Chromosomal fusions and subsequent rearrangements shaped octopus genomes

Masa-aki Yoshida[*,1], Emese Tóth[2], Koto Kon-Nanjo[2], Tetsuo Kon[2], Kazuki Hirota[3,4], Atsushi Toyoda[5], Hidehiro Toh[5], Hideyuki Miyazawa[6], Makoto Terauchi[6], Hideki Noguchi[6], Davin H. E. Setiamarga[*,3], Oleg Simakov[*,2]

Affiliations:

1. **Marine Biological Science Section, Education and Research Center for Biological Resources, Faculty of Life and Environmental Science, Shimane University, Oki, Shimane 685-0024, Japan**
2. **Department for Neuroscience and Developmental Biology, University of Vienna, Vienna 1030, Austria**
3. **Department of Applied Chemistry and Biochemistry, National Institute of Technology (KOSEN), Gobo, Wakayama 644-0023, Japan**
4. **Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan**
5. **Advanced Genomics Center, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan**
6. **Center for Genome Informatics, Joint Support-Center for Data Science Research, Research Organization of Information and Systems, 1111 Yata, Mishima, Shizuoka 411-8540, Japan**

***Correspondence; mayoshida@life.shimane-u.ac.jp (M.A.Y), davin@wakayama-nct.ac.jp (D.H.E.S), oleg.simakov@univie.ac.at (O.S.)**

Abstract

Why some animal groups retain ancestral chromosomal complements while others change significantly is a fundamental question in evolutionary genomics. Few systems exist where accumulations of chromosomal changes can be studied in the context of morphological innovation. In coleoid cephalopods (octopus, squid, cuttlefish), an ancient coleoid chromosomal rearrangement event (ACCRE) has led to a substantial increase in the chromosome number and a new set of chromosomal homologies. Compared to typical molluscan or bilaterian genomes, ACCRE has enabled the origin of many novel regulatory regions in coleoid cephalopods. However, the discrepancies between extant octopodiform (octopus, ~30 chromosomes) and decapodiform (squid and cuttlefish, ~46 chromosomes) karyotypes and the direction of these evolutionary changes remain unexplained. Here we provide a draft genome assembly of the vampire squid *Vampyroteuthis* sp., the largest cephalopod genome sequenced to-date (over 10 gigabasepairs). Through syntenic comparisons, we infer that this basally branching octopodiform species shows partial retention of the chromosomal complement of Decapodiformes, indicating its more ancestral state, while at the same time also suggest the derived nature of the octopod karyotype. Together with the analysis of a new chromosome-level assembly of the pelagic octopod *Argonauta hians*, we identified irreversible chromosomal fusion-with-mixing events followed by inter-chromosomal translocations in octopods after their separation from the vampire squid lineage. This secondary reduction and mixing within octopod chromosomes has enabled the origin of a more entangled genomic configuration, shedding light onto the early evolutionary transitions within the clade. Our results offer broader insights into general patterns of

chromosomal evolution following large-scale rearrangement events in animal genomes, and clarify the evolutionary trajectory leading to present-day coleoid diversity.

Keywords

Karyotype evolution, Genome synteny, Chromosome fusions, Genome rearrangements, Gene translocation, Orthology, Bioinformatics, Deep Sea creatures, Coleoid Cephalopods, Living fossils

Introduction

Cephalopods, such as octopuses and squids, are known for their remarkably advanced behaviors[1]. These include solving complex problems, learning by observation, adapting to different tasks, and using exceptional camouflage techniques. While their nervous systems are similar in size to mammals, their structure and composition are distinct. The origin of these complex behaviors, advanced nervous systems, and other traits seen in cephalopods has a long evolutionary history.

Cephalopods split into nautiloids and the lineage leading to coleoids during the Late Ordovician (ca. 450 Ma)[2] (Figure 1a). While extant nautiloids retained their external shells, coleoids internalized or lost theirs further down the line during their evolution, allowing them to adapt to diverse marine environments[3,4]. Shell loss may also have driven the evolution of fast decision-making and camouflage-dependent survival strategies, requiring enhanced neural processing, and thus possibly facilitating the development of adaptive and complex behaviors[5]. By the Late Permian to the Middle

Triassic (~260–235 Ma), coleoids had further diversified into Octopodiformes (vampire squids and octopuses) and Decapodiformes (squids and cuttlefishes)[6-9] (Figure 1a). The evolution of advanced nervous systems within these clades has likely been driven by clade-specific predation pressures and ecological adaptation[10-13]

These adaptations likely required major underlying genetic innovations. Recent genomic studies have begun to reveal how large-scale changes in genome structure contributed to the evolution of their nervous systems, cognitive abilities, and other unique traits. Comparative analyses of cephalopod genomes highlight key differences in genome organization between coleoids and their more primitive relatives, shedding light on the molecular basis of their adaptations[14,15]. While the early-branching nautilus retains the genomic structure typical of conchiferan mollusks[16-18], a major genome rearrangement event, the ancient coleoid chromosomal rearrangement event (ACCRE), occurred in the lineage leading to the coleoids[16]. This large-scale genome rearrangement has facilitated regulatory novelties, for example via the formation of new local gene linkages, whose regulatory interactions depend on the topological organization of the genome[19].

While ACCRE has significantly altered the coleoid genome architecture, restructuring gene linkages and regulatory landscapes[19], further comparative studies also revealed substantial divergence in chromosome structure between major coleoid lineages, which most likely happened post-ACCRE. Octopuses and squids differ not only in chromosome numbers[16], but also in repeat content[20], and the expansion of particular gene families, such as *Protocadherins*[21] and *GPCRs*[22]. Interestingly, while coleoid chromosomes have remained largely intact within both decapodiform and

octopodiform lineages, their homologies are not strictly one-to-one, with some chromosomes showing many-to-many correspondences[16]. The discrepancy between predicted typical chromosome numbers (~30 in Octopodiformes vs. ~46 in Decapodiformes) also raises questions about whether chromosome numbers contracted or expanded during early coleoid evolution (Figure 1b). Thus far, key genomic data from early branching Octopodiformes or Decapodiformes have been lacking, leaving the directionality of these changes uncertain.

As an early-branching "living fossil" within Octopodiformes, the vampire squid is phylogenetically crucial for understanding the evolutionary processes that shaped cephalopod diversification[23]. With its oldest known and closest fossil relative dating to the Middle Jurassic (ca. 165 Ma)[24,25] the extant vampire squid (*Vampyroteuthis* cf. *infernalis*) is a globally distributed deep-sea pelagic cephalopod with an opportunistic detritivorous and zooplanktivorous feeding strategy[25,26]. The species is the only extant representative of the vampyropods, an ancient clade within Octopodiformes (Figure 1) that has existed since the Mesozoic Era, from at least the Late Jurassic (e.g., *Vampyronassa*; ca. 165 Ma) [25] with some evidence suggesting its origins in the Early Jurassic (e.g., *Teudopsis* and *Simoniteuthis*; ca.183 Ma)[27-29]. Ancient vampyropods coexisted with the ammonites and the belemnites throughout the Jurassic (~201–145 Ma) and Cretaceous (~145–66 Ma), occupying various ecological niches across global marine environments[28,30,31]. However, most vampyropod lineages went extinct at the end of the Cretaceous (~66 Ma), eventually leaving extant *Vampyroteuthis* as the sole surviving species[32]. In this study, we present the draft genome and evolutionary analysis of *Vampyroteuthis*, offering insights that may help resolve long-standing

questions about the evolutionary trajectories of the lineages of Octopodiformes and Decapodiformes, including their earliest evolutionary transitions.

To provide a further resolution within Octopodiformes, particularly in the octopod lineage, we also present the chromosome-level genome sequence of the argonaut octopod *Argonauta hians*. Along with its congeneric species, *A. argo*, these two species form Argonautoidea, an octopodiform clade composed mainly of pelagic octopods with shell-like calcified eggcases that were secondarily acquired in the lineage through convergence and iterative evolution[33,34]. Including these two new genomes in the chromosomal rearrangement analyses reveals the directionality of chromosomal evolution, genomic structure changes, and size variation in cephalopod genomes, particularly in the lineage leading to extant coleoids and, eventually, the octopods.

