Ge e a i f la ge i ch d ial a d . clea . cle ide e . e ce a d h l ge e ic a al e . i g high- h . gh . h - ead da a e f e da ge ed Plac li ae ail f he . h e Pacific

Mathieu Quenu, Steven A. Trewick, Elizabeth E. Daly and Mary Morgan-Richards

relationships among Placostylinae species living on separate islands, and among other related species.

The Placostylinae belong to the superfamily Orthalicoidea (Breure *et al.* 2010; Breure and Romero 2012) and contain among others the three genera *Placostylus* (H. Beck, 1837), *Emeculus* (Martens in Albers, 1860) and *Placochasis* (Pilsbry, 1900) that are endemic to

of ancient DNA extracted from fossil shells. We compare phylogenetic inferences from few taxa with long sequences to phylogenetic inferences obtained using limited amounts of data but a larger taxon sample. To achieve this, three distinct phylogenetic analyses are generated. First, phylogenetic relationships between the five taxa are estimated using large mitochondrial and nuclear sequence datasets. Second, we extract a short COXI sequence for the five taxa and incorporate the new sequences in phylogenetic analysis of *Placo l* and other closely related taxa (Trewick *e al.* 2009). Third, a multi-locus phylogenetic analysis containing mitochondrial and nuclear sequences is generated to investigate the relationships of sampled Placostylinae snails in relation to other Orthalicoidea sub-families (Breure and Romero 2012).

Tissue collection

We used foot muscle tissue samples from four species of land snail of the genus *Placo l* . The samples of the New Zealand species, *P. ambagio* and *P. hongii*, were from a frozen tissue collection originally developed from whole body samples harvested for allozyme analysis (Triggs and Sherley 1993). *Placo l ambagio* (PS185; NZ National Arthropod collection code GS4 of Triggs and Sherley 1993) was collected by G. Sherley from Cape Maria van Diemen, New Zealand. The specimen of *P. hongii* (PS257; NZ National Arthropod collection code WG865) came from eastern Far North, North Island. Although the precise provenance of PS257 is not recorded it was confirmed as *P. hongii* on the basis of shell morphology and region of origin. The two New Caledonian specimens, *P. fib a* (PS28) and *P. o h o om* (PS45), were collected in the field from populations on the Isle of Pines (Dowle *e al.* 2015). The Solomon Island specimen, *E. cle i* (PS127), was collected in Honiara, Guadalcanal.

DNA extraction

Foot muscle samples of approximately 50 mg were cut from specimens using sterile scalpel blades. Each tissue was pressed in a clean paper towel to remove excess storage ethanol and cut into smaller pieces. Whole genomic DNA was extracted using incubation at 55°C in CTAB buffer (2% hexadecyltrimethyl ammonium bromide, 100 mM Tris-HCL pH 8.0 20 mM EDTA) with proteinase K (Trewick *e al.* 2009). Following tissue digestion the solution was purified using an equal volume of 24:1 chloroform-isoamyl alcohol and centrifugation. DNA was precipitated from the aqueous fraction using sodium acetate (3M NaOAC) and chilled 95% ethanol. This extraction method has been found to be the most effective for isolating high molecular weight DNA from neogastropods while avoiding the problems of mucopolysaccharide contamination (Winnepenninckx *e al.* 1993). DNA was re-suspended in 50 µl or 100 µl TE buffer (10 mM Tris, 0.1 mM EDTA) and quantified using Qubit fluorometry (Life Technologies, Thermo Fisher Scientific Inc).

Illumina sequencing

Total DNA extracts from the five Placostylinae specimens were processed through massive parallel, high-throughput sequencing using the ThruPLEX DNA-seq kit (Rubicon Genomics). Fragmented genomic DNA was pair-end sequenced on an Illumina Hiseq 2500. Reads were de-multiplexed using standard indexes. Resulting Illumina short reads were trimmed of adapters and passed through standard quality filters using the software fastp (Chen *e al.* 2018). Reads were paired in Geneious v8 (Kearse *e al.* 2012).

