

Commentary

# Classification Problems of Repetitive DNA Sequences

Eva Šatović-Vukšić \*  and Miroslav Plohl

Division of Molecular Biology, Ruđer Bošković Institute, 10000 Zagreb, Croatia; plohl@irb.hr

\* Correspondence: esatovic@irb.hr

**Abstract:** Repetitive DNA sequences, satellite DNAs (satDNAs) and transposable elements (TEs) are essential components of the genome landscape, with many different roles in genome function and evolution. Despite significant advances in sequencing technologies and bioinformatics tools, detection and classification of repetitive sequences can still be an obstacle to the analysis of genomic repeats. Here, we summarize how specificities in repetitive DNA organizational patterns can lead to an inability to classify (and study) a significant fraction of bivalve mollusk repetitive sequences. We suggest that the main reasons for this inability are: the predominant association of satDNA arrays with Helitron/Helentron TEs; the existence of many complex loci; and the unusual, highly scattered organization of short satDNA arrays or single monomers across the whole genome. The specificities of bivalve genomes confirm the need for introducing diverse organisms as models in order to understand all aspects of repetitive DNA biology. It is expected that further development of sequencing techniques and synergy among different bioinformatics tools and databases will enable quick and unambiguous characterization and classification of repetitive DNA sequences in assembled genomes.

**Keywords:** repetitive DNA classification; satellite DNA; transposable element; Helitron/Helentron; bivalves; genome assemblies



**Citation:** Šatović-Vukšić, E.; Plohl, M. Classification Problems of Repetitive DNA Sequences. *DNA* **2021**, *1*, 84–90. <https://doi.org/10.3390/dna1020009>

Academic Editor: Darren Griffin

Received: 10 August 2021

Accepted: 11 October 2021

Published: 2 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Despite the exponential number of genome sequencing projects arising and spanning all taxa, genomic regions largely composed of repetitive DNA sequences still present substantial technical issues in the assembly of genomes [1]. Repetitive DNAs are mainly constituted of satellite DNAs (satDNAs), formed by sequences repeated in tandem, and of mobile elements, interspersed throughout the genome [2]. According to the established classical view, satDNAs are associated with constitutive heterochromatin which is commonly located at pericentromeric and subtelomeric chromosomal domains and at interstitial loci of the chromosomal arms. They build long arrays of monomers repeated in tandem, comprised of hundreds to thousands of highly similar repeat units [3]. However, more recent work has introduced new data and showed that satDNA sequences can also be located outside of heterochromatin, where they can be found in different organizational forms: as monomers or monomer fragments, in arrays of diverse length or incorporated into mobile elements, for example [4–10]. In addition, many links show that satDNAs and mobile elements are often tightly interconnected. For example, tandem repeats can be created from mobile elements or their segments, or satDNAs can expand from short internal arrays carried by mobile elements (reviewed in [11]).