Results and Discussion

We sequenced the draft genome of *Vampyroteuthis* sp. from an individual collected in the West Pacific Ocean. Genome sequencing was performed using the PacBio HiFi platform at approximately 40X coverage and assembled with hifiasm, resulting in a 14 Gb assembly. Different assembly settings were tested, yielding similar assembly sizes (**Supplementary Note**). A more stringent setting to both avoids haplotype overestimation and remove duplicates still resulted in similar completeness (BUSCO score of 95%) and an assembly size of over 11 Gb. Therefore, regardless of the assembly settings, the total length estimates confirm that the *Vampyroteuthis* genome constitutes the largest cephalopod genome sequenced to date (**Supplementary Figure 1)**. Since we could not currently obtain additional samples for

this rare deep-sea species, chromosomal conformational data is not yet attainable. However, while our *Vampyroteuthis* draft genome assembly consists of numerous sub-chromosomal level contigs (about 5,600 contigs, N50 of 9.96 Mb, and the longest contig of 51.7 Mb), it still contains sufficient syntenic information[23], with 162 contigs containing 15 or more orthologous genes and the largest contig containing 130 orthologous genes. Besides the *Vampyroteuthis* genome, to obtain further information on karyotypic changes in the Octopodiformes, we also sequenced the genome of the argonautid octopod *Argonauta hians*. Similar to its congener *A. argo*[34], the genome size of *A. hians* (~1.57 Gb) is only half that of *Octopus* (~3 Gb)[15]. The draft genome of *A. hians* was assembled into putative 28 chromosomes, comparable in contiguity and resolution to other octopod genomes, which typically have 30 chromosomes (**Supplementary Figure 2**).

To investigate how these genome architectures relate to chromosomal evolution, we conducted comparative analyses across multiple cephalopod lineages using a modified synteny approach (see Methods). Importantly, this approach enabled us to profile the presence of irreversible fusion-with-mixing (FWM) chromosomal characters[34,36,37]. Since the mixing of genes on fused chromosomes cannot revert to its ancestral unfused state, any ancestral FWM would constitute a strong synapomorphic character shared across all members of the clade in question (**Figure 2**). Using this approach, we confirmed the highly rearranged nature of coleoid chromosomes relative to the nautilus (**Supplementary Figure 3**), consistent with a previous report[16]. We then analyzed our newly sequenced cephalopod genomes to obtain novel insights into their chromosomal evolution. Despite its gigantic size, due to the massive expansion of

repetitive elements, we found that the *Vampyroteuthis* genome structure predominantly adheres to the basic coleoid karyotype rather than the more ancestral nautilus or the general molluscan chromosomal homology complement[23] (**Supplementary Figure 3)**. This pattern remained consistent even when alternative mapping and orthology search strategies were applied (**Supplementary Figure 4**), confirming the robustness of this result.

Surprisingly, we found that the *Vampyroteuthis* chromosomal structure is more closely aligned with that of Decapodiformes, including those of the longfin squid *Doryteuthis pealeii*, the bobtail squid *Euprymna scolopes*, the diamond squid *Thysanoteuthis rhombus*, and the cuttlefish *Sepia lycidas* (**Figure 2**, **Supplementary Figure 5**), despite its well-supported phylogenetic placement in Octopodiformes and thus its affinities with the octopods. Despite the fragmented nature of the *Vampyroteuthis* genome assembly, its contigs often correspond to single *Doryteuthis* chromosomes, reflecting a high degree of chromosomal retention. Conversely, *Vampyroteuthis* contigs exhibited complex relationships with the chromosomes of the other octopodiform species (the muddy argonaut *Argonauta hians*, the curled octopus *Eledone cirrhosa*, and the common octopus *Octopus vulgaris*), indicating synapomorphic fusions and rearrangements shared across octopods (**Figure 2**, **Supplementary Figure 5**). Multiple *Doryteuthis* chromosomes and their one-to-one homologous *Vampyroteuthis* sets of contigs correspond to single octopod chromosomes and show highly entropic mixing. For example, *O. vulgaris* chromosome 22 corresponds to *Doryteuthis* chromosomes 8 and 9, each with its own non-overlapping set of *Vampyroteuthis* contigs, with the degree of entropic mixing of these

two states along the *Octopus* chromosome of 0.193, as measured by the normalized turbulence[38]. This implies that FWM of two ancestral decapodiform-like chromosomes (*Doryteuthis* chromosomes 8 and 9) occurred at the base of the octopod (*Eledone*, *Argonauta*, *Octopus*) lineage after its separation from *Vampyroteuthis*. For comparison, *O. vulgaris* chromosome 2 shows inter-chromosomal translocations with sharp syntenic boundaries and little mixing (<0.1 normalized turbulence for *Doryteuthis* chromosomes 9 and 29), suggesting a more recent origin of this translocation event. These chromosomal changes most likely accumulated gradually in the lineage leading to the octopods after their divergence from the *Vampyroteuthis* lineage (**Figure 2**, **Supplementary Figure 5**). Given the basal position of *Vampyroteuthis* among the Octopodiformes, the conservation of chromosomal structure across decapodiform lineages and *Vampyroteuthis*, along with the drastic chromosomal rearrangements shared by all octopods, is striking, as these observations suggest the ancestrality of a large part of the decapodiform karyotype.

To further explore the extent of chromosomal fusion and mixing in the lineage leading to octopods, we examined how *Doryteuthis* chromosomes were mapped onto the *Vampyroteuthis* genome. The results revealed chromosomal rearrangements with a high degree of mixing within the *Vampyroteuthis* contigs. For example, contig ptg000753l is mapped to *Doryteuthis* chromosomes 7, 37, and 35 with normalized turbulence measure of 0.285 for Dpe35 and Dpe37, 0.290 for Dpe07 and Dpe37, and 0.193 for Dpe07 and Dpe35[38]. The pattern is observed for 33 out of the 46 *Doryteuthis* chromosomes, where parts of the chromosomes were mapped onto the same set of *Vampyroteuthis* contigs (Fisher's exact test p-value <0.05), suggesting irreversible

fusion-with-mixing. As the same chromosomes were found to be fused and mixed in all other octopodiform species, this suggests an ancient synapomorphy supporting the affinity of *Vampyroteuthis* with octopods. On top of this synapomorphic signature, further substantial chromosomal fusion-with-mixing occurred in the lineage leading to octopods (*Eledone*, *Argonauta*, and *Octopus*) (**Figure 2**). Octopod homologs of only 3 *Doryteuthis* chromosomes (including the proposed sex chromosome[39]) remain unfused, suggesting that 43 out of 46 cases where chromosomes underwent fusions on the branch leading to octopods. Of these, at least 10 chromosomes were previously unfused in *Vampyroteuthis*, indicating additional fusions after the divergence of octopods from the vampyropod lineage. Comparison with *O. vulgaris* karyotype showed that the argonaut underwent further fusions between chromosomes 3 and 8, as well as 6 and 7, reducing the chromosome number from 30 to 28, suggesting a general trend of chromosomal number reduction in octopuses.

Our observations thus provide strong evidence that the ancestral coleoid had a decapodiform-like karyotype, which later underwent additional fusions and karyotype reduction, first at the base of Octopodiformes, and then later in the octopods (**Figure 2**). The alternative scenario of chromosomal fissions is highly unlikely, as it would require independent, convergent fissions of hundreds of genes to produce identical chromosomal splits observed in all Decapodiformes and *Vampyroteuthis*[36]. We could also confirm this pattern by partitioning the *O. vulgaris* genome into fragments containing the same gene content as those contained in our *Vampyroteuthis* assembly. Using the same syntenic analysis approach, this heavily rearranged, “fragmented”, *Octopus* genome was still recapitulating octopod fusions (**Supplementary Figure 6**).

This artificially fragmented *O. vulgaris* genome assembly also showed a signal for the fusion of 42 out of 46 *Doryteuthis* chromosome homologs, compared to the 43 fused *Doryteuthis* chromosome homologs found in the complete *O. vulgaris* genome (Fisher's exact test p-value <0.05). This suggests that the contiguity of the *Vampyroteuthis* assembly is unlikely to impact the inference of chromosomal fusions. These results suggest that *Vampyroteuthis* has a chromosomal structure partially shared with extant Decapodiformes, supporting the interpretation that it retains the more ancestral coleoid configuration, while octopods underwent additional fusion-with-mixing events.

