2017

Phylogeography and Cryptic Diversity of Intertidal *Ligia* Isopods (Crustacea, Isopoda, Ligiidae) Across the Southern Africa Coastline

Taylor M. Greenan
University of South Florida

Charles L. Griffiths
University of Cape Town

Carlos A. Santamaria
University of South Florida, csantamaria@usf.edu

A peer-reviewed version of this preprint was published in PeerJ on 19 April 2018.

Greenan TM, Griffiths CL, Santamaria CA. 2018. Molecular approaches uncover cryptic diversity in intertidal *Ligia* isopods (Crustacea, Isopoda, Ligiidae) across the southern Africa coastline. PeerJ 6:e4658 https://doi.org/10.7717/peerj.4658

Phylogeography and cryptic diversity of intertidal *Ligia* isopods (Crustacea, Isopoda, Ligiidae) across the southern Africa coastline

Taylor M Greenan [1], **Charles L Griffiths** [2], **Carlos A Santamaria** [Corresp. 1, 3]

[1] Biology Program, College of Science and Mathematics, University of South Florida Sarasota-Manatee, Sarasota, Florida, United States

[2] Department of Biological Sciences and Marine Research Institute, University of Cape Town, Rodenbosch, South Africa

[3] Department of Biological Sciences, Sam Houston State University, Huntsville, Texas, United States

Corresponding Author: Carlos A Santamaria
Email address: csantamaria@sar.usf.edu

The extensive coastlines of South Africa and Namibia extends from the Atlantic to the Indian Ocean and encompass several major biogeographic provinces, each characterized by unique faunal and floral assemblages. Recent biogeographic studies have led to competing biogeographic models of the southern African coastline. This has stimulated phylogeographic work to determine whether the distribution of genetic diversity within coastal invertebrate species match the proposed biogeographic regions. The lack of congruence between studies and the discovery of cryptic diversity indicating the possible existence of cryptic species in coastal isopods in the region underscore the need for additional phylogeographic research in southern Africa, particularly for organisms that have been shown to both harbor cryptic diversity and to retain signatures of past geological and oceanographic processes in their phylogeographic patterns. Isopods in the genus *Ligia* exhibit several biological traits that suggest they may be informative on phylogeographic patterns. They inhabit patchy rocky beaches, are direct developers, avoid the open water, and exhibit several biological traits that severely constrain their dispersal potential (e.g. poor desiccation resistance). These traits are thought to lead to long term isolation of populations, the retention of geological and oceanographic signatures in phylogeographic patterns of *Ligia*, and the presence of cryptic lineages. In this study, we used mitochondrial and nuclear markers to characterize *Ligia* collected in 18 localities across Namibia to the KwaZulu-Natal region of South Africa. We report the presence of cryptic lineages within *Ligia* species in the region, as well as distributional patterns that differ from those reported from other coastal taxa, but that broadly matches a widely used biogeographic model for the region.

Phylogeography and cryptic diversity of intertidal *Ligia* isopods (Crustacea, Isopoda, Ligiidae) across the southern Africa coastline

Taylor M. Greenan[1], Charles L. Griffiths[2], Carlos A. Santamaria[1,3]

[1:] Biology Program, College of Science and Mathematics, University of South Florida Sarasota-Manatee, Sarasota, Florida, United States of America

[2:] Department of Biological Sciences and Marine Research Institute, University of Cape Town, Rondebosch, South Africa

[3]: Department of Biological Sciences, Sam Houston State University, Huntsville, Texas, United States of America

Corresponding author:

Carlos A Santamaria

Email: santamaria.carlos.a@gmail.com

ABSTRACT

The extensive coastlines of South Africa and Namibia extends from the Atlantic to the Indian Ocean and encompass several major biogeographic provinces, each characterized by unique faunal and floral assemblages. Recent biogeographic studies have led to competing biogeographic models of the southern African coastline. This has stimulated phylogeographic work to determine whether the distribution of genetic diversity within coastal invertebrate species match the proposed biogeographic regions. The lack of congruence between studies and the discovery of cryptic diversity indicating the possible existence of cryptic species in coastal isopods in the region underscore the need for additional phylogeographic research in southern Africa. This is particularly true for organisms shown to both harbor cryptic diversity and to retain signatures of past geological and oceanographic processes in their phylogeographic patterns. Isopods in the genus *Ligia* exhibit several biological traits that suggest they may be informative on phylogeographic patterns. They inhabit patchy rocky beaches, are direct developers, avoid the open water, and exhibit several biological traits that severely constrain their dispersal potential (e.g. poor desiccation resistance). These traits are thought to lead to long term isolation of populations, the retention of geological and oceanographic signatures in phylogeographic patterns of *Ligia*, and the presence of cryptic lineages. In this study, we used mitochondrial and nuclear markers to characterize *Ligia* collected in 18 localities across Namibia to the KwaZulu-Natal region of South Africa. We report the presence of cryptic lineages within *Ligia* species in the region, as well as distributional patterns that differ from those reported from other coastal taxa, but that broadly match a widely used biogeographic model for the region.

INTRODUCTION

The coastlines of Namibia and South Africa together extend for over 4,700 km and incorporate a wide diversity of habitats across both the Atlantic and Indian Oceans. Namibia and the western coastline of South Africa are washed by the Benguela Current, which brings cool (10–18ºC), low salinity, and slow-moving (0.1–0.3 m s^{-1}) waters from the polar region in the Atlantic Ocean and transports these northwards towards the equator (Demarcq et al. 2003; Hutchings et al. 2009; Shannon & Nelson 1996). Close to shore, offshore winds and the flow of the Benguela Current causes upwelling and thus elevated biological productivity that supports high levels of biomass and rich commercial fisheries along both the South African and Namibian coastline (Crawford et al. 1987). In contrast, along the east coast of South Africa, the Agulhas Current brings warm (20–28ºC), nutrient-poor, fast-moving (up to 2 m s^{-1}) waters that flow southward from the tropical Indian Ocean (Lutjeharms 1998; Schumann 1987). The effects of the current are felt most intensely in northern KwaZulu-Natal, where it flows closest to the shoreline, before deflecting offshore and following the edge of the Agulhas Bank in the area south of East London. Along the south coast of South Africa, the coastlines between East London and Cape Point exhibit intermediate ranges of abiotic factors and support a warm-temperate fauna, rich in endemic species (Griffiths et al. 2010). These contrasting conditions result in the presence of very distinct faunal and floristic assemblages (or bioregions) occurring along the coastline. Many studies have analyzed biogeographic zonation patterns around the coastline and have recognized distinct coastal biogeographical provinces, but with some discrepancies with regard to the numbers of such provinces, their nomenclature, their exact boundaries and the recognition (or not) of ‘overlap zones’ (for brief historical reviews see Griffiths et al. 2010; Teske et al. 2011). The most recent and widely used biogeographic analysis (Sink et al. 2012), however,

proposes four main South African 'Ecoregions' the inshore and offshore boundaries of which differ. As regards the coastline itself, the recognized zones are: the Namaqua region, extending from mid-Namibia to Cape Point in South Africa; the Agulhas region from Cape Point to just east of East London; the Natal region from East London to northern KwaZulu-Natal and the Delagoa region, which stretches from northern KwaZulu-Natal across the border into Mozambique. An additional bioregion occurs in northern Namibia and is known as the Namib region. A similar model was developed by Lombard et al. (2004) and also used by Griffiths et al. (2010), but it differs by recognizing an additional South-Western Cape 'overlap' region between the Namaqua and Agulhas regions, making five distinct bioregions in all. Other similar models that differ mostly by recognizing additional overlap zones, also exist (see Teske et al. 2006). One positive aspect of these competing biogeographic models of the South Africa coastline has been to stimulate recent research focusing on whether the distributional patterns of genetic variance within coastal organisms match the proposed biogeographic regions (Baldanzi et al. 2016; Evans et al. 2004; Ridgway et al. 2001; Teske et al. 2006; Teske et al. 2007; Zardi et al. 2007). Incongruent results between studies; however, underscore the need for additional research that may help further our understanding of coastal processes and their role in driving diversification along southern African shores.

Mitochondrial markers have recently been used to evaluate whether genetic diversity within South African coastal invertebrate species is distributed according to the proposed bioregions (i.e., is genetic variance partitioned along proposed biogeographic breaks). Teske et al. (2006) studied the phylogeographic patterns for three coastal crustaceans (*Exosphaeroma hylecoetes*, *Iphinoe truncata,* and *Upogebia africana*) and reported not only the presence of multiple, deeply-divergent lineages within each of these species, but also a lack of

correspondence in the geographic distributional breaks between the species. More recently, Baldanzi et al. (2016) reported the presence of multiple evolutionary lineages within another coastal crustacean, the amphipod *Talorchestia capensis,* and found phylogeographic breaks that did not correspond with those observed by Teske et al. (2006). Similar observations have been made for other coastal invertebrates (Evans et al. 2004; Ridgway et al. 2001; Teske et al. 2007; Zardi et al. 2007). Although these studies failed to uncover congruent geographic patterns of genetic variance for the surveyed species, they revealed that several of these species represent complexes of deeply-divergent lineages indicating the presence of cryptic diversity among South African coastal invertebrates. Considering that cryptic diversity has been reported for other coastal invertebrates around the world (e.g. Chan et al. 2007; Hurtado et al. 2013; Radulovici et al. 2009; Santamaria et al. 2017; Santamaria et al. 2016; Santamaria et al. 2014; Santamaria et al. 2013; Varela & Haye 2012), the findings of Teske et al. (2006) and Baldanzi et al. (2016) suggest other coastal organisms in South Africa, particularly those with low vagility, may harbor previously unreported cryptic diversity. Studying such organisms may thus further our understanding of the biogeographic patterns of southern Africa and possibly uncover new taxa.

Isopods of the genus *Ligia* are one such group of organisms characterized by low vagility. Although found along rocky coastlines throughout the world (Schmalfuss 2003), the biology of these supralittoral isopods is marked by traits that severely limit their dispersal potential. As all other peracarids, they lack planktonic larvae, the embryos developing instead inside a marsupium, or brood pouch, on females until hatching as fully-formed juveniles (termed manca). Once hatched, *Ligia* isopods exhibit low desiccation and submergence resistance (Barnes 1936; Barnes 1938; Todd 1963; Tsai et al. 1997; Tsai et al. 1998; Zhang et al. 2016), avoid open water and quickly attempt to regain the shore when dislodged from rocks (Barnes

1932; Barnes 1935), as well as exhibiting poor locomotion on non-rocky substrata. These traits limit both their overland and overwater dispersal potential, which may lead to severely restricted gene flow between populations, long term isolation, and in turn allopatric and cryptic speciation, as has been reported for *Ligia hawaiensis* (Santamaria et al. 2013; Taiti et al. 2003), *L. exotica* and *L. cinerascens* (Yin et al. 2013), *L. occidentalis* (Hurtado et al. 2010), *L. baudiniana* (Santamaria et al. 2014), *L. oceanica* (Raupach et al. 2014), as well *L. vitiensis* and *L. dentipes* (Santamaria et al. 2017). Phylogeographic studies of *Ligia* have led to the discovery of cryptic speciation in areas where marine diversification was not thought to occur (Santamaria et al. 2013), as well as the discovery of distributional patterns incompatible with reigning phylogeographic paradigms (Hurtado et al. 2010). Molecular characterization of previously unstudied species of *Ligia* may thus not only uncover deeply divergent lineages, representing putative cryptic species, but also be informative on the biogeography of the region under study.

