

Two parallel chromosome-level reference genomes to support restoration and aquaculture of European flat oyster *Ostrea edulis*

Tim P. Bean¹  | Arnaud Tanguy² | Carolina Peñaloza¹ | Manu Kumar Gundappa¹  |
Isabelle Boutet² | Ross D. Houston¹ | Daniel J. Macqueen¹ | Pierre Boudry³ 

¹The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, Edinburgh, UK

²CNRS, UMR 7144, Station Biologique de Roscoff, Sorbonne Université, Roscoff, France

³Département Ressources Biologiques et Environnement, Ifremer, Plouzané, France

Correspondence

Tim P. Bean, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, Edinburgh, UK.

Email: tim.bean@roslin.ed.ac.uk

Funding information

Biotechnology and Biological Sciences Research Council, Grant/Award Number: BB/S004181/1, BB/S004343/1 and BBS/E/D/10002070; Fonds Européen pour les Affaires Maritimes et la Pêche, Grant/Award Number: PFEA470017FA1000016

Abstract

This volume of *Evolutionary Applications* sees the publication of two genomes for the European native flat oyster *Ostrea edulis*, a species of significant evolutionary, ecological and commercial value. Each is a highly contiguous chromosome-level assembly from individuals of different genetic backgrounds, which have been benchmarked against one another. This situation has resulted from the serendipitous discovery that two independent research groups were both deep into the process of building, annotating and investigating separately produced assemblies. Due to constraints with funder requirements and the need to recognize early career researchers for their work, alongside the technical challenge of integrating assemblies from two very different genomes, there was limited capacity to merge the sequences into one publication at the stage of discovery. This issue is likely to become very common over the next few years until the technologies for working with multiple genomes at once, for example, graph genomes, become commonplace in nonmodel species. Consequently, both of our teams have decided to collaborate rather than compete, recognizing the benefit to copublishing two separate genome resources for the research community, each with distinct scientific investigations, and working collaboratively to benchmark the assemblies.

KEYWORDS

aquaculture, bivalve, genomics, mollusc, restoration, shellfish

This volume of *Evolutionary Applications* includes the deliberate simultaneous publication of two highly contiguous, chromosome-level reference genome assemblies for the native European flat oyster *Ostrea edulis* (Boutet et al., 2022; Gundappa et al., 2022). *Ostrea edulis* is an iconic native species across Europe where it was heavily fished and consumed through to the early 1900s. Natural beds extend in shallow waters from the Norwegian coast of North Sea, along the Atlantic Ocean down to the Mediterranean and through

to the Black Sea. Non-native *O. edulis* populations have naturalized in Atlantic waters of North America, namely in Nova Scotia, New Brunswick and Maine. The commercial success of this species in Europe and heavy mortality resulting from two parasitic diseases eventually resulted in widespread collapse of fisheries, and stocks have remained in a depleted state ever since (Laing et al., 2006). Scientific interest in this species has, however, recently seen a resurgence due to burgeoning activities associated with ecological

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.

© 2022 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

restoration (Pogoda et al., 2019; Zu Ermgassen et al., 2020), and a revival in aquaculture production in Europe, as a complement to the Pacific oyster *Crassostrea gigas*. As such, it is expected that the timing of these genome releases will provide an extremely useful foundation for the purposes of comparative genomics (Regan et al., 2021), population genetics (Šegvić-Bubić et al., 2020), selective breeding (Gutierrez et al., 2018; Naciri-Graven et al., 1998; Peñaloza et al., 2022) and deep explorations into the biology of this species (see Colsoul et al., 2021 for review). Moreover, the joint effort to publish these resources at the same time, in the same journal issue, with overlapping authorship, shows the collaborative nature of those working in this field. However, we also feel there is value in sharing details of the circumstances that led to the publishing of two separate genomes and the scientific potential of having access to multiple genomes for a single species.

