animal kingdom (even compared to much larger genomes, such as axolotl[91]). Further studies are needed to assess their function in respect to the unique Hox expression domains reported for cephalopods.[92]

Overall, genome reorganization represents a major contributor to genome evolution in cephalopods, having a strong impact on the regulatory context of other genomic innovations. While it played a key role in the ancestral diversification of the clade, its role on the evolution of more recently diverged lineages remains to be investigated.

6.2. Novel Gene Emergence during Key Evolutionary Transitions in Cephalopods: Largely Unexplored

Contrary to gene duplications, novel genes were enigmatic, and not considered as a major evolutionary driver until recently.[5,37,93,94] These can arise either de novo from noncoding sequences or through accelerated evolution after a gene duplication event. Functions of novel genes have been studied in a few select animal model species (reviewed in refs. [94–96]) and have been hypothesized to be relevant for taxon-specific adaptations.[38,97] Some genomic studies estimate the high propensity of novel gene formation in animal genomes, yet also show a contribution of significant loss, likely due to selection.[98] While most studies have associated gene duplication with either sub- or neo-functionalized scenarios, evolutionarily fixed novel genes are likely to be associated with neo-functionalization[99] (Box 1). However, it is unclear to what extent novel genes can be incorporated into the already existing networks to facilitate sub-functionalization[98,100] and eventually perform key organismal functions in specific tissues.[101]

In cephalopods, we find distinct signatures of novel gene contribution in certain organs. The expression of reflectins, as one example of a unique cephalopod gene family, is responsible for the iridescence of the skin.[102] Additionally, a symbiotic organ of *E. scolopes*, the ANG, shows a higher expression of novel genes (Figure 2) than other tissues.[49] Unlike in other species, where functional approaches can disentangle the function of those novel genes, very little is known about them in cephalopods. It is not clear whether they contribute to the evolution of new or, rather, act as an addition to the already existing core gene regulatory networks. Their genomic location, however, provides additional insight into the likely mechanism: novel genes expressed in the ANG in *Euprymna* are located on the same scaffold, suggesting co-regulation.[49] While the function of those genes remains to be investigated, their correlated emergence and expression with other types of novelty (e.g., gene duplications) may help establish functional links between and within tissues.

Taken together, emerging genomic data indicates the importance of looking simultaneously across different levels of genome organization, as many of them have been shown to be associated with cephalopod novelties. Their co-evolutionary signatures may help reveal functional connectivity.

7. Toward the Quantification of Modes of Genome Evolution

To be able to further study genomic character evolution in cephalopods, more complete genomes and more sophisticated tools are required to enable deeper comparative and functional analysis. Although the recent genomic studies using cephalopods have shown great insight into the evolution of genomic novelty, currently only a handful of cephalopod genomes have been published. The great majority of them are representatives of Octopodiformes (i.e., *O. bimaculoides*,[30] *Octopus vulgaris*,[50] *Callistoctopus minor*,[48] and the Decapodiform *E. scolopes*[49]). Given that the divergence time between Decapodiformes and Octopodiformes is estimated to be around 270 mya, much higher taxonomic sampling is needed (particularly within the decapodiform

lineage) in order to identify their modes of genome evolution. Within Octopodiformes, more basally branching species (such as cirrate octopods) would be required to reconstruct how independent gene family expansions occurred, and the extent to which gene co-localization is affected by genome reorganization within each clade.

Recent advances in sequencing approaches allow for complete chromosomal reconstructions of the genome.[103,104] It is becoming more affordable to generate such data for any species of interest. This leads to an unprecedented opportunity to identify how different genomic characters evolve in distinct clades. Two major approaches in evolutionary dating of genomic characters are using empirical dating estimates and parsimony approaches (e.g., phylostratigraphic dating[105]). One of the widely used methods to study the evolutionary history of genes is through phylogenetic approaches. To this end, the timing of gene duplication can be directly estimated. Using the example of large gene families, such as zinc-finger transcription factors in cephalopods, we can identify many species-specific duplication events and estimate their timing. Additionally, branch length estimation indicates the total amount of amino acid or nucleotide substitutions at a given branch in a phylogenetic tree enabling the quantification of evolutionary rates.

