A Reference Genome from the Symbiotic Hydrozoan, *Hydra viridissima*

Mayuko Hamada,*†‡,1 Noriyuki Satoh,* and Konstantin Khalturin*

*Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan, †Ushimado Marine Institute, Okayama University, Setouchi, Okayama 701-4303, Japan, and ‡Zoological Institute, Kiel University, Kiel 24118, Germany

ORCID IDs: 0000-0001-7306-2032 (M.H.); 0000-0002-4480-3572 (N.S.); 0000-0003-4359-2993 (K.K.)

**ABSTRACT** Various Hydra species have been employed as model organisms since the 18th century. Introduction of transgenic and knock-down technologies made them ideal experimental systems for studying cellular and molecular mechanisms involved in regeneration, body-axis formation, senescence, symbiosis, and holobiosis. In order to provide an important reference for genetic studies, the *Hydra magnipapillata* genome (species name has been changed to *H. vulgaris*) was sequenced a decade ago (Chapman et al., 2010) and the updated genome assembly, Hydra 2.0, was made available by the National Human Genome Research Institute in 2017. While *H. vulgaris* belongs to the non-symbiotic brown hydra lineage, the green hydra, *Hydra viridissima*, harbors algal symbionts and belongs to an early diverging clade that separated from the common ancestor of brown and green hydra lineages at least 100 million years ago (Schwentner and Bosch 2015; Khalturin et al., 2019). While interspecific interactions between *H. viridissima* and endosymbiotic unicellular green algae of the genus *Chlorella* have been a subject of interest for decades, genomic information about green hydras was nonexistent. Here we report a draft 280-Mbp genome assembly for *Hydra viridissima* strain A99, with a scaffold N50 of 1.1 Mbp. The *H. viridissima* genome contains an estimated 21,476 protein-coding genes. Comparative analysis of Pfam domains and orthologous proteins highlights characteristic features of *H. viridissima*, such as diversification of innate immunity genes that are important for host-symbiont interactions. Thus, the *H. viridissima* assembly provides an important hydrozoan genome reference that will facilitate symbiosis research and better comparisons of metazoan genome architectures.

KEYWORDS green hydra *Hydra viridissima* A99 whole genome sequencing de novo assembly symbiosis

The Cnidaria is an evolutionarily ancient and well-defined phylum, characterized by the possession of nematocytes (Brusca et al. 2016). Cnidarian species belong to the Medusozoa, which comprises the Hydrozoa, Scyphozoa, Cubozoa, and Anthozoa (Figure 1A). Although cnidarian morphology exhibits astonishingly diverse forms and life styles, those of fresh water hydrozoans of the genus *Hydra* are relatively simple. *Hydra* possess only a polyp stage, while other medusozoans exhibit alternation of polyp and medusa stages.

With its simple body structure and easy laboratory cultivation, *Hydra* has been an experimental model for studying cellular and molecular mechanisms underlying the formation of the body axis (Bode 2011), regeneration (Trembley et al. 1744; Bode 2003; Holstein et al. 2003), and also holobiotic relationships with microbiota (Deines and Bosch 2016). Introduction of transgenic and knock-down technologies further promoted these studies (Wittlieb et al. 2006). In order to provide genetic information for these studies, the *Hydra magnipapillata* (now classified as *H. vulgaris*) genome was sequenced in 2010 (Chapman et al. 2010), and an improved version was published in 2017 (Hydra 2.0 Web Portal: https://research.nhgri.nih.gov/hydra/). While *H. vulgaris* belongs to the non-symbiotic brown hydra lineage, the green hydra, *Hydra viridissima*, establishes a mutualistic relationship with microalgae and exchanges metabolites with its symbionts (Figure 1C) (Muscatine 1965; Cernichiari et al. 1969; Mews 1980; McAuley 1991). While symbiosis with dinoflagellates...
belong to the basally branching lineage in the genus Hydra. According to several phylogenetic reconstructions, Hydra viridissima (A) Phylogenetic position of H. viridissima (red) within the phylum Cnidaria. (B) Relationship of Hydra viridissima strain A99 (red) with other H. viridissima strains and brown hydra species, based on phylogenetic analysis with the NJ method using cytochrome c oxidase subunit I (COI) gene sequences. The genomic region in H. viridissima strain A99 as another high-quality Hydra reference genome. We report significant characteristics of the green hydra genome, including transposable elements, innate immunity-related genes, and genes that determine its body plan.

MATERIALS AND METHODS

Hydra and extraction of DNA

The Australian Hydra viridissima strain A99, which was kindly provided by Dr. Richard Campbell, at the University of California at Irvine, was used in this study. Polyps were maintained at 18°C on a 12-hour light/dark cycle and fed with Artemia nine times a week. DNA for genome sequencing were isolated from about 1000 polyps that were clonally cultured. Before genomic DNA extraction, symbiotic algae in H. viridissima were removed by photo-bleaching with 5 μM DCMU (3-(3,4-dichlorophenyl)-1,1-dimethyleurea), as described previously (Pardy 1976; Habetha et al. 2003). To remove contamination from other organisms, polyps were starved and treated with antibiotics (50 mg/L ampicillin, rifampicin, neomycin, and streptomycin) for one week.

After several rounds of washing in sterilized culture medium, polyps were lysed in DNA extraction buffer (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 25 mM EDTA, pH 8.0, 0.5% SDS) and digested with 100 mg/L Proteinase K. Genomic DNA was extracted using the standard phenol-chloroform method with 100 mg/L RNaseA treatment. The quantity of DNA was determined using a NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA), and the quality of high molecular-weight DNA was checked using agarose gel electrophoresis.

Sequencing of genomic DNA

In paired-end library preparations for genome sequencing, genomic DNA was fragmented with a Focused-ultrasonicator M220 (Covaris Inc., Woburn, MA, USA), following the manufacturers’ protocols. These libraries were sequenced on the Illumina HiSeq system with 600-cycle chemistry (2 × 300 bp). Genome sequencing statistics is shown in Table S1A.
RNA extraction and sequencing

Total RNA was extracted from about 1000 polyps in six different conditions (with or without symbiotic algae; in light or dark conditions, and treated with antibiotics or DMCU with symbiotic algae) using Trizol reagent (Thermo Fisher Scientific) and an RNaseasy Mini kit (Qiagen, Hilden, Germany). The quantity of RNA was determined with a NanoDrop (Thermo Fisher Scientific). Quality of total RNA was checked with a BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA). For mRNA-seq, libraries were produced using an Illumina TruSeq Stranded mRNA Sample Prep Kit and were sequenced on HiSeq 2000 instruments using 2 × 150-cycle chemistry. mRNA-sequencing statistics are shown in Table S1B.

