



# Allopatric mosaics in the Indo-West Pacific crab subfamily Chlorodiellinae reveal correlated patterns of sympatry, genetic divergence, and genitalic disparity

Robert M. Lasley Jr<sup>a,b,c,\*</sup>, Nathaniel Evans<sup>a,c,d</sup>, Gustav Paulay<sup>a</sup>, Francois Michonneau<sup>a</sup>, Amanda Windsor<sup>c</sup>, Irwansyah<sup>e</sup>, Peter K.L. Ng<sup>f</sup>

<sup>a</sup> Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA

<sup>b</sup> University of Guam, EPSCoR, Mangilao 96926, Guam

<sup>c</sup> National Museum of Natural History, Smithsonian Institution, Suitland, MD 20746, USA

<sup>d</sup> Tropical Biosphere Research Center, University of the Ryukyus, Nishihara, Okinawa 903-0213, Japan

<sup>e</sup> Syiah Kuala University, Aceh 23111, Indonesia

<sup>f</sup> Lee Kong Chian Natural History Museum, National University of Singapore, 117377, Singapore

## ARTICLE INFO

### Keywords:

Biogeography  
Phylogenetics  
Reproductive isolation  
Speciation  
Genitalia  
Selection

## ABSTRACT

Molecular studies have revealed that many species once thought to be wide-ranging in the Indo-West Pacific contain allopatric mosaics of endemic lineages. These lineages provide compelling evidence that substantial time is needed to evolve isolating mechanisms sufficient to permit successful secondary sympatry, and that divergence is initiated in allopatry. In this context, questions arise regarding the nature, timing, and origin of isolating mechanisms that permit secondary sympatry. We present a phylogeny of the crab subfamily Chlorodiellinae which displays allopatric mosaics within species. These allopatric lineages typically do not have divergent male genitalia, while older sympatric lineages do. We tested the relationship between genetic distance (proxy for time), sympatry, and the divergence of male genitalic morphology. Our results suggest that male genitalic divergence is not involved in the initiation of speciation in chlorodielline crabs, having likely occurred only after isolation began in allopatry. However, morphological evolution of genitalia seemingly does play an important role in completing the process of speciation in these crabs.

## 1. Introduction

Over the last several decades molecular phylogenetic studies have challenged traditional concepts regarding the diversity and distribution of marine species. Historically, researchers were left perplexed at how “marine speciation on a small planet” was possible given the broad ranges of many species, the prevalence of long-lived planktonic larvae, and the interconnectedness of marine habitats (Palumbi, 1992). Yet in the proceeding years, as predicted by Palumbi (1992) and others, molecular studies continued to reveal increasingly finer levels of geographic structuring and cryptic species in many marine taxa (e.g., Knowlton, 1993).

Much of the work on geographic differentiation in the marine realm has focused on fauna of the tropical Indo-West Pacific (IWP), the largest and most diverse marine region. Spanning two thirds of the marine

tropics from East Africa to the Eastern Pacific barrier, this region has long been recognized as a single, vast biogeographic entity with much of its fauna ranging across its entire extent (Forest & Guinot, 1961, Kay, 1984, Myers, 1994, Briggs & Bowen, 2012). Genetic studies have demonstrated that many marine species do maintain genetic connectivity across the IWP (e.g., Crandall et al., 2019, and references within). However, molecular studies have also revealed that many species once thought to be wide-ranging and interconnected contain allo- or parapatric mosaics of lineages with relatively restricted ranges (e.g., Meyer et al., 2005, Drew & Barber, 2009, Malay & Paulay, 2010, Titus et al., 2018).

Meyer et al. (2005) discussed this pattern for the gastropod genus *Astrarium*, a clade where two diverse subclades coexist in sympatry, but each comprises numerous allopatric lineages despite up to 6.2 % COI sequence divergence (an estimated 7 Ma of isolation) among sister

\* Corresponding author at: Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA.

E-mail address: [rlasleyjr@gmail.com](mailto:rlasleyjr@gmail.com) (R.M. Lasley).

<https://doi.org/10.1016/j.ympev.2023.107710>

Received 18 October 2022; Received in revised form 31 December 2022; Accepted 18 January 2023

Available online 24 January 2023

1055-7903/© 2023 Elsevier Inc. All rights reserved.

terminal taxa. What is surprising in *Astraliium*, is that allopatry seemingly persists for millions of years despite great dispersal capacity in evolutionary time, as demonstrated by the colonization of virtually all IWP reefs and islands and low levels of gene flow. Occasional instances of mitotypes from one lineage occurring in an individual within the range of another lineage further demonstrate low levels of gene flow and suggest that isolation in these allopatric lineages is not complete. In other words, allopatric lineages diverge despite some degree of gene flow, but require substantial time to evolve isolating mechanisms sufficient enough to permit successful colonization and coexistence of lineages (i.e., to establish secondary sympatry). In *Astraliium*, this was achieved only between the two subclades estimated to have diverged 30 Ma. This process is consistent with both Mayr's traditional model of geographic speciation (Mayr 1942, 1963, Nosil, 2008) and contemporary theory on gene flow and speciation (Nosil, 2008, Mallet et al. 2009, Wang et al. 2020), whereby some degree of homogenizing gene flow does not bar speciation and may even be common throughout the speciation process. In this context, questions arise regarding the nature, timing, and origin of isolating mechanisms (reproductive barriers) that permit secondary sympatry, and how these are related to those that allow the persistence of allopatric sister lineages in the face of considerable gene flow. However, investigating this remains challenging as genetic drift, natural selection, and sexual selection may work in tandem toward the completion of reproductive isolation before, during, or after the establishment of secondary sympatry (Servodio & Boughman, 2017).

In many brachyuran crab families, genitalic divergence is relied upon heavily to differentiate between closely related species (e.g., Serène, 1984) and has been implicated in the establishment and maintenance of species boundaries (Guinot et al. 2013, Yao et al., 2020). The first male gonopod (G1), which is the predominant intromittent structure in crabs, often exhibits a complex morphology with relatively little intraspecific variation that can be strikingly different among closely related species. The G1 is a tubular structure that delivers spermatophores into the female seminal receptacle (McClay & Becker, 2015), and the female vulva also exhibits complex morphology and interspecific differentiation (Guinot et al. 2013). Coevolution between the G1 and female vulva may drive divergent morphological evolution and play a part in speciation. Literature on terrestrial arthropods favors sexual selection as a driver of rapid divergent evolution of genitalia, although the mechanism is contested—i.e., cryptic female choice versus sperm competition versus sexual conflict (Waage, 1979, Thornhill, 1983, Gage, 1992, Arnqvist and Rowe, 1995, 2005, Eberhard, 1996, Arnqvist, 1999, Briceño & Eberhard, 2009). Very little is known about the mechanism driving the divergence of genitalia in crabs.

Like much of the fauna in the region, crabs in the IWP are exceptionally diverse with large species ranges, but their biogeography remains poorly documented. There have been virtually no genetic studies that use multiple true crab (Brachyura) species to test IWP biogeographic patterns. Given their abundance and ecological importance, studies using brachyuran species as models will enhance our understanding of IWP biogeography, especially since marine taxa exhibit diverse distributional patterns (e.g., Meyer et al., 2005, Malay & Paulay, 2010, Titus et al., 2019, Lasley et al., 2022). In addition to knowledge gaps in taxon-specific biogeography, few studies have investigated the selective forces driving speciation, including isolating mechanisms that permit secondary sympatry. To our knowledge, the connection between male genitalic divergence and biogeography in the IWP has only been studied in one taxon, the marine gastropod subfamily Littorininae (Hollander et al., 2013, 2018).

Here we investigate whether genetic divergence (a proxy for time) and genitalic divergence are correlated with the emergence of sympatric distributions among related lineages in the IWP crab subfamily Chlorodiellinae. Chlorodielline crabs are some of the most abundant crustaceans inhabiting coral reef environments in the IWP (Peyrot-Clausade 1977, 1979, Plaisance et al. 2011, Leray et al. 2012, Lasley et al. 2013, Lasley et al. 2015). Despite their ubiquity, a thorough taxonomic

revision of the group is still needed. Previous studies have focused on subfamilial and generic relationships (Lai et al., 2011, Lasley et al. 2013, 2015, Mendoza et al. 2022) but have not clarified several species-level problems that confound identification and lessen the taxon's utility as a model for other evolutionary studies. Through thorough taxonomic and geographic sampling, we develop an updated molecular phylogeny of the subfamily. Using several standard molecular species delineation methods, we first group specimens into Molecular Operational Taxonomic Units (MOTUs). We then use the resulting MOTUs to develop a robust multilocus phylogeny. Finally, using a phylogenetic comparative approach we test for correlation between allopatry, timing of divergence, and genitalic divergence.

## 2. Material and methods

### 2.1. Voucher material and taxon sampling

Specimens for morphological and molecular analyses were obtained from the Florida Museum of Natural History, University of Florida, Gainesville, Florida, USA (UF); Zoological Reference Collection of the Lee Kong Chian Natural History Museum, National University of Singapore, Singapore (ZRC); Queensland Museum, Brisbane, Australia (QM); Muséum national d'histoire Naturelle, Paris, France (MNHN), and Senckenberg Museum, Frankfurt am Main, Germany (SMF). All voucher material and sequence data used in this study are listed in SM1 and SM2.

We examined a total of ca. 1,000 specimens representing all 31 well-established and two undescribed chlorodielline species. Four species of uncertain status that have not been reported since their original descriptions were not available for study: *Chlorodiella quadrilobata* Dai, Cai, & Yang, 1996, *Chlorodiella ohshimai* Miyake & Takeda, 1967, *Cyclodius perlatus* Nobili, 1905, and *Pilodius kauaiensis* Edmondson, 1962. A further species, *Cyclodius drachi* Guinot, 1964, was not included either, as it most likely is a junior synonym of *Cyclodius granulatus* (Targioni Tozzetti, 1877). These taxa will be commented on in upcoming taxonomic revisions (Lasley, in prep.). Morphological and sequence data were obtained from all the remaining 33 species (Table 1). Identifications were based on keys and diagnoses from Forest & Guinot (1961), Guinot (1964), Serène (1984), Dai & Yang (1986), and Clark & Galil (1993).

### 2.2. Generation and composition of molecular datasets

For this study we generated and compiled two molecular data sets: the first, composed of mitochondrial COI sequences, was used for species and MOTU delineation analyses; the second, composed of four standard loci (mitochondrial COI, 12S rRNA, 16S rRNA, and nuclear H3), was used for phylogenetic analyses of Chlorodiellinae and subsequent generation of an ultrametric topology for phylogenetic comparative analyses.

