

REVIEW ARTICLE OPEN ACCESS

Transforming Life Science Through Chromosome-Level Genome Assemblies

Tetsuo Kon¹  | Kosuke Kataoka^{2,3}  | Yi-Jyun Luo⁴  | Johannes Nicolaus Wibisana⁵  | Kouhei Toga^{6,7}  | Narumi Uno⁸  | Yuki Monden⁹ | Mayuko Hamada¹⁰ 

¹Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, Austria | ²Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan | ³Comprehensive Research Organization, Waseda University, Tokyo, Japan | ⁴Biodiversity Research Center, Academia Sinica, Taipei, Taiwan | ⁵Genomics and Regulatory Systems Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa, Japan | ⁶Laboratory of BioDX, PtBio Co-Creation Research Center, Genome Editing Innovation Center, Hiroshima University, Hiroshima, Japan | ⁷Laboratory of Genome Informatics, Graduate School of Integrated Sciences for Life, Hiroshima University, Hiroshima, Japan | ⁸Laboratory of Bioengineering, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan | ⁹Graduate School of Environmental, Life, Natural Science, and Technology, Okayama University, Okayama, Japan | ¹⁰Ushimado Marine Institute, Okayama University, Okayama, Japan

Correspondence: Tetsuo Kon (kont34@univie.ac.at) | Kosuke Kataoka (kataokak@go.tuat.ac.jp) | Mayuko Hamada (hamadam@okayama-u.ac.jp)

Received: 17 April 2026 | **Revised:** 15 July 2026 | **Accepted:** 7 August 2026

Transmitting Editor: Tadashi Uemura

Keywords: chromosome-level genome assembly | genome evolution | Hi-C scaffolding | long-read sequencing | non-model animals

ABSTRACT

Advances in long-read sequencing and Hi-C scaffolding have made chromosome-level genome assembly increasingly accessible to individual laboratories, shifting genome research from large consortium-led projects toward investigator-driven studies across diverse taxa. This transition allows researchers to select organisms based on biological questions rather than the prior availability of genomic resources. In this review, we summarize the core experimental and computational steps for generating, evaluating, and annotating chromosome-level assemblies, and examine how they have advanced research in non-model organisms and genetically complex systems. Representative case studies illustrate four major contributions: resolving structural variation and lineage-specific genome architecture, linking genome organization to phenotypic innovation and plasticity, reconstructing deep chromosome evolution and macrosynteny, and distinguishing homologous and homoeologous chromosomes in polyploid genomes. These examples show that chromosome-level assemblies provide more than complete reference sequences. They establish a continuous genomic coordinate system through which genes, regulatory elements, transposable element insertions, sequence variants, and cellular states can be interpreted within broader chromosomal, population, and evolutionary contexts. We describe this integrative perspective as “glocal biology.” Future progress will require pangenomic, population-scale, and haplotype-resolved resources integrated with multi-omics and functional analyses. Collectively, chromosome-level genomics is reshaping life science by embedding molecular functions within chromosomal and evolutionary contexts.

1 | Transitioning From Consortium-Based Genomics to Investigator-Driven Chromosome-Level Genomics

Early chromosome-scale genome sequencing was exemplified by the Human Genome Project (HGP), which required

extensive international collaboration, clone-based sequencing, and physical mapping to reconstruct genome organization (Collins et al. 2003; International Human Genome Sequencing Consortium 2004; Lander et al. 2001). Because of the substantial infrastructure, cost, and coordination required, chromosome-level assemblies were initially restricted

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Genes to Cells* published by Molecular Biology Society of Japan and John Wiley & Sons Australia, Ltd.

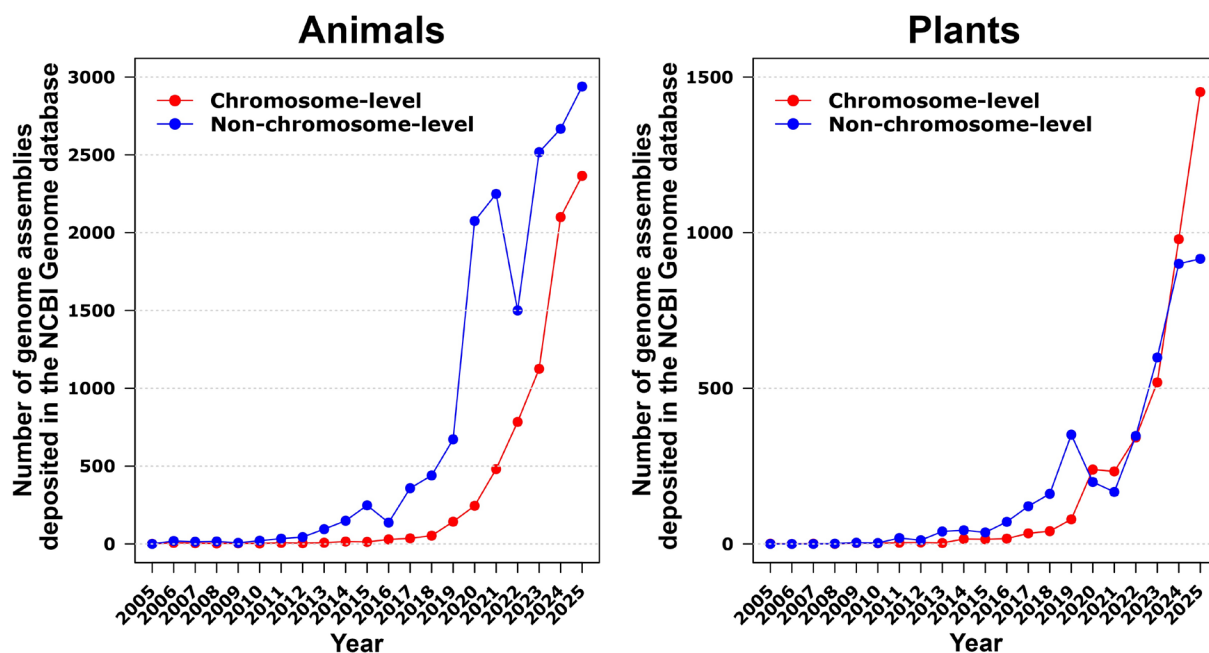


FIGURE 1 | The rapid accumulation of chromosome-level genome assemblies. These plots show the annual number of genome assemblies deposited in the NCBI Genome database (2005–2025) for animals (left) and plants (right). Red indicates chromosome-level assemblies, and blue indicates non-chromosome-level (contig-level and scaffold-level) assemblies. Note the exponential growth of chromosome-level assemblies since around 2018.

to a small number of intensively studied model organisms. The subsequent introduction of short-read sequencing greatly increased throughput and expanded the number of available draft genomes. However, short reads generally could not resolve long repetitive elements, segmental duplications, and structural variants, and therefore often produced fragmented assemblies (van Dijk et al. 2018; Mardis 2008; Pollard et al. 2018).

A fundamental transition occurred with the development of long-read sequencing and proximity ligation-based scaffolding. Long reads can span repetitive and structurally complex regions, allowing the construction of highly contiguous sequences, whereas Hi-C provides long-range linkage information that enables contigs to be ordered, oriented, and assigned to chromosomes (Burton et al. 2013; Kaplan and Dekker 2013). Their combination has made chromosome-level genome assembly increasingly accessible to individual laboratories and small research groups (Ito et al. 2025; Kon-Nanjo et al. 2023; Kon et al. 2025) (Figure 1, Table 1). Consequently, chromosome-level genome assembly has shifted from a predominantly centralized, consortium-driven model to an investigator-driven approach applicable across the tree of life.

This shift allows researchers to select organisms on the basis of distinctive biological traits and evolutionary questions rather than the prior availability of genomic resources. More importantly, chromosome-level assemblies provide a continuous genomic coordinate system in which local molecular features, such as genes, regulatory elements, transposable element insertions, and sequence variants, can be interpreted within broader chromosomal, regulatory, and evolutionary contexts. Together with advances in genetic manipulation and multi-omics approaches, this makes it possible to investigate how local gene functions and cell-type-specific regulatory states are embedded

within global chromosome architecture. We refer to this integrative framework as “glocal biology.” Here, the term does not denote a separate subdiscipline but describes an approach that connects multiple layers of biological information through chromosome-resolved genome coordinates. In this sense, glocal biology complements comparative genomics, regulatory genomics, 3D genome biology, and evolutionary genomics by integrating their findings across cell types, populations, species, and evolutionary timescales. In this review, we first summarize the core workflow required to generate and evaluate chromosome-level genome assemblies. We then highlight representative studies showing how chromosome-level assemblies have connected local genomic findings with broader chromosome-scale contexts in non-model organisms across diverse taxa, before discussing their implications for genome evolution and experimental chromosome-scale genomics.

2 | Technological Foundations of Chromosome-Level Genome Assembly

2.1 | Core Steps of Modern Genome Assembly

Despite the diversity of sequencing platforms and analytical tools currently available, most modern genome assembly projects follow a common conceptual framework. At its core, chromosome-level genome assembly can be understood as a three-step process: contig construction, scaffolding, and gap filling or polishing (Figure 2). In practice, however, these steps are often iterative and may be interleaved, with polishing or gap filling occurring both during and after contig construction depending on the assembly strategy and data types used.

Before any computational assembly begins, high-molecular-weight (HMW) DNA extraction constitutes a foundational

TABLE 1 | Numbers of chromosome-level genome assemblies.

Taxonomic group	Number of chromosome-level genome assemblies ^a	Examples of species with chromosome-level genome assemblies
Metazoa (animals)	7435	
Bilateria/Deuterostomia	3239	<i>Homo sapiens</i> (Nurk et al. 2022), <i>Protopterus annectens</i> (Schartl et al. 2024), <i>Ambystoma mexicanum</i> (Nowoshilow et al. 2018), <i>Garra rufa</i> (Kon et al. 2025), <i>Oikopleura dioica</i> (Bliznina et al. 2021), <i>Branchiostoma floridae</i> (Simakov et al. 2020)
Bilateria/Protostomia	3995	<i>Gryllus bimaculatus</i> (Kataoka, Sanno, et al. 2026), <i>Apis cerana</i> (Lee et al. 2025), <i>Octopus vulgaris</i> (Destanović et al. 2023), <i>Placopecten magellanicus</i> (Xie et al. 2023), <i>Alentia gelatinosa</i> (Adkins et al. 2023)
Cnidaria	116	<i>Hydra vulgaris</i> (Cazet et al. 2023; Kon-Nanjo et al. 2025; Simakov et al. 2022), <i>Hydractinia symbiolongicarpus</i> (Kon-Nanjo et al. 2023), <i>Nanomia septata</i> (Ahuja et al. 2026), <i>Nematostella vectensis</i> (Zimmermann et al. 2023), <i>Acropora palmata</i> (Locatelli et al. 2024)
Placozoa	0	Not available
Porifera	76	<i>Spongilla lacustris</i> (Leys et al. 2025), <i>Diacarnus erythraeanus</i> (Steindler et al. 2025)
Ctenophora	5	<i>Hormiphora californensis</i> (Schultz et al. 2021), <i>Bolinopsis microptera</i> (Schultz et al. 2023), <i>Mnemiopsis leidyi</i> (Kim et al. 2025)
Viridiplantae (green plants)	3987	
Spermatophyta	3695	<i>Triticum aestivum</i> (Walkowiak et al. 2020), <i>Glycine max</i> (Zhang, Shao, et al. 2024), <i>Solanum lycopersicum</i> (Li et al. 2023), <i>Citrullus lanatus</i> (Zhang, Zhao, et al. 2024), <i>Coffea arabica</i> (Salojärvi et al. 2024), <i>Gossypium hirsutum</i> (Hu et al. 2019)

^aThe numbers represent the total number of chromosome-level genome assemblies deposited in the NCBI Genome Database by the end of 2025. Subcategories shown are not exhaustive.

prerequisite for successful long-read-based genome assembly (Figure 2). Regardless of sequencing platform, DNA integrity directly constrains achievable read length and therefore the ability to resolve repetitive and structurally complex genomic regions. Equally important is minimizing contamination from symbionts, microbiota, or environmental DNA, as contaminant sequences can complicate assembly graphs and compromise downstream analyses. In practice, however, complete removal of symbiont-derived DNA is not always feasible, particularly in organisms with intimate host–microbe associations, and such sequences are often identified and separated at the contig or scaffold level during downstream analyses. Routine quality control using electrophoresis systems such as TapeStation (Agilent Technologies), which assesses size distribution and fragmentation, and fluorometers such as Qubit (Thermo Fisher Scientific) and Quantus (Promega), which provide accurate double-stranded DNA quantification, allows researchers to evaluate whether extracted DNA meets the requirements for long-read sequencing. In practice, careful optimization of sample handling and extraction protocols often has a greater impact on downstream assembly outcomes than subsequent computational choices.

