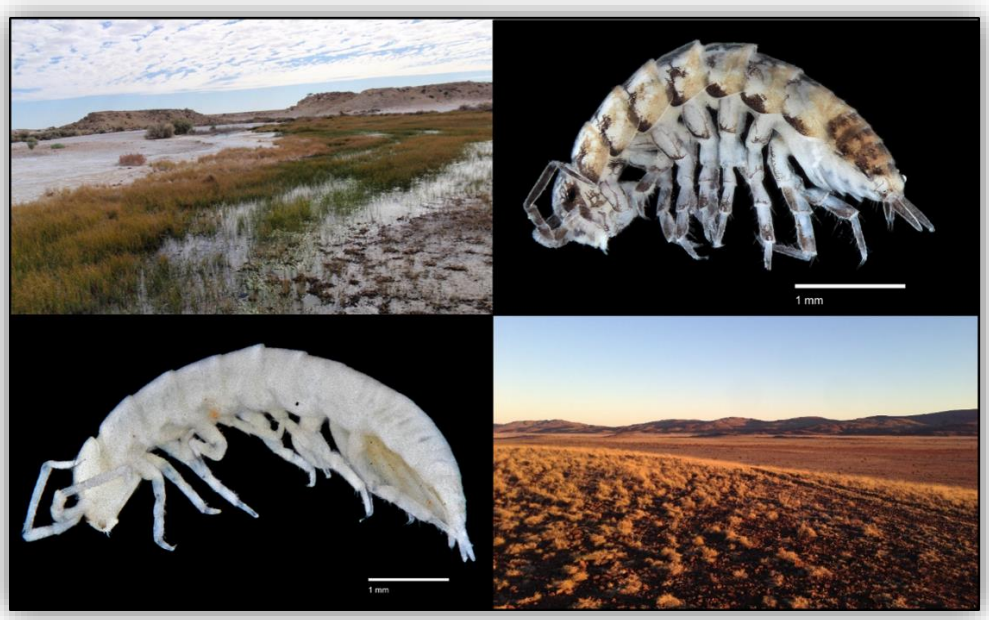
The systematics and biogeographic history of Australian arid zone oniscidean isopods (Philosciidae)

Danielle Nicole Stringer



A thesis submitted for the Degree of Doctor of Philosophy
Department of Ecology and Evolutionary Biology
School of Biological Sciences, Faculty of Sciences
The University of Adelaide

July 2019

Cover page images: Top left, Great Artesian Basin springs in the Lake Eyre region of South Australia; top right, *Haloniscus fontanus* from Strangways springs, South Australia; bottom left, a subterranean species of *Haloniscus* (undescribed) from the Laverton calcrete aquifer in the Yilgarn region of Western Australia; bottom right, an exemplar image of the Western Australian arid zone. *Haloniscus* images by Alana Delaine.

TABLE OF CONTENTS

ABSTRACT

Groundwater-dependent ecosystems in the Australian arid zone consist of highly diverse and relictual endemic invertebrates with complex evolutionary histories. Recent molecular phylogenetic studies on *Haloniscus* isopods, in particular, have identified significant levels of short-range endemicity, revealed extensive diversity with 26 new putative species, and have uncovered preliminary findings for a pattern of a shared evolutionary history amongst *Haloniscus* species from disparate and isolated groundwater regions. However, molecular datasets were restricted to either a single mitochondrial (*cytochrome "c" oxidase subunit I* (*COI*)) or two genes (*COI* and *18S rRNA*), which resulted in poor topological resolution for internal branches, and evolutionary connections were not assessed in detail with divergence dating analyses. In this study, we aimed to generate a substantial and informative phylogenomic dataset with a transcriptome-based exon capture approach to examine the evolution and biogeographic history of *Haloniscus* from three major Australian arid zone groundwater-dependent ecosystems: subterranean calcrete aquifers of the Yilgarn region in Western Australia and Ngalia Basin region, Northern Territory, and surface springs fed by the Great Artesian Basin in South Australia. In Chapter 2, we generated an effective methodological framework to infer an isopod-specific orthologous marker set and bait design targeting 469 single-copy protein-coding genes, provided empirical data and post-processing scripts to improve future exon capture experiments, and produced a well-resolved *Haloniscus* isopod phylogeny for further phylogenetic and biogeographic inference.

In Chapter 3, we implemented this dataset, together with additional phylogenetic analyses, divergence time dating and ancestral area reconstructions, to highlight significant historical connections between *Haloniscus* from the three groundwater regions and the influence of two major aridification intervals, one in the late Miocene and a second, following a temporary return to warmer and wetter conditions, in the Pliocene, on the isolation and ensuing diversification of the fauna. These findings contribute key insights into the biogeographic history of the Australian continent, and provide support for important hypotheses regarding the aridification of Australia. Lastly, in Chapter 4, we described four new species of *Haloniscus*, presented a revised generic diagnosis and key to the genus, transferred the genus from Scyphacidae to Philosciidae, and also transferred two species from *Andricophiloscia* to *Haloniscus*. The exploitation of groundwater for industrial, agricultural, and domestic uses represents a serious threat to these important taxa, and the formal documentation and naming of species (beyond just molecular results) is critical to the successful conservation management of these climate relicts and their refugial ecosystems.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

Full citations of publications that appear in this thesis

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33:** 556–574. https://doi.org/10.1071/IS18070

Stringer, D.N., King, R.A., Taiti, S., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2019. Systematics of *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae), with description of four new species from threatened Great Artesian Basin springs in South Australia. *Journal of Crustacean Biology*, **39**(5): 651–668. https://doi.org/10.1093/jcbiol/ruz044

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: **Date:** 29/07/2019

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisors, Dr Michelle Guzik, Dr Rachael King, Prof. Andy Austin, Prof. Steve Cooper and Dr Simon Tierney, for their continual support and guidance throughout this PhD process. Thank you especially to Michelle for her constant encouragement, invaluable advice, willingness to offer assistance, and continued belief in me when times became tough and I lost belief in myself. Thanks to Rachael for teaching me the taxonomic process, for always being enthusiastic and answering my many morphology-related questions, and for passing on a love of all things crustacean. Thank you to Steve for providing valuable expertise on genetic concepts and the subterranean realm, and for always being willing to meet and discuss ideas. Thanks to Simon for his valuable advice, good humour, and constructive feedback. Lastly, thank you to Andy for accepting my request to undertake a summer project in the invertebrate lab, for that original introduction to the world of taxonomy and systematics, and for continuing to provide me with fantastic opportunities ever since.

Thank you to the funding bodies who made this work possible: The University of Adelaide through the award of an Australian Government Research Training Program Scholarship, the Australian Biological Resources Study for a PhD top-up scholarship (CT214-11), the Nature Conservancy, kindly supported by the Thomas Foundation (Australian Conservation Taxonomy Award, administered by the Society of Systematic Biologists and the Australasian Systematic Botany Society), the Australian Research Council Linkage Scheme, and Bioplatforms Australia for sequencing costs.

I would also like to thank Dr Terry Bertozzi for his guidance with the bioinformatics side of the project and for always providing quick responses to my numerous questions. Many thanks also to Kathy Saint and Dr Tessa Bradford for important advice and training in the molecular labs, and to Tessa further for assistance with the initial round of exon capture and discussions about the results. Thanks to Dr Karen Meusemann for coming to Adelaide and teaching me the ways of orthology prediction, for introducing me to new software and for always being eager to help. Thank you to Dr Christoph Mayer for providing unpublished software and offering critical advice on the exon capture bait design. Thanks to Dr Steve Delean for offering much needed, and last minute, assistance with R, and helping to generate fantastic looking plots for the methodological chapter. Thanks to Prof. Bill Humphreys for sharing his wealth of expertise and passion for subterranean ecosystems, for accompanying me on field work, and teaching me how to sample calcrete aquifers. The collection trip to the Northern Territory was one of the most gratifying and rewarding aspects of my PhD, and I thank both Bill and Steve Cooper for this excellent opportunity.

Thank you to the wonderful people, both past and present, of the Austin lab group who made this PhD journey all the more enjoyable, particularly Erinn Fagan-Jeffries, Amelia Lewis, Barbara Langille, Gary Taylor, Sophie Harrison, Josie Hyde, and Alana Delaine. Thanks especially to Erinn for many supportive and helpful discussions, much needed lunch breaks, and for listening to my miseries with good grace and understanding.

Thanks to my family and friends for their constant support, love, and encouragement. To my parents, Julie and Mike, thanks for the never-ending reassurance and belief in me, and for our Sunday dinners where I could debrief about my week. Many thanks to my brother, Steve, for his talents and skills with Photoshop and Illustrator, and for always being willing to help fix up images at short notice. Finally, I am extremely grateful to Declan for his loving emotional and mental support, for enduring the good and the not-so-good times over the last few years, and for his continuous comfort and assurance.

CHAPTER 1:

Introduction and literature review

It has become increasingly apparent that groundwater-dependent ecosystems in the Australian arid zone comprise an extraordinarily diverse endemic invertebrate fauna with highly complex evolutionary histories. The onset of aridity and formation of inland deserts in the late Miocene is believed to have significantly impacted the evolution of these taxa, with the contraction of widespread mesic habitats and resultant isolation of once broadly distributed fauna within fragmented refugia, stimulating high levels of diversity and endemism (Byrne *et al*., 2008; Murphy *et al*., 2015a). This introductory chapter explores these scenarios in detail, and highlights emerging molecular phylogenetic methods, which are providing important insights into these historical processes and an increased understanding of climate-related faunal diversification in arid areas. The chapter further examines key groundwater-dependent ecosystems, current molecular studies on the invertebrate communities inhabiting these regions, and the limitations of these studies. Lastly, background information is provided on the focus isopod group, *Haloniscus*, the need for formal taxonomic descriptions, and next-generation sequencing approaches that enable the generation of large datasets to better address difficult evolutionary questions.

The Australian arid zone

The arid zone is the largest biome in Australia and represents one of the largest deserts in the world, occupying around 70% of the Australian continent's 7.5 million km^2 (Byrne *et al*., 2008). This biome is characterised by significantly low and unpredictable annual precipitation, with average rainfall of less than 250 mm per year, and high potential evaporation rates, generally exceeding 2,800 mm annually (Davis *et al*., 2013). Inland Australia, however, was not always arid and conditions were instead once warm and wet (Bowler, 1976; Alley & Lindsay, 1995; Martin, 2006). During much of the Tertiary, the environment consisted of extensive rainforest vegetation and substantial water bodies, with large permanent freshwater lakes, rivers and wetland systems, which supported species such as flamingos, dolphins and crocodiles (Martin, 2006). The onset of aridity in the late Miocene, however, completely transformed the Australian environment, with the contraction of once widespread mesic habitats and the subsequent expansion of sclerophyllous forests, salt lakes and dry open shrublands (Martin, 2006; Byrne *et al*., 2008), predominantly resulting from the northward movement of the Australian continent towards Asia (Beard, 1977).

It is believed that the early Pliocene interval saw a temporary return to warmer and wetter conditions, although never repeating the rich lakes of the mid-Miocene, prompted by a significant rise in sea level (Sniderman *et al*., 2007, 2016). This fluctuating phase was followed by a period of maximum aridity, which occurred during the glacial cycles of the Plio-Pleistocene, promoting the formation of sandy and stony deserts (Byrne *et al*., 2008) (see Fig. 1 for schematic summary of climate changes). Throughout these phases, sea level changes and a lack of limestone and coal sedimentation in the region resulted

in the loss of a continuous fossil record (Hill, 1994), which means that palaeoclimatic reconstructions for this interval have been difficult to assemble, and the influence of significant aridification events on evolutionary history remains poorly understood. Furthermore, the majority of studies regarding the origins and assembly of major biomes have predominantly focussed on arctic and forest regions in the Northern Hemisphere, with an emphasis on the evolutionary response of taxa to major glaciations and the spread of continental icesheets (Hewitt, 2000, 2004; Schafer *et al*., 2010). Nonetheless, molecular sequencing technologies are now helping to provide much needed insight into these complex historical processes.

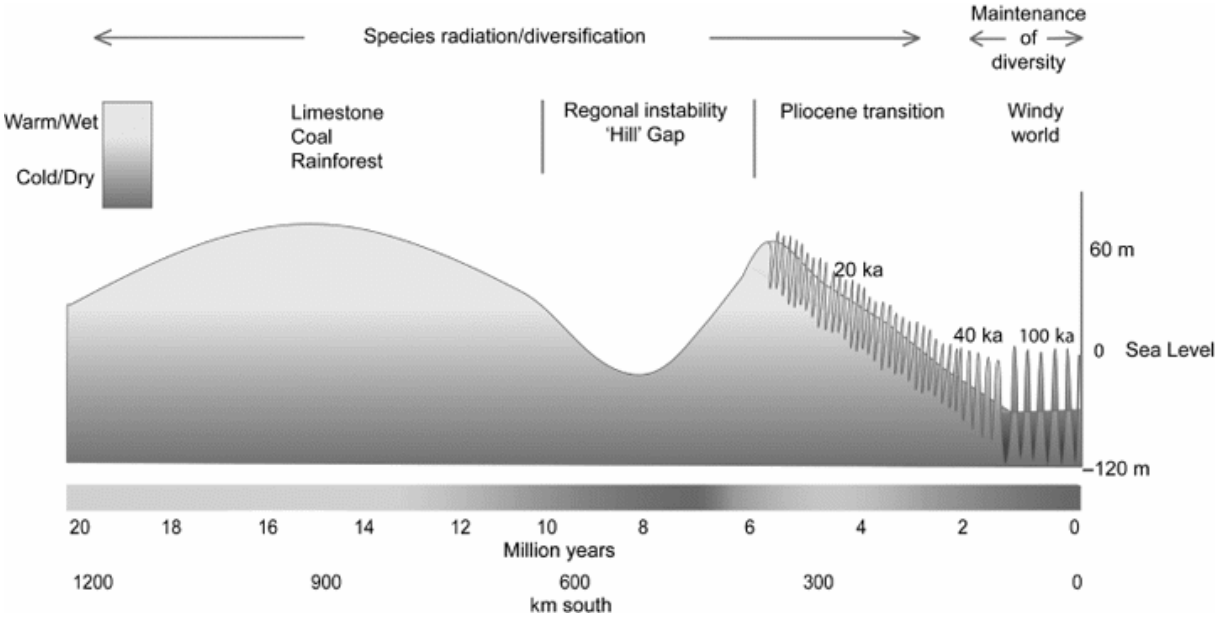


Figure 1: Summary of the palaeoclimatic conditions in Australia from 20 million years ago to present. Vertical axis representing sea level not to scale. Horizontal axes represent (i) time in the past, and (ii) distance that the continent of Australia was further south than present during the past. Shaded areas indicate warm/wet vs. cold/dry climate conditions (from Byrne *et al*., 2008: Fig. 1).

Phylogeography and refugia

In cases where fossil data are absent, species-level phylogenetics and phylogeographic studies of taxa can increase understanding of the biogeographic history of a region and the influence of climatic, and other environmental, changes on present species' distributions (Byrne, 2008; Pepper *et al*., 2011, 2018; Kleckova *et al*., 2015; Javidkar *et al*., 2016). Phylogeography involves investigating patterns of genetic variation, using gene trees along with divergence time dating techniques, in a geographic context to determine factors that have promoted diversification and the evolution of populations (Avise *et al*., 1987; Knowles & Maddison, 2002). Historical impacts may be inferred when changing conditions lead to corresponding changes in species' distributions, which then modify the genetic structure of extant species (Byrne, 2008). Dated molecular phylogenies of Australian arid zone lineages have highlighted

deep genetic divergences during the late Miocene and early Pliocene, with primarily population-level genetic effects during the later Plio-Pleistocene, coinciding with enhanced aridity (Chapple *et al*., 2004; Shepherd *et al*., 2004; Cooper *et al*., 2007; Hugall *et al*., 2008; Rix *et al*., 2018). Nevertheless, despite this recent research, a detailed understanding of climate-related faunal diversification in arid systems is largely wanting (Beheregaray, 2008).

Phylogeographic research can further assist in the identification of important refugial habitats (Byrne & Hopper, 2008; Nistelberger *et al*., 2014). Refugia are areas where components of biodiversity have retreated to, or persisted in and potentially expanded from under changing environmental conditions (Hewitt, 2000; Keppel *et al*., 2012). When conditions become unsuitable for species, their ranges may contract to limited regions that provide favourable habitats (Hewitt, 2000; Soltis *et al*., 2006; Davis *et al*., 2013). These isolated refugia offer stability as they retain environmental characteristics that were once prevalent across the landscape (Keppel *et al*., 2012). Detecting, managing, and preserving these areas is now considered an important priority for conservation, particularly under anticipated climate change, given that they have facilitated the survival of taxa over millennia through past unfavourable conditions, may promote diversification through genetic drift, and comprise many rare species (Heller & Zavaleta, 2009; Ashcroft *et al*., 2012; Moritz & Agudo, 2013; Murphy *et al*., 2015a).

Contraction, expansion and periods of isolation, often characteristic of biota present in refugia, leave genetic signatures in the molecular structure of populations and species, which can be analysed with phylogeographic techniques (Keppel *et al*., 2012). Survival in refugia encourages signatures of highly divergent lineages, which provides evidence of long-term isolation and persistence (Byrne & Hopper, 2008; Keppel *et al*., 2012). Refugial habitats can further be identified through the occurrence of relict species, which are defined as descendants of once widespread ancestors that now possess a narrow range and often originate from key climatic or other environmental changes (Habel *et al*., 2010). The presence of short-range endemic species, which inhabit significantly small ranges of less than 10,000 km^2 and demonstrate poor dispersal capabilities, is also suggestive of refugia (Harvey, 2002; Davis *et al*., 2013). Studying the phylogeography of these species, in particular, can broaden understanding of the evolutionary and biogeographical impacts of significant historical processes, such as aridification. Nonetheless, research into the role of refugia in the persistence and diversification of taxa, as for the studies into the origin of biomes, is correspondingly biased towards Northern Hemisphere ecosystems (Stewart *et al*., 2010).

Examples of acknowledged refugial habitats within the Australian arid zone include granite outcrops, isolated mountain ranges, and groundwater-dependent ecosystems (Pepper *et al*., 2011; Davis *et al*., 2013; Tapper *et al*., 2014). Groundwater-dependent ecosystems, in particular, have operated as vital

refugia within the arid zone for numerous freshwater aquatic taxa (Humphreys, 2008, 2012; Guzik *et al*., 2012; Murphy *et al.*, 2009, 2012, 2013, 2015a). With the onset of aridity in the late Miocene, and subsequently during the early Pliocene, once widespread taxa are believed to have become trapped within these fragmented groundwater habitats, following the drying of once permanent inland lakes (Murphy *et al*., 2012). These areas now contain relictual species, with very limited dispersal capabilities and small ranges, distinctive of short-range endemics (Cooper *et al.*, 2002; Guzik *et al*., 2012). Recent molecular studies focussing on these relictual species are beginning to increase understanding of the evolutionary and biogeographical impacts of aridification on the Australian arid zone biota, as well as the possible origins of groundwater biodiversity (Leys *et al*., 2003; Cooper *et al*., 2007; Murphy *et al*., 2012; Javidkar *et al*., 2017). However, the extent and complexity of the Australian arid zone indicates that further comparative phylogeographic studies integrating additional taxa from broad geographic ranges, and an increased number of molecular markers are required to better understand the impact of this aridification process.

Groundwater-dependent ecosystems

Nevill *et al*. (2010) defines groundwater-dependent ecosystems as aquatic habitats which are totally, partially or seasonally dependent on groundwater and further suggests that they can be divided into three distinct groupings: surface terrestrial, surface aquatic and subterranean. Subterranean regions, including wet caves and aquifers, are entirely dependent on groundwater and are frequently termed the ultimate groundwater-dependent ecosystems for their endemic and vastly abundant biodiversity (Humphreys, 2006). In contrast, surface aquatic habitats include water bodies, such as springs, lakes and wetlands, which depend on a connection to aquifers to sustain their water supply, while surface terrestrial environments involve plant communities extracting groundwater through their roots from the water table (Nevill *et al*., 2010). This review will largely focus on surface spring and subterranean groundwater-dependent ecosystems within the Australian arid zone.

Groundwater-dependent ecosystems are a geographically small, yet markedly diverse, and essential component of arid zone biodiversity (Murray *et al*., 2003; Tomlinson & Boulton, 2010). Groundwater supplies, however, are now increasingly exploited for domestic, agricultural and industrial processes, especially in extensive arid regions, and overuse of this essential water resource represents a critical threat to the biota inhabiting these systems (Nevill *et al*., 2010). Mining developments and coal seam gas extraction, in particular, may fundamentally impact water availability, and result in drawdown of the water table (Harrington & Cook, 2014). Groundwater overdraft, aquifer drawdown, and water diversions arising from this development may promote considerable declines in biodiversity and, thus, have a negative impact on ecosystem structure and function (Kingsford, 2000; Cramer & Hobbs, 2002).

The isolation and fragmentary nature of these habitats further indicates that faunal communities may be particularly threatened by changes to local conditions and more susceptible to extinction resulting from limited dispersal pathways (Gotch *et al*., 2008).

Effective groundwater management and a better understanding of the composition, distribution and evolutionary history of groundwater taxa are critical for conservation and future monitoring of these threatened refugial ecosystems. The South Australian Great Artesian Basin (GAB) springs and Western Australian subterranean aquifers are the foremost examples of groundwater-dependent ecosystems within the arid zone, which have been formally acknowledged as regions of biological, historical, and cultural significance. The springs are recognised as an "endangered ecological community" under the Australian Commonwealth's Environment Protection and Biodiversity Conservation Act of 1999, and the subterranean groundwater fauna are now a consideration for environmental impact assessment (EPA, 2003). Molecular phylogenetic studies on these systems are now working towards recognising and documenting their rich endemic biodiversity and increasing understanding of the relictual status and biogeographic history of these species (Guzik *et al*., 2011; Murphy *et al*., 2009, 2012).

The Great Artesian Basin springs

The GAB is a deep, regional groundwater system that underlies approximately one-quarter of Australia (Mudd, 2000). This continual desert aquifer is the largest freshwater basin in the world, spanning 1.76 million km^2 across areas of Queensland, New South Wales, South Australia, and the Northern Territory (Fig. 2) (Habermehl, 1980). Groundwater from the GAB is discharged through naturally flowing spring outlets, which form at fractures and fault lines around the boundaries of the basin, producing wetlands of variable sizes (Habermehl, 1982). These springs have been traditionally designated 'mound springs' since they are often recognisable by the presence of rounded cones, which form over time as water is released from the basin through geological pressure points, depositing minerals and carbonates on the desert surface (Thomson & Barnett, 1985). The GAB springs fall naturally into geographic hierarchical clusters where distinct springs form proximate 'groups' and 'complexes' that are hydrogeologically and hydrochemically comparable (Habermehl, 1980). The springs have been further grouped into 13 major 'supergroups' (Fig. 2), with the Lake Eyre and Dalhousie supergroups in South Australia including some of the most intact and diverse springs of the extensive GAB region (Ponder, 2002; Guzik *et al*., 2012). These ecosystems contain aquatic plants, crustaceans, molluscs, arachnids, insects and fish, which are thought to have been isolated within these regions since spring formation (Glover, 1979; Ponder *et al*., 1995; Perez *et al*., 2005; Framenau *et al*., 2006; Murphy *et al*., 2009).

Figure 2: A map of Australia highlighting the Great Artesian Basin and the locations of all major spring supergroups (from Mudd, 2000: Fig. 1). Individual springs are marked with black dots and supergroups are indicated with dashed circles.

Early research on the GAB spring fauna suggested that several of the endemic invertebrate taxa were monotypic (signifying a single species within each faunal group) and widespread across the complete range of GAB springs (Mitchell, 1985; Harris, 1992). However, more recent phylogenetic studies have revealed that these taxa in fact comprise multiple, genetically distinct, and (at times) morphologically cryptic lineages, each restricted in their distribution to discrete geographically isolated spring groups (Perez *et al*., 2005; Gotch *et al*., 2008; Murphy *et al*., 2009, 2012, 2013; Guzik *et al*., 2012). Molecular analyses of GAB spring chiltoniid amphipods, in particular, have uncovered high levels of species and

generic diversity, with each taxon occupying less than 1 km^2 of spring area, implying that these species are not just short-range endemics, but are instead best recognised as ultra-short-range endemic taxa (Murphy *et al*., 2013). The considerable levels of diversity and endemism for the spring invertebrates likely result from the geographic isolation of GAB springs, the timing of this isolation, and the limited dispersal pathways between spring groups and complexes (Ponder *et al*., 1995; Guzik *et al*., 2012). These molecular phylogenetic studies have facilitated the description of several new species from the GAB landscape (King, 2009; Murphy *et al*., 2015b); nonetheless, taxonomic work to formally describe new species has lagged behind molecular-based methods for species discovery, and, as such, further morphological taxonomic analyses are fundamental, particularly with respect to the threats faced by groundwater-dependent ecosystems (Witt *et al*., 2006; King *et al*., 2012).

Phylogeographic analyses have further proposed that the chiltoniid amphipods, along with hydrobiid snails (*Trochidrobia)* and *Phreatomerus* isopods, have multiple ancestries within the Lake Eyre spring supergroup (Murphy *et al*., 2009, 2012; Guzik *et al*., 2012). The estimated divergence times between (morphologically definable) species for each group coincide with the onset of aridification in the late Miocene, suggesting that previously widespread species became trapped within desert spring refugia (Murphy *et al*., 2012). A recent comparative phylogeographic study based on sequence data from the *cytochrome "c" oxidase subunit 1* gene (*COI*), unified these findings with those from further Lake Eyre invertebrate groups: *Fonscochlea* snails, *Haloniscus* isopods, *Ngarawa* ostracods, and *Gymnothebius* beetles (Murphy *et al*., 2015). This molecular study revealed a large degree of congruence among the evolutionary patterns for these taxa, with all groups having multiple lineages extending back to a time prior to the formation of present-day deserts, and considerable geographic-based diversification since adapting to refugial spring habitats. The chance of fauna adapting to survive in isolated refugia is low (Svenning, 2003) and, thus, the current diversity within the GAB springs may represent only a fraction of the aquatic biodiversity that originally existed preceding the aridification of Australia (Murphy *et al*., 2012).

Subterranean ecosystems

Subterranean ecosystems have been comprehensively studied in the Northern Hemisphere (Culver & Sket, 2000; Culver *et al*., 2006; Stoch & Galassi, 2010). However, due to the lack of karst habitats and predominant Pleistocene glaciation events (Barr, 1973), subterranean life in Australia was principally understudied by biospeleologists. Nevertheless, significant discoveries within the last two decades have indicated that the continent actually consists of an immense array of underground habitats and associated faunal groups, with areas of Australia now referred to as global hotspots for subterranean biodiversity (Humphreys, 2006, 2008; Guzik *et al*., 2011). Exploration of these systems has uncovered

a diverse endemic fauna with a range of higher order taxa, some of which are new to science or signify the first living relatives of lineages formerly known only from fossils (see Humphreys, 2008, 2012 for details). Subterranean taxa can be categorised into two key groups according to their habitat, namely stygofauna (or aquatic subterranean fauna) and troglofauna (or terrestrial subterranean fauna), which exhibit convergent morphologies, including reduced eyes, pigments and hardened body parts, as well as enhanced non-optic sense organs, that are associated with their evolution underground (Culver *et al*., 1995).

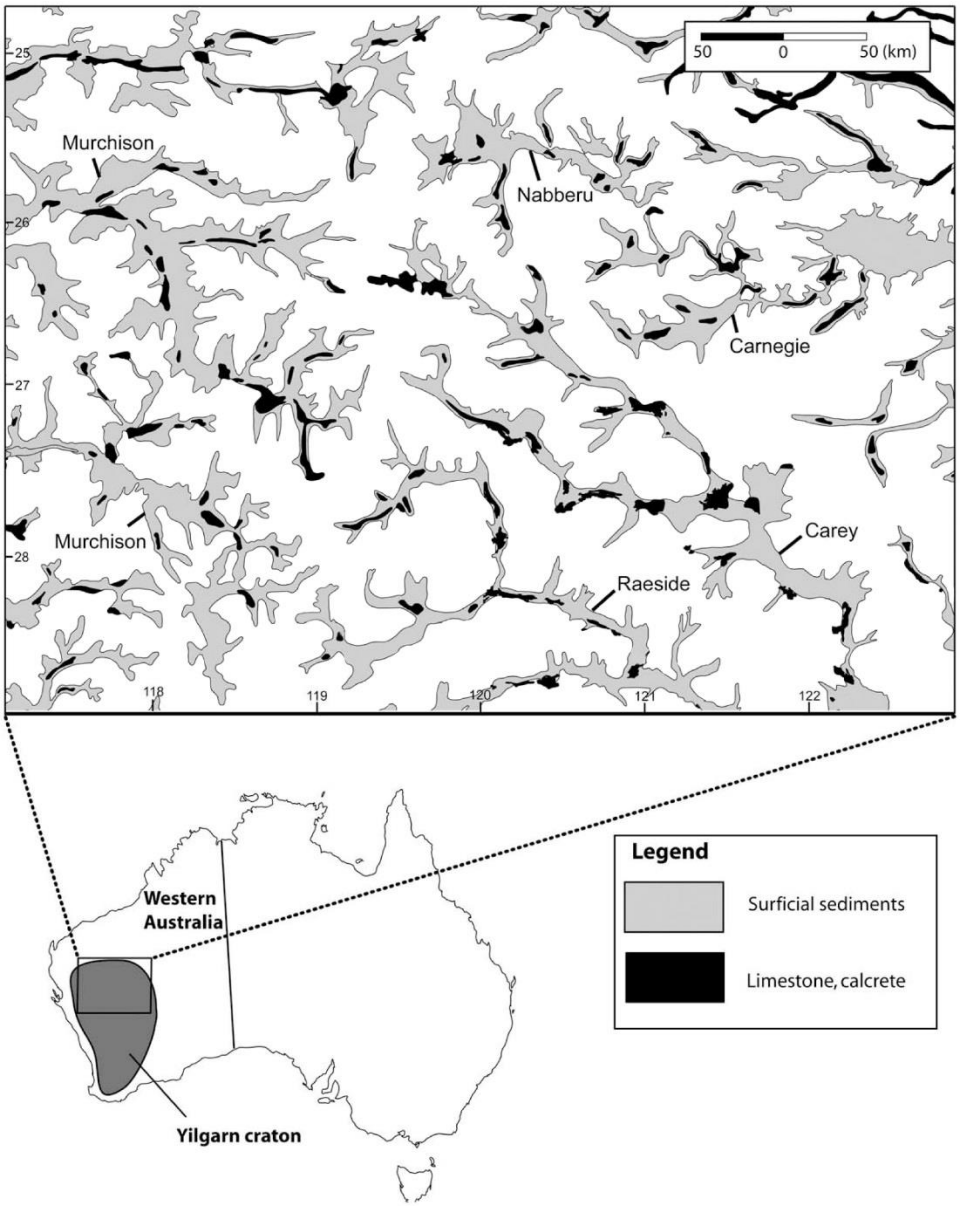


Figure 3: The calcrete aquifers of the Yilgarn region in Western Australia, with palaeodrainage valleys labelled (from King *et al*., 2012: Fig. 1).

Subterranean habitats in Australia exist primarily as groundwater aquifers and are found in a variety of matrices including: fractured rock, unconsolidated sedimentary deposits (comprising alluvium and colluvium), carbonate karst, and chemical deposits (e.g. calcretes) (Johnson & Wright, 2001). Calcrete

aquifers, in particular, are discrete, shallow and thin carbonates deposited from groundwater flow in palaeodrainage valleys, immediately upstream of salt lakes (playas) that represent the groundwater base level (Humphreys, 2006, 2008). The arid Yilgarn region in central Western Australia encompasses more than 200 isolated calcrete bodies (approximately 50–1000 km^2) associated with five distinctive palaeodrainage systems (Fig. 3). These calcretes are known to contain unique and diverse groundwater assemblages, consisting of around 200 described species as well as many additional undescribed taxa (Humphreys, 2008). These fragmented calcretes have been compared to isolated island habitats within a subterranean archipelago, with recent morphological and molecular studies highlighting a recurrent pattern of numerous phylogenetic lineages for distinct faunal groups, with one or more unique species commonly restricted to individual calcrete aquifers (Cooper *et al*., 2002, 2007, 2008; Guzik *et al*., 2008; Harrison *et al*., 2014; Javidkar *et al*., 2016). This evidence is suggestive of limited dispersal pathways and a lack of gene flow between calcrete populations, encouraging long-term isolation, speciation, and high levels of short-range endemism.

The evolution of subterranean fauna has been hypothesised to result from the colonisation of these underground regions by surface ancestral species, with diversification either occurring independently from surface ancestors due to extinction during significant climate change (climate relict hypothesis: Barr & Holsinger, 1985) or by adaptation to the subterranean habitat in the presence of reduced gene flow with surface ancestors (adaptive shift hypothesis: Desutter-Grandcolas & Grandcolas, 1996). In the most comprehensive study so far on the timing and mechanisms of colonisation within calcretes, Leys *et al*. (2003) estimated the times of transition into this subterranean environment by dytiscid diving beetles using mitochondrial *COI* sequence data and external calibrated clock rates, applying a similar method to Murphy *et al*. (2015a). The authors proposed a distinct lack of gene flow between Yilgarn calcrete populations roughly 9–4 million years ago, and supported the climate relict hypothesis, suggesting that the evolution of these beetles was driven by the aridification of Australia (Leys *et al*., 2003). These dates, nevertheless, coincide with both the late Miocene onset of aridification, and the early Pliocene return to wetter conditions, which was followed by increased aridity (Sniderman *et al*., 2007, 2016; Byrne *et al*., 2008). It, therefore, remains unclear whether these once widespread aquatic taxa became restricted to the subterranean calcretes during the first or second phase of aridification, and, as such, further research is required with clearer dates, a robust phylogeny, and additional taxa.

The Ngalia Basin region aquifers (see Fig. 4) in the Northern Territory are an additional example of an important subterranean groundwater ecosystem within the Australian arid zone. The Ngalia Basin is a small sedimentary basin, which underlies approximately 16,000 km^2 of central Australia. Unlike the Yilgarn, the Ngalia Basin aquifers (largely calcrete) are considerably less well studied, poorly mapped and the boundaries between distinct calcrete aquifers are not clearly established. Basement highs of

impermeable rock, nonetheless, ensures almost complete closure of the groundwater system, which contains steep salinity gradients that extend to hypersalinity (English *et al*., 2001). Few molecular or taxonomic studies have investigated the faunal communities that inhabit these aquifers (Balke *et al*., 2004; Cho *et al*., 2006; Watts & Humphreys, 2006; Leys & Watts, 2008). Nevertheless, it appears that they contain comparable invertebrate groups to the Yilgarn calcretes and GAB springs (e.g. *Haloniscus* isopods, dytiscid diving beetles, hydrobiid snails, bathynellaceans, and ostracods), which are known to commonly inhabit subterranean and spring ecosystems (Namiotko *et al*., 2004; Eberhard *et al*., 2005; Hancock & Boulton, 2008; Storey *et al*., 2011).

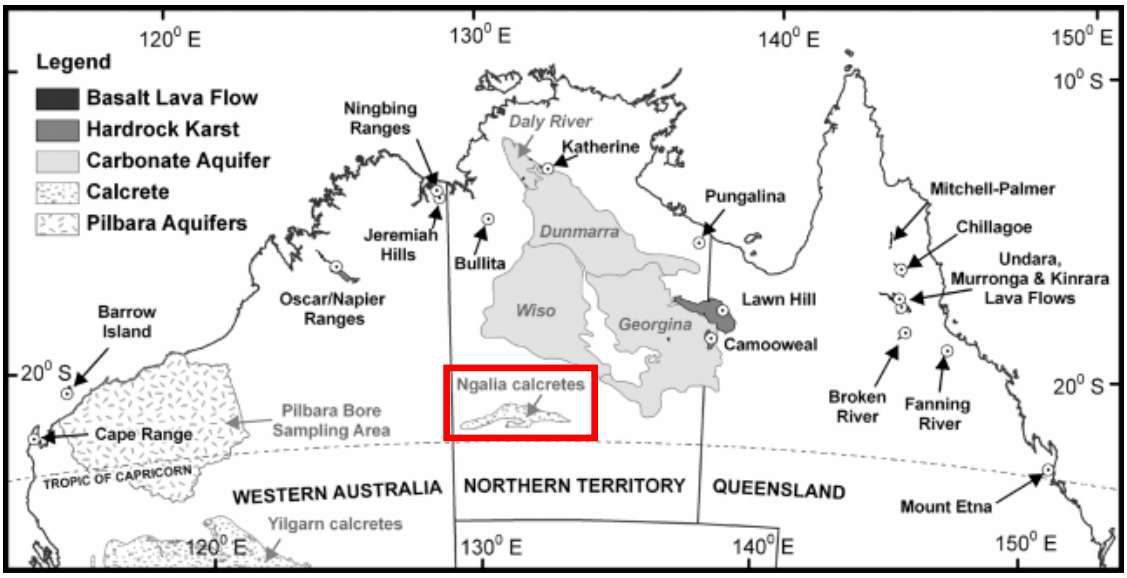


Figure 4: Map of the major aquifer basins in northern Australia. Ngalia Basin calcretes are highlighted in the red box (adapted from Moulds & Bannink 2012: Fig. 1).

Evolutionary links across arid zone groundwater ecosystems

Molecular phylogeographic studies have identified potential connections between the geographically disparate endemic invertebrate fauna of these groundwater-dependent ecosystems (Murphy *et al*., 2009; King *et al*., 2012; Guzik *et al*., 2019: Appendix). Murphy *et al*. (2009) revealed that stygobiontic chiltoniid amphipods from the Yilgarn calcretes were paraphyletic with respect to epigean chiltoniids from the GAB springs, which highlights a common ancestry and shared evolutionary history between the taxa from these apparently discrete areas of the arid zone. Approximate dating analyses proposed a Miocene origin for this aquatic fauna, where it is hypothesised that a widespread mesic and well-watered environment was once prevalent across the central Australian landscape (Martin, 2006; Byrne *et al*., 2008). The occurrence of multiple independent lineages and the connection between fauna from these disparate systems supports this hypothesis, and further suggests that a diverse and ubiquitous amphipod fauna existed in this wetter period (Murphy *et al*., 2009). However, with aridification, the

fauna has become extinct from inland lakes and river systems, but, as discussed earlier, likely survived and subsequently diversified in refugia that now represent the only permanent sources of freshwater in the arid zone (Murphy *et al*., 2015a).

Recent analyses by Guzik *et al*. (2019, see Appendix) further uncovered a repeated pattern for a shared evolutionary history between *Haloniscus* isopods from the GAB springs and Yilgarn calcrete aquifers. This study reveals connections between *Haloniscus* taxa from the Francis Swamp GAB spring complex and Yilgarn *Haloniscus*, and a Windimurra calcrete lineage with Dalhousie and Lake Eyre GAB springs taxa (Guzik *et al*., 2019). This study also integrated specimens from the Ngalia Basin aquifers, and found a potential link between these lineages and the Yilgarn (Laverton) *Haloniscus* within an *18S rRNA* only nuclear gene phylogeny; however, this link was not similarly recovered in a combined *COI* and *18S* tree. Furthermore, this study did not examine these connections in detail using a molecular divergence time dating analysis and well-resolved phylogeny. The analyses, together with those in all molecular-based studies on these groundwater ecosystems discussed so far, are limited to the small selection of readily available genes (with many only using the *COI* gene) employed in traditional Sanger sequencing, thus leading to reduced support for deeper branches. The *COI* gene, in particular, may not accurately reflect species relationships on account of the confounding effects of selection, incomplete lineage sorting or introgression (Moore, 1995). The historical connections and potential influence of aridity on *Haloniscus* (and chiltoniid amphipod) taxa remains unresolved, and requires further examination utilising multiple independent genetic markers and expanded sampling of arid zone relictual taxa.

Haloniscus

Isopods from the genus *Haloniscus* Chilton, 1920 have been recorded from the three refugial regions described above (GAB springs, Yilgarn calcretes, and Ngalia Basin aquifers), and are considered relict species, with ancestors once possibly widespread across inland Australia. *Haloniscus* from the aquifers are stygofaunal (obligate stygobionts: inhabitants of groundwater), whereas the GAB springs taxa are semi-terrestrial, occupying the moist sandy margins of springs. To date, *Haloniscus* encompasses five described species: the widespread type species, *H. searlei* Chilton, 1920, associated with salt lakes in Tasmania, Victoria, Western Australia, and South Australia (Williams, 1983); three species from the Yilgarn calcrete aquifers (Taiti & Humphreys, 2001); and one species, *H. anophthalmus* Taiti, Ferrara & Iliffe, 1995, which is found in anchialine cave waters (physicochemically stratified freshwater that has a subterranean connection to the ocean) within the Isle of Pines in New Caledonia. The three Yilgarn species were discovered from the Murchison region, with *H. tomentosus* Taiti & Humphreys, 2001 from the Cue aquifer (abandoned Cue water supply bores), and *H. stilifer* Taiti & Humphreys, 2001 and *H. longiantennatus* Taiti & Humphreys, 2001 both inhabiting the Uramurdah aquifer. There are currently

no described species from either the Ngalia Basin or the GAB springs, and many more species are known to occur within the Yilgarn calcretes.

The study by Guzik *et al*. (2019), mentioned above, examined the systematics of *Haloniscus* from the Yilgarn calcretes, Ngalia Basin aquifers, GAB springs and Australian lakes, and identified considerable endemicity, and a minimum of 26 new putative species (Appendix: Fig. 3). Their species delimitation analyses further revealed between three (Automatic Barcode Gap Discovery, ABGD (Puillandre *et al*., 2012)) and eight (Bayesian Poisson Tree Processes, bPTP (Zhang *et al*., 2013)) new species from GAB springs and between three (ABGD) and seven (bPTP) from the Ngalia Basin. As discussed previously, descriptive taxonomic work to formally identify species has lagged behind molecular-based methods for species delimitation and this is principally due to a lack of specialised taxonomists, but also owing to the morphological convergence of species and the detection of cryptic diversity (King *et al*., 2012). Alpha-taxonomic approaches to identification will, thus, likely lead to an underestimation of species diversity and so a combination of techniques and data (molecular, morphological, and geographical) are critical to assess these species boundaries (Guzik *et al*., 2011; King *et al*., 2012). Accurate species descriptions are essential for conservation purposes since legal protection and management is based on government legislation using conventional taxonomic distinctions (Harvey *et al*., 2011) and, thus, formal documentation is vital to protect these short-range, and likely relictual, species.

Next-generation sequencing

Recent developments in next-generation sequencing (NGS) technologies have enabled the rapid and cost-effective production of multi-locus sequence data for phylogeographic and systematics research (Mamanova *et al*., 2010; Zhang *et al*., 2011; Lemmon *et al*., 2012). Phylogeography and phylogenetic techniques require homologous genomic regions from numerous individuals to infer genealogies and phylogenetic trees (McCormack *et al*., 2013). With traditional Sanger sequencing methods (Sanger & Coulson, 1975), the practice of generating these overlapping regions has demanded labour-intensive marker development with single-locus polymerase chain reaction and DNA sequencing of individual samples at each locus. NGS approaches condense these time-consuming processes and produce data at considerably greater orders of magnitude (thousands of reads per locus), substantially reducing the cost per base, and increasing the amount of data that can be included in research projects (Carstens *et al*., 2012).

Using high throughput NGS methods, there are now a variety of approaches available to create large molecular datasets for the purpose of testing difficult phylogeographic hypotheses. The majority of these approaches fall within the reduced representation sequencing category, where orthologous sets

of markers from a subset of the genome are obtained across taxa of interest (Bragg *et al*., 2016). Commonly applied approaches are restriction site associated DNA (RAD) sequencing, which targets anonymous loci nearby to restriction enzyme sites (Miller *et al*., 2007), and those which target specific loci with DNA or RNA baits, such as ultra-conserved element sequencing (Faircloth *et al*., 2012, 2015), anchored hybrid enrichment (Lemmon *et al*., 2012), and transcriptome-based exon capture (Bi *et al*., 2012). This latter approach implements transcriptome sequencing to identify protein-coding exons across species, and is especially useful for organisms, such as the non-model isopod genus *Haloniscus*, that do not have a sequenced reference genome, and for sequencing degraded museum specimens and divergent lineages (Bi *et al*., 2013; Bragg *et al*., 2016; Portik *et al*., 2016; Wood *et al*., 2019).

This exon capture method has been used to successfully resolve phylogenetic relationships and infer the biogeographic history of non-model organisms, including vertebrates and invertebrates (Hugall et al., 2016; Abdelkrim *et al*., 2018; Moritz *et al*., 2018; Klopfstein *et al*., 2019; O'Hara et al., 2019; Reilly et al., 2019). The production of a capture bait set can be a time-consuming process; nonetheless, once this set has been developed for a group of organisms, it can be used again to rapidly obtain orthologous loci from additional taxa within that group (Lemmon *et al*., 2012). Therefore, the production of baits with a broad taxonomic applicability is of particular interest within a range of disciplines, including phylogenetics and species monitoring (Mayer *et al*., 2016). These bait sets could aid in the future management of the typically morphologically-conserved fauna of threatened groundwater-dependent systems, and provide the well-resolved phylogenies required to appropriately investigate the complex questions surrounding their evolution and status as climate relict species. These next-generation sequencing approaches, nevertheless, are still being developed for use in phylogenetic applications, with few readily available scripts for bioinformatic processing of the sequencing data.

Aims of the project

The overarching aim of this project was to investigate the systematics and biogeographic history of *Haloniscus* isopods from three groundwater-dependent ecosystems in the Australian arid zone using a combination of molecular, including traditional Sanger and next-generation sequencing methods, and morphological techniques.

The specific aims of the project were to:

1. Develop an effective targeted bait set for transcriptome-based exon capture of *Haloniscus* and more divergent isopod outgroup taxa (**Chapter 2**).

Chapter 2 assesses the utility of a custom bait design and the performance of transcriptome-based exon capture for the non-model isopod genus *Haloniscus,* and across distantly related outgroup taxa. The chapter outlines the methodological approach undertaken for selecting a suite of single-copy protein-coding loci using transcriptome data, designing the custom baits, laboratory protocols and pooling choices, and for processing the capture data, primarily with custom scripts, which are made available here.

2. Explore the evolution and biogeographic history of *Haloniscus* from three Australian arid zone groundwater-dependent ecosystems (Yilgarn calcretes, Ngalia Basin aquifers and GAB springs) using an exon capture approach (**Chapter 3**).

   Chapter 3 aimed to significantly broaden the scope of Guzik *et al*. (2019) (see Appendix) by investigating the biogeographic history of *Haloniscus* from the above arid zone groundwater-dependent ecosystems using the exon capture data generated as part of Chapter 2, together with a dating analysis and ancestral area reconstructions. Potential evolutionary connections between *Haloniscus* lineages from the three regions were investigated and discussed, as well as the influence of major aridification intervals on the evolution of the fauna and their status as climate relicts.

3. Use an integrative approach to describe new species of *Haloniscus* from Great Artesian Basin springs in South Australia with molecular and morphological analyses, and present a revised key to the genus (**Chapter 4**).

   In Chapter 4, four new species of *Haloniscus* were described from the South Australian GAB springs based on combined evidence from morphological assessments (conducted here), and phylogenetic and species delimitation analyses using two genes: *COI* and *18S rRNA* (study in Appendix). Based on these findings, the *Haloniscus* isopod genus was further transferred from the family Scyphacidae to Philosciidae, two additional species were transferred from the genus *Andricophiloscia* Vandel, 1973 to *Haloniscus*, and a revised key to the genus was presented.

These three results chapters have been formatted (including references) based on the requirements of particular journals, with the intention of submission following thesis completion. In the last chapter (**Chapter 5**), a general discussion is presented which provides a detailed synthesis of the research and the broader implications of this project (particularly for conservation), and further discusses limitations and likely avenues for future research. The Supplementary Material for Chapters 2 and 3 is located at the end of the thesis.

In addition to the central results chapters outlined above, I was involved in a larger collaborative and overarching research program within which my study was nested. This program generated a key paper, for which I am a co-author, directly relating to the work I present here for my PhD. The relevant paper is cited below and located in the Appendix.

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33:** 556–574. https://doi.org/10.1071/IS18070

References

Abdelkrim, J., Aznar-Cormano, L., Fedosov, A.E., Kantor, Y.I., Lozouet, P., Phuong, M.A., Zaharias, P. & Puillandre, N. 2018. Exon-capture-based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Molecular Biology and Evolution*, **35**: 2355–2374.

Alley, N.F. & Lindsay, J.M. 1995. Tertiary. In: *The geology of South Australia*. (J.F. Drexell & W.V. Preiss, eds.), pp. 151–217. South Australian Geological Survey, Bulletin 54, Adelaide, South Australia.

Ashcroft, M.B., Gollan, J.R., Warton, D.I. & Ramp, D. 2012. A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Global Change Biology*, **18**: 1866–1879.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**: 489–522.

Balke, M., Watts, C.H.S., Cooper, S.J.B., Humphreys, W.F. & Vogler, A.P. 2004. A highly modified stygobitic diving beetle of the genus *Copelatus* (Coleoptera, Dytiscidae): taxonomy and cladistic analysis based on mtDNA sequences. *Systematic Entomology*, **29**: 59–67.

Barr, T.C. 1973. Refuges of the Ice Age. *Natural History*, **26**: 26–35.

Barr, T.C. & Holsinger, J.R. 1985. Speciation in cave faunas. *Annual Review of Ecology and Systematics*, **16**: 313–337.

Beard, J.S. 1977. Tertiary evolution of the Australian flora in the light of latitudinal movements of the continent. *Journal of Biogeography*, **4**: 111–118.

Beheregaray, L.B. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**: 3754–3774.

Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R., & Moritz, C. 2013. Unlocking the vault: next-generation museum population genomics. *Molecular Ecology*, **22**: 6018–6032.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C. & Good, J. M. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, **13:** 403.

Bowler, J.M. 1976. Aridity in Australia: age, origins and expression in aeolian landforms and sediments. *Earth-Science Reviews*, **12**: 279–310.

Bragg, J.G., Potter, S., Bi, K. & Moritz, C. 2016. Exon capture phylogenomics: efficacy across scales of divergence. *Molecular Ecology Resources*, **16**: 1059–1068.

Byrne, M. 2008. Evidence for multiple refugia at different time scales during the Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews*, **27**: 2576–2585.

Byrne, M. & Hopper, S.D. 2008. Granite outcrops as ancient islands in old landscapes: evidence from phylogeography and population genetics of *Eucalyptus caesia* (Myrtaceae) in Western Australia. *Biological Journal of the Linnean Society*, **93**: 177–188.

Byrne, M., Yeates, D.K., Joseph, L., Kearney, M., Bowler, J., Williams, M.A.J., Cooper, S., Donnellan, S.C., Keogh, J.S., Leys, R., Melville, J., Murphy, D.J., Porch, N. & Wyrwoll, K-H. 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology*, **17**: 4398–4417.

Carstens, B., Lemmon, A.R. & Lemmon, E.M. 2012. The promises and pitfalls of next-generation sequencing data in phylogeography. *Systematic Biology*, **61**: 713–715.

Chapple, D.G. & Keogh, J.S. 2004. Parallel adaptive radiations in arid and temperate Australia: molecular phylogeography and systematics of *Egernia whitii* (Lacertilia: Scincidae) species group. *Biological Journal of the Linnean Society*, **83**: 157–173.

Chilton, C. 1920. On a new isopodan genus (family Oniscidae) from Lake Corangamite, Victoria. *Proceedings of the Linnean Society of New South Wales*, **44**: 723–734.

Cho, J-L., Humphreys, W.F. & Lee, S-D. 2006. Phylogenetic relationships within the genus *Atopobathynella Schminke*, 1973 (Bathynellacea, Parabathynellidae): with the description of six new species from Western Australia. *Invertebrate Systematics*, **20**: 9–41.

Cooper, S.J.B., Bradbury, J.H., Saint, K.M., Leys, R., Austin, A.D. & Humphreys, W.F. 2007. Subterranean archipelago in the Australian arid zone: mitochondrial DNA phylogeography of amphipods from central Western Australia. *Molecular Ecology*, **16**: 1533–1544.

Cooper, S.J.B., Hinze, S., Leys, R., Watts, C.H.S. & Humphreys, W.F. 2002. Islands under the desert: molecular systematics and evolutionary origins of stygobitic water beetles (Coleoptera: Dytiscidae) from central Western Australia. *Invertebrate Systematics*, **16**: 589–598.

Cooper, S.J.B., Saint, K.M., Taiti, S., Austin, A.D. & Humphreys, W.F. 2008. Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics*, **22**: 195–203.

Cramer, V.A. & Hobbs, R.J. 2002. Ecological consequences of altered hydrological regimes in fragmented ecosystems in southern Australia: impacts and possible management responses. *Austral Ecology*, **27**: 546–564.

Culver, D.C., Deharveng, L., Bedos, A., Lewis, J.L., Madden, M., Reddell, J.R., Sket, B., Trontelj, P. & White, D. 2006. The mid-latitude biodiversity ridge in terrestrial cave fauna. *Ecography*, **29**: 120–128.

Culver, D.C., Kane, T. & Fong, D.W. 1995. *Adaptation and natural selection in caves: the evolution of Gammarus minus*. Harvard University Press, Cambridge, MA, USA.

Culver, D.C. & Sket, B. 2000. Hotspots of subterranean biodiversity in caves and wells. *Journal of Cave and Karst Studies*, **62**: 11–17.

Davis, J., Pavlova, A., Thompson, R. & Sunnucks, P. 2013. Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Global Change Biology*, **19**: 1970–1984.

Desutter-Grandcolas, L. & Grandcolas, P. 1996. The evolution toward troglobitic life: a phylogenetic reappraisal of climatic relict and local habitat shift hypotheses. *Memories Biospéologie*, **23**: 57–63.

Eberhard, S.M., Halse, S.A. & Humphreys, W.F. 2005. Stygofauna in the Pilbara region, north-west Western Australia: a review. *Journal of the Royal Society of Western Australia*, **88**: 167–176.

English, P., Spooner, N.A., Chappel, J., Questiaux, D.G. & Hill, N.G. 2001. Lake Lewis basin, central Australia: environmental evolution and OSL chronology. *Quaternary International*, **83**: 81–101.