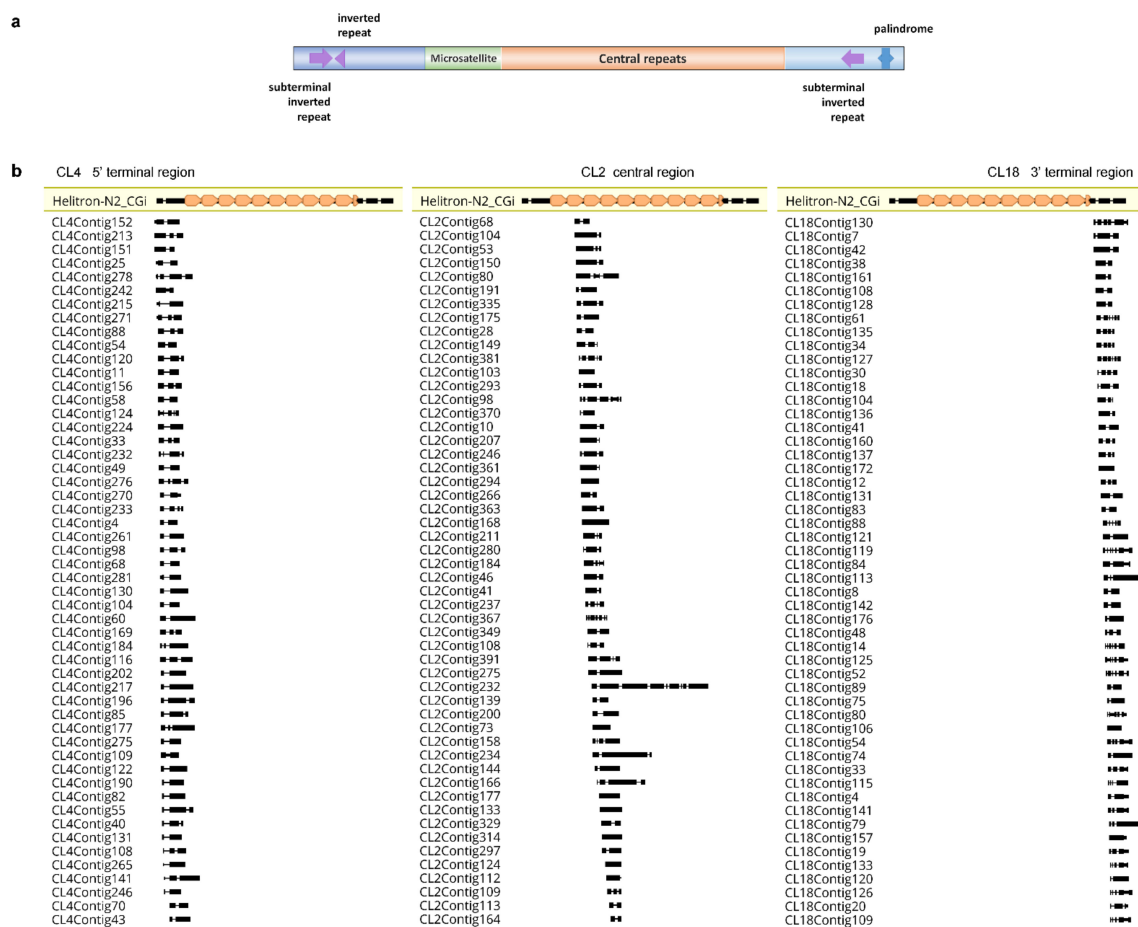
Sequencing problems arise in attempts to reconstruct repetitive genomic segments, and, subsequently, these regions are still regularly omitted or are misassembled in the available genomic data [12]. Ongoing improvements in sequencing technologies (e.g., long-read PacBio and Nanopore sequencing) are opening the possibility to obtain insights into these missing fractions of assembled genomes [13]. At the same time, a number of programs and software aimed to forward repeat detection and characterization are being generated and/or upgraded (reviewed in [14]), substantially changing our knowledge on

the inventory of repetitive sequences in genomes, satellitomes and repeatomes [15–20]. However, clear and unambiguous classifications of repetitive fractions of genomes still present a challenge, as described here for bivalve mollusks.

## 2. Repetitive Sequences in Bivalve Genomes

Bivalve genomes show many peculiarities in the content and composition of repetitive DNA sequences and heterochromatin and can be valuable sources of information as model organisms in studying repetitive DNA biology [21]. The overall contribution of repetitive DNAs to bivalve genomes is relatively high, in some species reaching 50% [22–25] or even 60% [26] of the genome. In contrast, the estimated content of sequences classified as satDNAs is very low according to genome sequencing projects, usually <2% [24,26–29]. This differs significantly from the organisms in which up to 40% of the genome is composed of satDNAs [30–32].

Interestingly, a large number of repetitive sequences, >70%, remain unassigned in many bivalve species ([28,29,33,34], etc.), differing again from the organisms where the success in repeat classification is significantly higher, reaching 98% in some plant species [35]. At the same time, it was noticed that Helitron/Helitron mobile elements constitute a substantial part of bivalve genomes, frequently being by far the most abundant type of DNA transposons in a certain genome [26,28,36]. Mobile elements of the Helitron/Helitron superfamily consist of conserved sequence segments at element ends that hold subterminal inverted repeats and a short palindromic sequence at the 3' end of the right flanking sequence, while in the central part, they frequently contain arrays of tandem repeats, preceded by a microsatellite sequence ([37], Figure 1a).