Genome and gene assemblies

Mitochondrial genomes were assembled from each of five sets of 101 bp paired-read data files. Initially the mitochondrial genome of *P. ambagio* PS185 was assembled starting with available Sanger sequenced partial sequences of *Placo l* mtDNA COXI as a reference using the Geneious v8 mapping tools (Kearse *e al.* 2012). We then used iterative remapping to form consensus sequences from each previous mapping round using medium-low sensitivity with 25 iterations to assemble the full circular mitochondrial genome. Once assembled, this mitochondrial genome served as the initial reference for mapping of paired-end reads of other individuals, which were then iteratively remapped until maximum read coverage was achieved.

We used the MITOS web server (Bernt *e al.* 2013) to estimate gene annotations for one of the novel mtDNA genomes and used amino acid translation tools in Geneious to verify that each protein-coding sequence had an uninterrupted translation frame and the expected start and stop codons. We used ARWEN (Laslett and Canbäck 2008) to confirm secondary folding of transfer RNAs then transferred annotations to the other four similar genomes. Annotations were further checked for homology across the five genomes.

The same mapping approach was used to assemble the 45S nuclear ribosomal DNA cassettes of the five species, using available 5.8S and 28S sequences as our starting reference, and then mapping reads at medium-high sensitivity in Geneious (Kearse *e al.* 2012).

Fragments of histone 3 (H3) were also reconstructed by mapping reads of the five Placostylinae species onto reference sequences. Five 267 bp H3 fragments

were reconstructed based on GenBank data from the Placostylinae species *P. angei* (JF514684, Breure and Romero 2012) and five 291 bp H4 fragments were reconstructed based on GenBank data from *Oc la doli m* (Draparnaud, 1801) (KY512728, Harl *e al.* 2014). To reconstruct whole histone H3/H4 complexes (containing their noncoding spacer region), we first used one of the H4 fragments as a reference (PS127, *E. cle i*) and iteratively mapped corresponding reads at low sensitivity until the H3 fragment of the same individual could be mapped to a reference sequence. We then mapped paired-end reads of other specimens to the PS127 consensus sequence, but only reads of the New Caledonian species *P. fib a.* and *P. o h o om* mapped continuously to the reference sequence. We identified coding direction of the histone genes using the Geneious amino acid translation tool.

Phylogenetic analysis

Mitogenome and combined nuclear markers analysis of five Placostylinae

Phylogenetic analysis was performed for mitochondrial genomes and nuclear markers separately. Whole mitochondrial genomes, ribosomal cassettes and histone genes were aligned using the software MUSCLE (Edgar 2004