Ligia populations along the southern Africa coastline have yet to be characterized using molecular approaches, leaving our understanding of the biodiversity of *Ligia* in this area incomplete. Currently, four valid *Ligia* species are thought to inhabit the region: the endemic *Ligia dilatata, L. glabrata*, and *L. natalensis*, and the introduced *L. exotica,* which to date is formally reported only from Durban harbour (Barnard 1932). Of the endemic species, *L. dilatata* and *L. glabrata* were first described by Brandt (1833) from specimens collected in the 'Cape of Good Hope' (a vague term used by early researchers to describe any location in the then Cape Colony). Due to the brevity of the initial descriptions, both species were re-described by Budde-Lund (1885). Inspection of specimens from the KwaZulu-Natal region led Collinge (1920) to describe *L. natalensis* from specimens collected from Umhlali and Winklespruit Beach. In the same work, Collinge cast doubt on the status of *L. glabrata,* suggesting it to be an immature form

of *L. dilatata*. However, Jackson (1922) and Barnard (1932) assessed all three species and considered them to be valid, based on differences in overall body shape, length of the 2nd antenna, and shape of the stylet of the 2nd pleopod in males. The last of these traits has been shown to be a useful character for distinguishing *Ligia* species, but not cryptic lineages (Santamaria et al. 2014 and references therein; see Taiti et al. 2003). Thus, similarities in the stylet of the 2nd pleopod between *Ligia* species in southern Africa (see Figure 2 of Barnard 1932) and the lack of any genetic characterization in the past, leaves it unclear whether these species are indeed valid taxa, or conversely, whether they harbor any cryptic diversity, and by extension cast doubt in their reported distributional ranges. This latter point is important, as distributional ranges of *Ligia* species and lineages may be informative in relation to the region's biogeography. The current accepted geographical ranges for *L. dilatata and L. glabrata* are similar: both being reported from Namibia to the Cape of Good Hope (Ferrara & Taiti 1979; Schmalfuss 2003); however, *L. dilatata* extends eastwards to Cape Agulhas, whereas *L. glabrata*'s range ends at the Cape of Good Hope (Figure 1A). *Ligia natalensis* is absent from the Atlantic coastline of South Africa, and is distributed from Victoria Bay (near George on the south coast of South Africa) to the KwaZulu-Natal region (Ferrara & Taiti 1979; Schmalfuss 2003). Thus, additional investigation of the ranges of these isopods along the South Africa coast may serve to further our understanding of the biogeography and biodiversity of the South Africa coastline.

In this study, we aim to determine: (1) whether the currently accepted species of *Ligia* from South Africa represent reciprocally monophyletic clades, (2) whether these species harbor deeply divergent lineages that may represent cryptic species in need of description, (3) the large scale distributional patterns of each of the *Ligia* species and lineages across southern Africa, and (4) whether distributions of these taxa/lineages along the southern African coastline match

proposed biogeographic regions. To this end we use characterized individuals collected from 18 localities spanning the area between Namibia and KwaZulu-Natal, using both mitochondrial and nuclear markers.

MATERIALS AND METHODS

Field sampling, preservation, and identification

We hand-collected *Ligia* individuals from 18 localities around the coastline of southern Africa between 2014–2017. Detailed locality information for each of the samples is provided in Table 1. Sampled localities span most of the bioregions proposed to date for South Africa (Figure 1). All samples were field-preserved and stored in 70% ethanol until molecular analyses were carried out. In the laboratory, specimens were identified to species by visual inspection of key characteristics (e.g. appendix masculina of the second pleopod of males) and comparing said traits to those reported for *Ligia* species in southern Africa (Barnard 1932; Ferrara & Taiti 1979). Field collections were carried out under Scientific Collection Permit RES2017/53 issued by the South African Department of Environmental Affairs.

Molecular laboratory methods

We extracted total genomic DNA from several pleopods for 2–15 *Ligia* individuals per location using the Quick g-DNA MiniPrep Kit (Zymo Research), following standard protocol instructions. For each individual, we PCR-amplified a 658-bp fragment of the Cytochrome Oxidase I (COI) mitochondrial gene using the LCO-1490 and HCO-2198 primers and previously published conditions (Folmer et al. 1994). We also PCR-amplified a 661-bp region of the sodium-potassium ATPase alpha subunit (NaK) gene using the NaKFb and NaKR2 primers and

standard conditions (Tsang et al. 2008). Positive PCR amplifications were determined by visualizing PCR products on 1% agarose gels stained using SYBR Safe (Invitrogen). Positive amplicons were sequenced at the University of Arizona Genetics Core, with sequences and assembled and edited (i.e. primer removal) using Geneious R8.0.5.

Sequence alignments, phylogenetic analyses, and estimation of molecular divergence

The mitochondrial COI and nuclear NaK sequence datasets were aligned independently using the MAFFT server (Katoh & Standley 2013) under standard settings for nucleotide sequences. Visual inspection of the resulting alignment produced no evidence suggestive of pseudogenes (e.g. stop codons, high rates of amino acid substitutions) or indels. Due to the limited phylogenetic signal within the NaK dataset, we did not concatenate the two datasets and carried out phylogenetic searches only on the COI resulting alignment. Relationships within the NaK dataset were estimated using haplotype network reconstructions (see below).

Phylogenetic searches were carried out under both Maximum Likelihood and Bayesian inference approaches. Maximum Likelihood phylogenetic searches were carried out in RAxML v8.1.2 (Stamatakis 2014; Stamatakis et al. 2008) and consisted of 1,000 thorough bootstrap replicates, followed by a thorough ML search under the GTR +Γ model. We produced a majority-rule consensus tree of all bootstrap replicates using the *Sumtrees* command of DendroPy v4.1.0 (Sukumaran & Holder 2010).

We carried out Bayesian phylogenetic searches in MrBayes v3.2.5 (Ronquist & Huelsenbeck 2003) and Phycas v2.2.0 (Lewis et al. 2015). Searches in MrBayes consisted of two simultaneous searches of four chains, each sampled every 5,000th tree, while Phycas searches consisted of a single search, sampled every 50th tree. All Bayesian searches were

196 carried out under the GTR +Γ model. For each Bayesian analysis, we estimated node support
197 values by discarding all samples prior to stationarity (10–25% of sampled trees) and calculating a
198 majority-rule consensus tree using the *Sumtrees* command of DendroPy v4.1.0 (Sukumaran &
199 Holder 2010).

200 Lastly, we used MEGA v7.0.7 (Kumar et al. 2016) to estimate COI Kimura 2-Parameter
201 distances (K2P) within and amongst sampled localities and major lineages observed in the above
202 phylogenetic reconstructions.

203

204 *Haplotype network reconstructions*

205 We used the ancestral parsimony algorithm proposed by Templeton et al. (1992) as implemented
206 in PopART v1.7 (Leigh & Bryant 2015) to visualize relationships between all COI haplotypes
207 recovered in this study, as well as geographic distributional patterns of genetic diversity. We
208 estimated branch connections using the TCS network option (Clement et al. 2000) of PopArt
209 with networks considered separate if connections between them exceeded 33 steps (i.e. a 95%
210 connection limit). We repeated this approach to visualize the relationships amongst NaK alleles.

211

212 *Population structure and geographical distribution of genetic diversity*

213 We carried out one-way AMOVA analyses in Arlequin v3.5 (Excoffier & Lischer 2010) to
214 explore patterns of population structure for *Ligia* across southern African coastlines. Pairwise
215 Φ_{ST} values for all localities sampled in this studied were estimated based on Tamura and Nei's
216 (1993) genetic distances, with significant deviations from a null hypothesis of no differentiation
217 among populations determined using a non-parametric permutation approach based on 10,000
218 permutations of our dataset (Excoffier et al. 1992).

To explore patterns of geographic distribution of genetic variance in *Ligia*, we carried out two-way AMOVAs under three different biogeographical hypotheses. Under Hypothesis 1 localities were separated into four groups according to their geographic location respective to the biogeographic breaks reviewed by Teske et al. (2006). Localities were grouped as follows: Group 1: A1–3, and B1; Group 2: B2–B5; Group 3: C1–2, D1, E1–E3, F1; Group 4: D2–D4. Hypothesis 2 clustered populations according to the biogeographic regions developed by Lombard et al. (2004) and also used by Griffiths et al. (2010). It included the following groups: Group 1: A1–A2; Group 2: A3 and B1; Group 3: B2–B5, C1–2, D1, E1–E3, F1; Group 4: D2–D4. Specific locality information is given in Table 1. Finally, Hypothesis 3 tested whether phylogenetic relatedness best explained the geographic distribution of genetic variance and thus grouped localities according to the six major clades (*A–F*) identified in our phylogenetic reconstructions (see Results). For each hypothesis, we used Arlequin v3.5 (Excoffier & Lischer 2010) to estimate Φ statistics (Wright 1949) based on Tamura and Nei's (1993) genetic distances, with significance levels estimated using 10,000 permutations, and all other settings as default.

RESULTS

We successfully amplified 658-bp of the COI mtDNA gene for 99 *Ligia* individuals from 18 localities across southern Africa (Figure 1B). From these individuals, we recovered 60 haplotypes, which were separated by 162 parsimony informative sites. All new COI haplotypes and NaK alleles recovered in this study have been deposited in GenBank under accession numbers XXXXXXX (Table 1).

242 *Phylogenetic Results*

243 Preliminary analyses recovered the monophyly of southern Africa *Ligia* species; however,
244 resolution and support values within the ingroup were poor. As such, we present trees resulting
245 of analyses excluding distant outgroups and rooted using a midpoint-root approach (Figure 2).
246 All analytical approaches produced similar topologies and similar support values.

247 We observed a basal split between two well supported clusters of highly divergent clades:
248 a "Western" cluster [reds and greens in all figures; Bootstrap support (BS): 100; Maximum
249 Posterior Probability (MPP): 100%] with a geographic distribution from Namibia to the Cape
250 Agulhas region, and an "Eastern" cluster [blues, yellows, and purples in all figures; Bootstrap
251 support (BS): 100; Maximum Posterior Probability (MPP): 100%] that was distributed from
252 Knysna, on the south coast of South Africa (hereafter SA), to the KwaZulu-Natal region of SA.
253 Each of these clusters were composed of two or more highly divergent clades (clades *A–F*;
254 amongst clade COI K2P divergences 3.1–17.2%, Table 2).