Assembly and gene prediction

Sequencing reads of genomic DNA were assembled using the Newbler Assembler, version 2.8 (Roche, Penzberg, Germany), and subsequent scaffolding was performed with SPAdes (Boetzer et al. 2011). Gaps inside scaffolds were closed with paired-end and mate-pair data using GapCloser of the Short Oligonucleotide Analysis Package (Luo et al. 2012). Then one round of Haplomerger2 processing pipeline (Huang et al. 2017) was applied to eliminate redundancy in scaffolds and to merge haplotypes. For gene model prediction, we used a species-specific gene prediction model that was trained based on mapping of the Hydra viridissima transcriptome and raw RNAseq reads against the genome assembly. Mapping and gene structure annotation were performed using the PASA pipeline (Haas et al. 2003; Stanke et al. 2006). Genome completeness was evaluated using BUSCO (Benchmarking Universal Single-Copy Ortholog) (Sepey et al. 2019). RNA-Seq transcripts were mapped to the genome assembly with BWA.

Genome size estimation

Genome size was estimated from raw paired-end reads by k-mer distribution analysis. Jellyfish v2.0.0 was used to count k-mers and their frequencies (Marçais and Kingsford 2011). The Hydra viridissima genome size was estimated from k-mer distribution frequencies using the GenomeScope web tool (Vurture et al. 2017) (http://qb.cshl.edu/genomescope/).

Analysis of repetitive elements

Repetitive elements in the draft genome assembly of Hydra viridissima were identified de novo with RepeatScout version 1.0.5 (http://www.repeatmasker.org/RepeatModeler) and RepeatMasker version 4.0.6 (http://www.repeatmasker.org). Repetitive elements were filtered by length and occurrence so that only sequences longer than 50 bp and present more than 10 times in the genome were retained. The resulting sets of repetitive elements were annotated using BLASTN and BLASTX searches against RepeatMasker.lib (35,996 nucleotide sequences) and RepeatPeps.lib (10,544 peptides) bundled with RepeatMasker version 4.0.6. The results of both searches were combined, and BLASTX results were given priority in cases where BLASTN and BLASTX searches gave conflicting results.

Analysis of Hydra viridissima genes

For comparative analysis of H. viridissima genes among cnidarians, protein sequences were obtained from Hydra 2.0 web portal (https://research.nhgri.nih.gov/hydra/) and the Compagen server (http://www.compgen.org) for Hydra vulgaris (H. magnipapillata) and Nematostella vectensis (Nematostellidae) as well as other cnidarians (Table 1).

| Species                  | Geographical origin | Number of Scaffolds | Scaffold N50 (Mbp) | Number of genes | Mean gene length (bp) | BUSCO (complete) % |
|--------------------------|---------------------|---------------------|--------------------|-----------------|-----------------------|------------------|
| Hydra vulgaris           | Laboratory strain   | 2,677               | 0.1                | 31,452          | 6,873                 | 83.9             |
| Hydra magnipapillata     | Laboratory strain   | 20,914              | 0.1                | 33,820          | 12,378                | 80.2             |

Table 1: Comparison of the genome assembly statistics of cnidarians.
strain 105, from JGI (https://genome.jgi.doe.gov/Nemvel/Nemvel. home.html) for *Nematostella vectensis* (Putnam et al. 2007), from MARIMBA (available at http://marimba.obs-vlfr.fr/organism/ Clytia/hemisphaerica) for *Clytia hemisphaerica* (Leclère et al. 2019), from the genome project website of OIST Marine Genomics Unit (https://marinegenomics.oist.jp/gallery/gallery/index) for *Acropora digitifera* (Shinzato et al. 2011), for *Morbakka virulenta* and for the Atlantic Ocean strain of *Aurelia aurita* (Khalturin et al. 2019). We used protein models derived from the Hydra 2.0 assembly of the *H. vulgaris* genome for all comparative analyses with *H. viridissima* as this assembly has higher continuity (scaffold N50 ~1Mbp) and BUSCO values than the originally published assembly (Chapman et al., 2010). For comparative reasons, statistics and results obtained with the Hydra 1.0 assembly (Chapman et al., 2010) and Hydra 2.0 assembly (https://research.nihri.nih.gov/hydra/) are shown side by side in Table 1 and Tables 4-7.

In comparative analyses, domain searches against the Pfam database (Pfam-A.hmm) were performed using HMMER (Finn et al. 2016), and ortholog gene grouping employed OrthoFinder (Emms and Kelly 2015). To classify homeodomain-containing proteins, BLAST searches and phylogenetic analyses were performed. Homeodomain sequences in various animals were obtained from the Homeobox Database (http://homeodb.zoo.ox.ac.uk/families/get?og=All) (Zhong and Holland 2011).

For phylogenetic analysis, multiple alignments were produced with CLUSTALX (2.1) with gap trimming (Larkin et al. 2007). Sequences of poor quality that did not align well were deleted using BioEdit (Hall 1999). Phylogenetic analyses were performed using the Neighbor-Joining method (Saitou and Nei 1987) in CLUSTALX with default parameters (1,000 bootstrap tests and 111 seeds). Representative phylogenetic trees were drawn using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Gene/protein IDs used for phylogenetic analysis are shown in the trees (Figs S2 and S3).

**Data availability**

This whole-genome shotgun sequencing project has been deposited at DDBJ/ENA/GenBank under BioSample ID SAMN09635813 and BioProject ID: PRJNA480404. RNA-seq reads have been deposited at SRA of NCBI (SRX6792700-SRX6792705). Genome sequences, gene models, and a genome browser are also accessible at the website of the OIST Marine Genomics Unit Genome Project (https://marinegenomics.oist.jp/hydraviridissima_A99). A genome browser was established for assembled sequences using the JBrowse 1.12.3 (Skinner et al. 2009). Gene annotations from the protein
domain search and BLAST search are likewise shown on the site. Reagents, software and datasets used in this study are listed in the Reagent Table. k-mer frequency distribution plots in the Hydra viridissima A99 genome is found in Figure S1. Phylogenetic analysis of PRD genes is presented in Figure S3. Sequencing statistics for Hydra viridissima A99 are in Table S1. A summary of repetitive sequences in the Hydra viridissima A99 genome assembly are found in Table S2. Pfam domain-containing genes in the Hydra viridissima A99 genome are available in Table S3. Orthologs enriched in Hydra viridissima A99 (A) and Hydra (B) are in Table S4. Gene IDs of ANTP genes in Hydra viridissima A99 are in Table S5. Supplemental material available at figshare: https://doi.org/10.25387/g3.12911426.