The first computational step, contig construction, aims to reconstruct continuous genomic sequences directly from sequencing reads (Figure 2). In contemporary workflows, this step is typically driven by long-read sequencing technologies, which can span repetitive and structurally complex regions that previously fragmented assemblies. Contigs may be generated using long-read-only approaches or hybrid strategies that combine long reads with short reads. In diploid, hybrid, or polyploid genomes, haplotype-aware assembly and strategies such as trio binning can separate homologous chromosome sequences, thereby reducing heterozygosity-induced fragmentation and enabling analyses of allelic and subgenomic variation.

The second step, scaffolding, focuses on ordering and orienting contigs into larger structures that correspond to whole chromosomes (Figure 2). Proximity ligation-based methods, most notably Hi-C, provide long-range interaction information that reflects the three-dimensional organization of chromosomes in the nucleus. This information enables contigs to be linked across megabase scales, forming chromosome-level scaffolds. Reference-guided scaffolding methods can also be useful when

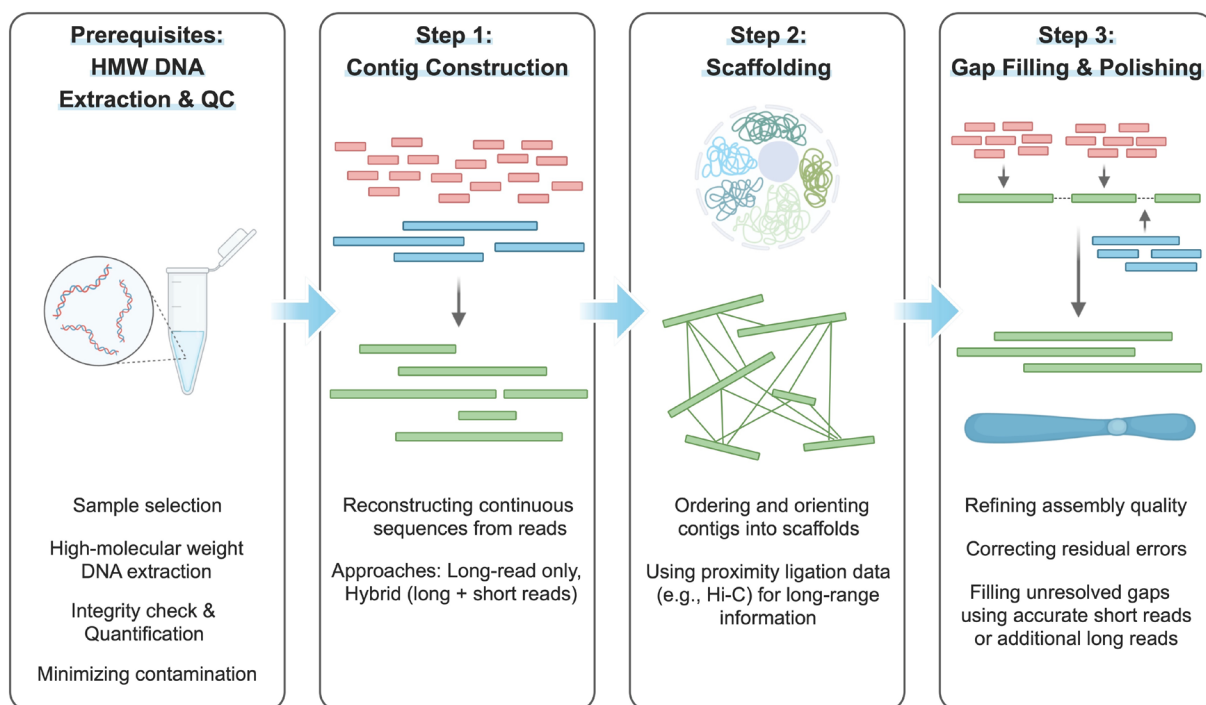


FIGURE 2 | Conceptual workflow of modern chromosome-level genome assembly Chromosome-level genome assembly proceeds through a series of coordinated experimental and computational steps. The process begins with the extraction and quality control of high-molecular-weight (HMW) DNA, where sample integrity, quantification, and contamination minimization are critical for maximizing long-read performance. In Step 1, contigs (green bars) are constructed by assembling sequencing reads into continuous sequences, typically using long-read-based or hybrid assembly strategies incorporating long reads (blue bars) and short reads (red bars). Step 2 employs proximity ligation-derived information, such as Hi-C contact data, to order and orient contigs into chromosome-level scaffolds by leveraging long-range genomic interactions. Finally, Step 3 focuses on gap filling and polishing, where additional long reads or high-accuracy short reads are used to correct residual errors and improve assembly completeness. Together, these stages illustrate how advances in sequencing and scaffolding technologies enable the reconstruction of genome assemblies that approximate native chromosomal organization. This figure was created in BioRender (<https://BioRender.com/f8nu3xi>).

chromosome-level assemblies of closely related species are available, particularly in organisms with relatively conserved karyotypes.

Finally, gap filling and polishing refine the assembly by correcting residual errors and filling unresolved gaps (Figure 2). This step typically leverages the high accuracy of short reads or additional long-read data to improve base-level quality and overall completeness. Together, these three steps form the technological backbone of modern chromosome-level genome assembly.

2.2 | Long-Read Sequencing as the Foundation of Modern Genome Assembly

Long-read sequencing has become the backbone of modern genome assembly. By producing reads that span repetitive elements and structurally complex regions, long-read technologies allow complex regions to be resolved by physically linking sequences that short reads cannot bridge. This capability is particularly critical for repeat-rich or large genomes, in which transposable elements, segmental duplications, and long introns routinely exceed the length of short reads.

Two major long-read technologies are currently used for genome assembly: single-molecule real-time sequencing, implemented

by Pacific Biosciences (PacBio), and nanopore sequencing (Goodwin et al. 2016). PacBio HiFi sequencing generates long reads with high per-base accuracy by repeatedly sequencing circularized DNA molecules. Nanopore sequencing detects changes in ionic current as individual DNA molecules pass through a nanoscale pore and can generate ultra-long reads extending to hundreds of kilobases or more. Nanopore sequencing can often be performed within individual laboratories, providing flexibility for pilot experiments and protocol optimization, whereas PacBio sequencing is more commonly conducted at centralized facilities and is particularly well suited for accurate contig construction and haplotype-aware assembly.

Long-read assemblers convert these data into contiguous genome sequences. Hifiasm is widely used for accurate and haplotype-resolved assembly of PacBio HiFi data (Cheng et al. 2021), whereas Flye supports both PacBio and nanopore reads across diverse genome architectures (Kolmogorov et al. 2019). Canu remains an established option, particularly for noisy long-read datasets (Koren et al. 2017). Additional assemblers, including wtdbg2 (Ruan and Li 2020), NextDenovo (Hu et al. 2024), Verkko (Rautiainen et al. 2023), and NOVOloci (Dierckxsens et al. 2025), provide alternative strategies suited to different read types, genome structures, and levels of heterozygosity. In parallel, hybrid assembly approaches that combine long reads with high-accuracy short reads continue to be used. Assemblers

such as MaSuRCA (Zimin et al. 2013), the hybrid modes of SPAdes (Prijbelski et al. 2020), and graph-based methods such as WENGAN (Di Genova et al. 2021) integrate complementary data types to improve assembly contiguity and accuracy. The optimal platform and assembly strategy ultimately depend on genome architecture, DNA quality and quantity, and the required balance among read length, accuracy, contiguity, and haplotype resolution.

2.3 | Hi-C Scaffolding: From Contigs to Chromosomes

Hi-C-based scaffolding enables chromosome-level genome assembly by converting three-dimensional chromatin proximity information into one-dimensional genomic order (Figure 2). Genomic regions located near one another on the same chromosome generally interact more frequently than regions on different chromosomes or at distant positions, allowing Hi-C contact frequencies to be used to infer contig adjacency, orientation, and chromosomal assignment. In contact maps, correctly assembled chromosomes typically form dense interaction blocks, whereas misassemblies produce discontinuities or discordant interaction patterns. Hi-C, therefore, serves not only as a scaffolding method but also as an important quality-control approach for detecting and correcting assembly errors.

Several proximity ligation methods and computational tools are available for this purpose. Classical Hi-C relies on restriction enzyme digestion of crosslinked chromatin, whereas commercial implementations such as Arima Genomics optimize enzyme combinations for more uniform coverage. Alternative approaches include Chicago libraries, which use in vitro reconstituted chromatin, and Omni-C, which employs endonuclease-based fragmentation independent of restriction enzyme recognition sites. Computational pipelines commonly use Juicer to generate normalized contact maps (Durand, Shamim, et al. 2016), followed by scaffolders such as 3D-DNA (Dudchenko et al. 2017) or YaHS (Zhou et al. 2023) to construct chromosome-level assemblies. When Hi-C data are unavailable, reference-guided tools such as RagTag can provide an alternative scaffolding strategy (Alonge et al. 2022). Finally, interactive visualization with Juicebox remains important for manual curation and the resolution of complex assembly structures (Durand, Robinson, et al. 2016).

2.4 | Assembly Evaluation and Genome Annotation

Finally, chromosome-level genome assemblies must be evaluated for both structural continuity and biological completeness. Assembly quality is commonly assessed using scaffold- and contig-level statistics together with gene space completeness metrics provided by tools such as BUSCO (Tegenfeldt et al. 2025), compleasm (Huang and Li 2023), and OMArk (Nevers et al. 2025), which benchmark assemblies against conserved ortholog sets and expected gene architectures. Base-level accuracy and completeness can be further examined using k-mer-based approaches such as Merqury (Rhie et al. 2020), which compare sequencing reads and assembled genomes without requiring reference sequences. Optical mapping can provide independent,

ultra-long-range validation of scaffold structure, helping to confirm chromosome-scale organization and identify potential misassemblies. In parallel, contamination screening helps identify sequences derived from symbionts, microbiota, or environmental sources that may otherwise confound downstream analyses. Dedicated tools such as FCS-GX (Astashyn et al. 2024) and BlobToolKit (Challis et al. 2020) support taxonomic assignment and coverage-based visualization, facilitating the detection and removal of contaminant sequences in complex assemblies. Integrated QC pipelines such as GenomeQC (Manchanda et al. 2020) can also be useful for streamlining evaluation.

