Wilding, CS, Fletcher, N, Smith, EK, Prentis, P, Weedall, GD and Stewart, Z

The genome of the sea anemone Actinia equina (L.): Meiotic toolkit genes and the question of sexual reproduction.

The genome of the sea anemone *Actinia equina* (L.): meiotic toolkit genes and the question of sexual reproduction.

Craig S. Wilding[1], Nicola Fletcher[1], Eleanor K. Smith[1], Peter Prentis[2,3], Gareth D. Weedall[1] and Zac Stewart[2].

1. School of Biological and Environmental Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK.
2. School of Earth, Environmental and Biological Sciences, Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia
3. Institute for Future Environments, Queensland University of Technology, Brisbane, Australia

ABSTRACT

The beadlet anemone *Actinia equina* (L.) (Cnidaria: Anthozoa: Actiniaria: Actiniidae) is one of the most familiar organisms of the North European intertidal zone. Once considered a single, morphologically variable species across northern Europe, it is now recognised as one member of a variable species complex. Previous studies of distribution, aggression, allozymes and mitochondrial DNA suggest that the diversity in form and colour within *A. equina* may hide still unrecognised species diversity. To empower further study of *A. equina* population genetics and systematics, we sequenced (PacBio Sequel) the genome of a single *A. equina* individual to produce a high-quality genome assembly (contig N_{50} = 492,607bp, 1,485 contigs, number of protein coding genes = 47,671, 97% BUSCO completeness). There is debate as to whether *A. equina* reproduces solely asexually, since no reliable, consistent evidence of sexual reproduction has been found. To gain further insight, we examined the genome for evidence of a 'meiotic toolkit' – genes believed to be found consistently in sexually reproducing organisms – and demonstrate that the *A. equina* genome appears not to have this full complement. Additionally, Smudgeplot analysis, coupled with high haplotype diversity, indicates this genome assembly to be of ambiguous ploidy, suggesting that *A. equina* may not be diploid. The suggested polyploid nature of this species coupled with the deficiency in meiotic toolkit genes, indicates that further field and laboratory studies of this species is warranted to understand how this species reproduces and what role ploidy may play in speciation within this speciose genus.

KEYWORDS

Meiotic toolkit; ploidy; phylum Cnidaria; cryptic species; speciation

INTRODUCTION

Genomic resources open up fantastic opportunities in population, speciation and comparative genomics [1-3]. However, non-model organisms typically lack such resources. This is particularly true for members of the phylum Cnidaria which, with currently only 26 genomes

(https://www.ncbi.nlm.nih.gov/genome) available for over 11,000 species [4], the majority from coral species [5], is particularly underrepresented in genome databases. Of these, only three are from the 1,100+ species [6] of Actiniarian sea anemones: *Nematostella vectensis* [7], *Exaiptasia pallida* [8] and *Anemonia viridis* [9], with an additional genome from *Actinia tenebrosa* recently completed [10]. Although transcriptomes are available for some other anemone species e.g. [11, 12], these lack corresponding genome assemblies. Yet Cnidarian genomes are extremely variable in terms of size, base composition, transposable element content, and gene conservation [13], and much can be learnt of the developmental transcriptional machinery from them [14] and so invite further study.

A common northern European Cnidarian is the littoral anthozoan *Actinia equina*. Despite its familiarity, *A. equina* has a complex and unresolved taxonomic history [15-17]. What was once considered a single polymorphic species with a wide geographic range is now defined as a species complex [16], with *A. equina sensu lato* split into at least *A. equina* (L.), *A. prasina* [18], *A. fragacea* [19] (but see [20]), *A. nigropunctata* [21], *A. ebhayiensis* [22], *A. schmidti* and *A. sali* [23], mostly based upon allozyme electrophoresis. Some members of the genus remain poorly described and others likely still contain cryptic species diversity [16]. For example, *A. equina sensu stricto* exists as a number of differently coloured morphs with a, typically, red or red-brown column but pedal discs (the structure used to attach to rocks) that can be red, pink, orange, green or grey. Animals with red/pink discs differ from those with green/grey discs in a variety of ways, including intertidal distribution [24-26], adhesiveness [25], aggression [27, 28], and nematocyst [29] and acrorhagial [30] morphology. In addition, both allozyme [24, 25, 30] and mitochondrial DNA [31] studies suggest genetic differentiation among these morphs. It seems likely that further diversity awaits discovery within this 'species'.

Cnidaria exhibit a range of reproductive strategies from fully sexual to asexual – employing pedal laceration, fission, budding, parthenogenesis or somatic embryogenesis [32-35] – though a mixture of both sexual and asexual strategies is common. Asexual reproduction in some species is associated

with periods of environmental stress, small body size or poor nutrition [32] and such facultative asexual reproduction has been reported in numerous species e.g. *Actinia tenebrosa* [36], *Anthopleura elegantissima* [37], *Haliplanella luciae* [38] and *Sagartia elegans* [39]. Although *A. equina* has been widely studied in an ecological context, there remain doubts about whether it reproduces sexually, asexually or uses a combination of both strategies, although much of the literature fails to distinguish between facultative and obligate asexuality. Thus, whilst Schama *et al.* [40] concluded that 'the binomial *A. equina* was retained for the asexually reproducing British samples' and Spaulding [41] reports *A. equina* as an obligate brooder, various studies [42-44] have detected gonadal tissue in *A. equina* and indeed described the nature of the sperm of this species [45].

Many of the studies reporting evidence of sexual reproduction (presence of eggs/sperm) predate the recognition of cryptic species among *A. equina s.l.* Due to the morphological similarity, and overlapping distributions of some of these species it is unclear whether samples claimed to be sexually reproducing definitively concern *A. equina* or other species, unrecognised at the time of study, which may have different reproductive strategies. For example, two species previously regarded as varieties of *A. equina*: *A. fragacea* [46] and *A. cari* [16] are non-brooding. Population genetics does suggest sexuality in *A. equina*, since populations examined through allozyme electrophoresis are typically in Hardy-Weinberg equilibrium [16] and Chomsky *et al.* [47] detected unique AFLP profiles for all individuals examined from the Mediterranean coast of Israel, arguing against a clonal origin. However, the samples studied by Chomsky *et al.* [47] have now been recognised as *A. schmidti* [48], again reinforcing the problem in interpreting studies of this species which pre-date the application of genetic studies. It would also be expected that if *A. equina* reproduced sexually, planktonic planulae would be seen in plankton tows. However, no *Actinia* samples were found from metabarcoding Adriatic Sea plankton samples [49] although six other Actiniarians and 12 hydrozoans were successfully identified.

It is certainly the case that *A. equina* are frequently encountered with young brooding within the coelenteron [42, 43]. These juvenile offspring have been suggested to arise through parthenogenesis or perhaps somatic embryogenesis [46, 50] or internal budding [51], a strategy rare in other genera

[33]. It has also been argued that those brooded young may represent sexually reproduced individuals which have subsequently re-entered an adult anemone. Although *A. equina* are capable of holding allogeneic individuals if these are introduced artificially [52] evidence from colouration matching the brooding 'parent' [53], from allozyme data [42, 51] and from DNA evidence [35] confirms that juveniles within the coelenteron are clonal individuals. However, this does not necessarily mean that *A. equina* are obligately asexual. A variety of anemones including *A. schmidti* [48] and *A. tenebrosa* [54] employ both sexual and asexual phases dependent upon the ecological context. *Actinia* may employ a mixed reproductive strategy with sporadic sexual recruitment [55], although there remains a lack of definitive, conclusive laboratory or field evidence for the existence of sexually reproduced larvae.

Genomics may shed light on this issue. Though there appear to be no consistent genomic signatures of asexuality across diverse taxa [56], species which undergo sexual reproduction require a series of genes involved in meiotic recombination and DNA repair, and the presence and expression of these within sequenced genomes may imply that a species undergoes sexual reproduction [57-59]. This set of genes has been termed 'the meiotic toolkit' and used previously to study the modes of reproduction in a variety of groups including arthropods [60], diatoms [61] and protists [62]. The absence of the full meiotic toolkit complement may suggest obligate asexuality in *A. equina*.

Here, we sequenced a single individual of *A. equina* and, from the resultant genome, annotated the genes of the meiotic toolkit, estimated ploidy, and designed PCR primers to amplify a polymorphic toxin locus, using these to provide additional evidence that what is currently considered as a single species (*A. equina*) is composed of more than a single genetic entity. This genomic resource promises much for the future detailed unravelling of this species.