**Figure 1.** The structural characteristics of Helitron mobile elements and the result of RepeatExplorer clustering. (a) Helitron mobile elements contain conserved sequence segments at element ends that hold subterminal inverted repeats, an inverted

repeat in the 5' conserved segment and a short palindromic sequence at the 3' conserved segment. The central part of the element frequently contains arrays of tandem repeats preceded by a microsatellite sequence. (b) In RepeatExplorer analysis, different parts of Helitron mobile elements are placed in separate clusters. In this example, central repeats were assigned to cluster CL2 and classified as low-confidence satellite DNAs (satDNA), while sequence segments flanking central repeats were assigned to clusters CL4 and CL18, both remaining unclassified. The position of contigs, obtained after the clustering, is shown in respect to the Helitron-N2\_Cgi consensus sequence.

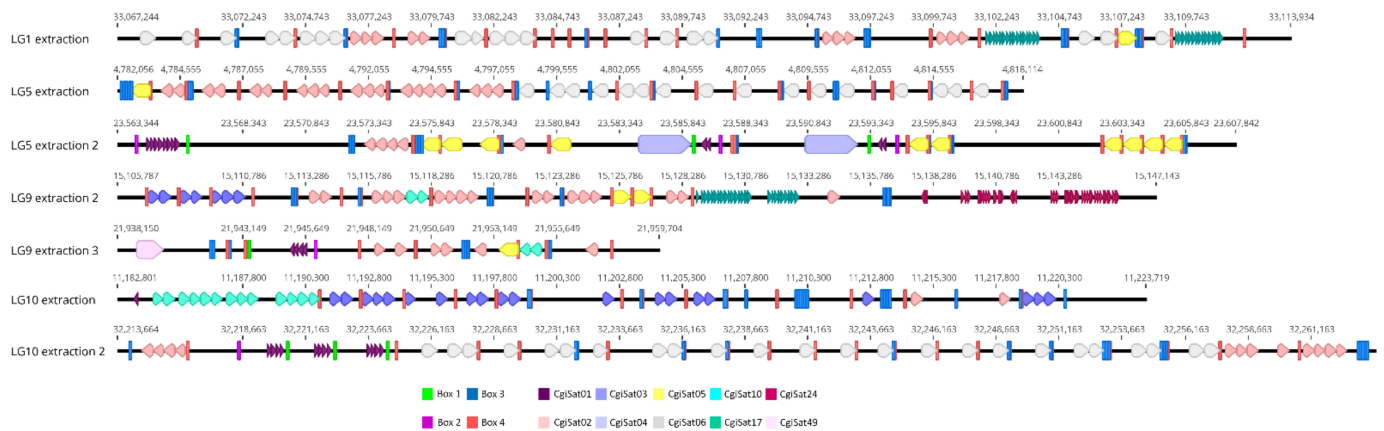
Whole elements can also be repeated in tandem, with frequent truncation at the 3' end of the element [38]. The number of central repeats found within Helitron/Helitron elements in bivalves can vary significantly from 1 to ~90 [39], with arrays holding up to 6 monomers being the most frequent [4,9,40–42].