COI sequence data was obtained from 892 specimens across 33 species. This included hundreds of sequences generated for this study during several large DNA barcoding initiatives including the Moorea Biocode project, MarBOL, and the Southern Line Island survey. These sequences were combined with additional sequences that we generated to provide a more thorough, representative sampling of species, morphological variants, and localities. Specimens from each of these previous sampling campaigns are housed at UF and sequence identifications were confirmed through morphological examination when specimens were available. Analyses of our COI data set led to the delineation of 38 reciprocally monophyletic, well-supported terminal clades (i.e., MOTUs, discussed below), which were then each targeted for development of our phylogenetic data set (listed in SM2).

For newly generated sequence data, DNA was isolated from ambulatory leg muscle tissue using an AutoGeneprep 965 at the Smithsonian Institution Laboratories of Analytical Biology (LAB). PCR amplifications were performed using the following primer sets: 12sf and 12slr (12S)

**Table 1**

List of chlorodielline species and MOTUs (in bold) in COXI gene trees with corresponding support values (posterior probabilities (pP) from the BI analysis, bootstrap values (BS) from the ML analyses); geographic overlap and G1 uniqueness; and delineation results. N = number of sequences included in analyses. “=” means cluster equivalent to clade. Clades recovered as more than one cluster indicated by number of clusters.

Taxon	Intraspecific variation	Smallest Interclade Distance	pP	BS	N	# haplotypes	Geographic Overlap With Closest Clade	G1 uniqueness	Species Identifier	ABGD	GMYC	PTP
<i>Chlorodiella barbata</i>	0–5.3	10.3	1	100	64	11	n/a	n/a	2 clusters	2 clusters	2 clusters	2 clusters
<b><i>Chlorodiella barbata</i> MOTU 1 (1)</b>	0–1.1	4.0	1	88	59	8	narrowly sympatric	shared	=	=	=	=
<b><i>Chlorodiella barbata</i> MOTU 2 (2)</b>	0–1.3	4.0	1	83	5	3	narrowly sympatric	shared	=	=	=	=
<b><i>Chlorodiella coclearis</i> (3)</b>	0–1.3	11.7	1	100	10	7	sympatric	unique	=	=	=	=
<i>Chlorodiella cytherea-davaoensis-crispipleopa</i>	0–12.2	9.3	1	98	99	52	n/a	n/a	5 clusters	5 clusters	7 clusters	6 clusters
<b><i>Chlorodiella cytherea</i> MOTU 1 (4)</b>	0–2.9	7.7	1	100	40	28	sympatric	shared	=	=	2 clusters	=
<b><i>Chlorodiella cytherea</i> MOTU 2 (5)</b>	0	6.1	1	100	14	1	allopatric	shared	=	=	=	=
<b><i>Chlorodiella cytherea</i> MOTU 3 (6)</b>	1.6–1.9	6.9	1	100	3	3	allopatric	shared	=	=	=	=
<b><i>Chlorodiella cytherea</i> MOTU 4 (7)</b>	0–2.7	6.1	1	95	37	15	allopatric	shared	=	=	2 clusters	2 clusters
<b><i>Chlorodiella cytherea</i> MOTU 5 (8)</b>	0–0.5	9.0	1	100	5	5	allopatric	shared	=	=	=	=
<b><i>Chlorodiella laevissima</i> (9)</b>	0–4.5	6.6	1	97	60	20	sympatric	unique	2 clusters	2 clusters	2 clusters	5 clusters
<i>Chlorodiella</i> sp. 1	0–7.4	6.6	1	99	37	24	n/a	n/a	2 clusters	3 clusters	3 clusters	3 clusters
<b><i>Chlorodiella</i> sp. 1 MOTU 1 (10)</b>	0–1.3	3.2	1	88	18	10	allopatric	shared	n/a	=	=	=
<b><i>Chlorodiella</i> sp. 1 MOTU 2 (11)</b>	0–1.1	3.2	1	98	7	4	allopatric	shared	n/a	=	=	=
<b><i>Chlorodiella</i> sp. 1 MOTU 3 (12)</b>	0–1.9	4.8	1	98	12	10	narrowly sympatric	shared	=	=	=	=
<i>Chlorodiella</i> sp. 2	0–4.5	7.7	1	97	112	48	n/a	n/a	1 cluster	3 clusters	3 clusters	3 clusters
<b><i>Chlorodiella</i> sp. 2 MOTU 1 (13)</b>	0–0.8	2.9	1	82	38	18	narrowly sympatric	shared	n/a	=	=	=
<b><i>Chlorodiella</i> sp. 2 MOTU 2 (14)</b>	0–4.0	2.9	1	87	74	30	narrowly sympatric	shared	n/a	2 clusters	2 clusters	2 clusters
<i>Chlorodiella nigra-spinimera</i>	0–4.8	6.6	0.9	99	40	25	n/a	n/a	1 cluster	1 cluster	3 clusters	3 clusters
<b><i>Chlorodiella nigra</i> MOTU 1 (15)</b>	0–3.2	2.1	0.9	n/a	13	12	narrowly sympatric	shared	n/a	n/a	2 clusters	2 clusters
<b><i>Chlorodiella nigra</i> MOTU 2 (16)</b>	0–1.6	2.1	0.8	40	27	13	narrowly sympatric	shared	n/a	n/a	=	=
<b><i>Chlorodiella xishaensis</i> (17)</b>	0–0.3	6.6	1	100	28	4	sympatric	unique	=	=	=	=
<b><i>Cyclodius granulatus</i> (18)</b>	0–1.4	7.6	1	100	3	2	sympatric	unique	=	=	=	=
<b><i>Cyclodius granulatus</i> (19)</b>	0–0.1	5.8	1	100	9	7	sympatric	unique	=	=	=	=
<i>Cyclodius nitidus</i>	0–9.1	5.8	n/a	n/a	53	20	n/a	n/a	2 clusters	2 clusters	2 clusters	2 clusters
<b><i>Cyclodius nitidus</i> MOTU 1 (20)</b>	0–0.4	5.8	1	100	34	6	allopatric	shared	=	=	=	=
<b><i>Cyclodius nitidus</i> MOTU 2 (21)</b>	0–2.8	5.8	1	78	19	14	allopatric	shared	=	=	=	=
<i>Cyclodius obscurus</i>	0–5.8	11.5	1	100	30	14	n/a	n/a	2 clusters	2 clusters	2 clusters	2 clusters
<b><i>Cyclodius obscurus</i> MOTU 1 (22)</b>	0	5.4	1	100	3	1	allopatric	shared	=	=	=	=
<b><i>Cyclodius obscurus</i> MOTU 2 (23)</b>	0–0.8	5.4	1	76	27	13	allopatric	shared	=	=	=	=
<b><i>Cyclodius paumotensis</i> (24)</b>	0	10.1	1	100	3	1	sympatric	unique	=	=	=	=
<b><i>Cyclodius unguatus</i> (25)</b>	0–3.2	10.7	1	100	57	37	sympatric	unique	=	=	=	=
<b><i>Luniella scabriculus</i> (26)</b>	0–1.2	12.8	1	100	15	6	sympatric	unique	=	=	=	=
<b><i>Luniella spinipes</i> (27)</b>	0–0.2	11.4	1	92	16	4	sympatric	unique	=	=	=	=
<b><i>Luniella pubescens</i> (28)</b>	0–0.6	8.3	1	75	29	20	sympatric	unique	=	=	=	=
<b><i>Luniella pugil</i> (29)</b>	0–0.8	8.3	1	91	21	8	sympatric	unique	=	=	=	=
<b><i>Pilodius nigrocrinitus</i> (30)</b>	0–0.4	11.4	1	100	11	3	sympatric	unique	=	=	=	=
<b><i>Pilodius pilumnoides-concours-cephalalgicus</i> (31)</b>	0–2.9	10.2	1	100	12	9	sympatric	unique	=	=	=	=
<b><i>Pilodius maotieni</i> (32)</b>	0–1.2	10.2	1	100	11	5	sympatric	unique	=	=	=	=
<b><i>Pilodius miersi</i> (33)</b>	0–0.8	10.6	1	100	14	3	sympatric	unique	=	=	=	=
<b><i>Pilodius granulatus-philippinensis</i> (34)</b>	0–2.3	11.2	1	100	12	8	sympatric	unique	=	=	=	=
<b><i>Pilodius moranti</i> (35)</b>	0–0.8	9.4	1	100	4	4	narrowly sympatric	unique	=	=	=	=
<b><i>Pilodius areolatus</i> (36)</b>	0–2.7	9.4	1	99	58	8	narrowly sympatric	unique	=	=	=	=
<b><i>Soliella flava</i> (37)</b>	0–1.4	12.0	1	100	45	6	sympatric	unique	=	=	6 clusters	=
<b><i>Soliella melanospinis</i> (38)</b>	0–0.8	12.0	1	100	39	16	sympatric	unique	=	=	2 clusters	=
Totals									36	39	49	45

(Mokady and Graur, 1994; Shull et al., 2005), crust16sf1 and crust16sr2 (16S) (Lai et al., 2009); jgHCO2198 and jgLCO1490 (COI) (Geller et al., 2013); and H3af and H3ar (H3) (Colgan et al., 1998). Amplifications consisted of 19- $\mu$ L reactions with 10  $\mu$ L of Promega GoTaq G2 Hot Start Master Mix, 0.6  $\mu$ L of each 10  $\mu$ M primer, and 0.2  $\mu$ L of 20 mg/mL BSA. The following PCR cycling parameters were used: initial denaturation at 95 °C for 5 min; 4 cycles at 94 °C for 30 s; 57 °C (12S), 52 °C (16S), 50 °C (COI), or 50 °C (H3) for 45 s; 72 °C for 1 min; then 34 cycles at 94 °C for 30 s; 52 °C (12S), 47 °C (16S), 45 °C (COI), or 47 °C (H3) for 45 s; and a final extension at 72 °C for 8 min. Successful PCR products were cleaned up and bidirectionally sequenced using standard LAB protocols as described in Evans (2018). Contig consensus sequences were constructed using Geneious version 7.1.4 (Biomatters Ltd.). COI sequences generated for this study were trimmed to 581 bp to match the length of sequences used from previous studies (Lasley et al., 2015). GenBank accession numbers for newly generated sequences are listed in SM1 and SM2).

Alignments for COI and H3 were generated using MAFFT (Katoh & Standley, 2013) and absence of spurious stop codons was confirmed. Ribosomal RNA sequences (12S and 16S) were aligned using Guidance2 (Sela et al., 2015) implementing MAFFT's G-INS-i settings (–globalpair –maxiterate 1000). For each rRNA marker, Guidance2 evaluated 400 alternative alignments from 100 alternative guide trees. Columns in the alignment with confidence scores below 0.93 were trimmed resulting in a final alignment for 12S of 355 bp (reduced from 389 bp) and for 16S of 504 bp (reduced from 522 bp). A final concatenated alignment of all four loci was 1768 bp in length and initially partitioned by locus with the protein coding genes partitioned by codon (SM3).