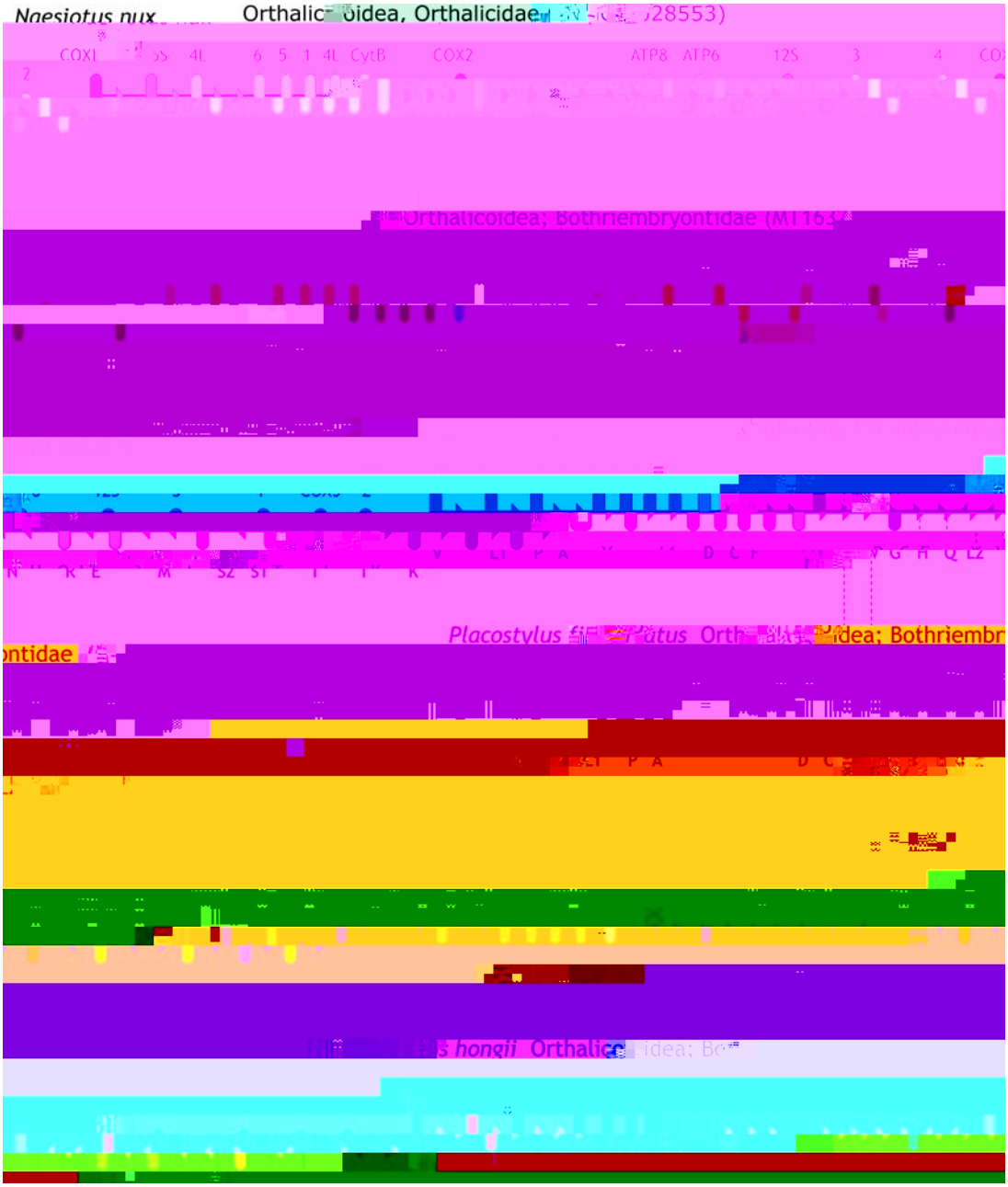

Figure 1. Comparison of mitochondrial genome gene order in five Placostylinae snail species. Full name of tRNA and CDS genes can be found in Table 3.

Histone genes H3 and H4 were reconstructed for the five Placostylinae species based on GenBank fragments from related species. Whole histone cassettes containing histone genes H3 and H4 and their noncoding spacer region could only be reconstructed for the Solomon Island species *E. cle i* (GenBank accession number: MT726982) and the two New Caledonian species *P. fib a* and *P. o h o om* (GenBank accession numbers: MT726983 and MT726984). In these cassettes both genes were coded on the same strand, and were separated by a noncoding spacer region of 707 bp (*P. o h o om*), 772 bp (*E. cle i*) or 800 bp (*P. fib a*). Histone genes H3 and H4 have been reported to be orientated in opposing directions in gastropod genomes (Armbruster *e al.* 2005; Harl *e al.* 2014), but other configurations have been reported in bivalves (Albig *e al.* 2003). Our findings confirm that histone gene cassettes are not configured consistently in gastropods. For the two New Zealand species, *P. ambagio* and *P. hongii*, the configuration of histone genes could not be retrieved from our data. Mapping reads from New Zealand species to consensus sequences of other species only led to discontinuous mapping reconstructions, with gaps and poor mapping resolution in the noncoding spacer region. It is hard to know if this result reflects read limitation in our data or indicates real biological information. If it were real, it would imply that histone genes H3 and H4 have physically distinct genome locations in