Beyond assembly evaluation, accurate gene prediction and annotation are essential for downstream biological analyses but remain challenging in non-model organisms, particularly when transcriptomic resources, curated proteins, or closely related reference genomes are limited. Recent pipelines such as NCBI EGAPx (Goldfarb et al. 2025) and BRAKER (Gabriel et al. 2024) integrate protein homology, RNA-seq evidence, and ab initio prediction. For species with well-annotated reference genomes, annotation transfer tools such as Liftoff (Shumate and Salzberg 2021) and LiftOn (Chao et al. 2025), as well as comparative annotation frameworks such as TOGA (Kirilenko et al. 2023), provide complementary approaches for projecting gene models onto newly assembled genomes. However, organisms with unusual gene architectures, including trans-splicing, very short introns, or compact intergenic regions, may still require species-specific optimization. Annotation strategies should also reflect the intended downstream applications. Cap analysis of gene expression (CAGE) data can accurately identify transcription start sites, thereby improving the annotation of 5' transcript ends and enabling the inference of alternative promoter usage, whereas long-read RNA sequencing can resolve full-length transcript isoforms and improve the annotation of 3' untranslated regions. Such improvements are especially important for single-cell RNA-seq because incomplete gene models can reduce read assignment and obscure cell-type-specific expression patterns. Functional annotation is also challenging for lineage-specific genes because experimentally characterized homologs are often absent from existing databases.

3 | Case Studies: Conceptual Advances Enabled by Chromosome-Level Genome Assemblies

The following case studies cover representative taxa across the tree of life (Figure 3) and are organized into four conceptual categories based on the biological questions addressed by chromosome-level genome assemblies. First, chromosome-resolved genomes reveal structural variation and lineage-specific genome architecture that are obscured in fragmented genome assemblies. Second, they provide a framework for linking genome organization to phenotypic innovation and developmental plasticity. Third, comparisons of chromosome-level assemblies enable the reconstruction of deep chromosome evolution and conserved macrosynteny. Finally, chromosome-level assemblies make it possible to analyze genetically complex systems, including highly heterozygous and polyploid genomes. Together, these examples illustrate how chromosome-level genome assemblies provide biological insights that cannot be obtained from gene inventories or fragmented draft genomes alone.

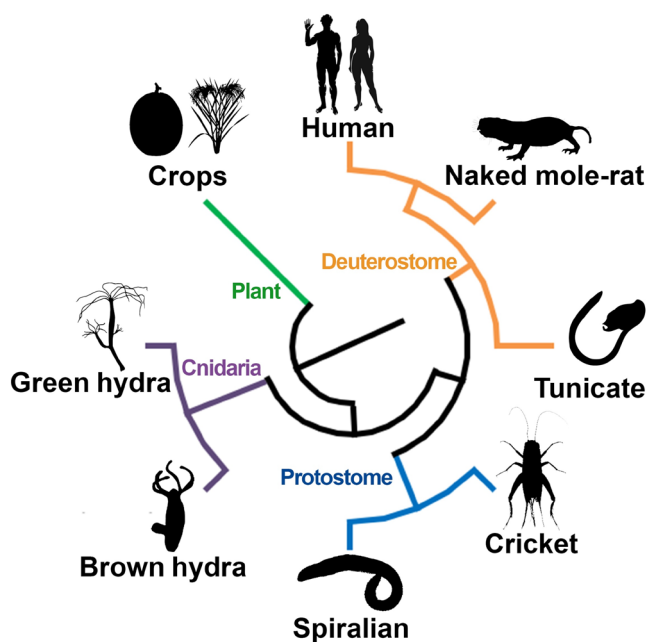


FIGURE 3 | Cladogram of representative taxa in which chromosome-level genome assemblies have advanced molecular and evolutionary analyses. Animal silhouettes were obtained from the PhyloPic database (<https://www.phylopic.org/>). The cladogram was constructed using MEGA12 software (Kumar et al. 2024).

3.1 | Resolving Structural Variation and Lineage-Specific Genome Architecture

3.1.1 | Genome Scrambling and Hidden Lineage Diversity in Larvaceans

Larvaceans (appendicularians) are tunicates and, therefore, represent an important comparative reference lineage for understanding chordate genome evolution before vertebrate-specific whole-genome duplications (Dehal and Boore 2005; Donoghue and Purnell 2005). Among them, *Oikopleura dioica* is particularly informative because it combines an extremely compact genome with extensive lineage-specific gene loss, gene duplication, and chromosome rearrangement (Albalat and Cañestro 2016; Ferrández-Roldán et al. 2019; Martí-Solans et al. 2021). This raises the question of how morphologically similar lineages can retain a conserved chordate body plan despite profound differences in genome content and chromosome organization.

The first *O. dioica* reference genome revealed an unusually compact genome of approximately 60 Mb containing more than 15,000 genes, widespread operons, and highly reduced introns (Denoeud et al. 2010; Ganot et al. 2004). Subsequent long-read and chromosome-level assemblies from geographically distinct populations revealed that *O. dioica* is not a single genetically uniform species but a complex of deeply divergent lineages (Bliznina et al. 2021; Masunaga et al. 2022; Plessy et al. 2024; Wang et al. 2020). Whole-genome comparisons identified extensive differences in genome content and chromosome organization despite limited morphological differentiation, supporting the existence of multiple cryptic lineages and raising fundamental questions concerning

species boundaries and phenotypic conservation (Masunaga et al. 2022; Plessy et al. 2024).

Accurate interpretation of these differences depends on assemblies that preserve native chromosome structure. *O. dioica* presents several technical challenges, including low DNA yield, uneven heterozygosity between chromosome arms, frequent inversions, and long homopolymeric sequences in the mitochondrial genome (Bliznina et al. 2021; Dierckxsens et al. 2024; Plessy et al. 2024). Fragmented assemblies and reference-guided scaffolding may obscure genuine inversions and other lineage-specific rearrangements. NOVOloci has recently been used with Oxford Nanopore data, even from earlier R9 chemistry, to generate near-complete or complete chromosome arms as single contigs, reducing reliance on reference-guided scaffolding and enabling more direct comparisons of chromosome structure among lineages (Dierckxsens et al. 2025).

Chromosome-level assemblies have therefore transformed *O. dioica* research from the analysis of a single compact reference genome into the comparative study of structurally divergent genomes. However, currently available assemblies do not yet capture the full geographic diversity of the species complex, particularly in the Southern Hemisphere. Existing resources, including ANISEED, TUNOME, and lineage-specific genome databases, remain distributed across separate platforms (Brozovic et al. 2018; Shito et al. 2025). The development of integrated pangenomic and multi-omics resources will be essential for determining how chromosome-scale rearrangements reshape local gene content and regulation, and how these changes are integrated to preserve a conserved body plan while permitting lineage-specific adaptation.

3.1.2 | Transposable Element-Driven Genome Expansion and Stem-Cell Genome Diversification in *Hydra*

The genomes of brown hydra, including *Hydra vulgaris*, are unusually large for cnidarians and are strongly enriched in transposable elements. The first *H. vulgaris* genome was generated through a large international consortium using clone-based Sanger sequencing and consisted of more than 20,000 sequence fragments (Chapman et al. 2010). Although this fragmented assembly revealed substantial expansion of chicken repeat 1 retrotransposons, it could not fully resolve the chromosomal distribution, copy-number structure, or lineage-specific activity of transposable elements.

Subsequent chromosome-level and telomere-to-telomere assemblies, together with single-cell and epigenomic datasets, have provided a chromosome-resolved view of transposable element evolution in *Hydra* (Cazet et al. 2023; Kon-Nanjo et al. 2025; Siebert et al. 2019; Simakov et al. 2022). These assemblies revealed that genome expansion is not attributable to a single repeat family but involves multiple anciently active transposable element families, including DNA transposons such as hAT-Charlie and TcMar-Tc1. They further showed that the three continuously self-renewing stem-cell lineages of *Hydra* have accumulated distinct transposable element insertions during prolonged asexual propagation,

and suggested that some of these insertions may influence lineage-specific gene expression (Kon-Nanjo et al. 2025).

These findings depended on highly contiguous reference genomes because transposable elements and their insertion sites cannot be reliably assigned to genomic context in fragmented assemblies. Chromosome-level and telomere-to-telomere assemblies made it possible to distinguish genuine lineage-specific insertions from assembly fragmentation, compare repeat distributions across entire chromosomes, and investigate how transposable element activity contributes to both genome expansion and somatic genome diversification. Together, the *Oikopleura* and *Hydra* examples illustrate how chromosome-level assemblies connect large-scale genome restructuring, including inversions and repeat expansion, with local changes in gene content, transposable element insertions, and gene regulation, thereby revealing how lineage-specific genome architecture may shape biological variation.

3.2 | Linking Genome Organization to Phenotypic Innovation and Plasticity

3.2.1 | Genome Architecture and the Evolution of Photosymbiosis in Green Hydra

Hydra can serve not only as a model for the transposon-driven evolution of genome architecture described above, but also as a model for understanding how interspecies interactions shape phenotypic and genomic evolution. *Hydra viridissima*, commonly known as green hydra, forms a stable intracellular photosymbiosis with the green alga *Chlorella*. The symbiont resides within endodermal epithelial cells and supplies photosynthetically derived carbohydrates to the host, whereas the host supplies nitrogen-containing metabolites to the alga (Davy et al. 2012). Unlike brown hydra, which possess genomes of approximately 1Gb enriched in transposable elements, the green *Hydra* genome is approximately 300Mb and contains a substantially smaller proportion of retrotransposons (Hamada et al. 2020; Wong et al. 2019). This contrast provides a comparative system for studying how genome architecture, transposable element expansion, and symbiotic innovation evolved within a single genus.

Short-read genome assemblies of *H. viridissima* strain A99 and its symbiotic *Chlorella* enabled initial analyses of gene content and metabolic interdependence (Hamada et al. 2018, 2020). Photosynthetically produced maltose is associated with increased expression of host genes involved in nutrient assimilation and transport, including glutamine synthetase and phosphate transporters. Conversely, the symbiotic alga lacks several genes required for autonomous inorganic nitrogen assimilation and instead depends on host-derived amino acids.

Chromosome-level genomes are now becoming available for green *Hydra*, including an assembly from a US population (NCBI Genome accession: GCF_964215635.1), with additional resources for strain A99 under development. These assemblies will allow direct chromosome-scale comparisons between green and brown hydra, including the reconstruction of lineage-specific transposable element expansions, structural rearrangements, and regulatory landscapes. They will also provide more

accurate gene models for single-cell and spatial transcriptomic analyses, which are particularly important for distinguishing host and symbiont transcription within individual cell types. The green *Hydra* example therefore extends chromosome-level genomics from the description of genome size differences to mechanistic questions concerning photosymbiosis, species diversification, and phenotypic evolution.

3.2.2 | Structural and Intraspecific Genomic Variation Underlying Naked Mole-Rat Phenotypes

The naked mole-rat, *Heterocephalus glaber*, exhibits a distinctive combination of exceptional longevity, cancer resistance, hypoxia tolerance, reduced thermoregulation, and eusociality. These traits make it a valuable comparative system for identifying genomic changes associated with physiological innovations that are not readily modeled in conventional laboratory rodents (Buffenstein 2008; Buffenstein and Pinto 2009).

Initial genomic studies identified several species-specific gene changes potentially associated with these phenotypes. These include substitutions in the telomere-associated protein TERF1, alterations in the thermogenic gene *UCP1*, natural loss-of-function mutations in the necroptosis genes *RIPK3* and *MLKL*, and naked mole-rat-specific substitutions in *cGAS* that enhance homologous recombination repair (Chen et al. 2025; Oka et al. 2022). Such findings demonstrate the value of gene-level comparative analyses, but they do not reveal how larger-scale genomic changes, including structural variants, tandem duplications, and population-specific variation, contribute to naked mole-rat phenotypes.