### 2.3. Gene-based species delineation

Using COI sequence data, we investigated the validity of all 31 well-established chlorodielline species and attempted to identify all distinct MOTUs within the group.

The monophyly of each described species was evaluated using unrooted COI gene trees generated for each chlorodielline genus. Phylogenetic analyses of COI sequence data were carried out separately for each genus using both maximum-likelihood (ML) and Bayesian Inference (BI) methods. ML analyses were carried out using Randomized Accelerated Maximum Likelihood (RAxML) version 7.7.7 (Stamatakis, 2014) on the CIPRES Science Gateway (Cyberinfrastructure for Phylogenetic Research project; Miller, Pfeiffer, & Schwartz, 2011). Searches were carried out using a GTR + gamma substitution model and confidence was assessed by generating 1000 non-parametric bootstrap replicates. BI analyses were carried out using MrBayes 3.2.2 (Ronquist & Huelsenbeck, 2003), also on CIPRES. Substitution models for BI analyses were selected using JModeltest version 2.1.4 (Posada, 2008) employing the Akaike information criterion. Selected models for *Chlorodiella*, *Cyclodius*, *Luniella*, *Pilodius*, and *Soliella* were TrN + G, TrN + I, HKY + G, HKY + G, and HKY + G, respectively. Each BI Markov chain Monte Carlo (MCMC) analysis was run for 10 million generations (with nruns = 2 nchains = 4 temp = 0.2 samplefreq = 1000) with a burn-in value of 20 %. Average standard deviation of split frequencies for each analysis was confirmed to be  $\leq 0.01$ . Convergence of the runs were assessed using Tracer version 1.5 (Rambaut & Drummond, 2009) including confirmation that ESS values were above 200.

Genetically distinct MOTUs within each chlorodielline genus were identified using COI data and four delineation approaches: (1) General Mixed Yule Model Coalescent (GMYC) (Pons et al., 2006, Monaghan et al., 2009); (2) SpeciesIdentifier from the TaxonDNA version 1.6.2 package (Meyer et al., 2006); (3) the Poisson Tree Processes model (PTP) (Zhang et al., 2013); and (4) Automatic Barcoding Gap Discovery (ABGD) (Puillandre et al., 2012). Default settings in the programs were used, except as described below.

For delineations using GMYC, an ultrametric, bifurcating input tree was created for each genus using BEAST version 1.10.2 on CIPRES (Drummond & Rambaut, 2007, Suchard & Rambaut, 2009). Input files were generated using BEAUti version 1.10.2. Each BEAST MCMC analysis was run for 20 million generations under a strict clock and HKY substitution model. HKY was chosen instead of the models selected by JModeltest (above) because of poor mixing due to over-parameterization in initial runs. Tree prior was set to “Speciation: Yule Process” (Yule, 1925, Gernhard, 2008). Trees were sampled every 2000 generations. Tracer version 1.7.1 was used to confirm that ESS values were above 200. Sampled trees were summarized using TreeAnnotator version 1.10.2 with a burn-in value of 10 % (Drummond & Rambaut, 2007). Trees were used to fit the GMYC model using the R package splits (Ezard et al., 2013).

For delineation using SpeciesIdentifier, we used the “cluster” module and iteratively tested threshold values of 1–4 %, ultimately using a value of 3 % to analyze each genus. PTP analyses were run on the webportal (<https://species.h-its.org/>), using our ML topologies generated in RAxML as inputs, 500,000 MCMC generations, a thinning value of 100 and a burn-in of 0.1. ABGD analyses were carried out using the webportal (<https://www.wabi.snv.jussieu.fr/public/abgd/abgdweb.html>). X values were iteratively tested and a value of X = 1 was selected. Other values were set to default.

MOTUs were designated for clades that were delineated by all methods (SpeciesIdentifier, GMYC, PTP, and ABGD) equivalently. When delineation methods gave contradictory results, MOTU designations were assigned based on the least inclusive terminal-most clades that were both reciprocally monophyletic and well-supported in the phylogenetic analyses (i.e., both the COI-only and four-loci concatenated analyses, described below). One exception, regarding a *Chlorodiella nigra* + *Chl. xishaensis* clade, is discussed in the results.

#### 2.4. Species and MOTU gonopod morphology

In addition to considering the gross morphology of each specimen, we evaluated our species- and MOTU-level delineations by

characterizing the morphological distinctiveness of the male first gonopods (G1). Like many crabs, the G1s of Chlorodiellinae comprise a diverse set of structurally complex morphotypes that do not lend themselves easily to quantitative characterization or complex, multi-character coding schemes. Yet, these structures can be strikingly distinct and highly conserved within—and diagnostic of—species or specific lineages. Consequently, here we identified all chlorodielline G1 morphotypes and for each MOTU determined if their respective morphotypes were ‘shared’ (same morphology as another MOTU) or ‘unique’ (morphology unique to MOTU, not shared). The appropriateness of this discrete characterization of G1 morphotypes was evaluated through careful examination of variation within and between MOTUs.

G1s of all male specimens used in this study were carefully examined using light microscopy and sorted to morphotype. Representatives of each G1 morphotype were further examined with a Leica Stereoscan 440 Scanning Electron Microscope (SEM). Mucus and debris were removed from G1s following Felgenhauer (1987). Samples were dehydrated through a graded ethanol series, followed by two changes in HMDS (hexamethyldisilazane), mounted on stubs using Elmer’s glue, and coated with 25 nm 60:40 gold:palladium using a Cressington Sputter Coater 108auto.

#### 2.5. Phylogenetic analyses

To assess the phylogenetic relationships of all chlorodielline species and MOTUs, ML and BI analyses were performed on the partitioned concatenated data set of four markers (COI, 12S, 16S, and H3) with *Etisus demani* Odhner, 1925 as outgroup, following Lasley et al. (2015). For many species, this data set included multiple specimens, some with only partial sequence data. Consequently, two concatenated datasets were constructed and analyzed: the first consisted of 79 (78 ingroup) taxa representing all sequenced 1; the second was a reduced set composed of 39 (38 ingroup) taxa representing the best (most complete) single exemplar for each species and MOTU. Composition of these data sets, including GenBank accession numbers for each marker, are listed in SM2. Partition schemes and substitution models used in these analyses are listed in SM3 and were selected in PartitionFinder2 (Lanfear et al., 2017) using the Bayesian Information Criterion. Selected models were evaluated against all models available in MrBayes (for BI analyses) and RAxML (for ML analyses), with the exception that + I models were not evaluated for ML analyses.

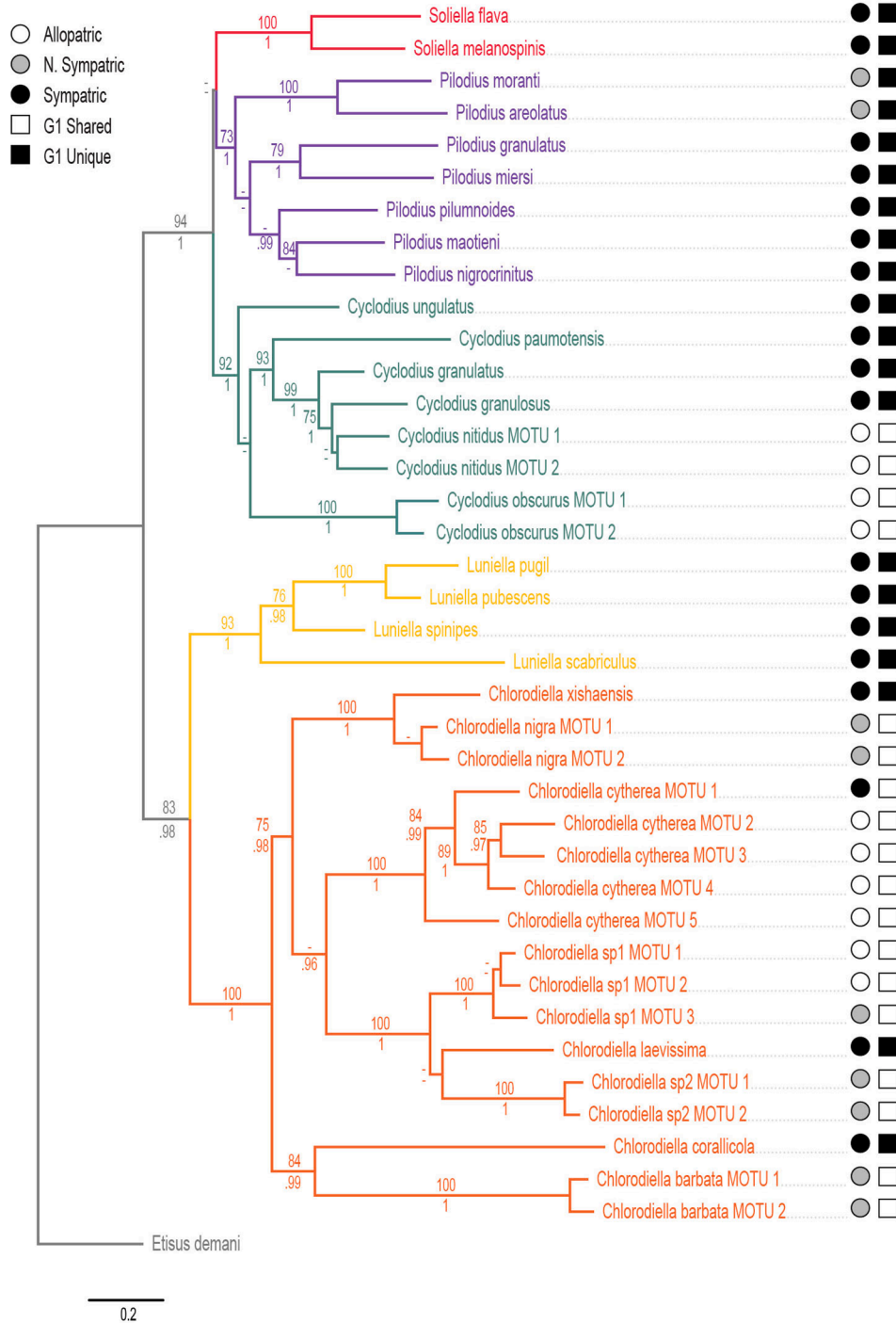
Analyses of the 79 taxa and 39 taxa datasets were carried out using the same search parameters. BI analyses were carried out using MrBayes version 3.2.7a on CIPRES and run 25 million generations (with nruns = 2 nchains = 4 temp = 0.175 samplefreq = 10000) with a burn-in value of 10 %. Average standard deviation of split frequencies was confirmed to be  $\leq 0.01$  and convergence of the runs was assessed using Tracer version 1.7.2, including confirmation that ESS values were above 200. Summary topologies were constructed using both a majority-rule consensus (MRC) tree and a maximum clade credibility (MCC) topology and compared. ML analyses were carried out using RAxML v.8.2.12 on CIPRES. For each dataset two independent runs were performed, each generating 1000 non-parametric bootstrap replicates summarized on the best scoring topology from 10 independent searches. Final nodal support values were summarized on the best scoring topologies using 2000 bootstrap replicates combined from the two separate runs. Additionally, 100 independent searches for the best scoring tree were performed for both data sets. The 100 independent searches recovered a slightly different “best scoring” topology than the first set of analyses (discussed below). To evaluate if one of these topologies was significantly more likely a log likelihood Shimodaira-Hasegawa (SH) test (Shimodaira & Hasegawa, 1999) was performed in RAxML.

