CellPress

Current Biology
Dispatches

Genome Biology: Unconventional DNA Repair in an Extreme Genome

David E.K. Ferrier* and Shunsuke Sogabe*
The Scottish Oceans Institute, Gatty Marine Laboratory, University of St Andrews, East Sands, St Andrews KY16 8LB, UK
*Correspondence: dekf@st-andrews.ac.uk (D.E.K.F.), ss358@st-andrews.ac.uk (S.S.)
https://doi.org/10.1016/j.cub.2018.09.004

DNA is under constant assault and needs to efficiently repair breaks. A tiny marine relative of vertebrates makes do with an alternative mechanism to the canonical repair system, which coincides with it possessing one of the most extreme animal genomes.

Organisms need to be able to repair their DNA given the barrage that the molecule has to endure, including exposure to radiation, whether from the disaster of a nuclear accident, the therapeutic requirement of getting an X-ray or simply sun-bathing in a summer heat-wave and being blasted by UV radiation. One sobering thought is that a 10-hour transatlantic flight exposes a person to enough cosmic rays to induce approximately 0.002 double strand breaks (DSBs) per cell [1] — which adds up when one considers the number of cells in our body. It is not just external factors that need to be guarded against either. Errors in DNA metabolism, presence of DNA-digesting nuclease enzymes and reactive oxygen species all take their toll. In addition, DNA can be 'deliberately' broken, with programmed genome rearrangements such as the V(D)J recombination process that is a vital component of our immune system and protects us from pathogens. So DSBs are not all bad. However, given this constant onslaught, molecular systems exist that efficiently repair DSBs, and the main proteins of this system are widely conserved from humans to yeast. So it was rather shocking when the genome of a tiny marine invertebrate, from a group of animals closely related to vertebrates, was sequenced and found to lack the genes that encode these core repair proteins. Now, as reported in this issue of *Current Biology*, Deng *et al.* [2] show that this animal, *Oikopleura dioica*, uses an alternative repair system, which has features similar to a system also present in ourselves. Athough this alternative system often leads to high rates of mutation and genome rearrangement, leading to problems such as cancer in humans, it seems that this animal has turned this to its advantage and enabled the evolution of an extremely compact and unusual genome.

The larvacean *O. dioica* is a remarkable animal in many respects. The tadpole-like adult is easy to find in the wild, if one is prepared to drag a plankton net for a few minutes through almost any marine surface waters. Larvaceans exist in almost astronomical numbers in the sea and play a significant role in the marine habitat. They create beautiful cellulose-based 'houses' [3] that efficiently channel seawater and filter out their phytoplankton food. This makes them major grazers of marine primary producers [4]. They also discard and rebuild their houses several times a day, which along with their high feeding levels and consequent abundant production of faecal material makes them major contributors to the carbon flux via marine snow (the constant rain of organic matter from the upper ocean layers to the depths).

O. dioica also has a remarkably short life cycle of only four days, and this 'live-fast-die-young' approach to existence correlates with possibly the smallest known genome of any animal, at under 70 megabases (Mb) [5]. This is dwarfed by our own genome of around 3,000 Mb. This extreme genome compaction has coincided with extensive rearrangement of the genes [5], which is exemplified by the Hox genes. These genes are renowned for patterning the development of the anterior–posterior (or head-to-tail) axis of animal embryos and 'telling' our developing foetus how many ribs to develop and where to put them, or telling a fly embryo where to put its wings, for example. These genes are also famed for often being found in a cluster in animal genomes: the Hox gene cluster. However, they are completely scattered around the *O. dioica* genome [6], the cluster being 'atomized' in the parlance of the Hox classification system of Duboule [7]. Further *O. dioica* genome oddities include high rates of intron turnover, use of non-canonical splice sites and intron insertion via reverse splicing [5]. All together, *O. dioica* does not have a conventional animal genome.