RESULTS AND DISCUSSION

Genome architecture of Hydra viridissima

Hydra viridissima appears green because of the symbiotic Chlorella that inhabit endodermal epithelial cells, and it is smaller than the brown hydra, Hydra vulgaris (Figure 1C). We decoded the genome of H. viridissima strain A99, which is closely related to strain A14c, Wuerzburg and J8 (Figure 1B). We previously reported the genome of its specific symbiotic alga, Chlorella sp. A99, and demonstrated that

| A. Domain | Hvir | Hvul | Ch | Aa | Mv | Nv | Ad | Chi test* |
|-----------|------|------|----|----|----|----|----|-----------|
| NACHT     | 161  | 75   | 42 | 23 | 45 | 39 | 458| 1E-53  |
| NB-ARC    | 106  | 28   | 18 | 4  | 20 | 6 | 220| 5E-53  |
| ATPase_2  | 64   | 19   | 28 | 13 | 14 | 19 | 36 | 7E-35  |
| TIR_2     | 49   | 11   | 19 | 15 | 24 | 17 | 49 | 7E-16  |
| DUF4218   | 47   | 13   | 16 | 12 | 4  | 5  | 2E-66 |
| Endonuclease_7 | 46   | 9    | 3  | 5  | 1  | 1  | 0  | 1E-189 |
| RAG1      | 42   | 13   | 1  | 0  | 4  | 1  | 0  | 2E-160 |
| TIR_3     | 41   | 7    | 7  | 3  | 11 | 15 | 12 | 3E-36  |
| CbiA      | 23   | 7    | 6  | 6  | 7  | 6  | 2  | 8E-19  |
| MarR_2    | 21   | 6    | 4  | 0  | 4  | 3  | 3  | 6E-40  |
| HTH_Tnp_IS630 | 15   | 5    | 1  | 1  | 1  | 0  | 0  | 4E-40  |
| DUF2961   | 14   | 5    | 2  | 3  | 3  | 1  | 3  | 5E-15  |
| TEMEM151  | 10   | 4    | 2  | 0  | 3  | 3  | 3  | 6E-06  |
| DUF1294   | 5    | 1    | 1  | 2  | 0  | 0  | 0  | 6E-07  |

| B. Domain | Hvir | Hvul | Ch | Aa | Mv | Nv | Ad | Chi test* |
|-----------|------|------|----|----|----|----|----|-----------|
| DDE_3     | 365  | 481  | 16 | 31 | 16 | 6  | 17 | 0E+00   |
| Dimer_Tnp_hAT | 296      | 376  | 11 | 44 | 52 | 21 | 27 | 0E+00   |
| HTH_Tnp_Tc3_2 | 266   | 268  | 15 | 23 | 13 | 0  | 3  | 0E+00   |
| HTH_22    | 185  | 176  | 14 | 25 | 4  | 4  | 1  | 1E-307 |
| ANAPC3    | 182  | 157  | 61 | 37 | 36 | 46 | 44 | 2E-117 |
| HTH_23    | 166  | 200  | 6  | 33 | 15 | 8  | 20 | 1E-225 |
| zf-BED    | 116  | 114  | 14 | 8  | 9  | 19 | 5  | 1E-156 |
| HTH_29    | 95   | 121  | 7  | 19 | 1  | 2  | 5  | 8E-148 |
| HTH_psq   | 93   | 165  | 3  | 3  | 6  | 1  | 1  | 3E-189 |
| HTH_28    | 85   | 87   | 2  | 6  | 5  | 6  | 11 | 5E-130 |
| SRP_TPR_like | 83  | 87   | 6  | 2  | 1  | 1  | 1  | 2E-163 |
| DUF4806   | 69   | 54   | 6  | 0  | 1  | 0  | 0  | 3E-157 |
| PAX       | 61   | 57   | 3  | 23 | 4  | 9  | 8  | 4E-63  |
| BTAD      | 32   | 33   | 7  | 1  | 3  | 3  | 3  | 4E-39  |
| Sigma70_r4_2 | 28  | 31   | 2  | 0  | 2  | 1  | 3  | 3E-44  |
| HTH_Tnp_Tc3_1 | 26  | 13   | 1  | 0  | 0  | 0  | 0  | 3E-87  |
| HTH_7     | 21   | 14   | 2  | 0  | 4  | 1  | 1  | 5E-33  |
| CD225     | 21   | 19   | 9  | 2  | 4  | 7  | 7  | 4E-11  |
| DUF2738   | 21   | 12   | 1  | 0  | 0  | 1  | 0  | 3E-56  |
| Sigma70_r4 | 17   | 16   | 1  | 1  | 0  | 2  | 2  | 2E-26  |
| IGFBP     | 15   | 21   | 7  | 5  | 4  | 1  | 3  | 1E-09  |
| Sulfate_transp | 15  | 18   | 2  | 4  | 5  | 4  | 5  | 3E-08  |
| DUF1280   | 13   | 31   | 4  | 0  | 4  | 3  | 2  | 3E-17  |
| HTH_3     | 13   | 11   | 1  | 5  | 2  | 2  | 1  | 4E-11  |
| Polysacc_deac_1 | 13  | 14   | 3  | 6  | 1  | 4  | 2  | 5E-08  |
| DUF4817   | 12   | 9    | 2  | 0  | 1  | 0  | 0  | 6E-21  |
| Ca_chan_IQ | 12   | 16   | 5  | 4  | 5  | 2  | 3  | 1E-05  |
| Torsin    | 11   | 11   | 2  | 2  | 2  | 5  | 5  | 2E-05  |
| CIDEN     | 10   | 7    | 1  | 2  | 3  | 2  | 1  | 1E-07  |
| GRDP-like | 9    | 7    | 2  | 2  | 1  | 2  | 3  | 2E-05  |
| Transposase_mut | 8  | 10   | 1  | 1  | 3  | 0  | 0  | 2E-08  |

Hvir: Hydra viridissima A99, Hvul: Hydra vulgaris (Hydra 2.0), Ch: Clytia hemisphaerica, Mv: Morbakka virulenta, Aa: Aurelia aurita, Nv, Nematostella vectensis, Ad: Acropora digitifera, *Chi-test: evalue of Chi-square test.
metabolic co-dependency exists between *H. viridissima* A99 and the symbiont (Hamada et al. 2018).

The genome of *H. viridissima* A99 was sequenced using the Illumina platform with paired-end and mate pair libraries. Statistics of sequence reads, the assembly, and genome architecture are shown in Table 1. We obtained ~7,070 Mbp of paired-end sequences, and 4,765, 4,769, 3,669 and 3,551 Mbp for 3.2k, 4.6k, 7.8k, and 15.2k insert-size mate-pair sequences, respectively, comprising a total of ~23,826 Mbp (Table S1). The size of the *H. viridissima* genome was estimated at ~254 Mbp using k-mer analysis (k-mer = 19) based on paired-end sequence data (Fig. S1). This indicates that we achieved more than 90-fold sequence coverage of the genome. On the other hand, the total length of the genome sequence assembly reached 284,265,305 bp. That is, the total assembly closely matched the estimated genome size.

Although genomic DNA was extracted from a clonally propagated culture of hydra polyps maintained in the laboratory, heterozygosity was comparatively high (25.4%). Using 67,339,858,036 nucleotides of RNA-seq data, the genome accounts for 91% of the metazoan reference gene set (Table 1). Comparison of *H. viridissima* genome statistics with those of other cnidarian genomes showed that the *H. viridissima* genome assembly is comparable or of even better quality in regard to the scaffold N50 and BUSCO completeness (Table 1).