EPA. 2003. Consideration of subterranean fauna in groundwater and caves during environmental impact assessment. In: *Guidance statement for the assessment of environmental factors no. 54*. Environmental Protection Authority, Perth, Australia.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T. & Glenn, T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, **61**: 717–726.

Faircloth, B.C., Branstetter, M.G., White, N.D. & Brady S.G. 2015. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Molecular Ecology Resources*, **15**: 489–501.

Framenau, V.W., Gotch, T.B. & Austin, A.D. 2006. The wolf spiders of artesian springs in arid South Australia, with a revalidation of *Tetralycosa* (Araneae, Lycosidae). *The Journal of Arachnology*, **34**: 1–36.

Glover, C.J.M. 1979. Studies on central Australian fishes: further observations and records, part 1. *South Australian Naturalist*, **53**: 58–62.

Gotch, T.B., Adams, M., Murphy, N.P. & Austin, A.D. 2008. A molecular systematic overview of wolf spiders associated with Great Artesian Basin springs in South Australia. *Invertebrate Systematics*, **22**: 151–165.

Guzik, M.T., Abrams, K.M., Cooper, S.J.B., Humphreys, W.F., Cho, J-L. & Austin, A.D. 2008. Phylogeography of the ancient Parabathynellidae (Crustacea: Bathynellacea) from the Yilgarn region of Western Australia. *Invertebrate Systematics*, **22**: 205–216.

Guzik, M.T., Adams, M.A., Murphy, N.P., Cooper, S.J.B. & Austin, A.D. 2012. Desert springs: deep phylogenetic structure in an ancient endemic crustacean (*Phreatomerus latipes*). *PLoS ONE*, **7**(7) e37642.

Guzik, M.T., Austin, A.D., Cooper, S.J.B., Harvey, M.S., Humphreys, W.F., Bradford, T., Eberhard, S.M., King, R.A., Leys, R., Muirhead, K.A. & Tomlinson, M. 2011. Is the subterranean fauna uniquely diverse? *Invertebrate Systematics*, **24**: 407–418.

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33:** 556–574.

Habel, J.C., Assmann, T., Schmitt, T. & Avise, J.C. 2010. Relict species: from past to future. In: *Relict species: Phylogeography and conservation biology*. (J.C. Habel & T. Assmann, eds.), pp. 1–8. Springer, Heidelberg, Germany.

Habermehl, M.A. 1980. The Great Artesian Basin, Australia. *BMR Journal of Australian Geology and Geophysics*, **5**: 9–38.

Habermehl, M.A. 1982. Springs in the Great Artesian Basin, Australia- their origin and Nature. *Bureau of Mineral Resources, Geology and Geophysics Report No. 235*.

Hancock, P.J. & Boulton, A.J. 2008. Stygofauna biodiversity and endemism in four alluvial aquifers in eastern Australia. *Invertebrate Systematics*, **22**: 117–126.

Harrington, N. & Cook, P. 2014. *Groundwater in Australia*. National Centre for Groundwater Research and Training, Australia.

Harris, C.R. 1992. Mound springs: South Australian conservation initiatives. *The Rangeland Journal*, **14**: 157–173.

Harrison, S.E., Guzik, M.T., Harvey, M.S. & Austin, A.D. 2014. Molecular phylogenetic analysis of Western Australian troglobitic chthoniid pseudoscorpions (Pseudoscorpiones: Chthoniidae) points to multiple independent subterranean clades. *Invertebrate Systematics*, **28**: 386–400.

Harvey, M.S. 2002. Short-range endemism among the Australian fauna: some examples from non-marine environments. *Invertebrate Systematics*, **16**: 555–570.

Harvey, M.S., Rix, M.G., Framenau, V.W., Hamilton, Z.R., Johnson, M.S., Teale, R.J., Humphreys, G. & Humphreys, W.F. 2011. Protecting the innocent: studying short-range endemic taxa enhances conservation outcomes. *Invertebrate Systematics*, **25**: 1–10.

Heller, N.E. & Zavaleta, E.S. 2009. Biodiversity management in the face of climate change: a review of 22 years of recommendations. *Biological Conservation*, **142**: 14–32.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, **405**: 907–913.

Hewitt, G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B*, **359**: 183–195.

Hill, R.S. 1994. The history of selected Australian taxa. In: *History of Australian vegetation: cretaceous to recent*. (R.S. Hill, ed.), pp. 390–419. Cambridge University Press, Cambridge, UK.

Hugall, A.F., Foster, R., Hutchinson, M. & Lee, M.S.Y. 2008. Phylogeny of Australian agamid lizards based on nuclear and mitochondrial genes: implications for morphological evolution and biogeography. *Biological Journal of the Linnean Society*, **93**: 343–358.

Hugall, A.F., O'Hara, T.D., Hunjan, S., Nilsen, R. & Moussalli, A. 2016. An exon-capture system for the entire class Ophiuroidea. *Molecular Biology and Evolution*, **33**: 281–294.

Humphreys, W.F. 2006. Aquifers: the ultimate groundwater-dependent ecosystems. *Australian Journal of Botany*, **54**: 115–132.

Humphreys, W.F. 2008. Rising from down under: developments in subterranean biodiversity in Australia from a groundwater fauna perspective. *Invertebrate Systematics*, **22**: 85–101.

Humphreys, W.F. 2012. Diversity patterns in Australia. In: *Encyclopedia of caves, 2nd Edition*. (W.B. White & S.C. Culver, eds.), pp. 203–219. Academic Press, San Diego, USA.

Javidkar, M., Cooper, S.J.B., Humphreys, W.F., King, R.A., Judd, S. & Austin, A.D. 2017. Biogeographic history of subterranean isopods from groundwater calcrete islands in Western Australia. *Zoologica Scripta,* **47**: 206–220.

Javidkar, M., Cooper, S.J.B., King, R.A., Humphreys, W.F., Bertozzi, T., Stevens, M.I. & Austin, A.D. 2016. Molecular systematics and biodiversity of oniscidean isopods in the groundwater calcretes of central Western Australia. *Molecular Phylogenetics and Evolution*, **104**: 83–98.

Johnson, S.L. & Wright, A.H. 2001. *Central Pilbara groundwater study*. Water and Rivers Commission, Perth, Australia.

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yares, C.J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D. & Franklin, S.T. 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change*. Global Ecology and Biogeography*, **21**: 393–404.

King, R.A. 2009. Two genera and species of chiltoniid amphipods (Crustacea: Amphipoda: Talitroidea) from freshwater mound springs in South Australia. *Zootaxa*, **2293**: 35–52.

King, R.A., Bradford, T., Austin, A.D., Humphreys, W.F. & Cooper, S.J.B. 2012. Divergent molecular lineages and not-so-cryptic species: the first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. *Journal of Crustacean Biology,* **32**: 465–488.

Kingsford, R.T. 2000. Ecological impacts of dams, water diversions and river management on floodplain wetlands in Australia. *Austral Ecology*, **25**: 109–127.

Kleckova, I, Cesanek, M., Fric, Z., Pellissier, L., 2015. Diversification of the cold-adapted butterfly genus *Oeneis* related to Holarctic biogeography and climate niche shifts. *Molecular Phylogenetics and Evolution*, **92**: 255–265.

Klopfstein, S. Langille, B., Spasojevic, T., Broad, G.R., Cooper, S.J.B., Austin, A.D. & Niehuis, O. 2019. Hybrid capture data unravel a rapid radiation of pimpliform parasitoid wasps (Hymenoptera: Ichneumondae: Pimpliformes). *Systematic Entomology*, **44**: 361–383.

Knowles, L.L. & Maddison, W.P. 2002. Statistical phylogeography. *Molecular Ecology*, **11**: 2623–2635.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenetics. *Systematic Biology*, **61**: 727–744.

Leys, R. & Watts, C.H. 2008. Systematics and evolution of the Australian subterranean hydroporine diving beetles (Dytiscidae), with notes on *Carabhydrus*. *Invertebrate Systematics*, **22**: 217–225.

Leys, R., Watts, C.H.S., Cooper, S.J.B. & Humphreys, W.F. 2003. Evolution of subterranean diving beetles (Coleoptera: Dytiscidae: Hydroporini, Bidessini) in the arid zone of Australia. *Evolution*, **57**: 2819–2834.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J. & Turner, D.J. 2010. Target-enrichment strategies for next-generation sequencing. *Nature Methods*, **7**: 111–118.

Martin, H.A. 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments*, **66**: 533–563.

Mayer, C., Sann, S., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., Petersen, M., Meusemann, K., Liere, K., Wäggle, J-W., Misof, M., Bleidorn, C., Ohl, M. & Niehuis, O. 2016. BaitFisher: a software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution*, **33**, 1875–1886.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**: 526–538.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A. & Johnson, E.A. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, **17**: 240–248.

Mitchell, B.D. 1985. Limnology of mound springs and temporary pools, south and west of Lake Eyre. In: *South Australia's mound springs*. (J. Greenslade, L. Joseph & A. Reeves, eds.), pp. 51–63. Nature Conservation Society SA Inc., Adelaide, Australia.

Moore, W.S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, **49**: 718–726.

Moritz, C. & Agudo, R. 2013. The future of species under climate change: resilience or decline? *Science*, **341**: 504–508.

Moulds, T. & Bannink, P. 2012. Preliminary notes on the cavernicolous arthropod fauna of Judbarra/Gregory Karst Area, northern Australia. *Helictite*, **41**: 75–85.

Moritz, C., Pratt, R.C., Bank, S., Bourke, G., Bragg, J.G., Doughty, P., Keogh, J.S., Laver, R.J., Potter, S., Teasdale, L.C., Tedeschi, L.G. & Oliver, P.M. 2018. Cryptic lineage diversity, body size divergence, and sympatry in a species complex of Australian lizards (*Gehyra*). *Evolution*, **72**: 54–66.

Mudd, G.M. 2000. Mound springs of the Great Artesian Basin in South Australia: a case study from Olympic Dam. *Environmental Geology*, **39**: 463–476.

Murphy, N.P., Adams, M. & Austin, A.D. 2009. Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. *Molecular Ecology*, **18**: 109–122.

Murphy, N.P., Adams, M., Guzik, M.T. & Austin, A.D. 2013. Extraordinary micro-endemism in Australian desert spring amphipods. *Molecular Phylogenetics and Evolution*, **66**: 645–653.

Murphy, N.P., Breed, M.F., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2012. Trapped in desert springs: phylogeography of Australian desert spring snails*. Journal of Biogeography*, **39**: 1573–1582.

Murphy, N.P., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2015a. Desert spring refugia: museums of diversity or evolutionary cradles? *Zoologica Scripta*, **44**: 693–701.

Murphy, N.P., King, R.A. & Delean, S. 2015b. Species, ESUs or populations? Delimiting and describing morphologically cryptic diversity in Australian desert spring amphipods. *Invertebrate Systematics*, **29**: 457–467.

Murray, B.R., Zeppel, M.J.B., Hose, G.C. & Eamus, D. 2003. Groundwater-dependent ecosystems in Australia: it's more than just water for rivers. *Ecological Management and Restoration*, **4**: 110–113.

Namiotko, T., Wouters, K., Danielopol, D.L. & Humphreys, W.F. 2004. On the origin and evolution of a new anchialine stygobitic *Microceratina* species (Crustacea, Ostracoda) from Christmas Island (Indian Ocean). *Journal of Micropalaeontology*, **23**: 49–59.

Nevill, J.C., Hancock, P.J., Murray, B.R., Ponder, W.F., Humphreys, W.F., Phillips, M.L. & Groom, P.K. 2010. Groundwater-dependent ecosystems and the dangers of groundwater overdraft: a review and an Australian perspective. *Pacific Conservation Biology*, **16**: 187–208.

Nistelberger, H., Gibson, N., MacDonald, B., Tapper, S-L. & Byrne, M. 2014. Phylogeographic evidence for two mesic refugia in a biodiversity hotspot. *Heredity*, **113**: 454–463.

O'Hara, T.D., Hugall, A.F., Cisternas, P.A., Boissin, E., Bribiesca-Contreras, G., Sellanes, J., Paulay, G. & Byrne, M. 2019. Phylogenomics, life history and morphological evolution of ophiocomid brittlestars. *Molecular Phylogenetics and Evolution*, **130**: 67–80.

Pepper, M., Fujita, M.K., Moritz, C. & Keogh, J.S. 2011. Palaeoclimate change drove diversification among isolated mountain refugia in the Australian arid zone. *Molecular Ecology*, **20**: 1529–1545.

Pepper, M., Sumner, J., Brennan, I.G., Hodges, K., Lemmon, A.R., Lemmon, E.M., Peterson, G., Rabosky, D.L., Schwarzkopf, L., Scott, I.A.W., Shea, G. & Keogh, J.S. 2018. Speciation in the mountains and dispersal by rivers: Molecular phylogeny of *Eulamprus* water skinks and the biogeography of Eastern Australia. *Journal of Biogeography*, **45**: 2040–2052.

Perez, K.E., Ponder, W.F., Colgan, D.J., Clark, S.A. & Lydeard, C. 2005. Molecular phylogeny and biogeography of spring-associated hydrobiid snails of the Great Artesian Basin, Australia. *Molecular Phylogenetics and Evolution*, **34**: 545–556.

Ponder, W.F. 2002. Desert springs of the Great Artesian Basin. In: *Proceedings of the meeting on spring-fed wetlands: important scientific and cultural resources of the intermountain region, May 2002, Las Vegas, NV*. (D.W. Sada & S.E. Sharpe, eds.), pp. 1–13. Desert Research Institute, Nevada, USA.

Ponder, W.F., Eggler, P. & Colgan, D.J. 1995. Genetic differentiation of aquatic snails (Gastropoda: Hydrobiidae) in artesian springs in arid Australia. *Biological Journal of the Linnean Society* **56**: 553–596.

Portik, D.M., Smith, L.L. & Bi, K. 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Molecular Ecology Resources*, **16**: 1069–1083.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**: 1864–1877.

Reilly, S.B., Stubbs, A.L., Karin, B.R., Bi, K., Arida, E., Iskandar, D.T. & McGuire, J.A. 2019. Leap-frog dispersal and mitochondrial introgression: phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. *Journal of Biogeography*, **46**: 757–769.

Rix, M.G., Cooper, S.J.B., Meusemann, K., Klopfstein, S., Harrison, S.E., Harvey, M.S. & Austin, A.D. 2017. Post-Eocene climate change across continental Australia and the diversification of Australian spiny trapdoor spiders (Idiopidae: Arbanitinae). *Molecular Phylogenetics and Evolution*, **109**: 302–320.

Sanger, F. & Coulson, A.R. 1975. A rapid method for determining sequencing in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, **94**: 441–446.

Schafer, A.B.A, Cullingham, C.I., Cote, S.D. & Coltman, D.W. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**: 4589–4621.

Shepherd, K.A., Waycott, M. & Calladine, A. 2004. Radiation of the Australian Salicornioideae (Chenopodiaceae) – based on evidence from nuclear and chloroplast DNA sequences. *American Journal of Botany*, **91**: 1387–1397.

Sniderman, J.M.K., Pillans, B., O'Sullivan, P.B. & Kershaw, A.P. 2007. Climate and vegetation in southeastern Australia respond to Southern Hemisphere insolation forcing in the late Pliocene-early Pleistocene. *Geology*, **35**: 41–44.

Sniderman, J.M.K., Woodhead, J.D., Hellstroma, J., Jordan, G.J., Drysdale, R.N., Tyler, J.J. & Porch, N. 2016. Pliocene reversal of late Neogene aridification. *Proceedings of the National Academy of Sciences*, **13***:* 1999–2004.

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S. & Soltis, P.S. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**: 4261–4293.

Stewart, J.R., Lister, A.M., Barnes, I. & Dalén, L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B*, **277**: 661–671.

Stoch, F. & Galassi, D.M.P. 2010. Stygobitic crustacean species richness: a question of numbers, a matter of scale. *Hydrobiologia*, **653**: 217–234.

Storey, A.W., Halse, S.A., Shiel, R.J. & Creagh, S. 2011. Aquatic fauna and water chemistry of the mound springs and wetlands of Mandora Marsh, north-western Australia. *Journal of the Royal Society of Western Australia*, **94**: 419–437.

Svenning, J-C. 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters*, **6**: 646–653.

Taiti, S., Ferrara, F. & Iliffe, T.M. 1995. A new species of *Haloniscus* Chilton, 1920 from New Caledonia (Isopoda: Oniscidea). *Crustaceana*, **68**: 321–328.

Taiti, S. & Humphreys, W.F. 2001. New aquatic Oniscidea (Crustacea, Isopoda) from groundwater calcretes of Western Australia. *Records of the Western Australian Museum*, **64**: 133–151.

Tapper, S-L., Byrne, M., Yates, C.J., Keppel, G., Hopper, S.D., Ven Niel, K., Schut, A.G.T., Mucina, L. & Wardell-Johnson, G.W. 2014. Prolonged isolation and persistence of a common endemic on

granite outcrops in both mesic and semi-arid environments in south western Australia. *Journal of Biogeography*, **41**: 2032–2044.

Thomson, R. & Barnett, S. 1985. Geology, geomorphology and hydrogeology. In: *South Australia's mound springs*. (J. Greenslade, L. Joseph & A. Reeves, eds.), pp. 3–26. Nature Conservation Society SA Inc., Adelaide, Australia.

Tomlinson, M. & Boulton, A.J. 2010. Ecology and management of subsurface groundwater dependent ecosystems in Australia – a review. *Marine and Freshwater Research*, **61**: 936–949.

Vandel, A. 1973. Les Isopodes terrestres (Oniscoidea) de la Mélanésie. *Zoologisches Verhandelingen,* **125**: 1–160.

Watts, C.H.S. & Humphreys, W.F. 2006. Twenty-six new Dytiscidae (Coleoptera) of the genera *Limbodessus* Guignot and *Nirripirti* Watts & Humphreys, from underground waters in Australia. *Transactions of the Royal Society of South Australia*, **130**: 123–185.

Williams, W.D. 1983. On the ecology of *Haloniscus searlei* (Isopoda, Oniscoidea), an inhabitant of Australian salt lakes. *Hydrobiologia*, **105**: 137–142.

Witt, J.D.S., Threloff, D.L. & Hebert, P.D.N. 2008. Genetic zoogeography of the *Hyalella Azteca* species complex in the Great Basin: rapid rates of molecular diversification in desert springs. In: *Late Cenozoic drainage history of the southwestern Great Basin and Lower Colorado River region: geological and biotic perspectives*. (R. Hershler, M. Reheis & D. Miller, eds), pp. 103–114, Geological Society of America Special Paper 439.

Wood, H.M., González, V.L., Lloyd, M., Coddington, J. & Scharff, N. 2018. Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). *Molecular Phylogenetics and Evolution*, **127**: 907–918.

Zhang, J., Chiodini, R., Badr, A. & Zhang, G. 2011. The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, **38**: 95–109.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**: 2869–2876.

CHAPTER 2:

Development and evaluation of a custom bait design based on 469 single-copy protein-coding genes for exon capture of isopods (Philosciidae: *Haloniscus*)

Statement of Authorship

Title of Paper	Development and evaluation of a custom bait design based on 469 single-copy protein-coding genes for exon capture of isopods (Philosciidae: *Haloniscus*)
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Intention to submit to *Molecular Ecology Resources*

Principal Author

Name of Principal Author (Candidate)	Danielle Stringer		
Contribution to the Paper	Organised field trip and collected specimens, assembled *Haloniscus* transcriptome, conducted orthology assignment and bait design, completed all laboratory work and generated data, conducted bioinformatics processing and analyses, interpreted results, wrote manuscript and compiled figures and tables.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	8/7/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author (Candidate)	Terry Bertozzi		
Contribution to the Paper	Assembled three isopod transcriptomes, guided post-processing bioinformatics, wrote some of the crucial scripts for data analysis, and critically reviewed manuscript.		
Signature		Date	8.vii.19

Name of Co-Author	Karen Meusemann		
Contribution to the Paper	Provided guidance and assisted with orthology assignment, conducted MARE analysis, and critically reviewed manuscript.		
Signature		Date	28 June 2019

Name of Co-Author	Steven Delean
Contribution to the Paper	Wrote R scripts and assisted with data analysis, and critically reviewed manuscript.
Signature	Date 23/7/19

Name of Co-Author	Michelle Guzik
Contribution to the Paper	Supervised development of work, collected specimens, and critically reviewed manuscript.
Signature	Date 8/7/19

Name of Co-Author	Simon Tierney
Contribution to the Paper	Supervised development of work, provided pipeline scripts, guidance and important advice for transcriptome assembly, and critically reviewed the manuscript.
Signature	Date 28 JUNE 2019

Name of Co-Author	Steven Cooper
Contribution to the Paper	Supervised development of work, specimen collection, and critically reviewed manuscript.
Signature	Date 8/7/2019

Name of Co-Author	Andreas Zwick
Contribution to the Paper	Provided two assembled transcriptomes and critically reviewed aspects of the manuscript.
Signature	Date 28/6/19

Name of Co-Author	Andrew Austin
Contribution to the Paper	Supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

Name of Co-Author	Christoph Mayer		
Contribution to the Paper	Provided pre-release version of BaitFisher software and important advice for bait design, and critically reviewed the manuscript.		
Signature		Date	July 2nd, 2019

Development and evaluation of a custom bait design based on 469 single-copy protein-coding genes for exon capture of isopods (Philosciidae: *Haloniscus*)

Running title: Exon capture bait design for *Haloniscus* isopods

Danielle N. Stringer[1], Terry Bertozzi[2], Karen Meusemann[3,4,5], Steven Delean[6], Michelle T. Guzik[1], Simon M. Tierney[1,7], Christoph Mayer[5], Steven J.B. Cooper[1,2], Andreas Zwick[4], Andrew D. Austin[1]

[1] Australian Centre for Evolutionary Biology and Biodiversity, School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia
[2] Evolutionary Biology Unit, South Australian Museum, North Terrace, Adelaide, SA, Australia
[3] Evolutionary Biology and Ecology, Institute for Biology I, University of Freiburg, Freiburg, Germany
[4] Australian National Insect Collection, CSIRO National Research Collections Australia, Acton, ACT, Australia
[5] Center for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, Bonn, Germany
[6] School of Biological Sciences and the Environment Institute, The University of Adelaide, Adelaide, SA, Australia
[7] Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW, Australia

Corresponding author email: danielle.stringer@adelaide.edu.au

Abstract

Transcriptome-based exon capture approaches, along with next-generation sequencing, are allowing for the rapid and cost-effective production of extensive and informative phylogenomic datasets from non-model organisms for phylogenetics and population genetics research. These approaches generally utilise a reference genome to infer the intron-exon structure of targeted loci, and preferentially select longer exons. However, in the absence of an existing and well-annotated genome, we applied this exon capture method directly, without initially identifying intron-exon boundaries for bait design, to a group of highly diverse *Haloniscus* (Philosciidae) and more divergent paraplatyarthrid and armadillid isopods, and examined the performance of our methods and bait design for phylogenetic exploration. Here, we inferred an isopod-specific set of single-copy protein-coding genes and produced a custom bait design using recently published software to capture target regions from 469 genes, and analysed the resulting sequence data with a mapping approach and newly-created post-processing scripts. We successfully recovered a large and informative dataset comprising both short (<100 bp) and longer (>300 bp) exons

with high uniformity in sequencing depth, as well as a considerable amount of flanking sequence data. We were further able to effectively capture sequence data from historical museum specimens as well as outgroup taxa, and efficiently pool samples prior to capture. Finally, our well-resolved phylogenies highlight the overall utility of this methodological approach and custom bait design, offering enormous potential for application to future isopod, and related crustacean, molecular studies.

Keywords: DNA enrichment, genomics, Isopoda, phylogenomics**,** transcriptome-based exon capture

Introduction

Phylogenetic and population genetic research on non-model organisms has largely relied on a limited selection of readily available genetic markers to address fundamental, and often difficult, evolutionary questions. Recent molecular studies have, nonetheless, highlighted that a significantly larger number of independently evolving loci are required to produce robust, well-resolved phylogenies and explore complex phylogenetic and biogeographic scenarios (Leaché & Rannala, 2011; Salichos & Rokas, 2013; Wortley, Rudall, Harris, & Scotland, 2005). Continual advances and improvements in high-throughput next-generation sequencing (NGS) technologies are now helping to alleviate this issue by allowing for the rapid and cost-effective production of large molecular datasets for phylogenetic, systematics and population genetic investigations (Lemmon & Lemmon, 2013; McCormack, Hird, Zellmer, Carstens, & Brumfeld, 2013; Peters et al., 2017). A variety of approaches are now available to help produce these substantial datasets, with the majority classified as reduced representation sequencing, where sets of preferably orthologous loci (clusters of orthologous groups) from a subset of the genome are obtained across taxa of interest (Bragg, Potter, Bi, & Moritz, 2016). Reduced representation approaches include RAD sequencing that targets unspecified loci associated with restriction enzyme sites (Miller, Dunham, Amores, Cresko, & Johnson, 2007), and those targeting highly specific loci with designed DNA and RNA baits (also termed probes), which are homologous to targeted DNA regions, including ultra-conserved element (UCE) sequencing (Faircloth et al., 2012), anchored hybrid enrichment (AHE) (Lemmon, Emme, & Lemmon, 2012; Wolfe et al., 2019), and transcriptome-based exon capture (Bi et al., 2012).

Transcriptome-based exon capture, in particular, employs the transcript sequences from orthologous groups (OGs or orthologues) to infer custom baits, which target protein coding exons across taxa, and is especially useful for acquiring sequence data from non-model organisms lacking reference genomes (Bi et al., 2012; Bragg et al., 2016; Dietz, Dömel, Leese, Mahon, & Mayer, 2019; Hugall, O'Hara, Hunjan, Nilsen, & Moussalli, 2016; Klopfstein et al., 2019; McCartney-Melstad, Mount, & Shaffer, 2016; O'Hara et al., 2019; Teasdale, Köhler, Murray, O'Hara, & Moussalli, 2016). This method can further be used to procure genomic data from historical museum specimens, typically comprising degraded DNA making

traditional sequencing techniques challenging, since it targets short DNA fragment sizes (100–400 bp) (Abdelkrim et al., 2018; Bi et al., 2013; Wood, González, Lloyd, Coddington, & Scharff, 2018). However, these methods commonly utilise a closely related genome to help identify intron-exon boundaries, and to preferentially select longer exon regions (>120 bp) in bait design (Bi et al., 2012; Bragg et al., 2016). Long exons are generally targeted because they exceed the length of standard baits (typically 120 bp), permitting increased tiling (overlap) to improve capture efficiency (Bi et al., 2012; Mayer et al., 2016). However, intron-exon identification becomes problematic when genomic references are too divergent from the species of interest due to issues aligning the exons, and since intron-exon structure may not always be preserved in distantly related taxa (Roy, Fedorov, & Gilbert, 2003).

In these instances, transcriptome sequences can be used directly to infer orthologues and design baits, precluding the need to differentiate exons using a genome reference *a priori*. A recent study by Portik et al. (2016) effectively employed this transcriptome-based exon capture method to generate a large and informative phylogenomic dataset across divergent frog lineages. This approach enabled exons of various lengths to be captured (because bait tiling may span multiple short exons), together with highly variable non-coding flanking sequences. Nevertheless, very few studies have used this direct approach, and further baseline information and empirical data, as well as detailed and reproducible bioinformatic methods, are required for the successful design of future capture experiments. In this study, we assess the performance and efficiency of transcriptome-based exon capture for the non-model isopod genus *Haloniscus* Chilton, 1920, and the application of our bait design across more distantly related outgroup isopod species for phylogenetic analysis. Orthologue and bait sets with broad taxonomic applicability are of significant interest in phylogenetics, specifically since this allows for consistency and comparison across multiple studies (Mayer et al., 2016; Teasdale et al., 2016; Wolfe et al., 2019).

Earlier phylogenetic studies on *Haloniscus* from Australian groundwater-dependent ecosystems have revealed extensive levels of diversity and short-range endemism for the genus, and further proposed that this isopod group represents a relictual fauna with an extraordinarily complex evolutionary history (Cooper, Saint, Taiti, Austin, & Humphreys, 2008; Guzik et al., 2019; Murphy, Guzik, Cooper, & Austin, 2015). Molecular datasets, however, have been restricted to either one mitochondrial (*cytochrome "c" oxidase subunit I* (*COI*)) or two genes (*COI* and *18S rRNA*), encouraging poor topological resolution for internal branches. Therefore, questions concerning the origins and biogeographic history of the genus remain unresolved and require further exploration with additional independent markers. The selection of a large orthologue gene set and development of a bait design targeting these *Haloniscus*, as well as distantly related isopods, will promote a better understanding of their evolution and the relationships between species. Nonetheless, there are currently no substantial multi-gene phylogenomic resources

available for this isopod genus, with exon capture studies, as well as appropriate orthologue sets and bait designs targeting a large number of loci, non-existent for any members of the Order Isopoda.

We, therefore, aimed to produce an effective and thorough methodological framework for: inferring an isopod-specific set of single-copy protein-coding orthologues from newly produced transcriptomes, filtering the loci for putative phylogenetic informativeness, developing a custom bait design to capture exons from several hundred loci, conducting all laboratory-based protocols, pooling selections, and for processing the capture sequence data using a mapping approach rather than assemblies (providing all bioinformatics scripts). We used the transcript sequences for bait design without first predicting exon boundaries due to the lack of an existing, closely related, and well-annotated reference genome. The overall success and effectiveness of our baits was evaluated by investigating: i) the number and length of exons, ii) sequencing depth (coverage per base) for targeted exons, iii) the percentage of missing data across exons incorporated into the final alignment, and iv) the applicability of the baits across divergent outgroups, including Paraplatyarthridae and Armadillidae. We further aimed to examine the influence of specimen preservation age and pool sizes prior to both capture and sequencing on depth of coverage, as well as the effect of pooling sizes on raw sequencing yield and PCR duplication. The use of older and poorly preserved samples is of significant interest in phylogenetics of non-model taxa as specimen re-collection, particularly of rare (or now extinct) lineages and the associated expenses can be prohibitive, while pooling improves sequence capture efficiency and reduces overall costs. Finally, we make available our assembled transcriptomes, selected OGs, bait design, concatenated alignments, and automated post-processing scripts for a completely reproducible framework, without the need for outsourcing to external providers.

Materials and methods

The methodological pipeline for this study was completed in eight major steps (Fig. 1): transcriptome sequencing; contig assembly; transcript assignment to OGs; an estimation of the putative phylogenetic informativeness of OGs and selection of OGs for downstream analyses; bait design and filtering; DNA extraction of preserved specimens for exon capture, library preparation of each sample and pooling; sequence capture laboratory procedures and Illumina MiSeq sequencing; and capture data processing, evaluation and phylogenetic analyses. These steps are outlined in detail in the subsequent sections. Custom Perl and Linux shell scripts were employed in the post-processing of the capture data (available online on Bitbucket at https://bitbucket.org/tbertozzi/scripts/src/master/ unless otherwise specified), with all analyses run on the University of Adelaide's Phoenix High Performance Computing Facility.

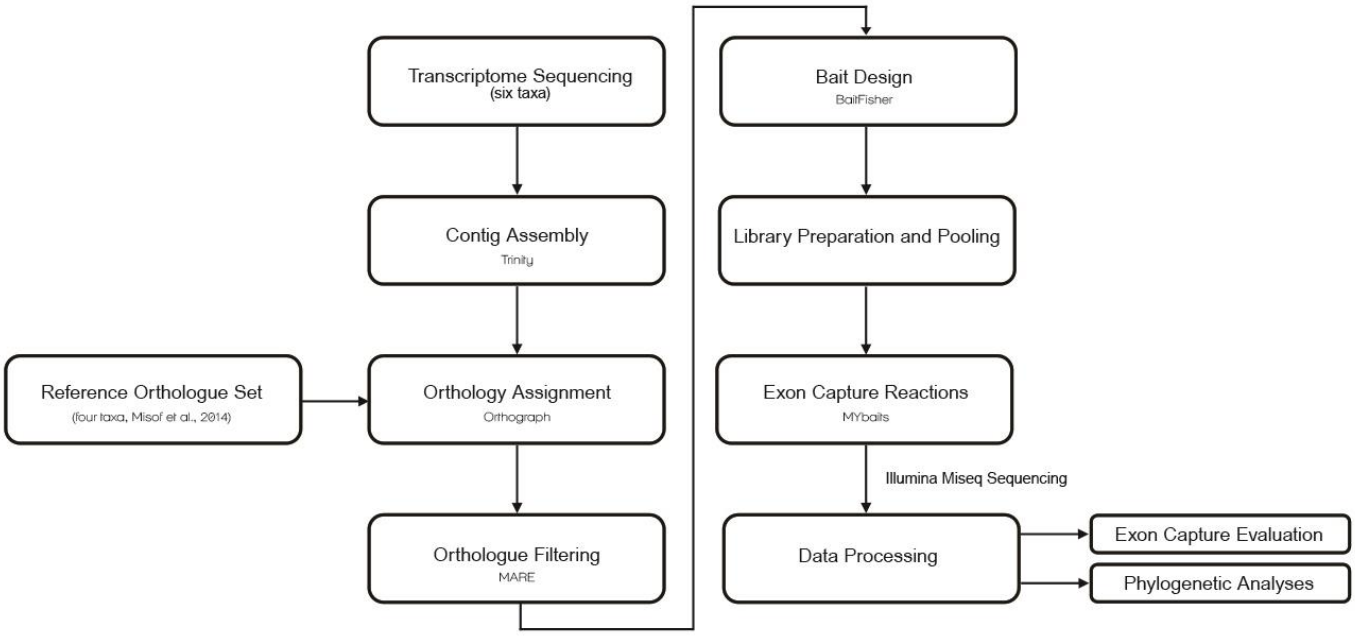


Figure 1: Schematic overview detailing the methodological framework used for orthology assignment, bait design, and sequence data generation. See text for further information.

Transcriptome sequencing and contig assembly

Six isopod species were selected for transcriptome sequencing, including one undescribed *Haloniscus* species (Philosciidae), and five more distantly related species: *Paraplatyarthrus subterraneus* Javidkar & King, 2015 (Paraplatyarthridae), *Paraplatyarthrus* sp. (Paraplatyarthridae), *Porcellionides pruinosus* (Brandt, 1833) (Porcellionidae), *Ceratothoa* sp. (Cymothoidae), and *Armadillidium vulgare* (Latreille, 1804) (Armadillidiidae) (see Table S1 for details). Specimens were preserved in RNAlater (Qiagen) and total RNA was successively extracted from whole isopod bodies with an RNeasy Plus Micro or Mini Kit (Qiagen) according to the standard protocol for tissue. RNA (pooled from multiple specimens in some cases due to very small individuals, Table S1) was quantified using a Quantus Fluorometer (Promega). Double-stranded cDNA was synthesised and PCR-amplified utilising the SMARTer cDNA Synthesis and Advantage 2 PCR Kits (Clontech), with PCR optimisation procedures verified using agarose gels. cDNA libraries were sent to the Australian Genome Research Facility (AGRF) in Adelaide, South Australia for *Haloniscus* sp., *Paraplatyarthrus* sp., *Pa. subterraneus* and *Po. pruinosus* and GATC (Eurofins Genomics) in Constance, Germany for *A. vulgare* and *Ceratothoa* sp. to be sequenced on the Illumina HiSeq2000 platform with TruSeq adapters, generating 100 bp paired-end reads.

For the *Haloniscus* sp. assembly (conducted by DNS), raw RNA-seq reads were quality controlled with FastQC v0.11.4 (Andrews, 2010), and transcripts were filtered and trimmed in Cutadapt v1.1 (Martin, 2011) to remove low quality reads (Phred scores <30), Illumina TruSeq barcoded adapters, SMARTer

adapters, poly-A tails, and sequences less than 25 bp. Transcripts were then *de novo* assembled using Trinity v2013-08-14 with default settings (Grabherr et al., 2011; Haas et al., 2013). Assembled contigs were quality assessed using Bowtie (Langmead, Trapnell, Pop, & Salzberg, 2009), which aligns contigs back against raw reads to specify the proportion of proper paired reads obtained. The five remaining isopod transcriptomes were contributed to this project as part of a collaboration, and were processed by co-authors TB and AZ using the methods detailed in Supplementary File S1.

Orthology assignment

We targeted single-copy, protein-coding orthologues, as recognised by The Hierarchical Catalogue of Orthologs (OrthoDB: http://www.orthodb.org/), from a previously published orthologue set based on 12 reference arthropod species (Misof et al., 2014: Table S3). This reference set, built on OrthoDB v5.0 (Waterhouse, Zdobnov, Tegenfeldt, Li, & Kriventseva, 2011), consisted of 1,478 orthologues from the crustacean *Daphnia pulex* Leydig, 1860, the arachnid *Ixodes scapularis* Say, 1821, as well as 10 hexapod species for which complete genomes and official gene sets were readily available. We optimised this reference orthologue set for our study by reducing the number of species to four, *D. pulex*, *I. scapularis*, the red flour beetle *Tribolium castaneum* (Herbst, 1797), and the termite *Zootermopsis nevadensis* Hagen, 1853, removing the remaining hexapods to lessen the emphasis on insects.

We inferred gene orthology in our six *de novo* sequenced and assembled transcriptome libraries with a pre-release of Orthograph (beta4.1 available at https://mptrsen.github.io/Orthograph/, Petersen et al., 2017) that assigned transcripts to orthologues in the reference set outlined above. Orthograph is a graph-based approach that utilises profile hidden Markov models to map transcripts to user-provided orthologous groups, followed by a best reciprocal hit (BRH) search of all candidate hits against the full official gene set of each reference species. We used Orthograph defaults except for the following: the maximum number of blast searches and blast hits=50, minimum transcript length=25, and also enabled the extension of the open reading frame (ORF) with a minimum overlap of 30% (extend-orf=1 and orf-overlap-minimum=0.3). The transcript sequences, where the orthologue criteria were fulfilled, were summarised into separate files per OG, and two reference species, *T. castaneum* and *I. scapularis,* were removed entirely since the majority of best reciprocal hits were to *D. pulex* and *Z. nevadensis*. Internal stop codons, along with Selenocysteine "U", were masked with "X" (and "NNN" on a nucleotide level) using a custom Perl script (provided with the Orthograph package). For downstream analyses, we used only OGs that consisted of hits to all six transcriptomes, totalling 531 protein-coding single-copy genes. Amino acid sequences for each OG were aligned with MAFFT v.7.220 (Katoh & Standley, 2013) and the corresponding nucleotide multiple sequence alignments were inferred using a modified version of the

software PAL2NAL v14.1 (Suyama, Torrents, & Bork, 2006, see Misof et al. (2014) for further details on the modification) with the amino acid alignments as blueprints.

Assessing phylogenetic informativeness of OGs

We applied the matrix reduction program MARE (Misof et al., 2013) v0.1.2-rc (Meyer, Meusemann, & Misof, 2011) to determine (putative phylogenetic) informativeness of all 531 OGs, and then assessed which of those OGs revealed the highest informativeness (i.e. information content (IC)). MARE utilises extended geometry quartet mapping to infer informativeness (or the "tree-likeness": the number of resolved quartets divided by the number of all quartets drawn for each partition or an OG in this study) from amino acid alignments in a user-provided supermatrix. The analysis yielded a reduced (optimised) matrix of a smaller size, but with an increased overall IC. Here, we retained only OGs with an IC >0.5 (479 OGs in total) for use in downstream analyses.

Bait design

The dataset, comprising 479 aligned OGs each with all eight reference species (the six transcriptomes, together with *D. pulex* and *Z. nevadensis*) on a nucleotide level, was used to design baits for targeted sequence capture with the software package BaitFisher (Mayer et al., 2016). BaitFisher consists of two programs: BaitFisher v1.2.7 and BaitFilter v1.0.4. BaitFisher infers baits using the nucleotide sequence information of target loci from reference species in a multiple sequence alignment, thereby targeting a diverse variety of species. This software generates all potential bait designs, which may be suitable for enriching a particular gene (or exon of a gene) with a user-specified tiling design. This output may then be passed to BaitFilter to select a more specific bait set by choosing the optimal start position for a given tiling design, by either minimising the number of baits required to capture a target locus, or by maximising the number of nucleotide sequences from which baits were inferred.

For bait design, we specified a bait length of 120 bp and a tiling design of seven baits spanning a total region of 300 bp, with a bait offset every 30 bp. The clustering threshold was also set to 0.15. BaitFisher removed 10 OGs from the design that were not suitable for bait construction as the sequences were either too short (<300 bp) or a suitable bait region was not detected since the region likely consisted of too many gaps or Ns and BaitFisher could not place a full bait region within the OG alignment. It was possible to include a reference genome and annotation file at this stage to split the cDNA sequences into known exons. Nevertheless, due to the lack of a closely related isopod genome at the time of bait design, we generated the baits directly from the transcriptome multiple sequence alignments. The bait design was optimised with BaitFilter by maximising the number of sequences, which resulted in 15,053

baits for 24,258 sequences (37.95% of baits saved, with respect to a bait design not generating baits for clusters of target sequences, but for all target sequences). Custom RNA baits were manufactured by MYcroarray (now Arbor Biosciences, Ann Arbor, MI, U.S.A.) for use with a MYbaits (v3) 12 reaction kit.

Library preparation and pooling

Genomic DNA was extracted from 36 whole specimens (31 *Haloniscus* taxa, representing the majority of known lineages as inferred using *COI* mitochondrial sequence data, with *H. anophthalmus* unlikely to actually belong to *Haloniscus* (Cooper et al., 2008; Guzik et al., 2019), and five further isopod species, see Table S2) using the Gentra® Puregene® DNA Purification protocol (Gentra Systems Inc.) according to the manufacturer's protocol. An additional *Haloniscus* taxon from Windimurra (Yilgarn calcrete, WA) was included using pooled DNA extracts from three individuals (Table S2). The DNA was quantified by Quantus Fluorometer (Promega) with the QuantiFlour dsDNA System Kit (Promega), and each sample was diluted to 1–10 ng/µL (reliant on the initial DNA concentration) in 100 µL of molecular grade water. A Bioruptor Pico (Diagenode) was employed to shear the DNA for 1–4 min using 30 s on/30 s off cycling. Each sonicated sample was then analysed by electrophoresis on an Agilent 2200 TapeStation (Agilent Technologies) to determine whether fragments were appropriately sized (average size of 300–500 bp) for later sequencing. For samples consisting of a broad fragment size distribution, a size selection step was completed with the SPRI bead method and polyethylene glycol to remove fragments less than 150 bp (protocol outlined by Li, Hofreiter, Straube, Corrigan, & Naylor, 2012).

Genomic libraries were then prepared following Meyer and Kircher (2010), with some modifications to the indexing PCR reaction. A unique combination of i7 and i5 indexes (1–10 from Meyer & Kircher, 2010, and 1–23 from Glenn et al., 2016) was added to each library in 25 µL reactions containing 1.5 µL H_2O, 12.5 µL KAPA HiFi Taq Ready Mix (2X), 0.5 µL of each indexing primer, and 10 µL of library (the remaining library retained as a back-up). Thermal cycling conditions involved an initial denaturation step at 98 ˚C for 45 s, then 18 cycles of 98 ˚C for 15 s, an annealing temperature of 65 ˚C for 30 s, and an elongation of 72 ˚C for 60 s. A final elongation phase of 72 ˚C for 10 min completed the reaction. Libraries were purified and the concentrations measured by a Qubit Fluorometer (Life Technologies) and qPCR amplification (KAPA Library Quantification Kit, Illumina). The resulting 25 µL of amplified library product had a concentration of at least 10 ng/µL, but the final results extensively varied due to starting concentration and sample quality.

These library preparation steps were completed across three rounds of laboratory work. For the first run of eight libraries (or samples), we aimed to optimise the number of libraries which could be pooled

prior to exon capture. The eight libraries were separated into four distinct pools (one pool comprising one library, one with two libraries, one with three libraries, and one with four libraries), where one library (BES 18659.1) was split across three different pools (each with different dual indexes, making 10 libraries total) to determine whether increasing the number of libraries within a single pool would influence the final read coverage obtained. For the second round, we combined eight libraries across two pools of four libraries, and for the final round, we split 22 libraries across six pools: four pools of four libraries and two pools with three libraries.

Exon capture reactions and sequencing

Pooled libraries were concentrated down to 7 µL using a CentriVap DNA Concentrator (Labconco) for sequence capture. MYbaits capture reactions were performed following the v3.01 manual, with heat denatured concentrated library pools combined with the designed baits and universal blocking oligos (included in the MYbaits kit), and hybridised for approximately 16–20 h at 65 ˚C. Reactions were then purified using Dynabeads MyOne Streptavidin C1 beads (Life Technologies) and post-capture products were amplified using KAPA HiFi DNA Polymerase (Kapa Biosystems) with the following protocol: 98 ˚C for 2 min, followed by 12 cycles of 98 ˚C for 20 s, 60 ˚C for 30 s, and 72 ˚C for 30 s, and final extension of 5 min at 72 ˚C. Pools were purified with 90 µL of AMPure XP beads (Agencourt), resuspended in 30 µL of elution buffer, and quantified with the Qubit Fluorometer (Life Technologies) and/or a standard quantitative PCR run with the LightCycler 96 Real-Time PCR System (Roche Diagnostics) for equimolar pooling. The fragment size distribution for each pool was, additionally, visualised on the TapeStation. Following the first round of capture, the four pools (with 10 libraries total) were combined in equimolar ratios and sequenced on the Illumina MiSeq platform with 300 bp paired-end reads. For the second (two pools each containing four libraries) and third capture rounds (six pools total containing four pools of four libraries and two pools with three libraries = 22 libraries), equimolar pools were sequenced on the Illumina MiSeq platform, but, in these latter capture runs, 150 bp paired-end reads were obtained due to the low average fragment size (<300 bp) in the final pools. Illumina sequencing of the captured DNA libraries was conducted by AGRF in Adelaide, South Australia.

Exon capture data processing

Raw sequence reads were quality-checked with FastQC v0.11.4 (Andrews, 2010) and filtered using the BBDuk v35.92 software package (BBTools: https://sourceforge.net/projects/bbmap/files) by trimming adapters, and removing low quality reads. Overlapping paired reads were merged using PEAR v0.9.10 (Zhang, Kobert, Flouri, & Stamatakis, 2014) to avoid inflated coverage estimates. For each sample, the resulting clean reads were mapped to the *Haloniscus* transcript orthologues used for bait design with

BWA v.0.7.15 (Li & Durbin, 2009) and SAMtools v1.3.1 (Li et al., 2009). The targeted orthologues were concatenated into one continuous sequence of all 469 targets, each separated by a string of 1000 Ns using a custom script, catFasta.pl. The script provides the option of simultaneously generating a BED4 file, which defines the start and end position of each target, the target sequence length and the name of each target orthologue. Output BAM files were assessed with the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir, Robinson, & Mesirov, 2013). Since an annotated reference genome could not be used during bait design to determine the positions of intron-exon boundaries, reference targets (or exons) were split manually (with BAM alignments in IGV, see Fig. S1) to reflect the boundaries. The reads were then mapped to the revised reference, which was similarly generated using the concatenation script, and duplicate reads were removed with Picard tools v2.2.4 (http://broadinstitute.github.io/picard/). Sequencing depth (coverage per base) files were then produced with BEDTools v2.25.0 (Quinlan & Hall, 2010). See Supplementary File S2 for an automated script to complete the above processes.

Variant calling was performed using FreeBayes v1.0.2 (Garrison & Marth, 2012) after initially trialling SAMtools v1.3.1 with BCFtools v1.4.1 (Li et al., 2009) and HaplotypeCaller in GATK v.3.7 (McKenna et al., 2010). BCFtools frequently reported lower values for variant sequencing depth than expected, i.e. differing from those calculated with BEDTools and as seen in alignments with IGV, which led to issues in later steps when filtering based on depth. The GATK variant caller, however, provided higher depth results but the VCF file contained considerably less variants, even after adjusting parameters. Overall, VCF files produced by FreeBayes included expected variant numbers and sequencing depth values to those in the BAM alignments. Complex polymorphisms in FreeBayes output files were decomposed to individual SNP and indel calls using vcfallelicprimitives in vcflib (https://github.com/vcflib/vcflib), and low depth variants (i.e. <10x coverage) were removed with vcffilter (vcflib). Heterozygous calls were filtered using a custom Perl script, filterVCF.pl, based on a minimum minor allele frequency of 0.2. For heterozygotes with a minimum alternate allele frequency below the threshold, the site was retained as homozygous to the reference; however, heterozygotes with a minimum reference allele frequency below threshold were changed to homozygous alternate alleles (GT field modified to 1/1). Consensus sequences were then generated using the Perl script, applyVariants.pl. The script produced consensus sequences inferred by the mapped reads by applying variants in a VCF file to the reference sequence used in mapping. The script also masked bases below 10x coverage using the 'per base' coverage files generated with BEDTools, and additionally included IUPAC ambiguity codes for heterozygous sites. See Supplementary File S3 for an automated script to complete the above processes.

A custom Perl script, groupTargets.pl, then used the fasta files created with applyVariants.pl to group the same target sequence from different isopod samples into individual files ready for alignment. The Bash script, runMuscle.sh, along with MUSCLE v3.8.31 (Edgar, 2004), was used to align sequences in

each target file produced with groupTargets.pl. Potential paralogues were identified and removed from the target alignment files based on an elevated proportion of heterozygous sites (>3%) with the custom script, FilterMergedLoci.pl. Sequences including an excess of variable sites were replaced with a string of Ns equal in length to the original sequence.

Exon capture evaluation

To examine capture efficiency, sequencing depth (coverage per base) values were calculated for each position (i.e. base pair) along the mapped transcripts for all samples with BEDTools (discussed above). The per base pair coverage estimates corresponding to each of the 469 target orthologues were plotted for all 40 samples (ingroup and outgroup species), with intron-exon boundaries delineated by vertical lines. Using these outputs, the median sequencing depth values were calculated across samples for all exons within each orthologue, and separated based on the sequencing batch (runs 1–3). Results were summarised for the 'targeted' exons (selected from exons with the highest median sequencing depths for each orthologue) from 50 randomly-chosen orthologues. Average differences in median sequencing depth between runs were estimated using a generalised linear mixed model fitted to the data on a log link scale and negative binomial variance distribution. Individual exon and specimen identifiers were integrated as random effects to account for average differences in sequencing depth values between each of these factors. Marginal mean sequencing depth (with 95% confidence intervals) was estimated for each run, and contrasts were used to infer differences in sequencing depth across the runs.

Differences in median sequencing depth for exons of a random subset of 50 targeted orthologues were examined. Pool sizes (with 1–4 samples) prior to exon capture, specimen preservation age (number of years since collected and preserved in 100% ethanol), raw paired-end data yield, and the percentage of missing data across exons within the threshold 50 concatenated alignment (see below for details) were used as additional covariates. The percentage of PCR duplication amongst reads for each sample was similarly examined against pool size prior to capture, with the amount of duplication calculated by dividing the number of duplicate reads (obtained with Picard tools as described above) by the total number of raw paired-end reads (as filtered reads were merged using PEAR). We did not formally test for the relationships with specimen preservation age or pooling sizes prior to capture. The distribution of specimen preservation ages varied considerably between runs, sometimes over a short range, and, for some runs, a single sample differed substantially in age from the remaining samples in that run. For pool sizes, some runs consisted of merely one or two different pooling selections, while additional runs comprised pools with a larger range of sizes.

The exon sequencing depth uniformity across samples, and among sequencing runs, was assessed by calculating the median absolute deviation and robust coefficient of dispersion (the median absolute deviation divided by the median) for targeted exons from the 50 randomly-selected orthologues, and plotting the results. This calculation examines the amount of variation in sequencing depth across all bases within exons. All processes for capture evaluation were conducted using R v3.6.0 (R Core Team, 2019) and are detailed in the included script (Supplementary File S4).

Phylogenetic analyses

The target alignments discussed previously were concatenated using a custom script, catAlignedLoci.pl, for phylogenetic analysis. The concatenation order was based on the BED4 candidate file, which was produced with the artificial reference prior to mapping. The script created a "candidate partition" file, including the target/exon boundaries, and further allowed for a threshold to be indicated to filter out targets that contained too many missing sequences (i.e. all Ns). Three distinct datasets were produced, each with differing thresholds for missing data: the first with a threshold of 25 (dataset A), which only included exon alignments with up to 25% missing sequences, the second with a threshold of 50 (dataset B), and the third dataset with a threshold of 75 (dataset C), which included exons from taxa with up to 75% of sequences missing. We employed PartitionFinder v2.1.1 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) to determine the most appropriate partitioning schemes for the three different datasets with the rcluster algorithm (Lanfear, Calcott, Kainer, Mayer, & Stamatakis, 2014) (parameters: models= all, model_selection= aicc, branchlengths= linked, rcluster-percent= 10.0, rcluster-max= 1000, raxml), and the exon partition files discussed above. Maximum likelihood (ML) phylogenies were then inferred using RAxML v8.2.10 (Stamatakis, 2014) with the deduced partitions, the GTRGAMMA model and 1,000 nonparametric bootstraps (remaining settings as default). PartitionFinder and RAxML were run on the CIPRES Science Gateway v3.3 (Miller, Pfeiffer, & Schwartz, 2010).

Results

Transcriptomes, orthology and bait design

An average of 24.1M (21.3M–27.9M) paired-end reads were sequenced for each transcriptome library, which assembled into approximately 6.2×10^4 contigs (4.0×10^4–1.37×10^5) (Table 1). The larger contig values acquired for *A. vulgare* and *Ceratothoa* sp. (Table 1), despite the comparable number of paired-end raw reads to the other transcriptomes, can be explained by the inclusion of many short contigs by IDBA-Tran, which remain ambiguous and are unlikely to be true transcripts. Searching for 1,478 single-copy OGs, similar numbers were inferred amongst our isopod transcriptomes (806–1272, see Table 1).