This lack of conventionality has been particularly intriguing since *O. dioica* is a member of the same phylum as ourselves, the Chordata (Figure 1A). *O. dioica* is an appendicularian larvacean, in the subphylum Urochordata, which together with another invertebrate subphylum, the Cephalochordata (containing the amphioxus lancelets), and the vertebrates like ourselves constitute the Chordata. Larvaceans, like *O. dioica*, occupy a key position in the chordate phylogeny for helping to understand the evolution of vertebrates, since the urochordates are the closest invertebrates to the vertebrates themselves (urochordates and vertebrates together often being called the Olfactores; Figure 1A). The whole genome duplicated twice at the origin of the vertebrates [8], which has resulted in vertebrates having more genes than invertebrates and more complicated gene networks due to extra duplicates of various genes. Invertebrates are thus useful systems for pulling apart gene networks without the hindrance of this extra complexity, and since the urochordates are the closest invertebrates to vertebrates, their networks are more likely to have similarities based on homology with vertebrates than the networks of more conventional invertebrate study systems like fruit flies and nematodes. The development of urochordates as

important study systems for genetic mechanisms in evolutionary and developmental biology has thus progressed dramatically in recent years, with not only larvaceans being studied but several ascidians as well, including the sea squirt *Ciona* [9,10].

This logic of a closer phylogenetic position likely equating to greater similarity of gene networks is not always realised in urochordates, however, due to their relatively elevated rates of molecular evolution [11]. The mechanism for dealing with repair of DSBs is a case in point. There are various DSB repair processes, one of which is non-homologous end joining (NHEJ). This typically involves the proteins Ku70, Ku80, DNA-PKcs, Lig4, XRCC4, XLF and Artemis (Figure 1B) [2], and is sometimes called classical or canonical NHEJ (cNHEJ). Several of the cNHEJ genes are absent from the unusual genome of *O. dioica*, and it now turns out that the cNHEJ machinery is absent from larvaceans as a whole (judging from the absence of the genes from half a dozen appendicularian genomes sequenced by Deng *et al.* [2]), making this another intriguing example of evolution via gene loss [12]. To understand how larvaceans can cope without this cNHEJ machinery, Deng *et al.* [2] assessed the DSB repair abilities of *O. dioica* in a couple of elegant ways. First, they injected a series of linear pieces of plasmid DNA into fertilised eggs, then allowed some time for repair to occur, which often involves circularization of the plasmids such that they can be extracted from the animals and then transformed into bacterial cells and the site of the repair sequenced and characterised. Their second approach involved adapting the technique of inducing DNA breaks by the CRISPR/Cas9 system, specifically targeting the *brachyury* gene in this case. Isolating and sequencing the *brachyury* gene sequence from many targeted animals allowed the characterisation of the repaired sites in endogenous *O. dioica* DNA. A wider comparison to sites of insertion/deletion between Norwegian and Japanese *O. dioica* as well as to other urochordates, such as the sea squirt *Ciona intestinalis*, which does have cNHEJ, helped to further support the conclusions that *O. dioica* tends to repair its DSBs with an end-joining mechanism that can exploit short stretches of sequence similarity (micro-homologies) of only a few base-pairs, and that the repair often involves deletion of some intervening nucleotides. This all has intriguing resemblance to the alternative NHEJ (aNHEJ) process of humans and other animals. This aNHEJ (or simply alternative end-joining (aEJ) [13]) is less understood than cNHEJ and the component proteins are only in the earliest stages of being revealed [14].

An intriguing point here is that aNHEJ is often associated with genomic rearrangements, such as translocations. Translocations are a common feature of cancers and so there is much interest in understanding the mechanisms of aNHEJ for potential therapeutic exploitation [14]. The types of rearrangements associated with aNHEJ also tend to include deletions that reduce the number of micro-homology repeats (Figure 1B) [2,14]. This raises the interesting possibility that the reliance of *O. dioica* on aNHEJ rather than