### 3. Problems in Classification of Repetitive DNA Sequences: The Case of the Pacific Oyster *Crassostrea gigas*

The “hybrid” structure of mobile elements incorporating tandem repeats could, at least partially, explain the difficulties in the precise classification of repetitive sequences, and in determining the exact contribution of each type of repeat to the genome. For example, in a RepeatExplorer [43] analysis of the Pacific oyster *Crassostrea gigas*, contigs based on short-read sequences (Supplementary Data S1–S3) built separate clusters, which corresponded to different parts of Helitron mobile elements. Tandem repeats derived from the central parts of Helitrons were placed into one cluster and classified as satDNA, while sequences corresponding to flanking sequences were allocated into two separate clusters (Figure 1b) that remained unclassified [44]. Similar assignment problems were also observed in *Drosophila virilis* and *Drosophila americana*, species containing structurally similar elements of the same superfamily [45]. The inability to detect three formerly known satDNAs of *Rhodnius prolixus* with RepeatExplorer analysis [46] could also be a consequence of their association with certain sequence segments (potentially of transposable element origin), resulting in their placement within clusters that remained unclassified. While such problems are noticeable during analyses based on short-read data, repeat classification on assembled genomes encounters different obstacles. Programs based solely on structural characteristics found at the element ends (subterminal inverted repeats, short palindromic sequence) would fail to characterize large numbers of truncated variants, while homology-based classifications could be biased by several other factors. Among these factors, we include: the extreme variation in the number of central repeats (observed in [39]), sequence variations, the existence of two different types of repeats within an array (reported in [42]) and the existence of many complex loci. The Pacific oyster *C. gigas* genome is replete with the last factor (Figure 2). Here, tandemization of repeats and of conserved boxes that are usually found at the element ends, together with insertion, deletion and recombination events, generates a very complex network of tandem repeats and Helitron components, presenting a significant classification challenge.

In *C. gigas*, satellitome analysis revealed an unusual, highly scattered organization of relatively short satDNA arrays across the whole genome, a pattern followed by all 52 satDNAs detected in this species [39]. Interestingly, the same species contains an extremely low amount of heterochromatin, limited to the centromeric region of one chromosome pair and the telomeric region of another, and predominantly composed of DNA transposons [44,47]. For this species, in strong contrast with the established concept of satDNA genomic organization, no significant accumulation of satDNAs has been observed in any chromosomal loci. Additionally, a substantial number of satDNA sequences has been found in the form of single, interspersed monomers. They were, in the same manner as arrays of monomers, found either in standalone form, or associated with conserved boxes of Helitron mobile elements, from one or both sides [39]. We believe that this pattern, differing significantly from the classical organizational concept of satDNA, is a significant contributor to the inability to classify a large number of repetitive DNA sequences, especially standalone interspersed monomers. If examined from the aspect of interspersed sequences, their classification would be omitted due to a lack of some recognizable structural characteristics, while on the

other hand, they would most probably remain unrecognized by satDNA-oriented tandem repeat finders.



**Figure 2.** An example of complex loci found within the *Crassostrea gigas* genome. Vertical bars denote conserved boxes of Helitron mobile elements. Boxes 1 and 2 represent conserved segments found at the 5' and 3' ends of Helitron-N2\_Cgi, while boxes 3 and 4 represent 5' and 3' conserved segments, respectively, shared among several Helitron mobile elements of *C. gigas*. CgiSat01–06 and CgiSat10, 17, 24, 49 denote monomers belonging to some of the satDNAs described in [46]. Each arrowhead represents one monomer, with arrowhead sizes varying in respect to the monomer length of a particular satDNA. In silico localization was performed on linkage groups/chromosomes of the currently representative genome assembly GCA\_902806645.1 [36].

#### 4. Conclusions

While there are other circumstances that can contribute to the difficulties in the classification of repetitive elements in bivalves, e.g., the existence of species-specific variants of known elements [33], we believe that a significant role is played by: (i) the absence of classical satellite DNAs, (ii) the predominant existence of satDNA arrays of different sizes within the complete and truncated Helitron/Helitron elements, (iii) the existence of combined types of arrays within such elements, (iv) the presence of a large number of single, interspersed monomers throughout the genome and (v) the existence of complex loci generated by insertion, deletion and recombination events in combination with tandemization of repeats and mobile element conserved boxes.

Comprehensive and accurate annotation and characterization of repetitive sequences are necessary, as the contribution of this part of the genome is important for the understanding of the genomic/chromosomal architecture and function as a whole. We believe that the combination of quickly evolving sequencing technologies, followed by the constant development and improvement of bioinformatics tools and databases, supplemented with manual curation and adaptation according to the specificities of each model organism of interest, will significantly forward our abilities to correctly characterize and classify repetitive DNA sequences and obtain novel insights into these important genomic components.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/dna1020009/s1>, Data S1: Contigs belonging to cluster CL2, obtained after RepeatExplorer clustering, Data S2: Contigs belonging to cluster CL4, obtained after RepeatExplorer clustering, Data S3: Contigs belonging to cluster CL18, obtained after RepeatExplorer clustering.