Nevertheless, only 531 OGs were present within all six transcriptomes and the two included reference genomes, *D. pulex* and *Z. nevadensis*. These 531 candidate OGs were filtered based on their putative phylogenetic informativeness using MARE: see a summary of results on the Figshare online repository (doi:10.25909/5d3678273b4f0). Overall, for only one OG, the IC was zero, while three OGs comprised an IC=1, with all quartets entirely resolved. 479 OGs (=90% of the original set of 531 OGs) were selected based on a specified IC threshold of >0.5 for bait design and exon capture. During bait design, 10 OGs were removed by BaitFisher (as outlined in Materials and methods), resulting in an optimised bait set, which targeted regions (after filtering with BaitFilter) from 469 genes. See the Figshare repository for the six assembled isopod transcriptomes (doi:10.25909/5d3674926d717), the final selected 469 OGs (doi:10.25909/5d3672cf76c28), and the bait nucleotide sequences (doi:10.25909/5d3548f059aed).

Table 1: Summary statistics for sequencing and *de novo* assembly of the six isopod transcriptomes used in orthology assignment.

Isopod species	**Pairs of raw reads**	***De novo* assembled transcripts**	**Number of identified orthologues**
Haloniscus sp.	21,354,576	43,455	942
Paraplatyarthrus sp.	24,795,047	40,461	895
Paraplatyarthrus subterraneus	21,896,830	46,114	1,011
Porcellionides pruinosus	23,941,206	37,368	806
Armadillidium vulgare	27,943,392	66,407	1,272
Ceratothoa sp.	24,786,465	137,713	1,260

Exon capture data

An average of 833,844 (73,241–2,898,504) paired-end reads were sequenced for all 40 samples, with run 1 (10 pooled samples) averaging 1,656,743 (764,956–2,898,504) PE reads, run 2 (8 pooled samples) averaging 1,261,684 (747,169–2,058,751) PE reads, and run 3 (22 pooled samples, including outgroup taxa) averaging 304,221 (73,241–717,722) PE reads (see Table 2). Following the removal of low quality reads and adapters, the percentage of clean paired-end reads retained ranged from ~92–98.5% (Table 2). Mapping the cleaned reads directly to the *Haloniscus* sp. transcript orthologues used for bait design revealed copious exon sequences of various lengths and their associated non-coding (intron) flanking sequences (see e.g. Fig. S2 for example). Since the bait region was constrained to a length of 300 bp, the complete transcript sequences were not generally captured (see sequencing depth summary plots, Fig. S3). Nonetheless, a total of approximately 1,150 exons (median: 798 exons captured across distinct samples) were captured across all the targeted orthologues and samples, with a median of two exons captured per orthologue (range: 1–4 exons), and only nine OGs were not captured across any samples. The length of exons captured varied substantially, ranging from 15–2,013 bp, with a median length of 153 bp for individual captured exons (Fig. S4).

Table 2: Exon capture sample statistics. Pool size 1 refers to the number of individuals pooled prior to exon capture, and pool size 2 is prior to final sequencing.

Specimen ID	Sequencing ID	Run	Ingroup or outgroup	Raw paired-end reads	Clean paired-end reads retained (%)	Duplication (%)	Missing data (%)	Preservation ages (years)	Pool size 1	Pool size 2
BES18775	27809	1	Ingroup	1,152,943	95.4	11.2	4.6	3.5	3	10
BES18774	27810	1	Ingroup	820,103	95.8	7.1	4.4	3.5	4	10
BES6573	27811	1	Ingroup	764,956	95.8	6.5	15.5	18	2	10
BES18645	27812	1	Ingroup	2,267,263	95.6	18.8	2.7	4	4	10
BES18659	27813	1	Ingroup	2,898,504	95.6	17.8	2.8	4	2	10
BES18659	27814	1	Ingroup	2,305,287	95.7	17.4	2.8	4	4	10
BES18659	27815	1	Ingroup	1,285,005	95.6	9.8	3.7	4	1	10
BES18601	27816	1	Ingroup	2,678,874	95.8	19.4	2.8	4	3	10
BES18754	27817	1	Ingroup	963,373	95.7	3.5	7.4	3.5	3	10
BES18644	27818	1	Ingroup	1,431,124	96.1	15.6	3.6	4	4	10
BES16434	28076	2	Ingroup	747,169	97.0	9.1	5.9	7.5	4	8
GAB01433	28077	2	Ingroup	875,801	97.3	20.0	7.9	10	4	8
GAB01616	28078	2	Ingroup	897,834	97.7	20.2	5.3	10	4	8
GAB00736	28079	2	Ingroup	2,058,751	94.3	29.0	4.5	11.5	4	8
GAB00764.1	28080	2	Ingroup	1,574,884	97.2	17.7	2.8	11.5	4	8
BES17062	28081	2	Ingroup	1,207,963	97.4	9.9	3.9	7	4	8
GAB01007.1	28082	2	Ingroup	1,696,723	97.3	24.0	3.7	11	4	8
GAB00764.1	28083	2	Ingroup	1,034,347	97.6	14.0	3.8	11.5	4	8
BES18773	1	3	Ingroup	282,187	96.9	0.8	37.5	3.5	4	22
BES18759.3	2	3	Ingroup	73,241	91.9	0.5	86.1	3.5	4	22
BES6655	3	3	Ingroup	135,325	94.2	1.4	83.7	18	4	22
BES16348	4	3	Ingroup	246,978	98.1	4.8	29.0	7.5	3	22
BES8623.1	5	3	Ingroup	230,473	95.2	0.6	70.0	18	4	22
BES16400.2	6	3	Outgroup	374,913	96.3	0.1	71.8	7.5	4	22
BES13246	7	3	Ingroup	620,284	95.9	0.7	19.4	12	4	22
BES13396	8	3	Ingroup	271,685	93.7	1.2	19.0	12	4	22
BES14385	9	3	Ingroup	109,362	98.4	0.7	52.8	12	4	22

Specimen ID	Sequencing ID	Run	Ingroup or outgroup	Raw paired-end reads	Clean paired-end reads retained (%)	Duplication (%)	Missing data (%)	Preservation ages (years)	Pool size 1	Pool size 2
BES13314	10	3	Ingroup	280,826	98.5	1.2	21.2	12	4	22
GAB00795	11	3	Ingroup	391,058	98.0	2.9	12.4	11.5	4	22
GAB00765	12	3	Ingroup	332,292	97.7	4.3	11.2	11.5	3	22
BES10201	13	3	Outgroup	89,157	98.4	0.1	99.4	16	4	22
BES6601.2	15	3	Ingroup	376,860	94.5	0.6	98.1	18	4	22
BES10410	16	3	Ingroup	509,729	92.1	1.8	40.5	15	4	22
BES6667.2	17	3	Ingroup	310,648	96.5	0.2	97.1	18	4	22
BES13452	18	3	Ingroup	382,589	97.8	0.2	44.6	12	4	22
BES8956, BES13133.1, BES13133.2	19	3	Ingroup	122,173	97.8	3.8	70.4	18	3	22
Ja243	20	3	Outgroup	346,847	98.3	0.6	66.8	8	3	22
B002	21	3	Outgroup	249,046	96.6	0.1	89.0	2.5	3	22
BES15525.10	22	3	Outgroup	239,462	98.4	1.3	38.8	9	4	22
BES15537.2	23	3	Outgroup	717,722	98.5	2.4	48.6	9	3	22

Sequencing depth, duplication and missing data

Sequencing depth (coverage per base) summary plots (Fig. 2A–D for examples, and Fig. S3) highlighted variation between exons, orthologues, and across individual samples. The most substantial changes in exon sequencing depth, nevertheless, occurred amongst samples from the three different sequencing runs ($\chi^2 = 99.7$, df = 2, $p < 0.0001$; Fig. 2E–F and Fig. S5). The third run comprised samples with an 11.6 fold lower median coverage (95% CI = 7.2, 18.9) across exons than the average median coverage values of runs 1 and 2 (which have equivalent median sequencing depth; ratio of coverage = 1.1, 95% CI = 0.5, 2.2). In addition, partitioning random variation in sequencing depth among orthologues versus among samples in the analysis (after accounting for differences resulting from sequencing run) revealed that 60% of the variation can be explained by individual sample differences, while only 15% is caused by gene to gene variation. Certain samples, including 27813, 27816 and 12, consistently encompassed the highest sequencing depth values across exons within their respective runs, whereas samples, such as 27809, 27811, 28076 and 15, repeatedly comprised some of the lowest values across exons (see Figs. S3 and S5). Run 1 samples further appeared to consist of greater variation in sequencing depth values within distinct exons than runs 2 and 3 (see Fig. 2E–F for example).

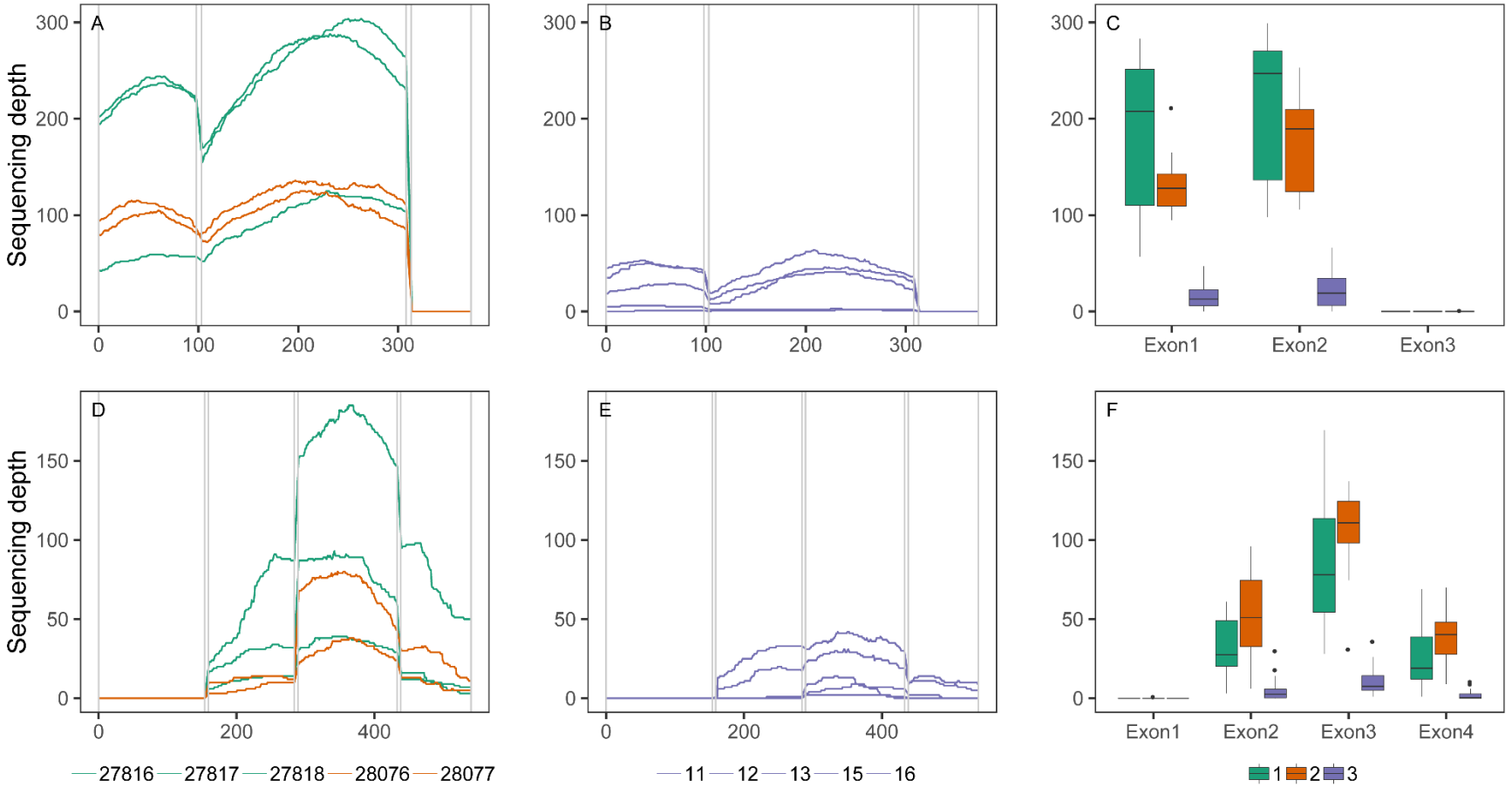


Figure 2: Examples of sequencing depth across orthologues EOG54MW8B (3 exons; upper row A–C) and EOG54MW8B (4 exons; lower row D–F). Examples from sequencing run 1 (specimens 27816, 27817, 27818) and run 2 (specimens 28076, 28077) are given in A and D, and from sequencing run 3 (specimens 11, 12, 13, 15, 16) in B and E. Introns are delineated by vertical lines in A, B, D and E. Boxplots in C and F highlight the distribution of sequencing depth for each exon within the orthologue grouped by the three sequencing runs (run 1 is indicated in green, run 2 in orange, and run 3 in purple). Horizontal lines in C and F are median sequencing depths, vertical lines show boxplot whiskers, and solid points represent outliers.

An assessment of pooling selections (pool size 1, Table 2) prior to exon capture experiments and their impact on median sequencing depth revealed no evident differences across particular pool sizes (Fig. S6). Although sample sizes for smaller pools were low and, as such, not rigorously tested here, a pool size of 2 revealed distinctly different results for sequencing depth across the two specimens included, with sequencing depth values constantly higher for one sample over the other (Fig. S6). Furthermore, the replicate samples from Laverton, WA (27813, 27814 and 27815) revealed consistently high depth values for 50 randomly-selected exons from distinct orthologues despite pool size; nevertheless, the sequencing depth values were persistently lower across exons for the single sample (27815) in pool 1 (Fig. S6), which is consistent with the number of raw paired-end reads acquired for each of these three samples (Table 2). A comparison of the median sequencing depth values across exons compared with the number of raw paired-end reads for each sample identified a largely positive linear relationship for the first two runs, but highlighted a less prominent pattern for run 3 (Fig. S7). Furthermore, the number of raw paired-end reads obtained was not correlated with pool sizes prior to capture (Fig. 3A).

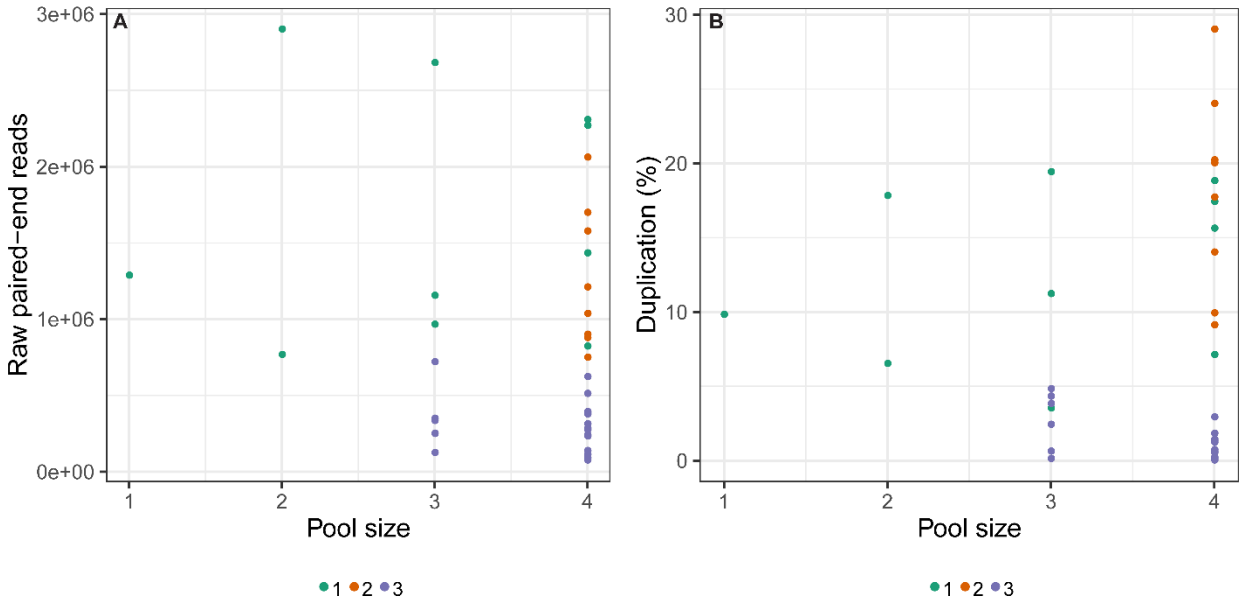


Figure 3: Plots of pooling sizes (1–4) prior to capture experiments against (A) raw paired-end reads and (B) duplication levels (%). Points are colour-coded by sequencing run (run 1 is indicated in green, run 2 in orange, and run 3 in purple).

Levels of PCR duplication were reasonably low within runs, ranging from 3.5–19.4% in run 1, 9.1–29% in run 2 and 0.1–4.8% in run 3, and did not differ substantially amongst ingroup and outgroup species (Table 2). An assessment of the relationship between percentage duplication levels versus pool size 1 revealed no noticeable correlation, with similar values and no apparent pattern across pools 1–4 (Fig. 3B). The amount of missing data in terms of coverage across exons (calculated using the threshold 50 concatenated alignment), however, differed considerably between sequencing run 3 and the batches with fewer pooled samples, ranging from 2.7–15.5% for run 1, 2.8–7.9% for run 2 and 11.2–99.4% for

sequencing run 3 (Table 2). The raw paired-end data yield for samples in run 3 did not appear to directly correspond to the amount of data acquired in the final alignment, with samples 11 and 12 comprising 391,058 and 332,292 raw paired-end reads and 12.4 and 11.2% missing data, respectively, and samples 15 and 17 consisting of a similar raw data yield, but a substantially larger amount of missing data (98.1 and 97.1%, respectively) (Table 2). The six outgroup samples, which are more distantly related to the reference used for mapping, revealed similar levels of missing data to some of the ingroup samples included in run 3 (Table 2).

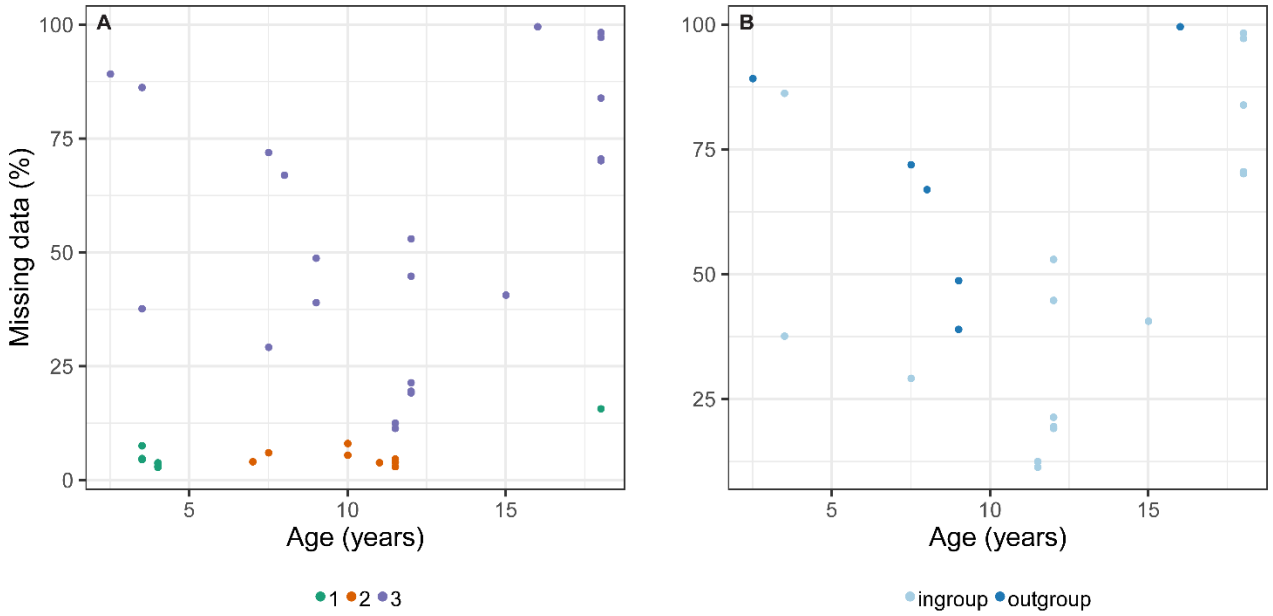


Figure 4: Preservation age of specimens (years) included in the exon captures against (A) missing data (%), colour-coded by sequencing run (run 1 is indicated in green, run 2 in orange, and run 3 in purple), and (B) missing data (%) for run 3 samples only, coloured by ingroup (light blue) and outgroup (dark blue) status.

We assessed whether the differences in missing data within sequencing runs (particularly run 3) were related to the preservation age of specimens. For run 1, the majority of samples were collected 3.5–4 years prior and included a similarly low percentage of missing data; however, one older sample, which was collected around 18 years prior, consisted of the highest level of missing data for the run at 15.5% (Fig. 4A, Table 2). For run 2, specimens ranged from 7–11.5 years old, and the level of missing sequence data for all samples corresponded to that of the more freshly collected and preserved specimens from run 1 (Fig. 4A, Table 2). For the third run, the percentage of missing data varied and did not appear to correlate with specimen preservation age as some of the more recently collected specimens contained a high amount of missing data; nevertheless, the oldest samples (16–18 years old) all comprised a large degree of missing data (Fig. 4A, Table 2). Run 3 samples were further compared on the basis of ingroup and outgroup status (Fig. 4B), with most outgroups consisting of a reasonably large amount of missing data (38.8–99.4%) despite preservation age. Furthermore, sequencing depth was assessed against age

for exons from 50 randomly-selected orthologues, with no apparent relationship (Fig. S8). This lack of a correlation was particularly evident for run 1, which exhibited a large degree of variation in median sequencing depth among the samples despite similar preservation ages, and highlighted (as above) the individual specimen effect on variation (Fig. S8).

Exon sequencing depth uniformity

Coefficient of dispersion (CD) across sequencing runs for the targeted (regions of the orthologue where overlapping baits were placed) exons for 50 randomly chosen orthologues was consistently higher and more variable in run 3 for sequencing depth (Fig. S9). The principally low CD values of roughly 10–15% for samples in runs 1 and 2, and 20–30% for samples in run 3, nonetheless, revealed largely consistent sequencing depth along positions within the majority of exons examined, indicating high uniformity in sequencing depth across the length of these exons (Fig. S9). These exons were typically short, between 94–234 bp long; however, CD values for the much smaller number of long exons (EOG5FXPQ3: 852 bp and EOG505QG8: 876 bp, Fig. S7) were considerably higher (>60%), signifying greater levels of variation and less uniformity.

Phylogenetic analyses

The three final alignments (containing up to 25% (A), 50% (B), and 75% (C) missing sequences) included sequence data for all 40 samples, and consisted of 420, 807, and 1026 exons, respectively (alignments available on Figshare at doi:10.25909/5d354467c921e). For dataset A, these exons were included from 335 of the 469 targeted OGs, constituting 88,402 bp of DNA sequence data. Datasets B and C consisted of exons from 440 (143,445 bp) and 451 (174,006 bp) OGs, respectively. The inferred ML trees revealed identical topologies (Fig. 5), with the majority of branches obtaining full support (statistical bootstrap support (BS) values of 100%). Support values were commonly consistent across the three phylogenies, excluding some of the more recent splits (see A–F, Fig. 5), where BS varied across all trees and did not necessarily increase with the addition of further sequence data. BS values for clades A and B decreased considerably between dataset A and the remaining two datasets (B and C), whereas support for clade F increased. The phylogenetic positions of the six outgroup taxa were principally well-resolved, except for clade G (Fig. 5), which is likely due to the large amount of missing data for *H. anophthalmus* (sample 13, Table 2).

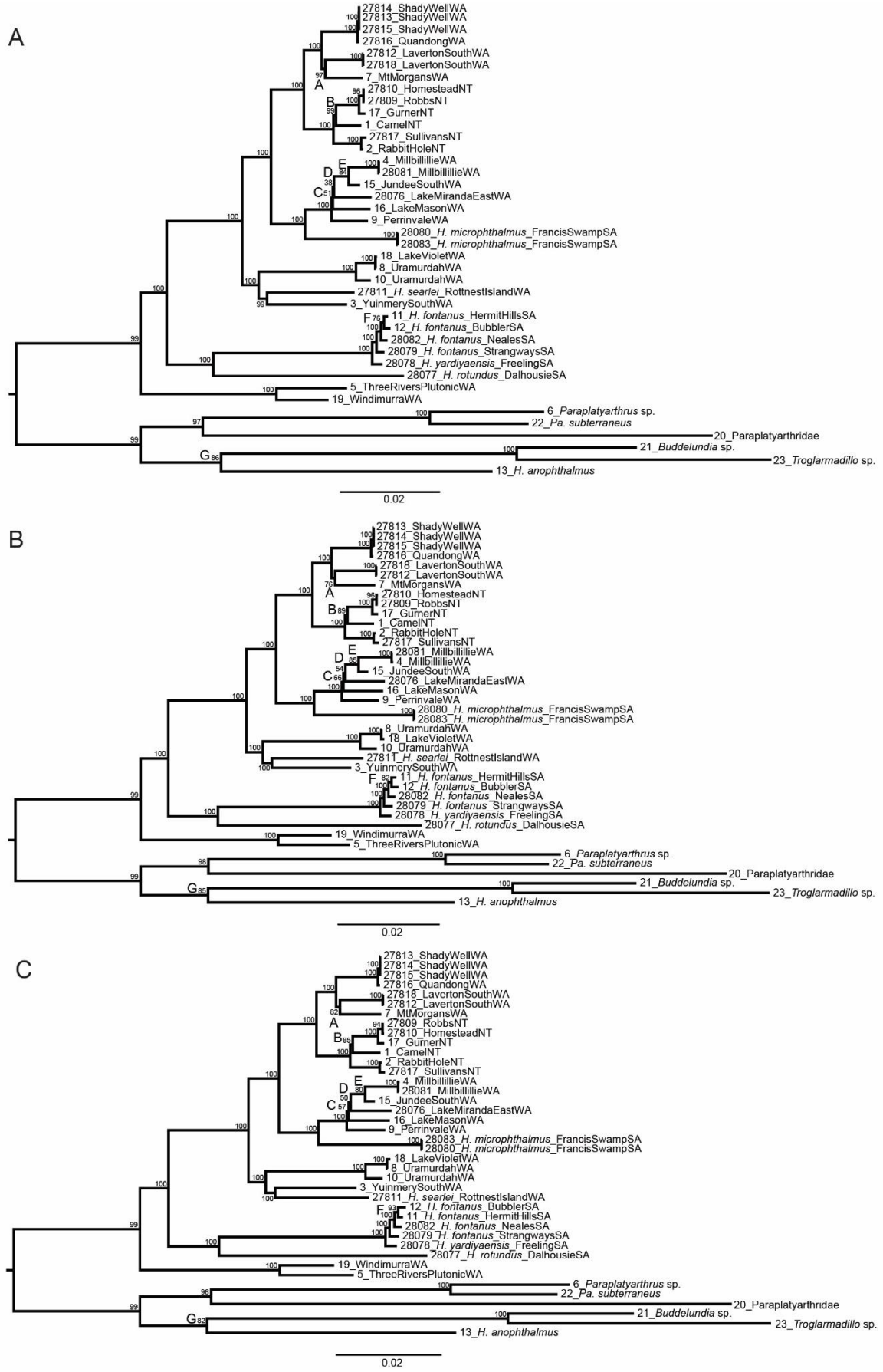


Figure 5: Phylogenies of *Haloniscus* (specified by name and/or collection locality) and outgroup isopods inferred using RAxML maximum likelihood analysis and varying levels of missing data: 25% threshold (dataset A), 50% threshold (dataset B), and 75% threshold (dataset C) of missing sequences. Bootstrap support values are indicated at all nodes.

Discussion

Here we used a custom transcriptome-based exon capture approach without a reference genome to effectively generate a substantial and informative phylogenomic dataset for *Haloniscus*, and distantly related paraplatyarthrid and armadillid isopod, phylogenetic analyses. Transcriptome sequences were used without initially distinguishing intron-exon boundaries prior to bait design, which resulted in the recovery of numerous coding exons of various lengths along with a considerable amount of non-coding (intron) flanking sequence data. The inferred bait set represents a significant step forward from earlier molecular datasets used to explore *Haloniscus* evolution and systematics that integrated either a single mitochondrial gene (*COI*) (Cooper et al., 2008; Murphy et al., 2015) or two genes (mtDNA *COI* and *18S rRNA*) (Guzik et al., 2019), where statistical support was poor for internal *Haloniscus* relationships. The largely well-resolved phylogenies produced here, therefore, indicate that this custom bait set provides enormous potential for application in future *Haloniscus* and other isopod molecular studies, especially in light of the baits' broad taxonomic applicability (Mayer et al., 2016). Nevertheless, we examine the overall effectiveness of our bait design and the performance of this transcriptome-based exon capture approach below.

The transcriptome-based exon capture approach implemented here has several advantages, helping to overcome the lack of an available reference genome to recognise intron-exon boundaries preceding bait design. Tiling across these boundaries has effectively allowed for the recovery of many short (>100 bp) and longer (>300 bp) exons, with high uniformity in sequencing depth (Fig. S9), and a large amount of highly variable non-coding flanking (intron) data. We recovered a substantial quantity of short exons (median length of 153 bp), which would not typically be targeted in exon capture studies that utilise a reference genome and larger exons for bait design (Bragg et al., 2016), but that dominate the majority of arthropod genes. Moreover, calculations for the coefficient of dispersion (CD) have highlighted that the bait design (which spans intron boundaries) promoted consistent sequencing depth values across, especially short, exons (Fig. S9). Portik et al. (2016) reported similar findings, contrasting with studies that employed a reference genome in bait design to define and tile along exons, which have uncovered an 'edge effect', where fewer tiled baits towards the ends of exons leads to a reduction in sequencing depth at contig edges (Bragg et al., 2016; Klopfstein et al., 2019). The methodological pipeline outlined here, nevertheless, differs from that applied by Portik et al. (2016) by implementing BaitFisher, which permits a large number of loci to be targeted for multiple reference species with a smaller number of baits (reducing overall costs) (Mayer et al., 2016), and the Orthograph software, which conveniently and reliably infers orthologues using a best reciprocal hit approach (Petersen et al., 2017). Lastly, the intron sequence data, while not examined in this study, may be highly valuable for population genetics, species delimitation, and phylogenetic analyses.

For our exon capture approach, sequencing depth was consistently high across targeted exons for all samples incorporated within the first (10 samples) and second (8 samples) sequencing runs, but were significantly (11 fold) lower for the third (22 samples) sequencing run (Fig. 2 for example, and Fig. S5). While the general coverage levels for the first two runs considerably outweighed the amount needed per exon for inclusion in the final alignments, the overall amount of missing data (across exons in the dataset B alignment) for the majority of samples included in the third run was considerably larger than for samples within the previous sequencing runs (see Table 2). These differences in sequencing depth and missing data recovered in the third run likely resulted from the higher number of pooled specimens prior to sequencing (Table 2), and, consequently, we recommend either a more conservative pooling selection or an alternative high-throughput platform, such as the Illumina HiSeq or NovaSeq, especially when the amount of intron sequence data is unknown. The third run, however, also consisted of many isopod specimens stored for a long period of time (collected and preserved in ethanol >14 years prior) with likely degraded DNA as well as more divergent outgroup taxa, which may have further influenced the ultimate success of this exon capture run (Abdelkrim et al., 2018; Bragg et al., 2016; Portik et al., 2016).

Therefore, we explored whether differences in sequencing depth and missing data within the distinct runs were correlated with specimen storage time (specimen age), and the ingroup or outgroup status of samples. We included specimens with a range of preservation ages (preserved 3.5–18 years prior) and uncovered no clear relationship between the age of preserved samples and either the level of sequencing depth across exons or missing data (Figs. 4A, S8). Overall, the percentage of missing data varied considerably across samples within the third sequencing run, with some of the more recently collected specimens (especially sequencing ID: 2 (BES18759.3), which contained a corresponding low number of raw paired- end reads, Table 2) comprising few recovered exons (Fig. 4). Nevertheless, the oldest specimens in the first and third runs all consisted of the highest quantities of missing data and, therefore, it is probable that specimen preservation age (potentially combined with storage conditions (Abdelkrim et al., 2018), although this was not investigated here) played a role in the success of these captures. Furthermore, the more divergent outgroup species exhibited an expected higher percentage of missing data (see Bragg et al., 2016; Portik et al., 2016), especially for sample 21 (Armadillidae, Table 2). However, this may be due more to the mapping approach used (mapped to *Haloniscus* orthologues) in data processing than to the efficacy of the baits, since comparable numbers of raw paired-end reads were obtained for these outgroup species and the *Haloniscus* taxa included in the run, suggesting that data assemblies may have been preferential to mapping for these outgroup samples. Nonetheless, the baits and this capture protocol successfully enriched sequence data from older and distantly related isopod specimens, which is equivalent to findings from previous phylogenomics studies (Bailey et al., 2016; Bi et al., 2013; Guschanski et al., 2013; Wood et al., 2018).

We further provide additional empirical data on the question surrounding how many samples may be pooled in a single reaction prior to capture without overly impacting the quality of the sequence data obtained (as in Portik et al., 2016). We examined pools containing 1–4 individuals (12 reactions total), and considered potential effects on sequencing depth, raw paired-end data yield and PCR duplication levels. While low sample sizes precluded rigorous testing, we uncovered no discernible patterns in the exon sequencing depth or raw sequencing yield across samples from different pool sizes (Figs. S6, 3A). However, unlike Portik et al. (2016), we uncovered no trend in duplication levels across pooling sizes (Fig. 3B), but instead duplication was principally lower across samples in the third run (Table 2), which included a lower number of raw paired-end reads for individual samples. Therefore, our results suggest that at least four samples (potentially more) may be pooled together in a single capture reaction, which has important implications for reducing the costs of a study by improving the efficiency of experiments, and increasing the number of samples that can be included in the project (Bi et al., 2012). Nonetheless, the limits of this pooling strategy have not been tested here and should be examined in future studies. Rather than sample pooling selections or the preservation age of specimens, our results have indicated that undetermined characteristics of the isopod specimens included in the capture runs, such as field handling or storage conditions, accounted for the vast majority of variation (after excluding sequencing run) in general sequence data quality.

Overall, the exon capture methods and bioinformatics data processing approach used here have been effective in obtaining a large set of single-copy orthologous groups, successful bait design that enriches target orthologues from diverse *Haloniscus* and other distantly related isopod species, and generating a large and informative phylogenomic dataset. While the final three superalignments used for ML tree inference contained differing levels of missing data, the phylogenies revealed identical topologies, and were largely consistent with previous taxonomic and phylogenetic research (Cooper et al., 2008; Guzik et al., 2019; Stringer et al., 2019). However, most of the phylogenetic relationships inferred here were entirely resolved with maximal statistical support, particularly at internal branches, which provides us with further confidence in this approach. By providing our transcriptome assemblies, filtered OGs, bait design, and custom-made scripts with automated data post-processing steps, we make our approach transparent and, therefore, applicable in future transcriptome-based exon capture studies, especially those focussed on isopods. While detecting and separating out the numerous exons manually is a very time-consuming process, automated scripts have been recently published (Klopfstein et al., 2019) that determine intron-exon boundaries from alignments. Finally, our methodological outline enables these targeted capture techniques to be carried out completely in-house, without the need for outsourcing, where protocols may not be fully disclosed.

Acknowledgements

The authors would like to thank Tessa Bradford and Kathy Saint for invaluable advice with all aspects of the transcriptome and capture sequencing laboratory work; Malte Peterson and Oliver Niehuis for guidance and computational assistance with orthology prediction; Bill Humphreys, Rachael King, and Nick Murphy for specimen collection; Mohammad Javidkar for providing essential outgroup samples; and Barbara Langille, Seraina Klopfstein, and Erinn Fagan-Jeffries for valuable discussions concerning the project. Field work across Australia was completed under: Northern Territory Permit No. 54946, South Australian Permit No: Z25519 and Western Australian Permit No. SF0009792. We would like to thank the traditional owners of the GAB spring country, particularly Reg Dodd and Dean Ah Chee, for providing us with access to sample GAB springs, and Jeff Hulcombe and the Aṉangu Luritjiku Rangers for assistance with specimen collection at Central Mt Wedge. We would also like to thank the staff at Newhaven Sanctuary (NT), and the numerous station managers and mining officers who granted us access to carry out field work. This project was supported with supercomputing resources offered by the Phoenix HPC service at The University of Adelaide and funded by the Australian Research Council (LP0669062 and P140100555) with the following industry partners: Department for Environment and Water (SA), the South Australian Museum, BHP Billiton, Nature Foundation (SA), Biota Environmental Sciences, the Western Australian Museum, Department of Biodiversity, Conservation and Attractions (WA), and Bennelongia. Additional funding was provided by an Australian Biological Resources Study Capacity Building grant to DNS (CT214-11), the Nature Conservancy, kindly supported by The Thomas Foundation (to DNS), and Bioplatforms Australia (Research Contract to ADA and SJBC) for sequencing. DNS acknowledges the support of an Australian Government Research Training Program Scholarship.

References

Abdelkrim, J., Aznar-Cormano, L., Fedosov, A. E., Kantor, Y. I., Lozouet, P., Phuong, M. A., … Puillandre, N. (2018). Exon-capture-based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Molecular Biology and Evolution*, *35*, 2355–2374. https://doi.org/10.1093/molbev/msy144

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bailey, S. E., Mao, X., Struebig, M., Tsagkogeorga, G., Csorba, G., Heaney, L. R., … Rossiter, S.J. (2016). The use of museum samples for large-scale sequence capture: a study of congeneric horseshoe bats (family Rhinolophidae). *Biological Journal of the Linnean Society*, *117*, 58–70. https://doi.org/10.1111/bij.12620

Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R., & Moritz, C. (2013). Unlocking the vault: next-generation museum population genomics. *Molecular Ecology*, *22*, 6018–6032. https://doi.org/10.1111/mec.12516

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., & Good, J. M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, *13*, 403. https://doi.org/10.1186/1471-2164-13-403

Bragg, J. G., Potter, S., Bi, K., & Moritz, C. (2016). Exon capture phylogenomics: efficacy across scales of divergence. *Molecular Ecology Resources*, *16*, 1059–1068. https://doi.org/10.1111/1755-0998.12449

Chilton, C. (1920). On a new isopodan genus (family Oniscidae) from Lake Corangamite, Victoria. *Proceedings of the Linnaean Society of N.S.W*, *44*, 723–734.

Cooper, S. J. B., Saint, K. M., Taiti, S., Austin, A.D., & Humphreys, W. F. (2008). Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic isopods (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics*, *22*, 195–203. https://doi.org/10.1071/IS07039

Dietz, L., Dömel, J. S., Leese, F., Mahon, A. R., & Mayer, C. (2019). Phylogenomics of the longitarsal Colossendeidae: the evolutionary history of an Antarctic sea spider radiation. *Molecular Phylogenetics and Evolution*, *136*, 206–214. https://doi.org/10.1016/j.ympev.2019.04.017

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*, 717–726. https://doi.org/10.1093/sysbio/sys004

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv:* arXiv:1207.3907v2

Glenn, T. C., Nilsen, R., Kieran, T. J., Finger, J. W., Pierson, T. W., Bentley, K. E., … Faircloth, B. C. (2016). Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). *BioRxiv*. https://doi.org/10.1101/049114

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *15*, 644–652. https://doi.org/10.1038/nbt.1883

Guschanski, K., Krause, J., Sawyer, S., Valente, L.M., Bailey, S., Finstermeier, K., … Nagy, Z.T. (2013). Next-generation museomics disentangles one of the largest primate radiations. *Systematic Biology*, *62*, 539–554. https://doi.org/10.1093/sysbio/syt018

Guzik, M. T., Stringer, D. N., Murphy, N. P., Cooper, S. J. B., Taiti, S., King, R. A., Humphreys, W. F., & Austin, A. D. (2019). Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, *33*, 556–574. https://doi.org/10.1071/IS18070

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., … Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*, 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hugall, A. F., O'Hara, T. D., Hunjan, S., Nilsen, R., & Moussalli, A. (2016). An exon-capture system for the entire class Ophiuroidea. *Molecular Biology and Evolution*, *33*, 281–294. https://doi.org/10.1093/molbev/msv216

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Klopfstein, S., Langille, B., Spasojevic, T., Broad, G. R., Cooper, S. J. B., Austin, A. D., & Niehuis, O. (2019). Hybrid capture data unravel a rapid radiation of pimpliform parasitoid wasps (Hymenoptera: Ichneumonidae: Pimpliformes). *Systematic Entomology*, *44*, 361–383. https://doi.org/10.1111/syen.12333

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*(1), 82. https://doi.org/10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*, 772–773. http://doi.org/10.1093/molbev/msw260

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25

Leaché, A. D., & Rannala, B. (2011). The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology*, *60*, 126–137. https://doi.org/10.1093/sysbio/syq073

Lemmon, E. M., & Lemmon, A. R. (2013). High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 99–121.

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenetics. *Systematic* Biology, *61*, 727–744. https://doi.org/10.1093/sysbio/sys049

Li, C., Hofreiter, M., Straube, N., Corrigan, S., & Naylor, G. J. P. (2013). Capturing protein-coding genes across highly divergent species. *BioTechniques*, *54*, 321–326. https://doi.org/10.2144/000114039

Li, H., & Durbin, R. (2008). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

Mayer, C., Sann, S., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., … Niehuis, O. (2016). BaitFisher: a software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution*, *33*, 1875–1886. https://doi.org/10.1093/molbev/msw056

McCartney-Melstad, E., Mount, G. G., & Shaffer, H. B. (2016). Exon capture optimization in amphibians with large genomes. *Molecular Ecology Resources*, *16*, 1084–1094. https://doi.org/10.1111/1755-0998.12538

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfeld, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, *66*, 526–538.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A. … DePristo, M. A., (2010). The Genome Analysis Toolkit: A MapReduce framework for analysing next-generation sequencing data. *Genome Research*, *20*, 1297–1303. . https://doi.org/10.1101/gr.107524.110

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5448-prot5448. http://doi.org/10.1101/pdb.prot5448

Meyer, B., Meusemann, K., & Misof, B. (2011). MARE v0.1.2-rc. Retrieved from http://www.zfmk.de/bioinformatics/MARE_v0.1.2-rc.tar.gz

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, *17*, 240–248. http://doi.org/10.1101/gr.5681207

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA (pp 1–8).

Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., … Zhou, X. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*, 763–767. https://doi.org/10.1126/science.1257570

Misof, B., Meyer, B., von Reumont, B. M., Kück, P., Misof, K., & Meusemann, K. (2013). Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. *BMC Bioinformatics*, *14*, 348. https://doi.org/10.1186/1471-2105-14-348

Murphy, N. P., Guzik, M. T., Cooper, S. J. B., & Austin, A. D. (2015). Desert spring refugia: museums of diversity or evolutionary cradles? *Zoologica Scripta*, *44*, 693–701. https://doi.org/10.1111/zsc.12129

O'Hara, T. D., Hugall, A. F., Cisternas, P. A., Boissin, E., Bribiesca-Contreras, G., Sellanes, J., Paulay, G., & Byrne, M. (2019). Phylogenomics, life history and morphological evolution of ophiocomid brittlestars. *Molecular Phylogenetics and Evolution*, *130*, 67–80. https://doi.org/10.1016/j.ympev.2018.10.003

Peters, R. S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S. Meusemann, K., … Niehuis, O. (2017). Evolutionary history of the Hymenoptera. *Current Biology*, *27*, 1013–1018.

Petersen, M., Meusemann, K., Donath, A., Dowling, D., Liu, S., Peters, R. S., … Niehuis, O. (2017). Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics*, *18*, 1–10. https://doi.org/10.1186/s12859-017-1529-8

Portik, D. M., Smith, L. L., & Bi, K. (2016). An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Molecular Ecology Resources*, *16*, 1069–1083. https://doi.org/10.1111/1755-0998.12541

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org.

Roy, S. W., Fedorov, A., & Gilbert, W. (2003). Large-scale comparison of intron positions in mammalian genes shows intron loss but no gain. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 7158–7162. https://doi.org/10.1073/pnas.1232297100

Salichos, L., & Rokas, A. (2013). Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature*, *497*, 327–331. https://doi.org/10.1038/nature12130

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stringer, D. N., King, R. A., Taiti, S., Guzik, M. T., Cooper, S. J. B., & Austin, A. D., 2019. Systematics of the isopod genus *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae) with description of four new species from Great Artesian Basin springs in South Australia. *Journal of Crustacean Biology*, *39*(5), 651–668. https://doi.org/10.1093/jcbiol/ruz044

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, *34*, W609–W612. https://doi.org/10.1093/nar/gkl315

Teasdale, L. C., Köhler, F., Murray, K. D., O'Hara, T., & Moussalli, A. (2016). Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. *Molecular Ecology Resources*, *16*, 1107–1123. https://doi.org/10.1111/1755-0998.12552

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2012). Integrative Genomics Viewer (IGV): high-performance genomics data visualisation and exploration. *Briefings in Bioinformatics*, *14*, 178–192. https://doi.org/10.1093/bib/bbs017

Waterhouse, R. M., Zdobnov, E. M., Tegenfeldt, F., Li, J., & Kriventseva, E. V. (2011). OrthoDB: the hierarchical catalog of eukaryotic orthologs in 2011. *Nucleic Acids Research*, *39*, D283–D288. https://doi.org/10.1093/nar/gkq930

Wood, H. M., González, V. L., Lloyd, M., Coddington, J., & Scharff, N. (2018). Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). *Molecular Phylogenetics and Evolution*, *127*, 907–918. https://doi.org/10.1016/j.ympev.2018.06.038

Wolfe, J. M., Breinholt, J. W., Crandall, K. A., Lemmon, A. R., Lemmon, E. M., Timm, L. E., … Bracken-Grissom, H. D. (2019). A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans. *Proceedings of the Royal Society B*, *286*. https://doi.org/10.1098/rspb.2019.0079

Wortley, A. H., Rudall, P. J., Harris, D. J., & Scotland, R. W. (2005). How much data are needed to resolve a difficult phylogeny?: case study in Lamiales. *Systematic Biology*, *54*, 697–709. https://doi.org/10.1080/10635150500221028

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics, 30*(5), 614–620. http://doi.org/10.1093/bioinformatics/btt593

CHAPTER 3:

Exon capture-based phylogeny and biogeographic history of *Haloniscus* (Isopoda: Philosciidae) from Australian arid zone groundwater-dependent ecosystems

Statement of Authorship

Title of Paper	Exon capture-based phylogeny and biogeographic history of *Haloniscus* (Isopoda: Philosciidae) from Australian arid zone groundwater-dependent ecosystems
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Intention to submit to *Molecular Phylogenetics and Evolution*

Principal Author

Name of Principal Author (Candidate)	Danielle Stringer
Contribution to the Paper	Manuscript conception, collected specimens, generated data, conducted analyses, interpreted results, wrote manuscript and compiled all figures and tables.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 28/6/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author (Candidate)	Terry Bertozzi
Contribution to the Paper	Guided bioinformatics, wrote some of the crucial scripts for data analysis, and critically reviewed manuscript.
Signature	Date 8.vii.19

Name of Co-Author	Michelle Guzik
Contribution to the Paper	Supervised development of work, collected specimens, provided advice on data analysis and interpretation, and critically reviewed manuscript.
Signature	Date 8/7/19

Name of Co-Author	Steven Cooper
Contribution to the Paper	Supervised development of work, collected specimens, provided advice on data analysis and interpretation, and critically reviewed manuscript.
Signature	Date 28/6/2019

Name of Co-Author	William Humphreys
Contribution to the Paper	Collected specimens, provided knowledge and expertise of subterranean fauna, and critically reviewed manuscript.
Signature	Date 27/6/2019

Name of Co-Author	Andrew Austin
Contribution to the Paper	Supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

Exon capture-based phylogeny and biogeographic history of *Haloniscus* (Isopoda: Philosciidae) from Australian arid zone groundwater-dependent ecosystems

Running title: Biogeographic history of Australian arid zone *Haloniscus*

Danielle N. Stringer[a,*], Terry Bertozzi[b], Michelle T. Guzik[a], Steven J.B. Cooper[a,b], William F. Humphreys[c,d], Andrew D. Austin[a]

[a]Australian Centre for Evolutionary Biology and Biodiversity and Department of Ecology and Evolutionary Biology, School of Biological Sciences, The University of Adelaide, SA 5005, Australia
[b]Evolutionary Biology Unit, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia
[c]Western Australian Museum, Welshpool DC, WA 6986, Australia
[d]School of Animal Biology, The University of Western Australia, Crawley, WA 6009, Australia

*Corresponding author email: danielle.stringer@adelaide.edu.au

Abstract

Groundwater-dependent ecosystems, with their associated relictual fauna, are considered important model systems for testing hypotheses concerning biogeographical processes and the impact of major climatic events on the evolution of species. Here, we examine the biogeographic history of *Haloniscus* isopods from three arid zone groundwater habitats (the Yilgarn calcrete aquifers (Western Australia), Ngalia Basin aquifers (Northern Territory) and Great Artesian Basin springs (South Australia)) utilising targeted exon capture data from 437 genes. Phylogenetic, divergence time dating, and ancestral area reconstruction analyses were used to determine the timing and influence of major aridification periods on the evolution of this fauna. Our analyses revealed historical connections between *Haloniscus* taxa from the three disparate arid zone regions, providing evidence for once broadly distributed ancestral populations, with current fauna representing relict species. Divergence dating further suggested that isolation of these isopod species within the groundwater ecosystems coincided with two aridification phases: one that emerged during the late Miocene and a second that occurred following a temporary return to warmer and wetter conditions during the early Pliocene. These major climatic changes, and the resultant isolation of *Haloniscus* within groundwater-dependent refugia, furthermore, promoted the high levels of species diversification and short-range endemism apparent today.

Keywords: divergence time estimates, evolutionary history, oniscidean isopods, phylogenomics

1. Introduction

Investigating the evolutionary and biogeographic history of relictual taxa, which are characteristically poor dispersers and have survived major climatic or other environmental changes, can offer valuable insights into the past history of continents and, additionally, help to unravel the origins of unique taxa (Vences et al., 2009; Habel et al., 2010; Bauzà-Ribot et al., 2012; Rix et al., 2017; Pepper et al., 2018). Phylogenetic and phylogeographic analyses, in particular, employ molecular sequencing data to assess the spatial distribution of genetic lineages, identify putative species and deduce the influence of critical historical events and processes on the evolution of these relictual taxa (Byrne, 2008; Page et al., 2008; Kleckova et al., 2015; Javidkar et al., 2016). Groundwater-dependent fauna, especially those inhabiting subterranean aquatic systems (stygobionts), provide numerous examples of relictual taxa and, due to their isolation and specialised traits limiting dispersal, are considered valuable models for investigating hypotheses concerning historical biogeographic processes, and modes of evolution (Juan et al., 2010; Juan and Emerson, 2010). In this study, we implement molecular phylogenetic techniques to examine a longstanding hypothesis regarding the biogeographic history of Australia, and the influence of major climatic changes on groundwater-dependent species.

Subterranean and freshwater spring groundwater-dependent ecosystems in the Australian arid zone consist of an extraordinarily diverse, endemic fauna, thought to be relicts from a warmer and wetter time in Australia's history (Humphreys, 2006, 2008; Cooper et al., 2007; Finston et al., 2009; Guzik et al., 2012; Murphy et al., 2009, 2012; Javidkar et al., 2017). The onset of aridification and formation of deserts around the late Miocene, and subsequently during the Pliocene following a proposed return to the warm and wet conditions of the early Miocene (Sniderman et al., 2007, 2016), are believed to have significantly influenced the evolution of this aquatic fauna (Byrne et al., 2008; Cooper et al., 2002; Leys et al., 2003; Murphy et al., 2015). During these phases, the central Australian environment altered from a warm and wet rainforest habitat, with permanent freshwater lakes and rivers, and transformed into the dry, shrubland, and salt lake landscape present today (Alley and Lindsay, 1995; Byrne et al., 2008). With deepening aridity, once widespread aquatic fauna potentially became isolated within fragmented groundwater regions, representing the only permanent freshwater ecosystems in the arid zone (Davis et al., 2013). The magnitude and periodicity of these climate fluctuations, nonetheless, remain poorly defined through the loss of a continuous fossil record (Hill, 1994; Byrne et al., 2008), but developments in molecular sequencing techniques, together with phylogeographic and molecular clock analyses, are providing valuable information regarding these complex historical processes (Beheregaray, 2008).

Molecular-based studies on relictual taxa can also assist with the identification of important refugial habitats, which facilitate the persistence, as well as expansion, of faunal communities through major

environmental changes (Hewitt, 2000; Keppel et al., 2012). Identifying, understanding and managing these environments is considered a critical conservation priority, particularly given predicted climate change (Moritz and Agudo, 2013), as these systems both preserve ancient lineages through unstable conditions, and also support the generation of new diversity (Davis et al., 2013; Murphy et al., 2015). Three currently recognised examples of groundwater refugia within the Australian arid zone, and the focus areas of the current study, are the aquifers (principally calcrete) of the Yilgarn craton in central Western Australia (WA) and Ngalia Basin region in the Northern Territory (NT), and the Lake Eyre and Dalhousie supergroups of springs fed by the Great Artesian Basin (GAB) in South Australia (SA) (Fig. 1 for map of locations). These systems are physically disparate and highly fragmented, with individual calcretes and groups of closely connected springs described as closed island-like entities with unique, highly specialised, and likely relictual faunal assemblages (Cooper et al., 2002; Murphy et al., 2009).

These three groundwater systems consist of similar invertebrate communities, exhibiting adaptations for either a subterranean or epigean way of life, including water beetles (Watts and Humphreys, 2006; Murphy et al., 2015), hydrobiid snails (Ponder et al., 1995, Humphreys, 2008), and crustacean groups such as isopods (Taiti and Humphreys, 2001; Stringer et al., 2019), amphipods (King, 2009; King et al., 2012), and parabathynellids (Cho et al., 2006; Abrams et al., 2013). Both morphological and molecular research has revealed a recurring pattern of numerous phylogenetic lineages for distinct faunal groups, with one or more species consistently restricted to geographically isolated groups of springs or discrete calcrete aquifers (Cooper et al., 2002, 2007, 2008; Guzik et al., 2008, 2012; Murphy et al., 2009, 2012, 2013). This pattern is symptomatic of limited dispersal pathways and a lack of gene flow across discrete similar habitats, leading to long-term isolation and subsequent speciation.