Another approach is to study the turnover rate in the non-coding regions of the genomes. Many genomes are enriched in transposable elements and these can comprise over 50% of the total DNA.[106] Given that many of them are transcriptionally active, it is not surprising to observe different repeat landscapes even in closely related species. The measurement of transposon activity, however, is difficult due to their short length (for both detection and distance estimation) and the reliance on consensus-based methods. In cephalopods, we cannot achieve resolution beyond 100 million years ago,[30] much younger than the estimated coleoid divergence of around 270 million years. For many other genomic characters (e.g., novel genes), no models for their empirical dating exist yet and the only approach to estimate their gain and loss dynamics is through parsimony or phylostratigraphy.[5,20]

Taken together, quantifying the evolution of genomic characters across scales, from modification of individual nucleotide sequences, to changes of gene family sizes, to their order on the chromosomes, will allow for a complete picture of the various modes of genome evolution in cephalopods. Presence of different modes of evolution can be expected in different cephalopod lineages due to various selective regimes. Additionally, while in some species or even organs, one genomic novelty can dominate, another mode can contribute to a different organ. Cephalopods with their large and complex genomes and complex morphology provide for a good model to study this general property of metazoan genomes.

8. Conclusions and Prospects

How genomic information translates into phenotype is still largely unknown outside of a relatively few examples. Cephalopods, given their many phenotypic innovations and large and complex genomes, provide for a unique model system to study the effect of various genomic characters on the phenotype. We show that a transition to multi-scale genome analyses, considering how distinct levels of genomic organization are evolving at the same time, is required to achieve this goal. We show specific examples where organs in cephalopods were associated with distinct types of genomic novelties. We argue that deciphering genomic signatures of co-evolution in several organ systems will allow us to dissect their functional linkage. This may reveal general principles that connect genomic and organismal evolution and thereby provide key insights required to identify evolutionarily linked subsystems in cephalopods and beyond.

Acknowledgements

E.A.R. and O.S. are supported by the Austrian Science Fund (Grant No. P30686-B29). E.A.R. is supported by Stazione Zoologica Anton Dohrn (Naples, Italy) PhD Program. The authors wish to thank Graziano Fiorito (SZN, Italy), Hannah Schmidbaur (University of Vienna, Austria), Thomas Hummel (University of Vienna, Austria) for many insightful comments and reading of the draft manuscript. The authors would like to apologize to all colleagues whose work has been omitted due to space constraints.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

cephalopod, gene duplication, genome rearrangement, novel gene, organismal innovation

Received: May 2, 2019
Revised: September 5, 2019
Published online: October 30, 2019

[1] S. B. Carroll, *PLoS Biol.* **2005**, *3*, 1159.
[2] B. Prud'homme, N. Gompel, S. B. Carroll, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8605.
[3] L. S. Babonis, M. Q. Martindale, J. F. Ryan, *Biol* **2016**, *16*, 1.
[4] D. M. Krylov, Y. I. Wolf, I. B. Rogozin, E. V. Koonin, *Genome Res.* **2003**, *13*, 2229.
[5] T. Domazet-Lošo, D. Tautz, *Nature* **2010**, *468*, 815.
[6] O. Simakov, T. Kawashima, *Dev. Biol.* **2017**, *427*, 179.
[7] G. P. Wagner, M. Pavlicev, J. M. Cheverud, *Nat. Rev. Genet.* **2007**, *8*, 921.
[8] E. H. Davidson, *Nature* **2010**, *468*, 911.
[9] R. D. Acemel, I. Maeso, J. L. Gómez-Skarmeta, *Wiley Interdiscip. Rev.: Dev. Biol.* **2017**, *6*, e265.
[10] W. Enard, *Curr. Biol.* **2016**, *26*, R1109.
[11] S. Shigeno, P. L. R. Andrews, G. Ponte, G. Fiorito, *Front. Physiol.* **2018**, *9*, 952.
[12] H. Wang, M. La Russa, L. S. Qi, *Annu. Rev. Biochem.* **2016**, *85*, 227.
[13] T. Rowe, *Science* **1996**, *273*, 651.
[14] J. H. Kaas, *Brain Res. Bull.* **2008**, *75*, 384.
[15] M. A. Clark, N. A. Moran, P. Baumann, J. J. Wernegreen, *Evolution* **2000**, *54*, 517.