During assembly and gene annotation, we noticed that scaffold2223, composed of 18,375 base pairs (bp), contained almost the entire *H. viridissima* mitochondrial genome. The mitochondrial genome of *H. viridissima* strain A99 was linear, as reported by Bridge et al. (1992) and Pan et al. (2014b) for the other green hydras, while in brown hydra, *Hydra vulgaris*, mitochondrial genome is composed of two linear molecules (Bridge et al. 1992; Pan et al. 2014a; Pan et al. 2014b).

**Repetitive sequences in the Hydra viridissima genome**

Although the abundance of repetitive sequences in anthozoan genomes is generally low (15~17%), genomes of medusozoans and hydrozoans have comparatively high levels of repetitive sequences, 60% in *H. vulgaris*, 41% in *Clytia*, 45% in *Aurelia*, and 37% in *Morbakka* (Table 1). This was also true of *H. viridissima* (37.5%) (Table 1). DNA transposons were the most abundant type, accounting for approximately 22.4% of the genome (Figure 2A, Table S2). Of these, TcMariner, CMC, Maverick and hAT were the largest components of the genome. It was suggested that a
burst of retrotransposons occurred in the brown hydra lineage after divergence from the green hydra lineage, and may account for the large genomes of brown hydras (Chapman et al. 2010; Wong et al. 2019). Because *H. viridissima* occupies a basal position in the Hydra lineage, and the genome of another hydrozoan, *Clytia*, is smaller (~450Mbp) and has fewer repetitive elements (41%) than *H. vulgaris*, the ancestral Hydra genome was likely rather compact, with fewer retrotransposons. Molecular and evolutionary mechanisms involved in the insertion of LINE components in the *H. vulgaris* genome will be a subject of future studies in relation to diversification and speciation within the Hydra clade.

### Innate immunity-related protein genes in the *Hydra viridissima* genome

Using the Pfam-domain search method, we surveyed genes for protein domains in the *H. viridissima* genome. We found approximately 4,500 different Pfam domains in this species (Table S3), a number comparable to those of other cnidarians. To identify the domains that are enriched in the *H. viridissima* genome, we counted the number of genes with each Pfam domain in cnidarian genomes, and selected the domains of which number are ≥2x higher in the green hydra genome than those of non-symbiotic cnidarians and show significant difference based on Chi-Square test (p-value < 0.001) (Table 2A). Then we checked the number of *H. viridissima*-enriched domains in the genome of the coral, *Acropora digitifera*, since it is also a symbiotic cnidarian. NACHT and NB-ARC, which have similar structures and functions, were the two most highly enriched structures in the genome of the coral, *Acropora digitifera*, since they also have a symbiotic cnidarian. NACHT and NB-ARC, which have similar structures and functions, were the two most highly enriched structures in the genome of the coral, *Acropora digitifera*, since it is also a symbiotic cnidarian. NACHT and NB-ARC, which have similar structures and functions, were the two most highly enriched domains in the *H. viridissima* genome (Table 3A, Table S4). However, these orthologs were not scored in the *H. vulgaris* genome (Table 3A), because *H. vulgaris* has smaller genome and finds all four types of Nod-like receptor family proteins, and their ligand recognition region is a leucine-rich repeat (LLR). On the other hand, in Nod-like receptors of basal metazoans, not only LLR, but also tetratricopeptide repeats (TPR), WD40 repeats, and ankyrin repeats (Ank) are found as repeat domains. We previously showed that *Acropora* has all 4 types of Nod-like receptors, and that those with LLR are the most common (Table 4) (Hamada et al. 2013). In other cnidarians examined, only TPR and WD40 are found as repeat domains of Nod-like receptors, suggesting loss of the other types. Especially in *H. viridissima*, a larger number of genes for Nod-like receptors with TPR were found. In addition, their domain structures in *H. viridissima* vary widely, compared to those of *H. vulgaris*. In addition to NACHT-containing proteins, *H. viridissima* has more genes for TIR containing proteins than that in *H. vulgaris*.

**Figure 3** Schematic representation of domain structures of NACHT/NB-ARC with TIR-domain-containing proteins identified in Hydra. The domain structures and the number of NACHT/NB-ARC or TIR-domain-containing proteins in *Hydra viridissima* A99 (*Hvir*) and *H. vulgaris* (*Hvul*) are shown.
domain-containing proteins such as an interleukin-1 receptor (ILR), which are not found in *H. vulgaris*.

As mentioned above, diverse pattern-recognition receptor-related genes are found in both *H. viridissima* and *Acropora digitifera*. The irms most significant shared attribute is symbiosis, the former with *Chlorella* and the latter with the dinoflagellate, *Symbiodinium*. Therefore, it is likely that the evolutionary development of symbiosis by certain cnidarians required expansion and greater sophistication of innate immunity genes. They may participate in recognition and maintenance of symbiotic organisms in cnidarian tissues. On the other hand, the structures (e.g., repeat combination) of the Nod-like receptors most abundant in green hydras and corals are different. This indicates that species-specific adaptations to the environment and particular symbions occurred independently in these lineages.

**Table 5** Number of putative transcription factor genes (A) and signaling molecule genes (B) in the *Hydra viridissima* genome.

| A. Domain     | Hvir | Hvul v1 | Hvul v2 | Ch  | Aa  | Mv  | Nv  | Ad  |
|---------------|------|---------|---------|-----|-----|-----|-----|-----|
| ARID          | 8    | 7       | 9       | 10  | 10  | 8   | 5   | 8   |
| AT_hook       | 0    | 0       | 0       | 1   | 2   | 0   | 0   | 0   |
| bZIP_1        | 26   | 26      | 30      | 26  | 22  | 25  | 36  | 29  |
| bZIP_2        | 25   | 22      | 23      | 31  | 24  | 27  | 32  | 17  |
| CUT           | 1    | 0       | 1       | 3   | 1   | 1   | 2   | 1   |
| DM            | 6    | 5       | 5       | 7   | 9   | 11  | 12  | 7   |
| Ets           | 9    | 11      | 11      | 13  | 21  | 14  | 16  | 12  |
| Forkhead      | 17   | 17      | 16      | 19  | 15  | 26  | 34  | 22  |
| GATA          | 4    | 4       | 5       | 7   | 7   | 7   | 4   | 5   |
| Hairy_orange  | 0    | 0       | 0       | 1   | 4   | 2   | 6   | 7   |
| HHL           | 33   | 36      | 34      | 44  | 52  | 50  | 72  | 53  |
| HMG_box       | 30   | 33      | 33      | 31  | 30  | 29  | 33  | 27  |
| Homeobox      | 50   | 44      | 49      | 70  | 88  | 82  | 153 | 96  |
| Hormone_recep | 9    | 9       | 9       | 12  | 12  | 9   | 20  | 9   |
| PS3           | 2    | 3       | 3       | 3   | 4   | 4   | 6   | 4   |
| PAX           | 61   | 23      | 57      | 3   | 4   | 11  | 2   | 3   |
| Pou           | 2    | 3       | 3       | 3   | 4   | 4   | 6   | 4   |
| RHD_DNA_bind  | 2    | 1       | 3       | 5   | 2   | 2   | 3   | 2   |
| SRF-TF        | 2    | 2       | 2       | 4   | 3   | 2   | 3   | 1   |
| T-box         | 6    | 7       | 7       | 11  | 10  | 9   | 14  | 10  |
| TF_AP-2       | 1    | 1       | 1       | 2   | 3   | 2   | 1   | 2   |
| zf-C2H2       | 105  | 123     | 121     | 244 | 233 | 118 | 169 | 90  |
| zf-C2HC       | 1    | 2       | 2       | 3   | 2   | 3   | 3   | 4   |
| zf-C4         | 9    | 8       | 8       | 11  | 8   | 9   | 19  | 12  |