Murphy et al. (2009) uncovered potential evolutionary connections between the endemic chiltoniid amphipods of the SA GAB springs and the Yilgarn calcrete aquifers. The study found that stygobiontic chiltoniids from the Yilgarn (WA) were paraphyletic with respect to epigean chiltoniids from the GAB springs, which suggests a shared evolutionary history between the fauna from these isolated regions of the arid zone. Additionally, approximate dating specified a Miocene origin for this fauna, where it is hypothesised that a shared mesic environment once dominated (Martin, 2006; Byrne et al., 2008). The occurrence of multiple independent chiltoniid lineages and the close connection between fauna from these disparate regions seems to suggest that a rich and widespread amphipod fauna prevailed at this time and, following aridification, likely survived in groundwater refugial habitats (Murphy et al., 2009). Nevertheless, this study (together with the majority of molecular-based studies on the groundwater invertebrates from these arid locations) was hampered by the limited selection of accessible genetic markers and only used one locus, *cytochrome "c" oxidase subunit 1* (*COI*), leading to poorly supported internal branches. The historical connections between chiltoniid taxa and the hypotheses regarding a

historically shared habitat, therefore, remain unresolved and require further exploration with multiple independent genetic markers and expanded sampling of arid zone relictual taxa.

Isopods from the genus *Haloniscus* Chilton, 1920 have been recorded from the three refugial systems described above, and are considered relict species, once possibly widespread across inland Australia. The genus comprises specialised species derived from terrestrial ancestors secondarily adapted to an aquatic existence (Taiti and Xue, 2012). Species from subterranean calcretes are obligate stygobionts that exhibit adaptations, including loss of eyes and pigments, for their underground way of life, while the SA GAB springs *Haloniscus* are semi-terrestrial, inhabiting the moist sandy margins of springs. To date, *Haloniscus* comprises 11 described species, with four from the Lake Eyre and Dalhousie SA GAB springs (Stringer et al., 2019), four from the WA Yilgarn calcrete aquifers (Taiti and Humphreys, 2001), one species associated with lakes in WA (*H. stepheni* Nicholls and Barnes, 1926), the widespread salt lake species *H. searlei* Chilton, 1920 (the type species of the genus), and *H. anophthalmus* Taiti, Ferrara and Iliffe, 1995 found in limestone cave waters in the Isle of Pines, New Caledonia. Nevertheless, both the Yilgarn and Ngalia Basin aquifers are known to harbour a suite of primarily undescribed *Haloniscus* taxa (Cooper et al., 2008; Guzik et al., 2019).

A phylogenetic study by Guzik et al. (2019) used two genetic markers (*COI* and *18S rRNA*) to assess the systematics of *Haloniscus* from the three groundwater locations and identified significant endemicity and 26 new putative species. Moreover, the study uncovered preliminary findings for a similar pattern of a shared evolutionary history amongst *Haloniscus* species from the SA GAB springs and WA Yilgarn calcretes; however, these connections were not examined in detail using a divergence dating analysis and well-resolved phylogeny. Here, we aimed to broaden the scope of this study by investigating the evolution and biogeographic history of *Haloniscus*, employing a considerably more robust phylogeny for taxa from the three groundwater-dependent habitats using a multi-locus exon capture approach with loci generated in Chapter 2. This next-generation sequencing method has recently been used to successfully resolve phylogenetic relationships and understand the biogeography of vertebrates (Blom et al., 2017; Moritz et al., 2018; Reilly et al., 2019) and invertebrates (Hugall et al., 2016; Abdelkrim et al., 2018; Wood et al., 2018; O'Hara et al., 2019). We tested the following hypothesis using molecular phylogenetic, divergence dating and ancestral area reconstruction analyses: *Haloniscus* taxa inhabiting these groundwater-dependent ecosystems represent climate relicts, where evolutionary connections between species from each region date to either the late Miocene aridification or the early Pliocene, which saw a potential temporary return to wetter conditions followed by increased aridity. Our results provide insights into the influence of major climatic changes on the evolution of Australian *Haloniscus* fauna, and the potential origins of groundwater biodiversity in the continent's arid zone.

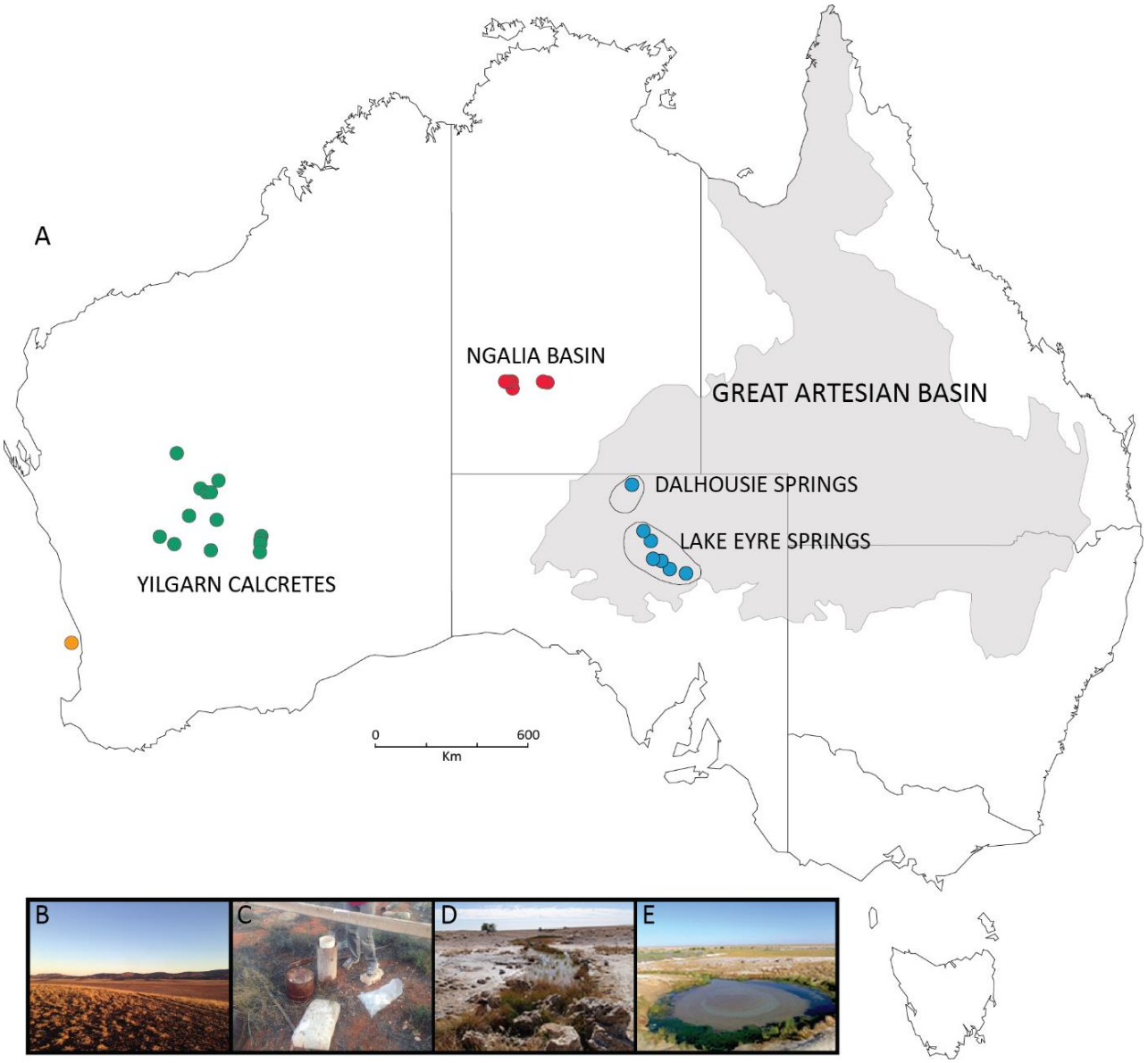


Figure 1: Map of Australia (A) with *Haloniscus* sampling locations from three groundwater habitats: calcrete aquifers in the Yilgarn craton, WA (green), boreholes to aquifers in the Ngalia Basin, NT (red), and the SA Lake Eyre and Dalhousie supergroup springs (blue) fed by the Great Artesian Basin (shaded grey). The location for the salt lake species, *H. searlei*, from Rottnest Island, WA is coloured orange. B–E are exemplar images of: (B) the Australian arid zone, (C) sampling of a borehole, and (D and E) GAB springs, SA.

2. Materials and methods

2.1 Taxon sampling and localities

We selected 31 *Haloniscus* samples, representative of key arid zone lineages informed by Cooper et al. (2008) and Guzik et al. (2019), and four outgroup isopod specimens, which were selected based on a previous study by Javidkar et al. (2016) assessing oniscidean isopod systematics (Table S1). *Haloniscus*

specimens were collected from three major arid zone groundwater-dependent ecosystems, including the SA GAB springs and aquifers (primarily calcrete) in the Yilgarn, WA and Ngalia Basin region, NT, as well as a salt lake on Rottnest Island, WA (see Fig. 1 for map of collection localities). The GAB springs are permanent freshwater systems, where groundwater from the Great Artesian Basin is discharged through outlets around areas of geological weakness (Habermehl, 1982). The springs occur naturally in geographic hierarchical clusters, with directly connected outlets forming 'groups' and, at broader scales, 'complexes', which are hydrochemically and hydrogeologically similar (Habermehl, 1980). The springs have been further grouped into 13 distinct 'supergroups', and, for this study, we focussed on the Lake Eyre and Dalhousie supergroups located in the southern and western areas of the Lake Eyre Basin in the central Australian desert. *Haloniscus* samples were collected opportunistically by hand from the wet margins of these springs and subsequently stored in 100% ethanol at -20°C.

The Yilgarn region comprises more than 200 subterranean calcrete bodies: discrete, shallow and thin (generally around 10 m thick) carbonate formations deposited from groundwater flow along ancient palaeodrainage channels (rivers that stopped flowing in the Palaeocene (Bowler, 1976)) (Humphreys, 2006). This area is comparable to a subterranean archipelago of isolated calcrete water bodies, with some calcretes larger than 100 km^2. The structure of the matrix separating discrete calcretes consists of fine alluvial deposits comprising clay, and is likely a barrier to dispersal (Cooper et al., 2002, 2007). Unlike the Yilgarn calcretes, the Ngalia Basin subterranean aquifers are significantly less well studied, poorly mapped, and the boundaries between particular aquifers are not clearly defined. In our study region, basement highs of impermeable rock allows essentially complete closure of the groundwater system, which exhibits steep salinity gradients that extend to hypersalinity (English et al., 2001). Our ability to sample the aquifers at both locations relied on the availability of pre-existing boreholes and pastoral wells. The sampling regime followed that of Cooper et al. (2008), with *Haloniscus* specimens stored in 100% ethanol at either 4°C or -20°C.

2.2 Exon capture

We sequenced and assembled the transcriptomes of six divergent isopod species (Chapter 2: Table 1 for specimen details) and employed a pre-release version of the software Orthograph (Petersen et al., 2017: beta4.1 available from https://mptrsen.github.io/Orthograph/), together with a published set of 1,478 nuclear genes derived from 12 arthropod genomes (Misof et al., 2014: Table S3), to identify 531 single-copy orthologous genes (or orthologous groups) for use in our exon capture approach. We utilised the software MARE (Misof et al., 2013) v0.1.2-rc (Meyer et al., 2011) to determine the most useful orthologues from our new set of 531 orthologous groups with the highest putative phylogenetic informativeness. Based on these results, we selected 469 orthologous groups, which consisted of an

information content greater than 0.5. Hybridisation baits for these targets were then designed using BaitFisher v1.2.7 (Mayer et al., 2016), and manufactured by MYcroarray (now Arbor Biosciences) for use with a MYbaits (v3) 12 capture reaction kit. Methods for DNA preparation, indexing, preparation of Illumina sequencing libraries, exon capture reactions, and pooling are described in detail in Chapter 2. Pooled samples were sequenced on the Illumina MiSeq platform with 150/300 bp paired-end reads by the Australian Genome Research Facility (AGRF) in Adelaide, SA.

The resulting raw reads were quality controlled, cleaned and merged following the protocol outlined in Chapter 2. Resulting clean reads were then mapped to the *Haloniscus* sp. transcript orthologues for each taxon with BWA v.0.7.15 (Li and Durbin, 2009) and SAMtools v1.3.1 (Li et al., 2009), and output BAM files were evaluated using the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013). Due to the lack of an annotated isopod genome at the time of bait design, the baits were generated directly from the transcriptome orthologue alignments and, consequently, reference targets were split manually (using BAM alignments in IGV) to reflect true intron-exon boundaries. PCR duplicates were removed with Picard tools v2.2.4 (http://broadinstitute.github.io/picard/), and variants were called using FreeBayes v1.0.2 (Garrison and Marth, 2012). Consensus sequences were generated by adding variants to the *Haloniscus* transcripts used during mapping. Bases below 10x coverage were masked using a per base coverage file generated with BEDTools v2.25.0 (Quinlan and Hall, 2010), and IUPAC ambiguity codes were included for heterozygous sites. Sequences for each target were aligned using MUSCLE v3.8.31 (Edgar, 2004) and potential paralogues were masked based on an elevated proportion of heterozygous sites (>3%). For further details on the methods outlined here, including scripts used in the data processing, see Chapter 2.

2.3 Phylogenetic and divergence time analyses

The resulting 786 exon alignments were concatenated with a custom Perl script (available on Bitbucket at: https://bitbucket.org/tbertozzi/scripts/src/master/alignment/), with a user-specified threshold of 50 (only including exon alignments with less than 50% missing data). We utilised PartitionFinder v2.1.1 (Lanfear et al., 2016) with the rcluster algorithm (Lanfear et al., 2014) (parameters: model_selection= aicc, models= all, branchlengths= linked, rcluster-percent= 10.0, rcluster-max= 1000, raxml) to select the most applicable partitioning scheme by exon and substitution models for phylogenetic analysis. A maximum likelihood (ML) phylogeny was estimated for the partitioned dataset using RAxML v8.2.10 (Stamatakis, 2014) employing the GTRGAMMA model of nucleotide substitution, together with 1,000 nonparametric bootstraps (remaining settings as default). The same partitions, as well as the models estimated in PartitionFinder, were also used in an additional ML analysis with IQ-Tree 1.6.10 (Nguyen

et al., 2014) and 1,000 ultra-fast bootstraps (Hoang et al., 2018). The PartitionFinder, RAxML and IQ-Tree analyses were run on the CIPRES Science Gateway v3.3 (Miller et al., 2010).

We inferred a Bayesian multispecies coalescent species tree and estimated species divergence times with StarBEAST2 (Ogilvie et al., 2017) directly from our multiple sequence alignment files. We tested numerous molecular clock combinations, including strict, uncorrelated lognormal, random local, and uncorrelated exponential, and varied the number of exon alignment files included, together with the processes applied for tree priors (see Table S2 for run details). The majority of parameters (including the prior, posterior, tree height and length, mutation rates and molecular clock rates) for the relaxed clock analyses did not reach convergence (ESS >200) after 500 million generations. Here, we present the maximum clade credibility tree based on the strict clock analysis, with all parameters of the model having ESS values >200. This analysis used a reduced dataset of 21 taxa (Table S1 for samples included) and 90 exon alignment files, selected based on the following criteria: ≥18 taxa, sequences 160–933 bp, and information content ≥0.75. The analysis was prepared in BEAUTi v2.5.0 (BEAST2: Bouckaert et al., 2014), run for one billion generations (sampling every 50,000 generations), and was conducted using the HKY substitution model, strict clock and Yule model process.

Haloniscus and other related oniscidean fossil data is distinctly lacking for the time period of interest. Known fossils span time periods outside of the dates predicted for *Haloniscus* divergences (Cooper et al., 2008; Murphy et al., 2015), with the youngest being from the Palaeogene–Eocene, and the oldest from the early Cretaceous (Broly et al., 2013). Thus, a biogeographic node calibration was used to help date the phylogeny. Javidkar et al. (2015) uncovered a strongly-supported sister relationship between South American and Australian Paraplatyarthridae isopods likely resulting from Gondwanan vicariance. Under this vicariance assumption, Javidkar et al. (2017) implemented a fixed minimum age to the root between these paraplatyarthrid taxa at 50 million years ago (Ma). This fixed age is based on the timing of separation between Gondwanan continents, according to the Australia- Antarctica split in the Early Eocene, indicated by the occurrence of a minor wind-driven current in the Australo-Antarctic Gulf (~50 Ma, Lawver and Gahagan, 2003). We used this calibration date of 50 Ma, with a lognormal distribution (mean = 3.91 and standard deviation = 0.02) to date the node between the outgroup paraplatyarthrids (*Paraplatyarthrus subterraneus* Javidkar and King, 2015 (Javidkar et al., 2015) from Laverton, WA, and an undescribed paraplatyarthrid species from Brazil). We, furthermore, added sequences from the *COI* mtDNA gene and applied the second calibration approach from Javidkar et al. (2017), which employed a borrowed rate of evolution for *COI* of 0.0125 substitutions per site per million years for subterranean stenasellid isopods (Ketmaier et al., 2003). However, this dating analysis failed to reach convergence and was not examined further (Table S2). Convergence diagnostics were evaluated with Tracer v1.7.1 (Rambaut et al., 2018) and all trees were visualised using FigTree v1.4.2 (Rambaut, 2014).

2.4Ancestral area reconstructions

Haloniscus ancestral regions were reconstructed with the biogeographic model Statistical Dispersal-Extinction-Cladogensis (S-DEC, Beaulieu et al., 2013) implemented in RASP v4.1 (Yu et al., 2015). This model employed our output StarBEAST2 species trees (1,000 random post burn-in trees and the final condensed tree to reduce optimisation uncertainty) estimated using the strict clock Yule model (SC1, Table S2) analysis, but with outgroups removed. The S-DEC analysis was run with default settings, and maximum areas = 4. The four areas were defined as: (NB) Ngalia Basin aquifers, NT; (Y) Yilgarn craton calcrete aquifers, WA; (GS) GAB springs from the Lake Eyre and Dalhousie supergroups, SA; and (SD) the southern distribution of *H. searlei* (SA, Tasmania (Tas), Victoria (Vic) and WA).

3. Results

3.1 Exon capture data

The final concatenated dataset comprised 786 exon alignments (for 437 of the targeted 469 loci) and 140,690 bp of DNA sequence data for 31 *Haloniscus* and four outgroup isopod samples (see Table S1): alignment available online on Figshare at doi:10.25909/5d3512c0a42be. Targeted exons that were not successfully captured for at least half of the individuals were not included in the final alignment. For the isopods sequenced in this study, an average of 794,220 paired-end reads of sequence data were generated (see Chapter 2 for coverage details). Two of the oldest museum samples, BES6601.2 (Jundee South, WA) and BES6667.2 (Gurner, NT), comprised very few captured exons, with 98 and 97% missing data, respectively. These capture samples were, nevertheless, retained in the below analyses as their phylogenetic position did not alter using three different analysis methods and was, additionally, similar to findings from previous studies on *Haloniscus* systematics (Cooper et al., 2008; Guzik et al., 2019).

3.2 Phylogenetic relationships

The ML (RAxML and IQ-Tree) and Bayesian multispecies coalescent (StarBEAST2) analyses generated almost identical topologies (Figs. 2 and 3), with the majority of splits completely resolved (bootstrap support values (BS) of 100% and posterior probabilities (PP) of 1.00). The Ngalia Basin, NT *Haloniscus* formed a well-supported (BS 100%, PP 1.00) monophyletic clade comprising three distinct groupings: two (Camel and Robbs/Homestead/Gurner: lineages correspond to the names of sampled boreholes) collected from a single, large calcrete aquifer accessed via bores in Newhaven Sanctuary, and a third lineage (termed Sullivan/Rabbit Hole) associated with an alluvial aquifer in the Napperby and Central Mt Wedge regions of the Ngalia Basin (Figs. 2 and 3).

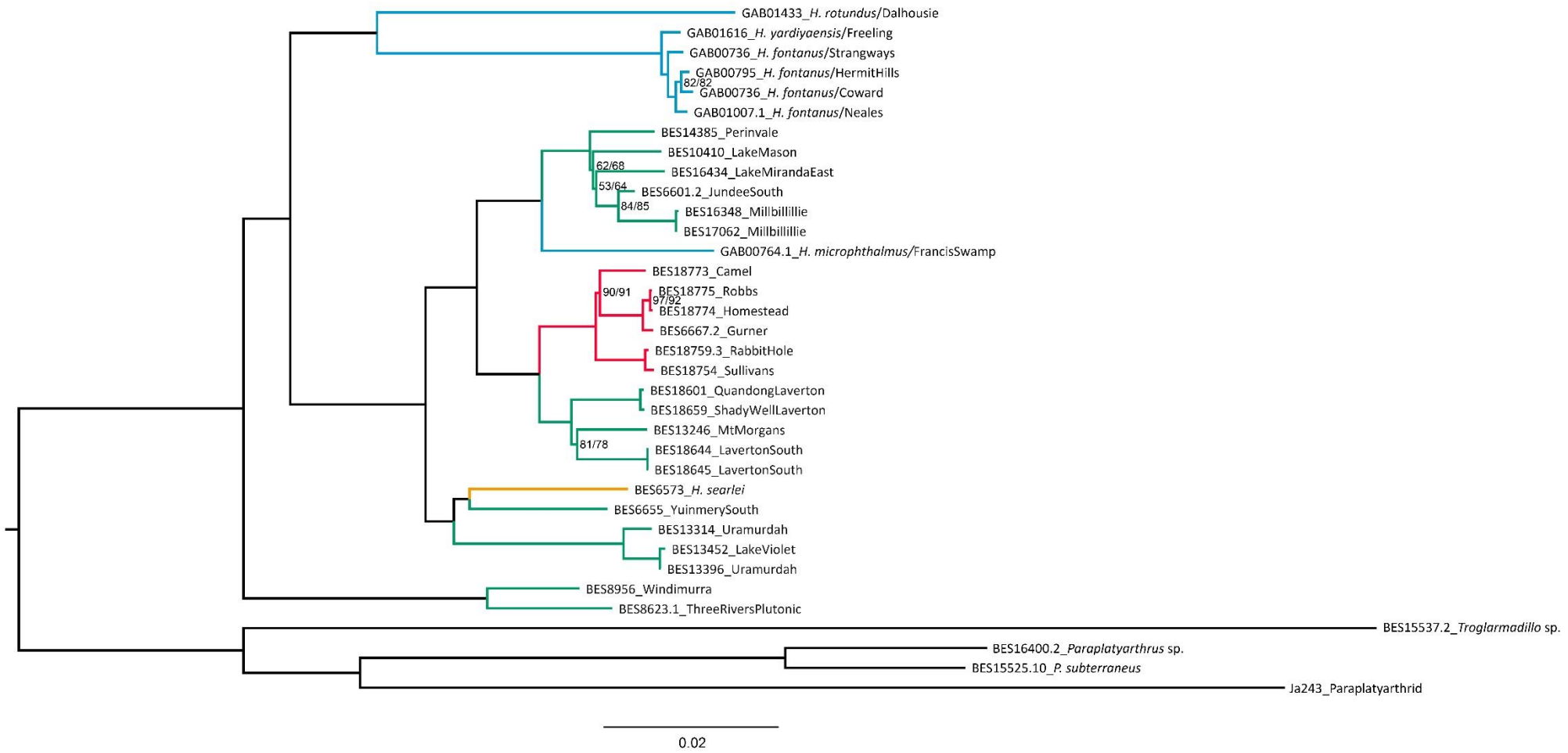


Figure 2: Phylogeny of *Haloniscus* from Australian groundwater-dependent ecosystems based on 786 exons and inferred using RAxML and IQ-Tree maximum likelihood analyses. Support values equal 100 unless otherwise stated and are specified at nodes as ML bootstraps/ultra-fast bootstraps from RAxML and IQ-Tree analyses, respectively. The colour-coded branches represent collection localities as in Fig. 1, with the Yilgarn calcretes (WA) depicted in green, Ngalia Basin aquifers (NT) in red, GAB springs (SA) in blue, Rottnest Island (WA) in orange, and outgroup taxa in black.

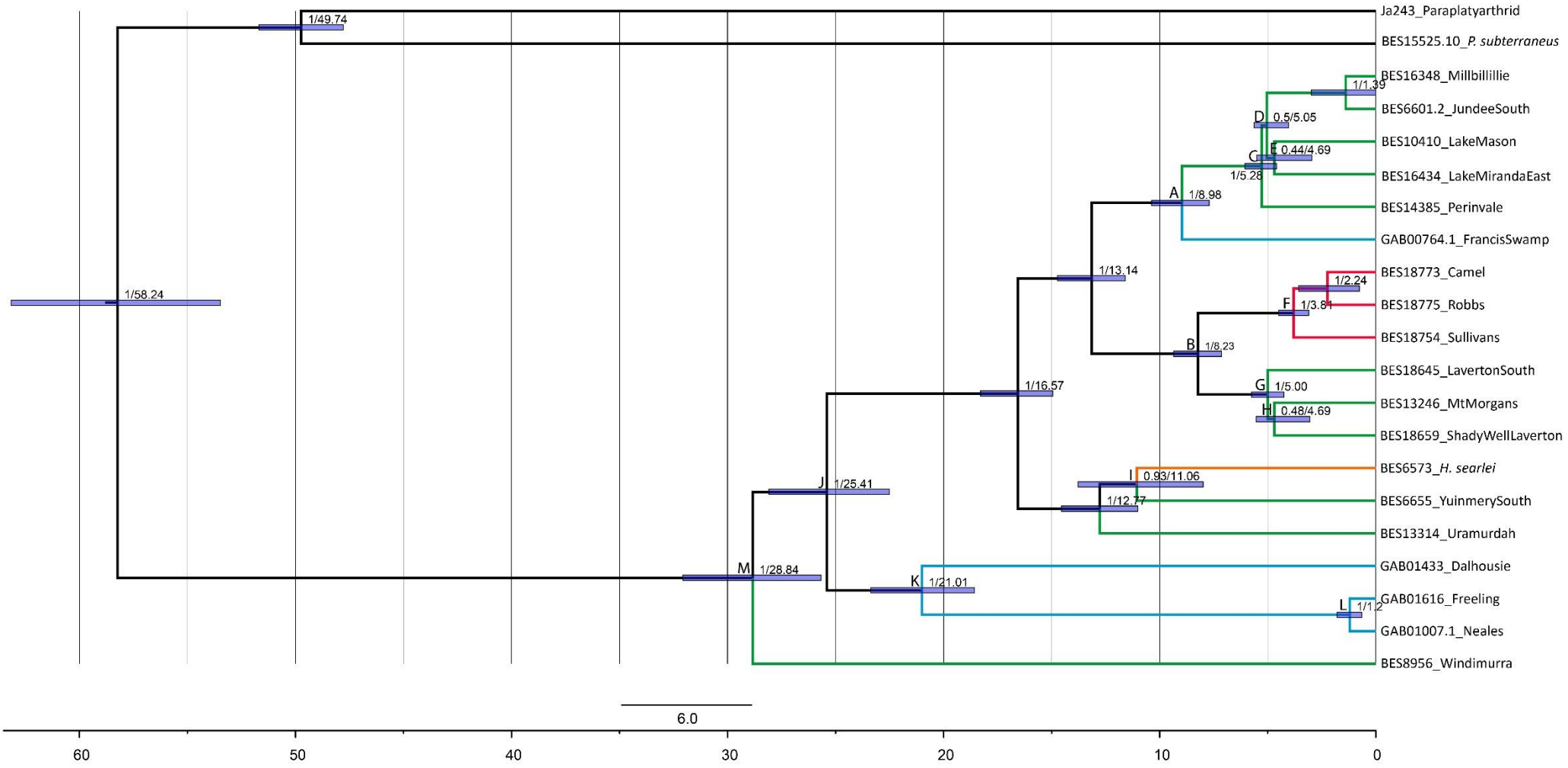


Figure 3: Calibrated evolutionary species tree inferred from StarBEAST2 for *Haloniscus* taxa based on 90 exons. Letters at nodes correspond to node values in Table 2, and colour-coded branches represent collection localities as in Fig. 1, with the Yilgarn calcretes (WA) in green, Ngalia Basin aquifers (NT) in red, GAB springs (SA) in blue, Rottnest Island (WA) in orange, and outgroup taxa in black. Posterior probabilities are stated at each node followed by the divergence time estimates (in millions of years). Blue bars indicate 95% highest probability density intervals.

The Ngalia Basin clade formed a sister relationship with *Haloniscus* from the Laverton Downs and Mt Morgans calcrete aquifers (Carey palaeodrainage valley) within the Yilgarn region (WA) (BS 100%, PP 1.00) and was nested inside a larger *Haloniscus* group (Figs. 2 and 3). The Yilgarn *Haloniscus* were, thus, not monophyletic and, further, consisted of four distinct clades (including the Laverton/Mt Morgans group). The Windimurra (Murchison)/Three Rivers Plutonic (Gascoyne) group consisted of the deepest divergence within *Haloniscus*, with complete support for all analyses, appearing as sister to the other *Haloniscus* taxa (Fig. 2). An additional, primarily Yilgarn, clade contained *Haloniscus* from the Yuinmery South (Raeside), Uramudah (Carey) and nearby Lake Violet (Carey) calcrete aquifers, and the epigean salt lake species *H. searlei* from Rottnest Island, WA (Fig. 2). This relationship again indicates that the Yilgarn *Haloniscus* are not monophyletic, with the Yuinmery specimen more closely related to *H. searlei* than to all other Yilgarn *Haloniscus.*

The final Yilgarn clade was most closely related to the combined Ngalia Basin and Laverton Downs/Mt Morgans group (BS 100%, PP 1.00: Figs. 2 and 3). The position of lineages, including fauna from Jundee South (Carnegie), Millbillillie (Carey), Lake Miranda East (Carey), Perrinvale (Raeside), and Lake Mason (Raeside), was largely unresolved for all trees, excluding the split between a Perrinvale individual and the Yilgarn *Haloniscus* from the aforesaid calcretes (BS 100%, PP 1.00: Figs. 2 and 3). Interestingly, this Yilgarn clade grouped (BS 100%, PP 1.00) with an epigean individual from the Francis Swamp Springs within the Lake Eyre, SA supergroup (*H. microphthalmus* Stringer, King and Taiti, 2019; Stringer et al., 2019) (Figs. 2 and 3). The remaining GAB springs *Haloniscus* formed a well-resolved clade including two distinct groupings: a lineage (*H. rotundus* Stringer, King and Taiti, 2019; Stringer et al., 2019) from the Dalhousie supergroup, and a phylogeographically structured group (based on spring location) from the Lake Eyre springs supergroup, with the most divergent lineage from the northern Lake Eyre, Freeling South Springs (*H. yardiyaensis* Stringer, King and Taiti, 2019; Stringer et al., 2019), and the other related lineages (*H. fontanus* Stringer, King and Taiti, 2019; Stringer et al., 2019) sampled from the Strangways, Hermit Hills, Bubbler (Coward), and Neales Springs (Fig. 2). The SA GAB springs *Haloniscus* were, thus, paraphyletic, where *H. microphthalmus* was more closely related to Yilgarn (WA) calcrete *Haloniscus* than to other known GAB species.

3.3 Estimation of divergence times

The strict clock species phylogeny (Fig. 3) estimated using StarBEAST2 (Ogilvie et al., 2017) recovered high ESS values (>200) for all parameters with the Yule evolution model and 90 exon alignments, and produced a consistent topology to that of the ML phylogenies. The only exceptions were the position of more recently evolved lineages within two Yilgarn, WA clades; Mt Morgans/Shady Well (Laverton), and Lake Mason with Lake Miranda East, which were not well resolved in any phylogeny. To examine

the evolutionary history of Australian arid zone *Haloniscus*, the impact of aridification and the timing of diversification within groundwater habitats, we estimated divergence times with several different models at key nodes (Table 1): a common ancestor link between *Haloniscus* from Francis Swamp SA GAB springs and Yilgarn calcrete aquifers (node A), an evolutionary link between Yilgarn Laverton/Mt Morgans and Ngalia Basin, NT *Haloniscus* (node B), diversification among aquifers within groundwater regions (nodes C, D, E, F, G, H), the relationship between the widespread salt lake inhabiting *H. searlei* and *Haloniscus* from Yuinmery South (Yilgarn, WA) (node I), and the diversification of Dalhousie and Lake Eyre GAB spring *Haloniscus* (nodes J, K, L) (see Table 1 and Fig. 3 for all nodes).

All models produced relatively similar results; nonetheless, the relaxed clock analyses failed to provide adequate ESS values for the majority of parameters, and, consequently, the strict clock analysis (SC1, Table 1) was selected for examination in this study. The distinctive evolutionary connections amongst the Francis Swamp springs (Lake Eyre supergroup, SA) and the Yilgarn *Haloniscus* (node A), as well as between Ngalia Basin and Yilgarn *Haloniscus* individuals (node B), dated to extremely similar times in the late Miocene, 8.23 (95% HPD 7.14–9.36) and 8.98 (95% HPD 7.71–10.38) Ma (Fig. 3). In addition, diversification among taxa within the Ngalia Basin and Yilgarn aquifers was estimated to occur during the early Pliocene around 3–6 Ma (nodes C, D, E, F, G, H: Fig. 3). The split between salt lake *H. searlei* and the Yuinmery South *Haloniscus* taxon (node I, Fig. 3) was recovered as 11.06 Ma (95% HPD 7.99–13.79). The split between Windimurra and all remaining *Haloniscus* taxa (Node M: Fig. 3), which may represent the age of the radiation, was recovered as 28.8 Ma (95% HPD 95% 25.7–32.08). Lastly, the divergence times for *Haloniscus* taxa from the SA GAB supergroups were 25.41 Ma (95% HPD 22.51–28.09, node J) for the split between Francis Swamp and all other SA GAB *Haloniscus*, 21.01 (95% HPD 18.57–23.28, node K) for the node separating Dalhousie and Lake Eyre (not including Francis Swamp) supergroup taxa, and 1.2 Ma (95% HPD 0.64–1.80, node L) between the Lake Eyre (Neales and Freeling South complexes) *Haloniscus* (Fig. 3).

Table 1: Age (millions of years) associated with key nodes connecting distinct *Haloniscus* lineages and the 95% highest probability densities for different models used in StarBEAST2 molecular clock analyses with the exon capture data. Table also includes a comparison of GAB spring *Haloniscus* node ages from Murphy et al. (2015).

Run	Loci	Clock	Process	ESS	Node A	Node B	Node C	Node D	Node E	Node F	Node G	Node H	Node I	Node J	Node K	Node L
SC1	90 exons	Strict	Yule	>200	8.98 (7.71–10.38)	8.23 (7.14–9.36)	5.05 (4.60–6.06)	5.05 (4.04–5.64)	4.69 (2.96–5.51)	3.81 (3.10–4.50)	5.00 (4.26–5.74)	4.69 (3.05–5.54)	11.06 (7.99–13.79)	25.41 (22.51–28.09)	21.01 (18.57–23.38)	1.20 (0.64–1.80)
SC2	90 exons	Strict	Birth-death	>200	8.93 (7.68–10.33)	8.19 (7.05–9.25)	5.29 (4.52–6.05)	4.95 (3.80–5.80)	4.61 (2.70–5.50)	3.83 (3.11–4.53)	4.99 (4.25–5.71)	4.71 (3.10–5.46)	10.96 (8.13–13.50)	25.37 922.76–28.09)	20.94 (18.65–23.45)	1.22 (0.59–1.82)
SC3	90 exons	Strict	Calibrated Yule	>200	8.80 (7.56–10.07)	8.06 (7.01–9.11)	5.21 (4.48–5.93)	4.90 (3.63–5.54)	4.59 (2.66–5.33)	3.73 (3.02–4.39)	4.90 (4.18–5.63)	4.65 (3.24–5.48)	10.90 (7.82–13.44)	24.88 (22.26–17.51)	20.57 (18.16–22.89)	1.20 (0.64–1.79)
RLC1	10 exons	Random Local Clock	Yule	<100	7.93 (6.03–10.07)	6.46 (4.92–8.04)	5.95 (4.42–7.44)	5.48 (3.81–6.85)	5.01 (2.86–6.59)	3.22 (2.08–4.50)	4.43 (3.14–5.70)	3.94 (1.75–5.21)	8.79 (3.19–11.86)	22.42 (18.35–26.42)	18.51 (14.63–22.37)	0.88 (0.00–1.67)
UCLN2	5 exons	Uncorrelated Lognormal	Birth-death	<100	7.63 (4.70–10.73)	6.48 (4.05–8.89)	4.52 (2.04–6.73)	4.26 (1.55–5.07)	3.91 (0.76–4.13)	3.98 (1.77–6.13)	3.83 (2.09–5.89)	2.95 (0.24–4.46)	6.70 (1.86–10.25)	22.66 (16.64–28.74)	17.69 (10.87–25.23)	1.11 (0.00–2.32)
UCED2	5 exons	Uncorrelated Exponential	Birth-death	<100	7.58 (3.55–12.54)	8.46 (4.00–13.52)	4.57 (1.86–7.65)	4.05 (1.24–6.37)	3.05 (0.03–4.72)	3.45 (1.20–6.19)	4.58 (1.77–7.76)	3.95 (0.81–6.49)	11.19 (0.35–17.20)	26.12 (14.04–40.83)	19.59 (8.45–33.71)	1.25 (0.00–3.01)
Murphy et al. (2015)	*COI*	Uncorrelated Lognormal	Yule	>200										22.80 (14.01–32.09)		7.00 (4.80–9.70)

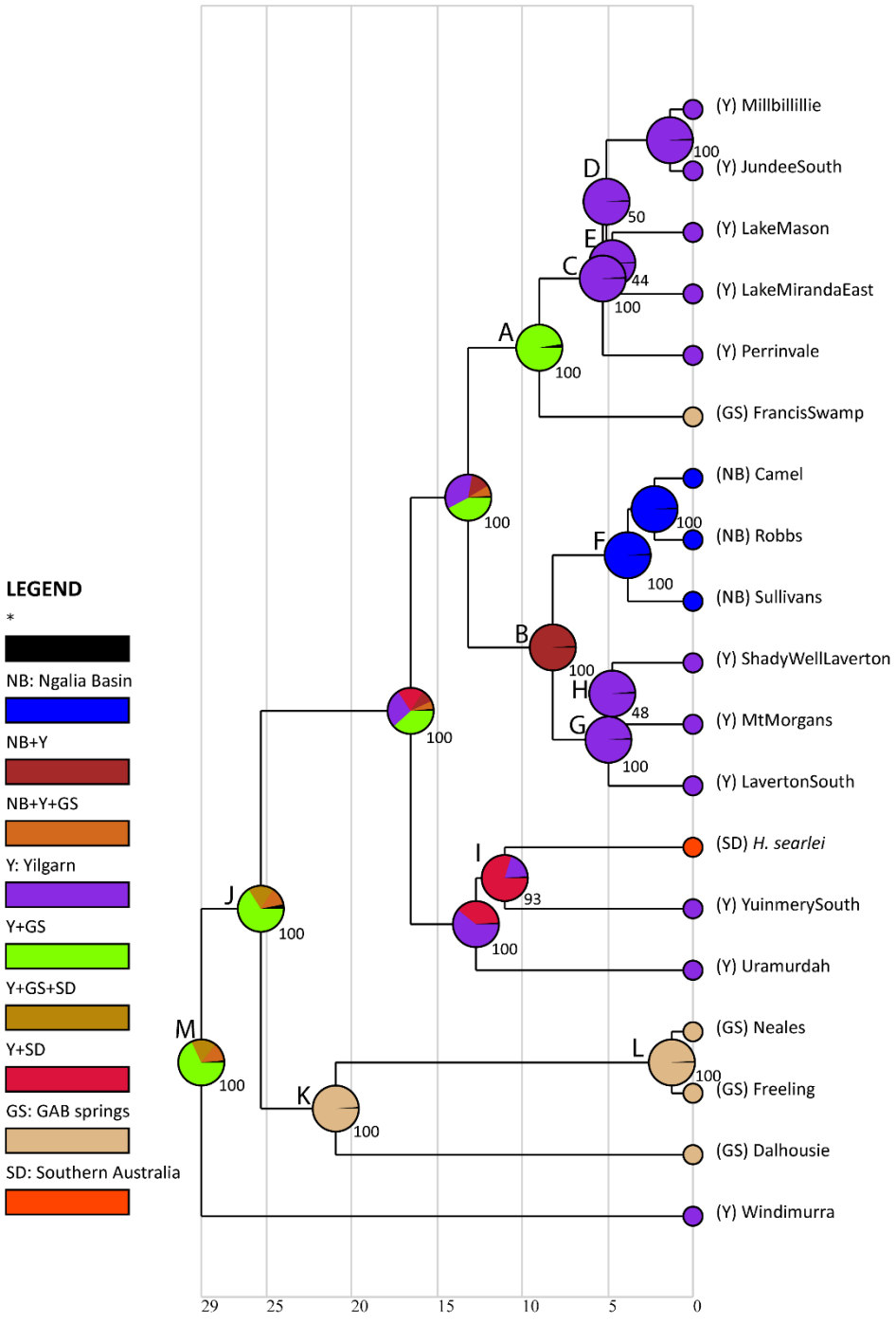


Figure 4: Biogeographic inference recovered with Statistical Dispersal-Extinction-Cladogensis (S-DEC) in RASP. Pie charts represent the marginal probabilities for each alternative ancestral area. The colour-coded legend displays possible ancestral ranges at different nodes, where black represents all other possible ancestral areas. Legend codes are as follows: NB: Ngalia Basin (NT), Y: Yilgarn (WA), GS: GAB springs (SA), SD: southern Australian distribution (SA, WA, Tas and Vic).

3.3 Ancestral areas

The results of the ancestral state analysis using S-DEC in RASP are summarised in Fig. 4, and revealed a broad ancestral range for the Australian *Haloniscus* fauna. S-DEC reconstructed the Yilgarn and the SA GAB springs (Y+GS) as the most likely ancestral area for the Australian *Haloniscus*, with the Ngalia

Basin, Yilgarn and GAB springs (NB+Y+GS) or possibly the Yilgarn, GAB springs and southern Australia (Y+GS+SD) as potential ancestral areas (Fig. 4). The analysis further reconstructed a shared GAB springs and Yilgarn (Y+GS) ancestral area at node A and a comparably shared Ngalia Basin and Yilgarn (NB+Y) region at node B (Fig. 4). Lastly, the ancestral area for the widespread salt lake species *H. searlei* and the Yuinmery South *Haloniscus* lineage (Yilgarn) (node I) was reconstructed as the Yilgarn region and the southern portion of Australia (the distribution of *H. searlei,* which includes WA) (Y+SD, Fig. 4).

4. Discussion

The aridification of Australia and the resultant contraction and fragmentation of the continent's once widespread mesic landscape was an immensely transformative event in the biogeographic history of Australia and considerably impacted the evolution of faunal assemblages (Byrne et al., 2008; Murphy et al., 2015; Rix et al., 2017). In this study, we present the most comprehensive examination into the evolution and biogeographic history of *Haloniscus* isopods from three arid zone groundwater refugia. Our phylogenetic analyses uncovered fully-resolved historical connections between *Haloniscus* fauna from the three disparate regions (Yilgarn calcretes, Ngalia Basin aquifers and GAB springs), indicating a shared evolutionary history and offering evidence for ancestral populations once occupying a much broader distribution, with current taxa representing relict species. Divergence dating analysis further suggested that the isolation of these species within the major groundwater regions coincided with late Miocene aridification, with additional diversification occurring in the Pliocene following the temporary return to warmer and wetter conditions (Byrne et al., 2008; Sniderman et al., 2016). These aridification events and the resultant isolation of *Haloniscus* within groundwater-dependent refugia, furthermore, promoted the high levels of species diversification and short-range endemism apparent today.

4.1 Exon capture and diversity

The exon capture bait set developed here for *Haloniscus* isopod phylogenetics was highly effective in enriching 786 exons from 437 loci for the ingroup and more divergent outgroup paraplatyarthrid and armadillid isopod taxa. This bait set marks a significant progression from previous molecular datasets used to explore *Haloniscus* evolution and systematics, which integrated either a single mitochondrial gene (*COI*) (Cooper et al., 2008; Murphy et al., 2015) or two genes (mtDNA *COI* and *18S rRNA*) (Guzik et al., 2019), with the end result being poor topological resolution for internal branches. Our analyses confirmed significant phylogeographic structuring, with one or more *Haloniscus* lineages confined to distinct aquifers or groups of geographically proximate GAB springs, indicative of long-term isolation and limited dispersal. These lineages are restricted to exceedingly narrow geographic ranges and are, accordingly, short-range (Harvey et al., 2002) or even ultra-short-range (Guzik et al., 2019) endemics.

These results were consistent across our analyses (showing high support) and with earlier *Haloniscus* taxonomic and phylogenetics research (Cooper et al., 2008; Guzik et al., 2019; Stringer et al., 2019), which emphasises the utility and effectiveness of our bait design for phylogenetic inference. This bait set, therefore, offers significant potential for application to future isopod phylogenetic studies.

4.2 Historical connections and the onset of aridity

Our phylogenetic and molecular dating analyses further provide evidence for a shared evolutionary history amongst *Haloniscus* fauna from the three distinct groundwater ecosystems, and highlight the influence of two major aridification events on the isolation and resultant species-level diversification of the genus within these refugial regions. The onset of aridity in the late Miocene (10–6 Ma) greatly transformed the mesic Australian environment, with the contraction of extensive rainforest habitats, large inland lakes and river systems, and the consequent expansion of sclerophyllous vegetation, salt lakes and dry, open shrublands (Alley and Lindsay, 1995; Martin, 2006), principally resulting from the northward movement of the Australian continent (Bowler, 1976). The early Pliocene phase (5–3 Ma), however, is believed to have comprised a temporary return to the warm and wet conditions, although never recovering the substantial lakes of the mid-Miocene, prompted by a significant rise in sea level (Sniderman et al., 2007, 2016). This fluctuating interval was, nevertheless, likely followed by a period of intensive aridity, with the formation of sandy and stony inland deserts (Byrne et al., 2008). The early Pliocene return-to-wet hypothesis is currently based on results from a limited number of geographic locations and not often considered in studies concerning the climate-induced evolution of Australian fauna (Leys *et al*., 2003; Cooper *et al*., 2007, 2008), but overall our findings support this idea, revealing that *Haloniscus* taxa were present on the surface and able to colonise aquifers during this time period. Nonetheless, with the onset of aridity following the early Pliocene, taxa likely either became extinct or adapted to survive within fragmented groundwater refugia.

Our phylogenetic analyses uncovered a distinct connection between *Haloniscus* from Francis Swamp springs in the GAB wetland and the Yilgarn calcrete aquifers. Three SA GAB spring *Haloniscus* groups, explicitly from the Dalhousie springs supergroup, Lake Eyre supergroup, and Francis Swamp complex (Lake Eyre supergroup), were inferred from all phylogenetic analyses (Figs. 2 and 3), with the Francis Swamp species, *H. microphthalmus*, highly divergent from all other currently known GAB *Haloniscus*. Surprisingly, the next closest relatives to *H. microphthalmus* were sampled from calcretes (Perrinvale, Lake Miranda East, Lake Mason, Jundee South, and Millbillillie) in the Yilgarn (WA), a finding similarly highlighted by Guzik et al. (2019), but with markedly lower support for internal branches (Figs. 2 and 3). Guzik et al. (2019) explained this observation of GAB spring *Haloniscus* paraphyly using a scenario of independent colonisation events by divergent ancestors in discrete, unconnected regions, and this

interpretation is supported here by a considerably more robust phylogeny. The presence of multiple unrelated lineages, as well as the common ancestry of *Haloniscus* from the GAB springs and Yilgarn calcrete aquifers, indicates that a widespread *Haloniscus* fauna was once prevalent across the central Australian landscape (Davis et al., 2013). This result of a widely distributed *Haloniscus* ancestor was reinforced using ancestral state reconstruction (RASP), which proposed a combined SA GAB and WA Yilgarn ancestral area for the clade (node A, Fig. 4) containing *H. microphthalmus* and the *Haloniscus* taxa from the above-listed calcretes.

Further examination of phylogeographical patterns amongst *Haloniscus* from the three groundwater regions revealed additional historical connections between the Ngalia Basin, NT and Yilgarn calcrete, WA *Haloniscus*. The Mt Morgans and Laverton Downs (collected from the Shady Well, Quandong and Laverton South bore holes) Yilgarn clade was recovered as sister to a monophyletic group of *Haloniscus* taxa from the Ngalia Basin region, with this relationship gaining maximum support (Figs. 2 and 3). This connection has been previously suggested by Guzik et al. (2019), based on the analyses of a single *18S rRNA* gene tree, but this relationship was not apparent in either a *COI* only or combined *COI* and *18S* phylogeny. In addition, a lineage from the Yuinmery South calcrete in the Yilgarn region was recovered as sister to the salt lake-associated species *H. searlei*, (broadly distributed across southern Australia – WA, SA, Tas and Vic) (BS 100%, PP 1.00: Figs. 2 and 3). These phylogeographic patterns, together with results from ancestral state reconstructions (nodes B and I, Fig. 4), further highlight a broad ancestral distribution for the Australian *Haloniscus* fauna, with subsequent isolation of these ancestral species principally within groundwater-dependent ecosystems.

Estimated divergence times inferred from our analyses provide an approximate timeframe for the evolution of groundwater-dependent *Haloniscus* lineages, their isolation inside refugial groundwater habitats and the close evolutionary connections between the disparate regions. Divergence dating of the major nodes linking the Francis Swamp GAB springs *Haloniscus* with the Yilgarn taxa (node A, Fig. 3) and the Yilgarn region *Haloniscus* with Ngalia Basin (node B, Fig. 3) specified a late Miocene origin for the diversification, with highly comparable estimates of 8.98 Ma (7.71–10.38) and 8.23 Ma (7.14–9.36), respectively. These estimates correspond with the first critical period of aridification across the Australian continent, where there was a major cessation of the warm and wet conditions of the early Miocene phase (Beard, 1977; Bowler, 1976; Byrne et al., 2008). These results, therefore, represent a repeated evolutionary and biogeographic pattern since Murphy et al. (2009) similarly suggested that divergences between distinct clades of SA GAB spring and WA Yilgarn chiltoniid amphipods occurred during the late Miocene. Here, we propose that once ubiquitous ancestors to present day *Haloniscus* (and other groundwater-associated taxa) became isolated within disparate refugial habitats, namely

the SA GAB springs and subterranean aquifers (NT and WA), as a result of late Miocene aridification, and that the current taxa represent climate relicts (Leys et al., 2003; Murphy et al., 2015).

4.3 Relictualisation: diversification within groundwater refugia

Further diversification within the distinct groundwater regions appears to have transpired during the early Pliocene. Divergence times inferred at nodes C, D, E, F, G and H (Fig. 3, Table 1) concur with the isolation and subsequent diversification of *Haloniscus* lineages within discrete calcretes in the Yilgarn and Ngalia Basin and these times (5.28, 5.05, 4.69, 3.81, 5.00 and 4.69 Ma, respectively) overlap with the second intensive period of aridification, which followed a proposed return to wetter conditions during the Pliocene (Sniderman et al., 2016). These time estimates suggest that, following the Miocene aridification, *Haloniscus* were able to move around the landscape, with the early Pliocene wet period likely facilitating dispersal and eventual colonisation of nearby aquifers. These ancestral species may have been wholly epigean (surviving through Miocene aridity) or potentially partial-eyed troglophiles that colonised and evolved within the subterranean calcretes during initial Miocene aridification, but were able to return to the surface during the wet phase (as for paraplatyarthrid isopods, see Javidkar et al. (2017)). Nonetheless, with aridity in the Pliocene, species likely became trapped within aquifers, with limited dispersal both between aquifers and within palaeodrainage valleys, which promoted the high levels of diversification and short-range endemism now apparent (Fig. 2). In addition, the lower support values for some of the more recent nodes (e.g. D, E, H: Fig. 3) likely suggest a hard polytomy, with rapid diversification in subterranean refugia following Pliocene aridification, which isolated the aquifer populations.

The time-point for the node (I, Fig. 3) connecting *Haloniscus* from the Yuinmery South aquifer with *H. searlei* of 11.06 Ma (7.99–13.79) may, nonetheless, suggest an early colonisation into a subterranean calcrete during the first phase of aridification in the Miocene. Previous divergence time estimates for Yilgarn (calcrete-inhabiting) dytiscid diving beetles and chiltoniid amphipods of 9–4 Ma and 15–4 Ma, respectively (Leys et al., 2003; Cooper et al., 2007) are consistent with the results here, which affords some confidence in our estimates. The significantly later date for the diversification between Jundee South and Millbillillie of 1.39 Ma (0.00–2.98, Fig. 3), however, may suggest that connections between nearby aquifers were possible through major flooding events, resulting from the wetter conditions of interglacials during Pleistocene climate oscillations.

Diversification of the Windimurra Yilgarn calcrete taxon (node M, Fig. 3), nonetheless, does not appear to be associated with either late Miocene or early Pliocene aridification. Divergence dating suggested that the *Haloniscus* lineage from Windimurra diversified from all other included *Haloniscus* 28.8 Ma

(25.67–32.08, Fig 3.), which corresponds to the Oligocene. The phylogeny in Fig. 2 further highlighted an entirely supported (BS 100%) sister relationship between lineages from the Windimurra and Three Rivers Plutonic WA calcretes, which has not been uncovered in any previous *Haloniscus* phylogeny as only partial *COI* sequences of the Three Rivers taxa were generated (Cooper et al., 2008). Differences in the evolutionary history of the Windimurra and Three Rivers *Haloniscus* taxa from those discussed above may be related to their distribution within the Yilgarn region. The Windimurra calcrete is sited within the Murchison palaeodrainage and Three Rivers in the Gascoyne, and these distinct drainages are separated from all other Yilgarn palaeodrainage valleys (in which all additional Yilgarn *Haloniscus* taxa inhabit) by a drainage divide, with the Murchison and Gascoyne river valleys draining to the Indian Ocean and the remaining palaeovalleys draining to the Southern Ocean (what was previously the Eucla Basin) (Beard, 1998).