Compared with virtually all the classic model species, molluscs are understudied and are fundamentally different from all model taxa on which animal genomic techniques were pioneered (Hedgecock et al., 2005). This is perhaps best highlighted amongst the bivalve molluscs, where several key observations exemplify extreme genetic complexity. Primarily, bivalves exhibit exceptionally high levels of polymorphism and genomic plasticity, which appears to persist even with high levels of consanguineous mating (Zhang et al., 2012). This variation has been observed from the level of marker heterozygosity (Hedgecock et al., 2005) through to significant structural variation (Jiao et al., 2021) including measures of interindividual genomic divergence of 0.21 (calculated by dividing the total length of the divergent regions by genome size, Qi et al., 2021), and gene presence-absence variation suggesting one-third of genes in a mussel pan-genome (includes core genes essential for all individuals, and dispensable genes found only in some) are dispensable (Gerdol et al., 2020). This high level of variability results in high heterozygosity, including in the *Ostrea* genome published in this volume of Evolutionary Applications, which has a heterozygosity rate of 1.07% (Gundappa et al., 2022). As such, the assembly of high-quality reference genomes for bivalve species has historically proved challenging (Davison & Neiman, 2021), in part due to the difficulty in identifying or breeding inbred animals (Dégremont et al., 2022; Zhang et al., 2012).

An associated peculiar observation is that wild cohorts of oyster are often observed to have a low effective population size relative to census population sizes (e.g. N_e/N of lower than 10^{-5} in *C. gigas* but also low ratios in *O. edulis*) (Hedgecock, 1994; Lallias et al., 2010), leading to 'genetic patchiness' in nature. This is most likely caused by a phenomenon known as a 'sweepstakes reproduction strategy', which occurs when few parents successfully mate and suggests that a very small proportion of gametes produced by a population are viable in wild situations, either through physical or through genetic barriers to reproduction (Plough, 2016). From a physical perspective, the peculiar characteristics in *O. edulis* include sperm clustered in spermatozeugmata (O'Foighil, 1989; Suquet et al., 2018), consecutive hermaphroditism, where individuals are able to change sex during reproduction (Cole, 1942; Zapata-Restrepo et al., 2019)

and brooding of larvae, are known to influence the genetics and the reproductive dynamics of the species (Lallias et al., 2010). From a genomics perspective, incompatibilities between parent animals caused by aforementioned structural variation (Gerdol et al., 2020) could potentially be a factor. There are further complications; severe segregation distortion in the form of heterozygous or homozygous deficiencies is frequently observed at nuclear markers in progenies of pair crosses (Dégremont et al., 2022; Plough, 2016). Regardless, the interplay between mating success or failure and population level outcomes is hugely important from the perspective of selective breeding and will likely require a great deal of genomic analysis to fully resolve.

Finally, the way through which this assorted diversity accumulates and is maintained in bivalves is not yet fully elucidated, but there are indications this is achieved by multiple mechanisms, often at play during the gametogenesis and larval phases. This encompasses multiple examples of variation being generated de novo, such as transposon activity (Peñaloza et al., 2021; Zhang et al., 2012), and the suggestion of spontaneous replication errors in the meiotic process associated with high fecundity (Plough & Hedgecock, 2016). The result of these mechanisms is an increase in genomic variation and a high genetic load, some of which is expressed temporally and purged through purifying selection during larval development (Plough, 2016). As recent studies have suggested that genetic diversity may be maintained via temporally balancing selection during the larval phase of oyster development (Durland et al., 2021), further studies that combine genomic approaches with classical Mendelian experimental designs will be key to resolve the remaining uncertainty in these systems, and unravel the function they hold in the oyster life cycle (Durland et al., 2022; Hedgecock, 2022).

All these phenomena come together to form highly successful life histories. In total, there are over 20,000 known bivalve species, which together inhabit virtually all known aquatic (both freshwater and marine) ecosystems (Vaughn & Hoellein, 2018). A key to unraveling the genetics underlying these life histories is the development of reference genomes for the key species in these clades. A reference genome is usually a haploid approximation of a genome sequence for a single species. In some animal groups, such as mammals, more than one individual sequence may be merged into one contiguous sequence. For example, the human genome reference (GRCh38) is composed of DNA sequences from 338 individuals and includes sequences of over 100,000 insertions (Wong et al., 2020). The clear benefit of having one, high-quality reference genome is that of consensus and convenience. However, in species such as oysters, we must consider whether this method is the most appropriate, or even possible. The hugely variable genomes in bivalves, especially the exceptionally high heterozygosity and structural variation, means that a single reference genome may poorly represent the extent and organization of the polymorphism landscape contained in bivalve genomes.