255 COI haplotypes assigned to the "Western" cluster were further divided into two highly
256 divergent clades (amongst clade COI K2P divergences: 8.5–10.7%, Table 2). *Clade A* (reds in
257 all figures; BS: 93%; MPP: 100%) included all *Ligia* individuals sampled in Namibia (A1), as
258 well as from two locations in South Africa: Jacob's Bay (A2) and Ganzekraal (A3) and
259 corresponds to the species morphologically identified as *L. glabrata*. Within this clade, we
260 observed three lineages that correspond with the sampled localities and that were moderately
261 divergent from each other (COI K2P: 5.1–5.6%; Table 3). The relationships between these
262 lineages were not well supported; however, our analyses suggest a sister-taxon relationship
263 between the lineage found in *Ligia* from Luderitz, Namibia (A1) and that found in Jacob's Bay
264 (A2) (BS: 62; MPP: <50–59). The second clade part of the "Western" cluster, *Clade B* (greens

in all figures; BS: 100; MPP: 100), comprised all *Ligia* individuals collected from localities between the Cape of Good Hope and Cape Agulhas (B1–B5) and morphologically corresponds to the species *L. dilatata*. *Clade B*, contrary to the *Clade A*, does not appear to be composed of any further divergent lineages and within-clade divergences within it were low (COI K2P: 0.0–1.2%; Table 2).

The "Eastern" cluster, which contained all *Ligia* collected from Knysna to the KwaZulu-Natal region of South Africa, was composed of three highly divergent and well supported monophyletic clades (*C–E*; COI K2P 3.1–12.2%; Table 2) which morphologically correspond to the established species *L. natalensis*. Within this cluster, clades *D*, *E,* and *F* (blues and purples in all figures) are placed in a well-supported clade (BS: 82; MPP: 100) with *Clade C* (yellows in all figures) in turn sister to this group. Relationships between *D*, *E*, and *F* are not well resolved. *Clade C* containing all *Ligia* individuals collected in the Port Elizabeth area (C1–2), was highly supported across analyses (BS: 87; MPP: 98–100), and exhibited low within-clade divergences (COI K2P 0.0–1.1%, Table 2). This clade was highly divergent from all other clades in the "Eastern" cluster (COI K2P 9.4–12.2%; Table 2) and appears genetically distinct enough to be considered a separate and previously unrecognized species within the *natalensis* group. *Clade D* (BS: 85; MPP: 100) includes most COI haplotypes obtained from *Ligia* individuals collected in Knysna (D1) and the Port Edward area (D2–4). Within clade divergence for *Clade D* was low (COI K2P 0.0–1.9%, Table 2). *Clade E* contained all COI haplotypes recovered from individuals from the Kenton-on-Sea area (E1–2), those from Kidd's Beach (E3), as well as one each from Knysna (D1) and Salmon Bay (D2). Although this clade was not strongly supported by any phylogenetic analyses (BS < 50; MPP < 50), we denote it as a separate clade given the very low levels of divergence between all haplotypes in it (COI K2P average ~1.3%; Table 2),

moderate amongst clade divergence when compared to haplotypes from clades *D* and *F* (COI K2P 3.5%–6.4%; Table 2), and the results of haplotype network reconstructions (Figure 3). Lastly, the well supported *Clade F* (cyan in all figures, BS: 100; MPP: 100) contained all individuals collected at the East London Harbor (F1) and exhibited low levels of within clade divergence (COI K2P 0.0–0.6%; Table 2)

COI haplotype network reconstructions

The results of our COI haplotype network reconstructions (Figure 3) largely match patterns produced by our phylogenetic analyses, as we recovered four separate networks (i.e. connections of <95%) largely corresponding to clades observed in phylogenetic reconstructions.

Network I (Panel I of Figure 3) contained four haplotypes recovered from *Ligia* individuals from Luderitz (A1), Jacob's Bay (A2) and Ganzekraal (A3). In Luderitz, we recovered a single haplotype that diverged by 29–30 steps from haplotypes recovered from Ganzekraal, which in turn diverged by 31–32 steps from the single haplotype recovered in Jacob's Bay. This network closely parallels the patterns observed for *Clade A* in our phylogenetic reconstructions and contains all individuals morphologically identified as *L. glabrata*.

Network II (Panel II of Figure 3) contained all 14 haplotypes recovered from *Ligia* collected between the Cape of Good Hope and Cape Agulhas (B1–B5) and closely matches *Clade B*. Divergences in this network were low, with most connections between haplotypes being only 1–2 steps and the maximum connection between haplotypes being 11 steps. These correspond to *L. dilatata*. Despite such short connections, the network suggests some isolation between localities, as no sharing of haplotypes is observed. Furthermore, haplotypes recovered

within a single location appear to be much more similar (1–2 steps) than to those found at other locations in the region (~4 steps).

Network III (Panel III of Figure 3) consisted of six haplotypes recovered from four *Ligia* collected in localities near Port Elizabeth (C1–2) and corresponds with *Clade C* from our phylogenetic findings. As observed in *Network II*, connections between haplotypes are very short, as most haplotypes are connected by 1–2 steps and the maximum span between haplotypes is 6 steps. This group appears to represent a previously undescribed species most closely related to *L. natalensis*.

Lastly, *Network IV* (Panel IV of Figure 3) contained 32 haplotypes divided into three sub-networks separated by <17 steps. These sub-networks appear to correspond with clades *D–F* from our phylogenetic results and represent the *L. natalensis* species complex. One sub-network (blues in Figure 3) contained all but two haplotypes from localities around Knysna (D1) and Port Edward (D2–4). Another sub-network (purples in Figure 3) contained all the haplotypes collected in the localities of Kenton-on-Sea (E1–2) and Kidd's Beach (E3). Intermediate to these two subnetworks is a small subnetwork of four haplotypes recovered from individuals collected in East London (F1; cyan in Figure 3). In general, haplotypes collected from the same locality are much more similar to each other (<6 steps) than those from others (>10 steps), with two exceptions. A COI haplotype recovered from a *Ligia* individual collected in Salmon Bay (D2) was much more similar to those found in Kidd's Beach (E3; 4–6 steps) than others from its own location (>26 steps). This haplotype was not observed in any other *Ligia* individual from any other locality. The other exception was a COI haplotype collected from an individual collected in Knysna (D1) that was shared with individuals from the Kenton-on-Sea area (E1–2). These

patterns are congruent with the amongst-locality divergences where these lineages were found (Tables 4 and 5).

NaK haplotype network reconstructions

NaK haplotype network reconstructions (Figure 4) were congruent with the above results; however, they produced much simpler patterns. We uncovered four NaK alleles separated by 1–10 steps: one allele shared by all surveyed individuals within *Clade A* (*L. glabrata*), one shared by all individuals within *Clade B* (*L. dilatata*), and two alleles from individuals from within the "Eastern" Cluster (*L. natalensis*). These latter two alleles were much more similar to each other (1 step) than to the other two recovered alleles (7–9 steps). The allele founds in the other clades were also highly divergent from those found in other clades, with the *Clade A* allele separated by 5–10 steps from other alleles and that found in *Clade B* being separated by 5–9 steps. These patterns reinforce mitochondrial findings and are concordant with the deep divergences observed in the COI dataset.

Population structure and geographical distribution of genetic diversity

Our initial one-way AMOVA produced evidence of strong population sub-division amongst all localities included in this study ($\Phi_{ST} = 0.92006$; $p < 0.001$). Furthermore, most pairwise Φ_{ST} were significant (data not shown), with only 11 exceptions. Of these, seven exceptions involved comparisons between Luderitz (A3) and other localities (A2–3, B1–2, B3, C1, E1). Two other exceptions occurred when comparing Kenton-on-Sea (E2) to Boesmansriviermond (E1) and to Salmon Bay (D2). The other exceptions were observed between Jacobsbaai (A2) and Skoenmakerskop (C1), and Skoenmakerskop (C1) and Summerstrand (C2). Despite not

achieving significance, all pairwise comparisons, with the exception of those between C1–C2, D2–E2, and E1–E2, produced Φ_{ST} values above 0.80. These findings suggest a strong pattern of population sub-division across *Ligia* populations.

Two-way AMOVA results (Table 6) suggest that Hypotheses 1 and 3 are appropriate explanations for the geographic distribution of the genetic variance for *Ligia* in southern Africa. Under these hypotheses, amongst group variance (V_A) explained 64.55% (H1) and 85.22% (H3) of the total variance in the COI dataset. Under Hypothesis 2, V_A explained 30.71% of the total variance. Lastly, hypotheses 1 and 3 produced high Φ_{CT} values (H1: Φ_{CT} = 0.64552; H3: Φ_{CT} = 0.85218). These values more than doubled those observed for hypothesis 2 (H2: Φ_{CT} = 0.30711).

DISCUSSION

Currently, three *Ligia* species endemic to the southern Africa coastlines are accepted as valid: *L. glabrata, L. dilatata,* and *L. natalensis* (Schmalfuss 2003); however, their morphological similarity (Barnard 1932; Collinge 1920) and reports of possible cryptic species in other *Ligia* species (Hurtado et al. 2010; Raupach et al. 2014; Santamaria et al. 2017; Santamaria et al. 2014; Santamaria et al. 2013; Taiti et al. 2003; Yin et al. 2013) cast doubt on whether these taxa in fact represent true biological species, junior synonyms, or cryptic species complexes in need of further description and taxonomic revision. By applying molecular and morphological approaches to *Ligia* individuals sampled from Namibia and around the coastline of South Africa to the KwaZulu-Natal region, we report evidence suggesting that two of the three currently valid species of *Ligia* in the region appear to be cryptic species complexes. Thus, the current

taxonomic standing of *Ligia* isopods in southern Africa underrepresents the true biodiversity of these isopods.

Our phylogenetic reconstructions consistently place individuals putatively identified to a given species based on morphological characteristics in highly supported and highly divergent reciprocally monophyletic clades. All individuals putatively identified as *L. glabrata* were placed in *Clade A* (red and pink in all figures), those identified as *L. dilatata* in *Clade B* (greens in all figures), and all putative *L. natalensis* individuals with the "Eastern" cluster (clades *C–E*; yellow, blues and purples in all figures). Two of these clades; however, are composed of highly divergent lineages that may represent cryptic species in need of taxonomic description. *Clade A*, which can be identified as *L. glabrata* based on morphology, and the "Eastern" cluster, morphologically identified as *L. natalensis*, are composed of moderately to highly divergent (COI K2P >4%) lineages that exceed within-species levels of divergence reported for other isopods (Hurtado et al. 2013; Hurtado et al. 2014; Hurtado et al. 2010; Santamaria et al. 2014; Santamaria et al. 2013; Xavier et al. 2012) and invertebrate species (Hebert et al. 2003). For instance, *Clade A* is composed of three highly divergent lineages (COI K2P 5.1–5.6%; Table 3) that are geographically disjunct: one found in Namibia (A1), with the other two found in localities north of Cape Town, South Africa (A2, A3). Within the "Eastern" cluster we observe a deep split between *Clade C* and the other lineages within the cluster (COI K2P 9.4–12.2%; Table 2), as well as a moderately divergent split between clades *D*, *E,* and *F* (COI K2P 3.1–6.4%; Table 2). Although within-clade divergences between localities in *Clade B* was much lower than other clades (COI K2P 0.0–1.2%), differentiation between populations appears to be ongoing, as there were no shared haplotypes between sampled locations (Figure 3).