**Author Contributions:** Conceptualization, M.P. and E.Š.-V.; formal analysis, E.Š.-V.; writing—original draft preparation, E.Š.-V.; writing—review and editing, M.P., E.Š.-V.; visualization, E.Š.-V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Croatian Science Foundation, grant number IP-2019-045522.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Short reads obtained by Illumina sequencing of *Crassostrea gigas* genomic DNA, used for RepeatExplorer clustering, can be found in NCBI under the BioSample accession number: SAMN15184427, BioProject: PRJNA638244. *C. gigas* chromosome assembly used for in silico localization can be found in NCBI, assembly accession: GCA\_902806645.1. Nucleotide sequence of Helitron-N2\_Cgi is available at: Repbase database, Repbase Reports 13(1), 530–530 (2013). Contigs belonging to clusters CL2, CL4 and CL18, obtained after RepeatExplorer clustering, are available as Supplementary Data S1–S3, respectively.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Lower, S.S.; McGurk, M.P.; Clark, A.G.; Barbash, D.A. Satellite DNA Evolution: Old Ideas, New Approaches. *Curr. Opin. Genet. Dev.* **2018**, *49*, 70–78. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, B.; Sniegowski, P.; Stephan, W. The Evolutionary Dynamics of Repetitive DNA in Eukaryotes. *Nature* **1994**, *371*, 215–220. [[CrossRef](#)] [[PubMed](#)]
- Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. *Genes* **2017**, *8*, 230. [[CrossRef](#)] [[PubMed](#)]
- Gaffney, P.M.; Pierce, J.C.; Mackinley, A.G.; Titchen, D.A.; Glenn, W.K. Pearl, a Novel Family of Putative Transposable Elements in Bivalve Mollusks. *J. Mol. Evol.* **2003**, *56*, 308–316. [[CrossRef](#)] [[PubMed](#)]
- Brajković, J.; Feliciello, I.; Bruvo-Mađarić, B.; Ugarković, D. Satellite DNA-Like Elements Associated With Genes Within Euchromatin of the Beetle *Tribolium castaneum*. *G3—Genes Genom. Genet.* **2012**, *2*, 931–941. [[CrossRef](#)]
- Dias, G.B.; Svartman, M.; Delprat, A.; Ruiz, A.; Kuhn, G.C.S.S. Tetris Is a Foldback Transposon That Provided the Building Blocks for an Emerging Satellite DNA of *Drosophila virilis*. *Genome Biol. Evol.* **2014**, *6*, 1302–1313. [[CrossRef](#)]
- Dias, G.B.; Heringer, P.; Svartman, M.; Kuhn, G.C.S.S. Helitrons Shaping the Genomic Architecture of *Drosophila*: Enrichment of DINE-TR1 in  $\alpha$ - and  $\beta$ -Heterochromatin, Satellite DNA Emergence, and PiRNA Expression. *Chromosom. Res.* **2015**, *23*, 597–613. [[CrossRef](#)] [[PubMed](#)]
- Luchetti, A. TerMITes: Miniature Inverted-Repeat Transposable Elements (MITes) in the Termite Genome (Blattodea: Termitoidea). *Mol. Genet. Genom.* **2015**, *290*, 1499–1509. [[CrossRef](#)] [[PubMed](#)]
- Šatović, E.; Vojvoda Zeljko, T.; Luchetti, A.; Mantovani, B.; Plohl, M. Adjacent Sequences Disclose Potential for Intra-Genomic Dispersal of Satellite DNA Repeats and Suggest a Complex Network with Transposable Elements. *BMC Genom.* **2016**, *17*, 997. [[CrossRef](#)]