[16] A. H. Moeller, A. Caro-Quintero, D. Mjungu, A. V. Georgiev, E. V. Lonsdorf, M. N. Muller, A. E. Pusey, M. Peeters, B. H. Hahn, H. Ochman, *Science* **2016**, *353*, 380.
[17] D. Juan, F. Pazos, A. Valencia, *FEBS Lett.* **2008**, *582*, 1225.
[18] E. Teppa, D. J. Zea, C. Marino-Buslje, *Protein Sci.* **2017**, *26*, 2438.
[19] D. Barker, M. Pagel, *PLoS Comput. Biol.* **2005**, *1*, e3.
[20] V. Ruano-Rubio, O. Poch, J. D. Thompson, *BMC Bioinf.* **2009**, *10*, 383.
[21] H. Zhou, E. Jakobsson, *PLoS One* **2013**, *8*, e81100.
[22] C. E. O'Brien, K. Roumbedakis, I. E. Winkelmann, *Front. Physiol.* **2018**, *9*, 700.
[23] B. Kröger, J. Vinther, D. Fuchs, *BioEssays* **2011**, *33*, 602.
[24] R. Hanlon, J. Messenger, *Cephalopod Behaviour*, 2nd ed., Cambridge University Press, Cambridge, UK **2018**.
[25] I. Winkelmann, P. F. Campos, J. Strugnell, Y. Cherel, P. J. Smith, T. Kubodera, L. Allcock, M. L. Kampmann, H. Schroeder, A. Guerra, M. Norman, J. Finn, D. Ingrao, M. Clarke, M. T. Gilbert, *Proc. R. Soc. B: Biol. Sci.* **2013**, *280*, 20130273.
[26] T. Wollesen, S. V. Rodríguez Monje, C. McDougall, B. M. Degnan, A. Wanninger, *EvoDevo* **2015**, *6*, 41.
[27] K. M. Koenig, P. Sun, E. Meyer, J. M. Gross, *Development* **2016**, *143*, 3168.
[28] S. Shigeno, R. Parnaik, C. B. Albertin, C. W. Ragsdale, *Zoolog. Lett.* **2015**, *1*, 26.
[29] T. Shomrat, A. L. Turchetti-Maia, N Stern-Mentch, J. A. Basil, B. Hochner, *J. Comp. Physiol. A* **2015**, *201*, 947.
[30] C. B. Albertin, O Simakov, T Mitros, Z. Y. Wang, J. R. Pungor, E. Edsinger-Gonzales, S. Brenner, C. W. Ragsdale, D. S. Rokhsar, *Nature* **2015**, *524*, 220.
[31] F. Henning, A. Meyer, *Annu. Rev. Genomics Hum. Genet.* **2014**, *15*, 417.
[32] O. Simakov, F. Marletaz, S-J. Cho, E. Edsinger-Gonzales, P. Havlak, U. Hellsten, D.-H. Kuo, T. Larsson, J. Lv, D. Arendt, R. Savage, K. Osoegawa, P. de Jong, J. Grimwood, J. A. Chapman, H. Shapiro, A. Aerts, R. P. Otillar, A. Y. Terry, J. L. Boore, I. V. Grigoriev, D. R. Lindberg, E. C. Seaver, D. A. Weisblat, N. H. Putnam, D. S. Rokhsar, *Nature* **2013**, *493*, 526.
[33] Drosophila 12 Genomes Consortium, A. G. Clark, M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver, T. A. Markow, T. C. Kaufman, M. Kellis, W. Gelbart, V. N. Iyer, D. A. Pollard, T. B. Sackton, A. M. Larracuente, N. D. Singh, J. P. Abad, D. N. Abt, B. Adryan, M. Aguade, H. Akashi, W. W. Anderson, C. F. Aquadro, D. H. Ardell, R. Arguello, C. G. Artieri, D. A. Barbash, D. Barker, P. Barsanti, P. Batterham, S. Batzoglou, et al., *Nature* **2007**, *450*, 203.