400 The presence of possible cryptic *Ligia* species in southern Africa is in line with recent
401 studies of other species in this genus from other regions. Hurtado et al. (2010) reported the
402 presence of seven major clades (amongst clade divergences: 7.3–29.9% COI K2P) in the area
403 from Central California to Central Mexico, an area previously thought to harbor a single endemic
404 species: *Ligia occidentalis*. The presence of multiple species of *Ligia* in that region is also
405 supported by past experimental reciprocal crosses that produced no viable offspring between
406 localities now known to be highly divergent (McGill 1978). Santamaria et al. (2013) found that
407 *Ligia hawaiensis*, the single intertidal *Ligia* species considered endemic to the Hawaiian
408 archipelago, is a paraphyletic complex of at least four highly divergent lineages (amongst clade
409 divergences: 10.5–16.7% COI K2P). The presence of multiple divergent lineages has also been
410 reported for *L. baudiniana* (Santamaria et al. 2014), *L. exotica* (Yin et al. 2013), and *L. oceanica*
411 (Raupach et al. 2014) as well as other *Ligia* species in the Indian Ocean (Santamaria et al. 2017).
412 Our findings thus represent another example of underreported biodiversity for *Ligia* isopods and
413 underscore the necessity for the molecular characterization of *Ligia* populations from other
414 regions of the world and taxonomic descriptions of cryptic lineages as species.

415 The presence of cryptic lineages in southern African *Ligia* are also in accordance with
416 recent studies reporting the presence of possible cryptic species among other groups of coastal
417 invertebrates along the coastline of South Africa (Baldanzi et al. 2016; Evans et al. 2004;
418 Ridgway et al. 2001; Teske et al. 2006; Teske et al. 2007; Zardi et al. 2007). To date, deeply
419 divergent mitochondrial lineages have been reported for four other South African coastal
420 crustacean species. Teske et al. (2006), using the same region of the mitochondrial COI gene as
421 used in this study, uncovered three cryptic lineages each for the coastal isopod *Exosphaeroma*
422 *hylecoetes* and for the cumacean *Iphinoe truncata*, and two such lineages for the decapod

Upogebia africana. More recently, Baldanzi et al. (2016) used a similar approach to Teske et al. (2006) to study the coastal amphipod *Talorchestia capensis*, reporting the presence of at least three deeply-divergent lineages in this species. As with *Ligia*, the lineages reported for these species were largely geographically isolated and non-overlapping.

The distributional limits of the lineages reported by these authors are not congruent amongst previously studied species and/or with *Ligia* patterns reported herein. The three *E. hylecoetes* lineages reported by Teske et al. (2006) were a "South" lineage distributed from east of Cape Agulhas to KwaZulu-Natal (near Port St. Johns), a "South-West" lineage distributed between Cape Agulhas and the Cape of Good Hope, and a "West" lineage found west of the Cape of Good Hope. The *I. truncata* lineages exhibited dissimilar distributional ranges: a "South-West" lineage was reported from locations east of Cape Agulhas to the Goukou area, a "South" lineage ranged from the Touws Estuary to the Sundays Estuary, while the "East" lineage was reported from the Boknes Estuary to KwaZulu-Natal (Kosi Estuary). Lastly, for *U. africana,* a "West" lineage was found in the west and southeast coastlines (Olifants Estuary to Mbhanyana Estuary) with the "East" lineage reported from southeast and east coastline locations (Haga-Haga Estuary to Mkomazi Estuary). Meanwhile, the *T. capensis* lineages reported by Baldanzi et al. (2016) included a "South-West" lineage spanning the region between Port Nolloth in the west coast of South Africa to Gouritzmound in the south, a "South" lineage occurring from Glentana to Gouritzmound, and a "South-East" lineage distributed largely from Port Alfred to Port St. Johns. The distributional limits for these lineages thus contrast with those we report for *Ligia* species in the region. Broadly speaking, we observe three major distributional breaks: the Cape of Good Hope area appears to be the southernmost limit to *Clade A* or the *L. glabrata* species complex and the westernmost limit for *Clade B* or *L. dilatata*, Cape Agulhas which

446 represents the easternmost range for *Clade B*, and the Knysna area where the westernmost limit
447 for the "Eastern" cluster lies. Note, however, that the exact transition point between these two
448 clusters cannot be accurately determined from the data at hand, given the approximately 370 km
449 of yet to be sampled coastline between Cape Agulhas and Knysna.

450 The lack of congruence amongst species suggest that the evolutionary histories of coastal
451 species in southern Africa may have been shaped by different forces, rather than by a shared
452 evolutionary history. In the case of *Ligia* from southern Africa, the cryptic genetic diversity
453 herein reported may be the result of the interplay between biological traits severely limiting
454 dispersal ability (direct development, poor desiccation tolerance, no planktonic life stages),
455 patchiness of the rocky habitats preferred by *Ligia* isopods, and ecological differences amongst
456 regions along the expansive geographic area covered in this study. The absence of haplotype
457 sharing between localities, even at fairly small geographic scales (e.g. *Clade B* localities), are
458 indicative of severely restricted dispersal and allopatric isolation of populations, while the
459 geographic distribution of lineages largely along biogeographic regions suggests ecological
460 differences have played a role in the evolution of *Ligia* in the region. Most lineages appear to
461 exhibit geographical limits constrained by previously proposed biogeographic breaks (see Teske
462 et al. 2006 and references therein). Furthermore, AMOVA analyses under a widely applied
463 biogeographic hypothesis of South Africa (H1) explain the partitioning of genetic variation at a
464 comparable rate to a hypothesis where groups were clustered by phylogenetic relatedness (H3).

465 The breaks observed in southern Africa *Ligia* largely correspond with the limits of areas
466 known to differ in their physical characteristics due to the varying influences of the Atlantic
467 Ocean and Indian Ocean, and their corresponding currents: the Benguela and Agulhas Currents.
468 Colder currents are associated with the Atlantic Ocean and the west coast of southern Africa,

while the warmer currents of the Indian Ocean wash the coastal habitats of eastern southern Africa. The region between exhibits a gradient where the oceans meet. As such, these areas are known to differ in both abiotic factors, such as Sea Surface Temperatures (SST) and their seasonal variability, as well as biotic factors, such as surface chlorophyll (Demarcq et al. 2003) and primary productivity (Bustamante et al. 1995). These differences may affect the distribution of *Ligia* lineages in southern Africa. *Ligia* species are known to be differentially affected by physical factors, such as salinity, pH, and substrate moisture levels (Barnes 1932; Barnes 1934; Barnes 1935; Tsai et al. 1997; Tsai et al. 1998; Zhang et al. 2016), all of which are known to correlate with mean SST (Kawai & Wada 2007; Rayner et al. 2003; Tang 2012; van den Dool & Nap 1985). Variability amongst lineage differences in susceptibility to ecological differences have been proposed as a mechanism for the distributional patterns observed for *Ligia* lineages along the western coast of the United States (Eberl et al. 2013), suggesting differences in abiotic factors, such as ocean current and SST, across the southern Africa coastline affects the geographic distribution of *Ligia* lineages in the region. Additional work remains necessary to determine which ecological covariates shape the distribution of the lineages.

Future research may also prove helpful in determining the causal agent of local patterns that appear to be exceptions to large scale patterns herein reported; in particular: the distribution of *Clade C* amidst the geographic range of *Clade DEF*, the close relationship between *Ligia* from Knysna (D1) and those from the Port Edward area (D2–4), as well as the rare sharing of haplotypes between distant populations. Further sampling along biogeographic breaks may help elucidate the geographic ranges of the lineages reported herein and help fully resolve their distributional extents. Similarly, expanding sampling to additional areas of the southern Africa coastline may help determine whether additional cryptic lineages exist in yet to be sampled

localities. This is particularly true for the area between Knysna and Port Elizabeth, as well as the area between Cape Town and Namibia, where our sampling was not as extensive. As our sampling efforts consisted of a single visit to each site, future work should consider sampling different microhabitats or tidal levels at given localities and/or re-sampling localities included in this study at different periods of the year. Combining such efforts with sampling of new localities will not only help determine the geographic ranges of *Ligia* lineages and species in the region, but also help determine whether these are sympatric or truly allopatric. Lastly, additional work will be needed to formally describe the cryptic lineages into species.

CONCLUSIONS

By using both nuclear and mitochondrial markers, we have detected the presence of several cryptic lineages within nominal *Ligia* species in southern Africa. Lineages herein reported may represent putative cryptic species in need of description. Thus additional taxonomic work is needed to identify potential diagnostic characters. Further work may also help fully discern the distributional patterns for *Ligia* lineages and the drivers of diversification for the genus in the region.

ACKNOWLEDGEMENTS

Our thanks to Kolette Grobler for collecting the specimens from Luderitz and Alan Hodgson for those from several sites in the Eastern Cape Province. Clara Schellberg, Bekah Weatherington and Roberta Griffiths also assisted CG in collecting trips to various other sites, while Beatriz E. Santamaria and Juan G. Higuera assisted CAS in sampling several localities.

514 Christopher Randle and Sybil Bucheli kindly allowed the use of their laboratory facilities at Sam
515 Houston State University.

REFERENCES

Baldanzi S, Gouws G, Barker NP, and Fratini S. 2016. Molecular evidence of distinct evolutionary units in the sandhopper *Talorchestia capensis* (Amphipoda, Talitridae) along South African coasts. *Hydrobiologia* 779:35-46.

Barnard KH. 1932. Contributions to the Crustacean fauna of South Africa. No. 11-Terrestial Isopods. *Annals of the South African Museum* 30:179-388.

Barnes TC. 1932. Salt requirements and space orientation of the littoral isopod *Ligia* in Bermuda. *The Biological Bulletin* 63:496-504.

Barnes TC. 1934. Further observations on the salt requirements of *Ligia* in Bermuda. *The Biological Bulletin* 66:124-132.

Barnes TC. 1935. Salt requirements and orientation of *Ligia* in Bermuda. III. *The Biological Bulletin* 69:259-268.

Barnes TC. 1936. Experiments on *Ligia* in Bermuda IV. The effects of heavy water and temperature. *The Biological Bulletin* 70:109-117.

Barnes TC. 1938. Experiments on *Ligia* in Bermuda V. Further effects of salts and of heavy sea water. *The Biological Bulletin* 74:108-116.