| B. Domain     | Hvir | Hvul v1 | Hvul v2 | Ch  | Aa  | Mv  | Nv  | Ad  |
|---------------|------|---------|---------|-----|-----|-----|-----|-----|
| Cbl_N         | 1    | 1       | 1       | 2   | 0   | 1   | 0   | 1   |
| Cbl_N2        | 1    | 1       | 1       | 1   | 0   | 1   | 0   | 1   |
| Cbl_N3        | 1    | 1       | 1       | 2   | 2   | 1   | 0   | 1   |
| DIK           | 1    | 1       | 3       | 3   | 3   | 3   | 4   | 2   |
| FGF           | 16   | 12      | 16      | 10  | 18  | 13  | 13  | 13  |
| Focal_AT      | 1    | 2       | 3       | 1   | 1   | 1   | 0   | 0   |
| G-alpha       | 29   | 28      | 27      | 29  | 29  | 32  | 37  | 22  |
| G-gamma       | 2    | 2       | 3       | 1   | 7   | 5   | 4   | 3   |
| IL3           | 0    | 0       | 0       | 1   | 0   | 0   | 0   | 0   |
| PDGF          | 1    | 2       | 2       | 5   | 2   | 3   | 1   | 6   |
| Phe_ZIP       | 1    | 0       | 1       | 1   | 1   | 1   | 0   | 1   |
| Rabaptin      | 1    | 2       | 1       | 1   | 1   | 1   | 1   | 1   |
| RGS           | 13   | 13      | 12      | 14  | 16  | 16  | 13  | 11  |
| RGS-like      | 2    | 3       | 2       | 1   | 2   | 1   | 0   | 1   |
| STAT_alpha    | 1    | 0       | 0       | 1   | 1   | 2   | 2   | 1   |
| STAT_bind     | 2    | 1       | 3       | 1   | 1   | 1   | 1   | 1   |
| STAT_int      | 2    | 1       | 1       | 1   | 0   | 1   | 0   | 1   |
| TGF beta      | 11   | 11      | 11      | 7   | 9   | 9   | 7   | 10  |
| TGFb_propeptide| 8    | 10      | 9       | 4   | 6   | 8   | 6   | 9   |
| wnt           | 10   | 13      | 11      | 12  | 15  | 17  | 26  | 15  |

Hvir: Hydra viridissima A99, Hvul v1: Hydra vulgaris (Chapman et al., 2010), Hvul v2: Hydra vulgaris (Hydra2.0), Ch: Clytia hemisphaerica, Mv: Morbakka virulenta, Aa: Aurelia aurita, Nv, Nematostella vectensis, Ad: Acropora digitifera.

Genes enriched in the genus Hydra

We further examined Pfam domains overrepresented specifically in *H. viridissima* and others present in both *H. viridissima* and *H. vulgaris*. This was done using the same criteria as above, that is, that the number of domains is ≥2x higher than those in other cnidarians and that the difference is significant by Chi-Square test (p-value < 0.001). (Table 2B). Pfam domain searches and ortholog protein grouping demonstrated that *H. viridissima* and *H. vulgaris* possess many genes encoding domains that function in DNA binding. For example,
genes containing transposase-related domain (DDE_3, Dimer_Tnp_hAT and Transposase_mut) and DNA-binding motif (HTH: helix-turn-helix, zf: zinc finger, Sigma70_r4, CIDE-N, RAG1) were overrepresented in both *H. viridissima* and *H. vulgaris* (Table 2B). In addition, ortholog protein grouping suggested that genes for HTH domain-containing transposase, ATP-dependent DNA helicase PIF1-like protein, DDE superfamily endonuclease, and zinc finger domain-containing transposase were overrepresented in both *Hydra* species (Table 3A). Although the functions of these genes are unknown, they may be involved in genome structure maintenance of *Hydra*, which contains many transposable elements.

Pfam domain searches also demonstrated that genes for proteins containing Sulfate_transp domain and those containing Polysacc_deac_1 domain are enriched in both *H. viridissima* and *H. vulgaris* (Table 2B). Sulfate_transp is found in the sulfate permease family, which is involved in uptake or exchange of inorganic anions, such as sulfate. So far, their functions in *Hydra* are unknown, but may be involved in genome structure maintenance as *Hydra*, which contains many transposable elements.

Figure 4 ParaHox, Hox and NK genes in cnidarians. The putative Homeobox megacluster in the last common ancestor of cnidarians (top) and homeobox genes and their cluster structures in extant cnidarians are represented. ParaHox genes (green boxes); Hox genes (pink boxes); NK genes (blue boxes). Empty boxes indicate lost genes. Horizontal lines (black) indicate chromosome fragments.

Table 6 Number of genes for the subclass of homeodomain-containing proteins in cnidarians

| Class             | Medusozoa | Anthozoa |
|-------------------|-----------|----------|
|                   | Hvir | Hvul v1 | Hvul v2 | Ch | Mv | Aa | Nv | Ad |
| ANTP-HOXL         | 5    | 7       | 7       | 6  | 13 | 13 | 17 | 9  |
| ANTP-NKL          | 9    | 8       | 11      | 20 | 21 | 20 | 65 | 33 |
| PRD               | 21   | 16      | 18      | 18 | 25 | 28 | 43 | 31 |
| LIM               | 4    | 4       | 5       | 5  | 5  | 5  | 5  | 4  |
| TALE              | 4    | 3       | 5       | 5  | 4  | 10 | 6  | 5  |
| SINE              | 2    | 2       | 4       | 4  | 5  | 4  | 6  | 4  |
| POU               | 2    | 2       | 3       | 3  | 4  | 4  | 6  | 4  |
| CERS              | 1    | 1       | 1       | 1  | 1  | 1  | 1  | 1  |
| CUT               | 0    | 0       | 0       | 0  | 0  | 0  | 1  | 2  |
| HNF               | 0    | 0       | 0       | 0  | 0  | 0  | 1  | 1  |
| PROS              | 0    | 0       | 0       | 0  | 0  | 0  | 0  | 0  |
| ZF                | 0    | 0       | 0       | 0  | 0  | 0  | 0  | 0  |
| Total             | 48   | 44      | 48      | 62 | 79 | 85 | 151 | 94 |