Divergence time estimates reiterate the dissimilar origins in evolutionary history between *Haloniscus* lineages in the Dalhousie supergroup, Lake Eyre supergroup and Francis Swamp complex. Our dating analysis indicates that *Haloniscus* taxa inhabiting Dalhousie and Lake Eyre springs have been isolated for roughly 21 Ma (18.57–23.38: node K, Fig. 3), which dates to the early Miocene period and, hence, suggests that their initial diversification was not linked to aridity. Dalhousie springs are, nevertheless, known to be geographically isolated from other GAB spring outlets and to hold evolutionarily distinct endemic fauna (Ponder et al., 1996; Murphy et al., 2009). The basis for this isolation is unknown, but Krieg (1989) suggests that the Dalhousie spring supergroup has been isolated from Lake Eyre since at least the late Pleistocene, and that the regions comprise different histories. Our time estimates here, however, propose that the Dalhousie Basin has been isolated for significantly longer, with *Haloniscus* dispersal pathways limited since the early Miocene phase. This restricted distribution may have been influenced by an inland sea or through the presence of multiple ancestral species with one, by chance, surviving in the south and the other in the Dalhousie Basin. The Freeling South and Neales complex springs (node L, Fig. 3) are both situated towards the northern end of the Lake Eyre supergroup, and were more recently connected (1.2 Ma: node L, Fig. 3), but taxa likely diversified within these springs following the formation of deserts and spring isolation between 1–4 Ma, resulting from aridification (Byrne et al., 2008; Murphy et al., 2012, 2015).

Our divergence dating estimates are consistent with previous studies and hypotheses concerning the timing of aridification, but they should, nevertheless, be regarded with some caution since they were based on a strict clock rather than relaxed molecular clock analyses. A strict clock assumes that every branch in a phylogenetic tree evolves at the same evolutionary rate, which may not be accurate, particularly for fauna within markedly different habitats, such as the Yilgarn and Ngalia aquifers with stygobionts, and the GAB springs with epigean taxa. However, the results we obtained for significant

nodes using different models, including those implementing a relaxed molecular clock (although these analyses did not reach convergence), recovered similar divergence time estimates (see Table 1). The divergence estimate at node J (Fig. 3) for the connection between the major GAB *Haloniscus* lineages from Dalhousie, Lake Eyre and Francis Swamp was recovered as approximately 25 Ma (22.51–28.09), which is analogous to the estimate of Murphy et al. (2015) of 22 Ma (14.01–32.09, Table 1), but with a narrower 95% HPD interval. The timing of isolation between the Lake Eyre springs (node L, Table 1) revealed here (1.2 Ma) was, however, significantly lower than the estimate of 7 Ma (4.80–9.70) from Murphy et al. (2015). Overall, the approach used by Murphy et al. (2015) (*COI* dataset and calibrated using a standard arthropod mtDNA molecular clock) differed from the approach used here, but both studies similarly highlight the impact of aridity on *Haloniscus* species-level diversification.

5. Conclusions

Late Miocene and early Pliocene climatic changes and the subsequent development of the Australian arid zone prompted substantial contraction and isolation of *Haloniscus* isopod populations in refugial groundwater-dependent ecosystems, and further promoted the extreme levels of diversification and short-range endemism observed today. This study signifies the most comprehensive exploration into longstanding questions of historical connectivity across fragmented Australian arid zone habitats and climate-induced faunal diversification (Byrne et al., 2008; Davis et al., 2013). Through examination of the evolution and biogeographic history of *Haloniscus* using exon capture techniques, we highlighted significant evolutionary links between fauna from the SA GAB springs and aquifers in the Yilgarn, WA and Ngalia Basin, NT regions, with current taxa representing relict species. These results offer crucial insights into the evolution of aquatic arid zone fauna through fluctuating and generally unfavourable conditions, and emphasise the significance of refugial environments in facilitating the persistence, as well as diversification, of once broadly distributed *Haloniscus* taxa and for testing theories regarding the biogeographic history of the Australian continent. The methods used here, together with the bait set comprising 469 loci, offer enormous potential for application to future phylogenetic studies, and contribute a phylogenetic framework for continued and essential taxonomic research of this relictual group of isopods.

Acknowledgements

The authors would like to thank Tessa Bradford and Kathy Saint for invaluable advice with all aspects of the sequencing laboratory work; Simon Tierney for computational support and assistance with the RNA-seq filtering and assembly; Karen Meusemann, Christoph Mayer, Malte Peterson, Oliver Niehuis for guidance with orthology prediction and bait design; Rachael King, Nick Murphy, and Steve Delean

for specimen collection; and Mohammad Javidkar for providing outgroup samples. Field work across Australia was completed under: Northern Territory Permit No. 54946, Western Australian Permit No. SF0009792, and South Australian Permit No: Z25519. We appreciate the access provided to us by the traditional owners of the GAB spring country, particularly Reg Dodd and Dean Ah Chee, as well as Jeff Hulcombe (and a group of rangers) for specimen collection at Central Mt Wedge (NT). We would like to thank the station managers, staff at Newhaven Sanctuary (NT) and mining officers who granted us access to carry out field work. This project was supported with supercomputing resources offered by the Phoenix HPC service at The University of Adelaide and funded by the Australian Research Council (LP0669062 and P140100555) with the following industry partners: Department for Environment and Water (SA), the South Australian Museum, BHP Billiton, Nature Foundation (SA), Biota Environmental Sciences, the Western Australian Museum, Department of Biodiversity, Conservation and Attractions (WA), and Bennelongia. Additional funding was provided by an Australian Biological Resources Study Capacity Building grant to DNS (CT214-11), the Nature Conservancy, kindly supported by The Thomas Foundation (to DNS), and Bioplatforms Australia (Research Contract to ADA and SJBC) for sequencing. DNS acknowledges the support of an Australian Government Research Training Program Scholarship.

References

Abdelkrim, J., Aznar-Cormano, L., Fedosov, A.E., Kantor, Y.I., Lozouet, P., Phuong, M.A., Zaharias, P., Puillandre, N., 2018. Exon-capture-based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). Mol. Biol. Evol. 35, 2355–2374. https://doi.org/10.1093/molbev/msy144

Abrams, K.M., King, R.A., Guzik, M.T., Cooper, S.J.B., Austin, A.D., 2013. Molecular phylogenetic, morphological and biogeographic evidence for a new genus of parabathynellid crustaceans (Syncarida: Bathynellacea) from groundwater in an ancient southern Australian landscape. Invertebr. Syst. 27, 146–172. https://doi.org/10.1071/IS12033

Alley, N.F., Lindsay, J.M., 1995. Tertiary, in: Drexell, J.F., Preiss, W.V. (Eds.), The Geology of South Australia. South Australian Geological Survey Bulletin 54, South Australia, pp. 151–217.

Bauza-Ribot, M.M., Juan, C., Nardi, F., Oromi, P., Pons, J., Jaume, D., 2012. Mitogenomic phylogenetic analysis supports continental-scale vicariance in subterranean thalassoid crustaceans. Curr. Biol. 22, 2069–2074. https://doi.org/10.1016/j.cub.2012.09.012

Beard, J.S., 1977. Tertiary evolution of the Australian flora in the light of latitudinal movements of the continent. J. Biogeogr. 4, 111–118. https://doi.org/10.2307/3038133

Beard, J.S., 1998. Position and development history of the central watershed of the Western Shield, Western Australia. J. R. Soc. West. Aust. 81, 157–164.

Beaulieu, J.M., Tank, D.C., Donoghue, M.J., 2013. A Southern Hemisphere origin for campanulid angiosperms, with traces of the break-up of Gondwana. BMC Evol. Biol. 13, 80. https://doi.org/10.1186/1471-2148-13-80

Beheregaray, L.B., 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol. Ecol. 17, 3754–3774. https://doi.org/10.1111/j.1365-294X.2008.03857.x

Blom, M.P.K., Bragg, J.G., Potter, S., Moritz, C., 2017. Accounting for uncertainty in gene-tree estimation: summary-coalescent species tree inference in a challenging radiation of Australian lizards. Syst. Biol. 66, 352–366. https://doi.org/10.1093/sysbio/syw089

Bouckaert, R.R., Heled, J., Kühnert, D., Vaughan, T., Wu, C-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A., 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bowler, J.M., 1976. Aridity in Australia: age, origins and expressions in aeolian landforms and sediments. Earth-Sci. Rev. 12, 279–310. https://doi.org/10.1016/0012-8252(76)90008-8

Byrne, M., 2008. Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. Quat. Sci. Rev. 27, 2576–2585. https://doi.org/10.1016/j.quascirev.2008.08.032

Byrne, M., Yeates, D.K., Joseph, L., Kearney, M., Bowler, J., Williams, M.A.J., Cooper, S., Donnellan, S.C., Keogh, J.S., Leys, R., Melville, J., Murphy, D.J., Porch, N., Wyrwoll, K-H., 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. Mol. Ecol. 17, 4398–4417. https://doi.org /10.1111/j.1365-294X.2008.03899.x

Chilton, C., 1920. On a new isopodan genus (family Oniscidae) from Lake Corangamite, Victoria. Proc. Linn. Soc. N.S.W. 44, 723–734.

Cho, J-L., Humphreys, W.F., Lee, S-D., 2006. Phylogenetic relationships within the genus *Atopobathynella Schminke*, 1973 (Bathynellacea, Parabathynellidae): with the description of six new species from Western Australia. Invertebr. Syst. 20, 9–41. https://doi.org/10.1071/IS05019

Cooper, S.J.B., Bradbury, J.H., Saint, K.M., Leys, R., Austin, A.D., Humphreys, W.F., 2007. Subterranean archipelago in the Australian arid zone: mitochondrial DNA phylogeography of amphipods from central Western Australia. Mol. Ecol. 16, 1533–1544. https://doi.org/10.1111/j.1365-294X.2007.03261.x

Cooper, S.J.B., Hinze, S., Leys, R., Watts, C.H.S., Humphreys, W.F., 2002. Islands under the desert: molecular systematics and evolutionary origins of stygobitic water beetles (Coleoptera: Dytiscidae) from central Western Australia. Invertebr. Syst. 16, 589–598. https://doi.org/10.1071/IT01039

Cooper, S.J.B., Saint, K.M., Taiti, S., Austin, A.D., Humphreys, W.F., 2008. Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. Invertebr. Syst. 22, 195–203. https://doi.org/10.1071/IS07039

Davis, J., Pavlova, A., Thompson, R., Sunnucks, P., 2013. Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. Global Change Biol. 19, 1970–1984. https://doi.org/10.1111/gcb.12203

Edgar, R. C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

English, P., Spooner, N.A., Chappel, J., Questiaux, D.G., Hill, N.G., 2001. Lake Lewis basin, central Australia: environmental evolution and OSL chronology. Quatern. Int. 83, 81–101. https://doi.org/10.1016/S1040-6182(01)00032-5

Finston, T.L., Francis, C.J., Johnson, M.S., 2009. Biogeography of the stygobitic isopod *Pygolabis* (Malacostraca: Tainisopidae) in the Pilbara, Western Australia: Evidence for multiple colonisations of the groundwater. Mol. Phylogenet. Evol. 52, 448–460. https://doi.org/10.1016/j.ympev.2009.03.006

Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. arXiv: arXiv:1207.3907v2

Guzik, M.T., Abrams, K.M., Cooper, S.J.B., Humphreys, W.F., Cho, J-L., Austin, A.D., 2008. Phylogeography of the ancient Parabathynellidae (Crustacea: Bathynellacea) from the Yilgarn region of Western Australia. Invertebr. Syst. 22, 205–216. https://doi.org/10.1071/IS07040

Guzik, M.T., Adams, M.A., Murphy, N.P., Cooper, S.J.B., Austin, A.D., 2012. Desert springs: deep phylogenetic structure in an ancient endemic crustacean (*Phreatomerus latipes*). PLoS ONE 7, e37642. https://doi.org/10.1371/journal.pone.0037642

Guzik, M.T., Austin, A.D., Cooper, S.J.B., Harvey, M.S., Humphreys, W.F., Bradford, T., Eberhard, S.M., King, R.A., Leys, R., Muirhead, K.A., Tomlinson, M., 2011. Is the subterranean fauna uniquely diverse? Invertebr. Syst. 24, 407–418. https://doi.org/10.1071/IS10038

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F., Austin, A.D., 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. Invertebr. Syst. 33, 556–574. https://doi.org/10.1071/IS18070

Habel, J.C., Assmann, T., Schmitt, T., Avise, J.C., 2010. Relict species: from past to future, in: Habel J.C., Assmann, T. (Eds.), Relict species: phylogeography and conservation biology. Springer, Heidelberg, Germany, pp. 1–8.

Habermehl, M.A., 1980. The Great Artesian Basin, Australia. BMR J. Aust. Geol. Geophys. 5, 9–38.

Habermehl, M.A., 1982. Springs in the Great Artesian Basin, Australia: their origin and nature, Report No. 235. Bureau of Mineral Resources, Australia.

Harvey, M.S., 2002. Short-range endemism among the Australian fauna: some examples from non-marine environments. Invertebr. Syst. 16, 555–570. https://doi.org/10.1071/IS02009

Hewitt, G.M., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913. https://doi.org/10.1038/35016000

Hill, R.S., 1994. The history of selected Australian taxa, in: Hill, R.S. (Ed.), History of Australian Vegetation: Cretaceous to Recent. Cambridge University Press, Cambridge, UK, pp. 290–419.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281

Hugall, A.F., O'Hara, T.D., Hunjan, S., Nilsen, R., Moussalli, A., 2016. An exon-capture system for the entire class Ophiuroidea. Mol. Biol. Evol. 33, 281–294. https://doi.org/10.1093/molbev/msv216

Humphreys, W.F., 2006. Aquifers: the ultimate groundwater-dependent ecosystems. Aust. J. Bot. 54, 115–132. https://doi.org/10.1071/BT04151

Humphreys, W.F., 2008. Rising from down under: developments in subterranean biodiversity in Australia from a groundwater fauna perspective. Invertebr. Syst. 22, 85–101. https://doi.org/10.1071/IS07016

Humphreys, W.F., 2009. Hydrogeology and groundwater ecology: does each inform the other? Hydrogeology 17, 5–21. https://doi.org/10.1007/s10040-008-0349-3

Javidkar, M., Cooper, S.J.B., Humphreys, W.F., King, R.A., Judd, S., Austin, A.D., 2017. Biogeographic history of subterranean isopods from groundwater calcrete islands in Western Australia. *Zool. Scripta 47, 206–220.* https://doi.org/10.1111/zsc.12265

Javidkar, M., Cooper, S.J.B., King, R.A., Humphreys, W.F., Austin, A.D., 2015. Molecular phylogenetic analyses reveal a new southern hemisphere oniscidean family (Crustacea: Isopoda) with a unique water transport system. Invertebr. Syst. 29, 554–577. https://doi.org/10.1071/IS15010

Javidkar, M., Cooper, S.J.B., King, R.A., Humphreys, W.F., Bertozzi, T., Stevens, M.I., Austin, A.D., 2016. Molecular systematics and biodiversity of oniscidean isopods in the groundwater calcretes of central Western Australia. Mol. Phylogenet. Evol. 104, 83–98. https://doi.org/10.1016/j.ympev.2016.07.026

Juan, C., Guzik, M.T., Jaume, D., Cooper, S.J.B., 2010. Evolution in caves: Darwin's 'wrecks of ancient life' in the molecular era. Mol. Ecol. 19, 3865–3880. https://doi.org/10.1111/j.1365-294X.2010.04759.x

Juan, C., Emerson, B.C., 2010. Evolution underground. Shedding light on the diversification of subterranean insects. J Biol. 9, 17. https://doi.org/10.1186/jbiol227

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yares, C.J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D., Franklin, S.T., 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. Glob. Ecol. Biogeogr. 21, 393–404. https://doi.org/10.1111/j.1466-8238.2011.00686.x

Ketmaier, V., Argano, R., Caccone, A., 2003. Phylogeography and molecular rates of subterranean aquatic stenasellid isopods with a peri-Tyrrhenian distribution. Mol. Ecol. 12, 547–555. https://doi.org/10.1046/j.1365-294X.2003.01734.x

King, R.A., 2009. Two new genera and species of chiltoniid amphipods (Crustacea: Amphipoda: Talitroidea) from freshwater mound springs in South Australia. Zootaxa 2293, 35–52. https://doi.org/10.5281/zenodo.191467

King, R.A., Bradford, T., Austin, A.D., Humphreys, W.F., Cooper, S.J.B., 2012. Divergent molecular lineages and not-so-cryptic species: the first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. J. Crust. Biol. 32, 465–488. https://doi.org/10.1163/193724012X626566

Kleckova, I, Cesanek, M., Fric, Z., Pellissier, L., 2015. Diversification of the cold-adapted butterfly genus *Oeneis* related to Holarctic biogeography and climate niche shifts. Mol. Phylogenet. Evol. 92, 255–265. https://doi.org/10.1016/j.ympev.2015.06.012

Krieg, G.W., 1989. Geology, in: Zeidler, W., Ponder, W.F. (Eds.), Natural history of Dalhousie Springs. South Australian Museum, Adelaide, Australia, pp. 19–26.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14(1), 82. https://doi.org/10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34(3), 772–773. http://doi.org/10.1093/molbev/msw260

Lawver, L.A., Gahagan, L.M., 2003. Evolution of Cenozoic seaways in the circum-Antarctic region. Palaeogeogr. Palaeoclimatol. Palaeoecol. 198, 11–37. https://doi.org/10.1016/S0031-0182(03)00392-4

Leys, R., Watts, C.H.S., Cooper, S.J.B., Humphreys, W.F., 2003. Evolution of subterranean diving beetles (Coleoptera: Dytiscidae: Hydroporini, Bidessini) in the arid zone of Australia. Evolution 57, 2819–2834. https://doi.org/10.1111/j.0014-3820.2003.tb01523.x

Li, H., Durbin, R., 2008. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Martin, H.A., 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. J. Arid Environ. 66, 533–563. https://doi.org/10.1016/j.jaridenv.2006.01.009

Mayer, C., Sann, S., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., Petersen, M., Meusemann, K., Liere, K., Wäggle, J-W., Misof, M., Bleidorn, C., Ohl, M., Niehuis, O., 2016. BaitFisher: a software package for multispecies target DNA enrichment probe design. Mol. Biol. Evol. 33, 1875–1886. https://doi.org/10.1093/molbev/msw056

Meyer, B., Meusemann, K., Misof, B., 2011. MARE v0.1.2-rc. http://www.zfmk.de/bioinformatics/MARE_v0.1.2-rc.tar.gz (accessed 24 April 2019).

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees, in: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA, pp. 1–8.

Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., Frandsen, P., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, S., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B. M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M., Zhou, X., 2014.

Phylogenomics resolves the timing and pattern of insect evolution. Science 346, 763–767. https://doi.org/10.1126/science.1257570

Misof, B., Meyer, B., von Reumont, B. M., Kück, P., Misof, K., Meusemann, K., 2013. Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. BMC Bioinformatics 14, 348. https://doi.org/10.1186/1471-2105-14-348

Moritz, C., Agudo, R., 2013. The future of species under climate change: resilience or decline? Science 341, 504–508. https://doi.org/10.1126/science.1237190

Moritz, C., Pratt, R.C., Bank, S., Bourke, G., Bragg, J.G., Doughty, P., Keogh, J.S., Laver, R.J., Potter, S., Teasdale, L.C., Tedeschi, L.G., Oliver, P.M., 2018. Cryptic lineage diversity, body size divergence, and sympatry in a species complex of Australian lizards (*Gehyra*). Evolution 72, 54–66. https://doi.org/10.1111/evo.13380

Murphy, N.P., Adams, M., Austin, A.D., 2009. Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. Mol. Ecol. 18, 109–122. https://doi.org/10.1111/j.1365-294X.2008.04007.x

Murphy, N.P., Adams, M., Guzik, M.T., Austin, A.D., 2013. Extraordinary micro-endemism in Australian desert spring amphipods. Mol. Phylogenet. Evol. 66, 645–653. https://doi.org/10.1016/j.ympev.2012.10.013

Murphy, N.P., Breed, M.F., Guzik, M.T., Cooper, S.J.B., Austin, A.D., 2012. Trapped in desert springs: phylogeography of Australian desert spring snails. J. Biogeogr. 39, 1573–1582. https://doi.org/10.1111/j.1365-2699.2012.02725.x

Murphy, N.P., Guzik, M.T., Cooper, S.J.B., Austin, A.D., 2015. Desert spring refugia: museums of diversity or evolutionary cradles? Zool. Scripta 44, 693–701. https://doi.org/10.1111/zsc.12129

Nicholls, G.E., Barnes, H.E., 1926. Description of a new species of terrestrial isopod, *Haloniscus stepheni*, from Western Australia. J. R. Soc. West. Aust. 12, 87–96.

Nguyen, L-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2014. IQ-Tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300

Ogilvie, H.A., Bouckhaert, R.R., Drummond, A.J., 2017. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. Mol. Biol. Evol. 34, 2101–2114. https://doi.org/10.1093/molbev/msx126

O'Hara, T.D., Hugall, A.F., Cisternas, P.A., Boissin, E., Bribiesca-Contreras, G., Sellanes, J., Paulay, G., Byrne, M., 2019. Phylogenomics, life history and morphological evolution of ophiocomid brittlestars. Mol. Phylogenet. Evol. 130, 67–80. https://doi.org/10.1016/j.ympev.2018.10.003

Page, T.J., Humphreys, W.F., Hughes, J.M., 2008. Shrimps down under: evolutionary relationships of subterranean crustaceans from Western Australia (Decapoda: Atyidae: *Stygiocaris*). PLoS ONE 3, e1618. https://doi.org/10.1371/journal.pone.0001618

Pepper, M., Sumner, J., Brennan, I.G., Hodges, K., Lemmon, A.R., Lemmon, E.M., Peterson, G., Rabosky, D.L., Schwarzkopf, L., Scott, I.A.W., Shea, G., Keogh, J.S., 2018. Speciation in the mountains and dispersal by rivers: Molecular phylogeny of *Eulamprus* water skinks and the biogeography of Eastern Australia. J. Biogeogr. 45, 2040–2052. https://doi.org/10.1111/jbi.13385

Petersen, M., Meusemann, K., Donath, A., Dowling, D., Liu, S., Peters, R.S., Podsiadlowski, L., Vasilikopoulos, A., Zhou, X., Misof, B., Niehuis, O., 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinformatics 18, 1–10. https://doi.org/10.1186/s12859-017-1529-8

Ponder, W.F., Colgan, D.J., Terzis, T., Clark, S.A., Miller, A.C., 1996. Three new morphologically and genetically determined species of hydrobiid gastropods from Dalhousie Springs, northern South Australia, with the description of a new genus. Moll. Res. 17, 49–109. https://doi.org/10.1080/13235818.1996.10673675

Ponder, W.F., Eggler, P., Colgan, D.J., 1995. Genetic differentiation of aquatic snails (Gastropoda: Hydrobiidae) from artesian springs in arid Australia. Biol. J. Linn. Soc. 56, 553–596. https://doi.org/10.1111/j.1095-8312.1995.tb01110.x

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rambaut, A., 2014. FigTree 1.4.2 software, Institute of Evolutionary Biology, Univ. Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/ (accessed 9 July 2014).

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Sys. Biol. Syy032. https://doi.org/10.1093/sysbio/syy032

Reilly, S.B., Stubbs, A.L., Karin, B.R., Bi, K., Arida, E., Iskandar, D.T., McGuire, J.A., 2019. Leap-frog dispersal and mitochondrial introgression: Phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. J. Biogeogr. 46, 757–769. https://doi.org/10.1111/jbi.13526

Rix, M.G., Cooper, S.J.B., Meusemann, K., Klopfstein, S., Harrison, S.E., Harvey, M.S., Austin, A.D., 2017. Post-Eocene climate change across continental Australia and the diversification of Australian spiny trapdoor spiders (Idiopidae: Arbanitinae). Mol. Phylogenet. Evol. 109, 302–320. https://doi.org/10.1016/j.ympev.2017.01.008

Sniderman, J.M.K., Pillans, B., O'Sullivan, P.B., Kershaw, A.P., 2007. Climate and vegetation in southeastern Australia respond to Southern Hemisphere insolation forcing in the late Pliocene-early Pleistocene. Geology 35, 41–44. https://doi.org/10.1130/G23247A.1

Sniderman, J.M.K., Woodhead, J.D., Hellstroma, J., Jordan, G.J., Drysdale, R.N., Tyler, J.J., Porch, N., 2016. Pliocene reversal of late Neogene aridification. Proc. Natl. Acad. Sci. U.S.A. 113, 1999–2004. https://doi.org/10.1073/pnas.1520188113

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stringer, D.N., King, R.A., Taiti, S., Guzik, M.T., Cooper, S.J.B., Austin, A.D., 2019. Systematics of the isopod genus *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae) with description of four new species from Great Artesian Basin springs in South Australia. J. Crust. Biol. 39(5), 651–668. https://doi.org/10.1093/jcbiol/ruz044

Taiti, S., Ferrara, F., Iliffe, T.M., 1995. A new species of *Haloniscus* Chilton, 1920 from New Caledonia (Isopoda: Oniscidea). Crustaceana 68, 321–328. https://doi.org/10.1163/156854095X00502

Taiti, S., Humphreys, W.F., 2001. New aquatic Oniscidea (Crustacea, Isopoda) from groundwater calcretes of Western Australia. Rec. West. Aust. Mus 64, 133–151. https://doi.org/10.18195/issn.0313-122x.64.2001.133-151

Taiti, S., Xue, Z., 2012. The cavernicolous genus *Trogloniscus* nomen novum, with descriptions of four new species from southern China (Crustacea, Oniscidea, Styloniscidae). Tropical Zool. 25, 183–209. https://doi.org/10.1080/03946975.2012.751240

*T*horvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2012. Integrative Genomics Viewer (IGV): high-performance genomics data visualisation and exploration. Briefings in Bioinformatics 14, 178–192. https://doi.org/10.1093/bib/bbs017

Vences, M., Wollenberg, K.C., Vieites, D.R., Lees, D.C., 2009. Madagascar as a model region of species diversification. Trends Ecol. Evol. 24, 456–465. https://doi.org/10.1016/j.tree.2009.03.011

Watts, C.H.S., Humphreys, W.F., 2006. Twenty-six new Dytiscidae (Coleoptera) of the genera *Limbodessus* Guignot and *Nirripirti* Watts & Humphreys, from underground waters in Australia. Trans. Royal Soc. South Aust. 130, 123–185. https://doi.org/10.1080/3721426.2006.10887055

Wood, H.M., González, V.L., Lloyd, M., Coddington, J., Scharff, N., 2018. Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). Mol. Phylogenet. Evol. 127, 907–918. https://doi.org/10.1016/j.ympev.2018.06.038

Yu, Y., Harris, A.J., Blair, C., He, X., 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46–49. https://doi.org/10.1016/j.ympev.2015.03.008

CHAPTER 4:

Systematics of *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae), with description of four new species from threatened Great Artesian Basin springs in South Australia

(published paper)

Statement of Authorship

Title of Paper	Systematics of *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae), with description of four new species from threatened Great Artesian Basin springs in South Australia
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Stringer, D.N., King, R.A., Taiti, S., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2019. Systematics of *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae), with description of four new species from threatened Great Artesian Basin springs in South Australia. *Journal of Crustacean Biology*. https://doi.org/10.1093/jcbiol/ruz044

Principal Author

Name of Principal Author (Candidate)	Danielle Stringer		
Contribution to the Paper	Conducted specimen identifications, morphological taxonomic analyses, and compiled/inked figures (under guidance from isopod taxonomic experts, Rachael King and Stefano Taiti). Also wrote the species descriptions, key and manuscript, and acted as corresponding author.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	8/7/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Rachael King		
Contribution to the Paper	Supervised development of work, provided training in taxonomic practices, offered advice on species descriptions and figures, and critically reviewed manuscript.		
Signature		Date	04/07/2019

Name of Co-Author	Stefano Taiti		
Contribution to the Paper	Provided preliminary information on morphological characters for new species, guidance on species descriptions, the generic key and figures, and critically reviewed manuscript.		
Signature		Date	28/6/2019

Name of Co-Author	Michelle Guzik
Contribution to the Paper	Contributed unpublished (at the time) results from molecular phylogenetics study of *Haloniscus*, which informed the morphological analyses (Appendix 1). Also supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

Name of Co-Author	Steven Cooper
Contribution to the Paper	Supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

Name of Co-Author	Andrew Austin
Contribution to the Paper	Supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

Journal of Crustacean Biology Advance Access published 1 July 2019

The Crustacean Society

Journal of Crustacean Biology

Journal of Crustacean Biology 39(5), 651–668, 2019. doi:10.1093/jcbiol/ruz044

Systematics of *Haloniscus* Chilton, 1920 (Isopoda: Oniscidea: Philosciidae), with description of four new species from threatened Great Artesian Basin springs in South Australia

Danielle N. Stringer[1], Rachael A. King[1,2], Stefano Taiti[3,4], Michelle T. Guzik[1], Steven J.B. Cooper[1,2] and Andrew D. Austin[1]

[1]*Australian Centre for Evolutionary Biology and Biodiversity, School of Biological Sciences, The University of Adelaide, North Terrace, South Australia, 5005, Australia;*
[2]*South Australian Museum, North Terrace, Adelaide, South Australia, 5000, Australia;*
[3]*Istituto di Ricerca sugli Ecosistemi Terrestri, Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Florence, Italy; and*
[4]*Museo di Storia Naturale dell'Università, Sezione di Zoologia La Specola, Via Romana 17, 50125 Florence, Italy*

Correspondence: D.N. Stringer; e-mail: danielle.stringer@adelaide.edu.au

(Received 1 February 2019; accepted 23 May 2019)

ABSTRACT

Recent surveys of Australian arid-zone groundwater ecosystems have uncovered considerable species diversity and extreme endemism for the oniscidean isopod genus *Haloniscus* Chilton, 1920. Phylogenetic and species delimitation analyses have recognised several distinct species from the Great Artesian Basin springs in South Australia, inspiring a morphological reassessment of the genus and examination of specimens from the iconic Lake Eyre and Dalhousie Springs. We present a revised diagnosis of *Haloniscus*, transfer the genus from the family Scyphacidae to Philosciidae and describe four new species, *H. fontanus* Stringer, King & Taiti **n. sp.**, *H. microphthalmus* Stringer, King & Taiti **n. sp.**, *H. rotundus* Stringer, King & Taiti **n. sp.**, and *H. yardiyaensis* Stringer, King & Taiti **n. sp.**, based on combined morphological and molecular evidence. We compare the results of molecular-based species delimitation analyses with morphological data, provide distribution information, and present a key to the described species of *Haloniscus*. Two species presently included in *Andricophiloscia* Vandel, 1973, *A. stepheni* (Nicholls & Barnes, 1926) and *A. pedisetosa* Taiti & Humphreys, 2001, from Western Australia are also transferred to *Haloniscus*.

Key Words: *Andricophiloscia*, arid zones, conservation, groundwater, integrated taxonomy

INTRODUCTION

Exceptionally diverse invertebrate communities are known to inhabit the unique groundwater-dependent springs of the Great Artesian Basin (GAB) in South Australia (SA), with molecular studies uncovering numerous undescribed species and high levels of isolation and short-range endemism (Murphy *et al.*, 2009, 2012, 2013, 2015a; Guzik *et al.*, 2012). These desert springs represent the most reliable source of freshwater in this arid region, and are now regarded as significant areas for biodiversity (Ponder *et al.*, 1995; Ponder, 2003). The springs, nonetheless, are one of Australia's most threatened ecosystems due to aquifer drawdown from extensive water extraction from the GAB for mining and pastoral activities, and are, therefore, listed as an "endangered ecological community" under the Australian Environmental Protection and Biodiversity Act (Harris, 1992; Fensham & Price, 2004). A comprehensive understanding of the taxonomy, distribution, composition, and number of species is, consequently, vital to the conservation of important desert spring invertebrates for potential listing as threatened species (Taylor *et al.*, 2018).

The oniscidean isopod genus *Haloniscus* Chilton, 1920, in particular, was recently discovered from the GAB springs, with broad sampling, a multi-gene phylogeny, and species delimitation analyses revealing significant genetic diversity across taxa (Fig. 1) (Guzik *et al.*, 2019). The study estimated between three and eight new putative *Haloniscus* species using the delineation methods, Automatic Barcode Gap Discovery (ABGD) (Puillandre *et al.*, 2012) and Bayesian Poisson Tree Processes (bPTP) (Zhang *et al.*, 2013), respectively. This evidence, together with the discovery of new *Haloniscus* lineages from two further arid-zone groundwater ecosystems (Cooper *et al.*, 2008; Guzik *et al.*, 2019), has encouraged

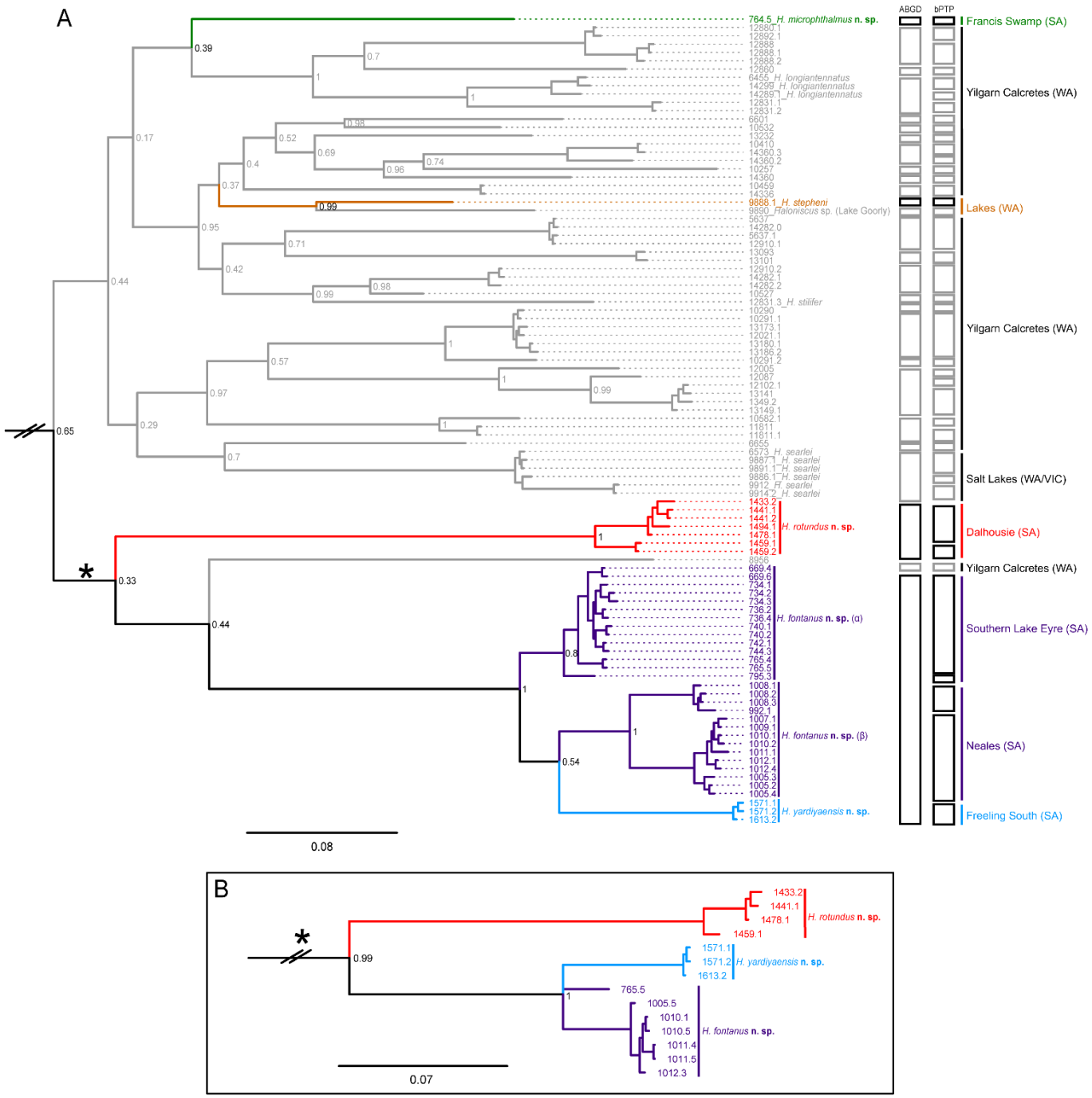


Figure 1. Molecular phylogenetic trees modified from Guzik *et al.* (2019) with colour-coded lineages corresponding to the new species of *Haloniscus*. Bayesian Inference COI phylogeny with *Haloniscus microphthalmus* **n. sp.** from Francis Swamp (green), *Haloniscus rotundus* **n. sp.** from Dalhousie Springs (red), *Haloniscus yardiyaensis* **n. sp.** from Freeling South Springs (blue), and *Haloniscus fontanus* **n. sp.** from (α) southern Lake Eyre and (β) northern Lake Eyre (Neales Complex) (purple) (**A**). The orange lineage corresponds to *Haloniscus stepheni*. Delineated species are indicated to the right of the phylogeny represented by outlined bars for the ABGD and bPTP methods, respectively. Subset (corresponding to the clade in the COI phylogeny denoted by * of the Bayesian Inference combined COI and 18S phylogeny with colour codes as above (**B**). Lineages not examined in this study are in grey. This figure is available in colour at *Journal of Crustacean Biology* online.

a morphological reassessment of the genus and formal descriptions of new species.

Haloniscus so far consists of five described species, with four from Australia and one from New Caledonia: *H. searlei* Chilton, 1920, with a wide distribution in salt lakes across Western Australia (WA), Victoria, Tasmania, and SA; *H. longiantennatus* Taiti & Humphreys, 2001, *H. stilifer* Taiti & Humphreys, 2001, and *H. tomentosus* Taiti & Humphreys, 2001 from subterranean calcrete aquifers in WA; and *H. anophthalmus* Taiti, Ferrara & Iliffe, 1995 found in anchialine limestone cave waters in the Isle of Pines, New Caledonia. We describe four new species of *Haloniscus* from SA GAB springs and compare the results of species delimitation analyses with morphological data. In light of the molecular phylogenetic results of Guzik *et al.* (2019) and our morphological study, we also reassess the family level placement of *Haloniscus* as well as the generic status of *Andricophiloscia stepheni*

(Nicholls & Barnes, 1926) and *A. pedisetosa* Taiti & Humphreys, 2001 from WA.

MATERIALS AND METHODS

Researchers from The University of Adelaide and the South Australian Museum undertook extensive sampling across the South Australian spring supergroups, Dalhousie and Lake Eyre (Fig. 2). The GAB springs fall naturally into geographic hierarchical clusters where proximate springs form 'groups' and, at broader scales, 'complexes' that are hydrochemically and hydrogeologically similar (Habermehl, 1980). The springs have been grouped further into 13 major clusters, called 'supergroups', in SA, Queensland, and New South Wales (Ponder, 2002). Specimens of *Haloniscus* were collected opportunistically by hand from the wet margins of springs and stored in 100% ethanol. Type specimens were dissected along the left side (where possible) and appendages were slide-mounted and illustrated with a camera lucida attachment to a Nikon Eclipse 80i (Nikon, Tokyo, Japan) compound microscope. Total body length was measured through the mid-line of the specimen, from the head to the telson. All type material of the new spring species is lodged at the South Australian Museum (SAMA). We also re-examined the type material for *H. stepheni* and *H. pedisetosus*, deposited in the Western Australian Museum (WAM).

Species were delineated using an integrated approach along with the general lineage species concept (de Queiroz, 1998, 2007); i.e. with fixed morphological differences (analysis carried out here) and phylogenetic analyses of molecular data (all sequencing and associated molecular analyses conducted by Guzik *et al.* (2019) (Fig. 1)), providing the operational criteria to delimit species. Guzik *et al.* (2019) generated sequences from the mitochondrial cytochrome c oxidase subunit I (COI) gene and the 18S rRNA gene, and further calculated COI nucleotide divergences using the Kimura 2-parameter (K2P) model (Kimura, 1980), as implemented in MEGA 6.0 (Tamura *et al.*, 2013), with pertinent results included here (see below). Guzik *et al.* (2019) provided a detailed outline of the DNA extraction and sequencing techniques, as well as the molecular phylogenetic, COI divergence, and species delimitation (ABGD and bPTP) analyses undertaken.

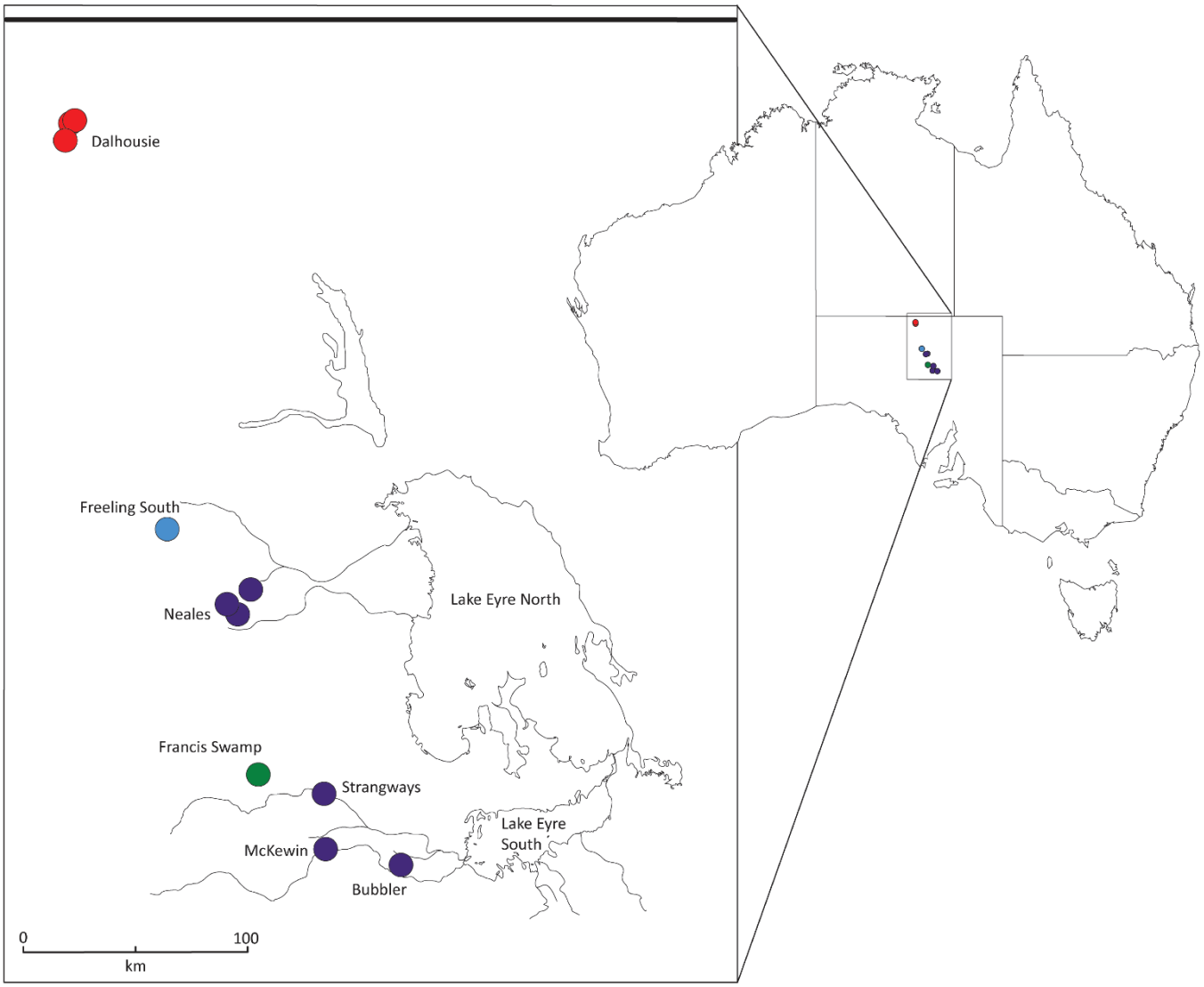


Figure 2. Map of sampled SA GAB spring locations where red represents the distribution of *Haloniscus rotundus* **n. sp.** across Dalhousie Springs (Main Pool, Kingfisher and Meeting Place), *Haloniscus yardiyaensis* **n. sp.** in Freeling South Springs (blue), *Haloniscus microphthalmus* **n. sp.** in Francis Swamp (green), and *Haloniscus fontanus* **n. sp.** across the Neales (Hawker and Fanny Springs), Strangways, McKewin and Coward (Bubbler Spring) Complexes (purple). This figure is available in colour at *Journal of Crustacean Biology* online.

SYSTEMATICS

Family Philosciidae Kinahan, 1857

Genus *Haloniscus* Chilton, 1920

Type species: *Haloniscus searlei* Chilton, 1920 by monotypy.

Species included: *Haloniscus anophthalmus* Taiti, Ferrara & Iliffe, 1995, *H. fontanus* Stringer, King & Taiti **n. sp.**, *H. longiantennatus* Taiti & Humphreys, 2001, *H. microphthalmus* Stringer, King & Taiti **n. sp.**, *H. pedisetosus* (Taiti & Humphreys, 2001) **n. comb.**, *H. rotundus* Stringer, King & Taiti **n. sp.**, *H. searlei* Chilton, 1920, *H. stepheni* Nicholls & Barnes, 1926, *H. stilifer* Taiti & Humphreys, 2001, *H. tomentosus* Taiti & Humphreys, 2001, and *H. yardiyaensis* Stringer, King & Taiti **n. sp.**

Amended diagnosis: Body elongated with pleon distinctly narrower than pereon. Noduli laterales on pereonites present in epigean and some subterranean species, secondarily absent in most subterranean species; when present, noduli laterales inserted at similar distance from lateral margins of pereonites. Cephalon with small, rounded lateral lobes not protruding frontward; supra-antennal line usually present (absent in *H. anophthalmus*); no frontal line. Pleon epimera with posterior points slightly reduced. Antennule of 3 articles with some aesthetascs at apex. Antennal flagellum of 3 articles, with short apical organ. Molar penicil of mandible dichotomised, consisting of tuft of plumose setae. Outer branch of maxillule with 10 or 11 teeth all with simple apex, flagelliform seta among outer group of teeth; inner branch with 2 subequal penicils, no posterior point. Endite of maxilliped setose with large penicil on medial corner. Pereopods with flagelliform dactylar, ungual seta. Pereopod 1 with cleaning device (or tuft of setae) for antennae slightly developed. Exopods of pleopods with marginal fringe of long, thin setae overlapping medially in aquatic forms (epigean and subterranean), no fringe of setae, not overlapping in epigean terrestrial forms; no respiratory structures. Uropod with protopod, exopod grooved on outer margin in some species, flattened in others; insertion of endopod slightly proximal to that of exopod.

Remarks: Despite sharing some characters with the family Philosciidae, *Haloniscus* was tentatively included in Scyphacidae by Taiti & Humphreys (2001), primarily based on the characters for the known, aquatic species: noduli laterales absent; uropods with protopods often enlarged, protopods and exopods with lateral margins not grooved. *Haloniscus stepheni*, described by Nicholls & Barnes (1926) from Kockatea Gully, Tenindewa, WA, was further transferred to *Andricophiloscia* Vandel, 1973 (type species: *A. melanesiensis* Vandel, 1973) from *Haloniscus* by Taiti & Humphreys (2001) and *A. pedisetosa* was described. These two Australian species were principally included in *Andricophiloscia* as they possessed characters listed in the generic diagnosis of Vandel (1973a), particularly in the presence of noduli laterales, and uropods with protopods and exopods grooved on the outer margin. These species are members of Philosciidae, but, according to molecular evidence from Guzik *et al.* (2019) (Fig. 1A), they do fall within *Haloniscus* and are transferred herein. *Haloniscus stepheni* is grouped with an undescribed *Haloniscus* species from Lake Goorly in WA with strong support (posterior probability of 99%) and is, furthermore, clustered in a larger clade of subterranean *Haloniscus* (containing *H. stilifer*) from calcrete aquifers in the Yilgarn region of WA (posterior probability of 95%) (Fig. 1A). It appears that the absence of noduli laterales and grooved lateral margins on the uropods, along with the occurrence of a marginal fringe of setae on the exopods of pleopods, are adaptations to an aquatic lifestyle and, as such, *Haloniscus* should be accommodated in the family Philosciidae rather than Scyphacidae.

Haloniscus now includes 11 species with some solely aquatic forms (*H. searlei*, *H. anophthalmus*, *H. longiantennatus, H. stilifer,* and *H. tomentosus*), whereas others are terrestrial or semi-terrestrial. The new SA GAB associated species (*H. fontanus* Stringer, King & Taiti **n. sp.**, *H. microphthalmus* Stringer, King & Taiti **n. sp.**, *H. rotundus* Stringer, King & Taiti **n. sp.**, and *H. yardiyaensis* Stringer, King & Taiti **n. sp.**) occur along the wet margins of springs, whereas *H. stepheni* is found beneath logs by creek banks. *Haloniscus pedisetosus* was collected within the same calcrete system in the Yilgarn (WA) as *H. longiantennatus*, but, according to Taiti & Humphreys (2001), it is unclear whether this species is equally aquatic. These semi-terrestrial species are all morphologically comparable (noduli laterales present on pereonites, uropods grooved on the lateral margins, and marginal fringe of setae on exopods of pleopods absent, characters in opposition to those in aquatic species as mentioned above), which again further seems to suggest that variation in morphological characters across *Haloniscus* species is likely associated with adaptations to different environments.

***Haloniscus fontanus* Stringer, King & Taiti n. sp.**

(Figs. 3–5)

Material examined: Holotype, male SAMA C13220 (GAB00765), Bubbler Spring, Coward Complex, South Australia, 29°26′46.9"S 136°51′28.8"E, coll. M. Guzik and N. Murphy, 3 November 2007. Paratypes: 4 males, 3 females SAMA C13221 (GAB00765), same collection data as holotype; 1 female SAMA C13222 (GAB00765.5; Genbank COI: KT236011, Genbank 18S: MK286387), same collection data as holotype; 3 males, 2 females SAMA C13223 (GAB00738), Strangways Springs, South Australia, 29°09′31.4"S 136°32’37.6”E, coll. M. Guzik and N. Murphy, 1 November 2007.

Additional material: 1 female SAMA C13224 (GAB00744.4; Genbank COI: KT236007), Strangways Springs, South Australia, 29°09′35.0"S 136°33′04.2"E, coll. M. Guzik and N. Murphy, 1 November 2007; 2 males, 1 female SAMA C13225 (GAB00669), McKewin Spring, South Australia, 29°23′10.3"S 136°32′48.8"E, coll. M. Guzik and N. Murphy, 3 November 2007; 1 female SAMA C13226 (GAB00669.4; Genbank COI: KT235998), McKewin Spring, South Australia, 29°23′10.3"S 136°32′48.8"E, coll. M. Guzik and N. Murphy, 3 November 2007; 1 male and 1 female SAMA C13228 (GAB01007), Hawker Springs, Neales Complex, South Australia, 28°25′30.2"S 136°11′09.9"E, coll. M. Guzik and N. Murphy, 27 August 2008; 1 male SAMA C13227 (GAB01007.4; Genbank COI: KT236017), Hawker Springs, Neales Complex, South Australia, 28°25′30.2"S 136°11′09.9"E, coll. M. Guzik and N. Murphy, 27 August 2008; 1 male SAMA C13230 (GAB01008.3; Genbank COI: KT236020), 1 female SAMA C13229 (GAB01008.1; Genbank COI: KT236018), Hawker Springs, Neales Complex, South Australia, 28°23′04.1"S 136°09′03.7"E, coll. M. Guzik and N. Murphy, 27 August 2008; 1 male SAMA C13232 (GAB01011.5; Genbank COI: KT236024, Genbank 18S: MK286387), 1 female SAMA C13231 (GAB01011.1; Genbank COI: KT236024), Fanny Springs, Neales Complex, South Australia, 28°19′22.2"S 136°14′16.1"E, coll. M. Guzik and N. Murphy, 29 August 2008.

Diagnosis: Antennae reaching past pereonite 2. Male pereopods 1–3 with carpus, merus bearing thick brush of long setae on sternal margin. Male pleopod 2 exopod with around 10 robust setae on outer margin.

Description: Maximum body length: male 4.5 mm, female 5 mm. Colour in alcohol pale, speckled with brown on anterior parts of

pereonites, pleonites, more evident in medial area of the body. Body (Fig. 3A) moderately convex, elongated, about 3× as long as wide, with pleon distinctly narrower than pereon. Dorsum smooth with triangular scale-setae (Fig. 3A, B); noduli laterales present on pereonites with one per side on each segment, inserted at similar distance from lateral margins of pereonites; b/c and d/c co-ordinates as in Fig. 3C.