Leading up to the co-publication of two distinct *O. edulis* reference genomes, we found ourselves in a situation where two distinct research teams were concurrently working towards the same goal,

which was realized at an advanced stage of analysis through a mutual network focussed on *O. edulis* restoration (Pogoda et al., 2019). With the continuing increase in accessibility of sequencing resources, we can perhaps assume that this situation is likely to be common. There are, for example, two chromosome-level genome assemblies of *C. gigas* (Peñaloza et al., 2021; Qi et al., 2021), with at least two more currently being sequenced (*pers comms*). In the case of *O. edulis*, we agreed from that point onwards to work collaboratively to ensure both genomes were co-published and co-released, providing a positive endpoint for the scientific community. Collaboration allowed both teams to share sequence data that improved the quality of both reference genomes. In particular, resequencing data from a pair crosses in France allowed the development of a medium-density SNP linkage map harnessing the expertise of the team in the UK. This linkage map has been used to confirm the orientation of large blocks of sequence in both genomes and provides confidence on the accuracy of both assembled sequences. We also coordinated submission of both genomes to the public archives (NCBI) to occur on the same day, such that neither 'came first', which we felt was a fair outcome for the early career researchers that spent a huge amount of effort building and analysing these resources.

Both *O. edulis* reference genomes are now publicly available. The production, presentation and analysis of multiple genomes is not straightforward, and it is common that one simply supersedes the last, but several techniques have been developed to deal with this situation, most notably 'graph genomes' that capture variation in multiple rather than one reference. Graph genomes allow mapping of individual variation from single-nucleotide polymorphisms through to major structural differences in the same species, which can comprehensively and accurately capture the differences between thousands of individuals (Rakocevic et al., 2019). This is particularly relevant for bivalves, as large-scale structural variation between individuals and populations is almost certainly not being fully captured by the current generation of SNP level analyses (Calcino et al., 2021). However, while graph genomes are likely to become a feature of oyster genomics in coming years, this strategy will require a great deal of computing power. The benefits may only truly be realized when hundreds of individuals representing different populations have been sequenced, allowing population genetic parameters and structure to be more thoroughly characterized.

The next steps for oyster genome biology are likely to include the development of a phased (haplotype-resolved) assembly of an individual (Cheng et al., 2022), capturing heterozygosity inclusive of structural differences and then most likely the development of a bivalve graph genome. Once in place, graph genomes will be added to and incrementally increase the accuracy of all genomic analyses. This includes transcriptomic analysis, studies of local genetic adaptation, genomic and epigenomic architectures of key traits, genome-assisted breeding and possibly an understanding of the scope of the pan-genome in bivalves. Most importantly, for the field to progress, the required collaborations must be forged from an early stage. Research communities for specific nonmodel organisms should coordinate and agree on a strategy for which specimens to

sequence and why, then communicate openly and widely about this strategy. Here, we see not only the production of the first reference sequences for *O. edulis*, but also the beginnings of collaborations for the production of one of the first bivalve graph genomes.

ACKNOWLEDGEMENTS

We thank the Native Oyster Restoration Alliance (NORA) for facilitating collaborative research and exchanges contributing to the protection and ecological restoration of the European native oyster. This study was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) under the AquaLeap consortium (grant codes: BB/S004181/1, BB/S004343/1), the Fonds Européen pour les Affaires Maritimes et la Pêche (FEAMP) (grant code: PFEA470017FA1000016), and received additional support from the BBSRC Institute Strategic Programme grant BBS/E/D/10002070. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

DATA AVAILABILITY STATEMENT

There is no data associated with this article.