MATERIALS AND METHODS

DNA extraction and sequencing

We have previously described the genome sequencing for this species [31]. Briefly, genomic DNA was extracted from a single individual *A. equina* with a red column and red pedal disk collected from Rhosneigr, Wales, UK, following grinding in liquid nitrogen in 20 ml 80mM EDTA (pH 8.0), 100mM Tris-HCl (pH 8.0), 0.5% SDS, 100 µg/ml proteinase K, and 40 µl RNaseA (100 mg/ml) and incubated at 60°C for 3 h. Salt-chloroform extraction [63] of DNA was undertaken, then DNA precipitated with 0.6 volumes of isopropanol, and dissolved in water, following which additional purification was undertaken using a Qiagen Genomic Tip 20/G and precipitated a second time with 0.6 volumes of isopropanol. 20 kb-insert PacBio sequencing libraries were sequenced on five SMRT cells on a Pacific Biosciences Sequel (Pacific Biosciences, Menlo Park, CA, USA) at the Centre for Genomic Research, University of Liverpool. Full MIxS details for this project are provided in Table 1.

Genome assembly

Assembly of subreads was undertaken using three separate assembly methods: CANU v1.7 [64], SMARTdenovo (https://github.com/ruanjue/smartdenovo) or WTDBG [65]. For diploid species with high levels of polymorphism, alternative haplotypes at the same genomic region may be assembled into multiple separate sequence contigs that appear to be different genomic regions in the final assembly, erroneously inflating the haploid genome size. To reduce the impact of this on assembly, for the SMARTdenovo assembly, the Purge Haplotigs procedure [66] was applied. For CANU v1.7 [64] all subreads were assembled in three steps: read correction, trimming and assembly. Read correction and trimming were run with default parameters for PacBio data and an estimated genome size of 503 Mb, based on [67]. Assembly of trimmed reads was run for a range of predicted error rates ('error' here also includes heterozygosity). The default value for corrected PacBio reads is 0.045, which was run in addition to 0.035, 0.055, 0.065, 0.075, 0.085, 0.095 and 0.105. Assembly quality was assessed for each assembly using N statistics and BUSCO analysis [68, 69], with the 978 gene 'Metazoa' reference set.

Meiotic toolkit gene model annotation

Gene models were produced using a transcriptome assembled from short-read Illumina transcript sequences of *Actinia equina* (accessions SRX4378330 and SRX4378325 [70]) using a combination of different assemblers: Trinity *de novo*/genome-guided [71, 72], Velvet-oases [73], SOAPdenovo-trans [74], and Scallop [75]) combined with EvidentialGene https://sourceforge.net/projects/evidentialgene/. Gene model annotation involved using EVidenceModeler to combine *Program to Assemble Spliced Alignments* (PASA) outputs [76] and BRAKER [77] with 2x PASA updates followed by downstream processing using a custom script (processing_pipeline.sh; https://github.com/zkstewart/Genome_analysis_scripts) in which gmap_gene_find.py was used to annotate extra genes missed by PASA+BRAKER, followed by automatic removal of transposons and rRNA models falsely annotated as coding genes. Gene models annotated on contigs that Purge Haplotigs identified as 'additional' haplotypes were then removed to avoid 'double-counting' of alleles as paralogues. The final set of gene models was used to produce a set of predicted transcripts that were analysed for genome 'completeness' using BUSCO [68, 69] (with the option '-m trans') to identify orthologues of the 978-gene 'Metazoa' single-copy orthologues reference gene set.

There is some variation in which genes are considered to be part of the meiotic toolkit. Patil *et al.* [61] lists 37, Malik *et al.* [62] 29, Schurko and Logsdon Jr [57] 12 and Hofstatter and Lahr [58] 14. Here, we combined the genes in these four studies to search for a total of 46 genes (Table 2).

Meiotic toolkit genes were verified in the *A. equina* genome through standalone tBLASTn searching of transcripts using sequences from the anthozoans *Nematostella vectensis*, *Exaiptasia pallida* or *Pocillipora damicornis* (Supp. Table 1) or other invertebrates where no orthologue in these Cnidarian species could be found. Meiotic toolkit transcripts were subsequently used in BLASTn searches of the final genome assembly with intron/exon structure manually annotated. Where no transcript was found, tBLASTn searches of the genome assembly was undertaken. Intron/exon structures of genes

were determined manually following transcript alignments to genomic contigs and, where necessary, extended using tBLASTn searches.

Variation in the Acrorhagin-1 *gene*

Anemones with red columns and red, orange, or green pedal discs were collected from 10 locations in England (New Brighton), Scotland (Millport), Wales (Abraham's Bosom – Holyhead, Llandudno, Marloes, Penbryn and Rhosneigr), Ireland (Portmarnock) and the Isle of Man (Peel and Niarbyl). Location and sample details are provided in Supp. Table 2. DNA was extracted from clips of tentacles or pedal disc using the GeneJet Genomic DNA purification kit following the manufacturer's instructions. Primers AcroF1 (5'-TTTGCGAGAAGTTGGATTTCC-3') and AcroR1 (5'-GCAGCGTCCTTTGAACATCA-3') to amplify *Acrorhagin-1* [78] (Genbank accession number AB212066) which is intron-less in this genome assembly and therefore amplifiable from genomic DNA without the issue of length variable introns disrupting sequencing quality, were designed using Primer3 [79]. PCR reactions were conducted for 35 cycles of 95°C for 30 seconds; 55°C for 30 seconds; and 72°C for 1 minute with successful PCR products cleaned using a GeneJet PCR purification kit and sequenced using both AcroF1 and AcroR1 by GATC Biotech (Constanz, Germany). Sequences were aligned and manually edited in CodonCode Aligner (CodonCode Corporation). Where indels resulted in heterozygous sequences containing strings of double peaks, haplotypes were separated using Poly Peak Parser [80].

Ploidy estimation

Smudgeplot v0.2.1 [81] was used to estimate ploidy levels from corrected reads generated using MECAT (a modified Canu) [82] using default settings excepting that a k-mer value of 31 was used instead of 21.

RESULTS

Genome Assembly Statistics

PacBio sequencing produced 3,507,426 'polymerase' reads (single reads that can cover the same insert multiple times) that were split into a total of 4,936,001 subreads (full or partial passes of the same insert). Of these subreads, 487,629 were longer than 20 kb and 1,409,598 longer than 10 kb. Raw FASTQ data have been submitted to the Sequence Read Archive (SRA) with accession number SRR7651651.

Assembly statistics from the three separate assembly methods utilised (Canu, SMARTdenovo and WTDBG) are shown in Table 3. Because of the high number of contigs in the WTDBG assembly, Purge Haplotigs was not run and this assembly was not further considered. For the Canu assembly, the effect of different error rates is shown in Supplementary Fig. S1 with the default error rate (0.045) accepted since this produced the optimum balance of genome size and number of contigs. The 'best' assembly, as adjudged by higher N_{50} and lower number of contigs is the SMARTdenovo assembly following Purge Haplotigs. This genome assembly has been deposited at DDBJ/ENA/GenBank under the accession WHPX00000000. The version described in this paper is version WHPX01000000 and has been used in all further BLAST analyses. It is also available on the ReefGenomics web resource [83].

The best assembly (SMARTdenovo + Purge Haplotigs) was taken forward for detailed annotation. The predicted proteome of *A. equina*, based upon genome annotation of this assembly, contains 47,671 proteins (55,607 including alternative isoforms). BUSCO analysis indicated a high level of completeness of these gene models. Of the 978 reference genes in the 'Metazoa' gene set, 949 (97%) were found with 618 (63.2%) complete single-copy orthologues, 331 (33.8%) complete but duplicated, 19 (1.9%) fragmented and 10 (1.1%) missing. We note that, despite the Purge Haplotigs contigs removal step, a large proportion of these apparently single-copy genes were duplicated, suggesting that either the process may incompletely remove such allelic contigs or that the genome is unusually repetitive and the gene duplication represents a real phenomenon. Regarding the missing genes, we also note that a higher BUSCO score (98.2% completeness) was found before the Purge Haplotigs

contig removal step, thus it may be possible that manual inspection of the contigs Purge Haplotigs flagged for removal could marginally elevate the BUSCO completeness score in this assembly.

