PUBLISHED: 3 APRIL 2017 | VOLUME: 1 | ARTICLE NUMBER: 0142

GENOME EVOLUTION

# Shellfish genes

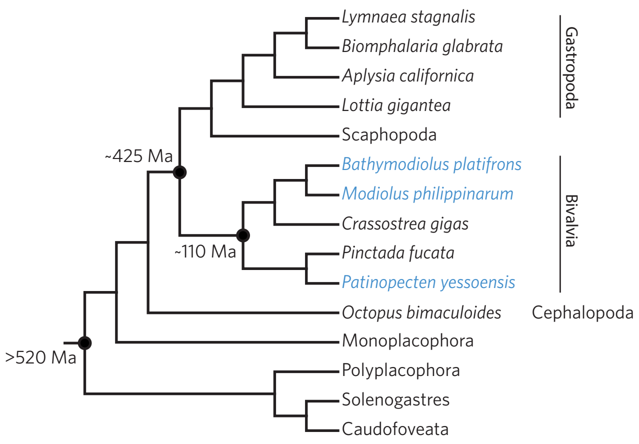
Three new bivalve genomes are resources for comparative genomics over broad timescales, providing a glimpse into the evolution of understudied marine animals and their adaptations to extreme environments.

Kenneth M. Halanych and Kevin M. Kocot

Mollusca is the second most species-rich animal phylum with over 100,000 recognized species, including many of importance to humans as food sources, disease vectors, damaging invasive species or aesthetic resources (for example, in jewellery manufacture). Despite their immense importance, only a few mollusc genomes are currently publicly available. Contributions by Sun *et al.*[1] and Wang *et al.*[2] in *Nature Ecology & Evolution* help rectify this problem by sequencing and analysing the genomes of three bivalve species. Both reports employ comparative and evolutionary genomic approaches to advance understanding of genomic evolution in bivalves. Sun *et al.* compare the mussel *Bathymodiolus platifrons*, which lives in deep-sea chemosynthetic environments, to its shallow-water relative *Modiolus philippinarum*. Wang *et al.* dig deeper by comparing the genome of the scallop *Patinopecten yessoensis* (also called *Mizuhopecten yessoensis*) to other lophotrochozoans as well as across animal diversity.

Free-spawning marine invertebrate genomes are typically challenging to assemble because of high levels of heterozygosity. Wang *et al.* mitigated this problem by sequencing a single male *P. yessoensis* sired from a self-fertilized hermaphrodite. Using Illumina short-insert paired-end and long-insert mate-pair libraries, they generated a reasonably high-quality assembly (Fig. 1). Using a 2b-RAD linkage map approach, they were able to assign most of the scaffolds to the 19 chromosomes of *P. yessoensis*. In addition to heterozygosity, marine invertebrate genomes are often as replete with repetitive elements as human genomes. Sun *et al.* also used a combined paired-end and mate-pair sequencing strategy to sequence two mussel genomes; the *B. platifrons* assembly totalled 1.64 Gb and *M. philippinarum* assembly totalled 2.38 Gb. Although the $N_{50}$ value for *M. philippinarum*, for example, is lower than those routinely reported using similar approaches, for example bird genomes[3], this assembly is commendable given the presence of over 700 Mb (~30%) of repetitive DNA in this genome.

Wang *et al.* compared the scallop genome to those of other animals to inform understanding of the ancestral bilaterian genome. They paid particular attention to identifying conserved macrosynteny regions. Given that the genomes in question probably diverged over 540 million years ago (Ma) and that recombination is known to significantly rearrange chromosomal content even between recently diverged lineages, the finding of 17 macrosynteny regions in scallops that are conserved across bilaterians[4] is remarkable. Moreover, the existence of these macrosynteny regions begs the question of why have some organisms undergone very little genomic rearrangement since they split off from the last common bilaterian ancestor (for example, scallop and lancelet) while others have relatively highly rearranged genomes (for example, *Drosophila*). Although the exact complement of genes in these conserved regions needs to be further investigated, one possible reason to maintain linked clusters of genes is regulation of gene pathways. Well-studied examples of transcriptional regulation of linked genes include the homeobox-containing gene clusters. Wang *et al.* show that in the scallop genome, the *Hox* and *ParaHox* clusters are intact, but collinear expression is restricted to subclusters of genes. Interestingly, oysters have similar subcluster expression patterns but their *Hox* cluster is no longer linked