Hvir: *Hydra viridissima* A99, Hvul v1: *Hydra vulgaris* (Chapman et al., 2010), Hvul v2: *Hydra vulgaris* (Hydra2.0), Ch: Clytia hemisphaerica, Mv: Morbakka virulenta, Aa: Aurelia aurita, Nv, Nematostella vectensis, Ad: Acropora digitifera.
Table 7 Number of homeodomain-containing genes in the *Hydra viridissima* genome.

| Class | Subclass | Family | Hvir | Hvul v1 | Hvul v2 | Ch | Aa | Mv | Nv | Ad |
|-------|----------|--------|------|---------|---------|----|----|----|----|----|
| ANTP  | HOXL     | Cdx    | 0    | 0       | 0       | 1  | 2  | 1  | 0  | 0  |
|       |          | Evx    | 0    | 0       | 0       | 0  | 0  | 0  | 1  | 1  |
|       |          | Gbx    | 0    | 0       | 0       | 0  | 0  | 0  | 1  | 1  |
|       |          | Gsx    | 1    | 1       | 1       | 1  | 1  | 1  | 1  | 1  |
|       |          | Hox1   | 1    | 1       | 2       | 1  | 1  | 1  | 1  | 1  |
|       |          | Hox2   | 0    | 0       | 0       | 0  | 0  | 0  | 3  | 1  |
|       |          | Hox9-13| 2    | 3       | 3       | 1  | 4  | 5  | 3  | 1  |
|       |          | Mlox   | 1    | 1       | 1       | 0  | 1  | 1  | 4  | 1  |
|       |          | Nlox/Pdx| 0   | 0       | 0       | 1  | 1  | 1  | 1  | 1  |
| NKL   | Barx     | 1    | 1    | 1       | 0       | 1  | 1  | 4  | 2  |    |
|       | Dbx      | 0    | 0    | 0       | 0       | 1  | 1  | 2  | 1  |    |
|       | Dlx      | 1    | 2    | 2       | 2       | 1  | 2  | 1  | 1  |    |
|       | Emx      | 0    | 0    | 0       | 0       | 1  | 1  | 2  | 1  |    |
|       | Hhex     | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Hlx      | 0    | 0    | 0       | 0       | 0  | 1  | 3  | 2  |    |
|       | Lbx      | 0    | 0    | 0       | 0       | 0  | 0  | 1  | 1  |    |
|       | Mnx      | 1    | 1    | 1       | 1       | 0  | 0  | 1  | 1  |    |
|       | Msx      | 0    | 0    | 0       | 0       | 1  | 1  | 1  | 2  | 0  |
|       | Nedx     | 0    | 0    | 0       | 0       | 0  | 0  | 2  | 2  |    |
|       | Nk1      | 1    | 1    | 1       | 1       | 2  | 1  | 1  | 1  |    |
|       | Nk2.1/2.2/4| 2  | 1    | 1       | 1       | 1  | 1  | 3  | 8  | 4  |
|       | Nk3      | 0    | 0    | 0       | 0       | 0  | 1  | 1  | 1  | 1  |
|       | Nk5/Hmx  | 0    | 0    | 0       | 0       | 1  | 0  | 1  | 1  |    |
|       | Nk6      | 0    | 0    | 0       | 0       | 1  | 1  | 1  | 1  |    |
|       | Nk7      | 0    | 0    | 0       | 0       | 1  | 1  | 1  | 1  |    |
|       | Noto     | 1    | 1    | 1       | 0       | 1  | 1  | 6  | 2  |    |
|       | Ro       | 0    | 0    | 0       | 0       | 0  | 0  | 1  | 0  |    |
|       | Tlx      | 0    | 0    | 0       | 0       | 1  | 1  | 0  | 0  |    |
| PRD   | Alx      | 0    | 0    | 0       | 0       | 1  | 1  | 1  | 1  |    |
|       | Arx      | 1    | 2    | 1       | 0       | 0  | 0  | 1  | 1  |    |
|       | Dmbx    | 1    | 0    | 1       | 1       | 0  | 2  | 1  | 2  |    |
|       | Gsc      | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Hbn      | 1    | 1    | 1       | 1       | 2  | 1  | 1  | 1  |    |
|       | Otp      | 2    | 2    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Otx      | 3    | 3    | 2       | 3       | 4  | 7  | 3  | 4  |    |
|       | Pax3/7   | 0    | 0    | 0       | 0       | 0  | 0  | 2  | 2  |    |
|       | Pax4/6   | 2    | 1    | 1       | 2       | 2  | 1  | 2  | 2  |    |
|       | Pitx     | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Prox     | 0    | 0    | 0       | 0       | 0  | 0  | 1  | 1  |    |
|       | Rx       | 0    | 0    | 0       | 0       | 0  | 1  | 1  | 1  |    |
|       | Repo     | 0    | 0    | 0       | 0       | 0  | 0  | 1  | 1  |    |
|       | Uncx    | 1    | 1    | 1       | 1       | 1  | 1  | 2  | 2  |    |
|       | Vsx      | 2    | 0    | 1       | 1       | 1  | 1  | 0  |    |    |
| LIM   | Isl      | 0    | 0    | 0       | 1       | 1  | 1  | 1  | 0  |    |
|       | Lhx1/5   | 1    | 0    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Lhx2/9   | 0    | 0    | 0       | 1       | 0  | 1  | 1  | 1  |    |
|       | Lhx6/8   | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Lmx      | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
| POU   | Hdx      | 0    | 0    | 0       | 1       | 0  | 0  | 0  | 0  |    |
|       | Pou1     | 0    | 0    | 0       | 0       | 1  | 1  | 1  | 1  |    |
|       | Pou3     | 0    | 0    | 0       | 0       | 1  | 1  | 3  | 2  |    |
|       | Pou4     | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
|       | Pou6     | 1    | 1    | 1       | 1       | 1  | 1  | 1  | 1  |    |
| SINE  | Six1/2   | 0    | 0    | 0       | 0       | 1  | 2  | 2  | 2  |    |
|       | Six3/6   | 1    | 1    | 1       | 1       | 1  | 2  | 1  | 1  |    |
|       | Six4/5   | 1    | 1    | 1       | 1       | 1  | 1  | 2  | 1  |    |

(continued)
domain-containing and 50 homeobox domain-containing genes (Homeodomain-containing genes of *H. viridissima* are discussed in the next section). A similar analysis of putative signaling molecule genes showed that the *H. viridissima* genome contains 16 fibroblast growth factor (FGF)-like domain genes, 11 transforming growth factor-beta (TGF-β) genes, and 10 Wnt genes (Table 5B). These numbers are comparable to those in *H. vulgaris*. In general, the number of transcription factor and signaling molecule family members appeared similar among cnidarians, although a few families, such as AT_Hook and hairy-orange of transcription factors (Table 5A) and Interleukin 3 (IL3) families of signaling molecules (Table 5B) were not found in *Hydra* genomes.