Independent sequencing of naked mole-rats maintained in Japan revealed substantial intraspecific variation and highlighted limitations of relying on a single reference genome (Toga et al. 2026). Comparative analyses identified 77 transcripts with clear variation, including genes encoding amine ligand-binding receptors such as *Htr2b* and *Gpr143*. Population divergence among naked mole-rats had previously been detected using mitochondrial markers, with some geographically separated groups approaching levels of divergence observed between recognized species of bathyergid rodents (Faulkes et al. 1997; Ingram et al. 2015). Chromosome-level assemblies from multiple populations will therefore be necessary to determine whether structural and allelic variants contribute to differences in physiology, behavior, or disease resistance.

More recent long-read chromosome-level assemblies have resolved genomic regions that were poorly represented in short-read references (Sokolowski et al. 2024). Thousands of repetitive elements were newly resolved, and the olfactory receptor gene repertoire was found to span approximately 17Mb rather than the 13Mb represented in the previous reference. The assembly also identified tandem duplications involving biologically relevant genes, including *TINF2*, which is associated with telomere biology and hypoxia tolerance, and *TCPI1*, which functions in protein folding and has been implicated in tumor progression. These structural features could not be reliably identified in fragmented short-read assemblies.

The naked mole-rat example therefore demonstrates that chromosome-level genomics extends phenotype-oriented comparative studies beyond individual coding mutations. By resolving gene-family expansions, tandem duplications, repetitive regions, and population-specific structural variation, chromosome-level assemblies provide a more complete framework for testing how exceptional physiological traits emerge and vary within a species.

3.2.3 | Chromosomal and Epigenetic Regulation of Developmental Fate Plasticity in Crickets

Crickets belong to Polyneoptera, an insect lineage that is phylogenetically distant from the holometabolous species that dominate classical insect genetics. They exhibit diverse environmentally responsive traits, including changes in wing morphology, reproduction, metabolism, stress tolerance, and developmental timing (Bonfoey et al. 2023; Harrison 1979; Kasumovic et al. 2016). These traits provide an opportunity to investigate how a single genome produces alternative phenotypes through context-dependent regulation.

The band-legged ground cricket *Dianemobius nigrofasciatus* is a particularly informative system because maternal photoperiod determines whether genetically similar embryos enter diapause or proceed through direct development (Masaki and Sugahara 1992; Shiga and Numata 1997). Diapause-destined and nondiapause embryos are initially morphologically indistinguishable but later diverge when nondiapause embryos initiate germ-band formation and diapause-destined embryos arrest at the cellular blastoderm stage (Tanigawa et al. 2009). The central question is therefore not which genes are present, but how alternative regulatory states are organized across the same genome before visible phenotypic divergence.

A highly contiguous genome assembly has enabled the gene regulatory landscape to be mapped and interpreted across continuous genomic regions. Whereas earlier studies relied primarily on transcriptome comparisons, integration of the chromosome-level assembly with ATAC-seq has revealed coordinated changes in chromatin accessibility at loci associated with metabolism, cell-cycle regulation, and neural development (Kataoka, Shimizu, et al. 2026). Importantly, these chromatin-state differences arise before strong transcriptional divergence becomes apparent.

This finding illustrates a limitation of purely gene-centric explanations of developmental plasticity. Diapause and nondiapause embryos share the same genomic sequence, yet local regulatory elements are activated or repressed in coordinated patterns across contiguous genomic regions. Chromosome-level assemblies are required to connect these regulatory elements to neighboring genes, repetitive sequences, and broader chromosomal context, thereby revealing how genome organization supports large-scale and reversible changes in developmental fate.

Chromosome-level references also improve functional testing. CRISPR/Cas9-based editing has become increasingly feasible in crickets, and a continuous reference genome improves guide

RNA design, target specificity, and genome-wide assessment of potential off-target sites (Matsuoka et al. 2025; Shimizu and Kataoka 2026; Shirai et al. 2022). Thus, the cricket case study illustrates how chromosome-level assemblies connect environmental cues and local regulatory changes with their broader genomic context and organismal developmental outcomes in a non-model insect.

Together, the green hydra, naked mole-rat, and cricket case studies illustrate how chromosome-resolved genomes connect local molecular changes and regulatory states to organism-level phenotypes.

3.3 | Reconstructing Deep Chromosome Evolution and Macrosynteny

Chromosome-level genome assemblies have revealed that chromosome architecture can remain conserved across hundreds of millions of years, even when gene sequences and organismal morphology have diverged extensively. Orthologous genes retained on the same chromosomes define conserved macrosyntenic blocks, and comparisons across metazoans have enabled the reconstruction of 24 ancestral linkage groups in the last common bilaterian ancestor (Simakov et al. 2013, 2020, 2022). These ancestral linkage groups provide a framework for identifying chromosome fusions, fissions, and fusion-with-mixing events across deep evolutionary time.

Spiralians are particularly informative for such analyses because they include numerous animal phyla with contrasting chromosome architectures. Chromosome-level assemblies from mollusks, annelids, nemerteans, bryozoans, and brachiopods have shown that many spiralian lineages retain substantial components of the ancestral bilaterian karyotype (Liao et al. 2023; Martín-Zamora, Davies, et al. 2023). For example, the scallop *Pecten maximus*, the nemertean *Lineus longissimus*, and the annelid *Owenia fusiformis* retain most ancestral linkage groups as separate chromosomes (Kenny et al. 2020; Kwiatkowski et al. 2021; Martín-Zamora, Liang, et al. 2023). Four fusion-with-mixing events shared across lophotrochozoans reduced the inferred ancestral chromosome number from 24 to 20 (Simakov et al. 2022).

Comparative chromosome analyses have also distinguished distinct modes of chromosome evolution, corresponding to low-rearrangement lineages and high-rearrangement lineages. In low-rearrangement lineages, ancestral chromosomes are primarily combined through progressive fusion. In high-rearrangement lineages, repeated fusion and fission events dismantle ancestral architecture (Lewin et al. 2024; Lewin, Liao, Luo 2025). The incidence of chromosome fission is a key feature distinguishing these two patterns of chromosome evolution.

Clitellate annelids represent an extreme example of genome restructuring. Whereas many marine annelids retain an ancestral complement of approximately 20 chromosomes, clitellate genomes show extensive dispersal of genes from all 24 ancestral bilaterian linkage groups across their chromosomes (Lewin et al. 2024). This pattern is consistent with a rapid episode of chromosome rearrangement rather than the gradual accumulation

of isolated events. Losses of genes involved in genome stability and cell division, transposable element expansion, putative neo-centromere formation, and whole-genome duplication may all have contributed to this architectural disruption (Jin et al. 2020; Vargas-Chávez, Benítez-Álvarez, et al. 2025; Vargas-Chávez, McLysaght, et al. 2025).

Extensive restructuring has also been identified in Bryozoa, a phylum for which chromosome-scale genome organization was previously unknown. Analysis of five chromosome-level genomes revealed that ancestral bilaterian linkage groups were extensively fused and rearranged before the diversification of living bryozoans, reducing the ancestral complement to approximately six linkage groups (Lewin, Liao, Chen, et al. 2025). Subsequent lineage-specific fissions and rearrangements occurred in both gymnolaemate and phylactolaemate bryozoans, demonstrating that chromosome fission has played a larger role in bilaterian evolution than previously recognized.

Chromosome rearrangements can also serve as phylogenetic characters. Fusion-with-mixing events involve the interspersal of genes from two ancestral chromosomes and are unlikely to arise independently in the same configuration. Such rare genomic changes have helped resolve the position of ctenophores at the base of the animal tree and branching relationships among teleost fishes (Parey et al. 2023; Schultz et al. 2023). Within Spiralia, shared derived chromosome fusions between bryozoans, brachiopods, and phoronids support the monophyly of Lophophorata and place phoronids and bryozoans as sister groups (Lewin, Liao, Chen, et al. 2025; Lewin, Sakagami, Kao, et al. 2025).

These studies demonstrate that chromosome-level assemblies contribute more than improved gene annotation. They reveal conserved ancestral chromosome units, identify irreversible structural changes, and provide phylogenetic information that complements sequence-based analyses. As chromosome-level assemblies accumulate across the tree of life, integration of macrosynteny, regulatory genomics, and epigenetic data will help explain why chromosome architecture remains stable in some lineages but undergoes rapid reorganization in others.

3.4 | Enabling Genomics in Complex and Polyploid Genomes

Polyploid genomes contain more than two sets of chromosomes and are common among crop species, including cotton (4×), potato (4×), bread wheat (6×), sweetpotato (6×), and strawberry (8×). Their analysis is challenging because homologous and homoeologous chromosomes contain highly similar sequences, while allelic dosage and subgenome-specific variation must be distinguished. Fragmented assemblies can collapse related chromosome copies, obscure structural variation, and prevent accurate assignment of genes and markers to individual subgenomes.

Historically, genomic analyses of polyploid crops often relied on diploid relatives. For example, the diploid *Fragaria vesca* genome was initially used as a reference for octoploid cultivated strawberry, and diploid *Ipomoea trifida* served as a reference

for hexaploid sweetpotato (Hirakawa et al. 2015; Li et al. 2019; Shulaev et al. 2011; Wu et al. 2018). However, diploid relatives do not necessarily contain the genes, structural variants, or allelic combinations underlying important traits in cultivated polyploids.

Long-read sequencing and Hi-C scaffolding have enabled direct chromosome-level reconstruction of polyploid crop genomes. A phased chromosome-level assembly of hexaploid sweetpotato, for example, resolved the architecture of individual chromosome copies and provides a substantially improved reference for breeding and genetic analysis (Wu et al. 2025). Similar advances in other polyploid crops have enabled subgenome assignment, detection of structural variation, and comparative analysis of chromosome evolution.

Chromosome-level references are also essential for dosage-aware genetic analysis. In polyploids, a biallelic locus can occur in multiple dosage states rather than as a single heterozygous genotype. Polyploid-specific methods therefore estimate allele dosage from sequencing depth and incorporate this information into linkage mapping and genome-wide association studies. Such approaches have been used to identify loci associated with disease resistance in hexaploid sweetpotato and to develop markers for breeding (Obata et al. 2022; Yamamoto et al. 2020). Tools such as MAPpoly additionally model the large number of possible gametic combinations and recombination patterns in polyploid populations (Mollinari et al. 2020).

The principal contribution of chromosome-level assemblies in polyploid genomics is the ability to distinguish homologous and homoeologous chromosome copies, assign variants to the correct genomic background, and connect genotype dosage to phenotype. These capabilities transform polyploid genomes from incompletely resolved collections of related sequences into chromosome-resolved resources that can support association mapping, marker-assisted selection, and genomic breeding.