ORCID

Tim P. Bean  <https://orcid.org/0000-0002-2544-9918>

Manu Kumar Gundappa  <https://orcid.org/0000-0003-4328-2178>

<https://orcid.org/0000-0003-4328-2178>

Pierre Boudry  <https://orcid.org/0000-0002-5150-2276>

REFERENCES

- Boutet, I., Alves Monteiro, H. J., Baudry, L., Takeuchi, T., Bonnard, E., Billoud, B., Farhat, S., Gonzales-Araya, R., Salaun, B., Andersen, A., Toullec, J.-Y., Lallier, F., Flot, J. F., Guiglielmoni, N., Guo, X., Allam, B., Pales-Espinoza, E., Hemmer-Hansen, J., Marbouty, M., ... Tanguy, A. (2022). Chromosomal assembly of the flat oyster (*Ostrea edulis* L.) genome as a new genetic resource for aquaculture. *Evolutionary Applications*. <https://doi.org/10.1111/eva.13462>
- Calcino, A. D., Kenny, N. J., & Gerdol, M. (2021). Single individual structural variant detection uncovers widespread hemizyosity in molluscs. *Philosophical Transactions of the Royal Society B*, 376(1825), 20200153.
- Cheng, H., Jarvis, E. D., Fedrigo, O., Koepfli, K. P., Urban, L., Gemmill, N. J., & Li, H. (2022). Haplotype-resolved assembly of diploid genomes without parental data. *Nature Biotechnology*, 1–4. <https://doi.org/10.1038/s41587-022-01261-x>
- Cole, H. A. (1942). Primary sex-phases in *Ostrea edulis*. *Journal of Cell Science*, 2(331), 317–356.
- Colsoul, B., Boudry, P., Pérez-Parallé, M. L., Bratoš Cetinić, A., Hugh-Jones, T., Arzul, I., Mérou, N., Wegner, K. M., Peter, C., Merk, V., & Pogoda, B. (2021). Sustainable large-scale production of European flat oyster (*Ostrea edulis*) seed for ecological restoration and aquaculture: A review. *Reviews in Aquaculture*, 13(3), 1423–1468.
- Davison, A., & Neiman, M. (2021). Pearls of wisdom—A Theo murphy issue on molluscan genomics. *Philosophical Transactions of the Royal Society B*, 376(1825), 20200151.
- Dégremont, L., Maurouard, E., Jourdan, A., Lamy, J. B., & Boudry, P. (2022). SNP-based parentage analyses over two successive generations demonstrates the feasibility of efficient production of inbred lines in the Pacific oyster (*Crassostrea gigas*) by self-fertilization of