### Hox and Para-Hox genes in *Hydra viridissima*

Among transcription factors, homeodomain-containing proteins have been intensively investigated as key molecules in the developmental toolkit. They are highly diversified and participate in a wide variety of developmental processes in metazoans. In particular, those in Cnidarians that are shared by the common ancestors of deuterostomes and protostomes are important to understand body plan evolution of bilaterians (Ferrier and Holland 2001; Chourrout et al. 2006; Ferrier 2016; DuBuc et al. 2018). While many orthologous genes of known homeodomain-containing proteins, including Hox and ParaHox genes, have been identified in cnidarians, cnidarian-specific specializations, such as loss of some homeodomain protein genes and fragmentation of the Hox cluster have been reported (Kamm et al. 2006; Steele et al. 2011; Chapman et al. 2010; Leclère et al. 2019). To understand the evolutionary trajectory of homeobox protein genes in the *Hydra* lineage, we classified them into ANTP (HOXL and NKL), PRD, LIM, POU, PROS, SINE, TALE, CERS, or ZF using bi-directional BLAST searches against sequences of homeodomains in other animals, using HomoeDB (Zhong and Holland 2011) (Table 6, Table S5) and phylogenetic analysis for ANPT- and PRD-class genes (Figs. S2 and S3), referring to the Hox genes previously identified in other cnidarians (Schummer et al. 1992; Chourrout et al. 2006; Leclère et al. 2019; Khalturin et al. 2019).

In the *H. viridissima* genome, we identified 48 homeodomain-containing genes in the genome, 5 ANTP-HOXL, 9 ANTP-NKL, 21 PRD, 4 LIM, 4 TALE, 2 SINE, 2 POU, and 1 CERS; however, we failed to find CUT, HNF, PROS, and ZF classes. This tendency toward gene loss is shared by the two other hydrozoans, *H. vulgaris* and *Clytia hemisphaerica* (Table 6). Among cnidarians, anthozoan genomes (*Nematostella* and *Acropora*) apparently contain the most homeodomain-containing genes, while scyphozoans (*Aurelia*) and cubozoans (*Morbakka*) have intermediate numbers, and hydrozoan genomes contain the fewest. CUT class genes are not found in medusozoan genomes and all cnidarian genomes lack PROS and ZF class genes altogether. In addition, NKL genes are less abundant in *Hydra* and HOXL genes are less abundant in hydrozoans generally, than in other cnidarians. *H. viridissima* and *H. vulgaris* possess the same ANTP genes (Figure 4, Table 7, Table S5), suggesting a reason for the same body plan in these *Hydra* species, although the body size of *H. viridissima* is smaller. As previously reported (Leclère et al. 2019; Khalturin et al. 2019; Gauchat et al. 2000; Quiquand et al. 2009), ParaHox genes Gsh and Mox are present in *Hydra*, whereas Xlox and Cbx are missing, unlike other medusozoans (Table 7, Figure 4). On the other hand, Hox gene composition is quite similar among medusozoans. They have Hox1 and Hox9-14, but lack Hox2, Evx, Gbx, Mnx, unlike anthozoans. Medusozoans have lost many NKL genes, Ndx, Hlx, Mnx, Mxs, and Lbx compared to anthozoans. In addition, Dbx, Hlx, Nk3, Nk6 and Nk7 are not found in hydrozoans, nor are Nk5, Exm or Msx-1 in *Hydra* (Table 7, Figure 4). In addition, some degree of synteny conservation of HOXL genes and NKL genes is found in Anthozoa, but not in Medusozoa (Figure 4), suggesting complete fragmentation of the homeobox gene cluster in the common ancestor of medusozoans. *Nematostella* expresses Gbx, Hlx, Nk3 and Nk6 in the pharyngeal or mesenteric region (Gbx in pharyngeal endoderm (Matus et al. 2006), Hlx and Nk6 in pharyngeal ectoderm, and Nk3 in nutrient-storing somatic gonads in mesenteries (Steinmetz et al. 2017)). Anthozoans have a pharynx and a mesentery that structurally supports the pharynx and serves as the site of gamete production in the gastrointestinal cavity, while these tissues are not found in *Hydra*. Loss of these genes reflects the simplification of body structure in the *Hydra* lineage.

### CONCLUSIONS

In this study, we report the first genome assembly of *H. viridissima*, which is one of the most basal species in the genus *Hydra* and the only species with symbiotic algae. Compared to *H. vulgaris*, *H. viridissima* has a compact genome one-third the size and with 36.5% fewer genes (Table 1). In addition, the *H. viridissima* genome has fewer repetitive sequences. RNA transposons, in particular, are almost absent (Figure 2). On the other hand, the repertoire of transcription factor genes, including homeodomain-containing genes in *H. viridissima* is quite similar to that in *H. vulgaris* (Table 5), reflecting the common body plan in these species. Comparative analysis of homeodomain genes among cnidarians indicates gradual simplification of the ANTP gene repertoire in the *Hydra* lineage (Tables 6 and 7, Figure 4), which is likely to reflect the simple body structure of *Hydra* and the absence of jellyfish and planula stages. In addition, we found diverse innate immunity genes in the *H. viridissima* genome that are also observed in corals (Table 4, Figure 3), indicating a common feature involved in algal symbiosis. The *H. viridissima* genome presented here provides a

---

**Table 7, continued**

| Class | Subclass | Family | Hvir | Hvul v1 | Hvul v2 | Ch | Aa | Mv | Nv | Ad |
|-------|----------|--------|------|---------|---------|----|----|----|----|----|
| TALE  |          | Hvir   | 1    | 1       | 1       | 1  | 2  | 1  | 1  | 1  |
|       |          | Hvul v1| 1    | 1       | 1       | 1  | 5  | 1  | 1  | 1  |
|       |          | Hvul v2| 1    | 2       | 1       | 1  | 1  | 1  | 1  | 1  |
|       |          | Ch     | 1    | 1       | 1       | 1  | 1  | 1  | 1  | 1  |
|       |          | Aa     | 1    | 1       | 1       | 1  | 1  | 1  | 1  | 1  |
| CERS  |          | Cers   | 2    | 0       | 0       | 0  | 0  | 0  | 1  | 1  |
|       |          | Tgf    | 0    | 0       | 1       | 2  | 1  | 1  | 1  | 1  |
| CUT   |          | Onecut | 0    | 1       | 0       | 0  | 0  | 0  | 1  | 1  |
| HNF   |          | Hnf1   | 0    | 0       | 0       | 0  | 1  | 0  | 1  | 1  |
Hydra genome comparable in quality to those of other cnidarians, including medusozoans and anthozoans, which will hopefully facilitate further studies of cnidarian genes, genomes, and genetics to understand basal metazoan evolution and strategies to support algal symbiosis in cnidarians.

ACKNOWLEDGMENTS

This research was supported by funding from the Okinawa Institute of Science and Technology (OIST) to the Marine Genomics Unit (NS) and by the Okayama University Dispatch Project for Female Faculty members (MH). MH was supported by the Japanese Society for Promotion of Science Funding (15K07173, 18K06364). We thank Dr. Steven D. Aird for editing the manuscript, Ms. Kanako Hisata for valuable discussion, and Prof. Thomas C. G. Bosch (Kiel University) for valuable discussion. We also thank the DNA Sequencing Section and the IT Section of OIST for excellent technical support. Computation for this work was partially performed on the NIG supercomputer at the ROIS National Institute of Genetics.