[34] U. Technau, S. Rudd, P. Maxwell, P. M. Gordon, M. Saina, L. C. Grasso, D. C. Hayward, C. W. Sensen, R. Saint, T. W. Holstein, E. E. Ball, D. J. Miller, *Trends Genet.* **2005**, *21*, 633.
[35] A. McLysaght, K. H. Wolfe, *Nat. Genet.* **2002**, *31*, 200.
[36] Y. Van de Peer, S. Maere, A. Meyer, *Nat. Rev. Genet.* **2009**, *10*, 725.
[37] K. Khalturin, F. Anton-Erxleben, S. Sassmann, J. Wittlieb, G. Hemmrich, T. C. G. Bosch, *PLoS Biol.* **2008**, *6*, 2436.
[38] K. Khalturin, G. Hemmrich, S. Fraune, R. Augustin, T. C. Bosch, *Trends Genet.* **2009**, *25*, 404.
[39] A. R. Tanner, D Fuchs, I. E. Winkelmann, M. T. P. Gilbert, M. S. Pankey, Â. M. Ribeiro, K. M. Kocot, K. M. Halanych, T. H. Oakley, R. R. da Fonseca, D. Pisani, J. Vinther, *Proc. R. Soc. B: Biol. Sci.* **2017**, *284*, 20162818.
[40] B. Kröger, *Paleobiology* **2005**, *31*, 253.
[41] S. von Boletzky, *Adv. Mar. Biol.* **2003**, *44*, 143.
[42] A. Naef, *Cephalopoda*, Smithsonian Institution Libraries, Washington D.C. **2000**.
[43] M. T. Nödl, A Kerbl, M. G. Walzl, G. B. Müller, H. Gert de Couet, *Front. Zool.* **2016**, *13*, 44.
[44] A. Packard, *Biol. Rev.* **1972**, *47*, 241.
[45] J. Z. Young, *Proc. Zool. Soc. Lond.* **1963**, *140*, 229.
[46] M. Nixon, J. Z. Young, Z. John, *The Brains and Lives of Cephalopods*, Oxford University Press, Oxford, UK **2003**.
[47] R. A. Cloney, E. Florey, *Z. Zellforsch. Mikrosk. Anat.* **1968**, *89*, 250.
[48] B.-M. Kim, S. Kang, D.-H. Ahn, S.-H. Jung, H. Rhee, J. S. Yoo, J. E. Lee, S. Lee, Y. H. Han, K. B. Ryu, S. J. Cho, H. Park, H. S. An, *Gigascience* **2018**, *7*. https://doi.org/10.1093/gigascience/giy119.
[49] M. Belcaid, G. Casaburi, S. J. McAnulty, H. Schmidbaur, A. M. Suria, S. Moriano-Gutierrez, M. S. Pankey, T. H. Oakley, N. Kremer, E. J. Koch, A. J. Collins, H. Nguyen, S. Lek, I. Goncharenko-Foster, P. Minx, E. Sodergren, G. Weinstock, D. S. Rokhsar, M. McFall-Ngai, O. Simakov, J. S. Foster, S. V. Nyholm, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3030.
[50] I. Zarrella, K. Herten, G. E. Maes, S. Tai, M. Yang, E. Seuntjens, E. A. Ritschard, M. Zach, R. Styfhals, R. Sanges, O. Simakov, G. Ponte, G. Fiorito, *Sci. Data* **2019**, *6*, 13.
[51] R. Styfhals, E. Seuntjens, O. Simakov, R. Sanges, G. Fiorito, *Front. Physiol.* **2019**, *9*, 1905.
[52] Z. Y. Wang, C. W. Ragsdale, *Semin. Cell Dev. Biol.* **2017**, *69*, 151.