In addition to the fusion-rich history within Octopodiformes, we also find that the observed syntenic pattern in octopod chromosomes is more complex than a simple fusion-with-mixing of complete ancestral chromosomes. The observed pattern (**Figure 3**) can be explained by two possible scenarios (**Figure 3b**). First, it may suggest that major inter-chromosomal translocations have occurred in the lineage leading to octopods (*Eledone*, *Argonauta*, and *Octopus*; **Figure 2b**). The second possibility is that a large post-ACCRE ancestral coleoid karyotype independently fused into different configurations in Decapodiformes and Octopodiformes (**Figure 3b**). If the second scenario was the case, we would expect to find a Decapodiformes species with an unfused chromosomal pair that corresponds to separate chromosomes in *Vampyroteuthis* or octopods, yet fused in other Decapodiformes. However, since we have thus far not found any decapodiform with chromosomal arrangements more ancestral to the detected fusions (**Supplementary Figure 7**), our data most strongly support the first scenario of translocation and rearrangement-rich early octopod evolution. Although similar syntenic patterns can also arise after whole genome

duplication[36], we found no evidence of paralogous enrichment on any of the octopodiform chromosomes. Besides that, the observed chromosomal correspondence pattern does not indicate the presence of a consistent genome-wide doubling of ancestral chromosomes. Therefore, we believe it is unlikely for any whole-genome duplication event to have happened and be the cause of the syntenic patterns we observed. Additionally, comparisons within the octopods (**Supplementary Figure 8**) show that *Eledone*, despite its basal position on the tree, showed additional modifications to its karyotype, including translocations to the otherwise highly conserved ancestral coleoid chromosomes Dpe03-Ovu12. On the other hand, genomes of *A. hians* and *O. vulgaris* were strikingly similar and shared more chromosomal characteristics with *Vampyroteuthis*, suggesting a more ancestral octopod karyotype (**Supplementary Figure 8**).

Why has the vampire squid maintained its more ancestral chromosomal state despite a dramatic increase in genome size? While several studies have implied the effects of transposable elements (TEs) accumulation on enhancing genome rearrangement rate, some of the largest genomes sequenced to date (e.g., lungfish[40]) show surprisingly well-conserved karyotypes. This suggests that the *Vampyroteuthis* genome provides another example where TE-driven genome expansion was decoupled from an increase in genome rearrangements. Future studies of both epigenetic states and genome topology in this species should provide fruitful insight into the role of TE accumulation in the maintenance of ancestral genomic configuration in animal genomes.

To further investigate the role of chromosomal fusion events onto putative regulatory landscapes as reflected by conserved non-coding element (CNE)

preservation, we have conducted whole genome alignments between *Vampyroteuthis* and other coleoid cephalopods (Methods). We found that the conservation of non-coding regions was higher between *Vampyroteuthis* and Decapodiformes (over 7.1 Mb aligned sequences) compared to *Vampyroteuthis* and *Octopus* (2.5 Mb total alignment, **Figure 4a**). The decrease in the overall alignment length is particularly visible in *Argonauta* and *Octopus*. This pattern, while correlated with the genome size, is also corroborated by the overall repeat composition of the *Vampyroteuthis* genome, with long interspersed nuclear elements (LINEs) comprising the major part of the genome (13.59%) and short interspersed nuclear elements (SINEs) less than 1%. This is similar to other Decapodiformes genomes and in contrast to the SINE-dominated octopod genomes[20], indicating *Vampyroteuthis*'s higher propensity to retain ancestral non-coding features.

Using chromosomal homologies inferred from gene-level comparisons (**Figures 2** and **3**), we traced their CNE complement evolution on the derived octopod chromosomes. *Vampyroteuthis* homologs of unfused *Doryteuthis* chromosomes (Dpe09 and Dpe08) remain unmixed as separate units, whereas *Vampyroteuthis* contigs that correspond to the predicted octopodiform FWM events show substantial mixing of both coding and non-coding elements (Dpe29 and Dpe23, **Figure 4b**). All of these contigs map to a set of *O. vulgaris* chromosomes, resulting in a high degree of mixing of genes and their putative regulatory sites. From this observation, we can infer that a large portion of the regulatory landscapes on modern *O. vulgaris* chromosomes has a complex evolutionary origin, shaped by fusion-with-mixing events combining segments derived from distinct chromosomal units in the ancestral octopodiform lineage.

Conclusions

Our findings reinforce the hypothesis that the decapodiform-like karyotype represents the ancestral state of coleoids more closely, while strongly supporting the basal positioning of *Vampyroteuthis* within Octopodiformes. This conclusion aligns well with the paleontological data and insights regarding cephalopod bauplan evolution. Stem coleoid species from the Paleozoic, such as *Gordoniconus* and *Bactrites*, exhibit internal shells, streamlined morphologies, and ten arms, long before the appearance of the oldest vampyropods in the Jurassic, such as *Teudopsis* and *Vampyronassa*. Interestingly, *Vampyroteuthis* also possesses two long feeding filaments, thought to be vestigial arms, indicating a possible morphological affinity to the 10-armed decapodiform. The retention of parts of the decapodiform-like chromosomal architecture in *Vampyroteuthis*, along with the extensive chromosomal rearrangements observed in octopods, provides genomic evidence supporting the transition from decapodiform-like ancestors to Octopodiformes. Our analyses also indicate a surge of inter-chromosomal translocations at the base of the octopods, potentially facilitated by demographic shifts or genomic instability. These ACCRE-like inter-chromosomal translocations, possibly stabilized by the notably large genome sizes in coleoids, may have driven morphological innovations and regulatory complexity. For instance, the specialization of arms into feeding tentacles in squids and the reduction of arms in octopuses likely result from modifications in gene regulation and its chromosomal context, and not gene content. Similarly, the presence of shell matrix protein-coding genes in shell-less

octopods suggests that regulatory changes, rather than structural ones, facilitated shell reduction. Fossil record and molecular data, defining the timing of these innovations (**Figure 1a**), provide a temporal framework to further study the accumulation of chromosomal rearrangements and their long-term impact on the phenotype evolution in the lineage leading to the radiation of octopods.

Taken together, our findings have important implications for understanding the directionality of chromosomal evolution in the animal kingdom, its long-term impact on the emergence of novel putative regulatory regions, and, eventually, morphological innovations. Further attempts to obtain a chromosomal-scale assembly of Vampyroteuthis will help to determine the exact karyotypic changes that occurred at the base of Octopodiformes.

Materials and Methods

Samples

Vampyroteuthis sp. used for sequencing was derived from a single individual (juvenile, unknown sex) caught in Suruga bay by T/V Hokuto of Tokai University. The *A. hians* used for sequencing was derived from a single female caught as a bycatch in the fixed nets set along the coast in Oki Island, Shimane Prefecture, Japan (36°17′20.6″ ″N 133°12′46.4″″E). The egg-shell of the *A. hians* was deposited in the University Museum, the University of Tokyo with voucher ID: UMUT RM34224.

Sample preparations

Whole-genome shotgun sequencing was performed using PacBio and Illumina sequencing platforms. Genomic DNA was extracted from the whole body of *Vampyroteuthis* sp. and the ovary of *A. hians* using a Genomic-tip Kit (QIAGEN, Hilden, Germany) and was sheared into fragment sizes ranging from 15 kb to 20 kb with a g-tube device (Covaris Inc., MA, USA). PacBio HiFi libraries for *Vampyroteuthis* sp. and *A. hians* were prepared using a SMRTbell Prep Kit 3.0 (Pacific Bioscience, CA, USA) according to the manufacturer's instructions and were size-selected using the SageELF system (Saga Science, MA, USA). Seventeen SMRT cells for *Vampyroteuthis* libraries and two SMRT cells for *A. hians* libraries were sequenced on the PacBio Sequel II/IIe systems with Binding Kit 3.2 and Sequencing Kit 2.0 (30 hours collection times). The consensus (HiFi) reads were generated from raw full-pass subreads using the DeepConsensus v1.1.0 program[41] with the default parameter settings.

For Illumina sequencing, genomic DNA was fragmented to an average size of 500 bp using the Focused-ultrasonicator M220 (Covaris Inc., MA. USA). Paired-end libraries were constructed with a TruSeq DNA PCR-Free Library Prep kit (Illumina, CA, USA) and size-selected on an agarose gel with a Zymoclean Large Fragment DNA Recovery Kit (Zymo Research, CA. USA). The final libraries were sequenced on the NovaSeq 6000 system (Illumina, San Diego, CA, USA) with 2 × 150 bp read length.