4 | Future Perspectives: Synthesizing Lessons From Chromosome-Level Genomics

Taken together, the case studies reviewed here highlight several general lessons. First, chromosome-level assemblies reveal structural features that cannot be reliably resolved or interpreted from fragmented assemblies, including large inversions, tandem duplications, chromosome-wide transposable element landscapes, ancestral linkage groups, and relationships among homologous or homoeologous chromosomes. Second, genes, regulatory elements, sequence variants, and copy-number differences can be interpreted more accurately when their positions and organization within chromosomes are known. Third, chromosome-level assemblies become most informative when they are integrated with comparative, population genomic, epigenomic, and functional data. Finally, these integrated analyses generate hypotheses concerning genome structure, regulation, and evolution that can increasingly be tested experimentally. These lessons point toward four related future directions: moving from individual reference genomes to comparative chromosome resources, connecting local molecular findings with global

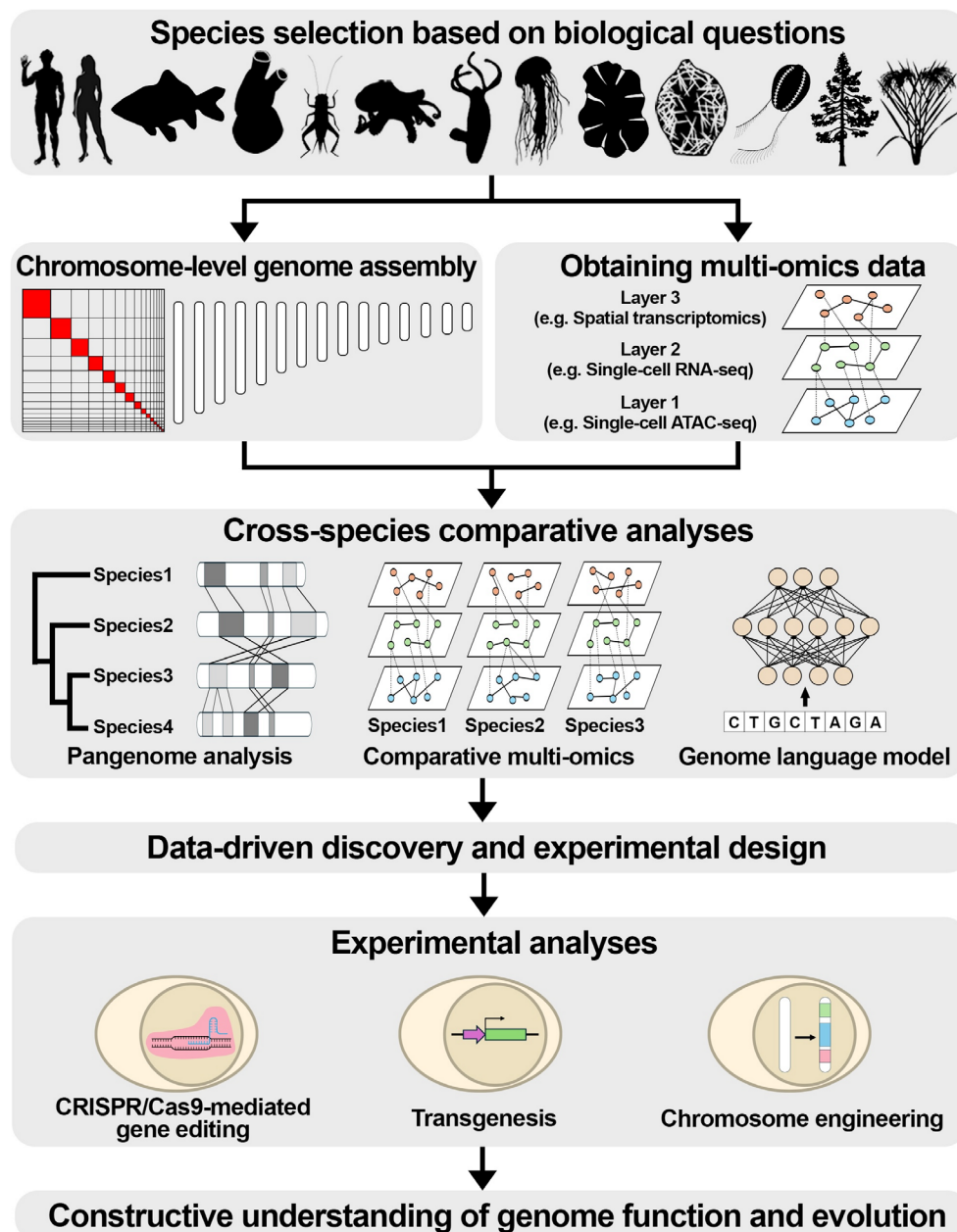


FIGURE 4 | From chromosome-level genome assembly to experimental chromosome-scale genomics The integration of chromosome-level genome assemblies with multi-omics data and pangenome analyses, together with genome language models, facilitates data-driven discovery and experimental design. Subsequent experimental validations, including CRISPR/Cas9-mediated gene editing, transgenesis, and chromosome engineering, lead to a constructive understanding of genome function and evolution. The animal silhouettes were obtained from the PhyloPic database (<https://www.phylopic.org/>).

chromosome organization, extending chromosome-scale comparisons to evolutionary inference, and experimentally testing chromosome-scale hypotheses (Figure 4).

4.1 | Beyond a Single Reference Genome: Capturing Genomic Diversity Within and Among Species

The increasing accessibility of long-read sequencing and Hi-C scaffolding has enabled individual laboratories to generate chromosome-level assemblies for species selected on the basis

of biological questions rather than the prior availability of genomic resources. The next step is to move beyond the generation of a single high-quality reference for each species and to develop comparative genomic resources that represent variation among individuals, populations, lineages, and related species.

The case studies discussed in this review illustrate why this transition is necessary. Comparisons among geographically distinct *Oikopleura* lineages revealed extensive differences in genome content and chromosome organization despite limited morphological divergence. Studies of naked mole-rats

showed that a single reference genome cannot fully represent population-specific sequence and structural variation. Comparisons between green and brown *Hydra* provide a framework for reconstructing lineage-specific transposable element expansion, genome-size evolution, and species diversification. In polyploid crops, multiple homologous and homoeologous chromosome copies must be distinguished to assign sequence variants and allele dosage to the correct genomic background.

Pangenomic, population-scale, and haplotype-resolved approaches will therefore become increasingly important. In highly variable, heterozygous, or polyploid genomes, a single linear reference may obscure structural variants, allelic differences, and subgenome-specific features. Generating chromosome-level assemblies for more individuals, populations, lineages, and related species will make it possible to distinguish conserved genome organization from individual, population, or lineage-specific variation. The future value of chromosome-level genomics will thus depend not only on the completeness of individual assemblies but also on their integration into comparative genomic resources that capture diversity within and among species.

4.2 | From Local Findings to Glocal Biology

Chromosome-level assemblies provide a common coordinate system through which local molecular observations can be interpreted within broader genomic contexts. The case studies reviewed here illustrate this principle at several levels. Transposable element insertions in *Hydra* can be related to chromosome-wide repeat distributions and lineage-specific gene regulation. Chromatin accessibility changes in crickets can be connected to neighboring genes, regulatory regions, and developmental programs. Structural variants and tandem duplications in naked mole-rats can be evaluated within population-specific chromosome architectures. In polyploid crops, allele dosage and sequence variation can be assigned to the appropriate homologous or homoeologous chromosome copies.

These examples define the practical basis of glocal biology. The term does not describe a separate field, but an integrative approach in which genes, regulatory elements, sequence variants, and cellular states are analyzed within chromosome-resolved structural and evolutionary frameworks. Chromosome-level assemblies make this integration possible by providing continuous genomic coordinates shared across gene annotation, epigenomic profiling, population variation, comparative genomics, and functional experiments.

Future studies should therefore combine chromosome-level assemblies with single-cell and spatial transcriptomics, chromatin accessibility and 3D genome data, population genomics, and phenotypic analyses. Such integration will allow cell-type-specific regulatory states and local molecular changes to be connected with organism-level phenotypes and evolutionary differences. A central question will be not only where genes and regulatory elements are located, but also how their chromosomal context constrains or enables their regulation,

variation, and evolutionary change. Addressing this question will help distinguish genomic features that merely record chromosome history from those that actively contribute to phenotypic innovation and plasticity.

4.3 | From Chromosome-Scale Comparisons to Evolutionary Inference

Chromosome-level comparisons provide information about evolutionary history that cannot be recovered reliably from individual genes or fragmented genome sequences alone. The *Oikopleura*, *Hydra*, and spiralian case studies show that chromosome architecture may either remain deeply conserved or undergo extensive lineage-specific reorganization. Comparisons of these contrasting patterns make it possible to reconstruct ancestral genome organization and identify the evolutionary processes that produced present-day chromosomes.

Ancestral linkage groups provide a framework for tracing chromosome fusions, fissions, and fusion-with-mixing events over hundreds of millions of years (Simakov et al. 2020, 2022). Because some of these rearrangements are unlikely to arise independently in the same configuration, they can also serve as rare genomic characters for resolving difficult phylogenetic relationships. Their application to early animal evolution and spiralian phylogeny demonstrates that chromosome structure can provide evidence complementary to conventional analyses based on sequence alignments (Lewin, Liao, Chen, et al. 2025; Lewin, Sakagami, Kao, et al. 2025; Schultz et al. 2023).

The next challenge is to move from reconstructing rearrangement histories to explaining why chromosome architecture remains stable in some lineages but changes rapidly in others. Denser taxonomic and population sampling will be required to determine when rearrangements occurred and whether they accumulated gradually or during short episodes of genome restructuring. Integration with information on transposable elements, centromeres, whole-genome duplication, chromatin organization, and gene regulation will also be necessary to identify the mechanisms and biological consequences of chromosome change. In this way, chromosome-scale comparisons can extend observations from individual species into broader models of genome evolution, phylogenetic history, and the constraints that shape chromosome architecture (Simakov and Wagner 2025).

4.4 | Toward Experimental Chromosome-Scale Genomics

Building on the comparative and glocal frameworks described above, the increasing accessibility of chromosome-level genome assemblies, together with advances in single-cell genomics and artificial intelligence tools such as genome language models (Consens et al. 2025; Merchant et al. 2025), is enabling local gene functions and cell-type-specific regulatory states to be interpreted within the context of global chromosome architecture. This integration provides not only new perspectives on genome function and evolution but also experimentally testable

hypotheses concerning the regulation and organization of extensive genomic regions. Chromosome-level genome assemblies can therefore serve not only as reference coordinates for interpreting genome function and evolution but also as design templates for experimentally testing hypotheses at the scale of large genomic regions.

Conventional genome editing generally targets individual genes or relatively short intervals, whereas chromosome-level analyses increasingly reveal structural variants, gene clusters, repetitive regions, centromeres, and long-range regulatory landscapes extending over hundreds of kilobases or megabases. Telomere-to-telomere assemblies have further resolved structurally complex regions that were inaccessible in earlier reference genomes, creating opportunities to move from reading chromosome sequences toward reconstructing and manipulating chromosome-scale structures (Altemose et al. 2022; Nurk et al. 2022).

Large genomic regions can be generated through chemical synthesis or isolated from existing genomes using long-fragment cloning methods, such as transformation-associated recombination cloning in yeast (Kouprina and Larionov 2016). Human and mouse artificial chromosomes provide additional platforms for maintaining megabase-sized genomic regions as independently segregating elements with functional centromeres and telomeres (Kouprina et al. 2013). These artificial chromosomes can be generated through bottom-up approaches, in which centromeric DNA is assembled de novo, or top-down approaches, in which native chromosomes are engineered while retaining their endogenous centromeres. Recent advances have enabled the generation of structurally defined human artificial chromosomes, while CRISPR/Cas9-based approaches have expanded chromosome engineering to include chromosome truncation, reciprocal translocation, and megabase-scale deletion (Egawa et al. 2025; Gambogi et al. 2024; Miyamoto et al. 2024; Uno et al. 2017). Engineered chromosomes can also be transferred between cells through methods such as microcell-mediated chromosome transfer, facilitating the establishment of cellular and animal models carrying large genomic regions (Kazuki et al. 2022; Uno et al. 2024).