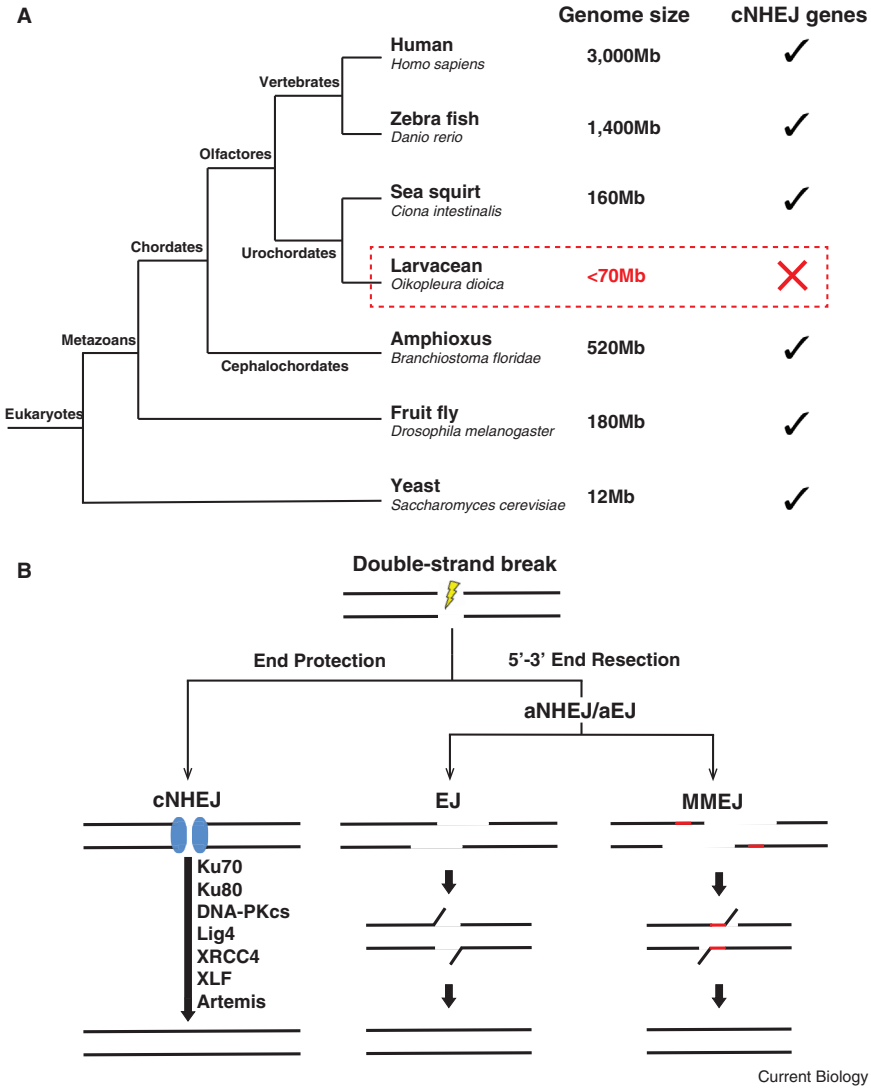

Figure 1. Phylogenetic location of *Oikopleura dioica* and selected double strand break (DSB) repair processes.
(A) *O. dioica* is a larvacean in the urochordate sister group to vertebrates within the chordate phylum of metazoan animals, with an unusually small genome for an animal that is distinctive (amongst other things) for loss of the genes encoding canonical non-homologous end-joining (cNHEJ) proteins. (B) Various mechanisms are used to repair DSBs, including cNHEJ and alternative NHEJ (also simply known as alternative end-joining (aEJ)), which can be further subdivided into end-joining (EJ) and microhomology mediated end-joining (MMEJ) (reviewed in [13,14]). *O. dioica* has lost the genes encoding proteins that perform cNHEJ. The cNHEJ proteins for vertebrates are shown here, with orthologues of many of these being used in other animals such as *Drosophila melanogaster*. *O. dioica* repairs its DSBs with a mechanism that results in joins resembling those formed by aNHEJ/aEJ processes [2].

cNHEJ could explain the unusual compaction of this animal's genome, perhaps in concert with the adaptation of this species to a life cycle of extreme brevity [2]. From the wider appendicularian genome sequencing done by Deng *et al.* [2], which will certainly become a valuable resource for further comparative genomics, it is clear that the larvacean ancestor was able to survive without the cNHEJ machinery, and that the consequent dramatic genome compaction and rearrangement that occurred (with this rearrangement still ongoing) is due to its likely dependence on an aNHEJ-like pathway. A prerequisite for this must have been that this ancestor's genome was presumably already 'primed' to be scrambled without fatally disrupting gene regulation at high frequency. Thus, the precursor to the appendicularian ancestor must already have evolved to a state in which its gene regulatory mechanisms were presumably focused on short-range gene-specific processes rather than the long-range multigenic mechanisms involving topologically associated domains and genomic regulatory blocks more typical of many other animal genomes [15,16]. Why this should be the case still remains an evolutionary mystery, which may be resolved by the burgeoning work on urochordates.