Cephalon (Fig. 3D–F) with small, rounded lateral lobes; supra-antennal line present, slightly sinuous. Eyes with 10–12 ommatidia. Pereonites 1–4 with straight posterior margins, right-angled

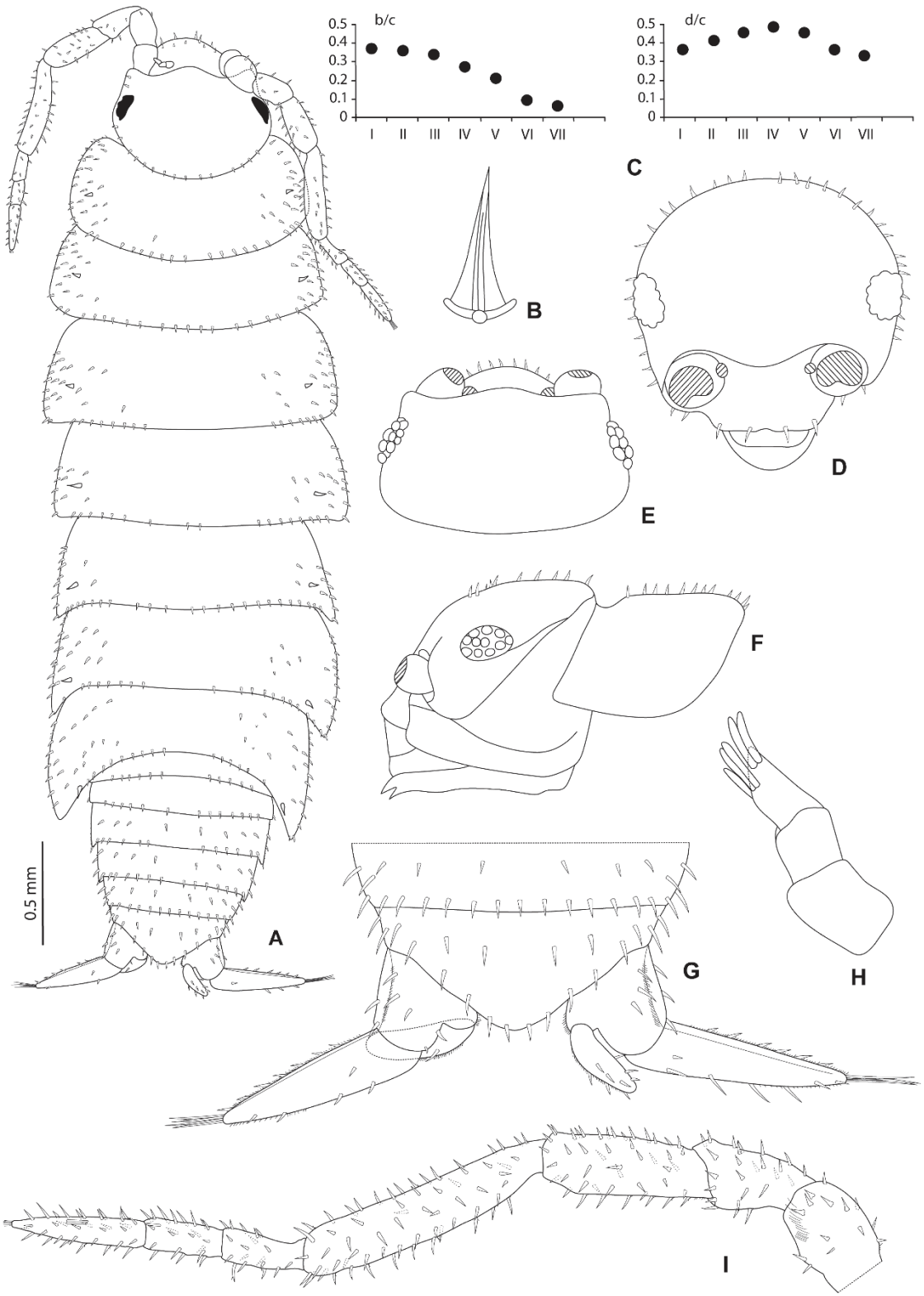


Figure 3. *Haloniscus fontanus* **n. sp.** holotype male (A, G–I), paratype female (B–F). Whole specimen, dorsal view (**A**); dorsal scale-seta (**B**); coordinates of the noduli laterales (**C**); cephalon, frontal view (**D**); cephalon, dorsal view (**E**); cephalon and pereonite 1, lateral view (**F**); telson and uropods (**G**); antennule (**H**); antenna (**I**).

posterior corners; pereonites 5–7 with posterior corners progressively more acute. Pleonites 3–5 with posterior points reduced (Fig. 3A, G). Telson (Fig. 3A, G) almost twice as wide as long, distal part with straight sides, rounded apex.

Antennule (Fig. 3H) with first, third articles longer than second, 2 aesthetascs at apex, tuft of 3 aesthetascs subapically. Antenna (Fig. 3I) short, reaching past pereonite 2; fifth article of peduncle slightly shorter than flagellum; flagellum with first, second articles subequal in length, third longer; second, third articles with 1, 2 aesthetascs, respectively.

Mandibles (Fig. 4A, B) with molar penicil consisting of 4 plumose setae, 2 + 1 free penicils on the left, 1 + 1 on the right mandible. Maxillule (Fig. 4C) inner branch apically rounded, with 2 large subequal penicils; outer branch with 4 + 6 simple large teeth, small tooth, flagelliform seta among outer group of teeth. Maxilla (Fig. 4D) with outer lobe about 1.5× wider than inner lobe, both covered with fine setae on distal ends. Maxilliped (Fig. 4E) with distal part of palp without visible transverse suture; proximal article of palp with 2 robust setae.

Uropod (Fig. 4F) protopod, exopod grooved on outer margin, insertion of endopod proximal to that of exopod, exopod about twice as long as endopod.

Male: Pereopods 1 (Fig. 5A)–3 with carpus, merus bearing brush of long setae on sternal margin. Pereopod 7 (Fig. 5B) ischium with sternal margin straight to slightly concave in proximal half.

Pleopod 1 (Fig. 5C) exopod apically rounded, with no setae along margin; endopod significantly longer than exopod, with pointed apical part bent outwards. Pleopod 2 (Fig. 5D) exopod

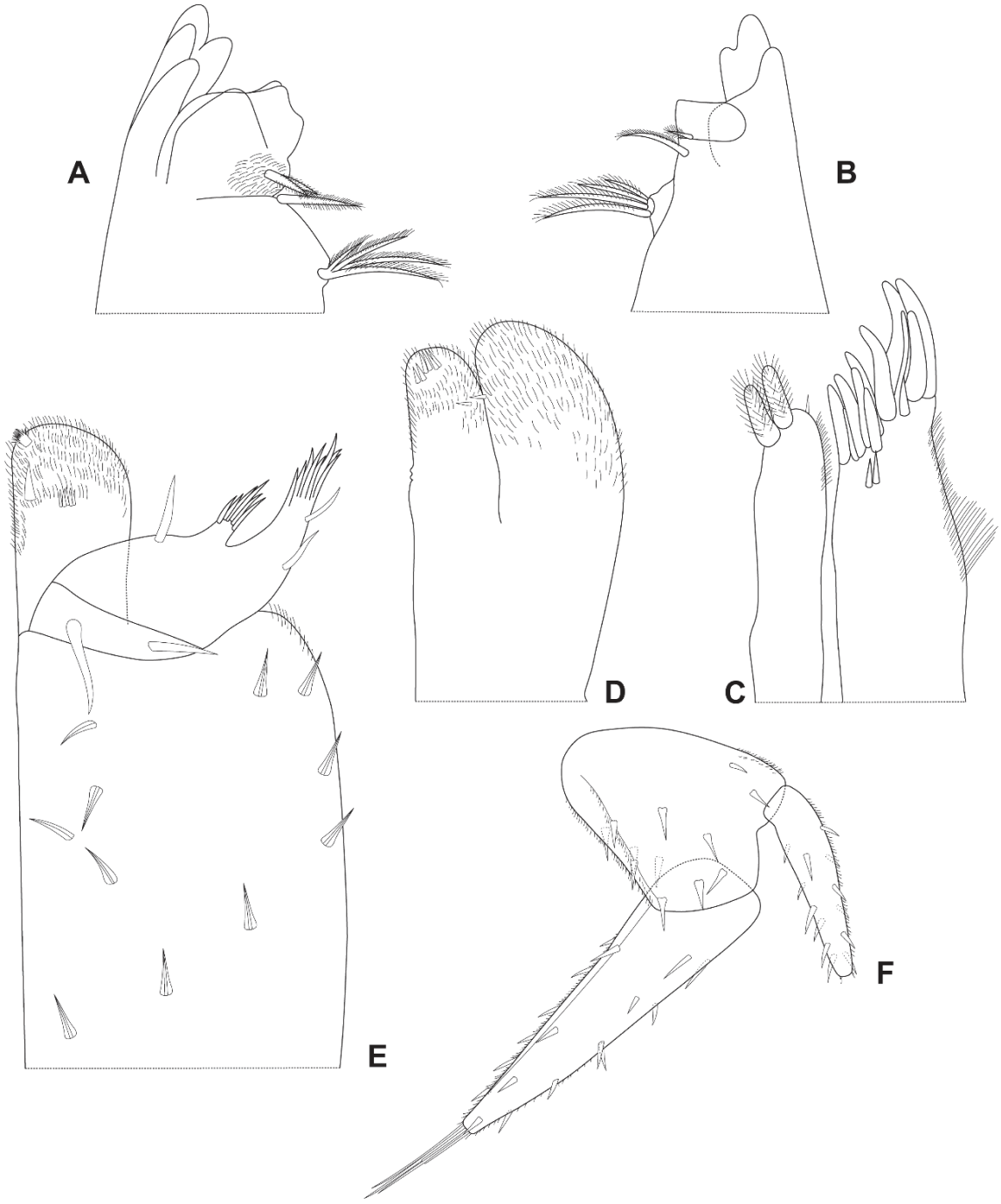


Figure 4. *Haloniscus fontanus* **n. sp.** holotype male. Left mandible (**A**); right mandible (**B**); maxillule (**C**); maxilla (**D**); maxilliped (**E**); uropod (**F**).

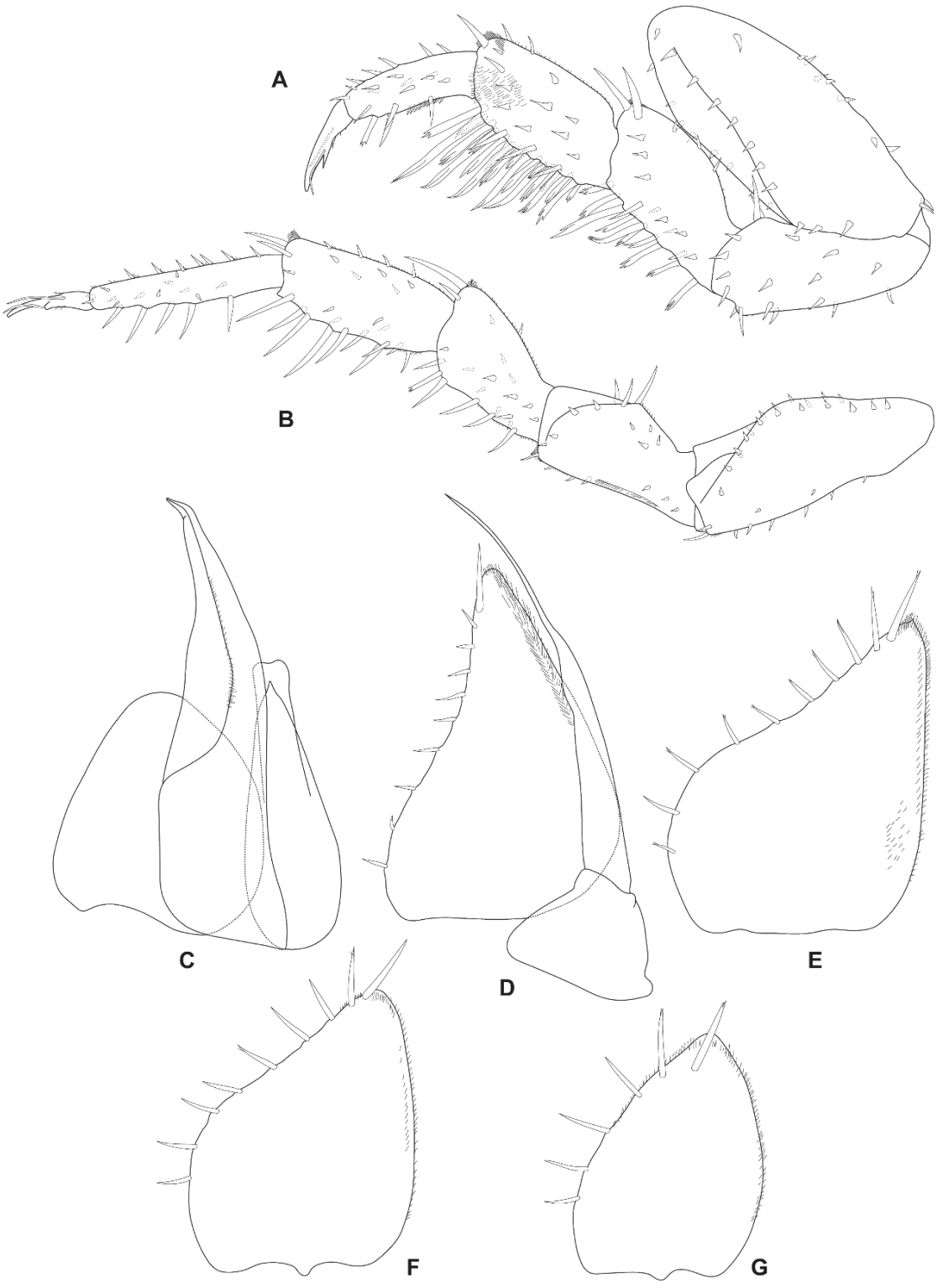


Figure 5. *Haloniscus fontanus* **n. sp.** holotype male. Pereopod 1 (**A**); pereopod 7 (**B**); pleopod 1 (**C**); pleopod 2 (**D**); pleopod 3 exopod (**E**); pleopod 4 exopod (**F**); pleopod 5 exopod (**G**).

triangular with 10 robust setae on outer margin, endopod distinctly longer than exopod, with flagelliform distal end. Pleopods 3, 4 (Fig. 5E, F) exopods triangular with outer margin comprising 9, 10 robust setae. Pleopod 5 (Fig. 5G) exopod triangular with 6 or 7 robust setae on outer margin. Pleopods 2–5 exopods with short setae on medial margin.

Etymology: The species name, *fontanus*, refers to the Latin word for 'related to spring', and is a reference to the Great Artesian Basin springs in South Australia where the species occurs.

Remarks: *Haloniscus fontanus* **n. sp.** is morphologically similar to *H. yardiyaensis* **n. sp.** and *H. stepheni*, but differs predominantly in

antenna length (reaching past pereonite 2) and in the male characters where pereopods 1–3 consist of longer, thicker brushes of setae on the carpus and merus, and the exopods of pleopods 2–5 comprise a larger number of robust setae on the outer margin. As well as being the most alike morphologically, the combined phylogeny (Fig. 1B) suggests that *H. fontanus* is most closely related to *H. yardiyaensis*, with a 10–12% COI divergence (Guzik *et al.*, 2019), of the newly described GAB species. *Haloniscus stepheni* presumably displays similar characters to the SA GAB species, in particular *H. fontanus* and *H. yardiyaensis*, because they occur in similar habitats.

The specimens from the northern Lake Eyre region (Neales Complex) also appear to be morphologically identical to the *H. fontanus* specimens from the southern Lake Eyre Springs (Bubbler, Strangways, and McKewin Springs), despite the apparent phylogenetic structuring (COI divergence estimates of 8–9%; Guzik *et al.*, 2019) evident in Fig. 1 as well as the geographic isolation between the spring complexes. There was also very little COI genetic divergence (2–3%; Guzik *et al.*, 2019) between the Bubbler, Strangways, and McKewin spring group populations.

Nomenclatural statement: A life science identifier (LSID) number was obtained for the new species: urn:lsid:zoobank.org:pub:71C7FCDD-FEB8-4F1D-A709-CE4E0F27A7DD.

Haloniscus microphthalmus Stringer, King & Taiti n. sp.

(Figs. 6, 7)

Material examined: Holotype, male SAMA C13233 (GAB0764.5; Genbank COI: KT236009, Genbank 18S: MK286391), Francis Swamp Springs, South Australia, 29°04′46.9"S 136°16′36.7"E, coll. M. Guzik and N. Murphy, 3 November 2007. Paratypes: 6 males SAMA C13234–C13239 (GAB0764.2–4 and GAB0764.6–8; Genbank COI: KT236009, Genbank 18S: MK286391), same collection data as holotype.

Diagnosis: Dorsum equipped with numerous broad based, apically frayed triangular scale setae. Eyes reduced to single spot of black pigment. Antennule with first, third articles longer than second; aesthetascs grooved longitudinally.

Description: Maximum body length 3.5 mm. Colour in alcohol pale with few traces of pigment. Body (Fig. 6A) moderately convex, elongated, about 3.5× as long as wide, with pleon distinctly narrower than pereon. Dorsum smooth with numerous frayed, triangular scale-setae (Fig. 6A, B); noduli laterales present, inserted at similar distance from lateral margins of pereonites; b/c, d/c co-ordinates as in Fig. 6C.

Cephalon (Fig. 6D–F) with small, rounded lateral lobes; supraantennal line present, slightly sinuous. Eyes reduced to spot of black pigment. Pereonites 1–4 with straight posterior margins, right-angled posterior corners; pereonites 5–7 with posterior corners progressively more acute. Pleonites 3–5 with posterior points reduced (Fig. 6A, G). Telson (Fig. 6A, G) almost twice as wide as long, distal part with straight sides, rounded apex.

Antennule (Fig. 6H) with first, third articles longer than second, 2 aesthetascs at apex, tuft of 4 aesthetascs subapically; all aesthetascs grooved longitudinally. Antenna (Fig. 6I) short, reaching to, but not past, pereonite 2; fifth article of peduncle slightly shorter than flagellum; flagellum with first, second articles subequal in length, third article longer; second, third articles with 1, 2 aesthetascs, respectively.

Mouth appendages as in *H. fontanus*.

Uropod (Fig. 6J) protopod, exopod grooved on outer margin, insertion of endopod proximal to that of exopod, exopod less than twice as long as endopod.

Pereopods 1 (Fig. 7A), 2 with carpus, merus bearing some long setae on sternal margin; pereopod 7 (Fig. 7B) ischium with sternal margin straight to slightly concave in proximal half.

Pleopod 1 (Fig. 7C) exopod apically rounded, with no setae along margin; endopod significantly longer than exopod, with pointed apical part bent outwards. Pleopod 2 (Fig. 7D) exopod triangular with 6 or 7 robust setae on outer margin; endopod distinctly longer than exopod, with flagelliform distal end. Pleopods 3, 4 exopods (Fig. 7E, F) triangular with rounded apex, 6, 7 robust setae on outer margin. Pleopod 5 (Fig. 7G) exopod triangular with 6 robust setae on outer margin. Pleopods 2–5 exopods with short setae on medial margin.

Etymology: The species name is composed of the Greek *micros* for 'small' and *ophthalmos* for 'eye' referring to the reduced eye visible as a spot of dark pigment.

Remarks: *Haloniscus microphthalmus* **n. sp.** is readily distinguished from the other SA GAB spring species described here by the reduced eyes, body setation, and shortened uropods. The new species is highly divergent genetically, showing approximately 20% COI divergence from the additional Lake Eyre and Dalhousie spring species and is, remarkably, more closely related to *Haloniscus* from subterranean calcrete aquifers in the Yilgarn, WA than to the other SA GAB species (Fig. 1A; Guzik *et al.*, 2019). *Haloniscus microphthalmus* is restricted to one spring group, Francis Swamp. No female specimens were collected at this location and, as such, potential female-only characters could not be recorded.

Nomenclatural statement: A life science identifier (LSID) number was obtained for the new species: urn:lsid:zoobank.org:pub:71C7FCDD-FEB8-4F1D-A709-CE4E0F27A7DD.

Haloniscus pedisetosus (Taiti & Humphreys, 2001) n. comb.

Andricophiloscia pedisetosa Taiti & Humphreys, 2001: 147, fig. 10; Cooper *et al.*, 2008: 197.

Material examined: Holotype, male WAM C25015 (BES 7207), Murchison Region, Lake Way, site 286, Western Australia, 26°41.256′S 120°17.868′E, coll. W.F. Humphreys and H.J. Hahn, 21 May 1999.

Remarks: *Haloniscus pedisetosus* **n. comb.** was only tentatively included in *Andricophiloscia* by Taiti & Humphreys (2001) as it possesses all the major characters of the generic diagnosis and is similar in morphology to *H. stepheni*, particularly in the occurrence of noduli laterales on the pereonites, uropods grooved on the lateral margin, and the exopod of male pleopods not fringed with fine setae. Taiti & Humphreys (2001) nevertheless also noted that it displayed a number of unusual characters, such as the petaliform aesthetascs of the antennule and a brush of setae on the carpus and merus of the male percopods. These characters, as described above, are now all known in species of *Haloniscus* and, therefore, *A. pedisetosa* is transferred here to the genus *Haloniscus*.

Haloniscus rotundus Stringer, King & Taiti n. sp.

(Figs. 8, 9)

Material examined: Holotype, male SAMA C13240 (GAB01459), Main Pool, Dalhousie Springs, South Australia, 26°25′16.3"S 135°30′11.6"E, coll. M. Guzik, R. King and L. Harsche, 6 July 2009. Paratypes: 6 males, 6 females SAMA C13241 (GAB01459), same collection data as holotype; 1 female C13242 (GAB01459.1; Genbank COI: KT236034, Genbank 18S: MK286388), same collection data as holotype; 5 males, 4 females SAMA C13243

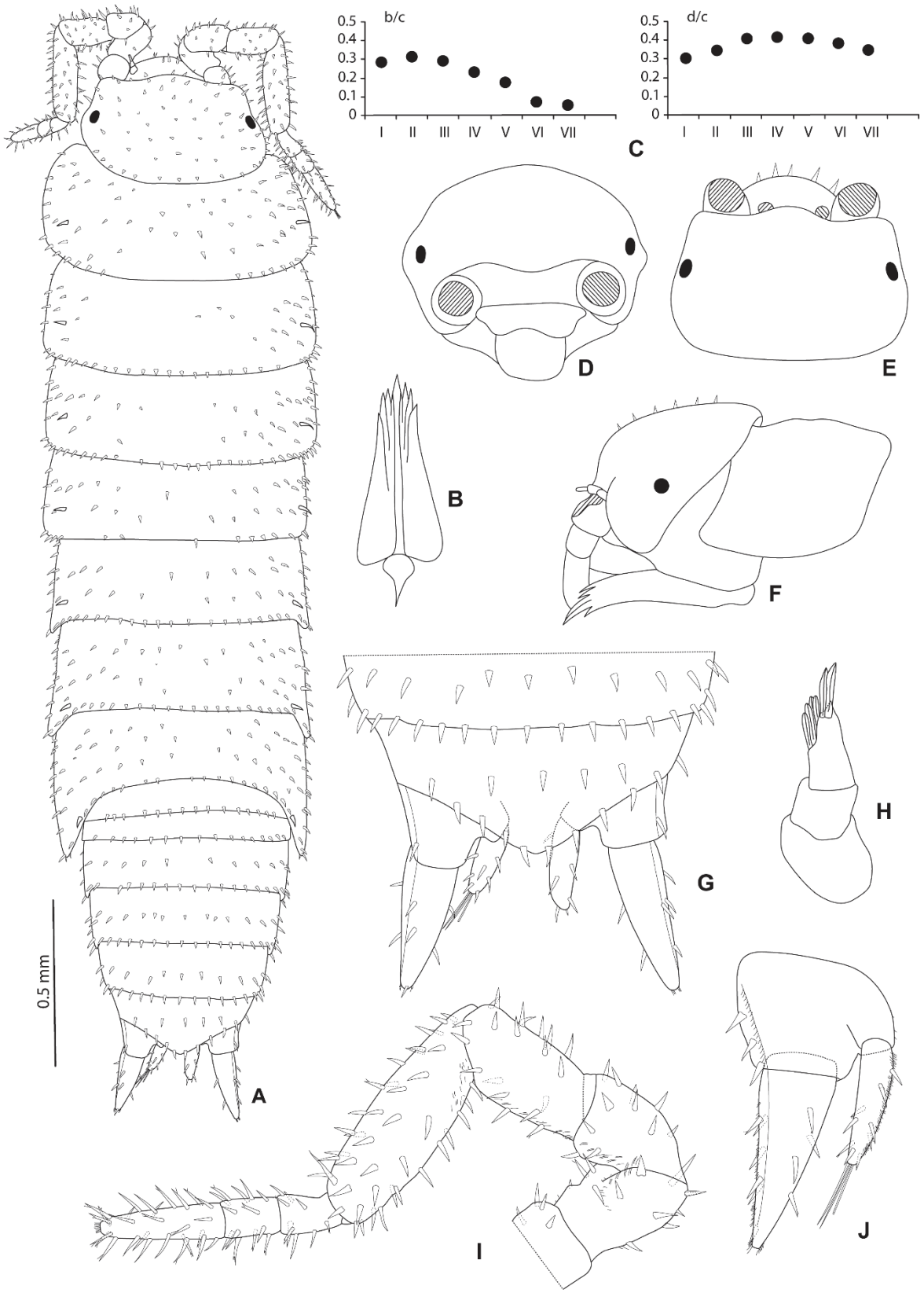


Figure 6. *Haloniscus microphthalmus* **n. sp.** holotype male (A, G–J), paratype male (B–F). Whole specimen, dorsal view (**A**); dorsal scale-seta (**B**); coordinates of the noduli laterales (**C**); cephalon, frontal view (**D**); cephalon, dorsal view (**E**); cephalon and pereonite 1, lateral view (**F**); telson and uropods (**G**); antennule (**H**); antenna (**I**); uropod (**J**).

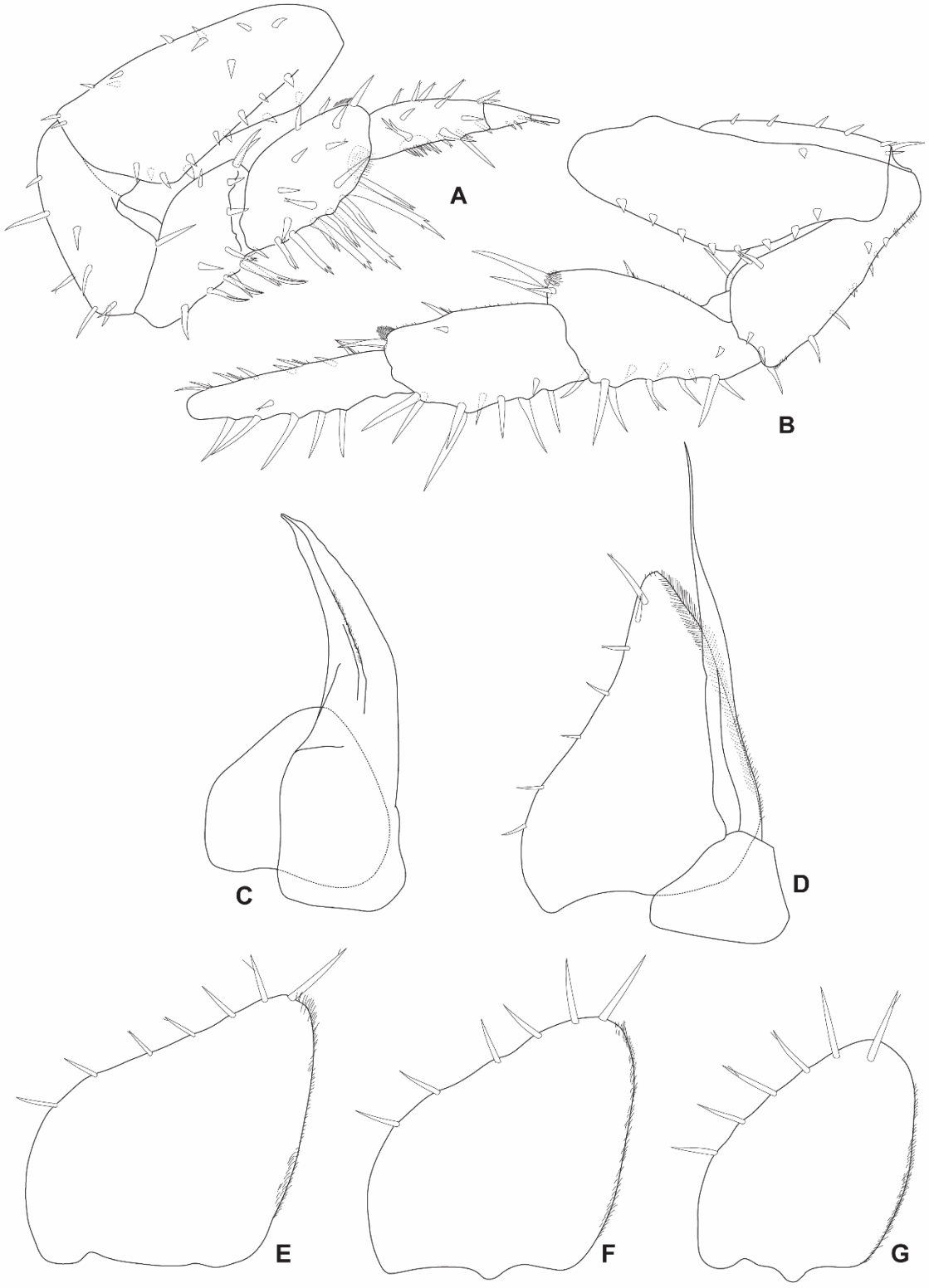


Figure 7. *Haloniscus microphthalmus* **n. sp.** holotype male. Pereopod 1 (**A**); pereopod 7 (**B**); pleopod 1 (**C**); pleopod 2 (**D**); pleopod 3 exopod (**E**); pleopod 4 exopod (**F**); pleopod 5 exopod (**G**).

(GAB01472), Meeting Place, Dalhousie Springs, South Australia, 26°29′18.3"S 135°29′15.6"E, coll. M. Guzik, R. King and L. Harsche, 9 July 2009.

Additional material: 1 male, 4 females, SAMA C13244 (GAB01433), Kingfisher, Dalhousie Springs, South Australia, 26°24′29.9"S 135°31′17.9"E, coll. M. Guzik, R. King and L. Harsche, 7 July 2009; 1 male SAMA C13245 (GAB01433.2; Genbank COI: KT236027, Genbank 18S: MK286388), Kingfisher, Dalhousie Springs, South Australia, 26°24′29.9"S 135°31′17.9"E, coll. M. Guzik, R. King and L. Harsche, 7 July 2009; 5 males, 5 females, SAMA C13246 (GAB01525),

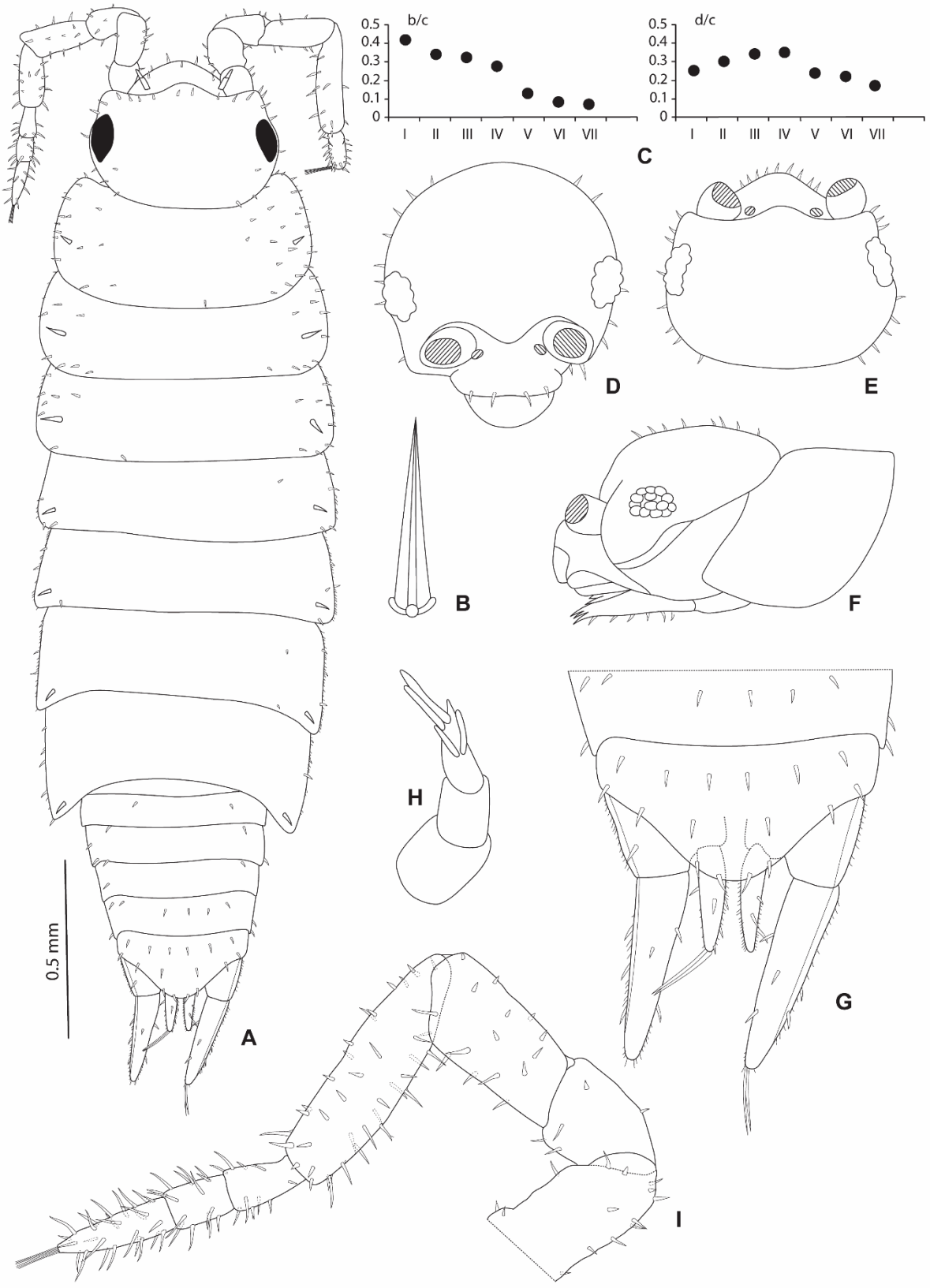


Figure 8. *Haloniscus rotundus* **n. sp.** holotype male (A, G–I), paratype female (B–F). Whole specimen, dorsal view (**A**); dorsal scale-seta (**B**); coordinates of the noduli laterales (**C**); cephalon, frontal view (**D**); cephalon, dorsal view (**E**); cephalon and pereonite 1, lateral view (**F**); telson and uropods (**G**); antennule (**H**); antenna (**I**).

Kingfisher, Dalhousie Springs, South Australia, 26°24′31.4"S 135°31′12.1"E, coll. M. Guzik, R. King and L. Harsche, 8 July 2009.

Diagnosis: Telson with broadly rounded apex. Male pleopod 1 exopod with outer margin slightly sinuous. Male pleopod 2 exopod triangular, distinctly concave towards the apex.

Figure 9. *Haloniscus rotundus* **n. sp.** holotype male (A, C), paratype male (B, D–H). Uropod (**A**); pereopod 1 (**B**); pereopod 7 (**C**); pleopod 1 (**D**); pleopod 2 (**E**); pleopod 3 exopod (**F**); pleopod 4 exopod (**G**); pleopod 5 exopod (**H**).

Description: Maximum body length: male, female 3 mm. Colour in alcohol light brown, with large pale muscle spots. Body (Fig. 8A) moderately convex, elongated, about 3× as long as wide, with pleon distinctly narrower than pereon. Dorsum smooth with some triangular scale-setae with narrow base (Fig. 8A, B); noduli laterales present, inserted at similar distance from lateral margins of pereonites; b/c, d/c co-ordinates as in Fig. 8C.

Cephalon (Fig. 8D–F) with small, rounded lateral lobes; supra-antennal line present, distinctly sinuous. Eyes with 13, 14 ommatidia. Pereonites 1–4 with straight posterior margins, right-angled posterior corners; pereonites 5–7 with posterior corners progressively more acute. Pleonites 3–5 with posterior points reduced (Fig. 8A, G). Telson (Fig. 8A, G) twice as wide as long with broadly rounded apex.

Antennule (Fig. 8H) with first, third articles longer than second, 2 long aesthetascs at apex, 2 in centre of third article. Antenna (Fig. 8I) short, reaching to, but not past, pereonite 2; fifth article of peduncle slightly shorter than flagellum; flagellum with first, second articles subequal in length, third longer.

Mouth appendages as in *H. fontanus*.

Uropod (Fig. 9A) protopod, exopod grooved on outer margin, insertion of endopod proximal to that of exopod, exopod less than twice as long as endopod.

Male: Pereopods 1 (Fig. 9B), 2 with merus bearing some long setae, a fringe of scales on sternal margin; carpus with several long setae on sternal margin, half frontal surface covered with fine setae or scales. Pereopod 7 (Fig. 9C) ischium with sternal margin straight.

Pleopod 1 (Fig. 9D) exopod with rounded apex, outer margin slightly sinuous, no setae along margin; endopod much longer than exopod, with short pointed apical part bent outwards. Pleopod 2 (Fig. 9E) exopod triangular, distinctly concave towards apex with 8 robust setae on outer margin; endopod distinctly longer than exopod, with flagelliform distal end. Pleopods 3, 4 (Fig. 9F, G) exopods triangular with outer margin bearing 6, 7 robust setae. Pleopod 5 (Fig. 9H) exopod triangular with 4 or 5 robust setae on outer margin. Pleopods 2–5 exopods with short setae on medial margin.

Etymology: Name composed of the Latin *rotundus* for "round" referring to the broadly rounded distal margin of the telson.

Remarks: *Haloniscus rotundus* **n. sp.** is primarily characterised by its broadly rounded telson. This new species is endemic to the Dalhousie Springs supergroup and is genetically distinct, showing COI divergences between 23–26%, from the more geographically distant Lake Eyre supergroup species: *H. fontanus* **n. sp.**, *H. microphthalmus* **n. sp.**, and *H. yardiyaensis* **n. sp.** (Guzik *et al.*, 2019).

Nomenclatural statement: A life science identifier (LSID) number was obtained for the new species: urn:lsid:zoobank.org:pub:71C7FCDD-FEB8-4F1D-A709-CE4E0F27A7DD.

Haloniscus stepheni Nicholls & Barnes, 1926

Haloniscus stepheni Nicholls & Barnes, 1926: 89, figs. 1–6, pl. 10, figs. 1–14; Vandel, 1973b: 101; Green, 1974: 245; Taiti *et al.*, 1995: 321.

Andricophiloscia stepheni - Taiti & Humphreys, 2001: 147; Cooper *et al.*, 2008: 197.

Material examined: Syntypes 10 males, 8 females WAM C25012, Kokatea Creek, Tenindewa, Western Australia, 8 January 1926.

Additional material: 3 males, 1 female WAM C74280 (BES 9888) (BES 9888.1; GenBank COI: EU364622), Martinjinni Nature Reserve Lake, Western Australia, 30.302°S 116.454°E, CALM Salinity Action Plan SPS155, 20 September 1999.

Remarks: *Haloniscus stepheni* was described by Nicholls & Barnes (1926) from specimens collected under logs by the banks of Kokatea Creek [= Kockatea Gully] near Tenindewa, WA, and later re-examined by Taiti & Humphreys (2001), with additional material from the type locality, and transferred to *Andricophiloscia*. Specimens subsequently collected from Martinjinni Nature Reserve (WA) (BES 9888) were found to be morphologically identical to *H. stepheni*, and were incorporated into the sequencing work of Guzik *et al.* (2019), revealing that this species does in fact belong to *Haloniscus* as originally described (Fig. 1A). The phylogeny suggests that *H. stepheni* is closely related to subterranean *Haloniscus* species from WA calcrete aquifers and, despite the similar morphology, is genetically distinct (COI divergence 18–23%; unpublished data) from the GAB species. With the exclusion of *H. stepheni* and *H. pedisetosus* from *Andricophiloscia*, this genus now comprises only the type species, *A. melanesiensis* Vandel, 1973, from Japen Island, New Guinea. The type material recorded by Vandel (1973a) should be re-examined to confirm the validity of the genus *Andricophiloscia*.

Haloniscus yardiyaensis Stringer, King & Taiti n. sp.

(Figs. 10, 11)

Material examined: Holotype, male SAMA C13247 (GAB01571), Freeling South Springs, Mount Dennison Complex, South Australia, 28°04′34.3"S 135°54′14.5"E, coll. M. Guzik, R. King and L. Harsche, 3 July 2009. Paratypes: 1 male, 3 females SAMA C13248 (GAB01571), same collection data as holotype; 1 female SAMA C13249 (GAB01571.1; GenBank COI: KT236029, GenBank 18S: MK286387), same collection data as holotype; 3 males, 2 females SAMA C13250 (GAB01614), Freeling South Springs, Mount Dennison Complex, South Australia, 28°04′17.7"S 135°54′14.3"E, coll. M. Guzik, R. King and L. Harsche, 3 July 2009; 2 males, 3 females SAMA C13251 (GAB01613), Freeling South Springs, Mount Dennison Complex, South Australia, 28°04′45.9"S 135°54′17.7"E, coll. M. Guzik, R. King and L. Harsche, 3 July 2009; 1 female SAMA C13252 (GAB01613.2; Genbank COI: KT236031, Genbank 18S: MK286387), Freeling South Springs, Mount Dennison Complex, South Australia, 28°04′45.9"S 135°54′17.7"E, coll. M. Guzik, R. King and L. Harsche, 3 July 2009.

Diagnosis: Antenna with fifth article of peduncle slightly swollen; flagellum with first, second articles subequal in length, third slightly longer. Male pereopod 1 carpus with several long setae on sternal margin, lines of short setae near sternal margin.

Description: Maximum body length: male 4.2 mm, female 5 mm. Colour in alcohol light brown with large pale muscle spots. Body (Fig. 10A) moderately convex, elongated, about 3.2× as long as wide, with pleon distinctly narrower than pereon. Dorsum smooth with triangular scale-setae (Fig. 10A, B); noduli laterales present on pereonites with one per side on each segment, inserted at similar distance from lateral margins of pereonites; b/c, d/c co-ordinates as in Fig. 10C.

Cephalon (Fig. 10D–F) with small, rounded lateral lobes; supra-antennal line present, distinctly sinuous. Eyes with 9, 10 ommatidia. Pereonites 1–4 with straight posterior margins, right-angled posterior corners; pereonites 5–7 with posterior corners progressively more acute. Pleonites 3–5 with posterior points reduced (Fig. 10A, G). Telson (Fig. 10A, G) almost twice as wide as long, distal part with straight sides, rounded apex.

Antennule (Fig. 10H) with third article slightly longer than first, second, 2 long aesthetascs subapically, tuft of 3 aesthetascs in central part of third article. Antenna (Fig. 10I) short, reaching to, but not past, pereonite 2; fifth article of peduncle slightly swollen, shorter than flagellum; flagellum with first, second articles subequal in length, third slightly longer; second, third articles with 1, 2 aesthetascs, respectively.

Mouth appendages as in *H. fontanus*.

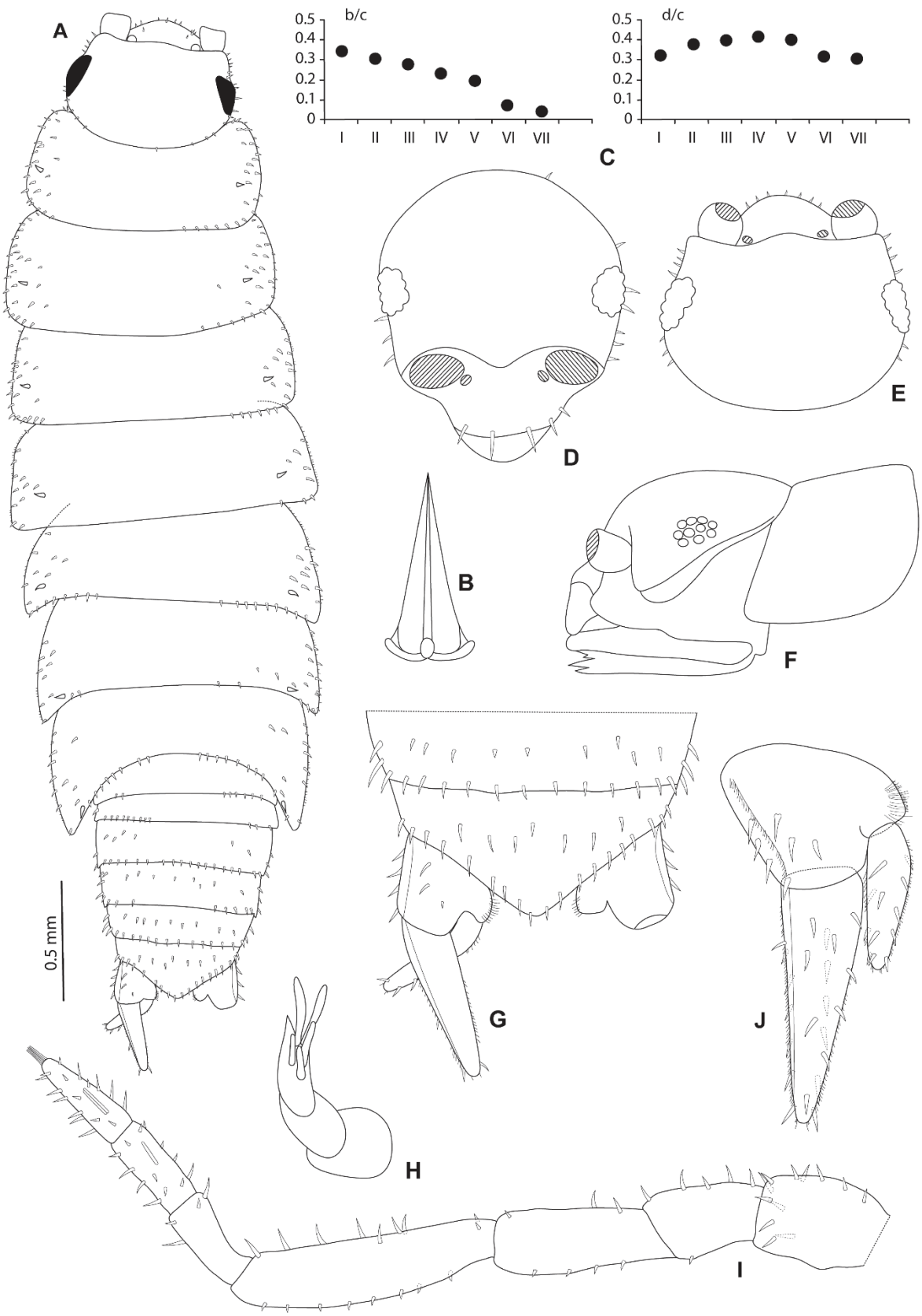


Figure 10. *Haloniscus yardiyaensis* **n. sp.** holotype male (A, G, H, J), paratype female (B–F), paratype male (I). Whole specimen, dorsal view (**A**); dorsal scale-seta (**B**); coordinates of the noduli laterales (**C**); cephalon, frontal view (**D**); cephalon, dorsal view (**E**); cephalon and pereonite 1, lateral view (**F**); telson and left uropod (**G**); antennule (**H**); antenna (**I**); uropod (**J**).

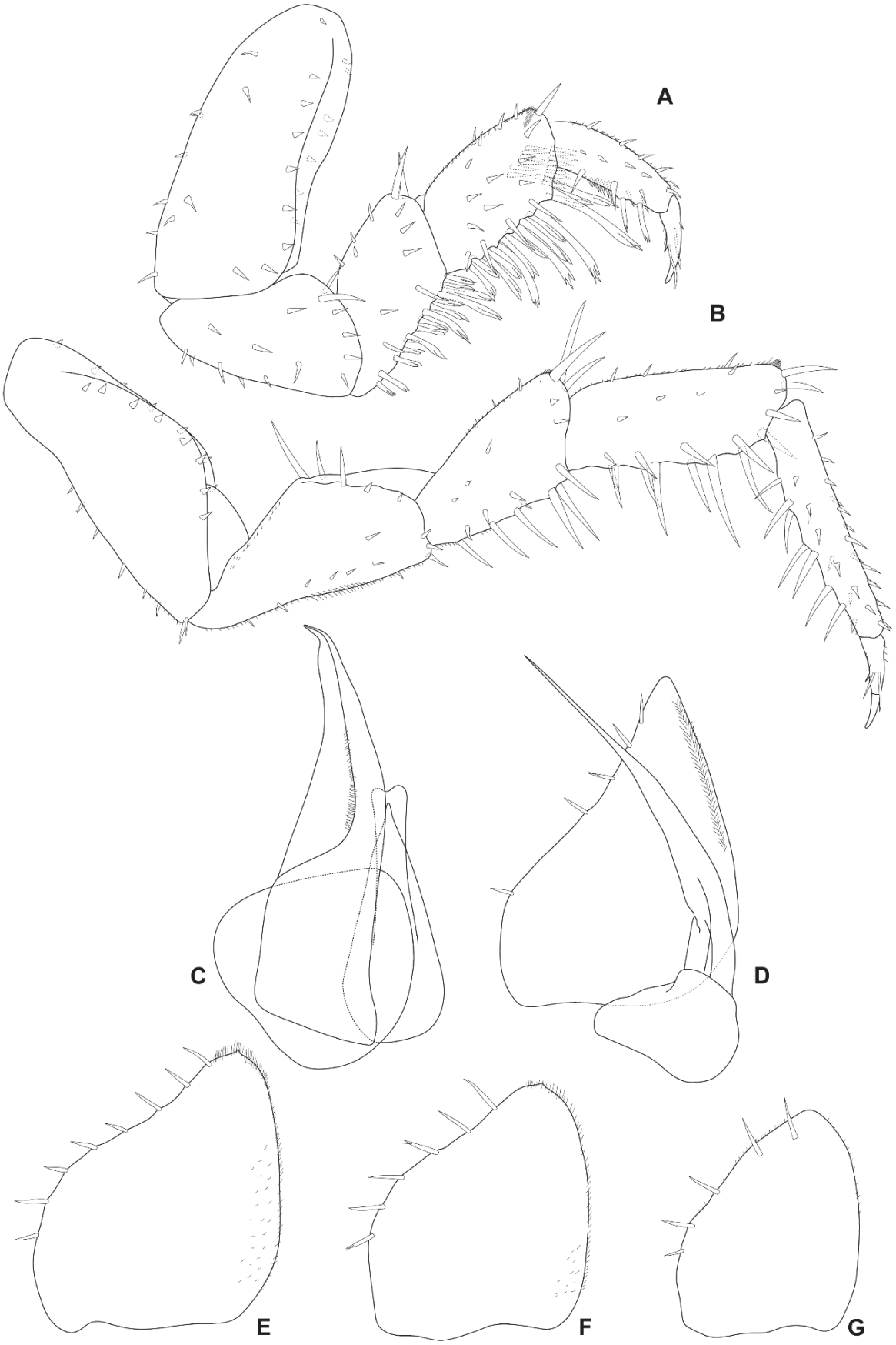


Figure 11. *Haloniscus yardiyaensis* **n. sp.** holotype male. Pereopod 1 (**A**); pereopod 7 (**B**); pleopod 1 (**C**); pleopod 2 (**D**); pleopod 3 exopod (**E**); pleopod 4 exopod (**F**); pleopod 5 exopod (**G**).

Uropod (Fig. 10J) with protopod, exopod grooved on outer margin, insertion of endopod proximal to that of exopod, exopod about twice as long as endopod.

Male: Pereopod 1 (Fig. 11A) with merus bearing some long setae on sternal margin; carpus with several long setae on sternal margin, lines of short setae. Pereopod 7 (Fig. 11B) ischium with sternal margin straight to slightly concave in proximal half.

Pleopod 1 (Fig. 11C) exopod apically rounded, with no setae along margin; endopod much longer than exopod, with pointed apical part bent outwards. Pleopod 2 (Fig. 11D) exopod triangular with outer margin slightly concave, bearing 5 robust setae. Pleopods 3, 4 (Fig. 11E, F) exopods triangular with outer margin comprising 6–8 robust setae. Pleopod 5 (Fig. 11G) exopod triangular with 5 robust setae on outer margin. Pleopods 2–5 exopods with short setae on medial margin.

Etymology: The species name refers to *Yardiya*, an Arabana language name for Freeling Springs.

Remarks: As discussed above, *H. yardiyaensis* **n. sp.** from Freeling South is morphologically similar to both *H. fontanus* **n. sp.** (southern Lake Eyre and Neales Complex), where it largely differs in antenna length, setation of the carpus and merus of male pereopod 1, and the length and number of robust setae on the exopods of male pleopods, and *H. stepheni* (WA), where it contrasts in the length of antennal flagellum articles. *Haloniscus yardiyaensis* and *H. stepheni* are nevertheless almost identical morphologically, but the phylogeny in Fig. 1A reveals that the two species are likely genetically distinct, with *H. yardiyaensis* more closely related to the GAB spring *Haloniscus* (excluding *H. microphthalmus* **n. sp.**) and *H. stepheni* grouping with a Yilgarn (WA) calcrete aquifer lineage. This morphology may, as mentioned previously, be convergent and an adaptation to their semi-terrestrial habitat.

Nomenclatural statement: A life science identifier (LSID) number was obtained for the new species: urn:lsid:zoobank.org:pub:71C7FCDD-FEB8-4F1D-A709-CE4E0F27A7DD.