[53] E. A. Ritschard, R. R. Fitak, O. Simakov, S. Johnsen, *Proc. R. Soc. B: Biol. Sci.* **2019**, *286*, 20182929.
[54] S. M. Peyer, M. S. Pankey, T. H. Oakley, M. J. McFall-Ngai, *Mech. Dev.* **2014**, *131*, 111.
[55] S. Garrett, J. J. C. Rosenthal, *Science* **2012**, *335*, 848.
[56] N. Liscovitch-Brauer, S. Alon, H. T. Porath, B. Elstein, R. Unger, T. Ziv, A. Admon, E. Y. Levanon, J. J. C. Rosenthal, E. Eisenberg, *Cell* **2017**, *169*, 191.e11.
[57] S. Ohno, *Evolution by Gene Duplication*, Allen & Unwin; Springer-Verlag, London, New York **1970**.
[58] D. M. Bear, J.-M. Lassance, H. E. Hoekstra, S. R. Datta, *Curr. Biol.* **2016**, *26*, R1039.
[59] M. Chakraborty, E. D. Jarvis, *Philos. Trans. R. Soc. B: Biol. Sci.* **2015**, *370*, 20150056.
[60] D. Arendt, H. Hausen, G. Purschke, *Philos. Trans. R. Soc. B: Biol. Sci.* **2009**, *364*, 2809.
[61] D. Arendt, J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko C, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler, G. P. Wagner, *Nat. Rev. Genet.* **2016**, *17*, 744.
[62] K. A. Adipietro, J. D. Mainland, H. Matsunami, *PLoS Genet.* **2012**, *8*, e1002821.
[63] S. Rouquier, A. Blancher, D. Giorgi, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2870.
[64] S. Serizawa, K. Miyamichi, H. Takeuchi, Y. Yamagishi, M. Suzuki, H. Sakano, *Cell* **2006**, *127*, 1057.
[65] S. Hasegawa, S. Hamada, Y. Kumode, S. Esumi, S. Katori, E. Fukuda, Y. Uchiyama, T. Hirabayashi, P. Mombaerts, T. Yagi, *Mol. Cell. Neurosci.* **2008**, *38*, 66.
[66] T. Shimizu, M. Hibi, *Dev., Growth Differ.* **2009**, *51*, 221.
[67] F. W. Grasso, in *Cephalopod Cognition*, Cambridge University Press, Cambridge, UK **2014**, pp. 94–122.
[68] L. Zullo, G. Sumbre, C. Agnisola, T. Flash, B. Hochner, *Curr. Biol.* **2009**, *19*, 1632.
[69] A.-S. Darmaillacq, L. Dickel, J. Mather, in *Cephalopod Cognition*, Cambridge University Press, Cambridge, UK **2014**.
[70] A. D. Hargreaves, M. T. Swaiin, M. J. Hegarty, D. W. Logan, J. F. Mulley, *Genome Biol. Evol.* **2014**, *6*, 2088.
[71] J. Reyes-Velasco, D. C. Card, A. L. Andrew, K. J. Shaney, R. H. Adams, D. R. Schield, N. R. Casewell, S. P. Mackessy, T. A. Castoe, *Mol. Biol. Evol.* **2015**, *32*, 173.
[72] B. G. Fry, *Genome Res.* **2005**, *15*, 403.
[73] S. Dutertre, A.-H. Jin, I. Vetter, B. Hamilton, K. Sunagar, V. Lavergne, V. Dutertre, B. G. Fry, A. Antunes, D. J. Venter, P. F. Alewood, R. J. Lewis, *Nat. Commun.* **2014**, *5*, 3521.