This highlights that not all animal genomes are necessarily working in the same way and that studying a diversity of species is important. Such an increased diversity of study species is extremely valuable [17], improving our capabilities to address pressing issues such as the various aspects of biology centred on genomics. For example, given the importance of DSB repair to understanding not only genome evolution but also human diseases with roots in gene rearrangements such as various types of cancers, having a whole animal study system that has done away with the cNHEJ machinery is likely to prove a valuable resource for the discovery and elucidation of alternative DSB repair pathways. Perhaps then larvaceans could become a useful new model for not only the evolution of development and genome rearrangement, but also oncogenesis, with a role in the search for the mechanisms of aNHEJ/aEJ and the consequent potential biomedical benefits.

REFERENCES

1. Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. Mol. Cell *40*, 179–204.
2. Deng, W., Henriet, S., and Chourrout, D. (2018). Prevalence of mutation-prone microhomology-mediated end joining in a chordate lacking the c-NHEJ pathway. Curr. Biol. *28*, 3337–3341.
3. Hosp, J., Sagane, Y., Danks, G., and Thompson, E.M. (2012). The evolving proteome of a complex extracellular matrix, the *Oikopleura* house. PLoS One *7*, e40172.
4. Lindsay, M.C.M., and Williams, G.D. (2010). Distribution and abundance of larvaceans in the Southern Ocean between 30 and 80°E. Deep-Sea Res. II *57*, 905–915.
5. Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J.-M., Da Silva, C., Brinkmann, H., Mikhaleva, J., Olsen, L.C., Jubin, C., Cañestro, C., *et al*. (2010). Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. Science *330*, 1381–1385.
6. Seo, H.C., Edvardsen, R.B., Maeland, A.D., Bjordal, M., Jensen, M.F., Hansen, A., Flaat, M., Weissenbach, J., Lehrach, H., Wincker, P., *et al*. (2004). Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. Nature *431*, 67–71.
7. Duboule, D. (2007). The rise and fall of Hox gene clusters. Development *134*, 2549–2560.
8. Putnam, N.H., Butts, T., Ferrier, D.E.K., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., *et al*. (2008). The amphioxus genome and the evolution of the chordate karyotype. Nature *453*, 1064–1071.
9. Lemaire, P. (2011). Evolutionary crossroads in developmental biology: the tunicates. Development *138*, 2143–2152.
10. Sasakura, Y. (2018). Transgenic Ascidians, First Edition (Singapore: Springer Nature).
11. Holland, L.Z. (2015). Genomics, evolution and development of amphioxus and tunicates: the Goldilocks principle. J. Exp. Zool. B Mol. Dev. Evol. *324*, 342–352.
12. Albalat, R., and Cañestro, C. (2016). Evolution by gene loss. Nat. Rev. Genet. *17*, 379–391.
13. Pannunzio, N.R., Watanabe, G., and Lieber, M.R. (2018). Nonhomologous DNA end-joining for repair of DNA double-strand breaks. J. Biol. Chem. *293*, 10512–10523.
14. Sallmyr, A., and Tomkinson, A.E. (2018). Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. J. Biol. Chem. *293*, 10536–10546.
15. Franke, M., and Gómez-Skarmeta, J.L. (2018). An evolutionary perspective of regulatory landscape dynamics in development and disease. Curr. Opin. Cell Biol. *55*, 24–29.
16. Kikuta, H., Fredman, D., Rinkwitz, S., Lenhard, B., and Becker, T.S. (2007). Retroviral enhancer detection insertions in zebrafish combined with comparative genomics reveal genomic regulatory blocks – a fundamental feature of vertebrate genomes. Genome Biol. *1* (8 Suppl.), S4.
17. Martinez, P. (2018). The comparative method in biology and the Essentialist Trap. Front. Ecol. Evol. *6*, 130.

Microbiology: Peeling Back the Layers of Bacterial Envelope Mechanics

Joshua W. Shaevitz
Department of Physics and Lewis-Sigler Institute, Princeton University, Princeton, NJ 08544, USA
Correspondence: shaevitz@princeton.edu
https://doi.org/10.1016/j.cub.2018.09.023

The Gram-negative cell envelope has two important mechanical elements. Whereas the cell wall bears the brunt of the turgor pressure during normal growth, the outer membrane also provides necessary rigidity under physical stress.

In the everyday life of a bacterium, physical insults — in the form of mechanical stresses encountered during motility, in the presence of flow, or during rapid changes in the osmotic environment —constantly challenge the