| | Assembly / total genome size (Mbp) | Scaffold $N_{50}$ (kbp) | Gene models |
|---|---|---|---|
| *Lymnaea stagnalis* | 997 / 1,190 | 45 | 78,075 |
| *Biomphalaria glabrata* | 916 / 929 | 48 | 36,675 |
| *Aplysia californica* | 927 / 1,760 | 918 | 27,591 |
| *Lottia gigantea* | 360 / 421 | 1,870 | 23,822 |
| *Bathymodiolus platifrons* | 1,640 / ? | 343.2 | 36,549 |
| *Modiolus philippinarum* | 2,380 / ? | 100.2 | 33,584 |
| *Crassostrea gigas* | 558 / 890 | 402 | 45,406 |
| *Pinctada fucata* | 1,000 / 1,100 | 1.6 | 23,257 |
| *Patinopecten yessoensis* | 988 / 1,430 | 804 | 26,415 |
| *Octopus bimaculoides* | 2,400 / 2,700 | 475 | 23,994 |

**Figure 1 |** Cladogram showing relationships among molluscs with publicly available sequenced genomes and summary statistics for those data. The species in blue are those sampled by Sun *et al.* and Wang *et al.*

in a single unit[5], suggesting the interplay between co-linear transcription and gene arrangement may be more nuanced than currently appreciated. The genome of a scallop will also be an important resource to the community of scientists studying the fascinating non-cephalic eyes[6] found along the mantle margin of these animals.

In contrast to spanning the diversity of bilaterian animals, Sun *et al.* sequenced and compared genomes of the relatively closely related shallow-water *M. philippinarum* and the deep-sea chemosynthetic *B. platifrons* to explore organismal adaption to different environments. Notably, they were inferred to have diverged around 110 Ma — around the same time placental mammals radiated. One significant difference between these species is that *B. platifrons* makes use of horizontally transmitted endosymbiotic bacteria to derive energy from chemosynthetic environments. Establishing such symbiosis is often linked to modified immune responses[7] as has been seen in the giant tubeworms[8] and the bobtail squid[9]. Sun *et al.* document an expanded repertoire of genes involved in immunorecognition (important for recognizing potential endosymbiont partners), endocytosis (important for allowing the endosymbiont to become intracellular) and apoptosis (potentially a mechanism for regulating endosymbiont growth or utilization). Additionally, the genome of *B. platifrons* shows expanded families of genes that may aid survival in extreme environments by dealing with toxicity and stabilizing protein structure (for example, heat shock proteins). Thus, differences in genetic composition highlighted by Sun *et al.* fit expectations of how the genome of a vent- or seep-dwelling species may differ from a shallow water relative.

Findings in both manuscripts highlight the need for additional metazoan genome sequencing efforts outside of vertebrates, insects and nematodes. Although these bivalve genomes are welcome new resources, their annotations are still largely dependent on comparison to lineages that contain model organisms limiting our ability to understand, for example, the 474 protein domains apparently unique to *B. platifrons*. In particular, Lophotrochozoa, the bilaterian subclade with the greatest variation of animal body plans, is considerably undersampled and even within Mollusca only three of eight major lineages have available genomes. In order to draw firm conclusions about genome evolution, more comparisons, like those undertaken by Wang *et al.* and Sun *et al.* are needed, especially from marine representatives. ❒

*Kenneth M. Halanych is in the Department of Biological Sciences and the Molette Biology Laboratory for Environmental and Climate Change Studies, Auburn University, Alabama 36849, USA. Kevin M. Kocot is in the Department of Biological Sciences and the Alabama Museum of Natural History, The University of Alabama, Tuscaloosa, Alabama 35487, USA.*
*e-mail:* ken@auburn.edu; kmkocot@ua.edu

## References

1. Sun, J. *et al. Nat. Ecol. Evol.* **1,** 0121 (2017).
2. Wang, S. *et al. Nat. Ecol. Evol.* **1,** 0120 (2017).
3. Zhang, G. *et al. GigaScience* **3,** 26 (2014).
4. Simakov, O. *et al. Nature* **493,** 526–531 (2013).
5. Zhang, G. *et al. Nature* **490,** 49–54 (2012).
6. Speiser, D. I., Gagnon, Y. L., Chhetri, R. K., Oldenburg, A. L. & Johnsen, S. *Integr. Comp. Biol.* **56,** 796–808 (2016).
7. Chu, H. & Mazmanian, S. K. *Nat. Immunol.* **14,** 668–675 (2013).
8. Nussbaumer, A. D., Fisher, C. R. & Bright, M. *Nature* **441,** 345–348 (2006).
9. Nyholm, S. V., Stewart, J. J., Ruby, E. G. & McFall-Ngai, M. J. *Environ. Microbiol.* **11,** 483–493 (2009).

## Competing interests

The authors declare no competing financial interests.