Meiotic toolkit

We searched the transcriptome and genome for 46 'meiotic toolkit' genes [57, 58, 61, 62] and found evidence for 41 of them (Table 2) with 27 completely annotated at the transcript and genomic level, four with partial transcript but complete genome annotations, and the remaining 10 being partial annotations. Of the core 13 meiotic proteins discussed in [57, 58, 61, 62], only eight were found, with *Hop1*, *Mer3*, *Msh4*, *Rec8* and *Zip4* all missing (Table 2). We note that for *Scc3* the best match is *Cohesin subunit SA-1*, and *Mlh2* was not found, with the closest match being *Pms2* (partial transcript and partial genome annotation). Some genes in the genome assembly (but not transcript sequence) had a single indel relative to the aligned transcript which would have shifted the reading frame. Since PacBio assemblies are known to be at risk of this [84], we adjusted the gene annotation based on the transcript sequence to produce full ORFs.

All meiotic toolkit gene annotations have been submitted to Genbank (Accession numbers MN307071-MN307111).

Polyploidy

Smudgeplot analysis using default settings indicated *A. equina* to have ambiguous ploidy status (Figure 1), with a diploid (AB) confidence of only 0.39, triploid (AAB) of 0.28, and tetraploid (AAAB) of 0.28. Additionally, BUSCO analysis indicated a high percentage of duplicated genes (33.8% duplicated genes - see above).

Toxin gene haplotypes

Primers AcroF1 and AcroR1 amplified only a single band from all anemones. There was discrete difference in size of amplified products from anemones with a red/orange pedal disc (389bp) versus

241 those with a green pedal disk (547-550 bp). Alignments of the full-length sequences seen in the
242 population are shown in Fig 2 and the haplotype network from the coding sequence in Supplementary
243 Fig. S2. From 57 individual anemones we identified seven haplotypes (Accession numbers MN605634-
244 MN605640) and eight separate genotypes. Regardless of collection locale, anemones with a green
245 pedal disc differed in length and sequence from those with a red/orange pedal disc; haplotypes 1-4
246 were seen only in anemones with a green pedal disc, whilst haplotypes 5-7 were seen in only
247 anemones with a red pedal disc (Table 4). Haplotype 6 was identical in sequence to the sequence of
248 *Acrorhagin-1* from [78]. Substantial variation was present in the 5'-UTR which displayed significant
249 length and sequence variation between haplotypes but variation was also present in the coding
250 sequence with both non-synonymous and frame-shift variation (a single base (A) frame-shift deletion
251 in the coding sequence of haplotype 1 results in truncation of the coding sequence).
252
253 **DISCUSSION**
254 Here, we provide a high-quality genome resource for the well-studied Cnidarian *A. equina* with an N_{50}
255 of 492,607bp and a BUSCO completeness of 97%. The estimated genome size of 409.0 MBp is
256 somewhat smaller than the 503MBp predicted by flow cytometric (FC) analysis [67] although since the
257 specimen of *A. equina* used for FC was collected in Japanese waters, it is not clear whether it truly was
258 *A. equina* (regarded as a North European species) or an unrecognised congeneric member of the
259 species complex and therefore not fully representative of *A. equina sensu stricto* (though we note that
260 the anemone sample of Honma *et al*. [78] was collected in Japanese waters and has an identical
261 *Acrorhagin-1* haplotype to those found in UK waters). The genome size is larger than that estimated
262 for the draft genome of its congener *A. tenebrosa* at 255Mbp, though this draft genome may not have
263 fully captured all repeat regions [10]. The genome of *A. equina* is available through ReefGenomics.org
264 [83], making it accessible to the research community. We utilise this to show that *A. equina* is missing
265 some of the genes regarded as critical for meiosis, providing evidence that *A. equina* may indeed be
266 asexual, as suggested by Schama *et al.* [40].

We were able to annotate only 40 of 46 genes suggested by various authors [57, 58, 61, 62] to form part of the meiotic toolkit. No matches in either the genome or transcriptome were found for *Hop1*, *Mer3*, *Msh4*, *Rec8* or *Zip4* whilst BLAST searches with *Mlh2* identified a partial match to *Pms2* only, yet the eukaryotic orthologue of *Mlh2* is *Pms1* not *Pms2* [85] and a partial annotation of *Pms1* was separately successfully completed. In stark contrast to *Actinia*, both *Exaiptasia pallida* and *Pocillipora damicornis* exhibit a full complement of these meiotic toolkit genes with the exception of *Mlh2* which appears absent in all Cnidaria (Supp. Table 1). In addition to *Mlh2*, *Rec8* could not be found within either the *Nematostella* genome [7] or transcriptome [86]. Whilst *Nematostella* appears not to display the full meiotic toolkit complement seen in *Exaiptasia* and *Pocillopora*, it does exhibit a more complete complement than *Actinia* despite the BUSCO completeness for *Nematostella* being lower than that of *Actinia* (93.8% vs. 97% for gene model statistics). Thus, *Rec8* may simply be missing in *Nematostella* due to a less complete assembly.

The apparent absence of six MT loci in the assembled *Actinia* genome could also result from incomplete assembly, whilst their absence from the transcriptomic dataset may be due to sampling of tissue/timepoints/developmental stages in which those genes are not expressed. However, the absence of these same six genes from two independent datasets – the genome reported here and the transcriptome built from the data of Waldron *et al.* [70] is suggestive that these genes are indeed absent from *A. equina*. Three other gene annotations contained apparent single indels in the genomic annotation that were not present in the corresponding transcripts, which could indicate the presence of pseudogenes. However, as indels are a common error in PacBio data [84, 87], we have corrected the gene model contingent upon the transcript data. For some genes we were able only to generate a partial annotation of both gene model and transcript. Identification of full-length orthologues may be hampered where genes are highly variable, which is particularly true in some gene regions of *BRCA2* where *N. vectensis* does have substantial differences from the human orthologue [88]. Thus, BLAST searches may not have identified the complete transcript leading to difficulty in annotating the full-length gene. Nevertheless, the fact that the 40 genes (41 with *Pms2*) fully or partially annotated are

functionally expressed (as evidenced from transcript data) suggests complete annotation should be possible for all identified genes if deeper transcript sequencing, or full-length isoform sequencing [89] is undertaken.

We identified 97% BUSCO completeness in the gene models from this genome, indicative of a high-quality genome. However, whilst largely complete, a significant proportion of genes (33.8%) are duplicated. This is much higher than that seen in other Cnidarian genomes [90]. This high level of duplicated loci could result from uncollapsed haplotypes in the data (despite the fact that Purge Haplotigs was run to remove these) or from complete or partial genome duplication. Contrary to what is expected in diploid species, Smudgeplot analysis of *A. equina* did not provide a clear indication of diploidy, with alternative ploidy statuses of triploidy or tetraploidy also being likely; this observation alongside the gene duplication suggested by BUSCO results indicates that this species may not be diploid. This ambiguous ploidy determination is in contrast to other Actiniarians which are confidently identified as diploid (Stewart and Prentis, unpublished). Questions have been raised previously about ploidy levels in the genus *Actinia*. Perrin *et al.* [16], reviewing the data from allozyme electrophoresis studies queried 'whether existence of multiple loci has resulted from duplication of restricted portions of the genome or from polyploidy'. Polyploidy is not unknown in phylum Cnidaria. Karyotype analysis indicates triploidy and tetraploidy occurs in the coral genus *Acropora* [91] and Shaw *et al.* [92] showed closely related *Sagartia* species have differing ploidy levels – *Sagartia troglodytes var. decorata* being diploid and *Sagartia troglodytes var. ornata* being tetraploid. Ploidy can be important in speciation [93, 94] and may be linked to asexual lineages [94]. Indeed, in 'asexual complexes', sexual species coexist and reproduce with asexual biotypes that have arisen from sexual ancestors producing derived biotypes of differing ploidy [94]. Thus, in any study of such a speciose genus as *Actinia* which has evidence of a mixture of sexual and asexual lineages it is important to consider further investigation of sample ploidy, and it remains to be seen whether *A. equina* exists as forms with >1 ploidy level. Ploidy estimation through karyotyping can be difficult in Cnidarians [95], especially where gametes are not easily accessible. However, since ploidy can be estimated from microsatellite [96], genotyping-

by-sequencing [97], and whole genome data [98], this genome will empower development of tools for this.