For subsequent phylogenetic comparative analyses (discussed below) one of our “best scoring” ML topologies was used to generate an ultrametric tree (as a chronogram) with the ‘chronos’ function in the R package ‘ape’ version 5.5 (Sanderson, 2002, Kim & Sanderson, 2008,



Paradis, 2013). This function was carried out with  $\lambda = 0$ , model = relaxed, and the following calibration intervals (with soft.bounds = true): age.max and age.min values for the Chlorodiellinae root were set to 33.9 Ma and 23.03 Ma respectively, while values for *Chlorodiella*'s root were set to 13.65 Ma and 12.70 Ma. These intervals are consistent with the putative Chlorodiellinae stem-group fossil *Prochlorodius* Müller & Collins, 1991 occurring during the late Eocene (Fraaye 1996, Müller & Collins, 1991) with reliable crown group fossils not appearing until the Miocene (Fraaye, 1996), among which the earliest reliable fossil *Chlorodiella* appears during the Badenian age (Middle Miocene, 13.65–12.70

Ma) (Müller, 1984, Collins, 2014). The resulting chronogram is provided in SM4, and again, was generated to function as an ultrametric input topology for subsequent analyses (discussed next). Given that this chronogram was generated using a simple penalized likelihood approach, we urge caution when interpreting any specific divergence dates. Nevertheless, this chronogram does provide a working hypothesis about the evolution of Chlorodiellinae, but the topic deserves a more thorough treatment, so we refrain from commenting any further on it here.



**Fig. 1.** Maximum likelihood tree inferred from combined 12S and 16S rRNA genes, COI and histone H3 sequences showing geographic overlap and genitalic divergence. Numbers above and below branches indicate maximum-likelihood bootstrap support (BS) and Bayesian inference posterior probability (PP), respectively. Values below 0.95 (PP) and 70 (BS) are represented by “-”. Missing PP values indicate clades not recovered in the Bayesian analysis.

## 2.6. Evolutionary patterns of geography, genitalia and genetic divergence

We investigated the evolutionary covariance of genetic divergence, geography and genitalic differentiation between chlorodielline MOTUs using a phylogenetic comparative implementation of the quantitative genetic threshold model (Felsenstein, 2012). Under this approach the evolution of discrete and continuous characters is modeled along a phylogeny as a function of an underlying continuous trait (a ‘liability’). Using MCMC, the covariance of liabilities is inferred between multiple traits and correlation is modeled between two discrete characters, or one discrete and one continuous character. For these analyses we used our chronogram and the character data listed in Table 1 and depicted in Fig. 1. This data included MOTU interclade genetic distances (a continuous character), G1 uniqueness (a discrete character), and geographic overlap (a discrete character).

Shortest interclade distances between MOTUs were obtained from COI distances calculated with TaxonDNA (Meier et al., 2006). As previously described, G1 morphology for each MOTU was scored as ‘shared’ or ‘unique’ based on our identification of distinct G1 morphotype.

For geographic overlap, the distributional ranges of each MOTU were compared with the ranges of its sister clade and scored as ‘sympatric’, ‘narrowly sympatric’, or ‘allopatric’. Distributional ranges of each MOTU were determined by mapping occurrence records from examined specimens, sequence data, and reliable, applicable records in the literature (SM2). Geocoordinates for literature records were obtained using Google Earth. Occurrence maps were generated with the R package ggplot2 (Wickham et al., 2016). When MOTUs were sister to clades with multiple MOTUs, the ranges of all MOTUs in the sister clade were combined. For example, *Chl. cytherea* MOTU 1 is sister to a clade comprising *Chl. cytherea* MOTUs 2–4. In this case, *Chl. cytherea* MOTU 1 was scored as ‘sympatric’ because it overlaps with the combined ranges of *Chl. Cytherea* MOTUs 2–4. ‘Narrowly sympatric’ was defined as sister MOTUs that co-occur at < 10 % of the MOTU with the larger range. Distributional data were also used to create heatmaps of species richness using the R package monographaR (Reginato, 2016). To gauge the importance of data generated for this study, separate and combined heatmap were generated for distributional data from literature and the material examined.

Evolutionary correlation of the above-described characters was investigated using a Bayesian implementation of the threshold model with the function ‘threshBayes’ in the R package phytools (Revell, 2012, 2014). Given that discrete characters must be binary for this approach, analyses of geography were performed twice with all ‘narrowly sympatric’ taxa coded as either ‘allopatric’ or ‘sympatric’. Details of each analysis and character coding schemes are listed in SM5. Each thresh-Bayes analysis was run for 10 million MCMC generations, sampled every 1000, with a burn-in of 20 percent. Posterior densities of estimated correlation coefficients ( $r$ ) were plotted, and mean values recorded. Stability of each run was confirmed by plotting log-likelihood values for all generations, and ESS values for each analysis were calculated and confirmed to be > 200 using the R package coda (Plummer et al., 2006). Using coda, we also calculated the 95 percent highest posterior density (95 % HPD) interval for estimated  $r$  values. When this interval of  $r$  values did not include 0 (i.e., no correlation) the mean estimated  $r$  value was considered significant.

## 3. Results

### 3.1. Species, MOTUs and G1 morphotypes

We used an integrative molecular and morphological approach to evaluate and identify chlorodielline crab species and MOTUs through combined analyses of COI data and G1 morphology. For most G1 morphotypes, specimens displayed no notable morphological variation. Within the species *Cyc. granulatus*, *Chl. nigra*, and *Chl. cytherea*, however, slight variation in G1 morphology was evident, but this was not

correlated with geography or phylogeny (MOTUs). For example, the *Chl. cytherea* morphotype showed subtle variation in the degree of rotation of its hooked tip, but it was always hooked versus spatulate in the *Chl. laevisissima* morphotype (compare Fig. 2 C-D with F).

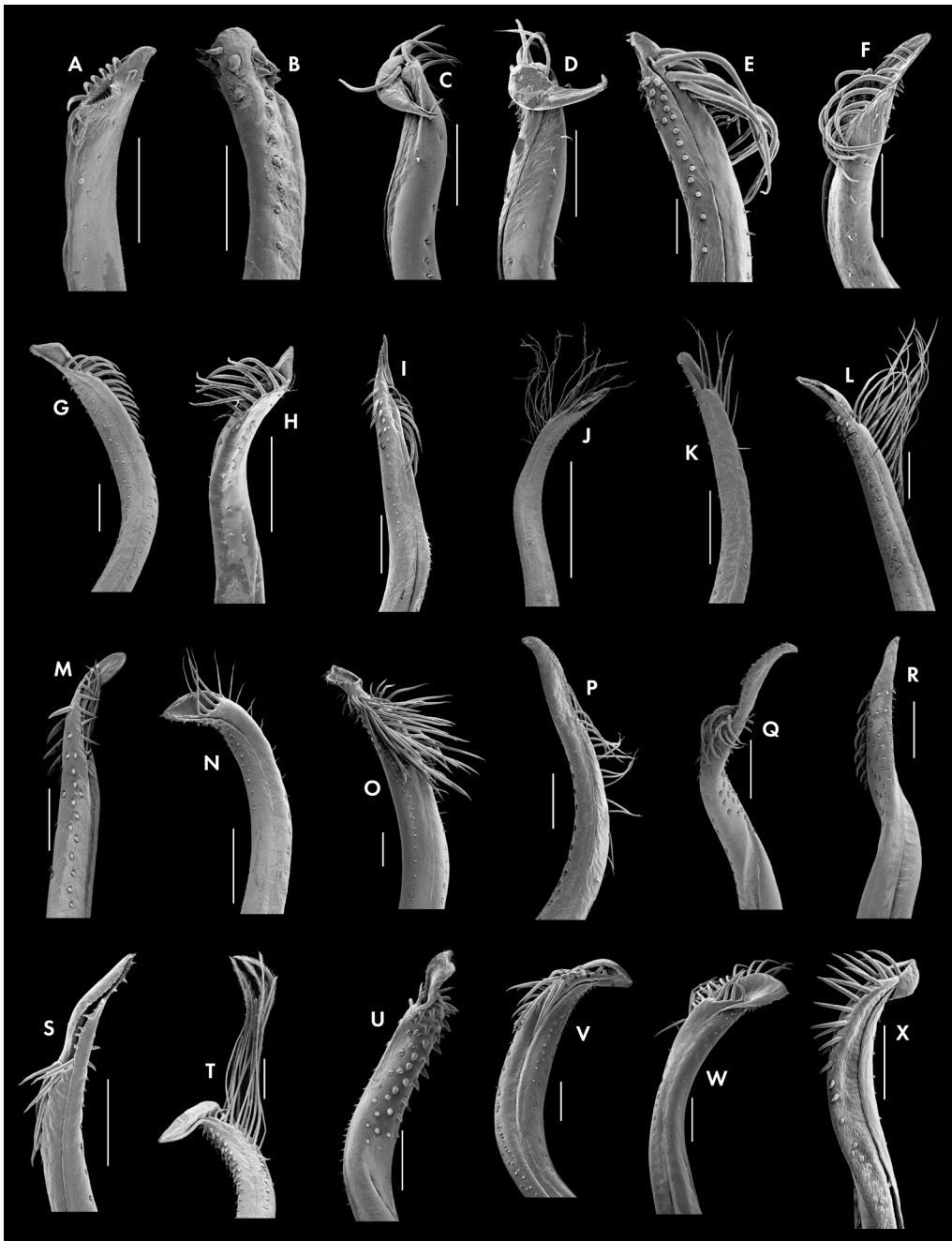
Phylogenetic analyses of COI sequence data revealed 36 reciprocally monophyletic, well-supported ( $BS \geq 70$  and  $pP \geq 0.95$ ) terminal clades (15 *Chlorodiella*, 8 *Cyclodius*, 4 *Luniella*, 7 *Pilodius*, and 2 *Soliella*). Of these 36 clades, 35 were designated as MOTUs using the aforementioned criteria. For the remaining clade, comprising *Chlorodiella nigra* and *Chl. xishaensis*, a total of three MOTUs were designated even though species delineation models did not recover these consistently and each was not well supported. *Chlorodiella nigra* and *Chl. xishaensis* are reliably separated by distinct G1s and external morphology, but the well-supported *Chl. xishaensis* clade/MOTU was sometimes recovered within two *Chl. nigra* clade/MOTUs, albeit with low support (see Discussion, and Fig. 3). Analyses of our four locus, concatenated dataset (discussed below) further highlights the unstable phylogenetic relationship among these two nominal taxa, but provide additional support for designating three MOTUs for the lineage. Including the two *Chl. nigra* MOTUs, the *Chl. xishaensis* MOTU, and the 35 others, we designated a total of 38 chlorodielline MOTUs for analyses (Table 1).