Brandt JF. 1833. Conspectus monographiae crustaceorum oniscodorum Latreillii. *Bulletin de la Societe Imperiale des Naturalistes de Moscou* 4:171-193.

Budde-Lund G. 1885. *Crustacea Isopoda Terrestria per Familias et Genera et Species Descripta*. Copenhagen: Sumtibus Auctoris.

Bustamante RH, Branch GM, Eekhout S, Robertson B, Zoutendyk P, Schleyer M, Dye A, Hanekom N, Keats D, Jurd M, and McQuaid C. 1995. Gradients of intertidal primary

productivity around the coast of South Africa and their relationships with consumer biomass. *Oecologia* 102:189-201.

Chan BKK, Tsang LM, and Chu KH. 2007. Cryptic diversity of the *Tetraclita squamosa* complex (Crustacea: Cirripedia) in Asia: description of a new species from Singapore. *Zoological Studies* 46:46-56.

Clement M, Posada D, and Crandall A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9:1657-1659.

Collinge WE. 1920. Contributions to a knowledge of the terrestrial Isopoda of Natal. Part III. *Annals of the Natal Museum* 4:471-490.

Crawford RJM, Shannon LV, and Pollock DE. 1987. The Benguela ecosystem. Part IV. The major fish and invertebrate resources. *Oceanography and Marine Biology: An Annual Review* 25:353-505.

Demarcq H, Barlow RG, and Shillington FA. 2003. Climatology and variability of Sea Surface Temperature and Surface Chlorophyll in the Benguela and Agulhas ecosystems as observed by satellite imagery. *African Journal of Marine Science* 25:363-372.

Eberl R, Mateos M, Grosberg RK, Santamaria CA, and Hurtado LA. 2013. Phylogeography of the supralittoral isopod *Ligia occidentalis* around the Point Conception marine biogeographical boundary. *Journal of Biogeography*:10.1111/jbi.12168 (doi).

Evans BS, Sweijd NA, Bowie RCK, Cook PA, and Elliott NG. 2004. Population genetic structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder events. *Marine Ecology Progress Series* 270:163-172.

559 Excoffier L, and Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform
560 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
561 10:564-567.
562 Excoffier L, Smouse PE, and Quattro JM. 1992. Analysis of molecular variance inferred from
563 metric distances among DNA haplotypes: application to human mitochondrial DNA
564 restriction data. *Genetics* 131:479-491.
565 Ferrara F, and Taiti S. 1979. A check-list of terrestrial isopods from Africa (south of the Sahara).
566 *Monitore Zoologico Italiano Supplemento* 12:89-215.
567 Folmer O, Black M, Hoeh W, Lutz R, and Vrijenhoek R. 1994. DNA primers for amplification
568 of mitochondrial Cytochrome C Oxidase subunit I from diverse metazoan invertebrates.
569 *Molecular Marine Biology and Biotechnology* 3:294-299.
570 Griffiths CL, Robinson TB, Lange L, and Mead A. 2010. Marine biodiversity in South Africa: an
571 evaluation of current states of knowledge. *PLoS ONE* 5:e12008.
572 Hebert PDN, Cywinska A, Ball SL, and deWaard JR. 2003. Biological identifications through
573 DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological*
574 *Sciences* 270:313-321.
575 Hurtado LA, Lee EJ, and Mateos M. 2013. Contrasting phylogeography of sandy vs. rocky
576 supralittoral isopods in the megadiverse and geologically dynamic Gulf of California and
577 adjacent areas. *PLoS ONE* 8:e67827.
578 Hurtado LA, Lee EJ, Mateos M, and Taiti S. 2014. Global diversification at the harsh sea-land
579 interface: mitochondrial phylogeny of the supralittoral isopod genus *Tylos* (Tylidae,
580 Oniscidea). *PLoS ONE* 9:e94081.

581 Hurtado LA, Mateos M, and Santamaria CA. 2010. Phylogeography of supralittoral rocky
582 intertidal *Ligia* isopods in the Pacific region from Central California to Central Mexico.
583 *PLoS ONE* 5:e11633.
584 Hutchings L, van der Lingen CD, Shannon LJ, Crawford RJM, Verheye HMS, Bartholomae CH,
585 van der Plas AK, Louw D, Kreiner A, Ostrowski M, Fidel Q, Barlow RG, Lamont T,
586 Coetzee J, Shillington F, Veitch J, Currie JC, and Monteiro PMS. 2009. The Benguela
587 Current: An ecosystem of four components. *Progress in Oceanography* 83:15-32.
588 Jackson HG. 1922. A revision of the isopod genus *Ligia* (Fabricius). *Proceedings of the*
589 *Zoological Society of London* 92:683-703.
590 Katoh K, and Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
591 improvements in performance and usability. *Molecular Biology and Evolution* 30:772-
592 780.
593 Kawai Y, and Wada A. 2007. Diurnal sea surface temperature variation and its impact on the
594 atmosphere and ocean: A review. *Journal of Oceanography* 63:721-744.
595 Kumar S, Stecher G, and Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis
596 version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870-1874.
597 Leigh JW, and Bryant D. 2015. popart: full-feature software for haplotype network construction.
598 *Methods in Ecology and Evolution* 6:1110-1116.
599 Lewis PO, Holder MT, and Swofford DL. 2015. Phycas: software for Bayesian phylogenetic
600 analysis. *Systematic Biology* 64:525-531.
601 Lombard AT, Struss T, Harris K, Sink K, Attwood C, and L. H. 2004. Volume 4: Marine
602 Component. South African National Spatial Biodiversity Assessment 2004. Technical
603 Report. Pretoria, South Africa: National Biodiversity Institute.

Lutjeharms JRE. 1998. Coastal hydrography. In: Lubke R, and De Moor I, eds. *Field Guide to the Eastern and Southern Cape Coasts*. Cape Town: University of Cape Town Press, 50–61.

McGill T. 1978. Genetic divergence of mainland and insular populations of *Ligia occidentalis* (Oniscoidea: Isopoda) Ph.D. dissertation. University of California, Santa Barbara.

Radulovici AE, Sainte-Marie B, and Dufresne F. 2009. DNA barcoding of marine crustaceans from the Estuary and Gulf of St Lawrence: a regional-scale approach. *Molecular Ecology Resources* 9:181-187.

Raupach MJ, Bininda-Emonds ORP, Knebelsberger T, Laakmann S, Pfaender J, and Leese F. 2014. Phylogeographical analysis of *Ligia oceanica* (Crustacea: Isopoda) reveals two deeply divergent mitochondrial lineages. *Biological Journal of the Linnean Society* 112:16-30.

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, and Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres* 108:4407.

Ridgway TM, Stewart BA, Branch GM, and Hodgson AN. 2001. Morphological and genetic differentiation of *Patella granularis* (Gastropoda: Patellidae): recognition of two sibling species along the coast of southern Africa. *Journal of Zoology* 245:317-333.

Ronquist F, and Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.

Santamaria CA, Bluemel JK, Bunbury N, and Curran M. 2017. Cryptic biodiversity and phylogeographic patterns of Seychellois *Ligia* isopods. *PeerJ* 5:e3894.

627 Santamaria CA, Mateos M, DeWitt TJ, and Hurtado LA. 2016. Constrained body shape among
628 highly genetically divergent allopatric lineages of the supralittoral isopod *Ligia*
629 *occidentalis* (Oniscidea). *Ecology and Evolution* 6:1537-1554.
630 Santamaria CA, Mateos M, and Hurtado LA. 2014. Diversification at the narrow sea-land
631 interface in the Caribbean: phylogeography of endemic supralittoral *Ligia* isopods.
632 *Frontiers in Ecology and Evolution* 2:42.
633 Santamaria CA, Mateos M, Taiti S, DeWitt TJ, and Hurtado LA. 2013. A complex evolutionary
634 history in a remote archipelago: Phylogeography and morphometrics of the Hawaiian
635 endemic *Ligia* Isopods. *PLoS ONE* 8:e85199.
636 Schmalfuss H. 2003. World catalog of terrestrial isopods (Isopoda: Oniscidea). *Stuttgarter*
637 *Beiträge zur Naturkunde Series A* 654:1-341.
638 Schumann EH. 1987. The coastal ocean off the east coast of South Africa. *Transactions of the*
639 *Royal Society of South Africa* 46:215-229.
640 Shannon LV, and Nelson G. 1996. The Benguela: Large Scale Features and Processes and
641 System Variability. In: Wefer G, Berger WH, Siedler G, and Webb DJ, eds. *The South*
642 *Atlantic: Present and Past Circulation*. Berlin, Heidelberg: Springer, 163-210.
643 Sink K, Holness S, Harris L, Majiedt P, Atkinson L, Robinson T, Kirkman S, Hutchings L,
644 Leslie R, Lambert S, and Kewarth S. 2012. National biodiversity assessment 2011:
645 marine and coastal component report. *Pretoria: South African National Biodiversity*
646 *Institute* 322.
647 Stamatakis A. 2014. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of
648 large phylogenies. *Bioinformatics* 30:1312-1313.

649 Stamatakis A, Hoover P, and Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML
650 web servers. *Systematic Biology* 57:758-771.
651 Sukumaran J, and Holder MT. 2010. DendroPy: a Python library for phylogenetic computing.
652 *Bioinformatics* 26:1569-1571.
653 Taiti S, Arnedo MA, Lew SE, and Roderick GK. 2003. Evolution of terrestriality in Hawaiian
654 species of the genus *Ligia* (Isopoda, Oniscidea). *Crustaceana Monographs* 2:85-102.
655 Tamura K, and Nei M. 1993. Estimation of the number of nucleotide substitutions in the control
656 region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and*
657 *Evolution* 10:512-526.
658 Tang Y. 2012. The effect of variable sea surface temperature on forecasting sea fog and sea
659 breezes: a case study. *Journal of Applied Meteorology and Climatology* 51:986-990.
660 Templeton AR, Crandall KA, and Sing CF. 1992. A cladistic analysis of phenotypic associations
661 with haplotypes inferred from restriction endonuclease mapping and DNA sequence data.
662 III. Cladogram estimation. *Genetics* 132:619-633.
663 Teske PR, McQuaid CD, Froneman PW, and Barker NP. 2006. Impacts of marine biogeographic
664 boundaries on phylogeographic patterns of three South African estuarine crustaceans.
665 *Marine Ecology Progress Series* 314:283-293.
666 Teske PR, Papadopoulos I, Zardi GI, McQuaid CD, Edkins MT, Griffiths CL, and Barker NP.
667 2007. Implications of life history for genetic structure and migration rates of southern
668 African coastal invertebrates: planktonic, abbreviated and direct development. *Marine*
669 *Biology* 152:697-711.
670 Teske PR, von der Heyden S, McQuaid CD, and Barker NP. 2011. A review of marine
671 phylogeography in southern Africa. *South African Journal of Science* 107:43-53.