Previous genetic research to look at population genetics or phylogenetics of this species and its close relatives has been limited by the low number of loci for consideration. The mtDNA, previously used to look at differentiation within *A. equina* [31] is a single locus and evolves slowly in the Anthozoa [99], and while multi-locus allozyme studies [18, 23-25, 30, 100] and rDNA sequencing [20, 35] have been undertaken, they are both limited in scope since they are restricted to a small number of loci [101]. This genome effectively removes these limits. We have utilised this resource to study genetic differentiation at a toxin locus (*Acrorhagin-1*) demonstrating additional nuclear DNA evidence (in concert with the mtDNA data of [31]) that there is consistent differentiation between pedal disc colour morphs collected from across the Western UK, with red-pedal disk and green-pedal disk morphs displaying highly divergent *Acrorhagin-1* haplotype lengths and sequences. The indel present in haplotype 1 from green-based anemones changes the reading frame substantially and it will be important to study whether this remains functional. It thus seems likely that what is currently recognised as *A. equina* is indeed >1 species and that further work using multi-locus data is needed to further investigate this. A replicated, structured ecological sampling coupled with genomic scale variant screening is ultimately necessary to quantify the variation between these two morphs. This genome provides the tools to undertake this, empowering understanding of the number of species in this common, familiar, but perhaps underappreciated genus.

ACKNOWLEDGEMENTS

This work was supported by an award from the Liverpool John Moores University – Technology Directorate Voucher Scheme providing access to University of Liverpool's Shared Research Facilities. We wish to thank Ríoghnach O'Neill and Síofra, Ruairí and Ronan Palliser for assistance with the anemone sampling at Portmarnock, Ireland.

REFERENCES

[1] C.R. Campbell, J.W. Poelstra, A.D. Yoder, What is speciation genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species, Biol. J. Linn. Soc., 124 (2018) 561-583.

[2] N. Galtier, Delineating species in the speciation continuum: a proposal, Evol. Appl., 12 (2019) 657-663.

[3] N.J. Nadeau, T. Kawakami, Population genomics of speciation and admixture, in: O.P. Rajora (Ed.) Population Genomics, Springer, 2018, pp. 1-41.

[4] M. Daly, M.R. Brugler, P. Cartwright, A.G. Collins, M.N. Dawson, D.G. Fautin, S.C. France, C.S. Mcfadden, D.M. Opresko, E. Rodriguez, S.L. Romano, J.L. Stake, The phylum Cnidaria: A review of phylogenetic patterns and diversity 300 years after Linnaeus, Zootaxa, 1668 (2007) 127-182.

[5] D. Bhattacharya, S. Agrawal, M. Aranda, S. Baumgarten, M. Belcaid, J.L. Drake, D. Erwin, S. Foret, R.D. Gates, D.F. Gruber, B. Kamel, M.P. Lesser, O. Levy, Y.J. Liew, M. MacManes, T. Mass, M. Medina, S. Mehr, E. Meyer, D.C. Price, H.M. Putnam, H. Qiu, C. Shinzato, E. Shoguchi, A.J. Stokes, S. Tambutté, D. Tchernov, C.R. Voolstra, N. Wagner, C.W. Walker, A.P.M. Weber, V. Weis, E. Zelzion, D. Zoccola, P.G. Falkowski, Comparative genomics explains the evolutionary success of reef-forming corals, eLife, 5 (2016) e13288.

[6] D.G. Fautin, T. Zelenchuk, D. Raveendran, Genera of orders Actiniaria and Corallimorpharia (Cnidaria, Anthozoa, Hexacorallia), and their type species, Zootaxa, 1668 (2007) 62.

[7] N.H. Putnam, M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V.V. Kapitonov, J. Jurka, G. Genikhovich, I.V. Grigoriev, S.M. Lucas, R.E. Steele, J.R. Finnerty, U. Technau, M.Q. Martindale, D.S. Rokhsar, Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization, Science, 317 (2007) 86-94.

[8] S. Baumgarten, O. Simakov, L.Y. Esherick, Y.J. Liew, E.M. Lehnert, C.T. Michell, Y. Li, E.A. Hambleton, A. Guse, M.E. Oates, J. Gough, V.M. Weis, M. Aranda, J.R. Pringle, C.R. Voolstra, The genome of *Aiptasia*, a sea anemone model for coral symbiosis, Proc. Natl. Acad. Sci. USA, 112 (2015) 11893-11898.

371 [9] I. Urbarova, H. Patel, S. Forêt, B.O. Karlsen, T.E. Jørgensen, J.M. Hall-Spencer, S.D. Johansen,
372 Elucidating the small regulatory RNA repertoire of the sea anemone *Anemonia viridis* based on whole
373 genome and small RNA sequencing, Genome Biol. Evol., 10 (2018) 410-426.
374 [10] J.M. Surm, Z.K. Stewart, A. Papanicolaou, A. Pavasovic, P.J. Prentis, The draft genome of *Actinia*
375 *tenebrosa* reveals insights into toxin evolution, Ecol. Evol., 9 (2019) 11314-11328.
376 [11] S.A. Kitchen, C.M. Crowder, A.Z. Poole, V.M. Weis, E. Meyer, *De novo* assembly and
377 characterization of four Anthozoan (phylum Cnidaria) transcriptomes, G3: Genes|Genomes|Genetics,
378 5 (2015) 2441-2452.
379 [12] J.M. Surm, H.L. Smith, B. Madio, E.A.B. Undheim, G.F. King, B.R. Hamilton, C.A. van der Burg, A.
380 Pavasovic, P.J. Prentis, A process of convergent amplification and tissue-specific expression dominates
381 the evolution of toxin and toxin-like genes in sea anemones, Mol. Ecol., 28 (2019) 2272-2289.
382 [13] R.E. Steele, C.N. David, U. Technau, A genomic view of 500 million years of cnidarian evolution,
383 Trends Genet., 27 (2011) 7-13.
384 [14] U. Technau, M. Schwaiger, Recent advances in genomics and transcriptomics of Cnidarians, Mar.
385 Genomics, 24 (2015) 131-138.
386 [15] R.L. Manuel, British Anthozoa, E.J. Brill, Leiden, 1988.
387 [16] M.C. Perrin, J.P. Thorpe, A.M. Solé-Cava, Population structuring, gene dispersal and reproduction
388 in the *Actinia equina* species group, Oceanog. Mar. Biol., 37 (1999) 129-152.
389 [17] T.A. Stephenson, The British sea anemones, The Ray Society, London, 1935.
390 [18] A.M. Sole-Cava, J.P. Thorpe, Further genetic evidence for the reproductive isolation of green sea
391 anemone *Actinia prasina* Gosse from common intertidal beadlet anemone *Actinia equina* (L.), Mar.
392 Ecol. Prog. Ser., 38 (1987) 225-229.
393 [19] M.A. Carter, J.P. Thorpe, Reproductive, genetic and ecological evidence that *Actinia equina* var.
394 *mesembryanthemum* and var. *fragacea* are not conspecific, J. Mar. Biol. Assoc. U. K., 61 (1981) 79-93.

[20] A.M. Pereira, C. Brito, J. Sanches, C. Sousa-Santos, J.I. Robalo, Absence of consistent genetic differentiation among several morphs of *Actinia* (Actiniaria: Actiniidae) occurring in the Portuguese coast, Zootaxa, 3893 (2014) 595-600.

[21] J.C. den Hartog, O. Ocaña, A new endemic *Actinia* species (Actiniaria: Actiniidae) from the central Macaronesian Archipelagos, Zool. Meded. (Leiden), 77 (2003) 1-14.

[22] R. Schama, M. Mitchell, A.M. Solé-Cava, *Actinia ebhayiensis* sp. nov., a new species of sea anemone (Anthozoa: Actiniaria: Actiniidae) from South Africa, J. Mar. Biol. Assoc. U. K., 92 (2012) 885-894.

[23] F.A. Monteiro, A.M. Solé-Cava, J.P. Thorpe, Extensive genetic divergence between populations of the common intertidal sea anemone *Actinia equina* from Britain, the Mediterranean and the Cape Verde Islands, Mar. Biol., 129 (1997) 425-433.

[24] D.L.J. Quicke, R.C. Brace, Evidence for the existence of a third, ecologically distinct morph of the anemone, *Actinia equina*, J. Mar. Biol. Assoc. U. K., 64 (1984) 531-534.

[25] D.L.J. Quicke, A.M. Donoghue, R.C. Brace, Biochemical-genetic and ecological evidence that red/brown individuals of the anemone *Actinia equina* comprise two morphs in Britain, Mar. Biol., 77 (1983) 29-37.

[26] D.L.J. Quicke, A.M. Donoghue, T.F. Keeling, R.C. Brace, Littoral distributions and evidence for differential post-settlement selection of the morphs of *Actinia equina*, J. Mar. Biol. Assoc. U. K., 65 (1985) 1-20.