- Feliciello, I.; Pezer, Ž.; Kordiš, D.; Mađarić, B.B.; Ugarković, Đ. Evolutionary History of Alpha Satellite DNA Repeats Dispersed within Human Genome Euchromatin. *Genome Biol. Evol.* **2020**, *14561197*, 2125–2138. [[CrossRef](#)] [[PubMed](#)]
- Meštrović, N.; Mravinac, B.; Pavlek, M.; Vojvoda-Zeljko, T.; Šatović, E.; Plohl, M. Structural and Functional Liaisons between Transposable Elements and Satellite DNAs. *Chromosom. Res.* **2015**, *23*, 583–596. [[CrossRef](#)] [[PubMed](#)]
- Tørresen, O.K.; Star, B.; Mier, P.; Andrade-Navarro, M.A.; Bateman, A.; Jarnot, P.; Gruca, A.; Grynberg, M.; Kajava, A.V.; Promponas, V.J.; et al. Tandem Repeats Lead to Sequence Assembly Errors and Impose Multi-Level Challenges for Genome and Protein Databases. *Nucleic Acids Res.* **2019**, *47*, 10994–11006. [[CrossRef](#)] [[PubMed](#)]
- Sedlazeck, F.J.; Lee, H.; Darby, C.A.; Schatz, M.C. Piercing the Dark Matter: Bioinformatics of Long-Range Sequencing and Mapping. *Nat. Rev. Genet.* **2018**, *19*, 329–346. [[CrossRef](#)] [[PubMed](#)]
- Šatović, E.; Tunjić Cvitanić, M.; Plohl, M. Tools and Databases for Solving Problems in Detection and Identification of Repetitive DNA Sequences. *Period. Biol.* **2020**, *121–122*, 7–14. [[CrossRef](#)]
- Kim, Y.B.; Oh, J.H.; McIver, L.J.; Rashkovetsky, E.; Michalak, K.; Garner, H.R.; Kang, L.; Nevo, E.; Korol, A.B.; Michalak, P. Divergence of *Drosophila melanogaster* Repeatomes in Response to a Sharp Microclimate Contrast in Evolution Canyon, Israel. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10630–10635. [[CrossRef](#)]
- Ruiz-Ruano, F.J.; López-León, M.D.; Cabrero, J.; Camacho, J.P.M. High-Throughput Analysis of the Satellitome Illuminates Satellite DNA Evolution. *Sci. Rep.* **2016**, *6*, 28333. [[CrossRef](#)] [[PubMed](#)]
- Ruiz-Ruano, F.J.; Cabrero, J.; López-León, M.D.; Sánchez, A.; Camacho, J.P.M. Quantitative Sequence Characterization for Repetitive DNA Content in the Supernumerary Chromosome of the Migratory Locust. *Chromosoma* **2018**, *127*, 45–57. [[CrossRef](#)] [[PubMed](#)]
- Utsunomia, R.; Ruiz-Ruano, F.J.; Silva, D.M.Z.A.; Serrano, É.A.; Rosa, I.F.; Scudeler, P.E.S.; Hashimoto, D.T.; Oliveira, C.; Camacho, J.P.M.; Foresti, F. A Glimpse into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes). *Front. Genet.* **2017**, *8*, 103. [[CrossRef](#)] [[PubMed](#)]
- Palacios-Gimenez, O.M.; Koelman, J.; Flores, M.P.; Bradford, T.M.; Jones, K.K.; Cooper, S.J.B.; Kawakami, T.; Suh, A. Comparative Analysis of Morabine Grasshopper Genomes Reveals Highly Abundant Transposable Elements and Rapidly Proliferating Satellite DNA Repeats. *BMC Biol.* **2020**, *18*, 199. [[CrossRef](#)]
- Sader, M.; Vaio, M.; Cauz-Santos, L.A.; Dornelas, M.C.; Vieira, M.L.C.; Melo, N.; Pedrosa-Harand, A. Large vs. Small Genomes in *Passiflora*: The Influence of the Mobilome and the Satellitome. *Planta* **2021**, *253*, 86. [[CrossRef](#)]