P. ambagio . and *P. hongii* rather than being organised in cassettes.

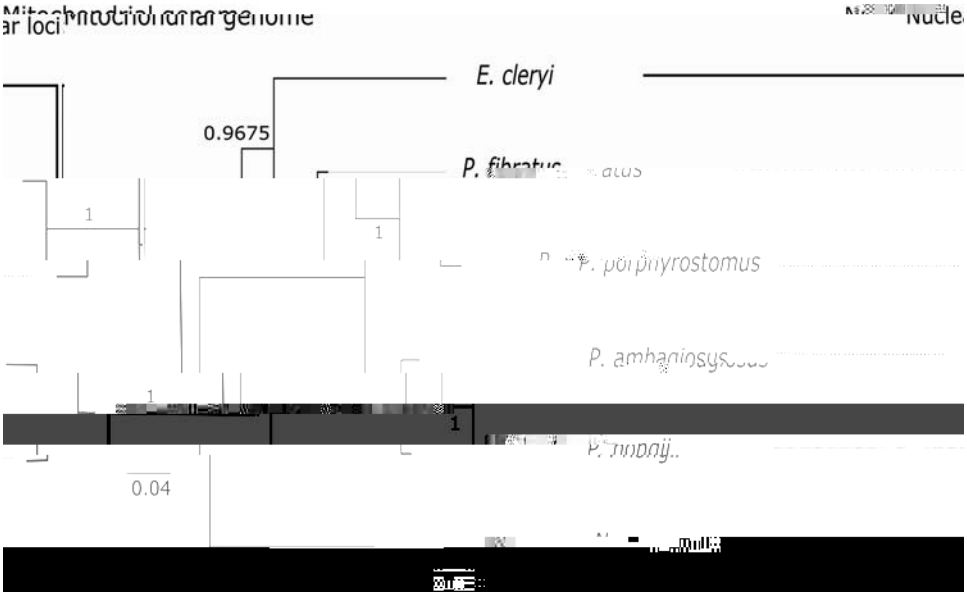

Figure 2. Phylogenetic relationships inferred for five giant terrestrial snails of the Pacific. *Naesiotus nux* (Bulimulidae) was used as the outgroup in the mitochondrial dataset analyses. *Placostylus fibratus* and *P. porphyrosotmus* are from New Caledonia, *P. ambagiosus* and *P. hongii* are from New Zealand and *E. cleryi* is from

redundancy, and using only one genus name for all species would seem more appropriate.

Multi-locus analysis of multiple Orthalicoidea species

Multi-locus analysis involving both mitochondrial (COXI) and nuclear (H3, 28S) sequences did not support the monophyly of the Placostylinae (Figure 4). Some Placostylinae species from the Solomon Islands (*P. angei*, *E. ligno*, *E. clei*) were grouped with species from other sub-families [*Bothriembryon d* (L. Pfeiffer, 1861); *Dicolpeus ameghinoi* (Ihering, 1908); *Peronella nuciali* (Melvill & Ponsonby, 1894); *Peronella bokei* (G.B. Sowerby III, 1890); *Plectostylus ianthinus* (Bruguière, 1789)], but support for the corresponding branch is low (BI posterior probability 0.87). This differs from previous findings in this group, which used the same combination of loci, but fewer taxa (Breure and Romero 2012).

Overall, all phylogenetic relationships reported here supported monophyly within geographical regions (or

Albig, W., Warthorst, U., Drabent, B., Prats, E., Cornudella, L. & Doenecke, D. (2003) *M il ed li* core histone genes are organized in two clusters devoid of linker histone genes. *Ja nal of Molec la E ol ion* 56, 597–606. doi:10.1007/s00239-002-2428-8.

Armbruster, G.F.J., Böhme, M., Bernhard, D. & Schlegel, M. (2005) The H3/H4 histone gene cluster of land snails (Gastropoda: Stylommatophora): TS/TV ratio, GC3 drive and signals in stylommatophoran phylogeny. *Ja nal of Moll can S. die* 71, 339–348. doi:10.1093/mollus/eyi038.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M. & Stadler, P.F. (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molec la Ph logene ic and E ol ion* 69, 313–319. doi:10.1016/j.ympev.2012.08.023.

Brescia, F.M., Pöllabauer, C.M., Potter, M.A. & Robertson, A.W. (2008) A review of the ecology and conservation of *Placo l* (Mollusca: Gastropoda: Bulimulidae) in New Caledonia. *Moll can Re ea ch* 28, 111–122.

Breure, A.S.H., Groenenberg, D.S.J. & Schilthuizen, M. (2010) New insights in the phylogenetic relations within the OrthalicoideaCalelim2(New)]TJ-4.97yor()T4g1fved