Among these 38 MOTUs, we identified 27 distinct G1 morphotypes. Twenty of these morphotypes were designated “unique”, each corresponding to only one MOTU (Table 1, Figs. 2-3). However, two of these unique G1 morphotypes include multiple nominal species within their respective single MOTUs: 1) *Pilodius pilumnoides*, *P. concors* and *P. cephalalgicus*; and 2) *Pilodius granulatus* and *P. philippinensis*. Each of the remaining seven G1 morphotypes were designated as “shared” as they are present in multiple closely related (typically sister) MOTUs, some of which include multiple nominal species. These seven shared G1 morphotypes include the following taxa: 1) *Chl. nigra* and *Chl. spinimera* (2 MOTUs, polyphyletic); 2) *Chl. cytherea*, *Chl. davaoensis*, and *Chl. crispipleopa* (5 MOTUs, polyphyletic); 3) *Chl. barbata* (2 MOTUs); 4) *Chl. sp. 1* (3 MOTUs); 5) *Chl. sp. 2* (2 MOTUs); 6) *Cyc. nitidus* (2 MOTUs); and 7) *Cyc. obscurus* (2 MOTUs) (18 total MOTUs designated with “shared” G1 morphotypes; Table 1, Fig. 2). These findings, combined with examination of external morphology, suggest that the six following nominal species are junior synonyms and will be dealt with in a future revision of the group (Lasley, in prep): *Chl. spinimera*, *P. concors*, *P. cephalalgicus*, *P. philippinensis*, *Chl. davaoensis*, and *Chl. crispipleopa*. Furthermore, for each MOTU with a shared G1 morphotype, we failed to identify any distinct, diagnostic morphology that would distinguish them from any other conspecific MOTUs, and thus assigning species-level status to any of these MOTUs is premature. However, two novel supra-MOTU clades with distinct G1s were supported in the COI-only and four locus analyses, listed above as *Chl. sp. 1* and *Chl. sp. 2*. *Chlorodiella sp. 1* and *Chl. sp. 2* are composed of three and two MOTUs, respectively.

### 3.2. Geographic overlap and species richness

Geographic overlap for each of the 38 chlorodielline MOTUs is listed in Table 1 and mapped on our summary phylogeny (Fig. 1). Among the 20 chlorodielline MOTUs with unique G1 morphotypes all but two display a sympatric distribution relative to their sister species or lineage. *Pilodius moranti* and *P. areolatus*, likewise possessing unique G1 morphotypes, are narrowly sympatric. Of the remaining 18 MOTUs comprising all “shared” G1 morphotypes, 10 are allopatric with their sister MOTUs, 7 are narrowly sympatric with their sister MOTUs, and 1 has a sympatric relationship with its sister MOTU lineage (Table 1, Fig. 3).

Species and MOTU richness was highest in the Indo-Australian Archipelago (IAA) and Great Barrier Reef (GBR) (Fig. 4), with a secondary diversity peak in the western Indian Ocean. A similar pattern was found in the literature-only heatmap (LO) (SM6), with a couple of notable exceptions—e.g., 1) the main peak of species richness in IAA and GBR is



**Fig. 2.** First male gonopods of chlorodielline species with size of scale bar in parentheses: A, *Chlorodiella barbata* (Borradaile, 1900) (200um); B, *Chlorodiella cochlearis* (Zehntner, 1894) (100um); C-D, *Chlorodiella cytherea* (Dana, 1852) (200um); E, *Chlorodiella laevis* (Dana, 1852) (100um); F, *Chlorodiella* sp. 1. (Krauss, 1843) (200um); G, *Chlorodiella nigra* (Forskål, 1775) (200); H, *Chlorodiella* sp. 2 (200um); I, *Chlorodiella xishaensis* Chen & Lan, 1978 (200um); J, *Cyclodius drachi* (Guinot, 1964) (1 mm); K, *Cyclodius granulatus* De Man, 1888 (200um); L, *Cyclodius nitidus* (Dana, 1852) (200um); M, *Cyclodius obscurus* (Hombron & Jacquinet, 1846) (200um); N, *Cyclodius paumotensis* (Rathbun, 1907) (200um); O, *Cyclodius unguatus* (H. Milne Edwards, 1834) (200um); P, *Luniella pubescens* (Dana, 1852) (200um); Q, *Luniella pugil* (Dana, 1852) (200um); R, *Luniella scabricula* (Dana, 1852) (200um); S, *Soliella melanospinis* (Rathbun, 1911) (200um); T, *Pilodius areolatus* (H. Milne Edwards, 1834) (200um); U, *Pilodius granulatus* Stimpson, 1859 (200um); V, *Pilodius maotieni* Serène, 1971 (200um); W, *Pilodius miersi* (Ward, 1936) (200um); X, *Pilodius nigrocrinitus* Stimpson, 1859 (200um).

patchy in the LO heatmap (versus more continuous throughout the region in the combined heatmap; and 2) there is a peak in diversity in French Polynesia in the LO heatmap (versus less distinct in the combined heatmap).

### 3.3. Multi-locus phylogenetic analyses

BI and ML analyses of the concatenated, partitioned, four locus dataset all recovered congruent topologies except where support values were low (Fig. 1, SM7-11). For BI analyses no significant topological or clade support differences were present between the MRC and MCC topologies. Monophyly for each genus (*Chlorodiella*, *Cyclodius*, *Luniella*, *Pilodius* and *Soliella*) was well supported in all analyses. Species-level lineage relationships were consistent with COI-only analyses (discussed above), with one exception. In both concatenated and H3-only ML and BI analyses, *Chlorodiella xishaensis* was recovered as a longer-

branched taxon that was poorly supported as derived within a polyphyletic *Chl. nigra* MOTU 1 + MOTU 2 clade (H3 results not shown). However, *Chl. nigra* was monophyletic in the COI-, 16S-, and 12S-only analyses (16S and 12S results also not shown). Yet subsequent ML analyses of both concatenated data sets, composed of 100 independent searches, resulted in a different, congruent, and only slightly worse scoring “best” topology in which a monophyletic *Chl. nigra* was recovered sister to a shorter-branched *Chl. xishaensis* (Figs. SM11-13). To explore if these competing ML topologies were significantly worse than the first set, we performed a SH-test in RAxML. For both the 79 and 39 taxa concatenated data sets, InL values for the monophyletic *Chl. nigra* ML topologies (-15776.866126 and -14478.420664, respectively) were not significantly different ( $p > 0.05$ ) than the better scoring topologies with a derived placement of *Chl. xishaensis* (InL = -15773.535752 and -14475.674861, respectively). Given overwhelming morphological similarity between *Chl. nigra*'s MOTUs (discussed above) and our SH-test

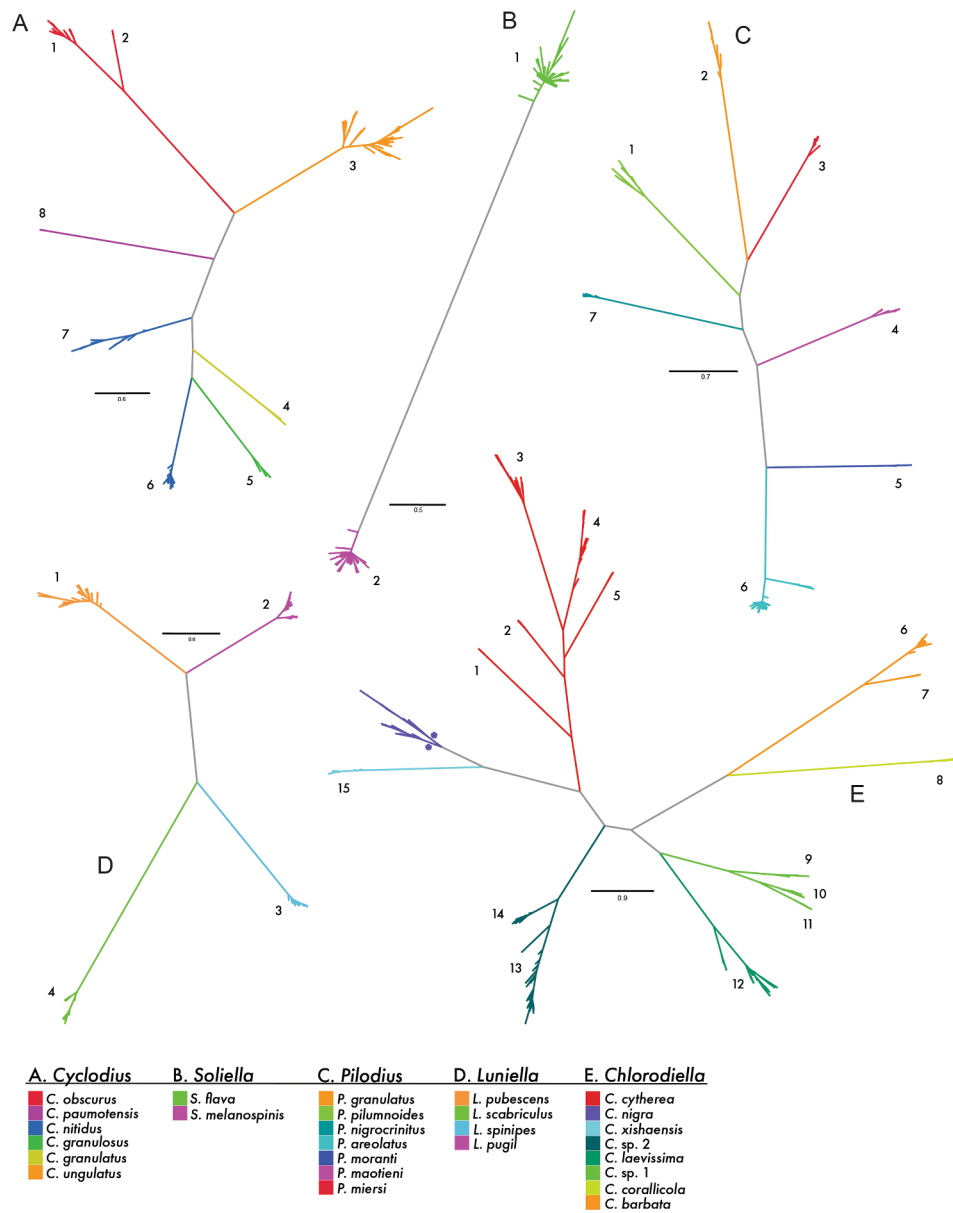


Fig. 3. Unrooted maximum likelihood tree for each chlorodielline genus from COI sequences. An asterisk indicates the clade was not recovered with high support (PP > 0.95) in the Bayesian analysis. Colors represent G1 morphotypes.