[27] R.C. Brace, H.A. Reynolds, Relative intraspecific aggressiveness of pedal disc colour phenotypes of the beadlet anemone, *Actinia equina*, J. Mar. Biol. Assoc. U. K., 69 (1989) 273-278.

[28] J.R. Collins, E.L. Vernon, J.S. Thomson, Variation in risk-taking and aggression in morphotypes of the beadlet anemone, *Actinia equina* (L.), and the green anemone, *Actinia prasina* (Gosse), J. Exp. Mar. Biol. Ecol., 496 (2017) 29-36.

[29] A.L. Allcock, P.C. Watts, J.P. Thorpe, Divergence of nematocysts in two colour morphs of the intertidal beadlet anemone *Actinia equina*, J. Mar. Biol. Assoc. U. K., 78 (1998) 821-828.

[30] A.M. Donoghue, D.L.J. Quicke, R.C. Brace, Biochemical-genetic and acrorhagial characteristics of pedal disc colour phenotypes of *Actinia equina*, J. Mar. Biol. Assoc. U. K., 65 (1985) 21-33.

[31] C.S. Wilding, G.D. Weedall, Morphotypes of the common beadlet anemone *Actinia equina* (L.) are genetically distinct, J. Exp. Mar. Biol. Ecol., 510 (2019) 81-85.

[32] F.-S. Chia, Sea anemone reproduction: patterns and adaptive radiations, in: G.O. Mackie (Ed.) Coelenterate Ecology and Behavior, Springer, New York, 1976, pp. 261-270.

[33] E.S. Bocharova, I.A. Kozevich, Modes of reproduction in sea anemones (Cnidaria, Anthozoa), Biol. Bull., 38 (2011) 849-860.

[34] A.M. Reitzel, D. Stefanik, J.R. Finnerty, Asexual reproduction in Cnidaria: comparative developmental processes and candidate mechanisms, in: T. Flatt, A. Heyland (Eds.) Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs, Oxford Scholarship Online, 2011, pp. 101-113.

[35] A.M. Pereira, E. Cadeireiro, J.I. Robalo, Asexual origin of brooding in the sea anemones *Actinia equina* and *A. schmidti*: molecular evidence from the Portuguese shore, N. Z. J. Mar. Freshw. Res., 51 (2016) 316-320.

[36] D.J. Ayre, The effects of asexual reproduction and inter-genotypic aggression on the genotypic structure of populations of the sea anemone *Actinia tenebrosa*, Oecologia, 57 (1983) 158-165.

[37] C.S. McFadden, R.K. Grosberg, B.B. Cameron, D.P. Karlton, D. Secord, Genetic relationships within and between clonal and solitary forms of the sea anemone *Anthopleura elegantissima* revisited: evidence for the existence of two species, Mar. Biol., 128 (1997) 127-139.

[38] J.M. Shick, A.N. Lamb, Asexual reproduction and genetics population structure in the colonizing sea anemone *Haliplanella luciae*, Biol. Bull., 153 (1977) 604-617.

[39] P.W. Shaw, Effects of asexual reproduction on population structure of *Sagartia elegans* (Anthozoa: Actiniaria), Hydrobiologia, 216 (1991) 519-525.

[40] R. Schama, A.M. Solé-Cava, J.P. Thorpe, Genetic divergence between east and west Atlantic populations of *Actinia* spp. sea anemones (Cnidaria: Actiniidae), Marine Biology, 146 (2005) 435-443.

[41] J.G. Spaulding, Embryonic and larval development in sea anemones (Anthozoa: Actiniaria), Integ. Compar. Biol., 14 (2015) 511-520.

[42] M.A. Carter, C.H. Thorp, The reproduction of *Actinia equina* L. var. *mesembryanthemum*, J. Mar. Biol. Assoc. U. K., 59 (1979) 989-1001.

[43] F.-S. Chia, M.A. Rostron, Some aspects of the reproductive biology of *Actinia equina* [Cnidaria: Anthozoa], J. Mar. Biol. Assoc. U. K., 50 (1970) 253-264.

[44] M.A. Rostron, J. Rostron, Fecundity and reproductive ecology of a natural population of *Actinia equina* L. (Cnidaria: Anthozoa), J. Exp. Mar. Biol. Ecol., 33 (1978) 251-259.

[45] A.U. Larkman, M.A. Carter, The spermatozoon of *Actinia equina* L. var. mesembryanthemum, J. Mar. Biol. Assoc. U. K., 60 (1980) 193-204.

[46] S.E. Gashout, R.F.G. Ormond, Evidence for parthenogenetic reproduction in the sea anemone *Actinia equina* L., J. Mar. Biol. Assoc. U. K., 59 (1979) 975-987.

[47] O. Chomsky, J. Douek, N.E. Chadwick, Z. Dubinsky, B. Rinkevich, Biological and population-genetic aspects of the sea anemone *Actinia equina* (Cnidaria: Anthozoa) along the Mediterranean coast of Israel, J. Exp. Mar. Biol. Ecol., 375 (2009) 16-20.

[48] E.A. Sales, L.F. Afonso, A.M. Pereira, Viviparous reproductive seasonality in *Actinia schmidti* (Cnidaria: Anthozoa) on the Portuguese west coast, Cah. Biol. Mar., 59 (2018) 143-147.

[49] S. Stefanni, D. Stanković, D. Borme, A. de Olazabal, T. Juretić, A. Pallavicini, V. Tirelli, Multi-marker metabarcoding approach to study mesozooplankton at basin scale, Sci. Rep., 8 (2018) 12085.

[50] D.G. Polteva, Regeneration and somatic embryogenesis of *Actinia equina* in different stages of ontogenic development, Acta Biol. Hung., 14 (1963) 199-208.

[51] J. Orr, J.P. Thorpe, M.A. Carter, Biochemical genetic confirmation of the asexual reproduction of brooded offspring in the sea anemone *Actinia equina*, Mar. Ecol. Prog. Ser., 7 (1982) 227-229.

[52] R. Lubbock, C. Allbut, The sea anemone *Actinia equina* tolerates allogeneic juveniles but alters their phenotype, Nature, 293 (1981) 474-475.

[53] A.J. Cain, Breeding system of a sessile animal, Nature, 247 (1974) 289-290.

[54] D.J. Ayre, The effects of sexual and asexual reproduction on geographic variation in the sea anemone *Actinia tenebrosa*, Oecologia, 62 (1984) 222-229.

[55] M.A. Carter, J. Miles, Gametogenic cycles and reproduction in the beadlet sea anemone *Actinia equina* (Cnidaria: Anthozoa), Biol. J. Linn. Soc., 36 (1989) 129-155.

[56] K.S. Jaron, J. Bast, T.R. Ranallo-Benavidez, M. Robinson-Rechavi, T. Schwander, Genomic features of asexual animals, bioRxiv, (2018) 497495.

[57] A.M. Schurko, J.M. Logsdon Jr, Using a meiosis detection toolkit to investigate ancient asexual "scandals" and the evolution of sex, Bioessays, 30 (2008) 579-589.

[58] P.G. Hofstatter, D.J.G. Lahr, All eukaryotes are sexual, unless proven otherwise, Bioessays, 41 (2019) 1800246.

[59] A.M. Schurko, M. Neiman, J.M. Logsdon, Jr., Signs of sex: what we know and how we know it, Trends Ecol. Evol., 24 (2009) 208-217.

[60] A.M. Schurko, D.J. Mazur, J.M. Logsdon Jr, Inventory and phylogenomic distribution of meiotic genes in *Nasonia vitripennis* and among diverse arthropods, Insect Mol. Biol., 19 (2010) 165-180.

[61] S. Patil, S. Moeys, P. von Dassow, M.J.J. Huysman, D. Mapleson, L. De Veylder, R. Sanges, W. Vyverman, M. Montresor, M.I. Ferrante, Identification of the meiotic toolkit in diatoms and exploration of meiosis-specific SPO11 and RAD51 homologs in the sexual species *Pseudo-nitzschia multistriata* and *Seminavis robusta*, BMC Genomics, 16 (2015) 930.

[62] S.-B. Malik, A.W. Pightling, L.M. Stefaniak, A.M. Schurko, J.M. Logsdon, Jr., An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*, PLoS ONE, 3 (2008) e2879.

[63] R. Müllenbach, P.J. Lagoda, C. Welter, An efficient salt-chloroform extraction of DNA from blood and tissues, Trends Genet., 5 (1989) 391.