672 Todd ME. 1963. Osmoregulation in *Ligia oceanica* and *Idotea granulosa*. *Journal of*
673 *Experimental Biology* 40:381-392.
674 Tsai M-L, Dai C-F, and Chen H-C. 1997. Responses of two semiterrestrial isopods, *Ligia exotica*
675 and *Ligia taiwanensis* (Crustacea) to osmotic stress. *Comparative Biochemistry and*
676 *Physiology Part A: Physiology* 118:141-146.
677 Tsai M-L, Dai CF, and Chen H-C. 1998. Desiccation resistance of two semiterrestrial isopods,
678 *Ligia exotica* and *Ligia taiwanensis* (Crustacea) in Taiwan. *Comparative Biochemistry*
679 *and Physiology - Part A: Molecular & Integrative Physiology* 119:361-367.
680 Tsang LM, Ma KY, Ahyong ST, Chan TY, and Chu KH. 2008. Phylogeny of Decapoda using
681 two nuclear protein-coding genes: origin and evolution of the Reptantia. *Molecular*
682 *Phylogenetics and Evolution* 48:359-368.
683 van den Dool HM, and Nap JL. 1985. Short and long range air temperature forecasts near an
684 ocean. *Monthly Weather Review* 113:878-887.
685 Varela AI, and Haye PA. 2012. The marine brooder *Excirolana braziliensis* (Crustacea: Isopoda)
686 is also a complex of cryptic species on the coast of Chile. *Revista Chilena de Historia*
687 *Natural* 85:495-502.
688 Wright S. 1949. The genetical structure of populations. *Annals of Eugenics* 15:323-354.
689 Xavier R, Santos AM, Harris DJ, Sezgin M, Machado M, and Branco M. 2012. Phylogenetic
690 analysis of the north-east Atlantic and Mediterranean species of the genus *Stenosoma*
691 (Isopoda,Valvifera, Idoteidae). *Zoological Scripta* 41:386-399.
692 Yin J, Pan D, He C, Wang A, Yan J, and Sun H. 2013. Morphological and molecular data
693 confirm species assignment and dispersal of the genus *Ligia* (Crustacea: Isopoda:

694 Ligiidae) along northeastern coastal China and East Asia. *Zoological Journal of the*
695 *Linnean Society* 169:362-376.
696 Zardi GI, McQuaid CD, Teske PR, and Barker NP. 2007. Unexpected genetic structure of mussel
697 populations in South Africa: indigenous *Perna perna* and invasive *Mytilus*
698 *galloprovincialis*. *Marine Ecology Progress Series* 337:135-144.
699 Zhang P, Sun J, Wang S, He D, and Zhao L. 2016. Influences of desiccation, submergence, and
700 salinity change on survival of *Ligia cinerascens* (Crustacea, Isopoda): high potential
701 implication for inland migration and colonization. *Hydrobiologia* 772:277-285.

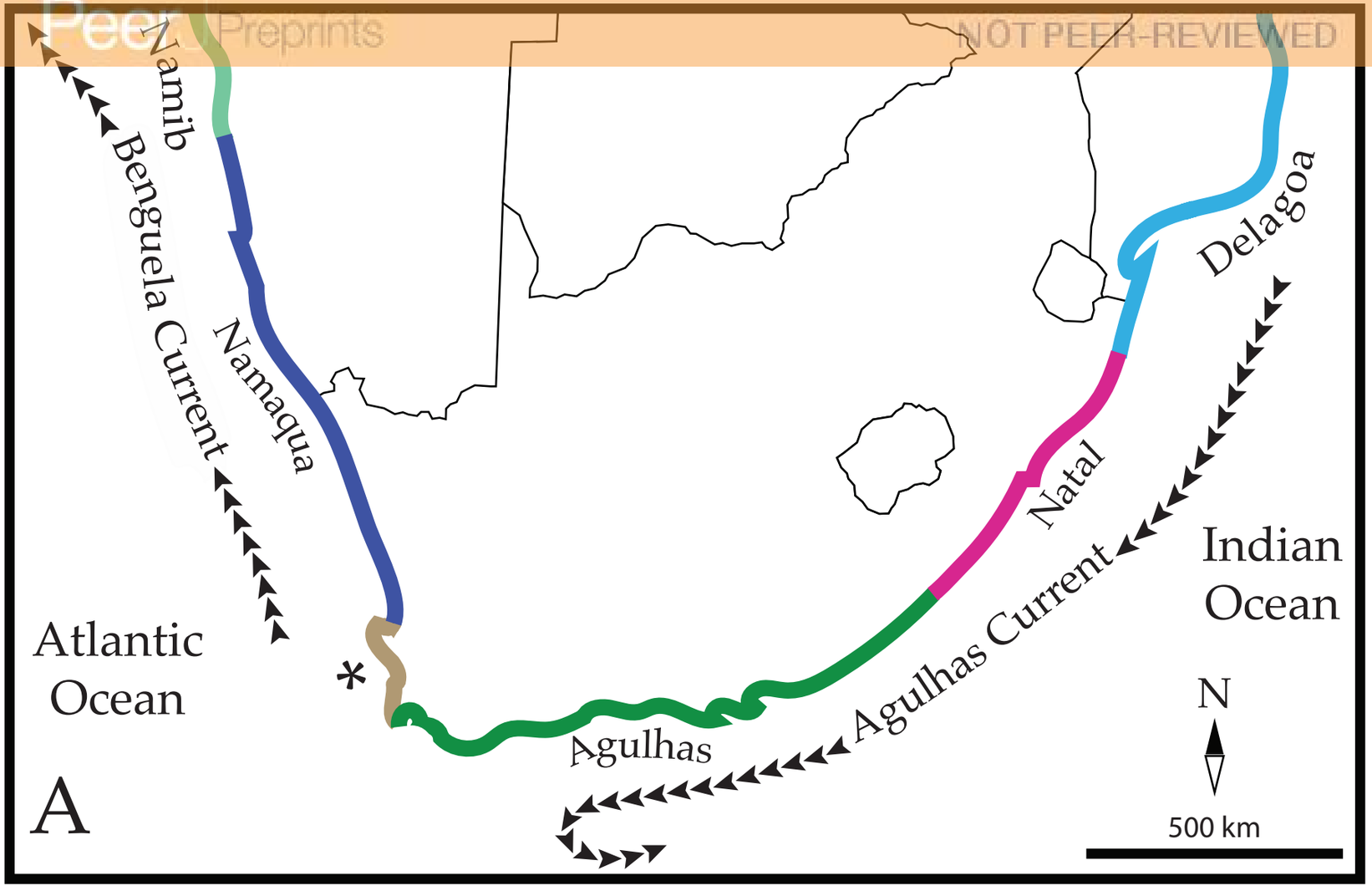
702

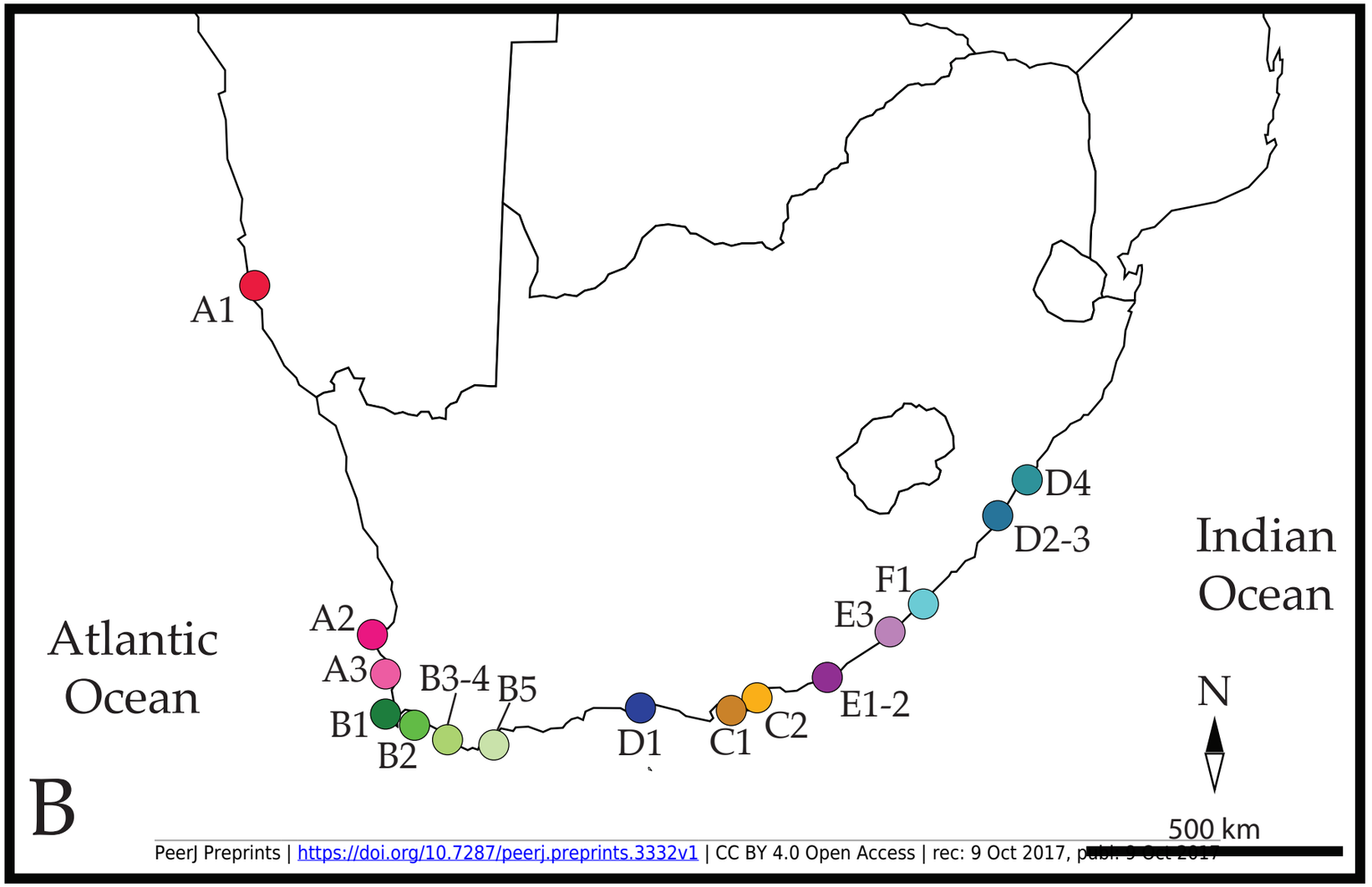
703

Figure 1(on next page)

Biogeographic regions of Southern Africa and sampled localities

Panel A: Bioregion breaks as described by Lombard et al. (2004) and Griffiths et al. (2010). This model differs from others by recognizing the South Western Cape bioregion, an overlap region between the Namaqua and the Agulhas bioregion. This is identified in Panel A by an asterisk. Proposed breaks under this model include Sylvia Hill (Namib and Namaqua), Cape Columbine (Namaqua and South Western Cape), Cape Point (South Western Cape and Agulhas), Moashe River (Agulhas and Natal), and Cape Vidal (Natal and Delagoa). Panel B: Locations sampled in southern Africa. Locations are as follows: (A1) Luderitz, (A2) Jacobsbaai, (A3) Ganzekraal, (B1) Kommetjie, (B2) Koeelbai, (B3) Onrus, (B4) Gansbaai, (B5) L'Agulhas, (C1) Skoenmakerskop, (C2) Summerstrand, (D1) Knysna, (E1, E2) Boesmansriviermond and Kenton-on-Sea, (E3) Kidd's Beach, (F1) East London Harbor, (D2-D3) Salmon Bay and Ivy Beach, (D4) Uvongo Beach.