Although chromosome engineering remains technically demanding and has so far been developed most extensively in mammalian cells and animal models, it represents a promising future extension of chromosome-level genomics. The case studies discussed in this review raise hypotheses concerning structural variation, transposable element expansion, long-range gene regulation, chromosome rearrangement, and genome dosage. By enabling large genomic regions to be deleted, transferred, reconstructed, or redesigned, chromosome engineering may provide a route to test these hypotheses experimentally rather than relying solely on comparative inference. Beyond its value for mechanistic studies, chromosome engineering may also contribute to clinical applications, including the development of therapeutic strategies for chromosomal aneuploidies such as Down syndrome (Hashizume et al. 2025). As genome assemblies and manipulation technologies become available across a broader range of species, experimental

chromosome-scale genomics may connect biodiversity-driven genome discovery with functional validation at the level of entire loci and chromosomes (Figure 4).

4.5 | Conclusion

Chromosome-level genome assemblies are transforming biological research not simply because they provide more complete reference sequences, but because they reveal the organization of genomes at the scale at which structural variation, gene regulation, chromosome evolution, and genome dosage can be meaningfully interpreted. The case studies reviewed here demonstrate how chromosome-resolved genomes reveal lineage-specific architecture, connect local molecular changes with phenotypic innovation and plasticity, reconstruct deep evolutionary history, and resolve genetically complex and polyploid systems.

Future progress will depend on moving beyond isolated reference genomes toward comparative, pangenomic, population-scale, and haplotype-resolved resources. These resources must in turn be integrated with multi-omics data and functional experiments (Figure 4). Within this framework, global biology provides an approach for connecting genes, regulatory elements, sequence variants, and cellular states with chromosome architecture, population diversity, organismal phenotypes, and evolutionary history. Chromosome-level assemblies can thereby move from descriptive genomic resources to platforms for generating and experimentally testing mechanistic hypotheses about genome function and evolution.

Author Contributions

Tetsuo Kon: conceptualization, writing – original draft, writing – review and editing, funding acquisition. **Mayuko Hamada:** conceptualization, writing – review and editing, writing – original draft, funding acquisition. **Kosuke Kataoka:** conceptualization, writing – review and editing, writing – original draft, funding acquisition. **Yuki Monden:** writing – original draft, writing – review and editing. **Johannes Nicolaus Wibisana:** writing – original draft, writing – review and editing. **Narumi Uno:** writing – review and editing, writing – original draft. **Kouhei Toga:** writing – original draft, writing – review and editing. **Yi-Jyun Luo:** writing – review and editing, writing – original draft.

Acknowledgments

This review was inspired by a symposium (3PS-12) organized at the 48th Annual Meeting of the Molecular Biology Society of Japan, held in December 2025. We thank the organizing committee of the society for providing the opportunity to hold the symposium. We thank Koto Kon-Nanjo (University of Vienna) and Nicholas Luscombe, Charles Plessy, and Gaspar Sánchez-Serna (Okinawa Institute of Science and Technology, OIST) for critically reading the manuscript and providing helpful comments and edits.

Funding

This work was supported by the JSPS KAKENHI (Grant Numbers JP24K09559 to M.H., JP26K02062 to K.K. and JP23K05867 to N.U.), the JSPS Overseas Research Fellowships to T.K., and the Bio-oriented Technology Research Advancement Institution (JPJ009237) to K.K.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

No new experimental data were generated in this study. The data underlying Figure 1 and Table 1 were retrieved from the NCBI Genome database.

References

- Adkins, P., R. Mrowicki, and Marine Biological Association Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. 2023. "The Genome Sequence of the Gelatinous Scale Worm, *Alentia gelatinosa* (Sars, 1835)." *Wellcome Open Research* 8: 542.
- Ahuja, N., D. T. Schultz, D. Destanović, et al. 2026. "Siphonophore Genome Structure and the Evolution of Functional Specialization." *PLoS One* 21: e0351247. <https://doi.org/10.1371/journal.pone.0351247>.
- Albalat, R., and C. Cañestro. 2016. "Evolution by Gene Loss." *Nature Reviews. Genetics* 17: 379–391.
- Alonge, M., L. Lebeigle, M. Kirsche, et al. 2022. "Automated Assembly Scaffolding Using RagTag Elevates a New Tomato System for High-Throughput Genome Editing." *Genome Biology* 23: 258.
- Altomose, N., G. A. Logsdon, A. V. Bzikadze, et al. 2022. "Complete Genomic and Epigenetic Maps of Human Centromeres." *Science* 376: eabl4178.
- Astashyn, A., E. S. Tvedte, D. Sweeney, et al. 2024. "Rapid and Sensitive Detection of Genome Contamination at Scale With FCS-GX." *Genome Biology* 25: 60.
- Bliznina, A., A. Masunaga, M. J. Mansfield, et al. 2021. "Telomere-To-Telomere Assembly of the Genome of an Individual *Oikopleura dioica* From Okinawa Using Nanopore-Based Sequencing." *BMC Genomics* 22: 222.
- Bonfoey, A. M., J. Chen, and Z. R. Stahlschmidt. 2023. "Stress Tolerance Is Influenced by Artificial Light at Night During Development and Life-History Strategy." *Journal of Experimental Biology* 226: jeb245195.
- Brozovic, M., C. Dantec, J. Dardaillon, et al. 2018. "ANISEED 2017: Extending the Integrated Ascidian Database to the Exploration and Evolutionary Comparison of Genome-Scale Datasets." *Nucleic Acids Research* 46: D718–D725.
- Buffenstein, R. 2008. "Negligible Senescence in the Longest Living Rodent, the Naked Mole-Rat: Insights From a Successfully Aging Species." *Journal of Comparative Physiology. B* 178: 439–445.
- Buffenstein, R., and M. Pinto. 2009. "Endocrine Function in Naturally Long-Living Small Mammals." *Molecular and Cellular Endocrinology* 299: 101–111.
- Burton, J. N., A. Adey, R. P. Patwardhan, R. Qiu, J. O. Kitzman, and J. Shendure. 2013. "Chromosome-Scale Scaffolding of de Novo Genome Assemblies Based on Chromatin Interactions." *Nature Biotechnology* 31: 1119–1125.
- Cazet, J. F., S. Siebert, H. Morris Little, et al. 2023. "A Chromosome-Scale Epigenetic Map of the Hydra Genome Reveals Conserved Regulators of Cell State." *Genome Research* 33: 283–298.
- Challis, R., E. Richards, J. Rajan, G. Cochrane, and M. Blaxter. 2020. "BlobToolKit—Interactive Quality Assessment of Genome Assemblies." *G3: Genes, Genomes, Genetics* 10: 1361–1374.
- Chao, K.-H., J. M. Heinz, C. Hoh, et al. 2025. "Combining DNA and Protein Alignments to Improve Genome Annotation With LiftOn." *Genome Research* 35: 311–325.
- Chapman, J. A., E. F. Kirkness, O. Simakov, et al. 2010. "The Dynamic Genome of Hydra." *Nature* 464: 592–596.
- Chen, Y., Z. Chen, H. Wang, et al. 2025. "A cGAS-Mediated Mechanism in Naked Mole-Rats Potentiates DNA Repair and Delays Aging." *Science* 390: eadp5056.
- Cheng, H., G. T. Concepcion, X. Feng, H. Zhang, and H. Li. 2021. "Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs With Hifiasm." *Nature Methods* 18: 170–175.
- Collins, F. S., M. Morgan, and A. Patrinos. 2003. "The Human Genome Project: Lessons From Large-Scale Biology." *Science* 300: 286–290.
- Consens, M. E., C. Dufault, M. Wainberg, et al. 2025. "Transformers and genome language models." *Nature Machine Intelligence* 7: 346–362.
- Davy, S. K., D. Allemand, and V. M. Weis. 2012. "Cell Biology of Cnidarian-Dinoflagellate Symbiosis." *Microbiology and Molecular Biology Reviews* 76: 229–261.
- Dehal, P., and J. L. Boore. 2005. "Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate." *PLoS Biology* 3: e314.
- Denoeud, F., S. Henriet, S. Mungpakdee, et al. 2010. "Plasticity of Animal Genome Architecture Unmasked by Rapid Evolution of a Pelagic Tunicate." *Science* 330: 1381–1385.
- Destanović, D., D. T. Schultz, R. Styfals, et al. 2023. "A Chromosome-Level Reference Genome for the Common Octopus, *Octopus vulgaris* (Cuvier, 1797)." *G3: Genes, Genomes, Genetics* 13: jkad220.
- Di Genova, A., E. Buena-Atienza, S. Ossowski, and M.-F. Sagot. 2021. "Efficient Hybrid de Novo Assembly of Human Genomes With WENGAN." *Nature Biotechnology* 39: 422–430.
- Dierckxsens, N., M. J. Mansfield, C. Plessy, N. M. Luscombe, and J. R. Vermeesch. 2025. "Unlocking the Full Potential of Oxford Nanopore Reads With NOVOloci." *bioRxiv*. <https://doi.org/10.1101/2025.08.08.669243>.
- Dierckxsens, N., K. Watanabe, Y. Tan, et al. 2024. "Tracing Homopolymers in *Oikopleura dioica*'s Mitogenome." *Genome Biology and Evolution* 16: evae182.
- Donoghue, P. C. J., and M. A. Purnell. 2005. "Genome Duplication, Extinction and Vertebrate Evolution." *Trends in Ecology & Evolution* 20: 312–319.
- Dudchenko, O., S. S. Batra, A. D. Omer, et al. 2017. "De Novo Assembly of the *Aedes aegypti* Genome Using Hi-C Yields Chromosome-Length Scaffolds." *Science* 356: 92–95.
- Durand, N. C., J. T. Robinson, M. S. Shamim, et al. 2016. "Juicebox Provides a Visualization System for Hi-C Contact Maps With Unlimited Zoom." *Cell Systems* 3: 99–101.
- Durand, N. C., M. S. Shamim, I. Machol, et al. 2016. "Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments." *Cell Systems* 3: 95–98.
- Egawa, M., N. Uno, R. Komazaki, et al. 2025. "Generation of Monosomy 21q Human iPS Cells by CRISPR/Cas9-Mediated Interstitial Megabase Deletion." *Genes to Cells* 30: e13184.
- Faulkes, C. G., D. H. Abbott, H. P. O'Brien, et al. 1997. "Micro- and Macrogeographical Genetic Structure of Colonies of Naked Mole-Rats *Heterocephalus glaber*." *Molecular Ecology* 6: 615–628.
- Ferrández-Roldán, A., J. Martí-Solans, C. Cañestro, and R. Albalat. 2019. "*Oikopleura dioica*: An Emergent Chordate Model to Study the Impact of Gene Loss on the Evolution of the Mechanisms of Development." *Results and Problems in Cell Differentiation* 68: 63–105.
- Gabriel, L., T. Brûna, K. J. Hoff, et al. 2024. "BRAKER3: Fully Automated Genome Annotation Using RNA-Seq and Protein Evidence With GeneMark-ETP, AUGUSTUS, and TSEBRA." *Genome Research* 34: 769–777.