[64] S. Koren, B.P. Walenz, K. Berlin, J.R. Miller, N.H. Bergman, A.M. Phillippy, Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation, Genome Res., 27 (2017) 722-736.

499 [65] J. Ruan, H. Li, Fast and accurate long-read assembly with wtdbg2, bioRxiv, (2019) 530972.
500 [66] M.J. Roach, S.A. Schmidt, A.R. Borneman, Purge Haplotigs: allelic contig reassignment for third-
501 gen diploid genome assemblies, BMC Bioinformatics, 19 (2018) 460.
502 [67] K. Adachi, H. Miyake, T. Kuramochi, K. Mizusawa, S.-i. Okumura, Genome size distribution in
503 phylum Cnidaria, Fisheries Sci., 83 (2017) 107-112.
504 [68] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing
505 genome assembly and annotation completeness with single-copy orthologs, Bioinformatics, 31 (2015)
506 3210-3212.
507 [69] R.M. Waterhouse, M. Seppey, F.A. Simão, M. Manni, P. Ioannidis, G. Klioutchnikov, E.V.
508 Kriventseva, E.M. Zdobnov, BUSCO applications from quality assessments to gene prediction and
509 phylogenomics, Mol. Biol. Evol., 35 (2018) 543-548.
510 [70] F.M. Waldron, G.N. Stone, D.J. Obbard, Metagenomic sequencing suggests a diversity of RNA
511 interference-like responses to viruses across multicellular eukaryotes, PLoS Genet., 14 (2018)
512 e1007533.
513 [71] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R.
514 Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren,
515 C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-
516 Seq data without a reference genome, Nat. Biotechnol., 29 (2011) 644-652.
517 [72] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D.
518 Eccles, B. Li, M. Lieber, M.D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman,
519 T. William, C.N. Dewey, R. Henschel, R.D. LeDuc, N. Friedman, A. Regev, *De novo* transcript sequence
520 reconstruction from RNA-seq using the Trinity platform for reference generation and analysis, Nat.
521 Protoc., 8 (2013) 1494-1512.
522 [73] M.H. Schulz, D.R. Zerbino, M. Vingron, E. Birney, Oases: robust *de novo* RNA-seq assembly across
523 the dynamic range of expression levels, Bioinformatics, 28 (2012) 1086-1092.

524 [74] Y. Xie, G. Wu, J. Tang, R. Luo, J. Patterson, S. Liu, W. Huang, G. He, S. Gu, S. Li, X. Zhou, T.-W. Lam,
525 Y. Li, X. Xu, G.K.-S. Wong, J. Wang, SOAPdenovo-Trans: *de novo* transcriptome assembly with short
526 RNA-Seq reads, Bioinformatics, 30 (2014) 1660-1666.
527 [75] M. Shao, C. Kingsford, Accurate assembly of transcripts through phase-preserving graph
528 decomposition, Nat. Biotechnol., 35 (2017) 1167.
529 [76] B.J. Haas, S.L. Salzberg, W. Zhu, M. Pertea, J.E. Allen, J. Orvis, O. White, C.R. Buell, J.R. Wortman,
530 Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to
531 Assemble Spliced Alignments, Genome Biol., 9 (2008) R7-R7.
532 [77] K.J. Hoff, S. Lange, A. Lomsadze, M. Borodovsky, M. Stanke, BRAKER1: unsupervised RNA-Seq-
533 based genome annotation with GeneMark-ET and AUGUSTUS, Bioinformatics, 32 (2016) 767-769.
534 [78] T. Honma, S. Minagawa, H. Nagai, M. Ishida, Y. Nagashima, K. Shiomi, Novel peptide toxins from
535 acrorhagi, aggressive organs of the sea anemone *Actinia equina*, Toxicon, 46 (2005) 768-774.
536 [79] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S.G. Rozen, Primer3 -
537 new capabilities and interfaces, Nucleic Acids Res., 40 (2012) e115.
538 [80] J.T. Hill, B.L. Demarest, B.W. Bisgrove, Y.-C. Su, M. Smith, H.J. Yost, Poly peak parser: method and
539 software for identification of unknown indels using sanger sequencing of polymerase chain reaction
540 products, Dev. Dyn., 243 (2014) 1632-1636.
541 [81] T.R. Ranallo-Benavidez, K.S. Jaron, M.C. Schatz, GenomeScope 2.0 and Smudgeplots: reference-
542 free profiling of polyploid genomes, bioRxiv, (2019) 747568.
543 [82] C.-L. Xiao, Y. Chen, S.-Q. Xie, K.-N. Chen, Y. Wang, Y. Han, F. Luo, Z. Xie, MECAT: fast mapping,
544 error correction, and de novo assembly for single-molecule sequencing reads, Nat. Methods, 14 (2017)
545 1072.
546 [83] Y.J. Liew, M. Aranda, C.R. Voolstra, Reefgenomics.Org - a repository for marine genomics data,
547 Database, 2016 (2016) baw152.
548 [84] M. Watson, A. Warr, Errors in long-read assemblies can critically affect protein prediction, Nat.
549 Biotechnol., 37 (2019) 124-126.

[85] S.D. Brown, O.D. Jarosinska, A. Lorenz, Genetic interactions between the chromosome axis-associated protein Hop1 and homologous recombination determinants in *Schizosaccharomyces pombe*, Curr. Genet., 64 (2018) 1089-1104.

[86] S. Tulin, D. Aguiar, S. Istrail, J. Smith, A quantitative reference transcriptome for *Nematostella vectensis* earlyembryonic development: a pipeline for de novo assembly in emerging model systems, EvoDevo, 4 (2013) 16.

[87] D. Wyman, A. Mortazavi, TranscriptClean: variant-aware correction of indels, mismatches and splice junctions in long-read transcripts, Bioinformatics, 35 (2018) 340-342.

[88] J.C. Sullivan, J.R. Finnerty, A surprising abundance of human disease genes in a simple "basal" animal, the starlet sea anemone (*Nematostella vectensis*), Genome, 50 (2007) 689-692.

[89] B. Wang, V. Kumar, A. Olson, D. Ware, Reviving the transcriptome studies: an insight into the emergence of single-molecule transcriptome sequencing, Frontiers Genet., 10 (2019).

[90] J.B. Jiang, A.M. Quattrini, W.R. Francis, J.F. Ryan, E. Rodríguez, C.S. McFadden, A hybrid *de novo* assembly of the sea pansy (*Renilla muelleri*) genome, GigaScience, 8 (2019).

[91] J.C. Kenyon, Models of reticulate evolution in the coral genus *Acropora* based on chromosome numbers: parallels with plants, Evolution, 51 (1997) 756-767.

[92] P.W. Shaw, J.A. Beardmore, J.S. Ryland, *Sagartia troglodytes* (Anthozoa: Actiniaria) consists of two species, Mar. Ecol. Prog. Ser., 41 (1987) 21-28.

[93] C.J. Rothfels, S.P. Otto, Polyploid speciation, in: R.M. Kliman (Ed.) The Encyclopedia of Evolutionary Biology, Academic Press, Oxford, 2016, pp. 317-326.

[94] L. Choleva, K. Janko, Rise and persistence of animal polyploidy: evolutionary constraints and potential, Cytogenet. Genome Res., 140 (2013) 151-170.

[95] L. Guo, A. Accorsi, S. He, C. Guerrero-Hernández, S. Sivagnanam, S. McKinney, M. Gibson, A. Sánchez Alvarado, An adaptable chromosome preparation methodology for use in invertebrate research organisms, BMC Biol., 16 (2018) 25.

575 [96] L.V. Clark, M. Jasieniuk, polysat: an R package for polyploid microsatellite analysis, Mol. Ecol. Res.,
576 11 (2011) 562-566.
577 [97] Z. Gompert, K.E. Mock, Detection of individual ploidy levels with genotyping-by-sequencing (GBS)
578 analysis, Mol. Ecol. Res., 17 (2017) 1156-1167.
579 [98] C.L. Weiß, M. Pais, L.M. Cano, S. Kamoun, H.A. Burbano, nQuire: a statistical framework for ploidy
580 estimation using next generation sequencing, BMC Bioinformatics, 19 (2018) 122.
581 [99] T.L. Shearer, M.J.H. van Oppen, S.L. Romano, G. Wörheide, Slow mitochondrial DNA sequence
582 evolution in the Anthozoa (Cnidaria), Mol. Ecol., 11 (2002) 2475-2487.
583 [100] A.M. Solé-Cava, J.P. Thorpe, Genetic divergence between colour morphs in populations of the
584 common intertidal sea anemones *Actinia equina* and *A. prasina* (Anthozoa: Actiniaria) in the Isle of
585 Man, Mar. Biol., 112 (1992) 243-252.
586 [101] J.P. Thorpe, A.M. Solé-Cava, The use of allozyme electrophoresis in invertebrate systematics,
587 Zool. Scr., 23 (1994) 3-18.