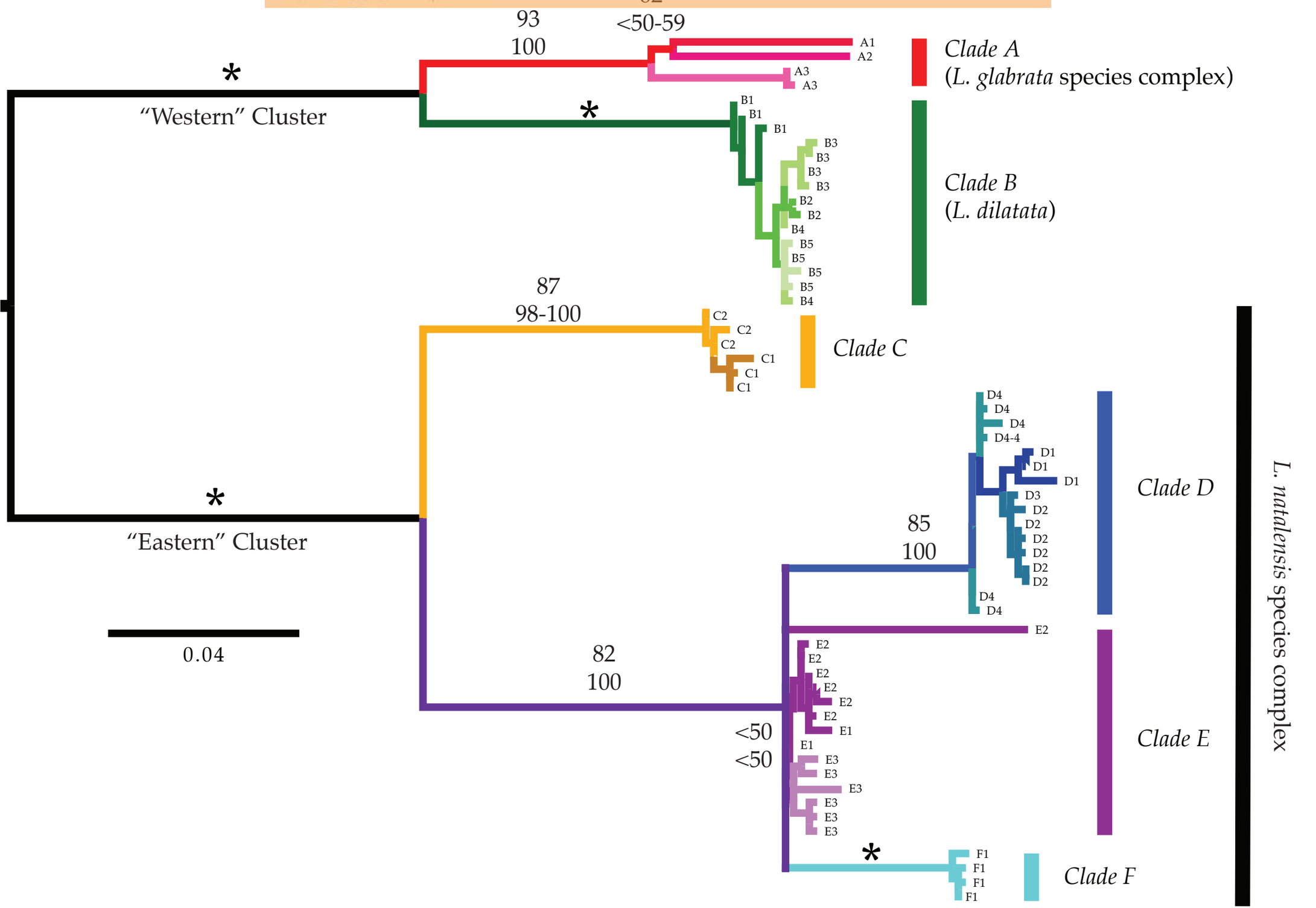


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3332v1 | CC BY 4.0 Open Access | rec: 9 Oct 2017, publ: 9 Oct 2017

Figure 2(on next page)

Phylogenetic patterns of *Ligia* from southern Africa

We observed three monophyletic groups that largely match currently valid species of *Ligia* in southern Africa; however, additional genetic divergence was observed within some of these groups. Six major clades were observed (*Clade A*: reds; *Clade B*: greens; *Clade C: yellows*; *Clade D*: blues; *Clade E*: purples; *Clade F:* cyan) containing seven moderately to highly divergent lineages. Most of the lineages contained haplotypes from geographically nearby localities. Clades and lineages exhibit mostly disjunct geographic distributions matching biogeographic regions; however, exceptions exist. Values above branches represent support values for the corresponding branch (top value: Bootstrap Support; bottom: Maximum Posterior Probablities; *: 100 in all analyses).

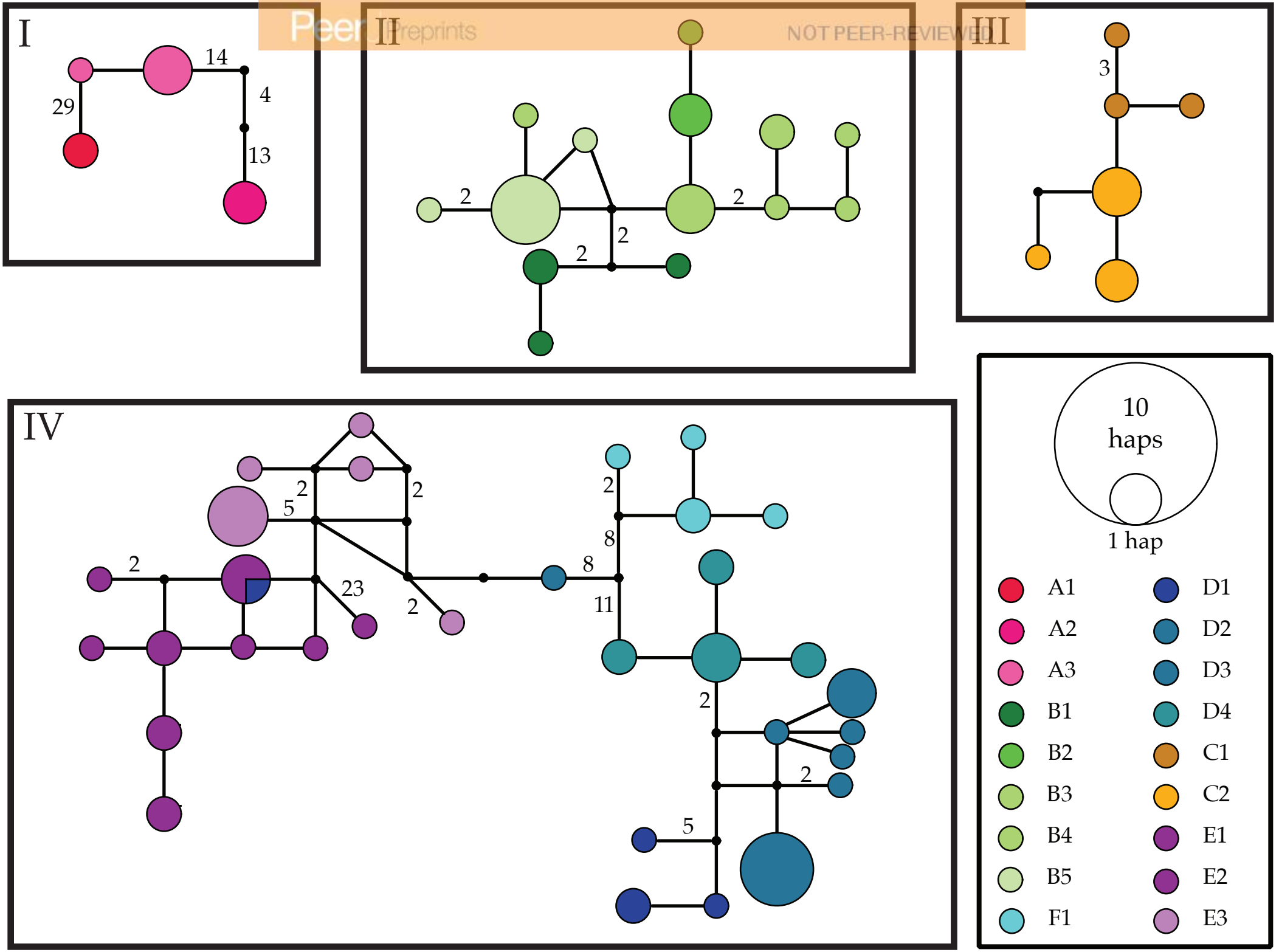


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3332v1 | CC BY 4.0 Open Access | rec: 9 Oct 2017, publ: 9 Oct 2017

Figure 3(on next page)

Haplotype networks for the COI mitochondrial gene fragment of *Ligia* from southern Africa.

Colors correspond with those used in other figures. Black circles represent inferred unsampled haplotypes with numbers along branches showing number of nucleotides differences between haplotypes. Frequency of haplotype recovery is represented through the relative sizes of the circles. Each panel (I, II, III, IV) represent networks which are more than 5% different. Locality labels correspond with those in Figure 1 and Table 1.