[74] L. C. H. W. Fingerhut, J. M. Strugnell, P Faou, Á. R. Labiaga, J. Zhang, I. R. Cooke, *J. Proteome Res.* **2018**, *17*, 3866.
[75] A. Kanda, K. Takuwa-Kuroda, M. Aoyama, H. Satake, *FEBS J.* **2007**, *274*, 2229.
[76] E. Baldascino, G. Di Cristina, P. Tedesco, C. Hobbs, C. Hobbs, T. J. Shaw, G. Ponte, P. L. R. Andrews, *Front. Physiol.* **2017**, *8*, 1001.
[77] L. N. Key, P. R. Boyle, M. Jaspars, *Toxicon* **2002**, *40*, 677.
[78] T. Ruder, S. A. Ali, K. Ormerod, A. Brust, M. L. Roymanchadi, S. Ventura, E. A. Undheim, T. N. Jackson, A. J. Mercier, G. F. King, P. F. Alewood, B. G. Fry, *Peptides* **2013**, *47*, 71.
[79] V. Erspamer, A. Anastasi, *Experientia* **1962**, *18*, 58.
[80] G. Ponte, M. V. Modica, *Front. Physiol.* **2017**, *8*, 580.
[81] B. L. Whitelaw, I. R. Cooke, J Finn, K. Zenger, J. M. Strugnell, *Aquat. Toxicol.* **2019**, *206*, 114.
[82] B. L. Whitelaw, J. M. Strugnell, P Faou, R. R. da Fonseca, N. E. Hall, M. Norman, J. Finn, I. R. Cooke, *J. Proteome Res.* **2016**, *15*, 3284.
[83] M. Heide, K. R. Long, W. B. Huttner, *Curr. Opin. Cell Biol.* **2017**, *49*, 22.
[84] P. H. Sudmant, J. Huddleston, C. R. Catacchio, M. Malig, L. W. Hillier, C. Baker, K. Mohajeri, I. Kondova, R. E. Bontrop, S. Persengiev, F. Antonacci, M. Ventura, J. Prado-Martinez; Great Ape Genome Project; T. Marques-Bonet, E. E. Eichler, *Genome Res.* **2013**, *23*, 1373.
[85] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, *Nature* **2012**, *485*, 376.
[86] B. D. Pope, T. Ryba, V. Dileep, F. Yue, W. Wu, O. Denas, D. L. Vera, Y. Wang, R. S. Hansen, T. K. Canfield, R. E. Thurman, Y. Cheng, G. Gülsoy, J. H. Dennis, M. P. Snyder, J. A. Stamatoyannopoulos, J. Taylor, R. C. Hardison, T. Kahveci, B. Ren, D. M. Gilbert, *Nature* **2014** *515*, 402.
[87] O. Symmons, L. Pan, S. Remeseiro, T. Aktas, F. Klien, W. Huber, F. Spitz, *Dev. Cell* **2016**, *39*, 529.
[88] N. H. Lazar, K. A. Nevonen, B. O'Connell, C. McCann, R. J. O'Neill, R. E. Green, T. J. Meyer, M. Okhovat, L. Carbone, *Genome Res.* **2018**, *28*, 983.
[89] C. Feschotte, *Nat. Rev. Genet.* **2008**, *9*, 397.
[90] C. B. Lowe, G. Bejerano, D. Haussler, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8005.
[91] S. Nowoshilow, S. Schloissnig, J.-F. Fei, A. Dahl, A. W. C. Pang, M. Pippel, S. Winkler, A. R. Hastie, G. Young, J. G. Roscito, F. Falcon, D. Knapp, S. Powell, A. Cruz, H. Cao, B. Habermann, M. Hiller, E. M. Tanaka, E. W. Myers, *Nature* **2018**, *554*, 50.
[92] P. N. Lee, P. Callaerts, H. G. deCouet, M. Q. Martindale, *Nature* **2003**, *424*, 1061.
[93] D. G. Knowles, A. McLysaght, *Genome Res.* **2009**, *19*, 1752.