21. Šatović Vukšić, E.; Plohl, M. Exploring Satellite DNAs: Specificities of Bivalve Mollusks Genomes. In *Satellite DNAs in Physiology and Evolution*; Ugarković, Đ., Ed.; Springer International Publishing: Cham, Switzerland, 2021; pp. 57–83. [\[CrossRef\]](#)
22. Bai, C.M.; Xin, L.S.; Rosani, U.; Wu, B.; Wang, Q.C.; Duan, X.K.; Liu, Z.H.; Wang, C.M. Chromosomal-Level Assembly of the Blood Clam, *Scapharca (Anadara) broughtonii*, Using Long Sequence Reads and Hi-C. *Gigascience* **2019**, *8*, giz067. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Dong, Y.; Zeng, Q.; Ren, J.; Yao, H.; Lv, L.; He, L.; Ruan, W.; Xue, Q.; Bao, Z.; Wang, S.; et al. The Chromosome-Level Genome Assembly and Comprehensive Transcriptomes of the Razor Clam (*Sinonovacula constricta*). *Front. Genet.* **2020**, *11*, 664. [\[CrossRef\]](#)
24. Song, H.; Guo, X.; Sun, L.; Wang, Q.; Han, F.; Wang, H.; Wray, G.A.; Davidson, P.; Wang, Q.; Hu, Z.; et al. The Hard Clam Genome Reveals Massive Expansion and Diversification of Inhibitors of Apoptosis in Bivalvia. *BMC Biol.* **2021**, *19*, 15. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Smith, C.H. A High-Quality Reference Genome for a Parasitic Bivalve with Doubly Uniparental Inheritance (Bivalvia: Unionida). *Genome Biol. Evol.* **2021**, *13*, evab029. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Sun, J.; Zhang, Y.; Xu, T.; Zhang, Y.; Mu, H.; Zhang, Y.; Lan, Y.; Fields, C.J.; Hui, J.H.L.; Zhang, W.; et al. Adaptation to Deep-Sea Chemosynthetic Environments as Revealed by Mussel Genomes. *Nat. Ecol. Evol.* **2017**, *1*, 0121. [\[CrossRef\]](#)
27. Takeuchi, T.; Kawashima, T.; Koyanagi, R.; Gyoja, F.; Tanaka, M.; Ikuta, T.; Shoguchi, E.; Fujiwara, M.; Shinzato, C.; Hisata, K.; et al. Draft Genome of the Pearl Oyster *Pinctada fucata*: A Platform for Understanding Bivalve Biology. *DNA Res.* **2012**, *19*, 117–130. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Mun, S.; Kim, Y.J.; Markkandan, K.; Shin, W.; Oh, S.; Woo, J.; Yoo, J.; An, H.; Han, K. The Whole-Genome and Transcriptome of the Manila Clam (*Ruditapes philippinarum*). *Genome Biol. Evol.* **2017**, *9*, 1487–1498. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Renaut, S.; Guerra, D.; Hoeh, W.R.; Stewart, D.T.; Bogan, A.E.; Ghiselli, F.; Milani, L.; Passamonti, M.; Breton, S. Genome Survey of the Freshwater Mussel *Venustaconcha ellipsiformis* (Bivalvia: Unionida) Using a Hybrid de novo Assembly Approach. *Genome Biol. Evol.* **2018**, *10*, 1637–1646. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Petitpierre, E.; Gatewood, J.M.; Schmid, C.W. Satellite DNA from the Beetle *Tenebrio molitor*. *Experientia* **1988**, *44*, 498–499. [\[CrossRef\]](#)
31. Meštrović, N.; Plohl, M.; Mravinac, B.; Ugarković, D. Evolution of Satellite DNAs from the Genus *Palorus*-Experimental Evidence for the “Library” Hypothesis. *Mol. Biol. Evol.* **1998**, *15*, 1062–1068. [\[CrossRef\]](#)
32. Gall, J.G.; Cohen, E.; Polan, M.L. Repetitive DNA Sequences in *Drosophila*. *Chromosoma* **1971**, *33*, 319–344. [\[CrossRef\]](#)
33. Murgarella, M.; Puiu, D.; Novoa, B.; Figueras, A.; Posada, D.; Canchaya, C. A First Insight into the Genome of the Filter-Feeder Mussel *Mytilus galloprovincialis*. *PLoS ONE* **2016**, *11*, e0151561. [\[CrossRef\]](#)