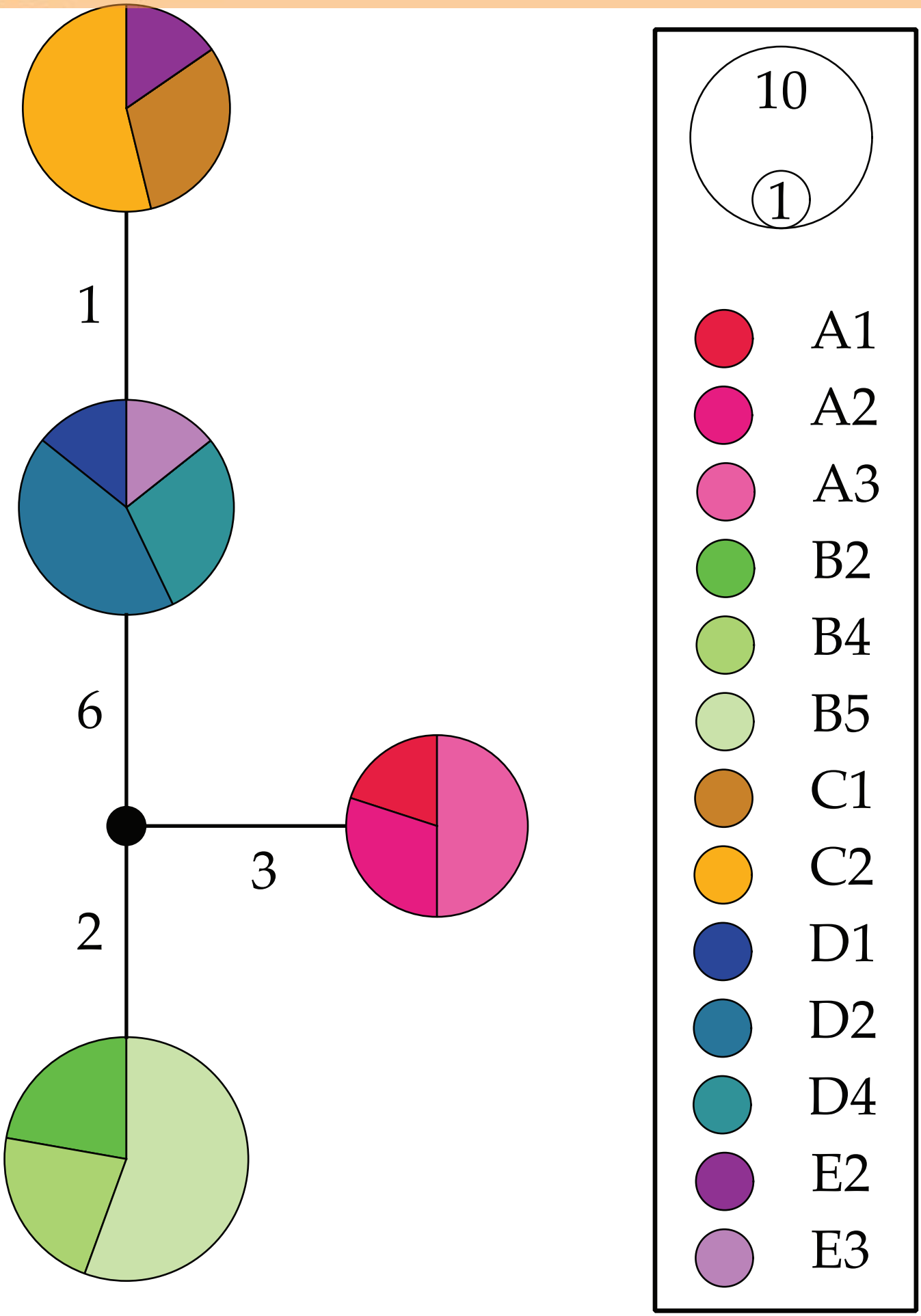


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3332v1 | CC BY 4.0 Open Access | rec: 9 Oct 2017, publ: 9 Oct 2017

Figure 4(on next page)

Haplotype networks for the nuclear gene NaK for *Ligia* from southern Africa

Colors correspond with those in all other figures with locality labels corresponding with those in other figures and Table 1. Unsampled or missing alleles are denoted by empty circles with numbers along branches indicating number of mutational steps separating alleles. Circle sizes and color proportions within them are relative to allele frequencies.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3332v1 | CC BY 4.0 Open Access | rec: 9 Oct 2017, publ: 9 Oct 2017

Table 1(on next page)

Localities included and corresponding GenBank Accession Numbers for all genetic markers used, latitude, and longitude.

Map labels correspond with other figures and tables.

Species	Locality	Map Label	N^A	$N_h{}^B$	COI Acc. Nos.	NaK Acc. No.	Latitude	Longitude
L. glabrata	Luderitz, Namibia	A1	2	1	XXXXXX	XXXXXX	26°39'47"S	15°04'55"E
L. glabrata	Jacobsbaai, South Africa	A2	3	1	XXXXXX	XXXXXX	32°58'26"S	17°53'06"E
L. glabrata	Ganzekraal, South Africa	A3	5	2	XXXXXX	XXXXXX	33°31'18"S	18°19'19"E
L. dilatata	Kommetjie, South Africa	B1	4	3	XXXXXX	N/A	34°08'17"S	18°19'24"E
L. dilatata	Koelbaai, South Africa	B2	4	2	XXXXXX	XXXXXX	34°14'51"S	18°51'15"E
L. dilatata	Onrus, South Africa	B3	5	4	XXXXXX	N/A	34°25'13"S	19°10'35"E
L. dilatata	Gansbaai, South Africa	B4	5	2	XXXXXX	XXXXXX	34°35'10"S	19°20'34"E
L. dilatata	L'Agulhas, South Africa	B5	10	4	XXXXXX	XXXXXX	34°49'26"S	20°01'01"E
L. natalensis	Knysna, South Africa	D1	4	3	XXXXXX	XXXXXX	34°02'16"S	23°01'09"E
L. natalensis	Skoenmakerskop, South Africa	C1	3	3	XXXXXX	XXXXXX	34°02'45"S	25°38'01"E
L. natalensis	Summerstrand, Port Elizabeth, South Africa	C2	4	1	XXXXXX	XXXXXX	33°59'01"S	25°40'16"E
L. natalensis	Boesmansriviermond, South Africa	E1	5	2	XXXXXX	N/A	33°40'51"S	26°39'20"E
L. natalensis	Kenton-on-Sea, South Africa	E2	10	7	XXXXXX	XXXXXX	33°41'41"S	26°39'54"E
L. natalensis	Kidd's Beach, South Africa	E3	10	5	XXXXXX	XXXXXX	33°08'50"S	27°42'10"E
L. natalensis	East London Harbor, South Africa	F1	5	4	XXXXXX	XXXXXX	33°01'28"S	27°53'26"E
L. natalensis	Salmon Bay, Port Edward, South Africa	D2	9	6	XXXXXX	XXXXXX	31°03'43"S	30°13'23"E
L. natalensis	Ivy Beach, Port Edward, South Africa	D3	9	1	XXXXXX	N/A	31°01'44"S	30°14'37"E
L. natalensis	Uvongo Beach, Margate, South Africa	D4	10	6	XXXXXX	XXXXXX	30°49'59"S	30°23'56"E

[A]: Number of individuals sampled in location
[B]: Number of unique COI haplotypes in location

Table 2 (on next page)

Pairwise amongst clade COI K2P divergences.

Ranges represent minimum and maximum values obtained when comparing individuals amongst clades, with values in parenthesis representing average divergences between members of various clades.

1

	Clade A	*Clade B*	*Clade C*	*Clade D*	*Clade E*	*Clade F*
Clade A	0.0-5.6% (3.7%)					
Clade B	8.5-10.7% (9.4%)	0.0-1.2% (0.5%)				
Clade C	13.2-15.3% (14.1%)	13.3-14.6% (13.8%)	0.0-1.1% (0.4%)			
Clade D	14.9-16.8% (15.7%)	15.4-17.0% (16.2%)	10.3-12.0% (11.2%)	0.0-1.9% (0.7%)		
Clade E	14.3-17.2% (15.4%)	15.1-16.6% (15.6%)	9.4-12.1% (10.0%)	3.5-6.3% (4.5%)	0.0%-5.4% (1.3%)	
Clade F	15.1-16.9% (15.6%)	15.5-16.5% (15.9%)	11.0-12.2% (11.5%)	3.6-6.4% (4.1%)	3.1%-6.4% (3.8%)	0.0-0.6% (0.4%)

2

3

Table 3(on next page)

Pairwise divergences for localities/lineages from *Clade A* as determined by COI K2P.

Ranges represent minimum and maximum values obtained when comparing individuals from different sampling localities, with values in parenthesis representing average divergences between members of said localities.

1

	A1	A2	A3
A1	0.0–0. 0% (0.0%)		
A2	5.6–5.6% (5.6%)	0.0–0. 0% (0.0%)	
A3	5.1–5.2% (5.2%)	5.2%–5.4% (5.2%)	0.0–0.2% (0.1%)

2
3

Table 4(on next page)

Within Clade divergences for populations from *Clade D* as determined by COI K2P.

Ranges represent minimum and maximum values obtained when comparing individuals from different sampling localities, with values in parenthesis representing average divergences between members of said localities.

1

	D1	D2	D3	D4
D1	0.0–4.6% (2.0%)			
D2	0.8–4.8% (1.9%)	0.0–4.8% (1.2%)		
D3	0.8–4.1% (1.6%)	0.3–4.6% (0.9%)	0.0–0.0% (0.0%)	
D4	0.9–4.1% (1.8%)	0.5–4.3% (1.2%)	0.8–1.2% (1.0%)	0.0–0.8% (0.4%)

2

3

Table 5(on next page)

Within Clade divergences for populations from *Clade E* as determined by COI K2P.

Ranges represent minimum and maximum values obtained when comparing individuals from different sampling localities, with values in parenthesis representing average divergences between members of said localities.

1

	E1	E2	E3
E1	0.0–0.5% (0.3%)		
E2	0.2–4.7% (0.9%)	0.0–4.9% (1.2%)	
E3	0.8–1.7% (1.2%)	0.8–5.4% (1.8%)	0.0–1.4% (0.8%)

2
3

Table 6(on next page)

Analysis of Molecular Variance (AMOVA) testing of the partitioning the genetic variation under three biogeographical hypotheses.

	Source of variation	d.f.[1]	SS[2]	Variance[3]	Var. %[4]	Φ-stats[5]	p>0.05[6]
Hypothesis 1: Biogeographic regions per Tesk et al. (2006)	Among groups	3	2100.826	23.85643	64.55	0.64552	***
	Among populations	14	918.605	10.60350	28.69	0.93244	***
	Within populations	93	232.200	2.49677	6.76	0.80941	***
Hypothesis 2: Biogeographic regions per Lombard et al. (2014)	Among groups	3	1036.466	11.01162	30.71	0.30711	0.01436
	Among populations	14	1982.964	22.34764	62.33	0.93037	***
	Within populations	93	232.200	2.49677	6.96	0.89950	***
Hypothesis 3: As per monophyletic lineages (This study)	Among groups	5	2785.060	30.90661	85.22	0.85218	***
	Among populations	12	234.370	2.86421	7.90	0.93116	***
	Within populations	93	232.200	2.49677	6.88	0.53427	***

[1]: Degrees of freedom
[2]: Sum of Squares
[3]: Variance
[4]: Percentage of variance explained by organizational level
[5]: Φ-statistics
[6]: *p*-values for Φ and components of variance (Φ_{CT} and V_A; Φ_{ST} and V_B; Φ_{SC} and V_C). ** $P < 0.01$; *** $P < 0.001$