Total RNA was extracted from three *Vampyroteuthis tissues* (brain and eyes) and four *A. hians* tissues (brain, eyes, heart, and embryo) using an E.Z.N.A. Mollusc RNA kit (Omega Bio-Tek, GA. USA) and a Nucleospin RNA clean-up XS (TaKaRa, Japan). RNA libraries were constructed using an Illumina Stranded mRNA Prep, Ligation (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Sequencing was

conducted on the NovaSeq 6000 system with 2 x 100 bp read length. The concentration and quality of the libraries were assessed using the Qubit 4 Fluorometer (Thermo Fisher Scientific, MA, USA), the 2100 Bioanalyzer system (Agilent Technologies, CA, USA), and the 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific, MA, USA).

Genome sequencing

The PacBio libraries were constructed using the genomic DNA from the *Vampyroteuthis* sample. In total, 571,504,779,822 bp of HiFi reads were obtained by PacBio Sequel II/IIe. The total sequencing depth was approximately 40X, and the average read length was 16.3 kb. Short reads by illumina NovaSeq6000 (PE500) were also obtained, and 1,303,449,174,900 bp of clean data were acquired. Two RNA-seq libraries were obtained from the same individuals of *Vampyroteuthis*.

We assembled the whole genome by Hifiasm[42,43] with an option --hg-size 14g since preliminary k-mer analysis showed a genome size of 14 Gb (**Supplemental Table 1**). We note that the results of the vampire squid stably reproduced 14G assembly size regardless of options. The v1.0 assembly we showed has a genome size of 14.7 Gbp and an N50 length of 7.6 Mbp. The genome size was 14.5 Gbp and N50 length 7.8 Mbp as a result of option 12G specification (hifiasm --hg-size 12g). Even using the latest version of hifiasm, the genome size was 14.4 Gbp and N50 length 8.5 Mbp with the 12G designation, and the N50 was extended a little further (hifiasm-0.24.0-r702 --hg-size 12g, 2024Dec_ver). Thus, regardless of the genome size setting option, the final assembly size remains the same across the currently available methods. BUSCO analysis

showed that its completeness was as high as 97.2. This score was comparable to other chromosome-level Molluscan genomes (**Supplemental Table 2**).

K-mer based genome size estimation was applied by GenomeScope2.0[44] (-k 35 -p 2). Jellyfish[45] run was performed with an option: high 10000000. The estimated genome size was 11.5 Gb, and heterozygosity was estimated to be 1.8-2% (**Supplemental figure 1**). Another k-mer condition (k=27) gave almost the same result (11.8 Gb). The Hifiasm standard output size is significantly larger than the genome size estimated by k-mer analysis of short-read data. The discrepancy in assembly size is probably the result of duplication due to haplotype diversity. BUSCO analysis also detected a high number of duplicate genes (**Supplemental Table 3**). Based on these findings, contig duplication was suspected. We applied purge_dups[46] to reduce the duplicated content to a haplotype consensus sequence. The purged assembly data (v2.1) set has a size of 11.7 Gb and 5,616 contigs with an N50 contig size of 9.96 Mbp.

GenomeScope 2.0 estimated an expected genome size of 11.5-11.8 Gb, which is close to the assembled size after applying purge_dups. The purge_dups collapses the haplotype-separated assembly and reduces the duplicated content to a haplotype consensus sequence, therefore, contigs purged in this process are probably duplicated contigs due to haplotype diversity. The purged contigs are not enriched with a high mask rate and appear to be a population of contigs derived from minor haplotigs with the same distribution as the whole (Supplemental figure 1). This is also in good agreement with the improved duplicate comparing BUSCO scores before and after purge_dups (7.1% to 3.5%).

Repetitive elements of the Vampyroteuthis genome were estimated with Repeatmodeler[47]. However, LTR structural analysis (LtrHarvest::gt command) was not finished for over a month, probably due to the large-sized genome. So, we applied an alternative method; the LTRs that were not detected by repeatmodeler (-LTRStruct off), and additionally estimated using ltrDetector[48]. This method is practical and finished within a day. 71.4% of the genome was estimated as repetitive regions and masked for the following analysis.

Gene prediction models were generated using GINGER 1.0.1[49]. In brief, this pipeline combines ab initio, RNA-seq-based, and homology-based prediction results for related species. For homology-based prediction, the amino acid sequences of *Octopus vulgaris*[50], *O. bimaculoides*[51], *Architeuthis dux*[52], *Crassostrea gigas*[53], and *Mizuhopecten yessoensis*[54] were utilized and mapped to the genome scaffolds. The predicted genes were evaluated using BUSCO[55], and resulted in 95.3% complete genes being marked, suggesting high accuracy of the annotation [C:95.3%, [S:91.8%, D:3.5%], F:3.2%, M:1.5%].

Genome assembly of *A. hians* was also conducted with standard parameters. We constructed Hi-C libraries for scaffolding using Dovetail Omni-C Kits (Cantata Bio, CA, USA) and sequenced them on an Illumina MiSeq system. The Hi-C reads were mapped to the contigs and filtered using Juicer v1.6[56]. Scaffolding was performed with 3D-DNA[57] v180419, and the resulting scaffolds were manually inspected and refined using Juicebox[58] v1.11.08. For *A. hians*, a repeat mask was followed by an estimation of the gene model. The repeat libraries were generated using RepeatModeler v2.0.5[47] with RECON v1.08[59], RepeatScout v1.0.6[60], and RMBlast v2.9.0

(http://www.repeatmasker.org/RMBlast.html). Long terminal repeat sequences were also identified with "-LTRStruct" option, using LtrHarvest v1.5.9[61] and Ltr_retriever v2.6[62]. Repeat annotation was performed with RepeatMasker v4.1.5[63] using the custom libraries.

Genome annotation was conducted with GINGER v1.0.1[48]. The conda environment was created with Python v3.7, and nextflow v21.10.0[64] was installed using conda. In the homology search of GINGER, the protein database consists of amino acid sequences from six mollusks, *Argonauta argo*[34], *Octopus vulgaris*[49], *O. bimaculoides*[50], *Architeuthis dux*[51], *Crassostrea gigas*[52], and *Mizuhopecten yessoensis*[53], and the parameter files of Spaln are all 'InsectDB'.

Orthology and syntenic analysis

For consistency in orthology handling, we mapped *Doryteuthis pealeii* and *Octopus vulgaris* peptides to all genomes using miniprot[65] 0.12 and using -gff option as an output. In total, 30432 *D. pealeii* and 25682 *O. vulgaris* genes were mapped. Custom scripts were used to parse the mapping results and convert them into pairwise synteny files (.psynt) for each species pair (available on the repository link below). Each .psynt file contains information on mutual best orthologs as well as their genomic locations in two species. Only contigs with 15 or more orthologous genes were used. Dotplots were made using a custom R (version 4.2.2) script psyntPlot.R using published procedures[7]. Significance of chromosomal homologies was tested using Fisher's exact test, and associations of adjusted p-value (Bonferroni correction) of 0.05 are shown as black colored dots. Clustering of contigs or chromosomes on dotplots was done either by

predefined order of chromosomes (Figure 2) to enable multi-species comparisons, or via Euclidean distance measure on the number of shared orthologs and ward.D2 clustering in R. Statistical analysis of mixing was done with TraMineR 2.2-11 package in R[66].

Whole genome alignments were done on hard-masked genomes (RepeatModeler 2.0.6 and RepeatMasker 4.1.8) with blastn 2.16.0+ using the following parameters: -task megablast -perc_identity 0 -template_length 16 -penalty -2 -word_size 11 -evalue 1 -template_type coding_and_optimal[67]. Transdecoder 5.7.1 was used to predict open reading frames on conserved aligned sequences (TransDecoder.LongOrfs -m10 parameter was used) and classify alignments into coding vs non-coding regions, in addition to available gene annotation overlap. Only alignments of 50 or more base pairs were reported. All dotplots for every figure and their supporting data can be generated by running the prepInput.sh and plot_figures.R scripts on the repository.

Acknowledgments

We sincerely thank Noriyoshi Sato (Tokai University) for providing the samples of *Vampyroteuthis* individuals for this study. We also thank all members of the Simakov lab (University of Vienna), in particular Thea Rogers and Darrin Schultz, for invaluable

discussions and comments on the manuscript. D.H.E.S. and K.H. would like to thank Nanami Tochino (Setiamarga Lab at NIT Wakayama), for her assistance on the phylogeny/timetree inference. We thank three anonymous reviewers for their valuable input on this manuscript.