- Gambogi, C. W., G. J. Birchak, E. Mer, et al. 2024. "Efficient Formation of Single-Copy Human Artificial Chromosomes." *Science* 383: 1344–1349.
- Ganot, P., T. Kallesøe, R. Reinhardt, D. Chourrout, and E. M. Thompson. 2004. "Spliced-Leader RNA Trans Splicing in a Chordate, *Oikopleura dioica*, With a Compact Genome." *Molecular and Cellular Biology* 24: 7795–7805.
- Goldfarb, T., V. K. Kodali, S. Pujar, et al. 2025. "NCBI RefSeq: Reference Sequence Standards Through 25 Years of Curation and Annotation." *Nucleic Acids Research* 53: D243–D257.
- Goodwin, S., J. D. McPherson, and W. R. McCombie. 2016. "Coming of Age: Ten Years of Next-Generation Sequencing Technologies." *Nature Reviews. Genetics* 17: 333–351.
- Hamada, M., N. Satoh, and K. Khalturin. 2020. "A Reference Genome From the Symbiotic Hydrozoan, *Hydra viridissima*." *G3: Genes, Genomes, Genetics* 10: 3883–3895.
- Hamada, M., K. Schröder, J. Bathia, et al. 2018. "Metabolic Co-Dependence Drives the Evolutionarily Ancient Hydra-Chlorella Symbiosis." *eLife* 7: e35122.
- Harrison, R. G. 1979. "Flight Polymorphism in the Field Cricket *Gryllus pennsylvanicus*." *Oecologia* 40: 125–132.
- Hashizume, R., S. Wakita, H. Sawada, et al. 2025. "Trisomic Rescue via Allele-Specific Multiple Chromosome Cleavage Using CRISPR-Cas9 in Trisomy 21 Cells." *PNAS Nexus* 4: pgaf022.
- Hirakawa, H., Y. Okada, H. Tabuchi, et al. 2015. "Survey of Genome Sequences in a Wild Sweet Potato, *Ipomoea trifida* (H. B. K.) G. Don." *DNA Research* 22: 171–179.
- Hu, J., Z. Wang, Z. Sun, et al. 2024. "NextDenovo: An Efficient Error Correction and Accurate Assembly Tool for Noisy Long Reads." *Genome Biology* 25: 107.
- Hu, Y., J. Chen, L. Fang, et al. 2019. "Gossypium Barbadosense and *Gossypium hirsutum* Genomes Provide Insights Into the Origin and Evolution of Allotetraploid Cotton." *Nature Genetics* 51: 739–748.
- Huang, N., and H. Li. 2023. "Compleasm: A Faster and More Accurate Reimplementation of BUSCO." *Bioinformatics* 39: btad595.
- Ingram, C. M., N. J. Troendle, C. A. Gill, S. Braude, and R. L. Honeycutt. 2015. "Challenging the Inbreeding Hypothesis in a Eusocial Mammal: Population Genetics of the Naked Mole-Rat, *Heterocephalus glaber*." *Molecular Ecology* 24: 4848–4865.
- International Human Genome Sequencing Consortium. 2004. "Finishing the Euchromatic Sequence of the Human Genome." *Nature* 431: 931–945.
- Ito, Y., R. Sanno, S. Ashikari, et al. 2025. "Chromosome-Scale Whole Genome Assembly and Annotation of the Jamaican Field Cricket *Gryllus assimilis*." *Scientific Data* 12: 826.
- Jin, F., Z. Zhou, Q. Guo, et al. 2020. "High-Quality Genome Assembly of *Metaphire vulgaris*." *PeerJ* 8: e10313.
- Kaplan, N., and J. Dekker. 2013. "High-Throughput Genome Scaffolding From In Vivo DNA Interaction Frequency." *Nature Biotechnology* 31: 1143–1147.
- Kasumovic, M. M., Z. Chen, and M. R. Wilkins. 2016. "Australian Black Field Crickets Show Changes in Neural Gene Expression Associated With Socially-Induced Morphological, Life-History, and Behavioral Plasticity." *BMC Genomics* 17: 827.
- Kataoka, K., R. Sanno, T. Gaczorek, et al. 2026. "Chromosome-Scale Genome Assembly and Annotation of the Two-Spotted Cricket *Gryllus bimaculatus* (Orthoptera: Gryllidae)." *G3: Genes, Genomes, Genetics* 16, no. 4: jkag036. <https://doi.org/10.1101/2025.10.31.685973>.
- Kataoka, K., Y. Shimizu, R. Sanno, et al. 2026. "Temporal Orchestration of Transcriptional and Epigenomic Programming Underlying Maternal Embryonic Diapause in a Cricket Model." *Communications Biology* 9: 825.
- Kazuki, Y., F. J. Gao, M. Yamakawa, et al. 2022. "A Transchromosomal Rat Model With Human Chromosome 21 Shows Robust Down Syndrome Features." *American Journal of Human Genetics* 109: 328–344.
- Kenny, N. J., S. A. McCarthy, O. Dudchenko, et al. 2020. "The Gene-Rich Genome of the Scallop *Pecten maximus*." *GigaScience* 9: g1aa037.
- Kim, I. V., C. Navarrete, X. Grau-Bové, et al. 2025. "Chromatin Loops Are an Ancestral Hallmark of the Animal Regulatory Genome." *Nature* 642: 1097–1105.
- Kirilenko, B. M., C. Munegowda, E. Osipova, et al. 2023. "Integrating Gene Annotation With Orthology Inference at Scale." *Science* 380: eabn3107.
- Kolmogorov, M., J. Yuan, Y. Lin, and P. A. Pevzner. 2019. "Assembly of Long, Error-Prone Reads Using Repeat Graphs." *Nature Biotechnology* 37: 540–546.
- Kon, T., K. Kon-Nanjo, K. S. Handayani, et al. 2025. "Chromosome-Level Genome Assembly of the Doctor Fish (*Garra rufa*)." *Scientific Data* 12: 765.
- Kon-Nanjo, K., T. Kon, H. R. Horkan, et al. 2023. "Chromosome-Level Genome Assembly of *Hydractinia symbiolongicarpus*." *G3: Genes, Genomes, Genetics* 13: jkad107.
- Kon-Nanjo, K., T. Kon, T. C.-T. K. Yu, et al. 2025. "The Dynamic Genomes of Hydra and the Anciently Active Repeat Complement of Animal Chromosomes." *Genome Biology* 26: 186.
- Koren, S., B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, and A. M. Phillippy. 2017. "Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation." *Genome Research* 27: 722–736.
- Kouprina, N., W. C. Earnshaw, H. Masumoto, and V. Larionov. 2013. "A New Generation of Human Artificial Chromosomes for Functional Genomics and Gene Therapy." *Cellular and Molecular Life Sciences* 70: 1135–1148.
- Kouprina, N., and V. Larionov. 2016. "Transformation-Associated Recombination (TAR) Cloning for Genomics Studies and Synthetic Biology." *Chromosoma* 125: 621–632.
- Kumar, S., G. Stecher, M. Suleski, M. Sanderford, S. Sharma, and K. Tamura. 2024. "MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing." *Molecular Biology and Evolution* 41: msae263.
- Kwiatkowski, D., M. Blaxter, and Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. 2021. "The Genome Sequence of the Bootlace Worm, *Lineus longissimus* (Gunnerus, 1770)." *Wellcome Open Research* 6: 272.
- Lander, E. S., L. M. Linton, B. Birren, et al. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409: 860–921.
- Lee, J.-H., B.-J. Kim, G. Han, O. Frunze, G. Nah, and H. W. Kwon. 2025. "Chromosome Level de Novo Hybrid Assembly of Asian Honeybee, *Apis cerana* Koreana." *Scientific Reports* 15: 26912.
- Lewin, T. D., I. J.-Y. Liao, M.-E. Chen, J. D. D. Bishop, P. W. H. Holland, and Y.-J. Luo. 2025. "Fusion, Fission, and Scrambling of the Bilaterian Genome in Bryozoa." *Genome Research* 35: 78–92.
- Lewin, T. D., I. J.-Y. Liao, and Y.-J. Luo. 2024. "Annelid Comparative Genomics and the Evolution of Massive Lineage-Specific Genome Rearrangement in Bilaterians." *Molecular Biology and Evolution* 41: msae172.
- Lewin, T. D., I. J.-Y. Liao, and Y.-J. Luo. 2025. "Conservation of Bilaterian Genome Structure Is the Exception, Not the Rule." *Genome Biology* 26: 247.