588

Figure Legends:

Figure 1.

Smudgeplot output with log scaling generated using default program settings. Diploid status (AB) is estimated to be marginally more likely (confidence = 0.39) than alternative ploidy status which includes triploidy (AAB, confidence = 0.28), tetraploidy (AAAB, confidence = 0.28) and others (variable, confidence = 0.05).

Figure 2.

Alignment of *Acrorhagin-1* haplotypes from geographic samples of *A. equina*. Seven haplotypes (1-7) were seen in total and are here aligned to *Acrorhagin-1* (Genbank accession number AB212066.1) and *Acrorhagin-1a* (AB212067.1) from Honma et al. [78]. Primers (Acro-F1 and Acro-R1) are shown above the sequence. Coding sequence is shown in bold.

Figure 1

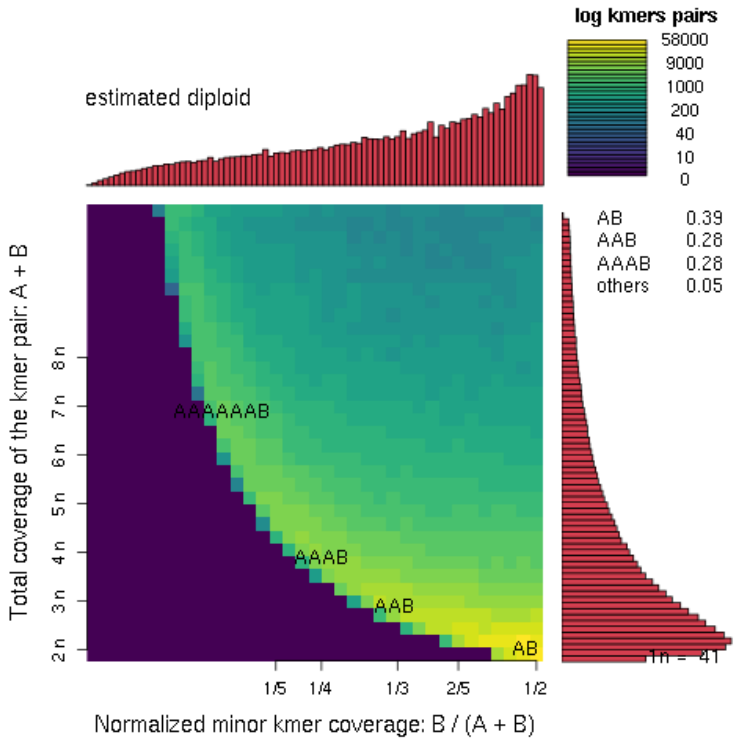
log kmers pairs
58000
9000
1000
200
40
10
0
estimated diploid
AB 0.39
AAB 0.28
AAAB 0.28
others 0.05
Total coverage of the kmer pair: A + B
2 n
3 n
4 n
5 n
6 n
7 n
8 n
AAAAAAB
AAAB
AAB
AB
1/5
1/4
1/3
2/5
1/2
Normalized minor kmer coverage: B / (A + B)

Figure 2

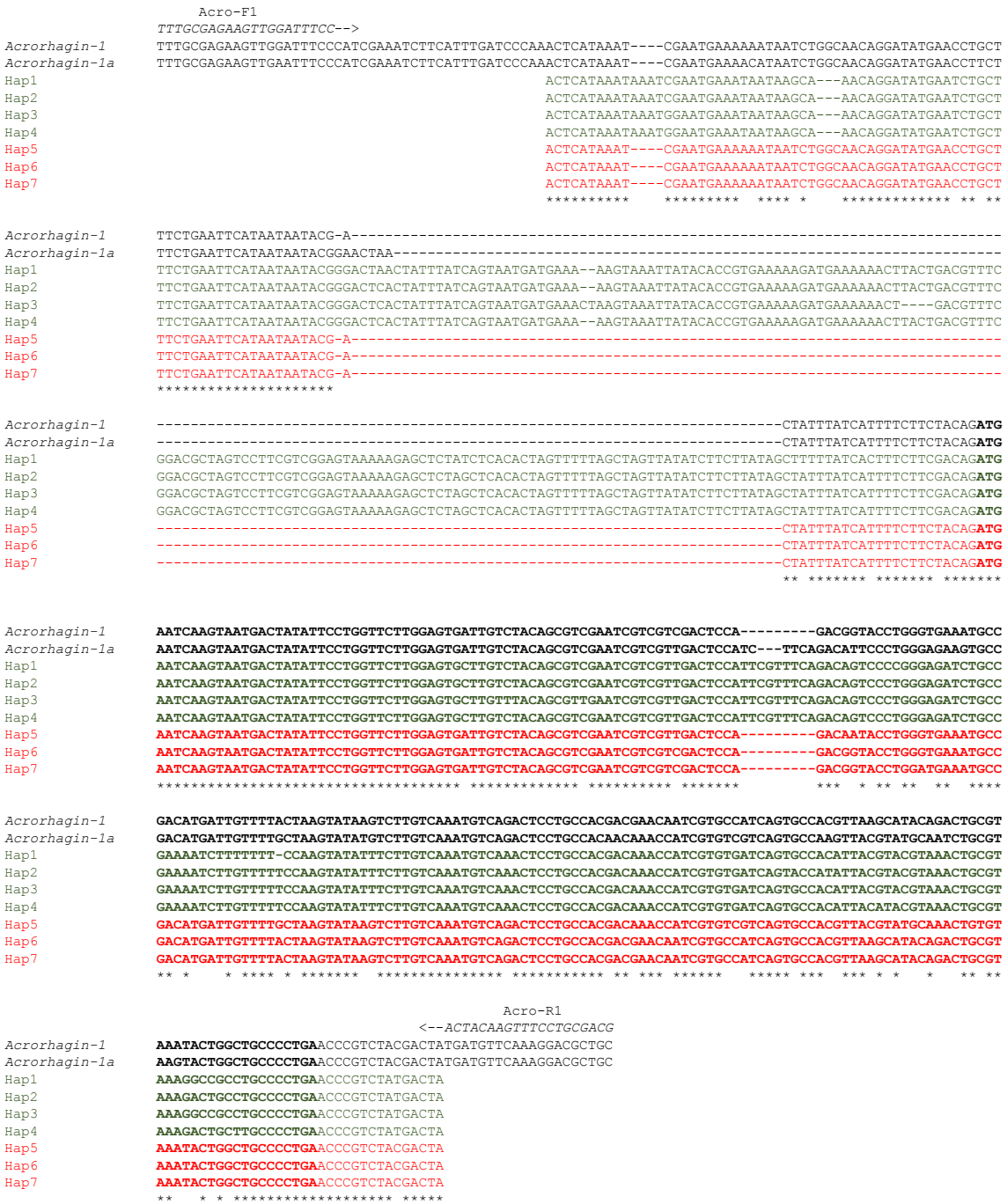
Acro-F1
TTTGCGAGAAGTTGGATTTCC-->
Acrorhagin-1
Acrorhagin-1a
Hap1
Hap2
Hap3
Hap4
Hap5
Hap6
Hap7
Acro-R1
<--ACTACAAGTTTCCTGCGACG

Table 1. MIxS descriptors

Investigation_type:	Eukaryote
Project_name:	The genome sequence of the sea anemone, *Actinia equina*
Lat_lon:	53.225889 N -4.524833 E
Geo_loc_name:	United Kingdom: Rhosneigr
Collected_by:	Craig Wilding
Collection_date:	10 April 2018
Environment	Intertidal zone
broad-scale environmental context :	ENVO:01000125
local-scale environmental context:	ENVO:01000428
environmental medium:	ENVO:00000319
Sample type:	Whole body
Developmental stage:	Adult
Sequencing method:	Pacbio sequel
Assembly method:	SMARTdenovo with Purge Haplotigs
Data accessibility:	BioProject: PRJNA479715
	BioSample: SAMN09602970
	Experiment: SRX4514416
	Raw read data: SRR7651651
	Genome: WHPX00000000

Table 2.