Grants

Genome sequencings of *Vampyroteuthis* sp. and *A*. *hians* were supported by JSPS KAKENHI Grant Number JP22H04925 (Platform for Advanced Genome Science). M.A.Y. was partially supported by Takeda Science Foundation Life Science Research Grants FY 2023, and KAKENHI Grants-in-aid for Basic Research (No. 22K06340) awarded to M.A.Y. and D.H.E.S.. M.A.Y. also thanks the Faculty of Life and Environmental Sciences at Shimane University for the financial support for publishing this report. D.H.E.S. was partially supported by KAKENHI Grants-in-aid for Basic Research (19K12424 and 23K11511), Takeda Science Foundation for Life Sciences Research Grant FY 2022, and the National Institute of Technology GEAR 5.0 Project for Agriculture, Forestry, and Fisheries to support KH's position in Setiamarga lab. K.H. was partially supported by Sasakawa Scientific Research Grant FY 2024 from the Japan Science Society and JST SPRING GX grant (No. 23A360), both of which are fellowships for graduate students. O.S. was supported by the European Research Council's Horizon 2020: European Union Research and Innovation Programme, grant No. 945026. The computational results of this work have been achieved using the Life Science Compute Cluster (LiSC) of the University of Vienna.

Author contributions

This work is a result of an equal collaboration among three labs led by three Principal Investigators (the corresponding authors: M.A.Y., D.H.E.S. and O.S.), all giving their shares equally toward the completion of this project. M.A.Y., D.H.E.S. and O.S. designed and led the study; M.A.Y., A. T., H. T., and H. N. contributed to the genome sequencing and data deposit; E.T., K.K., T. K., K. H., H. M., M. T., and H. N. conducted the bioinformatical work; M.A.Y., D.H.E.S. and O.S. produced the first draft of the manuscript; all authors contributed to the final manuscript.

Declaration of interests

The authors declare no competing interests.

Supplemental information

Figures S1–S8, Table S1-S3

References

1. Hanlon, R., and Messenger, J. (2018). Cephalopod Behaviour 2nd ed. (Cambridge University Press) https://doi.org/10.1017/9780511843600.
2. Tanner, A.R., Fuchs, D., Winkelmann, I.E., Gilbert, M.T.P., Pankey, M.S., Ribeiro, Â.M., Kocot, K.M., Halanych, K.M., Oakley, T.H., da Fonseca, R.R., et al. (2017). Molecular clocks indicate turnover and diversification of modern coleoid

cephalopods during the Mesozoic Marine Revolution. Proc. Biol. Sci. 284. https://doi.org/10.1098/rspb.2016.2818.

3. Kröger, B., Vinther, J., and Fuchs, D. (2011). Cephalopod origin and evolution: A congruent picture emerging from fossils, development and molecules: Extant cephalopods are younger than previously realised and were under major selection to become agile, shell-less predators. Bioessays *33*, 602–613. https://doi.org/10.1002/bies.201100001.
4. Clements, T., Colleary, C., De Baets, K., and Vinther, J. (2017). Buoyancy mechanisms limit preservation of coleoid cephalopod soft tissues in Mesozoic Lagerstätten. Palaeontology *60*, 1–14. https://doi.org/10.1111/pala.12267.
5. Amodio, P., Boeckle, M., Schnell, A.K., Ostojic, L., Fiorito, G., and Clayton, N.S. (2019). Shell Loss in Cephalopods: Trigger for, or By-Product of, the Evolution of Intelligence? A Reply to Mollo et al. Trends Ecol. Evol. *34*, 690–692. https://doi.org/10.1016/j.tree.2019.05.005.
6. López-Córdova, D.A., Avaria-Llautureo, J., Ulloa, P.M., Braid, H.E., Revell, L.J., Fuchs, D., and Ibáñez, C.M. (2022). Mesozoic origin of coleoid cephalopods and their abrupt shifts of diversification patterns. Mol. Phylogenet. Evol. *166*, 107331. https://doi.org/10.1016/j.ympev.2021.107331.
7. Young, R.E., Vecchione, M., and Donovan, D.T. (1998). The evolution of coleoid cephalopods and their present biodiversity and ecology. S. Afr. J. Mar. Sci./S.-Afr. Tydskr. Seewet. *20*, 393–420. https://doi.org/10.2989/025776198784126287.
8. Tanner, A.R., Fuchs, D., Winkelmann, I.E., Gilbert, M.T.P., Pankey, M.S., Ribeiro, Â.M., Kocot, K.M., Halanych, K.M., Oakley, T.H., da Fonseca, R.R., et al. (2017).

Molecular clocks indicate turnover and diversification of modern coleoid cephalopods during the Mesozoic Marine Revolution. Proc. Biol. Sci. *284*. https://doi.org/10.1098/rspb.2016.2818.

9. Benton, M.J., and Wu, F. (2022). Triassic Revolution. Front. Earth Sci. *10*, 899541. https://doi.org/10.3389/feart.2022.899541.
10. Amodio, P., Shigeno, S., and Ostojić, L. (2020). Evolution of Intelligence in Cephalopods. In eLS (Wiley), pp. 77–84. https://doi.org/10.1002/9780470015902.a0029004.
11. Mather, J.A., and Kuba, M.J. (2013). The cephalopod specialties: complex nervous system, learning, and cognition. Can. J. Zool. *91*, 431–449. https://doi.org/10.1139/cjz-2013-0009.
12. Schnell, A.K., Amodio, P., Boeckle, M., and Clayton, N.S. (2021). How intelligent is a cephalopod? Lessons from comparative cognition: How intelligent is a cephalopod? Biol. Rev. Camb. Philos. Soc. *96*, 162–178. https://doi.org/10.1111/brv.12651.
13. Jaitly, R., Ehrnsten, E., Hedlund, J., Cant, M., Lehmann, P., and Hayward, A. (2022). The evolution of predator avoidance in cephalopods: A case of brain over brawn? Front. Mar. Sci. *9*, 909192. https://doi.org/10.3389/fmars.2022.909192.
14. Belcaid, M., Casaburi, G., McAnulty, S.J., Schmidbaur, H., Suria, A.M., Moriano-Gutierrez, S., Pankey, M.S., Oakley, T.H., Kremer, N., Koch, E.J., et al. (2019). Symbiotic organs shaped by distinct modes of genome evolution in cephalopods. Proc. Natl. Acad. Sci. U. S. A. *116*, 3030–3035. https://doi.org/10.1073/pnas.1817322116.

15. Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., and Rokhsar, D.S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524, 220–224. https://doi.org/10.1038/nature14668.
16. Albertin, C.B., Medina-Ruiz, S., Mitros, T., Schmidbaur, H., Sanchez, G., Wang, Z.Y., Grimwood, J., Rosenthal, J.J.C., Ragsdale, C.W., Simakov, O., et al. (2022). Genome and transcriptome mechanisms driving cephalopod evolution. Nat. Commun. 13, 2427. https://doi.org/10.1038/s41467-022-29748-w.
17. Zhang, Y., Mao, F., Mu, H., Huang, M., Bao, Y., Wang, L., Wong, N.-K., Xiao, S., Dai, H., Xiang, Z., et al. (2021). The genome of Nautilus pompilius illuminates eye evolution and biomineralization. Nature Ecology & Evolution, 1–12. https://doi.org/10.1038/s41559-021-01448-6.
18. Huang, Z., Huang, W., Liu, X., Han, Z., Liu, G., Boamah, G.A., Wang, Y., Yu, F., Gan, Y., Xiao, Q., et al. (2021). Genomic insights into the adaptation and evolution of the nautilus, an ancient but evolving “living fossil.” Mol. Ecol. Resour. https://doi.org/10.1111/1755-0998.13439.
19. Schmidbaur, H., Kawaguchi, A., Clarence, T., Fu, X., Hoang, O.P., Zimmermann, B., Ritschard, E.A., Weissenbacher, A., Foster, J.S., Nyholm, S.V., et al. (2022). Emergence of novel cephalopod gene regulation and expression through large-scale genome reorganization. Nat. Commun. *13*, 2172. https://doi.org/10.1038/s41467-022-29694-7.
20. Marino, A., Kizenko, A., Wong, W.Y., Ghiselli, F., and Simakov, O. (2022). Repeat age decomposition informs an ancient set of repeats associated with coleoid

cephalopod divergence. Front. Genet. *13*, 793734. https://doi.org/10.3389/fgene.2022.793734.