34. Uliano-Silva, M.; Dondero, F.; Dan Otto, T.; Costa, I.; Lima, N.C.B.; Americo, J.A.; Mazzoni, C.J.; Prosdocimi, F.; Rebelo, M.d.F. A Hybrid-Hierarchical Genome Assembly Strategy to Sequence the Invasive Golden Mussel, *Limnoperna fortunei*. *Gigascience* **2018**, *7*, gix128. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Gaiero, P.; Vaio, M.; Peters, S.A.; Schranz, M.E.; De Jong, H.; Speranza, P.R. Comparative Analysis of Repetitive Sequences among Species from the Potato and the Tomato Clades. *Ann. Bot.* **2019**, *123*, 521–532. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Peñaloza, C.; Gutierrez, A.P.; Eory, L.; Wang, S.; Guo, X.; Archibald, A.L.; Bean, T.P.; Houston, R.D. A Chromosome-Level Genome Assembly for the Pacific Oyster (*Crassostrea gigas*). *Gigascience* **2021**, *10*, giab020. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Thomas, J.; Pritham, E.J. Helitrons, the Eukaryotic Rolling-Circle Transposable Elements. *Microbiol. Spectr.* **2015**, *3*, 3–4. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Xiong, W.; Dooner, H.K.; Du, C. Rolling-Circle Amplification of Centromeric Helitrons in Plant Genomes. *Plant J.* **2016**, *88*, 1038–1045. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Tunjić-Cvitanić, M.; Pasantes, J.J.; García-Souto, D.; Cvitanić, T.; Plohl, M.; Šatović-Vukšić, E. Satellitome Analysis of the Pacific Oyster *Crassostrea gigas* Reveals New Pattern of Satellite DNA Organization, Highly Scattered across the Genome. *Int. J. Mol. Sci.* **2021**, *22*, 6798. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Kourtidis, A.; Drosopoulou, E.; Pantzartzi, C.N.; Chintiroglou, C.C.; Scouras, Z.G. Three New Satellite Sequences and a Mobile Element Found inside HSP70 Introns of the Mediterranean Mussel (*Mytilus galloprovincialis*). *Genome* **2006**, *49*, 1451–1458. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Šatović, E.; Plohl, M. Tandem Repeat-Containing MITE Elements in the Clam *Donax trunculus*. *Genome Biol. Evol.* **2013**, *5*, 2549–2559. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Vojvoda Zeljko, T.; Pavlek, M.; Meštrović, N.; Plohl, M. Satellite DNA-like Repeats Are Dispersed throughout the Genome of the Pacific Oyster *Crassostrea gigas* Carried by Helitron Non-Autonomous Mobile Elements. *Sci. Rep.* **2020**, *10*, 15107. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Novák, P.; Neumann, P.; Pech, J.; Steinhaisl, J.; Macas, J. RepeatExplorer: A Galaxy-Based Web Server for Genome-Wide Characterization of Eukaryotic Repetitive Elements from Next-Generation Sequence Reads. *Bioinformatics* **2013**, *29*, 792–793. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Tunjić Cvitanić, M.; Vojvoda Zeljko, T.; Pasantes, J.J.; García-Souto, D.; Gržan, T.; Despot-Slade, E.; Plohl, M.; Šatović, E. Sequence Composition Underlying Centromeric and Heterochromatic Genome Compartments of the Pacific Oyster *Crassostrea gigas*. *Genes* **2020**, *11*, 695. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Silva, B.S.M.L.; Heringer, P.; Guilherme, B.D.; Svartman, M.; Kuhn, G.C.S. De Novo Identification of Satellite DNAs in the Sequenced Genomes of *Drosophila virilis* and *D. Americana* Using the RepeatExplorer and TAREAN Pipeline. *PLoS ONE* **2019**, *14*, e0223466. [\[CrossRef\]](#)

- 
46. Montiel, E.E.; Panzera, F.; Palomeque, T.; Lorite, P.; Pita, S. Satellitome Analysis of *Rhodnius Prolixus*, One of the Main Chagas Disease Vector Species. *Int. J. Mol. Sci.* **2021**, *22*, 6052. [[CrossRef](#)] [[PubMed](#)]
  47. Bouilly, K.; Chaves, R.; Leitao, A.; Benabdelmouna, A.; Guedes-Pinto, H. Chromosomal Organization of Simple Sequence Repeats in Chromosome Patterns. *J. Genet.* **2008**, *87*, 119–125. [[CrossRef](#)] [[PubMed](#)]