DISCUSSION

The four newly described species are endemic to the SA GAB springs and thus not found elsewhere in Australia. The addition of these species, as well as *H. stepheni* and *H. pedisetosus* **n. comb.**, to the genus elevates the total number of described *Haloniscus* from five to 11. Prior to the present study, *Haloniscus* species had never been described from GAB springs and, therefore, the inclusion of four new species to a slowly increasing list of endemic taxa living in this unique ecosystem is significant. The descriptions presented here are based on traditional, rigorous morphological comparison in conjunction with the molecular species delimitation results of Guzik *et al.* (2019). The molecular analyses proposed between

Key to species of *Haloniscus* (after Taiti & Humphreys (2001))

1. Eyes present 2
— Eyes absent 7
2. Uropod with protopod flattened, outer margin of protopod and exopod not grooved; pereopods 1–4 subchelate; male pereopod 1 merus wider than long; male pereopods 2–7 ischium with a brush of long setae on sternal margin; all pleopods with exopods bearing a fringe of fine setae; endopod and exopod of male pleopod 1 subequal in length; endopod of male pleopod 2 about ¾ that of exopod ..... *Haloniscus searlei*
— Uropod with protopod not flattened, outer margin of protopod and exopod grooved (Fig. 4G, F); pereopods not subchelate; male pereopod 1 merus longer than wide (Fig. 5A); male pereopods 2–7 ischium with none or some long setae on sternal margin (Fig. 5B); exopod of male pleopod 1 without setae on outer margin (Fig. 5C), exopods of male pleopods 2–5 fringed only on medial margin (Fig. 5D–G); endopod of male pleopods 1 and 2 distinctly longer than exopod (Fig. 5C, D) 3
3. Dorsum smooth with numerous broad based, apically frayed triangular scale-setae (Fig. 6B); eyes reduced to single spot of pigment (Fig. 6A, D–F); antennule with aesthetascs grooved longitudinally (Fig. 6H) ..... *Haloniscus microphthalmus* **n. sp.**
— Dorsum smooth with triangular scale-setae, not broad based and apically frayed (Figs. 3B, 8B); eyes with at least 9 ommatidia; antennule with aesthetascs not grooved longitudinally (Fig. 3H) 4
4. Telson with broadly rounded apex (Fig. 8G) ..... *Haloniscus rotundus* **n. sp.**
— Telson with apex not broadly rounded (Figs. 3G, 10G) 5
5. Antennae reaching past pereonite 2 (Fig. 3A); male pleopod 2 exopod with around 10 robust setae on outer margin (Fig. 5D) ..... *Haloniscus fontanus* **n. sp.**
— Antennae not reaching past pereonite 2; male pleopod 2 exopod with fewer than 8 robust setae on outer margin (Fig. 11D) 6
6. Antenna with fifth article of peduncle slightly swollen; flagellum with first and second articles subequal in length, third slightly longer (Fig. 10I); endemic to Freeling South Springs (SA) ..... *Haloniscus yardiyaensis* **n. sp.**
— Antenna with fifth article of peduncle not swollen; flagellum with first and third articles subequal in length, second shorter; known only from WA ..... *Haloniscus stepheni*
7. Cephalon with no supra-antennal line; male pleopod 1 endopod with stout, spoon-like apical part directed outwards, equipped with fine setae and a terminal acute spine ..... *Haloniscus anophthalmus*
— Cephalon with distinct supra-antennal line; male pleopod 1 endopod without spoon-like apical part and terminal acute spine 8
8. Noduli laterales distinct, one per side on each pereonite; antennule with petaliform shaped aesthetascs; uropod with protopod and exopod deeply grooved on outer margin; male pleopod 1 exopod without a marginal fringe of fine setae ..... *Haloniscus pedisetosus* **n. comb.**
— Noduli laterales absent, antennule without petaliform shaped aesthetascs; uropod with protopod flattened and not grooved on outer margin; male pleopod 1 exopod with a marginal fringe of fine setae 9
9. Antennae very long, reaching posterior margin of pereonite 6; anterior male pereopods without brush of scales on merus sternal margin; male pleopod 1 endopod with stout apical part ..... *Haloniscus longiantennatus*
— Antennae shorter, reaching posterior margin of pereonite 2 or 3; anterior male pereopods with a brush of scales on merus sternal margin; male pleopod 1 endopod with styliform apical part 10
10. Antennule with third article much shorter than second; exopods of male pleopods cordiform ..... *Haloniscus tomentosus*
— Antennule with third article as long as second; exopods of male pleopods ovoid ..... *Haloniscus stilifer*

three (ABGD) and eight (bPTP) putative species from the GAB springs. The ABGD analysis (Fig. 1A) suggested that the Francis Swamp, Dalhousie, and the larger Lake Eyre clade represent three distinct species, whereas bPTP (Fig. 1A) proposed one Francis and two Dalhousie species and further split the Lake Eyre lineages into five species, with two Neales, one Hermit Hills, one Freeling, and one combined Bubbler, Strangways, and McKewin species. A potential problem with these multi-species coalescent methods is that they cannot differentiate phylogeographic structure, resulting from isolation of populations, from species exhibiting long-term isolation (Sukumaran & Knowles, 2017) and, consequently, these estimates should ideally be used in combination with morphological assessment. Our morphological analysis largely supported the estimate of between three and eight new SA spring species, but the final species number presented here is nevertheless more comparable to the conservative estimate of the ABGD analysis.

The use of molecular systematics to help delineate putative species for taxonomic description is a widely accepted and successful practice (King, 2009; King *et al.*, 2012; Rix *et al.*, 2018). We combined evidence from a multi-gene phylogeny, species delimitation analyses (Fig. 1; Guzik *et al.*, 2019) and morphological assessment to identify four new species from the SA GAB springs. Overall, each of the new species was associated with a distinct geographic range: *Haloniscus fontanus* corresponded to the southern Lake Eyre (Strangways, McKewin, and Bubbler spring groups) and Neales populations, *H. yardiyaensis* is a seemingly closely related species from the proximate Freeling South Springs, *H. rotundus* occurs in the Dalhousie Springs supergroup, and *H. microphthalmus* is endemic to the Francis Swamp southern Lake Eyre region (see Fig. 1).

Specimens from the geographically distant northern Lake Eyre Neales region are morphologically identical to the specimens from Bubbler, McKewin, and Strangways Springs, which are within the southern region of the Lake Eyre Basin. This is a surprising result considering the evident isolation of the springs and some phylogeographic structuring of the Neales populations (Fig. 1), which may imply a cryptic species (Murphy *et al.*, 2015b). The COI-only phylogeny (Fig. 1A) furthermore suggested (although with low posterior support) that the Neales population forms a monophyletic grouping with the Freeling population rather than the morphologically identical southern Lake Eyre Springs populations. These relationships were nevertheless not reinforced by adding the more conserved 18S locus (Fig. 1B). The relationships amongst each of the reciprocally monophyletic southern Lake Eyre, Freeling, and Neales spring populations remain unresolved, and this may be due to relatively recent divergences between populations. Future next-generation sequencing work could help increase understanding of these complex species relationships. We have chosen a somewhat conservative approach and include the Neales population (with the Strangways, McKewin and Bubbler group) in *H. fontanus*. Our decision is based on this phylogenetic evidence along with COI-divergence estimates among the populations calculated below the 16% molecular threshold for species delimitation as proposed by Lefébure *et al.* (2006) as well as the apparent lack of morphological differences.

The molecular study by Guzik *et al.* (2019) reveals that the current taxonomic position of *Haloniscus* greatly underestimates the true species diversity within Australia, particularly across important groundwater-dependent ecosystems within the arid zone, with additional descriptions of species from subterranean aquifers in WA and the Northern Territory to be published at a later date. The description of four new species here has important implications for GAB springs conservation management and the protection of species in this unique and threatened system. Taylor *et al.* (2018) suggest *Phreatomerus latipes* (Chilton, 1922), an isopod species complex endemic to the SA GAB springs with similarly restricted distributions (Guzik *et al.*, 2012), as a flagship species of high scientific and social value, which should be listed under the EPBC Act. These newly described *Haloniscus* species are similarly diverse, short-range endemics, and potential relicts, and, as such, conservation management to maximise their genetic diversity is vital to avoid the extinction of these species.

ACKNOWLEDGEMENTS

The authors thank Nick Murphy (LaTrobe University) and Travis Gotch (Department for Environment and Water) for their invaluable advice at the GAB springs; Steve Delean (The University of Adelaide) and Lewis Harsche for field assistance; Bill Humphreys (Western Australian Museum) for his support with aspects of the manuscript; and Steven Stringer for his help with figure design. Field work at the springs was completed under South Australian Permit No: Z25519 to Dr Nick Murphy. We appreciate the access provided to us by the traditional owners of the GAB spring country, particularly Reg Dodd and Dean Ah Chee, and would like to thank Greg Campbell (Chief Executive Officer of S. Kidman & Co Ltd.) as well as station managers, Randall Crozier, Peter Paisley, and Bobby Hunter, for permission to sample on private property. This work was funded by the Australian Biological Resources Study (ABRS) National Taxonomy Research Grant Program (NTRGP) (Capacity-Building grant: CT214-11) to DNS, Australian Research Council Linkage Grants (LP0669062 and LP140100555) to ADA and SJBC with industry partners the South Australian Museum, the South Australian Department for Environment and Water (DEW), BHP Billiton and the Nature Foundation, SA (LP0669062), and the South Australian Museum, Western Australian Museum, WA Department of Biodiversity, Conservation and Attractions, Bennelongia Pty Ltd. and Biota Environmental Sciences Pty Ltd. (LP140100555), and the Nature Conservancy, generously supported by The Thomas Foundation to DNS. DNS also wishes to acknowledge the support of an Australian Government Research Training Program Scholarship through The University of Adelaide. Lastly, we thank two anonymous reviewers for providing constructive feedback on an earlier version of the manuscript.

REFERENCES

Chilton, C. 1920. On a new isopodan genus (family Oniscidae) from Lake Corangamite, Victoria. *Proceedings of the Linnean Society of New South Wales*, **44**: 723–734.

Chilton, C. 1922. A new isopod from Central Australia belonging to the Phreatoicidae. *Transactions of the Royal Society of South Australia*, **46**: 23–33.

Cooper, S.J.B., Saint, K.M., Taiti, S., Austin, A.D. & Humphreys, W.F. 2008. Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic isopods (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics*, **22**: 195–203.

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation. In: *Endless forms: species and speciation*. (D.J. Howard & S.H. Berlocher, eds.), pp. 57–75. Oxford University Press, New York.

de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*, **56**: 879–886.

Fensham, R.J. & Price, R.J. 2004. Ranking spring wetlands in the Great Artesian Basin of Australia using endemicity and isolation of plant species. *Biological Conservation*, **119**: 41–50.

Green, A.J.A. 1974. IX. Oniscidea (terrestrial Isopoda). In: *Biogeography and ecology in Tasmania* (W.D. Williams, ed.), pp. 229–249. Junk, The Hague, The Netherlands.

Guzik, M.T., Adams, M.A., Murphy, N.P., Cooper, S.J.B. & Austin, A.D. 2012. Desert springs: deep phylogeographic structure in an ancient endemic crustacean (*Phreatomerus latipes*). *PLoS One*, **7** [doi: org/10.1371/journal.ponc.0037642].

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular

phylogenetic analysis of Australian arid zone oniscidean isopods (Crustacea) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33**: 556–574.

Habermehl, M.A. 1980. The Great Artesian Basin, Australia. *BMR Journal of Australian Geology and Geophysics*, **5**: 9–38.

Harris, C.R. 1992. Mound Springs: South Australian conservation initiatives. *The Rangeland Journal*, **14**: 157–173.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**: 111–120.

King, R.A. 2009. Two genera and species of chiltoniid amphipods (Crustacea: Amphipoda: Talitroidea) from freshwater mound springs in South Australia. *Zootaxa*, **2293**: 35–52.

King, R.A., Bradford, T., Austin, A.D., Humphreys, W.F. & Cooper, S.J.B. 2012. Divergent molecular lineages and not-so-cryptic species: the first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. *Journal of Crustacean Biology*, **32**: 465–488.

Lefébure, T., Douady, C.J., Gouy, M. & Gibert, J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*, **40**: 435–447.

Murphy, N.P., Adams, M. & Austin, A.D. 2009. Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. *Molecular Ecology*, **18**: 109–122.

Murphy, N.P., Adams, M., Guzik, M.T. & Austin, A.D. 2013. Extraordinary micro-endemism in Australian desert spring amphipods. *Molecular Phylogenetics and Evolution*, **66**: 645–653.

Murphy, N.P., Breed, M.F., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2012. Trapped in desert springs: phylogeography of Australian desert spring snails. *Journal of Biogeography*, **39**: 1573–1582.

Murphy, N.P., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2015a. Desert spring refugia: museums of diversity or evolutionary cradles? *Zoologica Scripta*, **44**: 693–701.

Murphy, N.P., King, R.A. & Delean, S. 2015b. Species, ESUs or populations? Delimiting and describing morphologically cryptic diversity in Australian desert spring amphipods. *Invertebrate Systematics*, **29**: 457–267.

Nicholls, G.E. & Barnes, H.E. 1926. Description of a new species of terrestrial isopod, *Haloniscus stepheni*, from Western Australia. *Journal of the Royal Society of Western Australia*, **12**: 87–96.

Ponder, W.F. 2002. Desert springs of the Great Artesian Basin. In: *Proceedings of the Meeting on Spring-fed Wetlands: Important Scientific and Cultural Resources of the Intermountain Region, May 2002, Las Vegas, NV* (D.W. Sada & S.E. Sharpe, eds), pp. 1–13. Desert Research Institute, Reno, NV, USA.

Ponder, W.F. 2003. Endemic aquatic macroinvertebrates of artesian springs of the Great Artesian Basin – progress and future directions. *Records of the South Australian Museum Monograph Series*, **7**: 101–110.

Ponder, W.F., Eggler, P. & Colgan, D.J. 1995. Genetic differentiation of aquatic snails (Gastropoda: Hydrobiidae) in artesian springs in arid Australia. *Biological Journal of the Linnean Society*, **56**: 553–596.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**: 1864–1877.

Rix, M.G., Huey, J.A., Cooper, S.J.B., Austin, A.D. & Harvey, M.S. 2018. Conservation systematics of the shield-backed trapdoor spiders of the *nigrum*-group (Mygalomorphae, Idiopidae, *Idiosoma*): integrative taxonomy reveals a diverse and threatened fauna from south-western Australia. *Zookeys*, **756**: 1–121.

Sukumaran, J. & Knowles, L.L. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America*, **114**: 1607–1612.

Taiti, S. & Humphreys, W.F. 2001. New aquatic Oniscidea (Crustacea: Isopoda) from groundwater calcretes of Western Australia. *Records of the Western Australian Museum* Supplement No. **64**: 133–151.

Taiti, S., Ferrara, F. & Iliffe, T.M. 1995. A new species of *Haloniscus* Chilton, 1920 from New Caledonia (Isopoda: Oniscidea). *Crustaceana*, **68**: 321–328.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**: 2725–2729.

Taylor, G.S., Braby, M.F., Moir, M.L., Harvey, M.A., Sands, D.P.A., New, T.R., Kitching, R.L., McQuillan, P.B., Hogendoorn, K., Glatz, R.V., Andren, M., Cook, J.M., Henry, S.C., Valenzuela, I. & Weinstein, P. 2018. Strategic national approach for improving the conservation management of insects and allied invertebrates in Australia. *Austral Entomology*, **57**: 124–149.

Vandel, A. 1973a. Les Isopodes terrestres (Oniscoidea) de la Mélanésie. *Zoologisches Verhandelingen*, **125**: 1–160.

Vandel, A. 1973b. Les Isopodes terrestres de l'Australie. Étude systématique et biogéographique. *Mémoires du Muséum national d'Histoire naturelle*, Série A, Zoologie, **82**: 1–171.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**: 2869–2876.

CHAPTER 5:

General discussion

The overarching aim of this project was to increase knowledge of the systematics and biogeographic history of *Haloniscus* isopods, and, in particular, investigate how major changes in Australia's climate since the late Miocene have impacted the diversification and relictualisation of this group within three groundwater-dependent ecosystems in the Australian arid zone. This study has provided a wealth of new information on *Haloniscus* biodiversity and the evolution of these groundwater invertebrates, with four new species described, a revised key to the genus, an extensive and informative exon capture phylogenomic dataset for phylogenetic examination, and detailed insights into the timing and impacts of major aridification on the biogeographic and evolutionary history of species. This project is the first to use transcriptome-based exon capture methods for *Haloniscus* (or, in fact, any additional Australian groundwater invertebrate) phylogenetics research, with the phylogenomic resources generated here representing a significant step forward from earlier molecular datasets used to investigate *Haloniscus* evolution and systematics (Cooper *et al*., 2008; Murphy *et al*., 2015; Guzik *et al*., 2019). Our molecular dataset has allowed for a thorough exploration of complex evolutionary hypotheses with improved resolution, and highlights the overall significance of these groundwater isopods as climate relict taxa. In this Discussion, a synthesis of the thesis is presented, along with the contributions and limitations of the project and suggested future directions for *Haloniscus* and Australian groundwater ecosystem research.

Thesis synthesis and contributions

Exon capture phylogenomics

In Chapter 2, a substantial and highly informative exon capture phylogenomic dataset was generated for *Haloniscus* (and other isopod) phylogenetic analyses in order to investigate longstanding questions surrounding their evolution. Previous phylogenetic and phylogeographic studies regarding *Haloniscus* have relied on the limited selection of readily available molecular markers for crustaceans, resulting in poor resolution for internal branches of output phylogenies, and a lack of clarity concerning the origins and biogeography of the fauna (Cooper *et al*., 2008; Guzik *et al*., 2019). Exon capture approaches, with high-throughput next-generation sequencing, are now providing an alternate method, permitting the rapid production of extensive multi-locus molecular datasets for phylogenetic analysis (Bi *et al*., 2012; Li *et al*., 2013; Hugall *et al*., 2016; Teasdale *et al*. 2016; Klopfstein *et al*., 2019; Reilly *et al*., 2019). These protocols, nonetheless, are constantly developing and further baseline information and empirical data, as well as detailed and reproducible methodological guidelines, are critical for the effective design of future capture experiments. In this study, we explored the performance of transcriptome-based exon capture without the use of a reference genome, together with the efficacy of our custom bait design targeting 469 orthologous genes, for *Haloniscus* and more distantly related outgroup isopod species.

The findings presented in this chapter help to build on the results from previous studies, which aim to increase the efficiency and ultimate success of exon capture experiments (Bi *et al*., 2012; Bragg *et al*., 2016; Portik *et al*., 2016; Abdelkrim *et al*., 2018).

The transcriptome-based exon capture guideline and custom *Haloniscus* bait design developed in this study were successful in recovering an extensive dataset comprising short (<100 bp) and longer (>300 bp) exons with high uniformity in sequencing depth, as well as a substantial amount of flanking intron sequence data. The combination of coding and non-coding sequence data obtained using this capture technique offers significant flexibility for data analysis, with potential for application in phylogenetics, population genetics, and species delimitation analyses (Portik *et al*., 2016). Possible factors influencing the success of transcriptome-based exon capture, including specimen preservation age, capture pool size and final pooling size before sequencing, were examined, with only the final pooling size negatively impacting the quantity of sequence data obtained. The successful enrichment of *Haloniscus* specimens, which were preserved for >15 years and likely comprise degraded DNA, has significant implications and is of particular importance in phylogenetics (Bi *et al*., 2013; Abdelkrim *et al*., 2018; Wood *et al*., 2018), given the difficultly of re-sampling rare (or extinct) lineages, as well as the associated expenses (Bradley *et al*., 2014; Wen *et al*., 2015). Historical museum specimens, which have been collected over a long period of time, are essential for systematics and taxonomic research, signifying an invaluable resource, and targeted capture methods provide an opportunity to effectively acquire genomic data from these important collections.

In addition, new genetic resources, including six isopod transcriptomes, an isopod-specific orthologue set, custom bait design, and post-processing bioinformatics scripts have been provided here for use in future studies. The methodological pipeline is entirely reproducible, and the assembled transcriptomes and orthologue set (531 single-copy genes) could be employed in future isopod (not merely *Haloniscus*) phylogenetic, phylogeographic, and population genetic studies. Furthermore, the custom bait design, based on transcriptomes from six divergent isopod species, was able to successfully capture sequence data from distantly related outgroups (Paraplatyarthridae and Armadillidae), highlighting the general utility of this design for phylogenetic inference. The creation and availability of baits (and orthologues) capturing a broad taxonomic spectrum is of interest in numerous disciplines, including phylogenetics and biodiversity monitoring (Mayer *et al*., 2016; Teasdale *et al*., 2016). The established bait set can be used to rapidly obtain orthologues and capture data from more divergent isopod taxa (Lemmon et al., 2012), and could, furthermore, assist in continued research and biodiversity assessments of the short-range endemic *Haloniscus* fauna from threatened groundwater-dependent ecosystems.

Evolution and biogeographic history

In chapter 3, the exon capture phylogenomic dataset produced in chapter 2, along with phylogenetic, divergence dating and ancestral area reconstruction analyses, was used to explore the evolution and biogeographic history of *Haloniscus* fauna from three arid zone groundwater-dependent ecosystems, the Yilgarn calcrete aquifers (Western Australia (WA)), Ngalia Basin aquifers (Northern Territory (NT)) and Great Artesian Basin (GAB) springs (South Australia (SA)), highlighting the impact of major climatic change on diversification. Examining the evolutionary and biogeographic history of relict taxa, such as *Haloniscus*, that are poor dispersers and have survived significant environmental changes, can provide important insights into the past history of continents, and help to unravel the origins of these unique taxa (Habel *et al*., 2010; Bauzà-Ribot *et al*., 2012; Rix *et al*., 2017). In this chapter, phylogenetic analyses revealed fully resolved historical connections among *Haloniscus* from the three disparate groundwater habitats, suggesting a past shared evolutionary history, and offering evidence for ancestral populations once occupying a broader range. These connections dated to the late Miocene phase, corresponding to the onset of aridification across the Australian continent where there was a significant termination in the warmer and wetter conditions of the earlier Miocene interval (Bowler, 1976; Martin, 2006; Byrne *et al*., 2008). These findings, furthermore, represent a repeated evolutionary pattern since Murphy *et al*. (2009) similarly indicated that divergences between clades of GAB spring and WA Yilgarn chiltoniid amphipods coincided with the late Miocene. The current study, therefore, provides further evidence (more robust than in previous studies) that once widespread ancestors to present day *Haloniscus* (and other groundwater-dependent taxa) became restricted to these distinct habitats as a result of Miocene aridification, and that current taxa represent climate relict species (Cooper *et al*., 2007, 2008; Guzik *et al*., 2012; Murphy *et al*., 2012, 2015).

Further diversification within the discrete groundwater areas, nonetheless, appears to have occurred during the Pliocene. The early Pliocene is thought to have comprised a temporary return to the warm and wet conditions of the early Miocene, followed by a period of intensive aridity, with the formation of sandy and stony deserts (Sniderman *et al*., 2007, 2016; Byrne *et al*., 2008). The results in chapter 3 suggest that isolation and subsequent diversification of *Haloniscus* taxa within distinct aquifers in the Yilgarn (WA) and Ngalia Basin (NT) overlap with this second phase of aridity, indicating that ancestral *Haloniscus* were able to move around the surface following late Miocene aridification, with the early Pliocene wetter phase likely facilitating the dispersal and eventual colonisation of aquifers. Therefore, these findings lend important molecular support to hypotheses surrounding the early Pliocene return to wet conditions (Sniderman *et al*., 2016), which is currently based on data from a limited number of geographic locations, and not often considered in studies concerning the climate-induced evolution of

Australian fauna (Leys *et al*., 2003; Cooper *et al*., 2007, 2008). This study has, therefore, contributed further insights (to be viewed in combination with geological and palaeontological data, and additional molecular studies) into the assembly of the Australian arid zone, the evolutionary history of aquatic arid zone fauna, and the significance of groundwater ecosystems as refugia (Hewitt, 2000; Keppel *et al*., 2012; Davis *et al*., 2013) that have facilitated the survival and diversification of these isopods.

Species descriptions

In Chapter 4, four new species of *Haloniscus* isopod were described from the Lake Eyre and Dalhousie springs of the GAB in SA based on a combination of molecular and morphological analyses. The genus was, furthermore, transferred from the family Scyphacidae to Philosciidae, and two species previously belonging to *Andricophiloscia* Vandel, 1973 were transferred to *Haloniscus*. These new descriptions represent the first formally documented *Haloniscus* species from the GAB, and elevate (together with the transferred species) the total number of described *Haloniscus* from five to 11. The morphological taxonomic assessment of these species, which generally consist of few distinguishing characters, was informed by the phylogenetic and species delimitation analyses of Guzik *et al*. (2019; a study that I was intimately involved in – see attached Appendix), highlighting the importance of molecular methods to assist in the rapid evaluation of new species for description (King *et al*., 2009, 2012). Guzik *et al*. (2019) revealed considerable levels of diversity and short-range endemism, recognising a minimum of 26 new putative species from groundwater-dependent ecosystems in Australia, and the descriptions presented in this study, together with the key to species, represent essential progress towards formally identifying and naming all known *Haloniscus* species. An understanding of biodiversity and formal recognition of species is imperative in a rapidly changing world. Increasing the rate at which species are named is now deemed an essential priority in Australia, especially for conservation planning and legislation regarding threatened species (Taxonomy Decadal Plan Working Group, 2018), such as the isopods described here which are associated with habitats considered under threat (Environmental Protection and Biodiversity Conservation Act, 1999).

Limitations and future directions

A critical challenge in studying subterranean invertebrate fauna from calcrete aquifers is the capacity to effectively sample aquifers across the entire region. *Haloniscus* specimens have been collected and sequenced from 18 calcrete aquifers in the Yilgarn region, with *Haloniscus* taxa from only 12 calcretes included in the transcriptome-based exon capture analyses implemented in this study (due to budget constraints and specimen availability). These samples, however, likely represent only a fraction of the actual Yilgarn *Haloniscus* biodiversity, with new species likely yet to be discovered (Guzik *et al*., 2011).

Sampling of aquifers is currently limited to pre-existing access points, including boreholes installed for water extraction, groundwater monitoring and mineral exploration, along with pastoral wells, which were not established for the purpose of faunal surveys. Consequently unsampled lineages (both from the region and those not included in this study) as well as extinct taxa, and the resultant missing nodes in our output phylogenies may have impacted the results and inferences regarding the precise timings of *Haloniscus* calcrete colonisation. Further sampling into the future (with additional sequencing) may help to relieve this issue, but in order to effectively sample the region for a more complete assessment, a comprehensive sampling regime with additional calcrete access points is required to obtain a more thorough understanding of *Haloniscus* (and other calcrete-inhabiting species) diversity.

Enhancing sampling intensity is fundamental for accurate biodiversity estimates and the taxonomy of new species, particularly for conservation management as described above. In order to appropriately manage these groundwater-dependent *Haloniscus* species, it is essential to have an understanding of their diversity and distribution, especially since many taxa are short-range (or even ultra-short-range) endemics, and entirely restricted to habitats under increasing threat from pastoralism and the mining industry (Nevill *et al*., 2010; Guzik *et al*., 2019). In Chapter 4, four species of *Haloniscus* were formally described from the GAB springs in SA, and an identification key to all currently described species was presented, which represents an important contribution to the systematics of the genus. Nevertheless, only four species are presently described from the Yilgarn region in WA (from only two calcretes) and no species are described from the Ngalia Basin aquifers (NT) despite the molecular results of Guzik *et al*. (2019) revealing significant diversity. Due to time limitations, morphological taxonomic analyses of these species could not be undertaken during this project, but future research should aim to prioritise additional descriptions. Descriptive taxonomic work to formally describe species has generally lagged behind molecular-based species discovery, and this is largely due to a lack of specialised taxonomists (Tomlinson & Boulton, 2010; King *et al*., 2012). An important outcome from this study was my training and professional development in integrated taxonomic skills and practices (both alpha taxonomy and emerging molecular systematics methods). The training of students is critical to the continued formal documentation of Australia's largely undescribed invertebrate biodiversity.

The groundwater-dependent ecosystems discussed in this study are major refugial habitats for short-range endemic, relictual species that are bioindicators of ecosystem health and function (Humphreys, 2006). Identifying, understanding and managing these environments is now considered an important conservation priority given anticipated climate change (Moritz & Agudo, 2013) as these systems both preserve ancient lineages through fluctuating conditions, and also facilitate the generation of species diversity (Davis *et al*., 2013; Murphy *et al*., 2015). The findings presented in this study (see Chapter 3) have further emphasised the value of these groundwater refugia, and their endemic relictual species,

for examining theories regarding the biogeographic history of the Australian continent. Conservation measures and future planning, which incorporate a high level of protection for these refugial systems, are, thus, critical and an important future consideration. Additional research into further groundwater habitats in Australia (such as calcretes in the Amadeus Basin (NT)) should be undertaken to establish whether they too represent refugia, comprise unknown *Haloniscus* species and other taxa, and have the potential to provide increased understanding of groundwater-dependent faunal evolution and the biogeographic history of Australia.

Our understanding of the evolution of subterranean and other groundwater-associated fauna as well as the biogeographic history of the Australian continent may, furthermore, be better informed based on comparisons of the distribution and evolution of additional groups, such as dytiscid diving beetles, chiltoniid amphipods and copepods, inhabiting these regions. Similar target capture genomic datasets can be generated for taxa from these aforementioned faunal groups to facilitate stronger comparisons based on high resolution phylogenies, and the detailed methodological template and newly developed post-processing scripts provided in Chapter 2 may assist such future studies. Nevertheless, the exon capture approach and bait design utilised here may be improved by including additional transcriptome data from a broader range of isopod exemplars, and by potentially using an assembly approach (which should be tested in the future) for data processing, particularly for the outgroup paraplatyarthrid and armadillid outgroup taxa as reads were mapped to divergent *Haloniscus* orthologues. These extensive and informative datasets may, furthermore, be used to examine complex species, generic and family-level relationships (Hugall *et al*., 2016; Wood *et al*., 2018; O'Hara *et al*., 2019), and this is an area worth exploring for better understanding oniscidean isopod systematics, especially since current phylogenies remain speculative (Schmidt, 2008; Sfenthourakis & Taiti, 2015). Overall, the findings presented in this project signify an important contribution to knowledge surrounding Australian groundwater systems and their endemic climate relict *Haloniscus* fauna, with new species descriptions and further insights into the evolution of species. Nonetheless, these findings are just the beginning, with new molecular methods and technologies continuing to enhance our understanding of these important ecosystems.

References

Abdelkrim, J., Aznar-Cormano, L., Fedosov, A.E., Kantor, Y.I., Lozouet, P., Phuong, M.A., Zaharias, P. & Puillandre, N. 2018. Exon-capture-based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). *Molecular Biology and Evolution*, **35**: 2355–2374.

Bauza-Ribot, M.M., Juan, C., Nardi, F., Oromi, P., Pons, J. & Jaume, D. 2012. Mitogenomic phylogenetic analysis supports continental-scale vicariance in subterranean thalassoid crustaceans. *Current Biology*, **22**: 2069–2074.

Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R. & Moritz, C. 2013. Unlocking the vault: next-generation museum population genomics. *Molecular Ecology*, **22**: 6018–6032.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C. & Good, J. M. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, **13:** 403.

Bowler, J.M. 1976. Aridity in Australia: age, origins and expression in aeolian landforms and sediments. *Earth-Science Reviews*, **12**: 279–310.

Bradley, R.D., Bradley, L.C., Garner, H.J. & Baker, R.J. 2014. Assessing the value of natural history collections and addressing issues regarding long-term growth and care. *BioScience*, **64**: 1150–1158.

Bragg, J.G., Potter, S., Bi, K. & Moritz, C. 2016. Exon capture phylogenomics: efficacy across scales of divergence. *Molecular Ecology Resources*, **16**: 1059–1068.

Byrne, M., Yeates, D.K., Joseph, L., Kearney, M., Bowler, J., Williams, M.A.J., Cooper, S., Donnellan, S.C., Keogh, J.S., Leys, R., Melville, J., Murphy, D.J., Porch, N. & Wyrwoll, K-H. 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology*, **17**: 4398–4417.

Cooper, S.J.B., Bradbury, J.H., Saint, K.M., Leys, R., Austin, A.D. & Humphreys, W.F. 2007. Subterranean archipelago in the Australian arid zone: mitochondrial DNA phylogeography of amphipods from central Western Australia. *Molecular Ecology*, **16**: 1533–1544.

Cooper, S.J.B., Saint, K.M., Taiti, S., Austin, A.D. & Humphreys, W.F. 2008. Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics*, **22**: 195–203.

Davis, J., Pavlova, A., Thompson, R. & Sunnucks, P. 2013. Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Global Change Biology*, **19**: 1970–1984.

Guzik, M.T., Adams, M.A., Murphy, N.P., Cooper, S.J.B. & Austin, A.D. 2012. Desert springs: deep phylogenetic structure in an ancient endemic crustacean (*Phreatomerus latipes*). *PLoS ONE*, **7**(7) e37642.

Guzik, M.T., Austin, A.D., Cooper, S.J.B., Harvey, M.S., Humphreys, W.F., Bradford, T., Eberhard, S.M., King, R.A., Leys, R., Muirhead, K.A. & Tomlinson, M. 2011. Is the subterranean fauna uniquely diverse? *Invertebrate Systematics*, **24**: 407–418.

Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33:** 556–574.

Habel, J.C., Assmann, T., Schmitt, T. & Avise, J.C. 2010. Relict species: from past to future. In: *Relict species: Phylogeography and conservation biology*. (J.C. Habel & T. Assmann, eds.), pp. 1–8. Springer, Heidelberg, Germany.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, **405**: 907–913.

Hugall, A.F., O'Hara, T.D., Hunjan, S., Nilsen, R. & Moussalli, A. 2016. An exon-capture system for the entire class Ophiuroidea. *Molecular Biology and Evolution*, **33**: 281–294.

Humphreys, W.F. 2006. Aquifers: the ultimate groundwater-dependent ecosystems. *Australian Journal of Botany*, **54**: 115–132.

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yares, C.J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D. & Franklin, S.T. 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, **21**: 393–404.

King, R.A. 2009. Two genera and species of chiltoniid amphipods (Crustacea: Amphipoda: Talitroidea) from freshwater mound springs in South Australia. *Zootaxa*, **2293**: 35–52.

King, R.A., Bradford, T., Austin, A.D., Humphreys, W.F. & Cooper, S.J.B. 2012. Divergent molecular lineages and not-so-cryptic species: the first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. *Journal of Crustacean Biology,* **32**: 465–488.

Klopfstein, S. Langille, B., Spasojevic, T., Broad, G.R., Cooper, S.J.B., Austin, A.D. & Niehuis, O. 2019. Hybrid capture data unravel a rapid radiation of pimpliform parasitoid wasps (Hymenoptera: Ichneumondae: Pimpliformes). *Systematic Entomology*, **44**: 361–383.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenetics. *Systematic Biology*, **61**: 727–744.

Leys, R. & Watts, C.H. 2008. Systematics and evolution of the Australian subterranean hydroporine diving beetles (Dytiscidae), with notes on *Carabhydrus*. *Invertebrate Systematics*, **22**: 217–225.

Li, C., Hofreiter, M., Straube, N., Corrigan, S. & Naylor, G.J.P. 2013. Capturing protein-coding genes across highly divergent species. *BioTechniques*, **54**: 321–326.

Martin, H.A. 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments*, **66**: 533–563.

Mayer, C., Sann, S., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., Petersen, M., Meusemann, K., Liere, K., Wäggle, J-W., Misof, M., Bleidorn, C., Ohl, M. & Niehuis, O. 2016. BaitFisher: a software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution*, **33**: 1875–1886.

Moritz, C. & Agudo, R. 2013. The future of species under climate change: resilience or decline? *Science*, **341**: 504–508.

Murphy, N.P., Adams, M. & Austin, A.D. 2009. Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. *Molecular Ecology*, **18**: 109–122.

Murphy, N.P., Breed, M.F., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2012. Trapped in desert springs: phylogeography of Australian desert spring snails. *Journal of Biogeography*, **39**: 1573–1582.

Murphy, N.P., Guzik, M.T., Cooper, S.J.B. & Austin, A.D. 2015. Desert spring refugia: museums of diversity or evolutionary cradles? *Zoologica Scripta*, **44**: 693–701.

Nevill, J.C., Hancock, P.J., Murray, B.R., Ponder, W.F., Humphreys, W.F., Phillips, M.L. & Groom, P.K. 2010. Groundwater-dependent ecosystems and the dangers of groundwater overdraft: a review and an Australian perspective. *Pacific Conservation Biology*, **16**: 187–208.

O'Hara, T.D., Hugall, A.F., Cisternas, P.A., Boissin, E., Bribiesca-Contreras, G., Sellanes, J., Paulay, G. & Byrne, M. 2019. Phylogenomics, life history and morphological evolution of ophiocomid brittlestars. *Molecular Phylogenetics and Evolution*, **130**: 67–80.

Portik, D.M., Smith, L.L. & Bi, K. 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Molecular Ecology Resources*, **16**: 1069–1083.

Reilly, S.B., Stubbs, A.L., Karin, B.R., Bi, K., Arida, E., Iskandar, D.T. & McGuire, J.A. 2019. Leap-frog dispersal and mitochondrial introgression: phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. *Journal of Biogeography*, **46**: 757–769.

Rix, M.G., Cooper, S.J.B., Meusemann, K., Klopfstein, S., Harrison, S.E., Harvey, M.S. & Austin, A.D. 2017. Post-Eocene climate change across continental Australia and the diversification of Australian spiny trapdoor spiders (Idiopidae: Arbanitinae). *Molecular Phylogenetics and Evolution*, **109**: 302–320.

Schmidt, C. 2008. Phylogeny of the terrestrial Isopoda (Oniscidea): a review. *Arthropod Systematics and Phylogeny*, **66**: 191–226.

Sfenthourakis, S. & Taiti, S. 2015. Patterns of taxonomic diversity among terrestrial isopods. *Zookeys*, **515**: 13–25.

Sniderman, J.M.K., Pillans, B., O'Sullivan, P.B. & Kershaw, A.P. 2007. Climate and vegetation in southeastern Australia respond to Southern Hemisphere insolation forcing in the late Pliocene-early Pleistocene. *Geology*, **35**: 41–44.

Sniderman, J.M.K., Woodhead, J.D., Hellstroma, J., Jordan, G.J., Drysdale, R.N., Tyler, J.J. & Porch, N. 2016. Pliocene reversal of late Neogene aridification. *Proceedings of the National Academy of Sciences*, **13***:* 1999–2004.

Taxonomy Decadal Plan Working Plan. 2018. *Discovering biodiversity: A decadal plan for taxonomy and biosystematics in Australia and New Zealand 2018–2028*. Australian Academy of Science and Royal Society Te Apārangi, Canberra and Wellington.

Teasdale, L.C., Köhler, F., Murray, K.D., O'Hara, T. & Moussalli, A. 2016. Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. *Molecular Ecology Resources*, **16**: 1107–1123.

Tomlinson, M. & Boulton, A.J. 2010. Ecology and management of subsurface groundwater dependent ecosystems in Australia – a review. *Marine and Freshwater Research*, **61**: 936–949.

Vandel, A. 1973. Les Isopodes terrestres (Oniscoidea) de la Mélanésie. *Zoologisches Verhandelingen,* **125**: 1–160.

Wen, J., Ickert-Bond, S.M., Appelhans, M.S., Dorr, L.J. & Funk, V.A. 2015. Collections-based systematics: opportunities and outlook for 2050. *Journal of Systematics and Evolution*, **53**: 477–488.

Wood, H.M., González, V.L., Lloyd, M., Coddington, J. & Scharff, N. 2018. Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). Molecular Phylogenetics and Evolution, **127**: 907–918.

APPENDIX:

Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species

(published paper)

Statement of Authorship

Title of Paper	Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Guzik, M.T., Stringer, D.N., Murphy, N.P., Cooper, S.J.B., Taiti, S., King, R.A., Humphreys, W.F. & Austin, A.D. 2019. Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea: *Haloniscus*) reveals strong regional endemicity and new putative species. *Invertebrate Systematics*, **33**: 556–574.

Principal Author

Name of Principal Author	Michelle Guzik		
Contribution to the Paper	Conducted field work to collect specimens, conducted 75% of the data collection, conducted all of the phylogenetic analysis of the data, wrote and revised the manuscript, coordinated and contributed to production of the figures and tables, and acted as corresponding author.		
Overall percentage (%)	75%		
Signature		Date	8/7/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author (Candidate)	Danielle Stringer		
Contribution to the Paper	Conducted field work to collect rare and geographically inaccessible specimens, contributed new sequence data, important discussions that helped expand ideas, assisted with figures and tables and deposition of specimens to the South Australian Museum, and critically reviewed a number of versions of the manuscript.		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the second author of this paper.		
Signature		Date	8/7/19

Name of Co-Author	Nicholas Murphy		
Contribution to the Paper	As a major contributor to the broader original project on Great Artesian Basin springs, he helped conduct field work to collect specimens, important discussions that helped expand the concepts and critically reviewed a number of versions of the manuscript.		
Signature		Date	3/7/2019

Name of Co-Author	Steven Cooper
Contribution to the Paper	Provided some unpublished data and archival specimens, conducted field work to collect rare and geographically inaccessible specimens and critically reviewed a number of versions of the manuscript.
Signature	Date 8/7/2019

Name of Co-Author	Stefano Taiti
Contribution to the Paper	Provided taxonomic expertise and discussions on *Haloniscus* for which many of the species in this manuscript were previously unknown, and critically reviewed the manuscript.
Signature	Date 28/6/2019

Name of Co-Author	Rachael King
Contribution to the Paper	Conducted field work to collect specimens, provided important discussions on crustacean taxonomy and systematics that helped expand the ideas, and critically reviewed the manuscript.
Signature	Date 04/07/2019

Name of Co-Author	William Humphreys
Contribution to the Paper	Conducted field work to collect rare and geographically inaccessible specimens and provided knowledge and expertise of subterranean fauna, regional geology and climatic history, and critically reviewed a number of versions of the manuscript.
Signature	Date 27/6/2019

Name of Co-Author	Andrew Austin
Contribution to the Paper	Supervised development of work and critically reviewed manuscript.
Signature	Date 8/7/19

CSIRO PUBLISHING

Invertebrate Systematics, 2019, **33**, 556–574
https://doi.org/10.1071/IS18070

Molecular phylogenetic analysis of Australian arid-zone oniscidean isopods (Crustacea : *Haloniscus*) reveals strong regional endemicity and new putative species

Michelle T. Guzik[A,H], *Danielle N. Stringer*[A], *Nicholas P. Murphy*[B], *Steven J. B. Cooper*[A,C], *Stefano Taiti*[D,E], *Rachael A. King*[A,C], *William F. Humphreys*[F,G] *and Andrew D. Austin*[A]

[A]Australian Centre for Evolutionary Biology and Biodiversity, School of Biological Sciences, The University of Adelaide, North Terrace, SA 5005, Australia.

[B]Department of Ecology, Environment and Evolution, La Trobe University, Bundoora, Vic. 3086, Australia.
[C]South Australian Museum, North Terrace, Adelaide, SA 5000, Australia.
[D]Istituto di Ricerca sugli Ecosistemi Terrestri, Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, 50019 Sesto Fiorentino Florence, Italy.

[E]Museo di Storia Naturale dell'Università, Sezione di Zoologia La Specola' Via Romana 17, 50125 Florence, Italy.
[F]Western Australian Museum, Welshpool DC, WA 6986, Australia.
[G]School of Animal Biology, The University of Western Australia, Crawley, WA 6009, Australia.
[H]Corresponding author. Email: michelle.guzik@adelaide.edu.au

Abstract. During the Miocene, central and western Australia shared a warm–wet environment that harboured a mesic rainforest fauna. Now, although the area is within the arid climate zone, it provides a habitat for highly diverse groundwater-associated invertebrates. Periods of global cooling and aridification during the late Miocene resulted in isolated desert refuges that retained ancient lineages. We aimed to characterise oniscidean isopod crustaceans from three refugial locations in the arid zone, and salt lakes, to identify new putative species. Extensive sampling and sequencing of the mitochondrial *Cytochrome Oxidase* c *subunit 1* gene and the *18S* rRNA gene were conducted. A molecular phylogenetic analysis of the oniscidean genus *Haloniscus* showed results consistent with a relictualisation hypothesis of widespread populations from across South Australia to Western Australia with subsequent geographic isolation and diversification of new species within habitats. We observed significant regional endemicity, but some lineages were not regionally monophyletic, pointing to past connectivity. We expand the range of *Haloniscus* and identify at least 26 putative species from arid-zone locations in Australia, with substantial phylogeographic structure within locations. These findings highlight the importance of relictual groundwater habitats as refugia for a diverse fauna representing early climatic history in Australia's arid zone.

Additional keywords: *CO1*, *18S* rRNA gene, groundwater, isolation, refugia.

Received 23 August 2018, accepted 21 January 2019, published online 4 June 2019

Introduction

Refugial habitats are isolated localities that faunal communities retreat to and persist in during major climatic changes (Hewitt 2000; Provan and Bennett 2008; Keppel *et al*. 2012; Davis *et al*. 2013; Moritz and Agudo 2013). To date, a large number of groundwater-dependent faunal communities, harbouring an exceptional biodiversity, have been recognised as refugial in Yilgarn calcrete aquifers of Western Australia (WA) (subterranean) and the Great Artesian Basin (GAB) springs of South Australia (SA) (surface) (Davis *et al*. 2013). Part of the significance of these refugial communities is the sheer number of species that have evolved through the retention of ancient lineages and their subsequent expansion/diversification, primarily through geographic isolation and subsequent allopatric speciation (Murphy *et al*. 2015*a*). Many of the individual taxon groups from groundwater-dependent refugia in Australia are common to multiple groundwater regions. For instance, chiltoniid amphipods are found in both groundwater-fed surface habitats of SA GAB springs (King 2009; Murphy *et al*. 2009, 2013, 2015*b*) and subterranean groundwater of WA calcrete aquifers (Cooper *et al*. 2007), while parabathynellids (Guzik *et al*. 2008) and dystiscid diving beetles (Leys *et al*. 2003; Leys and Watts 2008) are ubiquitous throughout the different calcrete aquifers of WA and the Northern Territory (NT). Phylogenetic studies have shown that many of these taxon groups, e.g. paraplatyarthrid isopods (Javidkar *et al*. 2018) and chiltoniid amphipods (Murphy *et al*. 2009, 2013, 2015*a*; King *et al*. 2012), are likely to represent ancient lineages of

swamp and rainforest dwellers adapted to novel groundwater-dependent habitats once their surface water habitats disappeared (Humphreys 1993, 2000, 2001).

An additional taxonomic group that is poorly known from groundwater-dependent refugia in Australia and that has rainforest connections is the oniscidean isopod genus *Haloniscus* Chilton, 1920. Most of the aquatic or water-associated species of Oniscidea belong to the Synocheta (families Trichoniscidae and Styloniscidae) and only a few to the Crinocheta (higher Oniscidea), i.e. species of *Paradoniscus* Taiti & Ferrara, 2007 from Socotra Island, Yemen, and *Haloniscus*. All these species are considered to be specialised forms derived from terrestrial species and secondarily adapted to live in water (Taiti and Xue 2012). To date, *Haloniscus* comprises five nominal aquatic species (Taiti and Schotte 2016). *Haloniscus searlei* Chilton, 1920, the type species of the genus, occurs in inland lakes in southern Australia (WA, SA, Victoria (Vic.) and Tasmania) (Williams 1983), and *H. anophthalmus* Taiti, Ferrara & Iliffe, 1995 in anchialine cave waters (caves with physico-chemically stratified freshwater that has a subterranean connection to the ocean) in New Caledonia (Taiti *et al*. 1995). *Haloniscus searlei* alone is known to tolerate salinity levels up to twice that of sea water, and is considered an extremophile, able to survive in environments of physiological stress and physico-chemical instability (Bayly and Williams 1966). It is not surprising then that more recently *Haloniscus* species have been discovered and described from unusual arid-zone habitats such as subterranean calcrete aquifers in WA (Taiti and Humphreys 2001). Here, we document two new and distinctly different arid-zone habitats rich in *Haloniscus* diversity: (1) calcrete aquifers of the Ngalia Basin in the NT that have fully subterranean and aquatic (stygobiontic) *Haloniscus* living within them, and (2) groundwater-fed surface habitats of GAB springs in SA (Murphy *et al*. 2015*a*) that maintain populations of surface-dwelling *Haloniscus* sampled at the freshwater spring margins.

To date, the Yilgarn calcrete aquifers have revealed as many as 25 new endemic subterranean *Haloniscus* species based on morphological characters (three species described by Taiti and Humphreys (2001) and 22 unpublished species by Taiti and Humphreys) and molecular phylogeographic analyses (Cooper *et al*. 2008). The arid zone of the NT, particularly in its calcrete areas such as the Ngalia Basin, was recently found to harbour groundwater-dependent fauna similar to that of WA (Guzik *et al*. 2011; Davis *et al*. 2013). Calcrete aquifers of Ngalia Basin are geologically reminiscent of aquifers found in the Yilgarn of WA, yet very little investigation of their stygofauna has been undertaken to date (Humphreys 2006, 2008; Guzik *et al*. 2011). The most comprehensively studied taxon group from the Ngalia Basin is the stygobitic diving beetles with one described species of *Copelatus* (Balke *et al*. 2004) and six described *Paroster* (Watts and Humphreys 2006). *Haloniscus* individuals have also been newly identified from SA GAB springs (Murphy *et al*. 2015*a*). In contrast to the subterranean and aquatic individuals sampled from calcrete aquifers of the Yilgarn and Ngalian Basins, GAB spring species of *Haloniscus* are not fully aquatic, but semiterrestrial, found only on the wet margins of the springs rather than in the water. In the only study that has included GAB spring *Haloniscus*, Murphy *et al*. (2015*a*) conducted an analysis of evolutionary rates across several different taxon groups. A concordance in diversification rates between faunal groups was observed in that study, but individual taxa and potential new species were not examined in detail. Therefore, phylogeographic patterns and the distribution of genetically divergent lineages/species of GAB spring *Haloniscus* are yet to be explored.

Short-range endemic taxa (Harvey 2002), especially those that survive as relicts, can inform on the climatic history of the region and the subsequent structure of the landscape and its impact on extant taxa (Moritz and Agudo 2013). Endemic taxa can also inform on the conservation significance of species and their communities (IUCN 2012). It is not unusual to find distinctive isopod species that are locally endemic in groundwater-dependent habitats (Wilson and Keable 1999; Gouws and Stewart 2007, 2013; Wilson *et al*. 2009; Guzik *et al*. 2012). Phylogenetic studies using DNA evidence have revealed that subterranean aquatic and terrestrial isopods elsewhere in the world consistently show evidence of extreme endemism (Rivera *et al*. 2002; Verovnik *et al*. 2004, 2005). A previous study of arid-zone *Haloniscus* by Cooper *et al*. (2008) used a single locus, *Cytochrome Oxidase* c *subunit 1* (*CO1*), to explore phylogeographic patterns among individuals sampled throughout the Yilgarn and qualitatively identified a minimum of 24 putative species. Here we aimed to broaden the scope of that study and document all known *Haloniscus* locations from groundwater and groundwater-associated habitats in arid Australia, as well as other undescribed salt lake species of *Haloniscus* from around the continent to provide a broader context for the groundwater species. We aimed to document refugial communities in groundwater-dependent ecosystems beyond those that are currently known, extend the known range of *Haloniscus* and characterise species from groundwater habitats in the arid zone of Australia that we know are refugial areas (GAB springs SA, Ngalia Basin NT and Yilgarn WA), as well as investigate the relationships among them. In so doing we aimed to (1) undertake more comprehensive sampling to better assess species diversity of *Haloniscus*, (2) greatly expand on the currently available *CO1* data by including data from a nuclear rRNA gene, (3) develop a phylogenetic framework for the genus, and (4) use this phylogeny to examine regional differences among *Haloniscus* and document putative new species.

Materials and methods

Collection localities and specimen collection

Haloniscus samples for this study were collected from two major arid-zone regions (see Fig. 1*A–D* for images of exemplar *Haloniscus* treated in this study and habitat (e.g. GAB spring) and Fig. 2 for maps): (1) GAB springs, SA (surface-dwelling and groundwater-associated fauna) (Fig. 1*A*, *D*, Fig. 2*C*, *D*) and (2) Ngalia Basin, NT (stygobiontic fauna) (Figs 1*B*, 2*A*, *B*). *18S* rRNA gene sequences were obtained for another arid-zone region: (3) calcrete aquifers of the Yilgarn region, WA (Fig. 1*C*, 2*E*) and also for (4) salt lake *Haloniscus* species from around WA and Vic. (Fig. 2). Each of the two new arid-zone locations has its own geological and geographical structure (Fig. 2). We (MTG, NPM, RAK) sampled at the Lake Eyre and Dalhousie supergroups of GAB springs during 2007–10 where

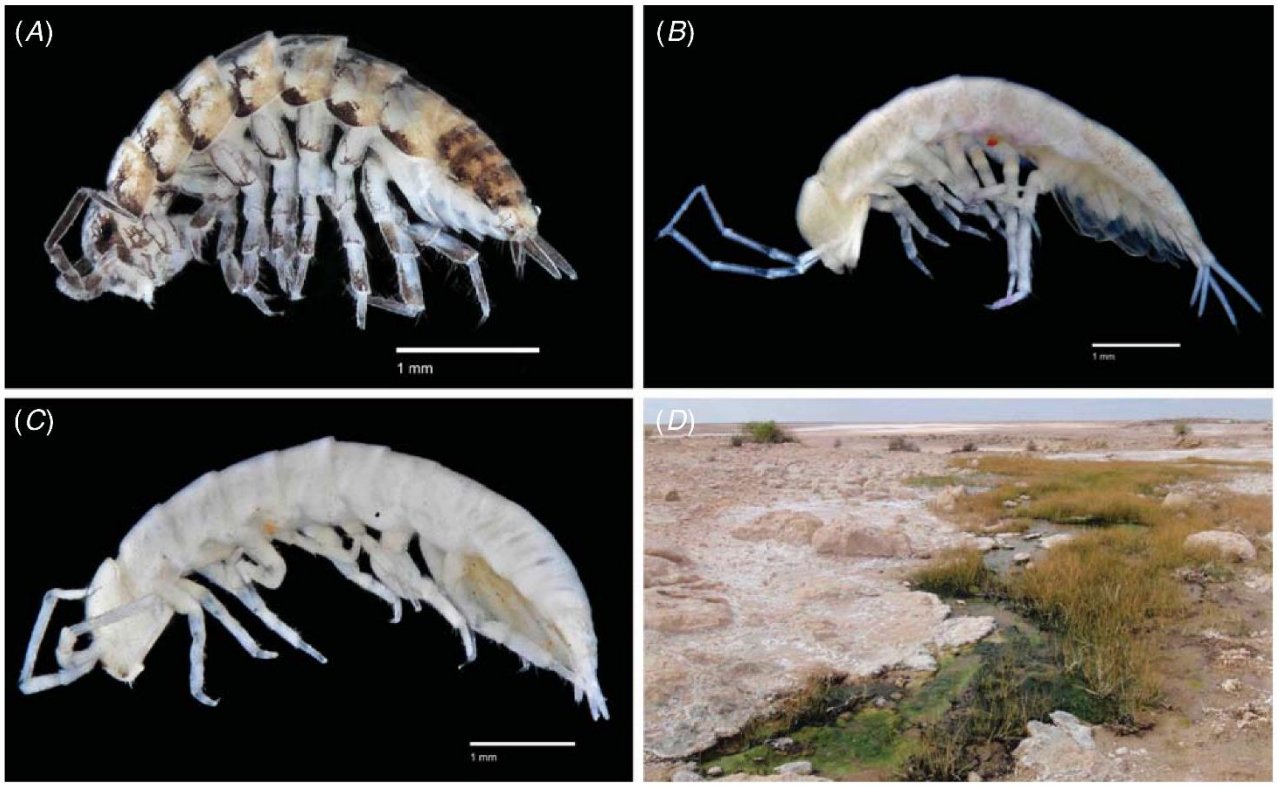


Fig. 1. Exemplar images of new *Haloniscus* species and one surface habitat. (*A*) GAB springs *Haloniscus* (Strangways springs); (*B*) Ngalia Basin calcrete aquifer stygobitic *Haloniscus* (Robbs Bore, Newhaven); (*C*) Yilgarn calcrete aquifer stygobitic *Haloniscus* (Shady Well, Laverton Downs); (*D*) GAB spring habitat (Coward Springs). Images by Alana Delaine.

specimens were all surface-dwelling fauna associated with the spring habitat, located throughout the southern and western portion of the Lake Eyre Basin in central Australia. Individual springs occur around areas of geological weakness (i.e. faults) and form large clusters of directly connected spring outlets, known as *spring groups*. Spring groups can be hierarchically categorised further at a broader geographical scale as *spring complexes* based on their hydrogeological and broader geographic location (Fig. 2*C*, *D*). The Ngalia Basin specimens were all stygobiontic and were collected in 2001–02 by WFH and in 2015 by DNS, SJBC and WFH at nine locations (Table 1, Fig. 2*A*, *B*). Unlike the calcrete aquifers of WA, those of the NT (Ngalia Basin) are poorly mapped and boundaries of specific calcretes are rarely clearly demarcated. However, in our study area, basement highs of impermeable basement rock ensure almost complete closure of the groundwater system, which exhibits very steep salinity gradients over several kilometres (English *et al.* 2001). In addition, the Basin is bisected by a north-westerly-trending Pre-Cambrian intrusion between Homestead and Rabbit Hole bores; Camel Bore is the only location to the south of the chain of salt lakes. The calcrete aquifers in Ngalia Basin maintain an intermediate alluvium (eroded and redeposited loose sediments through which groundwater can flow) between calcretes that connects the aquifers, making the groundwater habitat potentially continuous throughout the subterranean landscape. Here we sampled at four aquifers: three were calcrete aquifers (1) Newhaven 1 and 2 (these two locations comprised the same aquifer based on geological maps (Fig. 2*A*) but Camel Bore (Newhaven 2) may be isolated); (2) Napperby 2 (Fig. 2*B*) and (3) Napperby 3 (Fig. 2*B*); and (4) Napperby 1 which is an alluvial aquifer (Fig. 2*B*). At Ngalia Basin, individuals were sampled via bore holes, kept refrigerated until sorted and then stored in 100% ETOH at –20°C (after Cooper *et al.* 2007). Where possible, multiple individuals per location were sequenced to control for the possibility of sequencing errors and contamination. In contrast, the salt lakes were point location samples without duplicate individuals at a site. Details of the collection information are listed in Table 1. Calcrete aquifers in the Yilgarn are each isolated and discrete entities comprising groundwater calcretes (limestone-like substrate).