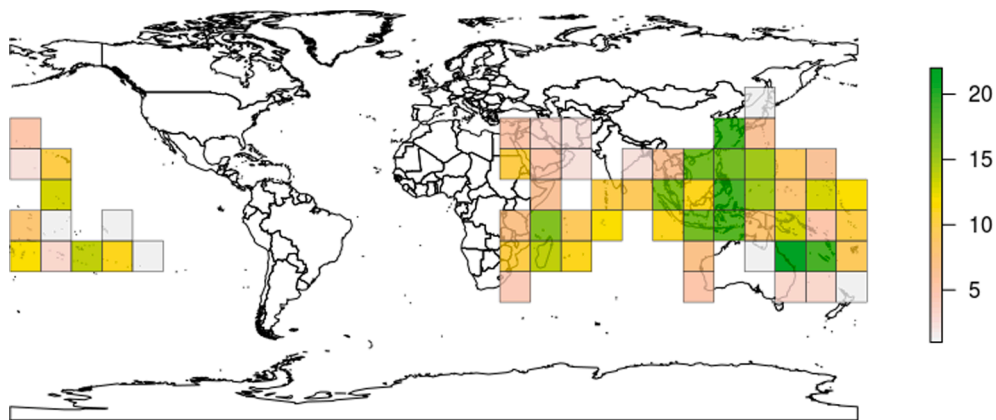


Fig. 4. Heatmap of chlorodielline species richness from literature and records in this study. Green represents highest diversity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



results, our summary topology (Figs. 1, 3) and subsequent comparative analyses used the slightly worse scoring ML topology in which *Chl. nigra* was monophyletic.

### 3.4. Evolutionary covariance of geography, genitalia, and genetic divergence

Results of our five threshBayes analyses are summarized in SM13. Posterior density plots of estimated  $r$  values appear in SM14. For each MCMC run ESS values were above 200 (range: 795.5–1997) and log-likelihood plots indicated all runs quickly reached stability (not shown). Four analyses recovered the following significant correlations (i.e., 95 % HPD intervals did not include 0): a moderate positive correlation ( $r = 0.6731$ ) was found between G1 uniqueness and genetic distance (i.e., shortest interclade genetic distance); a positive correlation between G1 uniqueness and geographic overlap was estimated to be strong ( $r = 0.7191$ ) or moderate ( $r = 0.634$ ) depending if “narrowly sympatric” MOTUs were coded, respectively, as “allopatric” or “sympatric”; and a moderate positive correlation ( $r = 0.6857$ ) was found between genetic distance and geographic overlap when “narrowly sympatric” MOTUs were coded as “allopatric”. In the fifth analysis, a weak positive correlation ( $r = 0.3624$ ) was found between genetic distance and geographic overlap when “narrowly sympatric” MOTUs were coded as “sympatric”. However, for this final analysis, the 95 % HPD interval of estimated  $r$  values included zero and thus should not be considered a significant result.

## 4. Discussion

We have demonstrated a striking, well-supported relationship between genetic distance (time), sympatry, and the divergence of genitalic morphology in chlorodielline crabs. To our knowledge, the relationship between genitalic divergence and biogeography has not been investigated in marine decapod crustaceans, and only rarely in marine invertebrates. These results suggest that differentiation is initiated in allopatry without genitalic differentiation, but sympatric distributions are strongly tied to genitalic differentiation. However, our results cannot reveal whether morphological divergence of genitalia was initiated in allopatry or during the establishment of secondary sympatry. MOTUs with identical gonopods can diverge > 9 % in COI in allopatry (Table 1). This likely represents millions of years of divergence in isolation despite the ability of MOTUs to disperse into the ranges of neighboring MOTU populations, as evidenced by (1) the IWP-wide ranges and ubiquity of most chlorodielline species that suggest high long-term dispersal ability, (2) relative genetic homogeneity across the distribution of MOTUs, and (3) evidence for low-levels of gene flow in other systems with similar allopatric mosaic distributions (Meyer et al. 2005). Millions of years combined with infrequent gene flow is likely sufficient to allow isolating mechanisms to evolve due to genetic drift or local adaptation, which would continue to decrease homogenizing gene flow (Coyne & Orr, 2004). Our data are insufficient to test whether genitalic divergence occurred in allopatry (e.g. through sexual selection) that then allowed establishment of secondary sympatry, or evolved after secondary sympatry (e.g., through reproductive character displacement). Given the complexity of this topic (e.g., Templeton, 1981, Noor, 1999, Coyne & Orr, 2004, Arnqvist & Rowe, 2005, Eberhard, 2010, Servedio & Boughman, 2017) and an absence of additional lines of evidence, we refrain from speculating about the exact nature and timing of these isolating mechanisms. Nevertheless, our results strongly suggest that gonopod (genitalic) divergence in chlorodielline crabs occurred after significant genetic isolation was established, and that distinct gonopod morphotypes appear in secondary sympatry. Pinpointing additional isolating mechanisms involved in initiating or completing the process of speciation—and the spatial and temporal aspects of gonopod divergence in this process—will require future research. However, the strong correlation between genetic distance (time), sympatry, and the divergence

of gonopods suggests that genitalic morphological evolution plays a significant role in completing the process of speciation and attaining sympatry in chlorodielline crabs, and likely for many other marine taxa as well.

With the potential exception of *Chlorodiella nigra*, all species-level clades recovered here and recognized in the literature and two additional, undescribed taxa, *Chl. sp. 1* and *Chl. sp. 2*, fulfill the requirements of the biological (Mayr, 1969, Dobzhansky, 1970, Coyne & Orr, 2004), phylogenetic (Rosen, 1978, 1979, McKittrick and Zink, 1988, Cracraft, 1983), Hennigian (Hennig, 1966), and morphological (Cronquist 1978) (also see Wheeler & Meier, 2000) species concepts. Some of these requirements include: 1) reciprocal monophyly, 2) reproductive isolation implied by divergent G1 morphology, and 3) and unique character states or diagnosability. MOTUs with significant genetic divergence but without divergence in G1 morphology, however, do not satisfy all these requirements. While some studies have recognized or described new species based on COI divergence alone, usually in conjunction with geographic information, we refrain from doing so here. It is clear that, even with substantial divergence in COI, isolating mechanisms may not have evolved sufficiently to provide reproductive isolation. Furthermore, the conflicting results of species delineation methods and the various genetic distances of subclades across the chlorodielline phylogeny indicate there is no single threshold of COI divergence for species assignment.

Species richness of chlorodielline crabs is highest in the IAA and adjacent areas, with secondary diversity peaks in the Western Indian Ocean (Seychelles and Madagascar) and southeastern Polynesia (Fig. 4). The first two diversity centers are typical of reef-associated organisms (Rosen, 1981, Briggs, 2000, 2005, Roberts et al., 2002, Mora et al., 2003). High species richness in southeastern Polynesia is less common but has been observed in other reef cryptobiota like *Calcinus* hermit crabs (Malay & Paulay, 2010), although this may partly reflect intensive sampling of this area.

Range boundaries among MOTUs can be considered evolutionary significant events (ESEs)—i.e., differentiation that potentially gives rise to full species (Malay & Paulay, 2010). The spatial and temporal distribution of ESEs inform on the importance of different areas and times in driving diversification. In our analyses, ESEs occur in peripheral areas of the IWP like the Red Sea, Hawaii and/or Line Islands, and the Southern Pacific Ocean, while several other ESEs overlap or abut in the IAA and adjacent areas (Fig. 5, SM15-17). Given the scattered distribution of ESEs, chlorodielline diversification appears to show no clear adherence to any proposed hypothesis explaining high diversity in the IAA—i.e., “Center of Origin”, “Center of Overlap”, “Center of Accumulation”, etc. (e.g., Briggs, 1992, Santini & Winterbottom, 2002, Mora et al., 2003, Barber & Bellwood, 2005, Carpenter & Springer, 2005, Malay & Paulay, 2010, Bellwood & Meyer, 2009, Hubert et al., 2012).

### CRediT authorship contribution statement

**Robert M. Lasley:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Nathaniel Evans:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization. **Gustav Paulay:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Francois Michonneau:** Software, Resources, Data curation, Visualization. **Amanda Windsor:** Resources, Writing – review & editing, Supervision. **Irwanyah:** Resources, Supervision. **Peter K.L. Ng:** Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