21. Styfhals, R., Zolotarov, G., Hulselmans, G., Spanier, K.I., Poovathingal, S., Elagoz, A.M., De Winter, S., Deryckere, A., Rajewsky, N., Ponte, G., et al. (2022). Cell type diversity in a developing octopus brain. Nat. Commun. *13*, 1–17. https://doi.org/10.1038/s41467-022-35198-1.
22. Ritschard, E.A., Fitak, R.R., Simakov, O., and Johnsen, S. (2019). Genomic signatures of G-protein-coupled receptor expansions reveal functional transitions in the evolution of cephalopod signal transduction. Proc. Biol. Sci. *286*, 20182929. https://doi.org/10.1098/rspb.2018.2929.
23. Simakov, O., Bredeson, J., Berkoff, K., Marletaz, F., Mitros, T., Schultz, D.T., O'Connell, B.L., Dear, P., Martinez, D.E., Steele, R.E., et al. (2022). Deeply conserved synteny and the evolution of metazoan chromosomes. Science Advances 8, eabi5884. https://doi.org/10.1126/sciadv.abi5884.
24. Košťák, M., Schlögl, J., Fuchs, D., Holcová, K., Hudáčková, N., Culka, A., Fözy, I., Tomašových, A., Milovský, R., Šurka, J., et al. (2021). Fossil evidence for vampire squid inhabiting oxygen-depleted ocean zones since at least the Oligocene. Commun Biol *4*, 216. https://doi.org/10.1038/s42003-021-01714-0.
25. Rowe, A.J., Kruta, I., Landman, N.H., Villier, L., Fernandez, V., and Rouget, I. (2022). Exceptional soft-tissue preservation of Jurassic Vampyronassa rhodanica provides new insights on the evolution and palaeoecology of vampyroteuthids. Sci. Rep. *12*, 8292. https://doi.org/10.1038/s41598-022-12269-3.

26. Golikov, A.V., Ceia, F.R., Sabirov, R.M., Ablett, J.D., Gleadall, I.G., Gudmundsson, G., Hoving, H.J., Judkins, H., Pálsson, J., Reid, A.L., et al. (2019). The first global deep-sea stable isotope assessment reveals the unique trophic ecology of Vampire Squid Vampyroteuthis infernalis (Cephalopoda). Sci. Rep. *9*, 19099. https://doi.org/10.1038/s41598-019-55719-1.
27. Schlögl, J., Košťák, M., and Hyžný, M. (2012). First record of a gladius-bearing coleoid Teudopsis bollensis Voltz (Cephalopoda, Coleoidea) in the Toarcian of the Western Carpathians (Slovakia). Palaontol. Z. *86*, 367–375. https://doi.org/10.1007/s12542-012-0139-z.
28. Fuchs, D., Keupp, H., and Schweigert, G. (2013). First record of a complete arm crown of the Early Jurassic coleoid Loligosepia (Cephalopoda). Palaontol. Z. *87*, 431–435. https://doi.org/10.1007/s12542-013-0182-4.
29. Fuchs, D., Weis, R., and Thuy, B. (2023). Simoniteuthis, a new vampyromorph coleoid with prey in its arms from the Early Jurassic of Luxembourg. Research Square. https://doi.org/10.21203/rs.3.rs-3331195/v1.
30. Iba et al. 2011 https://pubs.geoscienceworld.org/gsa/geology/article-abstract/39/5/483/130552/Belemnite-extinction-and-the-origin-of-modern
31. Tanabe, K., Misaki, A., Hikida, Y., and Nishimura, T. (2017). New records of coleoid cephalopod jaws from the Upper Cretaceous of Hokkaido, Japan, and their paleobiogeographic and evolutionary implications. Cretac. Res. *70*, 128–141. https://doi.org/10.1016/j.cretres.2016.10.009.

32. Albertin, C.B., and Simakov, O. (2020). Cephalopod Biology: At the Intersection Between Genomic and Organismal Novelties. Annu Rev Anim Biosci 8, 71–90. https://doi.org/10.1146/annurev-animal-021419-083609.
33. Setiamarga, D.H.E., Hirota, K., Yoshida, M.-A., Takeda, Y., Kito, K., Ishikawa, M., Shimizu, K., Isowa, Y., Ikeo, K., Sasaki, T., et al. (2021). Hydrophilic Shell Matrix Proteins of Nautilus pompilius and the identification of a core set of Conchiferan domains. Genes (Basel) *12*, 1925. https://doi.org/10.3390/genes12121925.
34. Yoshida, M.A., Hirota, K., Imoto, J., Okuno, M., Tanaka, H., Kajitani, R., Toyoda, A., Itoh, T., Ikeo, K., Sasaki, T., et al. (2022). Gene recruitments and dismissals in the argonaut genome provide insights into pelagic lifestyle adaptation and shell-like eggcase reacquisition. Genome Biol. Evol. https://doi.org/10.1093/gbe/evac140.
35. Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., and Rokhsar, D.S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524, 220–224. https://doi.org/10.1038/nature14668.
36. Schultz, D.T., Haddock, S.H.D., Bredeson, J.V., Green, R.E., Simakov, O., and Rokhsar, D.S. (2023). Ancient gene linkages support ctenophores as sister to other animals. Nature 618, 110–117. https://doi.org/10.1038/s41586-023-05936-6.
37. Mackintosh, A., Vila, R., Martin, S.H., Setter, D., and Lohse, K. (2023). Do chromosome rearrangements fix by genetic drift or natural selection? Insights from Brenthis butterflies. Mol. Ecol. https://doi.org/10.1111/mec.17146.
38. Ritschard, G. (2023). Measuring the nature of individual sequences. Sociol. Methods Res. *52*, 2016–2049. https://doi.org/10.1177/00491241211036156.

39. Coffing GC, Tittes S, Small ST, Songco-Casey JO, Piscopo DM, Pungor JR, Miller AC, Niell CM, Kern AD. Cephalopod sex determination and its ancient evolutionary origin. Curr Biol. 2025 Feb 24;35(4):931-939.e4. doi: 10.1016/j.cub.2025.01.005. Epub 2025 Feb 3. PMID: 39904337; PMCID: PMC11859755.
40. Meyer, A., Schloissnig, S., Franchini, P., Du, K., Woltering, J.M., Irisarri, I., Wong, W.Y., Nowoshilow, S., Kneitz, S., Kawaguchi, A., et al. (2021). Giant lungfish genome elucidates the conquest of land by vertebrates. Nature *590*, 284–289. https://doi.org/10.1038/s41586-021-03198-8.
41. Baid, G., Cook, D.E., Shafin, K., Yun, T., Llinares-López, F., Berthet, Q., Belyaeva, A., Töpfer, A., Wenger, A.M., Rowell, W.J., et al. (2023). DeepConsensus improves the accuracy of sequences with a gap-aware sequence transformer. Nat. Biotechnol. 41, 232–238. https://doi.org/10.1038/s41587-022-01435-7.
42. Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods 18, 170–175. https://doi.org/10.1038/s41592-020-01056-5.
43. Cheng, H., Jarvis, E.D., Fedrigo, O., Koepfli, K.-P., Urban, L., Gemmell, N.J., and Li, H. (2022). Haplotype-resolved assembly of diploid genomes without parental data. Nat. Biotechnol. 40, 1332–1335. https://doi.org/10.1038/s41587-022-01261-x.
44. Ranallo-Benavidez, T.R., Jaron, K.S., and Schatz, M.C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat. Commun. 11, 1432. https://doi.org/10.1038/s41467-020-14998-3.

45. Marçais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27, 764–770. https://doi.org/10.1093/bioinformatics/btr011.
46. Guan, D., McCarthy, S.A., Wood, J., Howe, K., Wang, Y., and Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics 36, 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025.
47. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. U. S. A. 117, 9451–9457. https://doi.org/10.1073/pnas.1921046117.
48. Valencia, J.D., and Girgis, H.Z. (2019). LtrDetector: A tool-suite for detecting long terminal repeat retrotransposons de-novo. BMC Genomics 20, 450. https://doi.org/10.1186/s12864-019-5796-9.
49. Taniguchi, T., Okuno, M., Shinoda, T., Kobayashi, F., Takahashi, K., Yuasa, H., Nakamura, Y., Tanaka, H., Kajitani, R., and Itoh, T. (2023). GINGER: an integrated method for high-accuracy prediction of gene structure in higher eukaryotes at the gene and exon level. DNA Res. 30, dsad017. https://doi.org/10.1093/dnares/dsad017.
50. Zarrella, I., Herten, K., Maes, G.E., Tai, S., Yang, M., Seuntjens, E., Ritschard, E.A., Zach, M., Styfhals, R., Sanges, R., et al. (2019). The survey and reference assisted assembly of the Octopus vulgaris genome. Sci Data 6, 13. https://doi.org/10.1038/s41597-019-0017-6.

51. Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., and Rokhsar, D.S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524, 220–224. https://doi.org/10.1038/nature14668.
52. da Fonseca, R.R., Couto, A., Machado, A.M., Brejova, B., Albertin, C.B., Silva, F., Gardner, P., Baril, T., Hayward, A., Campos, A., et al. (2020). A draft genome sequence of the elusive giant squid, Architeuthis dux. Gigascience 9. https://doi.org/10.1093/gigascience/giz152.
53. Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., et al. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49–54. https://doi.org/10.1038/nature11413.
54. Wang, S., Zhang, J., Jiao, W., Li, J., Xun, X., Sun, Y., Guo, X., Huan, P., Dong, B., Zhang, L., et al. (2017). Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat Ecol Evol 1, 120. https://doi.org/10.1038/s41559-017-0120.
55. Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
56. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 3, 95–98. https://doi.org/10.1016/j.cels.2016.07.002.

57. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., et al. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science 356, 92–95. https://doi.org/10.1126/science.aal3327.
58. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. 3, 99–101. https://doi.org/10.1016/j.cels.2015.07.012.
59. Bao, Z., and Eddy, S.R. (2002). Automated de novo identification of repeat sequence families in sequenced genomes. Genome Res. *12*, 1269–1276. https://doi.org/10.1101/gr.88502.
60. Price, A.L., Jones, N.C., and Pevzner, P.A. (2005). De novo identification of repeat families in large genomes. Bioinformatics *21 Suppl 1*, i351-8. https://doi.org/10.1093/bioinformatics/bti1018.
61. Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinformatics *9*, 18. https://doi.org/10.1186/1471-2105-9-18.
62. Ou, S., and Jiang, N. (2018). LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. Plant Physiol. *176*, 1410–1422. https://doi.org/10.1104/pp.17.01310.
63. Smit, A., Hubley, R., and Green, P. (2013). RepeatMasker Open-4.0. http://www.repeatmasker.org.

64. Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nat. Biotechnol. *35*, 316–319. https://doi.org/10.1038/nbt.3820.

65. Li, H. (2023). Protein-to-genome alignment with miniprot. Bioinformatics *39*. https://doi.org/10.1093/bioinformatics/btad014.

66. Gabadinho, A., Ritschard, G., Müller, N.S., and Studer, M. (2011). Analyzing and visualizing state sequences inRwithTraMineR. J. Stat. Softw. *40*, 1–37. https://doi.org/10.18637/jss.v040.i04.

67. Simakov, O., Kawashima, T., Marlétaz, F., Jenkins, J., Koyanagi, R., Mitros, T., Hisata, K., Bredeson, J., Shoguchi, E., Gyoja, F., et al. (2015). Hemichordate genomes and deuterostome origins. Nature *527*, 459–465. https://doi.org/10.1038/nature16150.

Figures and figure legends

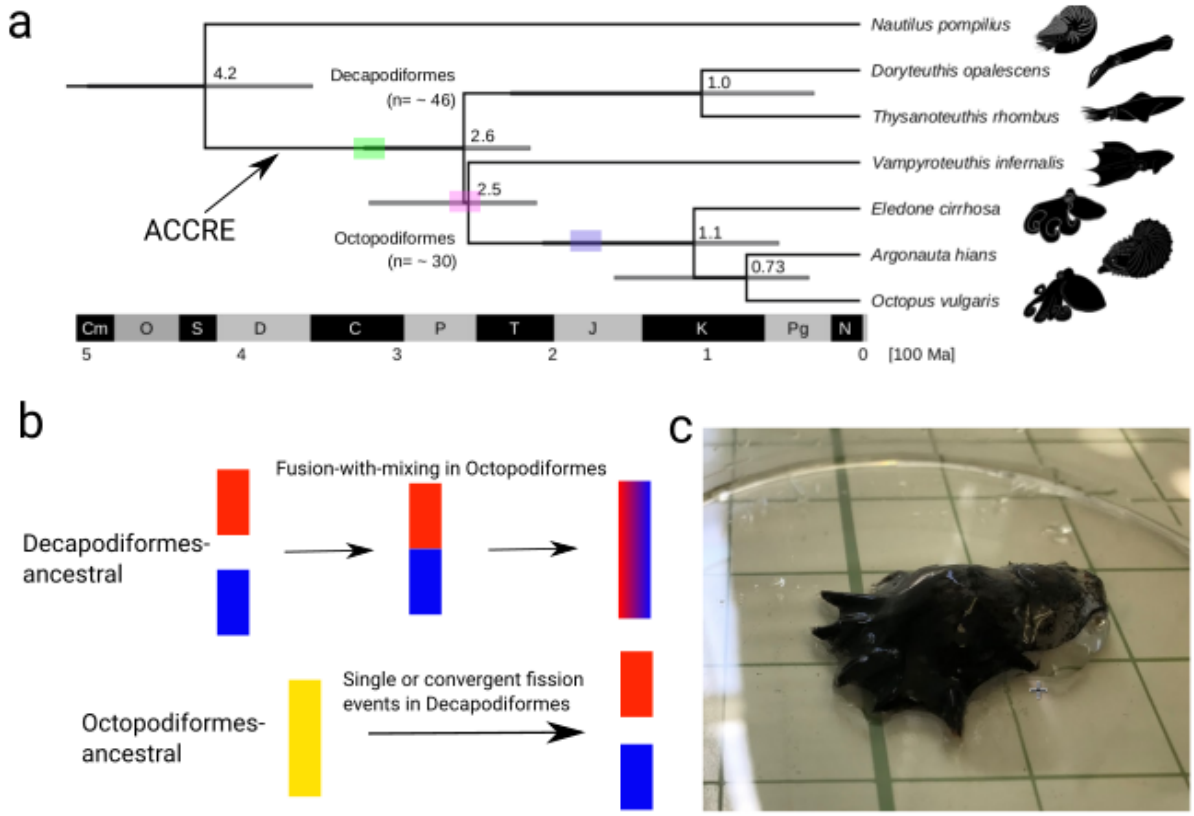

Figure 1. Early origins of coleoid cephalopods and conflicting chromosome evolution scenarios. (a) Phylogenetic time tree of mitochondrial genomes with the major coleoid cephalopod clades and their closest outgroups. Node branches are labelled according to the major cephalopod evolutionary transitions. (b) Two conflicting scenarios how coleoid karyotypes evolved, reduction and expansion (via fission) scenarios. Fusion-with-mixing (FWM) provides a powerful synapomorphic character to profile irreversible changes. (c) *Vampyroteuthis* specimen used for sequencing. ACCRE: ancient coleoid chromosomal rearrangement event.