LITERATURE CITED

Balasubramanian, P. G., A. Beckmann, U. Warnken, M. Schnöller, A. Schuler et al., 2012 Proteome of Hydra nematocyst. J. Biol. Chem. 287: 9672–9681. https://doi.org/10.1074/jbc.M111.328203

Baumgarten, S., O. Simakov, L. Y. Escherick, Y. J. Liew, E. M. Lehnert et al., 2015 The genome of Aiptasia a sea anemone model for coral symbiosis. Proc. Natl. Acad. Sci. USA 112: 11893–11898. https://doi.org/10.1073/pnas.1513318112

Bode, H., 2011 Axis Formation in Hydra. Annu. Rev. Genet. 45: 105–117. https://doi.org/10.1146/annurev-genet-102209-163540

Bode, H. R., 2003 Head regeneration in Hydra. Dev. Dyn. 226: 225–236. https://doi.org/10.1002/dvdy.10225

Boetzer, M., C. V. Henkel, H. J. Jansen, D. Butler, and W. Pirovano, 2011 Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27: 578–579. https://doi.org/10.1093/bioinformatics/btq683

Bridge, D., C. W. Cunningham, B. Schierwater, R. DeSalle, and L. W. Buss, 1992 Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. Proc. Natl. Acad. Sci. USA 89: 8750–8753. https://doi.org/10.1073/pnas.89.18.8750

Brusca, R. C., W. Moore, and S. M. Shuster, 2016 Invertebrates, Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts.

Cernichiarri, E., L. Muscatine, and D. C. Smith, 1969 Maltose Excretion by Boetzer, M., C. V. Henkel, H. J. Jansen, D. Butler, and W. Pirovano, 2011 Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27: 578–579. https://doi.org/10.1093/bioinformatics/btq683

Douglas, A. E., and V. A. R. Huss, 1986 On the characteristics and taxonomic position of symbiotic Chilorea. Arch. Microbiol. 145: 80–84. https://doi.org/10.1007/BF00413031

DuBuc, T. Q., T. B. Stephenson, A. Q. Rock, and M. Q. Martindale, 2018 Hox and Wnt pattern the primary body axis of an anhoozan cnidarian before gastrulation. Nat. Commun. 9. 2007. https://doi.org/10.1038/s41467-018-04184-x

Elieh Ali Komí, D., L. Sharma, and C. S. Dela Cruz, 2018 Chitin and Its Effects on Inflammatory and Immune Responses. Clin. Rev. Allergy Immunol. 54: 213–223. https://doi.org/10.1007/s11016-017-8600-0

Emms, D. M., and S. Kelly, 2015 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16: 157. https://doi.org/10.1186/s13059-015-0721-2

Ferrier, D. E. K., 2016 Evolution of Homeobox Gene Clusters in Animals: The Giga-Cluster and Primary vs. Secondary Clustering. Front. Ecol. Evol. 4: 36. https://doi.org/10.3389/fevo.2016.00036

Ferrier, D. E. K., and P. W. H. Holland, 2001 Ancient origin of the Hox gene cluster. Nat. Rev. Genet. 2: 33–38. https://doi.org/10.1038/35047605

Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry et al., 2016 The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44: D279–D285. https://doi.org/10.1093/nar/gkv1344

Gauchat, D., F. Mazet, C. Berney, M. Schummer, S. Kreger et al., 2000 Evolution of Antp-class genes and differential expression of <em>Hydra Hox</em> in <em>paraHox</em> genes in anterior patterning. Proc. Natl. Acad. Sci. USA 97: 4493–4498. https://doi.org/10.1073/pnas.97.9.4493

Gold, D. A., T. Katsuki, Y. Li, X. Yan, M. Regulski et al., 2019 The genome of the jellyfish Aurelia and the evolution of animal complexity. Nat. Ecol. Evol. 3: 96–104. https://doi.org/10.1038/s41559-018-0719-8

Haas, B. J., A. L. Delcher, S. M. Mount, J. R. Wortman, R. K. Smith, Jr. et al., 2003 Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 31: 5654–5666. https://doi.org/10.1093/nar/gkg770

Habetha, M., F. Anton-Erdleben, K. Neumann, and T. C. Bosch, 2003 The Hydra viridis/Chlorella symbiosis. Growth and sexual differentiation in polyps without symbionts. Zoology 106: 101–108. https://doi.org/10.1078/0944-2006-00104

Hall, T. A. 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series 41: 95–98.

Hamada, M., K. Schröder, J. Bathia, U. Kürn, S. Fraune et al., 2018 Metabolic co-dependence drives the evolutionarily ancient Hydra-Chlorella symbiosis. elife 7: e35122. https://doi.org/10.7554/elife.35122

Hamada, M., E. Shoguchi, C. Shinzato, T. Kawashima, D. J. Miller, et al., 2013 The complex NOD-like receptor repertoire of the coral Acropora digitifera includes novel domain combinations. Mol. Biol. Evol. 30: 167–176. https://doi.org/10.1093/molbev/msa213

Holstein, T. W., E. Hohmayer, and U. Technau, 2003 Cnidarians: An evolutionarily conserved model system for regeneration? Dev. Dyn. 226: 257–267. https://doi.org/10.1002/dvdy.10227

Huang, S., M. Kang, and A. Xu, 2017 Haplomerger2: rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. Bioinformatics 33: 2577–2579. https://doi.org/10.1093/bioinformatics/btx220

Huss, V. A. R., C. Holweg, B. Seidel, V. Reich, M. Rahat et al., 1993/ 1994 There is an ecological basis for host/symbiont specificity in Chlorella/Hydra symbioses. Endocytobiosis Cell Res. 10: 35.

Kham, K., B. Schierwater, W. Jakob, S. L. Delaporta, and D. J. Miller, 2006 Axial Patterning and Diversification in the Cnidaria Predate the Hox System. Curr. Biol. 16: 920–926. https://doi.org/10.1016/j.cub.2006.03.036

Khalturn, K., C. Shinzato, M. Khalturna, M. Hamada, M. Fujie et al., 2019 Medusozoan genomes inform the evolution of the jellyfish body plan. Nat. Ecol. Evol. 3: 811–822. https://doi.org/10.1038/s41559-019-0853-y

Kim, H.-M., J. A. Weber, N. Lee, S. G. Park, Y. S. Cho et al., 2019 The genome of the giant Nomura’s jellyfish sheds light on the early evolution of active predation. BMC Biol. 17: 28. https://doi.org/10.1186/s12915-019-0643-7

Kusserow, A., K. Pang, C. Sturm, M. Hrouda, J. Lentfer et al., 2005 Unexpected complexity of the Wnt gene family in a sea anemone. Nature 433: 156–160. https://doi.org/10.1038/nature03158

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan et al., 2007 Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948. https://doi.org/10.1093/bioinformatics/btm404