- Lewin, T. D., T. Sakagami, L.-J. Kao, et al. 2025. "Phoronid Genome Supports a Monophyletic Lophophorata." *Current Biology* 35: 5633–5645.e7.
- Leys, S. P., U. Hentschel, G. Oatley, et al. 2025. "The Chromosome-Level Genome Sequences of the Freshwater Sponge, *Spongilla lacustris* (Linnaeus, 1759) and the Chlorophyte Cobiont *Choricystis* sp., and the Associated Microbial Metagenome Sequences." *Wellcome Open Research* 10: 222.
- Li, M., S. Yang, W. Xu, et al. 2019. "The Wild Sweetpotato (*Ipomoea trifida*) Genome Provides Insights Into Storage Root Development." *BMC Plant Biology* 19: 119.
- Li, N., Q. He, J. Wang, et al. 2023. "Super-Pangenome Analyses Highlight Genomic Diversity and Structural Variation Across Wild and Cultivated Tomato Species." *Nature Genetics* 55: 852–860.
- Liao, I. J.-Y., T.-M. Lu, M.-E. Chen, and Y.-J. Luo. 2023. "Spiralian Genomics and the Evolution of Animal Genome Architecture." *Briefings in Functional Genomics* 22: 498–508.
- Locatelli, N. S., S. A. Kitchen, K. H. Stankiewicz, et al. 2024. "Chromosome-Level Genome Assemblies and Genetic Maps Reveal Heterochiasmy and Macrosynteny in Endangered Atlantic Acropora." *BMC Genomics* 25: 1119.
- Manchanda, N., J. L. Portwood, M. R. Woodhouse, et al. 2020. "GenomeQC: A Quality Assessment Tool for Genome Assemblies and Gene Structure Annotations." *BMC Genomics* 21: 193.
- Mardis, E. R. 2008. "Next-Generation DNA Sequencing Methods." *Annual Review of Genomics and Human Genetics* 9: 387–402.
- Martín-Zamora, F. M., B. E. Davies, R. D. Donnellan, K. Guynes, and J. M. Martín-Durán. 2023. "Functional Genomics in Spiralia." *Briefings in Functional Genomics* 22: 487–497.
- Martín-Zamora, F. M., Y. Liang, K. Guynes, et al. 2023. "Annelid Functional Genomics Reveal the Origins of Bilaterian Life Cycles." *Nature* 615: 105–110.
- Martí-Solans, J., H. Godoy-Marín, M. Diaz-Gracia, et al. 2021. "Massive Gene Loss and Function Shuffling in Appendicularians Stretch the Boundaries of Chordate Wnt Family Evolution." *Frontiers in Cell and Development Biology* 9: 700827.
- Masaki, S., and T. Sugahara. 1992. "Photoperiodic Control of Larval Development and Wing Form in *Modicogryllus* sp. (Orthoptera: Gryllidae)." *Ecological Research* 7: 25–30.
- Masunaga, A., M. J. Mansfield, Y. Tan, et al. 2022. "The Cosmopolitan Appendicularian *Oikopleura dioica* Reveals Hidden Genetic Diversity Around the Globe." *Marine Biology* 169: 157.
- Matsuoka, Y., T. Nakamura, T. Watanabe, et al. 2025. "Establishment of CRISPR/Cas9-Based Knock-In in a Hemimetabolous Insect: Targeted Gene Tagging in the Cricket *Gryllus bimaculatus*." *Development* 152: dev199746.
- Merchant, A. T., S. H. King, E. Nguyen, and B. L. Hie. 2025. "Semantic Design of Functional de Novo Genes From a Genomic Language Model." *Nature* 649: 749–758.
- Miyamoto, H., H. Kobayashi, N. Kishima, et al. 2024. "Rapid Human Genomic DNA Cloning Into Mouse Artificial Chromosome via Direct Chromosome Transfer From Human iPSC and CRISPR/Cas9-Mediated Translocation." *Nucleic Acids Research* 52: 1498–1511.
- Mollinari, M., B. A. Olukolu, G. d. S. Pereira, et al. 2020. "Unraveling the Hexaploid Sweetpotato Inheritance Using Ultra-Dense Multilocus Mapping." *G3: Genes, Genomes, Genetics* 10: 281–292.
- Nevers, Y., A. Warwick Vesztrocy, V. Rossier, et al. 2025. "Quality Assessment of Gene Repertoire Annotations With OMArk." *Nature Biotechnology* 43: 124–133.
- Nowoshilow, S., S. Schloissnig, J.-F. Fei, et al. 2018. "The Axolotl Genome and the Evolution of Key Tissue Formation Regulators." *Nature* 554: 50–55.
- Nurk, S., S. Koren, A. Rhie, et al. 2022. "The Complete Sequence of a Human Genome." *Science* 376: 44–53.
- Obata, N., H. Tabuchi, M. Kurihara, E. Yamamoto, K. Shirasawa, and Y. Monden. 2022. "Mapping of Nematode Resistance in Hexaploid Sweetpotato Using a Next-Generation Sequencing-Based Association Study." *Frontiers in Plant Science* 13: 858747.
- Oka, K., S. Fujioka, Y. Kawamura, et al. 2022. "Resistance to Chemical Carcinogenesis Induction via a Dampened Inflammatory Response in Naked Mole-Rats." *Communications Biology* 5: 287.
- Parey, E., A. Louis, J. Montfort, et al. 2023. "Genome Structures Resolve the Early Diversification of Teleost Fishes." *Science* 379: 572–575.
- Plessy, C., M. J. Mansfield, A. Bliznina, et al. 2024. "Extreme Genome Scrambling in Marine Planktonic *Oikopleura dioica* Cryptic Species." *Genome Research* 34: 426–440.
- Pollard, M. O., D. Gurdasani, A. J. Mentzer, T. Porter, and M. S. Sandhu. 2018. "Long Reads: Their Purpose and Place." *Human Molecular Genetics* 27: R234–R241.
- Prjibelski, A., D. Antipov, D. Meleshko, A. Lapidus, and A. Korobeynikov. 2020. "Using SPAdes De Novo Assembler." *Current Protocols in Bioinformatics* 70: e102.
- Rautiainen, M., S. Nurk, B. P. Walenz, et al. 2023. "Telomere-To-Telomere Assembly of Diploid Chromosomes With Verkko." *Nature Biotechnology* 41: 1474–1482.
- Rhie, A., B. P. Walenz, S. Koren, and A. M. Phillippy. 2020. "Mercury: Reference-Free Quality, Completeness, and Phasing Assessment for Genome Assemblies." *Genome Biology* 21: 245.
- Ruan, J., and H. Li. 2020. "Fast and Accurate Long-Read Assembly With wtdbg2." *Nature Methods* 17: 155–158.
- Salojärvi, J., A. Rambani, Z. Yu, et al. 2024. "The Genome and Population Genomics of Allopolyploid *Coffea arabica* Reveal the Diversification History of Modern Coffee Cultivars." *Nature Genetics* 56: 721–731.
- Schartl, M., J. M. Wolterring, I. Irisarri, et al. 2024. "The Genomes of All Lungfish Inform on Genome Expansion and Tetrapod Evolution." *Nature* 634: 96–103.
- Schultz, D. T., W. R. Francis, J. D. McBroome, L. M. Christianson, S. H. D. Haddock, and R. E. Green. 2021. "A Chromosome-Scale Genome Assembly and Karyotype of the Ctenophore *Hormiphora californensis*." *G3: Genes, Genomes, Genetics* 11: jkab302.
- Schultz, D. T., S. H. D. Haddock, J. V. Bredeson, R. E. Green, O. Simakov, and D. S. Rokhsar. 2023. "Ancient Gene Linkages Support Ctenophores as Sister to Other Animals." *Nature* 618: 110–117.
- Shiga, S., and H. Numata. 1997. "Seasonal Changes in the Incidence of Embryonic Diapause in the Band-Legged Ground Cricket, *Dianemobius nigrofasciatus*." *Zoological Science* 14: 1015–1018.
- Shimizu, Y., and K. Kataoka. 2026. "DIPA-CRISPR Mediated Knockout of Vermilion Generates a Visible Eye Color Marker for the Band-Legged Ground Cricket *Dianemobius nigrofasciatus*." *Archives of Insect Biochemistry and Physiology* 121: e70135.
- Shirai, Y., M.-D. Piulachs, X. Belles, and T. Daimon. 2022. "DIPA-CRISPR Is a Simple and Accessible Method for Insect Gene Editing." *Cell Reports Methods* 2: 100215.
- Shito, T. T., V. Jayakumar, K. Nishitsuji, et al. 2025. "Genomic Resources of *Ascidia aspersa* and Comparative Analysis Across Tunicates Reveal Conserved Class-Level Features and Evolutionary Diversification." *bioRxiv*. <https://doi.org/10.1101/2025.11.06.686920>.
- Shulaev, V., D. J. Sargent, R. N. Crowhurst, et al. 2011. "The Genome of Woodland Strawberry (*Fragaria vesca*)." *Nature Genetics* 43: 109–116.

- Shumate, A., and S. L. Salzberg. 2021. "Liftoff: Accurate Mapping of Gene Annotations." *Bioinformatics* 37: 1639–1643.
- Siebert, S., J. A. Farrell, J. F. Cazet, et al. 2019. "Stem Cell Differentiation Trajectories in Hydra Resolved at Single-Cell Resolution." *Science* 365: eaav9314.
- Simakov, O., J. Bredeson, K. Berkoff, et al. 2022. "Deeply Conserved Synteny and the Evolution of Metazoan Chromosomes." *Science Advances* 8: eabi5884.
- Simakov, O., F. Marletaz, S.-J. Cho, et al. 2013. "Insights Into Bilateral Evolution From Three Spiralian Genomes." *Nature* 493: 526–531.
- Simakov, O., F. Marlétaz, J.-X. Yue, et al. 2020. "Deeply Conserved Synteny Resolves Early Events in Vertebrate Evolution." *Nature Ecology & Evolution* 4: 820–830.
- Simakov, O., and G. P. Wagner. 2025. "The Application of Irreversible Genomic States to Define and Trace Ancient Cell Type Homologies." *EvoDevo* 16: 5.
- Sokolowski, D. J., M. Miclăuș, A. Nater, et al. 2024. "An Updated Reference Genome Sequence and Annotation Reveals Gene Losses and Gains Underlying Naked Mole-Rat Biology." *bioRxiv*. <https://doi.org/10.1101/2024.11.26.625329>.
- Steindler, L., M. A. Durán Canché, M. Ilan, et al. 2025. "The Chromosomal Genome Sequence of the Marine Sponge *Diacarnus Erythraeanus* Kelly-Borges & Vacelet, 1995, and Its Associated Microbial Metagenome Sequences." *Wellcome Open Research* 10: 466.
- Tanigawa, N., K. Matsumoto, K. Yasuyama, H. Numata, and S. Shiga. 2009. "Early Embryonic Development and Diapause Stage in the Band-Legged Ground Cricket *Dianemobius nigrofasciatus*." *Development Genes and Evolution* 219: 589–596.
- Tegenfeldt, F., D. Kuznetsov, M. Manni, M. Berkeley, E. M. Zdobnov, and E. V. Kriventseva. 2025. "OrthoDB and BUSCO Update: Annotation of Orthologs With Wider Sampling of Genomes." *Nucleic Acids Research* 53: D516–D522.
- Toga, K., K. Oka, H. Tanaka, et al. 2026. "Genome Assembly and Annotation of the Naked Mole Rat *Heterocephalus glaber* Reared in Japan." *Scientific Data* 13: 705.
- Uno, N., K. Hiramatsu, K. Uno, S. Komoto, Y. Kazuki, and M. Oshimura. 2017. "CRISPR/Cas9-Induced Transgene Insertion and Telomere-Associated Truncation of a Single Human Chromosome for Chromosome Engineering in CHO and A9 Cells." *Scientific Reports* 7: 12739.
- Uno, N., H. Miyamoto, K. Yamazaki, et al. 2024. "Microcell-Mediated Chromosome Transfer Between Non-Identical Human iPSCs." *Molecular Therapy - Nucleic Acids* 35: 102382.
- van Dijk, E. L., Y. Jaszczyszyn, D. Naquin, and C. Thermes. 2018. "The Third Revolution in Sequencing Technology." *Trends in Genetics* 34: 666–681.
- Vargas-Chávez, C., L. Benítez-Álvarez, G. I. Martínez-Redondo, et al. 2025. "An Episodic Burst of Massive Genomic Rearrangements and the Origin of Non-Marine Annelids." *Nature Ecology & Evolution* 9: 1263–1279.
- Vargas-Chávez, C., A. McLysaght, and R. Fernández. 2025. "Breaking Bad: When Clitellate Genomes Go Rogue." *Trends in Genetics* 41: 995–1004.
- Walkowiak, S., L. Gao, C. Monat, et al. 2020. "Multiple Wheat Genomes Reveal Global Variation in Modern Breeding." *Nature* 588: 277–283.
- Wang, K., R. Tomura, W. Chen, et al. 2020. "A Genome Database for a Japanese Population of the Larvacean *Oikopleura dioica*." *Development, Growth & Differentiation* 62: 450–461.
- Wong, W. Y., O. Simakov, D. M. Bridge, et al. 2019. "Expansion of a Single Transposable Element Family Is Associated With Genome-Size Increase and Radiation in the Genus *Hydra*." *Proceedings of the National Academy of Sciences of the United States of America* 116: 22915–22917.
- Wu, S., K. H. Lau, Q. Cao, et al. 2018. "Genome Sequences of Two Diploid Wild Relatives of Cultivated Sweetpotato Reveal Targets for Genetic Improvement." *Nature Communications* 9: 4580.
- Wu, S., H. Sun, X. Zhao, et al. 2025. "Phased Chromosome-Level Assembly Provides Insight Into the Genome Architecture of Hexaploid Sweetpotato." *Nature Plants* 11: 1951–1959.
- Xie, X., W. Teng, Z. Yu, et al. 2023. "Chromosome-Level Genome Assembly of Sea Scallop *Placopecten magellanicus* Provides Insights Into the Genetic Characteristics and Adaptive Evolution of Large Scallops." *Genomics* 115: 110747.
- Yamamoto, E., K. Shirasawa, T. Kimura, Y. Monden, M. Tanaka, and S. Isobe. 2020. "Genetic Mapping in Autohexaploid Sweet Potato With Low-Coverage NGS-Based Genotyping Data." *G3: Genes, Genomes, Genetics* 10: 2661–2670.
- Zhang, C., Z. Shao, Y. Kong, et al. 2024. "High-Quality Genome of a Modern Soybean Cultivar and Resequencing of 547 Accessions Provide Insights Into the Role of Structural Variation." *Nature Genetics* 56: 2247–2258.
- Zhang, Y., M. Zhao, J. Tan, et al. 2024. "Telomere-To-Telomere Citrullus Super-Pangenome Provides Direction for Watermelon Breeding." *Nature Genetics* 56: 1750–1761.
- Zhou, C., S. A. McCarthy, and R. Durbin. 2023. "YaHS: Yet Another Hi-C Scaffolding Tool." *Bioinformatics* 39: btac808.
- Zimin, A. V., G. Marçais, D. Puiu, M. Roberts, S. L. Salzberg, and J. A. Yorke. 2013. "The MaSuRCA Genome Assembler." *Bioinformatics* 29: 2669–2677.
- Zimmermann, B., J. D. Montenegro, S. M. C. Robb, et al. 2023. "Topological Structures and Syntenic Conservation in Sea Anemone Genomes." *Nature Communications* 14: 8270.