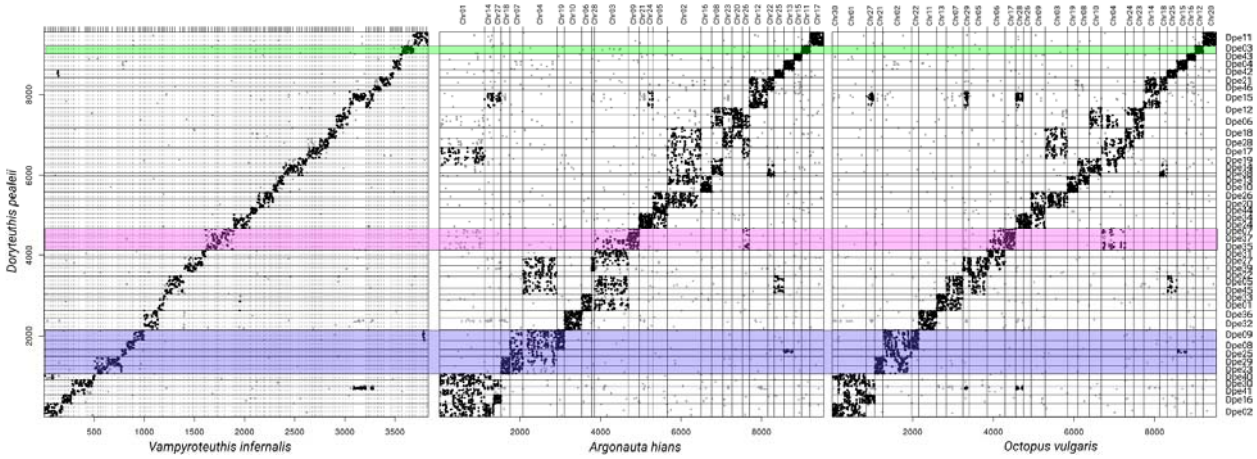

Figure 2. *Vampyroteuthis* genome suggests ancestral coleoid karyotype was decapodiform-like, followed by karyotype reduction in Octopodiformes. Synteny dotplots showing Octopodiformes and *Argonauta-Octopus* shared FWM characters, relative to the *Doryteuthis pealeii* chromosomes. *Doryteuthis* chromosomes are ordered in the same way across the three species comparisons. The color codings indicate representative examples of the ancestral coleoid chromosome (green), ancestral Octopodiformes mixed chromosome (pink), and fused-and-mixed chromosomes in octopods (blue), respectively (corresponding to the phylogenetic nodes highlighted in Figure 1).

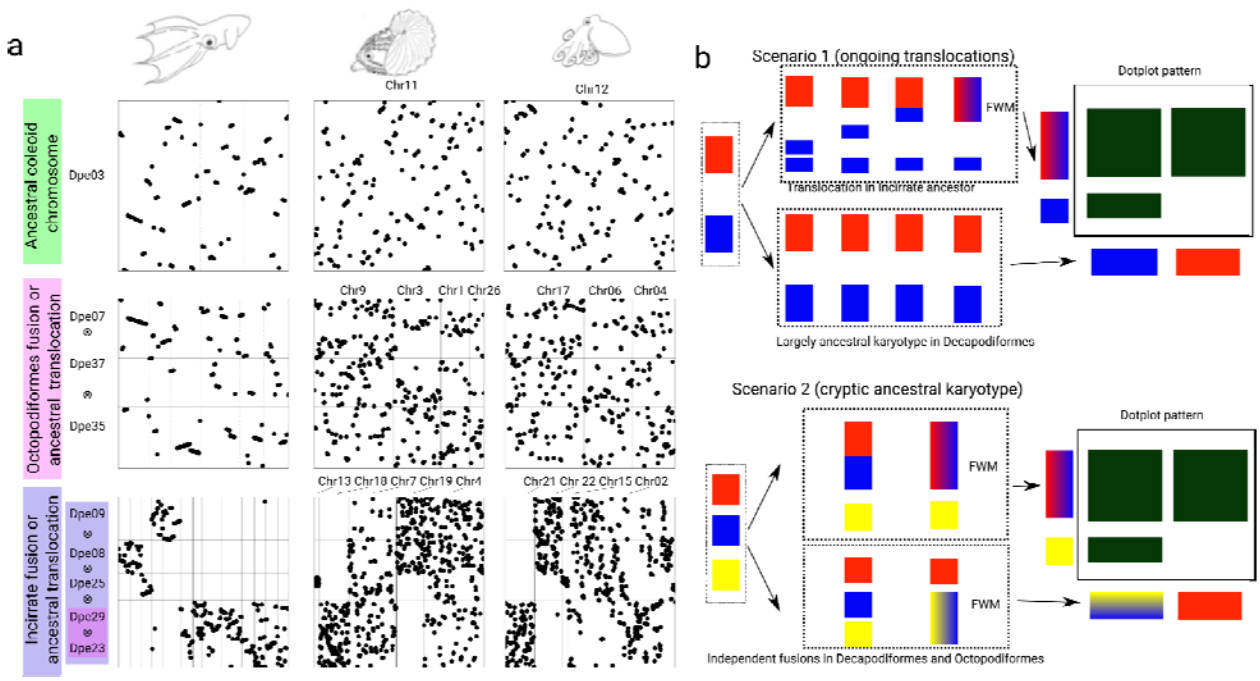

Figure 3. Chromosomal evolutionary pattern suggests additional ancestral translocations in octopods. (a) Dotplot representation for the three cases highlighted in Figure 1, with the Y axis indicating the chromosomes of *Doryteuthis*. Green indicates conserved coleoid chromosomes, as suggested by the conservation of genes in the chromosomes of all compared coleoids, despite translocation of the genes' location within the chromosomes (intrachromosomal rearrangements; "mixing in a chromosome"). Meanwhile, ancestral Octopodiformes (pink) indicate fusions of different *Doryteuthis* chromosomes, which happened in ancestral Octopodiformes. The third example (blue) shows the formation of complex patterns of fusion-with-mixing, where orthologous genes undergoing intrachromosomal translocations in specific regions of the chromosomes of *Doryteuthis* are present in multiple chromosomes in compared species. (b) Two main scenarios for the observed complex chromosomal evolutionary patterns. The first one involves a fusion-with-mixing of intact (in red) and partial (in blue)

chromosomes following inter-chromosomal translocations (blue) that occurred in Octopodiformes after divergence from the decapodiform karyotypes. A second possible scenario suggests that following the split of Octopodiformes and Decapodiformes, each lineage experienced distinct fusion-with-mixing events involving multiple ancestral chromosomes. In Octopodiformes, two specific chromosomes fused and mixed (red and blue), while in Decapodiformes, a different pair of chromosomes (blue and yellow) underwent a similar fusion-with-mixing process.

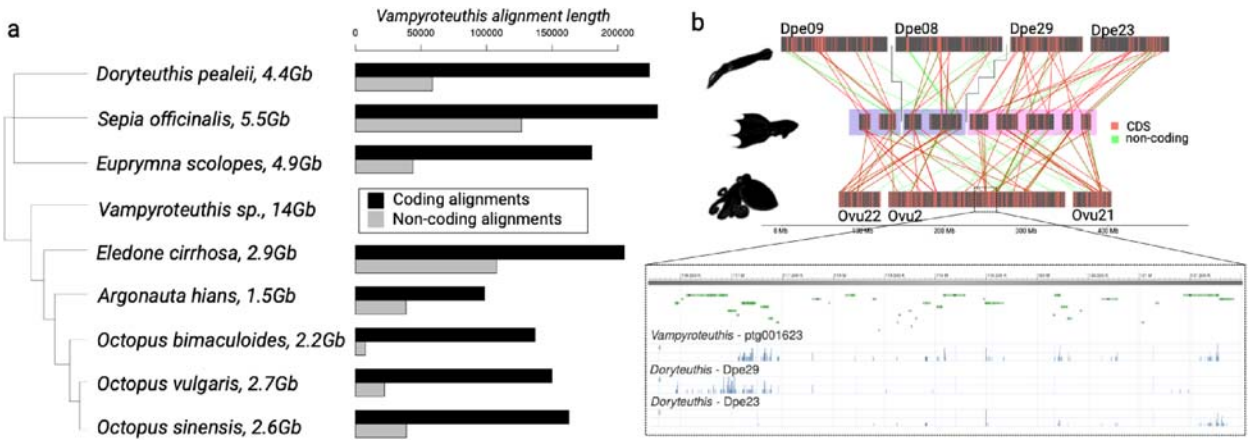

Figure 4. Conserved non-coding element complement in coleoid genomes. (a) *Vampyroteuthis*-centered whole genome alignments show the total content (in alignment numbers) of coding and non-coding alignments. (b) Ribbon diagram showing location of homologous alignments of coding (red, CDS) and non-coding (green) regions between *Doryteuthis* chromosomes (top) to *Vampyroteuthis* (middle) and *Octopus vulgaris* (bottom) for the complex FWM shown in Figures 2 and 3 (pink and blue colors). Zoom-in on a middle portion of *O. vulgaris* chromosome 2 (Ovu2) is highlighted below (NCBI genome browser) with gene track of annotation and conservation of both coding and non-coding elements derived from one homologous *Vampyroteuthis* contig and its pre-octopodiform unmixed state on two chromosomes Dpe29 and Dpe23 (pink color). Two *Vampyroteuthis* contigs homologous to Dpe09 and Dpe08, on the other hand, remain unmixed and their homologs undergo FWM only in octopods (lila color).