Meiotic toolkit genes studied in *A. equina*. Genes described as belonging to the meiotic toolkit in [57, 58, 61, 62] were examined. Gene models were complete (Ⓒ), partial at the 5' end (5'-Ⓟ), partial at the 3' end (Ⓟ-3'), or partial at both ends (5'-Ⓟ-3'). 1 = single base length variation seen between genomic model and transcript with genomic model corrected based upon transcript. 2 = closest match *Pms2*. 3= No Methionine at start. 4 = No stop codon. 5 = closest match *Cohesin subunit SA-1*. Genbank accession numbers of *A. equina* gene models are provided.

Gene	Description	[57]	[58]	[61]	[62]	Transcript	Gene	Accession
Brca2	Breast Cancer 2; DNA repair associated	●				Ⓟ	5'-Ⓟ-3'	MN307071
Dmc1	Meiotic recombination protein DMC1/LIM15 homolog		●	●	●	Ⓒ	Ⓒ	MN307072
Dna2	DNA replication factor *Dna2*	●				Ⓟ	5'-Ⓟ-3'	MN307073
Exo1	Exonuclease-1	●				Ⓟ	Ⓟ-3'	MN307074
Fancm	Fanconi anemia group M protein homolog	●				Ⓟ	Ⓒ	MN307075
Fen1	Flap endonuclease-1	●				Ⓒ	Ⓒ	MN307076
Hap2	Hapless 2				●	Ⓟ	Ⓟ-3'	MN307077
Hop1	HORMA domain-containing protein 1-like		●	●	●	x	x	
Hop2	Homologous pairing protein 2 homolog		●	●	●	Ⓒ	Ⓒ	MN307078
Mcm2	DNA replication licensing factor Minichromosome Maintenance Complex Component 2	●				Ⓒ	Ⓒ	MN307079
Mcm3	DNA replication licensing factor Minichromosome Maintenance Complex Component 3	●				Ⓒ	Ⓒ	MN307080
Mcm4	DNA replication licensing factor Minichromosome Maintenance Complex Component 4	●				Ⓒ	Ⓒ	MN307081
Mcm5	DNA replication licensing factor Minichromosome Maintenance Complex Component 5	●				Ⓒ	Ⓒ	MN307082
Mcm6	DNA replication licensing factor Minichromosome Maintenance Complex Component 6	●				Ⓒ	Ⓒ[1]	MN307083
Mcm7	DNA replication licensing factor Minichromosome Maintenance Complex Component 7	●				Ⓒ	Ⓒ	MN307084
Mcm8	DNA replication licensing factor Minichromosome Maintenance Complex Component 8	●				Ⓒ	Ⓒ	MN307085
Mcm9	DNA replication licensing factor Minichromosome Maintenance Complex Component 9	●				Ⓟ	Ⓟ-3'	MN307086
Mer3 = Hfm1	Helicase for Meiosis 1	●	●		●	x	x	
Mlh1	MutL Homolog 1	●	●			Ⓟ	Ⓒ	MN307087
Mlh2	MutL Homolog 2		●			Ⓟ	5'-Ⓟ-3'[2]	MN307088
Mlh3	MutL Homolog 3		●		●	Ⓒ	Ⓒ	MN307089
Mnd1	Meiotic nuclear divisions 1	●	●	●	●	Ⓟ	Ⓒ[3]	MN307090
Mre11	Meiotic Recombination 11 homolog	●	●			Ⓟ	Ⓒ[4]	MN307091
Msh2	mutS protein homolog 2-like	●	●	●		Ⓟ	Ⓟ-3'	MN307092
Msh4	mutS protein homolog 4-like	●	●	●	●	x	x	
Msh5	mutS protein homolog 5-like	●	●	●	●	Ⓟ	5'-Ⓟ	MN307093
Msh6	mutS protein homolog 6-like	●	●	●		Ⓒ	Ⓒ	MN307094
Mus81	Structure-specific endonuclease subunit MUS81	●			●	Ⓒ	Ⓒ	MN307095
Pch2	pachytene checkpoint 2				●	Ⓒ	Ⓒ	MN307096
Pds5	Precocious dissociation of sisters 5	●	●			Ⓟ	5'-Ⓟ-3'	MN307097
Pms1	Postmeiotic Segregation Increased 1	●	●			Ⓟ	Ⓟ-3'	MN307098
Rad1 (*Mei9*)	RAD1 cell cycle checkpoint protein	●	●			Ⓒ	Ⓒ	MN307099
Rad21	RAD21 Cohesin Complex Component	●	●	●		Ⓒ	Ⓒ	MN307100
Rad50	RAD50 Double Strand Break Repair Protein	●	●			Ⓒ	Ⓒ	MN307101
Rad51	RAD51 Recombinase	●	●	●		Ⓒ	Ⓒ	MN307102
Rad52	RAD52 DNA repair and recombination protein	●	●			Ⓒ	Ⓒ	MN307103
Rec8	Meiotic recombination protein REC8 homolog		●	●	●	x	x	
Scc3	Sister-chromatid cohesion protein 3/Stromalin	●	●			Ⓒ	Ⓒ[5]	MN307104
Smc1	Structural maintenance of chromosomes protein 1	●	●			Ⓒ	Ⓒ	MN307105
Smc2	Structural maintenance of chromosomes protein 2	●	●			Ⓒ	Ⓒ	MN307106
Smc3	Structural maintenance of chromosomes protein 3	●	●			Ⓒ	Ⓒ[1]	MN307107
Smc4	Structural maintenance of chromosomes protein 4	●	●			Ⓒ	Ⓒ	MN307108
Smc5	Structural maintenance of chromosomes protein 5	●	●			Ⓒ	Ⓒ	MN307109
Smc6 (*Rad18*)	Structural maintenance of chromosomes protein 6	●	●			Ⓒ	Ⓒ[1]	MN307110
Spo11	SPO11 Initiator Of Meiotic Double Stranded Breaks	●	●	●	●	Ⓒ	Ⓒ	MN307111
Zip4	Testis-expressed protein 11-like; "Meiosis protein SPO22/ZIP4 like"				●	x	x	

Table 3:
Assembly statistics from Canu, SMARTdenovo -/+ Purge Haplotigs, and WTDBG assemblers. BUSCO statistics refer to analysis of these genome assemblies (involving interim Augustus annotation) thus statistics differ from analysis of our detailed annotated gene models (see text for gene model BUSCO statistics).

	Canu*	SMRTdenovo	SMRTdenovo + PH	WTDBG2
Genome size	633,344,238	552,280,189	409,058,333	434,742,709
Number of contigs	8,123	2,705	1,485	5,621
Shortest contig	1,009	8,168	8,168	1,428
Longest contig	1,888,480	2,968,193	2,968,193	1,543,548
N50	134,191	381,457	492,607	208,156
Median	44,961	108,241	164,117	27,424
Mean	77,969	204,170	275,460	77,342
GC	38	37.62	38	
Complete BUSCOs (%)	93.2 (912)	94.1 (920)	94.0 (919)	
Complete and single-copy BUSCOs (%)	50.7 (496)	21.4 (209)	58.7 (574)	
Complete and duplicated BUSCOs (%)	42.5 (416)	72.7 (711)	35.3 (345)	
Fragmented BUSCOs (%)	1.1 (11)	0.5 (5)	0.6 (6)	
Missing BUSCOs (%)	5.7 (55)	5.4 (53)	5.4 (53)	

* with default error rate

Table 4:

Acrorhagin-1 haplotypes in anemone samples from UK and Irish collections. See Figure 2 for sequence of haplotypes 1-7. *N* = number of samples. Number of haplotypes assumes diploidy. Of 57 samples sequenced, 8 (14%) were repeated (including the two specimens demonstrating the singleton haplotypes 3 and 7) with identical results.

			Haplotype						
Location	Colour	*N*	1	2	3	4	5	6	7
New Brighton	Green	3	6						
	Red	6					6	6	
Holyhead	Green	2	4						
	Red	2					2	2	
Llandudno	Green	4	7	1					
	Red	5					4	6	
Marloes	Green	2	1		1	2			
	Red	2					2	2	
Rhosneigr	Green	3	4	2					
	Red	2					4		
Millport	Green	3	3	3					
	Red	3					1	5	
	Orange	3						6	
Niarbyl	Red	2						4	
Peel	Green	1	2						
	Red	4					4	4	
Penbryn	Red	2					2	1	1
Portmarnock	Green	4	8						
	Red	4					3	5	