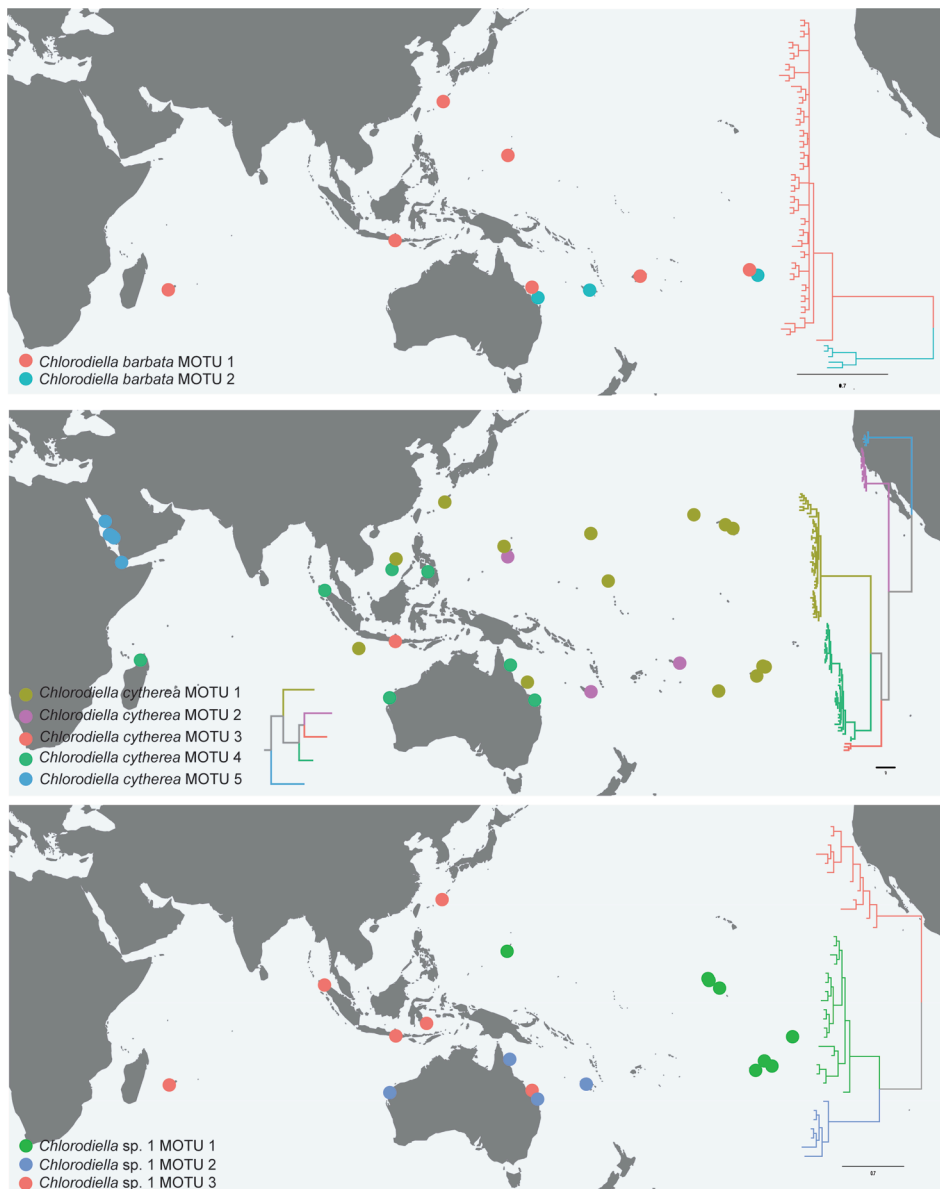


Fig. 5. Geographic distributions of select G1 morphospecies and their MOTUs: *Chlorodiella barbata*, *C. cytherea*, and *C. sp. 1*.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgements

We are grateful to numerous people that have aided in sequencing and other lab work, and the many more people that have helped in the field or provided specimens. We also thank Mandy Bemis and John Slapcinsky (UF) for their efforts in curation. This work was supported by the National Science Foundation [NSF DEB 1856245, GECCO 1457769], the Moore Foundation [Moorea BioCode], and the Sloan Foundation [MarBOL]. NE was partially supported by a Japan Society for the Promotion of Science (JSPS) post-doctoral fellowship (P22078)

#### Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2023.107710>.

#### References

- Arnqvist, G., Rowe, L., 1995. Sexual conflict and arms races between the sexes: a morphological adaptation for control of mating in a female insect. *Proc. R. Soc. Lond. B* 261, 123–127.
- Arnqvist, G., Rowe, L., 2005. *Sexual conflict*. Princeton University Press, Princeton, NJ.
- Barber, P.H., Bellwood, D.R., 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. *Mol. Phylogenet. Evol.* 35 (1), 235–253.
- Bellwood, D.R., Meyer, C.P., 2009. Endemism and evolution in the Coral Triangle: a call for clarity. *J. Biogeogr.* 36 (10), 2010–2012.
- Briceno, R.D., Eberhard, W.G., 2009. Experimental modifications imply a stimulatory function for male tsetse fly genitalia, supporting cryptic female choice theory. *J. Evol. Biol.* 22, 1516–1525.
- Briggs, J.C., 1992. The marine East Indies: centre of origin? *Glob. Ecol. Biogeogr.* 2, 149–156.
- Briggs, J.C., 2000. Centrifugal speciation and centres of origin. *J. Biogeogr.* 27 (5), 1183–1188.

- Briggs, J.C., 2005. The biogeography of otophysan fishes (Ostariophysi: Otophysi): a new appraisal. *J. Biogeogr.* 32 (2), 287–294.
- Briggs, J.C., Bowen, B.W., 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. *J. Biogeogr.* 39 (1), 12–30.
- Carpenter, K.E., Springer, V.G., 2005. The center of the center of marine shore fish biodiversity: the Philippine Islands. *Environ. Biol. Fishes* 72 (4), 467–480.
- Clark, P.F., Galil, B.S., 1993. A revision of the xanthid genus *Pilodius* Dana, 1851 (Crustacea: Brachyura: Xanthoidea). *J. Nat. Hist.* 27 (5), 1119–1206.
- Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Gray, M.R., 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Aust. J. Zool.* 46 (5), 419–437.
- Collins, J.S., 2014. Middle Miocene “Badenian” (Langhian) decapod crustaceans from the Retzneiquarry, Styrian Basin, Austria. *Bull. Mizunami Fossil Museum* 40, 29–50.
- Coyne, J.A., Orr, H.A., 2004. Speciation, Vol. 37. Sinauer Associates, Sunderland, MA.
- Cracraft, J., 1983. Species concepts and speciation analysis. In: *Current ornithology*. Springer, New York, NY, pp. 159–187.
- Crandall, E.D., Riginos, C., Bird, C.E., Liggins, L., Trembl, E., Beger, M., Gaither, M.R., 2019. The molecular biogeography of the Indo-Pacific: Testing hypotheses with multispecies genetic patterns. *Glob. Ecol. Biogeogr.* 28 (7), 943–960.
- Cronquist, A., 1978. The biota of the Intermountain Region in geohistorical context. *Great Basin Naturalist Memoirs* 3–15.
- Dai, A., Cai, Y., & Yang, S. (1996). New species and new records of crabs (Crustacea; Decapoda: Brachyura) from Nansha Islands, China. *Studies on Marine Fauna and Flora and Biogeography of the Nansha Islands and Neighbouring Waters*, 2, 234–257.
- Dobzhansky, T., 1970. Genetics of the Evolutionary Process. Columbia Univ. Press, New York.
- Drew, J., Barber, P.H., 2009. Sequential cladogenesis of the reef fish *Pomacentrus moluccensis* (Pomacentridae) supports the peripheral origin of marine biodiversity in the Indo-Australian archipelago. *Mol. Phylogenet. Evol.* 53 (1), 335.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7 (1), 1–8.
- Eberhard, W.G., 1996. *Females Control: Sexual Selection by Cryptic Female Choice*. Princeton Univ. Press, Princeton, NJ.
- Eberhard, W.G., 2010. Evolution of genitalia: theories, evidence, and new directions. *Genetica* 138, 5–18.
- Edmondson, C.H., 1962. Xanthidae of Hawaii. *Occasional Papers of Bernice P. Bishop Museum* 22 (13), 215–309.
- Evans, N., 2018. Molecular phylogenetics of swimming crabs (Portunoidea Rafinesque, 1815) supports a revised family-level classification and suggests a single derived origin of symbiotic taxa. *PeerJ* 6, e4260.
- Ezard, T., Fujisawa, T., & Barraclough, T. (2013). R package splits: SPecies' LImits by Threshold Statistics, version 1.0-18/r45. *R package splits: SPecies' LImits by Threshold Statistics, version 1.0-18/r45*.
- Felgenhauer, B.E., 1987. Techniques for preparing crustaceans for scanning electron microscopy. *J. Crustac. Biol.* 7, 71–76.
- Felsenstein, J., 2012. A comparative method for both discrete and continuous characters using the threshold model. *Am. Nat.* 179 (2), 145–156.
- Forest, J. & Guinot, D. (1961) Crustacés Décapodes Brachyours de Tahiti et des Tuamotu. In: *Expédition française sur les récifs coralliens de la Nouvelle-Calédonie. Volume préliminaire*. Editions de la Fondation Singer Polignac, Paris, 9–11, 1–195, pls. 1–18.
- Fraaye, R.H.B., 1996. A new Tethyan migrant: *Cretachlorodius enciensis* n. gen., n.sp. (Crustacea, Decapoda), from the Maastrichtian type area. *J. Paleo.* 70 (2), 293–296.
- Gage, M.J., 1992. Removal of rival sperm during copulation in a beetle, *Tenebrio molitor*. *Anim. Behav.* 44, 587–589.
- Geller, J., Meyer, C., Parker, M., Hawk, H., 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* 13 (5), 851–861.
- Gernhard, T., 2008. The conditioned reconstructed process. *J. Theor. Biol.* 253 (4), 769–778.
- Guinot, D., 1964. Crustacés décapodes brachyours (Xanthidae) des campagnes de la Calypso en Mer Rouge (1952), dans le Golfe Persique et à L'île Aldabra (1954). *Mémoires du Muséum national d'Histoire naturelle. Nouvelle Série. Série A, Zoologie* 32, 1–108.
- Guinot, D., Tavares, M., Castro, P., 2013. Significance of the sexual openings and supplementary structures on the phylogeny of Brachyuran crabs (Crustacea, Decapoda, Brachyura), with new nomina for high-ranked podotreme taxa. *Zootaxa* 3665, 1–414.
- Hennig, W., 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana.
- Hollander, J., Smadja, C.M., Butlin, R.K., Reid, D.G., 2013. Genital divergence in sympatric sister snails. *J. Evol. Biol.* 26 (1), 210–215.
- Hollander, J., Montaña-Rendón, M., Bianco, G., Yang, X., Westram, A.M., Duvaux, L., Butlin, R.K., 2018. Are assortative mating and genital divergence driven by reinforcement? *Evol. Lett.* 2 (6), 557–566.
- Hubert, N., Meyer, C.P., Bruggemann, H.J., Guerin, F., Komeno, R.J., Espiau, B., Planes, S., 2012. Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA barcoding provides new support to the centre-of-overlap hypothesis. *PLoS One* 7 (3), e28987.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780.
- Kay, E. A. (1984). Patterns of speciation in the Indo-west Pacific. Pages 15–31 in P. H. Raven, F. J. Radvosky and S. H. Sohmer, eds. *Bio-geography of the tropical Pacific*. Assoc. of Systematic Coli. and B. P. Bishop Museum. x + 221 pp.
- Kim, J., Sanderson, M.J., 2008. Penalized likelihood phylogenetic inference: bridging the parsimony-likelihood gap. *Syst. Biol.* 57 (5), 665–674.
- Knowlton, N., 1993. Sibling species in the sea. *Annu. Rev. Ecol. Syst.* 24 (1), 189–216.
- Lai, J.C., Ah Yong, S.T., Jeng, M.S., Ng, P.K., 2009. Are coral-dwelling crabs monophyletic? A phylogeny of the Trapezioida (Crustacea: Decapoda: Brachyura). *Invertebr. Syst.* 23 (4), 402–408.
- Lai, J.C., Mendoza, J.C.E., Guinot, D., Clark, P.F., Ng, P.K., 2011. Xanthidae MacLeay, 1838 (Decapoda: Brachyura: Xanthoidea) systematics: a multi-gene approach with support from adult and zoeal morphology. *Zoologischer Anzeiger-A J. Comparative Zool.* 250 (4), 407–448.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34 (3), 772–773.
- Lasley, R.M., Lai, J.C., Thoma, B.P., 2013. A new genus for *Chlorodiella longimana* (H. Milne Edwards) supported by morphology and molecular data, with a preliminary phylogeny of the Chlorodiellinae (Crustacea: Decapoda: Xanthidae). *Invertebr. Syst.* 27 (4), 379–390.
- Lasley Jr, R.M., Klaus, S., Ng, P.K., 2015. Phylogenetic relationships of the ubiquitous coral reef crab subfamily Chlorodiellinae (Decapoda, Brachyura, Xanthidae). *Zool. Scr.* 44 (2), 165–178.
- Lasley Jr, R.M., Smith, A., Paulay, G., Ng, P.K., 2022. Revision of the coral reef crab genus *Tweedia* Ward, 1935 (Decapoda: Brachyura: Xanthidae). *J. Crustac. Biol.* 42 (1), ruab081.
- Leray, M., Boehm, J.T., Mills, S.C., Meyer, C.P., 2012. Moorea BIOCDE barcode library as a tool for understanding predator-prey interactions: insights into the diet of common predatory coral reef fishes. *Coral Reefs* 31 (2), 383–388.
- Malay, M.C.M.D., Paulay, G., 2010. Peripatric speciation drives diversification and distributional pattern of reef hermit crabs (Decapoda: Diogenidae: Calcinus). *Evolution: International Journal of Organic. Evolution* 64 (3), 634–662.
- Mallet, J., Meyer, A., Nosil, P., Feder, J.L., 2009. Space, sympatry and speciation. *J. Evol. Biol.* 22 (11), 2332–2341.
- Mayr, E., 1963. *Animal species and evolution*. Harvard University Press, Cambridge, MA.
- Mayr, E., 1969. The biological meaning of species. *Biol. J. Linn. Soc.* 1 (3), 311–320.
- Mayr, E. (1942). *Systematics and the Origin of Species*, Columbia Univ. Press, New York, (re-issued, 1982).
- McKittrick, M.C., Zink, R.M., 1988. Species concepts in ornithology. *Condor* 90, 1–14.
- Meier, R., Shiyang, K., Vaidya, G., Ng, P.K., 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* 55 (5), 715–728.
- Mendoza, J.C., Chan, K.O., Lai, J.C., Thoma, B.P., Clark, P.F., Guinot, D., Ng, P.K., 2022. A comprehensive molecular phylogeny of the brachyuran crab superfamily Xanthoidea provides novel insights into its systematics and evolutionary history. *Mol. Phylogenet. Evol.* 177, 107627.
- Meyer, C.P., Geller, J.B., Paulay, G., 2005. Fine scale endemism on coral reefs: archipelagic differentiation in turbinid gastropods. *Evolution* 59 (1), 113–125.
- Miller, M.A., Pfeiffer, W., Schwartz, T., 2011. July). The CIPRES science gateway: a community resource for phylogenetic analyses. In: *Proceedings of the 2011 TeraGrid Conference: extreme digital discovery*, pp. 1–8.
- Miyake, S., Takeda, M., 1967. On some rare xanthid crabs from the Ryukyu Islands, with description of a new species. *J. Fac. Agric. Kyushu Univ.* 14 (2), 293–302.
- Mokady, O., Graur, S.R.D., 1994. Coral-host specificity of Red Sea Lithopagis bivalves: interspecific and intraspecific variation in 12S mitochondrial. *Mol. Mar. Biol. Biotech.* 3 (3), 158–164.
- Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J., Vogler, A.P., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst. Biol.* 58 (3), 298–311.
- Mora, C., Chittaro, P.M., Sale, P.F., Kritzer, J.P., Ludsin, S.A., 2003. Patterns and processes in reef fish diversity. *Nature* 421, 933–936.
- Müller, P., 1984. Decapod Crustacea from the Badenian. *Geologica Hungarica, Series Palaeontologica* 42, 1–317.
- Müller, P., Collins, J.S.H., 1991. Late Eocene coral-associated decapods (Crustacea) from Hungary. *Mededelingen van de Werkgroep voor Tertiaire en Kwartaire Geologie* 28 (2/3), 47–92.
- Myers, A.A., 1994. Biogeographic patterns in shallow-water marine systems and the controlling process at different scales. In: Giller, P.S., Hildrew, A.G., Raffaelli, D.G. (Eds.), *Aquatic Ecology, Scale, Pattern and Process*. British Ecological Society Symposium, Blackwell Scientific, Oxford, pp. 547–574.
- Nobili, G., 1905. *Decapodes nouveaux des cotes d'Arabie et du Golfe Persique*. (Diagnoses préliminaires). *Bulletin du Muséum national d'histoire naturelle* 11, 158–164.
- Noor, M.A.F., 1999. Reinforcement and other consequences of sympatry. *Heredity* 83, 503–508.
- Nosil, P., 2008. Speciation with gene flow could be common. *Mol. Ecol.* 17 (9), 2103–2106.
- Odhner, T. (1925). *Monographierte Gattungen der Krabbenfamilie Xanthidae. I. Goteborgs Kungliga Vetenskaps- och VitterhetsSsamhallets Handlingar*, 29, 3–92, pls 1–5.
- Palumbi, S.R., 1992. Marine speciation on a small planet. *Trends Ecol. Evol.* 7 (4), 114–118.
- Paradis, E., 2013. Molecular dating of phylogenies by likelihood methods: a comparison of models and a new information criterion. *Mol. Phylogenet. Evol.* 67, 436–444.
- Peyrot-Clausade, M., 1977. *Decapodes brachyours et anomours (à l'exception des Paguridae) de la cryptofaune de Tiahura Ile de Moorea*. *Extrait des Cahiers du Pacifique* n° 20, 211–222.
- Peyrot-Clausade (1979). *Cryptofaune annelidienne et carcinologique des récifs de l'île de la Réunion et de l'île Maurice*. *The Mauritius Institute Bulletin*, 8, 1–41.
- Plaisance, L., Caley, M.J., Brainard, R.E., Knowlton, N., 2011. The diversity of coral reefs: what are we missing? *PLoS One* 6 (10), e25026.

- Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. *R news* 6 (1), 7–11.
- Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55 (4), 595–609.
- Posada, D., 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25 (7), 1253–1256.
- Puillandre, N., Lambert, A., Brouillet, S., Achaz, G.J.M.E., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21 (8), 1864–1877.
- Rambaut, A. & Drummond, A. J. (2009). Tracer v1.5. Available at <http://beast.bio.ed.ac.uk/Tracer>.
- Reginato, M., 2016. MonographR: an R package to facilitate the production of plant taxonomic monographs. *Brittonia* 68 (2), 212–216.
- Revell, L.J., 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223.
- Revell, L.J., 2014. Ancestral character estimation under the threshold model from quantitative genetics. *Evolution* 68, 743–759.
- Roberts, C.M., McClean, C.J., Veron, J.E., Hawkins, J.P., Allen, G.R., McAllister, D.E., Werner, T.B., 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295 (5558), 1280–1284.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19 (12), 1572–1574.
- Rosen, D.E., 1978. Vicariant patterns and historical explanation in biogeography. *Syst. Zool.* 27, 159–188.
- Rosen, D.E., 1979. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative biogeography. *Bull. Am. Mus. Nat. Hist.* 162, 267–376.
- Rosen, B. R. (1981). The tropical high diversity enigma—the corals' eye view. Pp. 103–129. In Forey, P. L., ed. *The Evolving Biosphere*. Brit. Mus. (Nat. Hist.) and Cambridge Univ. Press; Cambridge and London.
- Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109.
- Santini, F., Winterbottom, R., 2002. Historical biogeography of Indo-western Pacific coral reef biota: is the Indonesian region a centre of origin? *J. Biogeogr.* 29 (2), 189–205.
- Sela, I., Ashkenazy, H., Katoh, K., Pupko, T., 2015. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* 43 (W1), W7–W14.
- Serène, R., 1984. Crustacés Décapodes Brachyours de l'océan Indien occidental et de la mer Rouge, Xanthoidea: Xanthidae et Trapeziidae. Avec un addendum par Crosnier A.: Carpillidae et Menippidae. *Faune tropicale* 24, 1–243.
- Servedio, M.R., Boughman, J.W., 2017. The role of sexual selection in local adaptation and speciation. *Annu. Rev. Ecol. Evol. Syst.* 48, 85–109.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16 (8), 1114.
- Shull, H.C., Pérez-Losada, M., Blair, D., Sewell, K., Sinclair, E.A., Lawler, S., Crandall, K. A., 2005. Phylogeny and biogeography of the freshwater crayfish *Euastacus* (Decapoda: Parastacidae) based on nuclear and mitochondrial DNA. *Mol. Phylogenet. Evol.* 37 (1), 249–263.
- Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30 (9), 1312–1313.
- Suchard, M.A., Rambaut, A., 2009. Many-core algorithms for statistical phylogenetics. *Bioinformatics* 25 (11), 1370–1376.
- Templeton, A.R., 1981. Mechanisms of speciation—a population genetic approach. *Annu. Rev. Ecol. Syst.* 12, 23–48.
- Thornhill, R., 1983. Cryptic female choice and its implications in the scorpionfly *Harpobittacus nigriceps*. *Am. Nat.* 122, 765–788.
- Titus, B.M., Daly, M., Hamilton, N., Berumen, M.L., Baeza, J.A., 2018. Global species delimitation and phylogeography of the circumtropical 'sexy shrimp' *Thor amboinensis* reveals a cryptic species complex and secondary contact in the Indo-West Pacific. *J. Biogeogr.* 45 (6), 1275–1287.
- Waage, J.K., 1979. Dual function of the damselfly penis: sperm removal and transfer. *Science* 203, 916–918.
- Wang, X., He, Z., Shi, S., Wu, C.L., 2020. Genes and speciation: is it time to abandon the biological species concept? *Natl. Sci. Rev.* 7 (8), 1387–1397.
- Wheeler, Q.D., Meier, R. (Eds.), 2000. *Species concepts and phylogenetic theory: a debate*. Columbia University Press.
- Wickham, H., Chang, W., Wickham, M.H., 2016. Package 'ggplot2'. Create elegant data visualizations using the grammar of graphics. Version 2 (1), 1–189.
- Yao, F., Shi, B., Wang, X., Pan, D., Bai, M., Yan, J., Sun, H., 2020. Rapid divergent coevolution of *Sinopotamon* freshwater crab genitalia facilitates a burst of species diversification. *Integrative zoology* 15 (3), 174–186.
- Yule, G. U. (1925). II.—A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, FR S. *Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character*, 213(402-410), 21-87.
- Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29 (22), 2869–2876.