[94] D. Tautz, T. Domazet-Lošo, *Nat. Rev. Genet.* **2011**, *12*, 692.
[95] H. Kaessmann, *Genome Res.* **2010**, *20*, 1313.
[96] C. Schlötterer, *Trends Genet.* **2015**, *31*, 215.
[97] A. McLysaght, L. D. Hurst, *Nat. Rev. Genet.* **2016**, *17*, 567.
[98] N. Palmieri, C. Kosiol, C. Schlötterer, *eLife* **2014**, *3*, e01311.
[99] Q. Zhou, G. Zhang, Y. Zhang, S. Xu, R. Zhao, Z. Zhan, X. Li, Y. Ding, S. Yang, W. Wang, *Genome Res.* **2008**, *18*, 1446.
[100] J. A. Capra, K. S. Pollard, M. Singh, *Genome Biol.* **2010**, *11*, R127.
[101] S. Chen, Y. E. Zhang, M. Long, *Science*. **2010**, *330*, 1682.
[102] A. Andouche, Y. Bassaglia, S. Baratte, L. Bonnaud, *Dev. Dynam.* **2013**, *242*, 560.
[103] N. H. Putnam, B. L. O'Connell, J. C. Stites, B. J. Rice, M. Blanchette, R. Calef, C. J. Troll, A. Fields, P. D. Hartley, C. W. Sugnet, D. Haussler, D. S. Rokhsar, R. E. Green, *Genome Res.* **2016**, *26*, 342.
[104] B. J. Matthews, O. Dudchenko, S. B. Kingan, S. Koren, I. Antoshechkin, J. E. Crawford, W. J. Glassford, M. Herre, S. N. Redmond, N. H. Rose, G. D. Weedall, Y. Wu, S. S. Batra, C. A. Brito-Sierra, S. D. Buckingham, C. L. Campbell, S. Chan, E. Cox, B. R. Evans, T. Fansiri, I. Filipović, A. Fontaine, A. Gloria-Soria, R. Hall, V. S. Joardar, A. K. Jones, R. G. G. Kay, V. K. Kodali, J. Lee, G. J. Lycett, et al., *Nature* **2018**, *563*, 501.
[105] T. Domazet-Lošo, J. Brajković, D. Tautz, *Trends Genet.* **2007**, *23*, 533.
[106] B. Chénais, A. Caruso, S. Hiard, N. Casse, *Gene* **2012**, *509*, 7.
[107] M. Srivastava, O. Simakov, J. Chapman, B. Fahey, M. E. A. Gauthier, T. Mitros, G. S. Richards, C. Conaco, M. Dacre, U. Hellsten, C. Larroux, N. H. Putnam, M. Stanke, M. Adamska, A. Darling, S. M. Degnan, T. H. Oakley, D. C. Plachetzki, Y. Zhai, M. Adamski, A. Calcino, S. F. Cummins, D. M. Goodstein, C. Harris, D. J. Jackson, S. P. Leys, S. Shu, B. J. Woodcroft, M. Vervoort, K. S. Kosik, G. Manning, B. M. Degnan, D. S. Rokhsar, *Nature* **2010**, *466*, 720.
[108] N. H. Putnam, M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, J. Jurka, G. Genikhovich, I. V. Grigoriev, S. M. Lucas, R. E. Steele, J. R. Finnerty, U. Technau, M. Q. Martindale, D. S. Rokhsar, *Science* **2007**, *317*, 86.
[109] J. Paps, P. W. H. H. Holland, *Nat. Commun.* **2018**, *9*, 1730.
[110] F. Raible, K. Tessmar-Raible, K. Osoegawa, P. Wincker, C. Jubin, G. Balavoine, D. Ferrier, V. Benes, P. de Jong, J. Weissenbach, P. Bork, D. Arendt, *Science* **2005**, *310*, 1325.
[111] K. Voordeckers, K. Pougach, K. J. Verstrepen, *Curr. Opin. Biotechnol.* **2015**, *34*, 180.