- simultaneous hermaphrodites despite severe inbreeding depression. *Aquaculture*, 547, 737443.
- Durland, E., De Wit, P., & Langdon, C. (2021). Temporally balanced selection during development of larval Pacific oysters (*Crassostrea gigas*) inherently preserves genetic diversity within offspring. *Proceedings of the Royal Society B*, 288(1958), 20203223.
- Durland, E., Wit, P. D., & Langdon, C. (2022). Genetic changes in larval oysters are more abundant and dynamic than can be explained by rare events or error: A response to Hedgecock (2022). *Proceedings of the Royal Society B*, 289(1976), 20220197.
- Gerdol, M., Moreira, R., Cruz, F., Gómez-Garrido, J., Vlasova, A., Rosani, U., Venier, P., Naranjo-Ortiz, M. A., Murgarella, M., Greco, S., & Balseiro, P. (2020). Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. *Genome Biology*, 21(1), 1–21.
- Gundappa, M. K., Peñaloza, C., Regan, T., Boutet, I., Tanguy, A., Houston, R. D., Bean, T. B., & Macqueen, D. J. (2022). A chromosome level reference genome for European flat oyster (*Ostrea edulis* L.). *Evolutionary Applications*. <https://doi.org/10.1111/eva.13460>
- Gutierrez, A. P., Matika, O., Bean, T. P., & Houston, R. D. (2018). Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): Potential of low-density marker panels for breeding value prediction. *Frontiers in Genetics*, 9, 391.
- Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms? In *Genetics and evolution of aquatic organisms* (pp. 122–134). Chapman and Hall.
- Hedgecock, D. (2022). No evidence for temporally balanced selection on larval Pacific oysters *Crassostrea gigas*: A comment on Durland et al. (2021). *Proceedings of the Royal Society B*, 289(1976), 20212579.
- Hedgecock, D., Gaffney, P. M., Goulletquer, P., Guo, X., Reece, K., & Warr, G. W. (2005). The case for sequencing the Pacific oyster genome. *Journal of Shellfish Research*, 24(2), 429–441.
- Jiao, Z., Tian, Y., Hu, B., Li, Q., & Liu, S. (2021). Genome structural variation landscape and its selection signatures in the fast-growing strains of the Pacific oyster, *Crassostrea gigas*. *Marine Biotechnology*, 23(5), 736–748.
- Laing, I., Walker, P., & Areal, F. (2006). Return of the native—is European oyster (*Ostrea edulis*) stock restoration in the UK feasible? *Aquatic Living Resources*, 19(3), 283–287.
- Lallias, D., Taris, N., Boudry, P., Bonhomme, F., & Lapègue, S. (2010). Variance in reproductive success of flat oyster *Ostrea edulis* L. assessed by parentage analyses in natural and experimental conditions. *Genetics Research*, 92, 175–187.
- Naciri-Graven, Y., Martin, A. G., Baud, J. P., Renault, T., & Gérard, A. (1998). Selecting the flat oyster *Ostrea edulis* (L.) for survival when infected with the parasite *Bonamia ostreae*. *Journal of Experimental Marine Biology and Ecology*, 224(1), 91–107.
- O'Foighil, D. (1989). Role of spermatozogmata in the spawning ecology of the brooding oyster *Ostrea edulis*. *Gamete Research*, 24, 219–228.
- Peñaloza, C., Barria, A., Papadopoulou, A., Hooper, C., Preston, J., Green, M., Helmer, L., Kean-Hammerson, J., Nascimento-Schulze, J. C., Minardi, D., Gundappa, M. K., Macqueen, D. J., Hamilton, J., Houston, R. D., & Bean, T. P. (2022). Genome-wide association and genomic prediction of growth traits in the European flat oyster (*Ostrea edulis*). *Frontiers in Genetics*, 13, 926638. <https://doi.org/10.3389/fgene.2022.926638>
- Peñaloza, C., Gutierrez, A. P., Eöry, L., Wang, S., Guo, X., Archibald, A. L., Bean, T. P., & Houston, R. D. (2021). A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. *GigaScience*, 10(3), giab020.
- Plough, L. V. (2016). Genetic load in marine animals: A review. *Current Zoology*, 62(6), 567–579.
- Pogoda, B., Brown, J., Hancock, B., Preston, J., Pouvreau, S., Kamermans, P., Sanderson, W., & Von Nordheim, H. (2019). The native oyster restoration Alliance (NORA) and the Berlin oyster recommendation: Bringing back a key ecosystem engineer by developing and supporting best practice in Europe. *Aquatic Living Resources*, 32, 13.
- Qi, H., Li, L., & Zhang, G. (2021). Construction of a chromosome-level genome and variation map for the Pacific oyster *Crassostrea gigas*. *Molecular Ecology Resources*, 21(5), 1670–1685.
- Rakocevic, G., Semenyuk, V., Lee, W. P., Spencer, J., Browning, J., Johnson, I. J., Arsenijevic, V., Nadj, J., Ghose, K., Suci, M. C., & Ji, S. G. (2019). Fast and accurate genomic analyses using genome graphs. *Nature Genetics*, 51(2), 354–362.
- Regan, T., Stevens, L., Peñaloza, C., Houston, R. D., Robledo, D., & Bean, T. P. (2021). Ancestral physical stress and later immune gene family expansions shaped bivalve mollusc evolution. *Genome Biology and Evolution*, 13(8), evab177.
- Šegvić-Bubić, T., Žužul, I., Talijančić, I., Ugrin, N., Lepen Pleić, I., Žuvić, L., Stagličić, N., & Grubišić, L. (2020). Translocation and aquaculture impact on genetic diversity and composition of wild self-sustainable *Ostrea edulis* populations in the Adriatic Sea. *Frontiers in Marine Science*, 7, 84.
- Suquet, M., Pouvreau, S., Queau, I., Boulais, M., Le Grand, J., Ratskol, D., & Cosson, J. (2018). Biological characteristics of sperm in European flat oyster (*Ostrea edulis*). *Aquatic Living Resources*, 31, 20.
- Vaughn, C. C., & Hoellein, T. J. (2018). Bivalve impacts in freshwater and marine ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 49, 183–208.
- Wong, K. H., Ma, W., Wei, C. Y., Yeh, E. C., Lin, W. J., Wang, E. H., Su, J. P., Hsieh, F. J., Kao, H. J., Chen, H. H., & Chow, S. K. (2020). Towards a reference genome that captures global genetic diversity. *Nature Communications*, 11(1), 1–11.
- Zapata-Restrepo, L. M., Hauton, C., Williams, I. D., Jensen, A. C., & Hudson, M. D. (2019). Effects of the interaction between temperature and steroid hormones on gametogenesis and sex ratio in the European flat oyster (*Ostrea edulis*). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 236, 110523. <https://doi.org/10.1016/j.cbpa.2019.06.023>
- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P. W. H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., ... Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, 490, 49–54. <https://doi.org/10.1038/nature11413>
- Zu Ermgassen, P. S., Bonačić, K., Boudry, P., Bromley, C. A., Cameron, T. C., Colsoul, B., Coolen, J. W., Frankić, A., Hancock, B., van der Have, T. M., & Holbrook, Z. (2020). Forty questions of importance to the policy and practice of native oyster reef restoration in Europe. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30(11), 2038–2049.

How to cite this article: Bean, T. P., Tanguy, A., Peñaloza, C., Gundappa, M. K., Boutet, I., Houston, R. D., Macqueen, D. J., & Boudry, P. (2022). Two parallel chromosome-level reference genomes to support restoration and aquaculture of European flat oyster *Ostrea edulis*. *Evolutionary Applications*, 15, 1709–1712. <https://doi.org/10.1111/eva.13465>