Novel GAB springs representatives

Haloniscus was unknown from the GAB springs before 2007, although other oniscoid families had been recorded (Greenslade 1985). These *Haloniscus* species occur in moist sandy soil under bushes on spring margins and, therefore, do not appear to be obligate aquatic organisms, making them distinct from Yilgarn species and some of the salt-lake species. *Haloniscus* did not occur at all GAB springs that were visited during field surveys in 2007–10. As well, *Haloniscus* species were collected at some non-spring water sources in this region (i.e. moist margins on the banks of rivers close to springs), although never in dry areas away from water. Individuals were examined by three of us (ST and subsequently by DNS and RAK) and identified as *Haloniscus*

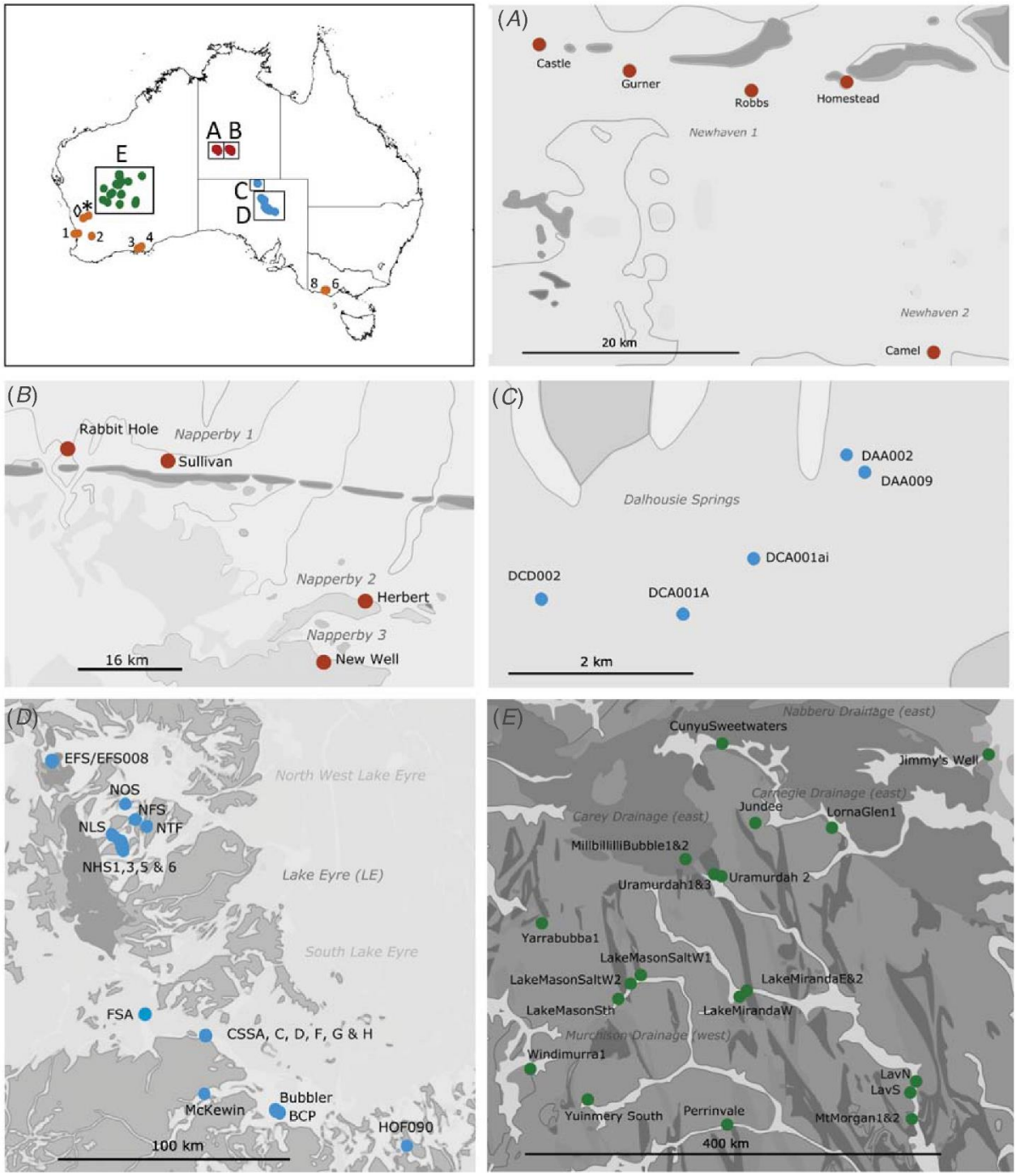


Fig. 2. Map of Australia showing grouped *Haloniscus* sample locations (*A–E*). Dots and labels represent sample locations and their name by region. Numbers 1–4, 6, 8 (orange dots) represent salt lake *Haloniscus searlei* from around Australia with numbers/identifiers (* and ♦ denote sample location details) from Table 1; Ngalia Basin bore (dark red dots) *A*, Mount Doreen (Newhaven 1 and 2 calcrete locations are shown); *B*, Napperby (Napperby 1–3 locations are shown)); GAB springs (blue dots) *C*, Dalhousie springs; *D*, Lake Eyre springs; calcrete aquifer bores in the Yilgarn, *E,* (green dots) (see full colour version online) and palaeodrainage channels are shown in light grey over regional geology (*E*).

Table 1. Locality information for groundwater habitats where *Haloniscus* individuals were sampled and subsequently DNA sequenced

CO1 and *18S* gene haplotype sequence information, GenBank Accession numbers, South Australian Museum Accession (SAMA) numbers and Western Australian Museum Biospeleology Collection numbers (BES) are given. Abbreviations for States within Australia that were sampled: SA, South Australia; WA, Western Australia; Vic., Victoria; NT, the Northern Territory. Ticks represent individuals that were successfully sequenced for each gene region (*CO1* and *18S*); bold indicates specimens that represent unique sequences submitted to South Australian Museum and Western Australian Museum as part of the present study

Locality	Locality code	Location ID (GAB/BES)	Latitude	Longitude	*CO1* Sequenced	*CO1* Haplotype name	*CO1* GenBank accession no.	*CO1* Catalogue or collection no.	*18S* Sequenced	*18S* Haplotype name	*18S* GenBank accession no.	*18S* Catalogue or collection no.
GAB springs (SA)[A]		**GAB**										
South of Lake Eyre												
Blanche Cup	BCP	682	–29.45292	136.85872					✓	1005.3		
McKewin	McKewin	669	–29.3862	136.54689	✓	**669.4** (2), **669.6**	KT235998, KT235999	SAMA C13226 SAMA C13180				
Strangways (A)	CSSA	732	–29.16363	136.55081					✓	732.1	MK286390	SAMA C13254
Strangways (C)	CSSC	734	–29.16317	136.55087	✓	**734.1**, **734.2** (2), **734.3**	KT236000, KT236001, KT236002	SAMA C13181 SAMA C13182 SAMA C13183				
Strangways (D)	CSSD	736	–29.16216	136.55167	✓	**736.2** (3), **736.4**	KT236003, KT236004	SAMA C13184 SAMA C13185				
Strangways (F)	CSSF	740	–29.16078	136.55132	✓	**740.1**, **740.2**	KT236005, KT236006	SAMA C13186 SAMA C13187				
Strangways (G&H)	CSSG&H	744	–29.15972	136.55116	✓	**742.1** (6), **744.3**	KT236007, KT236008	SAMA C13188 SAMA C13189				
Bubbler	Bubbler	765	–29.44637	136.85799	✓	**765.4**, **765.5** (2)	KT236010, KT236011	SAMA C13190 SAMA C13222	✓	1005.3		
Old Finnis (090)	HOF090	795	–29.58318	137.4408	✓	**795.3 (3)**	KT236012	SAMA C13191				
Francis Swamp	FSA	764	–29.07969	136.27686	✓	**764.7**(3)	KT236009	SAMA C13235	✓	**764.7** (6)	MK286391	SAMA C13235
North-west of Lake Eyre												
Loudon	NLS	922	–28.38122	136.14801	✓	**992.1** (2)	KT236013	SAMA C 13192				
Hawker (1)	NHS1	1005	–28.44288	136.19092	✓	**1005.2** (2), **1005.3**, **1005.4**	KT236014, KT236015, KT236016	SAMA C13193 SAMA C13194 SAMA C13195	✓	**1005.3** (2)	MK286387	SAMA C13193
Hawker (3)	NHS3	1007	–28.42505	136.18608	✓	**1007.5** (5)	KT236017	SAMA C13196				
Hawker (5)	NHS5	1009	–28.40578	136.18332	✓	**1009.1** (5)	KT236021	SAMA C13197				
Hawker (6)	NHS6	1008	–28.38448	136.15102	✓	**1008.1**, **1008.2**, **1008.3**	KT236018, KT236019, KT236020	SAMA C13229 SAMA C13198 SAMA C13230				
The Fountain	NTF	1010	–28.348	136.28271	✓	**1010.1** (3), **1010.2** (2)	KT236022, KT236023	SAMA C13199 SAMA C13200	✓	1005.3 (2)		
Fanny	NFS	1011	–28.32282	136.23781	✓	**1011.1** (5)	KT236024	SAMA C13231	✓	1005.3 (2)		
Outside	NOS	1012	–28.26256	136.19849	✓	**1012.1** (3), **1012.4**	KT236025 KT236026	SAMA C13201 SAMA C13202	✓	1005.3		
Freeling (EFS008)	EFS	1571	–28.0762066	135.9040301	✓	**1571.1**, **1571.2**, **1613.2**	KT236029, KT236030, KT236031	SAMA C13249 SAMA C13203 SAMA C13252	✓	1005.3 (6)		
Dalhousie springs[Λ]												
Kingfisher	DAA001	1437	–26.4083010	135.5216360					✓	1433.1		
Kingfisher	DAA002	1478	–26.4067164	135.5197814	✓	**1478.1**	KT236028	SAMA C13204	✓	1433.1		

(*continued next page*)

Table 1. (*continued*)

Locality	Locality code	Location ID (GAB/BES)	Latitude	Longitude	*CO1* Sequenced	*CO1* Haplotype name	*CO1* GenBank accession no.	*CO1* Catalogue or collection no.	*18S* Sequenced	*18S* Haplotype name	*18S* GenBank accession no.	*18S* Catalogue or collection no.
Kingfisher	DAA009	1433	–26.4083013	135.5216366	✓	**1433.2**	KT236027	SAMA C13245	✓	**1433.1** (2)	MK286388	SAMA C13255
Loveheart Pool	DCD002	1441	–26.4198392	135.4889121	✓	**1441.1, 1441.2**	KT236032, KT236033	SAMA C13205 SAMA C13206	✓	**1441.1**	MK286389	SAMA C13205
Main Pool	DCA001A	1459	–26.421179	135.5032145		**1459.1, 1459.2**	KT236034, KT236035	SAMA C13242 SAMA C13207	✓	1433.1		
Main Pool	DCA001ai	1494	–26.4161317	135.5104226	✓	**1494.1**	KT236036	SAMA C13208				
Ngalia (NT)[B]												
Newhaven (NH) 1												
Gurner Bore		6667	–22.716	130.98416	✓	**6667**	MK257746	**BES 6667**				
Gurner Bore		9461			✓	**9461, 9461.1, 9462.1**	MK257749 MK257750 MK257751	**BES 9461, BES 9461.1, BES 9462.1**				
Castle Bore		9458	–22.6959	130.90845	✓	**9458**	MK257748	**BES 9458**				
Homestead Bore		9465.1	–22.72519	131.16636	✓	**9465.1**	MK257752	**BES 9465.1**				
Homestead Bore		18774.1			✓	**18774.1** (3)	MK257760	**BES 18774.1** (3)	✓	**18774.1**	MK286394	**BES 18774.1**
Homestead Bore		18774.2							✓	18754.1		
Homestead Bore		18774.3			✓	**18774.3**	MK257761	**BES 18774.3**	✓	18754.1		
Homestead Bore		18774.4			✓	**18774.4**	MK257762	**BES 18774.4**	✓	18754.1		
Homestead Bore		18778.1							✓	18754.1		
Robbs Bore		18775.1	–22.73141	131.08667	✓	**18775.1**	MK257763	**BES 18775.1**	✓	18754.1		
Robbs Bore		18775.2							✓	18754.1		
Robbs Bore		18775.3			✓	**18775.3** (2)	MK257765	**BES 18775.3** (2)	✓	18754.1		
Robbs Bore		18775.4			✓	**18775.4** (2)	MK257764	**BES 18775.4** (2)	✓	18754.1		
Robbs Bore		18775.5							✓	18754.1		
Robbs Bore		10076			✓	**10076**	MK257755	**BES 10076**				
Newhaven (NH) 2												
Camel Bore		18773	–22.93439	131.23972	✓	**18773**	MK257759	**BES 18773**	✓	18754.1		
Napperby (NA) 1												
Sullivan Bore		18754	–22.73614,	132.46105	✓	**18754.1** (2)	MK257756	**BES 18754.1** (2)	✓	**18754.1** (2)	MK286392	**BES 18754.1**
Rabbit Hole (Central Mt Wedge) Bore		10064	–22.71776	132.32386	✓	**10064.1**	MK257754	**BES 10064.1**				
Rabbit Hole (Central Mt Wedge) Bore		18759.1			✓	**18759.1**	MK257757	**BES 18759.1**	✓	**18759.1**	MK286393	**BES 18759.1**
Rabbit Hole (Central Mt Wedge) Bore		18759.2			✓	**18759.2**	MK257758	**BES 18759.2**	✓	18754.1		
Rabbit Hole (Central Mt Wedge) Bore		18759.3			✓	**18759.2**	MK257758	**BES 18759.2**	✓	18754.1		
Napperby (NA) 2												
Herbert Bore		8086	–22.90891	132.72908	✓	**8086**	MK257747	**BES 8086**				
Napperby (NA) 3												
New Well Bore		10054	–22.98788	132.67552	✓	**10054**	MK257753	**BES 10054**	✓	18754.1		
Yilgarn calcrete aquifers (WA)[B]		**BES**										
Carnegie drainage (east)												
Jundee		6601	–26.26876	120.68094	✓		EU364563					
Lorna Glen (Site 1)		12880.2 12880.1 12892.1	–26.29826	121.40341	✓✓✓		EU3645681 EU3645651		✓✓✓	**12880.2** **12880.1** 12880.2	MK286396 MK286395	**BES 12880.2** **BES 12880.1**

(*continued next page*)

Table 1. (*continued*)

Locality	Locality code	Location ID (GAB/BES)	Latitude	Longitude	Sequenced	Haplotype name	*CO1* GenBank accession no.	Catalogue or collection no.	Sequenced	Haplotype name	*18S* GenBank accession no.	Catalogue or collection no.
		12888.1			✓✓✓				✓	12880.2		
		12888.2										
		12888										
Jimmy′s Well		H.nsp.12	–25.66058	122.87469			EU364569					
Nabberu drainage (east)												
Cunyu: Sweetwaters		12860	–25.59375	120.37241	✓		EU364570		✓	**12860**	MK286397	**BES 12860**
Carey drainage (east)												
Millbillillie (Bubble Well) (site 1)		5637 5637.1	–26.56072	120.0409	✓		EU364571					
Millbillillie (Bubble Well) (site 2)		14282.2 12910.1	–26.56362	120.04265	✓✓✓✓✓		EU364577		✓✓	**14282.2**	MK286399M	**BES 14282.2**
		12910.2								**14282.1**	K286398	**BES 14282.1**
		14282.0 14282.1										
Uramurdah (site 1)		12831.1 12831.2	–26.68763	120.30268	✓✓✓✓		EU364581		✓	12880.2		
H. stilifer		12831.3 6455										
H. longiantennatus												
Uramurdah (site 2)		10527	–26.68762	120.35283	✓		EU364582					
H. longiantennatus		14289.1 14289.0			✓		EU364579		✓✓	12880.2		
H. longiantennatus										12880.2		
Uramurdah (site 3)		14299	–26.6876	120.3078	✓		EU364585					
Lake Miranda East (site 1)		10459	–27.66407	120.61167	✓		EU364586		✓	**10459**	MK286403	**BES 10459**
Lake Miranda East (site 2)		14336	–27.6634	120.6124	✓		EU364587		✓	10459		
Laverton (Northern site)		10291.1 10291.2	–28.39652	122.19766	✓✓✓✓✓		EU364590		✓	**12011**	MK286400	**BES 12011**
		10290 13173.1										
		13186.2 12011										
		13180.1			✓		EU364593					
		12005			✓		EU364595					
Laverton (Southern site)		12087 12021.1	–28.48423	122.13336	✓✓✓✓✓		EU364596		✓✓	**12102.1**	MK286401	**BES 12102**
		12102.1 13141								12011		
		13149.1 13157										
		13149.2			✓		EU364600					
Mount Morgans (site 1)		10582.1 10582	–28.73177	122.1569	✓		EU364601		✓	**10582**	MK286402	**BES 10582**
Mount Morgans (site 2)		11811 11811.1	–28.73159	122.15884	✓✓		EU364602					
Lake Miranda West		10532	–27.71085	120.54332	✓		EU364604		✓	**10532**	MK286404	**BES 10532**
Lake Mason (Salt Well) (site 1)		10410	–27.53999	119.62427	✓		EU364605		✓	**10410**	MK286406	**BES 10410**
Lake Mason (Salt Well) (site 2)		14360.0	–27.586	119.5218	✓		EU364606		✓			
		14360.2			✓✓		EU364607		✓	10410		
		14360.3										
Lake Mason South		13232	–27.71371	119.39969	✓		EU364611		✓	**13232**	MK286405	**BES 13232**
Yuinmery South		6655	–28.54862	119.09113	✓		EU364609					
Perrinvale		10257	–28.77504	120.417	✓		EU364610		✓	**10257**	MK286407	**BES 10257**
Murchison drainage (west)												
Yarrabubba Nowthanna (site 1)		13093	–27.06683	118.67994	✓		EU364612		✓	**13093**	MK286408	**BES 13093**
		13101										
Windimurra (site 1)		8956	–28.2861	118.5743			EU364614					
		8956.2										
		13133										

(*continued next page*)

Table 1. (*continued*)

Locality	Locality code	Location ID (GAB/BES)	Latitude	Longitude	Sequenced	Haplotype name	*CO1* GenBank accession no.	Catalogue or collection no.	Sequenced	Haplotype name	*18S* GenBank accession no.	Catalogue or collection no.
Salt lakes												
Haloniscus searlei (1)		6573	–32.00	115.50	✓		EU364616					
Rottnest Island, WA												
Haloniscus searlei (2)		9886.1	–32.228	117.358	✓		EU364617		✓	**9886.1**	MK286409	**BES 9886.1**
Lake Mears, WA												
Haloniscus searlei (3)		9887.1	–33.86667	120.0489	✓		EU364618					
North Parriup Lake, WA												
Haloniscus searlei (4)		9891.1	–33.466	122.613	✓		EU364619					
Beaumont Nature Reserve, WA												
Haloniscus searlei (6)		9912	–38.20	143.10	✓		EU364620					
Lake Gnotuk, Vic.												
Haloniscus searlei (8)		9914.2	–38.200	142.867	✓		EU364621					
Keilambete Lake, Vic.												
Andricophiloscia stepheni		9888.1	–30.302	116.454	✓		EU364622					
Lake Martinjinni NR, WA[C]												
Haloniscus sp. 9890		9890	–29.983	117.000	✓		EU364623					
Lake Goorly, WA[D]												
Outgroup taxa												
Haloniscus anophthalmus		10201	–22.6166642	167.4833314			EU364626		✓			
(New Caledonia)												
Pygolabis humphreysi		11441	–23.32944	115.31531			EU364628		✓			
(Newman Borefield)[D]												

[A]Individuals sampled from surface springs.
[B]Individuals sampled from bores that access subterranean aquifers.
[C]Individuals sampled at salt lakes and marked as ♦ in Fig. 2.
[D]Individuals sampled at salt lakes and marked as * in Fig. 2.

using morphological characters to distinguish them from other oniscidean genera (Chilton 1920; Taiti *et al.* 1995; Taiti and Humphreys 2001).

DNA extraction, polymerase chain reaction (PCR) and sequencing

For all DNA extractions from GAB springs and Ngalia Basin specimens, every effort was made to remove one or two appendages from each individual on a single side so that the full set of appendages on the opposing side were retained for morphological examination.

We accessed additional sequences from GenBank to reconstruct our complete arid-zone *Haloniscus* phylogeny. These sequences included *CO1* sequences for individuals from GAB springs (KT235998–KT236036: Murphy *et al.* 2015*a*) and Yilgarn calcretes, as well as sequences for *H. longiantennatus*, *H. searlei*, *Andricophiloscia stepheni* and *H. stilifer* (EU364563–EU364628: Cooper *et al.* 2008) (Table 1). In addition, we were able to use archived DNA extracts from Cooper *et al.* (2008) to amplify and sequence *18S* fragments for selected individuals from the Yilgarn region and *H. searlei* individuals from WA and Vic.

To examine phylogenetic relationships among *Haloniscus*, we used partial DNA sequences of the mtDNA gene *CO1* and *18S* gene. DNA was extracted using the Gentra Systems PUREGENE DNA Purification Kit (Gentra Systems, Minneapolis, MN, USA) according to the manufacturer's protocol. A 621 base pair (bp) region of *CO1* was amplified with the universal oligonucleotide primers LCO1490 (5′-GGTCAACAAATCATAAAGATAT TGG-3′) and HCO2198 (5′-TAAACTTCAGGGTGACCAAA AAATCA-3′) (Folmer *et al.* 1994). A ~600-bp fragment of *18S* was PCR-amplified using the forward arthropod primer 1.2F (5′-TGCTTGTCTCAAAGATTAAGC-3′) (Whiting 2002) and the reverse b3.9 (5′-TGCTTTRAGCACTCTAA-3′) (Whiting *et al.* 1997). PCR-amplification of all sequences was carried out in 25-μL reactions containing 10x Amplitaq Gold Buffer, 2.5 mM Mg^{2+}, 2.5 mM of each deoxyribonucleotide triphosphate (dNTP), 0.5 μM of each primer, 0.1 units of Amplitaq Gold Polymerase and ~1 ng of DNA. Cycling conditions involved an initial denaturation at 95°C for 9 min, and 35 subsequent cycles of 94°C for 30 s, 47–50°C for 30 s and 72°C for 90 s. PCR products were purified using Agencourt AMPure XP PCR Purification and sequenced using the ABI PRISM Big Dye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA). DNA sequences were analysed commercially on an ABI PRISM 3700 (Applied Biosystems). All sequences were edited with reference to chromatograms using Geneious 5.0 and aligned with MUSCLE (Edgar 2004*a*, 2004*b*). All new and unique *CO1* and *18S* sequences for Ngalia Basin samples, and *18S* sequences for GAB springs, Yilgarn calcretes, *H. longiantennatus* and *H. searlei*, were submitted to GenBank (see Table 1).

Initially, we used the same outgroups as Cooper *et al.* (2008) to represent an array of isopod families and genera, including Tylidae (*Tylos neozelanicus* Chilton, 1901), Detonidae (*Deto marina* Chilton, 1884, *Armadilloniscus ellipticus* Harger, 1878), Philosciidae (*Laevophiloscia yalgooensis* Wahrberg, 1922, *Haloniscus anophthalmus* (New Caledonia) Taiti, Ferrara & Iliffe, 1995), Porcellionidae (*Porcellionides pruinosus* Brandt, 1833) and Flabellifera (*Pygolabis humphreysi* Wilson, 2003). However, several of these taxa were problematic in the phylogenetic analyses due to long branch attraction (see Cooper *et al.* 2008). Final outgroups were the evolutionarily distinct *H. anophthalmus* (EU364626) and *P. humphreysi* (EU364628).

Species delimitation

We implemented the Unified Species Concept from de Queiroz (1998, 2005, 2007), which states that species are 'a separately evolving metapopulation lineage' de Queiroz (2007: 881) with the categories that relate to our classification of a species being: phylogenetic, monophyletic and genealogical species. We used two quantitative species delimitation methods to help infer the putative species (lineages) across our phylogenetic tree: a Bayesian Poisson Tree Processes (bPTP) analysis implemented using the bPTP server (http://species.h-its.org/ptp/) (Zhang *et al.* 2013). The single-locus species delimitation method bPTP uses only nucleotide substitution information and implements a model assuming gene tree branch lengths generated by two independent Poisson process classes (within- and among-species substitution events). We ran this analysis on an unrooted phylogeny without outgroup taxa because distantly related taxa can unduly affect the results. For the bPTP analyses, we used the Bayesian 50% majority-rule consensus tree from MrBayes for both *CO1* and *18S* (data not shown). To compare the results from the tree-based method, bPTP, we used the Automatic Barcode Gap Discovery (ABGD) method (Puillandre *et al.* 2012*a*), which is a computationally efficient distance-based method of species delimitation. It performs well when compared with tree-based coalescent methods (Puillandre *et al.* 2012*b*; Kekkonen and Hebert 2014; Kapli *et al.* 2017) and other threshold techniques (Ratnasingham and Hebert 2013). This method seeks to quantify the location of the barcode gap that separates intra- from interspecific distances.

Phylogenetic analysis

Nucleotide sequence divergences (i.e. the number of base substitutions per site) were estimated, using the Kimura 2-parameter (K2P) model (Kimura 1980) as implemented by MEGA 6.0 (Tamura *et al.* 2013), within and between bPTP lineages (see below for detail on lineages) and one representative Yilgarn calcrete aquifer (Jundee). The K2P model was chosen to provide comparative data, with studies that have previously proposed threshold levels of divergence among crustacean species (Lefébure *et al.* 2006*b*). Given the low level of divergence found within bPTP lineages, mean divergences (i.e. divergences averaged over all sequence pairs) were calculated for inter-lineage comparisons (Table 2).

Phylogenetic reconstruction was undertaken using both Bayesian inference (BI) in MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) and maximum likelihood (ML) in RAxML (Randomised Axelerated Maximum Likelihood) 7.0 (Stamatakis 2014), and run on the CIPRES Science Gateway 3.3 (Miller *et al.* 2010). The RAxML analysis used 1000 rapid bootstrap inferences and the likelihood of the best tree was optimised and evaluated under the General Time Reversible (Tavaré 1986) (GTR) +Invariant Sites (I) +Gamma-Distributed

Substitution Rates (G) model. For use in MrBayes, the model that best fitted the data was estimated with ModelTest 3.7 (Posada and Crandall 1998) for nucleotide data under an Akaike information criterion framework. Models were tested for all three codon positions of *CO1*; the GTR +I +G model was favoured for *first*, Hasegawa–Kishono–Yano (Hasegawa *et al.* 1985) (HKY) model +I +G for the *second*, and GTR +I +G for the *third* position. Analyses were carried out separately for *CO1* and *18S* and on the concatenated dataset. For BI analyses, all parameters were unlinked and the rates were allowed to vary over the partitions. This approach allows each position to be treated independently and not assumed to be evolving at the same rate. Four chains were run simultaneously for 50 million generations, sampling trees every 1000 generations. To evaluate convergence to the stationary distribution, the program Tracer 1.4 (Rambaut *et. al.* 2018) was used. A burn-in of 25% was chosen and a 50% consensus tree was constructed from the remaining trees.

Results

A 621-bp fragment of the mtDNA gene *CO1* was sequenced for a total of 158 individuals and 119 haplotypes were obtained: 72 individuals (39 haplotypes) from GAB spring populations from western and southern Lake Eyre and 26 individuals (20 haplotypes) from Ngalia Basin. No stop codons or gaps were observed in any of the translated amino acid sequences for *CO1*, suggesting that the genuine mtDNA *CO1* gene was sequenced. We further added 60 *CO1* sequences from GenBank for individuals from the Yilgarn region, WA, and additional *Haloniscus* species representatives (i.e. salt lake *H. searlei* from around Australia) (Cooper *et al.* 2008), as well as two outgroup taxa. For *18S*, 69 individuals were sequenced for a maximum fragment size of 556 bp. For the final analyses, 20 haplotypes were suitable for concatenation with matching *CO1* sequences for a final concatenated dataset of 1177 bp. No *CO1* haplotypes were shared between Ngalia Basin bores. Furthermore, no *CO1* haplotypes were shared at a local scale (i.e. between GAB spring groups or between calcrete aquifers) or at a regional scale (i.e. between GAB springs and calcretes). However, evidence of haplotypes being shared between very closely located springs within spring groups was observed (e.g. CSSG and CSSH: see Table 1 for these codes) (Table 1). The conserved *18S* haplotypes were less discrete in their distribution than the more variable *CO1* haplotypes. A single *18S* haplotype was shared between Coward Springs, Neales and Freeling spring complexes.

Species delimitation

The bPTP analysis detected putative species on the input tree (marked by boxes beside the phylogenetic tree in Fig. 3). We observed Bayesian Support (BS) values >0.84 on each of the lineages from the bPTP analyses. BS values are an indication of confidence in delimited species where higher BS values for a node indicate that all descendants from this node are more likely to be from one species (Zhang *et al.* 2013). The total number of species delimited by ABGD was 26 compared with 44 with bPTP (four species were known). We observed three species delimited for both GAB springs and Ngalia Basin using ABGD, while seven species for Ngalia Basin and eight for GAB springs were delimited with bPTP.

From here on, we use the term 'lineages' to describe groups of closely related haplotypes from proximate geographic locations. We observed that both ABGD and bPTP delineated lineages best reflected phylogeographic structure, but we also observed that bPTP reflected the fine-scale geographic population structure especially well. We have, therefore, used bPTP 'lineages' to summarise delimitation results especially for genetic divergence estimates.

Sequence divergences

CO1 K2P divergences showed little divergence within bPTP lineages (0–0.04) but substantial diversification among lineages between the major arid-zone regions (GAB springs, Ngalia Basin aquifers and WA calcretes), with 17–26% divergence between all lineages (Table 2). Intra-regional divergences between bPTP lineages varied between the GAB springs and Ngalia Basin aquifers. Across the full extent of Lake Eyre supergroup springs, divergences between lineages ranged between 3% and 12% sequence divergence, excluding individuals from the southern Francis Swamp spring group. The Francis Swamp lineage revealed 20–21% sequence divergence from all other Lake Eyre spring group lineages. Between the Lake Eyre supergroup spring lineages and the Dalhousie supergroup spring lineages there was 23–26% divergence. At the intraregional level, lineages from all aquifers in the Ngalia Basin showed divergences of 2–16%. In one aquifer, Newhaven 1, where lineages were sampled across four bores ~30 km apart, divergences were variable (Table 2). For instance, between lineages from Gurner and Castle bores, divergences were only 2%, but from Gurner and Homestead sites they were 11%. In the Napperby 1 aquifer, lineages from Rabbit Hole Bore and Sullivan Bore were only 3% divergent. Divergences overall were very low for the *18S* fragment (1–3%) (data not shown).

Phylogenetic analyses

For the *CO1* BI tree (Fig. 3) seven major reciprocally monophyletic clades were resolved with strong support: Ngalia Basin clade (posterior probabililty (PP) 96%) (Fig. 3*A*), three Yilgarn clades (PP 95–100%) (Fig. 3*B–D*), *Haloniscus searlei* lineage (PP 100%) (Fig. 3*E*), Dalhousie clade (PP 100%) (Fig. 3*F*) and Lake Eyre spring clade (PP 100%) (Fig. 3*G*). The phylogenetic reconstruction of Yilgarn and salt lake haplotypes showed that these regions among *Haloniscus* individuals were not monophyletic. We observed three Yilgarn clades (Fig. 3*B–D*) with salt lake *Haloniscus* grouping within clade C and salt lake *H. searlei* (1–8) grouping with Yilgarn calcrete haplotype Yuinmery 6555 and Clade D. The addition of GAB spring sequences added another interesting layer to these relationships. We observed that the haplotypes from the FSA spring grouped most closely to the calcrete aquifer Clade B (Fig. 3) and haplotypes from the Yilgarn calcrete Windimurra grouped in among Dalhousie (Clade F) and Lake Eyre (Clade G) (Fig. 3). While few of the paraphyletic lineages were well supported in the *CO1* tree, the BI tree based on combined *CO1* + *18S* (Fig. 4) showed very strong similarities to the *CO1* tree but with improved support for deeper nodes. For instance, the

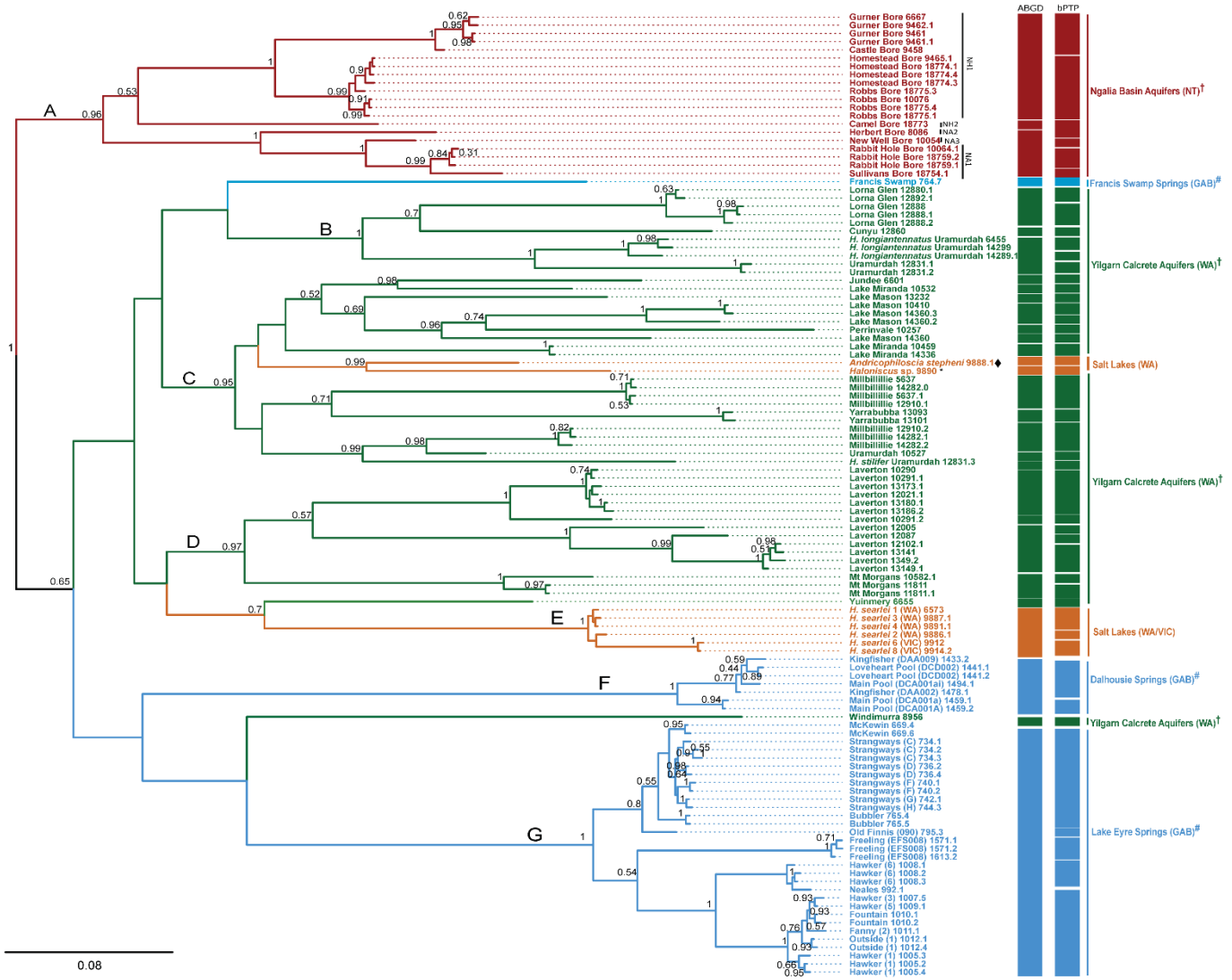


Fig. 3. Bayesian Inference *CO1* phylogeny of *Haloniscus* haplotypes from (1) Ngalia Basin, (2) GAB springs, (3) Yilgarn calcrete aquifers and (4) salt lakes. Haplotypes (branch tips) are labelled and colour-coded according to the arid-zone sample regions in the map in Fig. 2 (see full colour version online): Ngalia Basin (dark red); GAB springs (blue); Yilgarn calcrete aquifers (green); salt lake *Haloniscus searlei* (orange) from around Australia with species numbers/identifiers (numbers 1–4, 6, 8) (* and ♦ denote sample location details) from Table 1. Delineated species indicated to the right of the phylogeny are represented by bars for the ABGD and bPTP methods respectively. Clades representing Ngalia Basin (aquifer numbers Newhaven (NH) 1 and 2 and Napperby (NA) 1–3) and GAB springs (FSA, Francis Swamp) are annotated to the far right of the tree branches. Important nodes are marked A–G. Posterior probabilities are shown on nodes. Symbols: †, individuals sampled from bores that access subterranean aquifers; #, individuals sampled from surface springs.

grouping of Dalhousie/Lake Eyre (Clades F and G) was well supported at PP 99%, as was the FSA/Yilgarn calcrete grouping (Clade C, PP 97%) and the Ngalia clade (Clade A, PP 87%). For *CO1* + *18S*, weaker support was observed for Yilgarn calcrete and salt lake *Haloniscus* relationships (PP 55–100%), with lineages changing position on the tree, but still retaining the four major clades (B–E). For *18S* data, BI analyses (Supplementary Material 1) showed that the relationships among clades were consistent with *CO1*. The major differences between the individual gene trees were the position of the Ngalia Basin clade (Clade A), which appeared nested within the *18S* tree as a sister to a lineage from the Laverton calcrete, and *H. searlei* 2 was observed to be in a basal position in the *18S* tree when compared with the *CO1* tree. Data from ML analyses are not shown because results matched those of the BI analyses.

Ngalia Basin

The Ngalia Basin represents a new group of stygobiontic *Haloniscus* lineages that comprised individuals from four aquifers: three calcrete aquifers – (1) Newhaven 1 (NH1) and Newhaven 2 (NH2) (these two localities comprised the same aquifer based on geological maps (Fig. 1*A*), but Camel Bore (NH2) may be isolated (see Discussion), (2) Napperby 2 (NA2) (Fig. 2*B*) and (3) Napperby 3 (NA3) (Fig. 2*B*) – and one alluvial aquifer – (4) Napperby 1 (NA1) (Fig. 2*B*). All sequenced

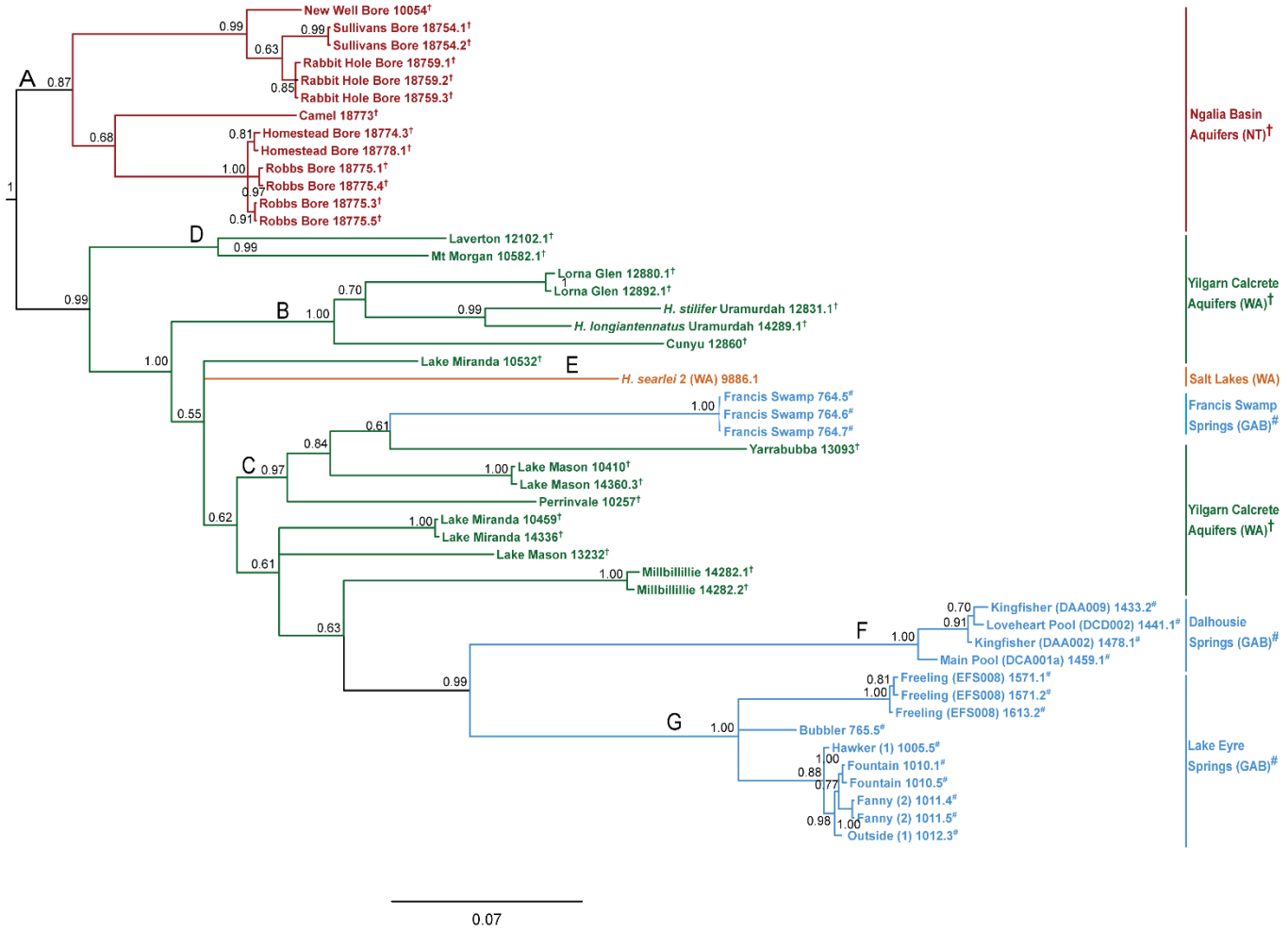


Fig. 4. Bayesian Inference *CO1* and *18S* combined tree. Haplotypes (branch tips) are labelled and colour-coded according to the arid-zone sample regions in the map in Fig. 2 (see full colour version online): Ngalia Basin (dark red); GAB springs (blue); Yilgarn calcrete aquifers (green); salt lake *Haloniscus searlei* (orange). Important nodes are marked A–G. Posterior probabilities are shown on nodes. Symbols: †, individuals sampled from bores that access subterranean aquifers; #, individuals sampled from surface springs.

haplotypes from Ngalia Basin individuals formed a well supported monophyletic group that was represented in both gene trees (*CO1* PP 96% and *18S* PP 99%) and in the combined tree (*CO1+18S* PP 87%). In the *CO1* tree there were three reciprocally monophyletic groups within the Ngalia Basin clade: Gurner/Castle, Homestead/Robbs, Sullivan/Rabbit Hole, and three lineages represented by single individuals: New, Herbert and Camel Bores. Each lineage was associated with a distinct aquifer and several lineages were found in the same aquifer (see Fig. 3). For instance, Gurner/Castle and Homestead/Robbs bores access the same large calcrete aquifer referred to as NH1.

GAB springs

Three major GAB spring clades were evident – Francis Swamp, Dalhousie and Lake Eyre (Fig. 3) – with substantial phylogeographic structure within two – Lake Eyre and Dalhousie. The Lake Eyre clade revealed no sharing of haplotypes among spring groups, except among proximate springs (e.g. CSSG and CSSH), while in the Dalhousie clade each sampled spring contained a unique haplotype. These results showed restricted genetic lineages to individual springs/spring groups. FSA represented an unusual result of being paraphyletic with clades from Yilgarn calcrete aquifers.

Discussion

At a time when anthropogenic climate change is putting substantial pressure on species viability, taxa that have successfully survived and adapted to global historic climate change are of particular interest (Moritz and Agudo 2013). Australia's arid zone is accepted as a region of exceptional biodiversity in refugial groundwater-dependent habitats and is a product of extended periods of fluctuating climatic change (Byrne *et al*. 2008; Davis *et al*. 2013; Humphreys 2017). Well known biodiversity hotspots are the groundwater aquifers in WA (Yilgarn and Pilbara) (Humphreys 1994, 1999, 2001, 2006, 2008; Guzik *et al*. 2011) and the GAB springs in SA (Guzik *et al*. 2011). We have further documented the biodiversity of

these locations and also characterised an additional location of new species diversity in the Ngalia Basin, NT. In total, we have identified a minimum of 26 new putative *Haloniscus* species within phylogenetic and geographical lineages from groundwater-dependent habitats of the arid zone. In particular, we observed three (ABGD) to seven (bPTP lineages) putative species from Ngalia Basin and three (ABGD) to eight (bPTP lineages) from GAB springs, and a minimum of 20 species from the calcrete aquifers and salt lakes. We also delineated four known and described species (i.e. *H. longiantennatus*, *H. searlei*, *H. stilifer*, and *Andricophiloscia stepheni*). Preliminary morphological examination supports our findings of delineated species (unpubl. data) and will be published elsewhere. We also observed *Andricophiloscia stepheni* (Location ID 9888.1, Lake Martinjinni) to be nested within the broader *Haloniscus* phylogeny and the taxonomy of this species will be revised in the future as a result of current findings. Each new putative species was found to be restricted to a small geographical range. The definition of short-range endemics (SRE) (10 000 km^2 (Harvey 2002), i.e. essentially the entire Ngalia Basin), is inappropriate for these groundwater fauna where even a definition for an ultra-SRE (uSRE) (term coined here) of 100 km^2 may be too coarse (Faille *et al.* 2015); Murphy *et al.* (2013) suggested 'microendemic' as a term to describe these taxa. The implications for taxa with extremely narrow ranges is significant because of the conservation challenges involved. All 26 putative species identified here fall into the uSRE category; a finding that makes the *Haloniscus* fauna of the arid zone especially speciose and geographically unique. Geographic isolation of SRE species can be indicative of refugial populations. That is, those lineages that historically shared a ubiquitous distribution across the landscape, but in the present day are restricted to remnant habitats. We discuss below the plausible hypothesis that the new putative species uncovered here represent relicts from the Miocene, as evidenced by paraphyletic clades between WA Yilgarn calcretes and SA GAB springs (Fig. 3). In summary, our analyses have (1) identified a minimum of 26 new putative species of *Haloniscus*, (2) extended the range of groundwater biodiversity to aquifers in the NT and GAB springs, and (3) further reinforced the findings of Murphy *et al.* (2015*a*), Cooper *et al.* (2008) and King *et al.* (2012) that groundwater-dependent faunal diversity in arid Australia is relictual and highly diverse.

Species delimitation

Our findings for species delimitation from GAB springs and calcrete aquifers in WA largely reflect the results of both Murphy *et al.* (2015*a*), who identified nine new lineages of *Haloniscus* (compared with our three to eight delineated by ABGD and bPTP, respectively) (Fig. 3), and Cooper *et al.* (2008) who identified at least 24 Yilgarn calcrete species. We note that Murphy *et al.* (2015*a*) implemented the GMYC model to detect changes from a Yule to a Coalescent process, revealing evidence for isolation of populations across taxon groups. However, their intention was not species delineation as GMYC is known to generate more Operational Taxonomic Units than other methods (Paz and Crawford 2012; Sauer and Hausdorf 2012; Miralles and Vences 2013; Talavera *et al.* 2013). We had some concern that the *CO1* locus is driving much of the phylogenetic signal in this study. However, the similarity in number of lineages between our study and those of Murphy *et al.* (2015*a*) and Cooper *et al.* (2008) provides us with confidence in our findings. We also have additional evidence from the nuclear gene *18S* data (Supplementary Material 1) that is largely in concordance with the lineages identified by *CO1* using different methods of species delineation (Fig. 3). Overall, the bPTP method returned ~50% more putative species than the more conservative estimates of the ABGD method (i.e. seven (bPTP) versus three (ABGD) delineated species for Ngalia Basin and eight (bPTP) versus three (ABGD) delineated species for GAB springs). For this reason, we primarily used the bPTP estimates as an indicator of 'lineages' rather than solely for defining putative species. It is most likely that the actual number of species that can be formally described from arid-zone groundwater habitats falls somewhere between these estimates. The information we have obtained in the current study combined with morphological examination of representative individuals from all populations will help to verify and determine the final suite of new *Haloniscus* species present in the GAB springs.

Haloniscus *species diversity in Australia's central arid zone*

Inhabitants of groundwater-dependent ecosystems, such as *Haloniscus*, are intrinsically (physiologically) adapted to groundwater habitats (Williams 1983) because they are repeatedly found in desert habitats that maintain sources of permanent water. These genetically divergent lineages range from subterranean and aquatic (blind species) to epigean species inhabiting damp habitats around desert freshwater springs, as well as the widespread *H. searlei* and unidentified *Haloniscus* from salt lakes across southern Australia (as discussed in Cooper *et al.* 2008). We consider these habitats physically extreme in nature and requiring intrinsic adaptations in the taxa that occupy them. As it stands, the genus *Haloniscus* is almost exclusively aquatic with one described species inhabiting salt lakes even though its ancestors are considered terrestrial (Williams 1983). *Haloniscus* can be considered an extremophile genus with *H. searlei* tolerant of salt concentrations of 8–160‰, where 35‰ is equivalent to sea water (Bayly and Williams 1966). The extremophile characteristics of *Haloniscus* are likely to have facilitated persistence of lineages in a variably arid environment over a long time-scale. Here we discuss each of the variable geographic locations of groundwater habitats in the arid zone that we have investigated and their respective putative species.

Ngalia Basin

The Ngalia Basin is a small (~15 000 km^2) sedimentary Basin in the southern NT that has its geological origins 850–340 million years ago (Wells and Moss 1983). This Basin was identified as a suitable habitat for stygofauna for hydrogeological reasons (Humphreys 2006, 2008). At this location we observed three species delineated by ABGD and possibly up to seven species of *Haloniscus* with bPTP lineages (Fig. 3). Each putative species (ABGD) corresponded to a region and Newhaven (NH1 and NH2) contained two species (see Fig. 3). These new putative species and their respective lineages are likely to represent intracalcrete diversification or

diversification by isolation and reconnection of different calcretes in the same system (see also Fig. 2 map). The first putative species was identified at four bores that access NH1 (Castle, Gurner, Robbs and Homestead) and their geographic distances are ~5 km apart with a maximum distance of ~15 km. The genetic structure observed in this putative species at NH1 is likely indicative of limited dispersal by *Haloniscus* and possible *in situ* diversification following isolation. The second putative species from Newhaven was identified at Camel Bore (NH2), a location that may have hydrogeological barriers south of the salt lakes (Humphreys 2009; Guzik *et al*. 2011). NH2 is the only location sampled to the south of a chain of salt lakes in the Ngalia Basin lying to the west of the Pre-Cambrian intrusion that may be a barrier to gene flow (Humphreys 2009; Guzik *et al*. 2011) and may make Camel Bore a geographically isolated location for subterranean fauna. In the Napperby region, a single putative species was identified by ABGD and phylogeographic differences were observed between the distinct calcrete aquifers (NA1, NA2 and NA3). Strong site fidelity of haplotypes to the geographic area at which they were sampled was the key observation at all Ngalia Basin aquifers. Hydrogeological barriers are likely to be the primary isolating factors responsible for narrow ranges observed in the new *Haloniscus* species identified here. While other stygofauna have been recorded from Ngalia Basin, in particular dytiscid diving beetles (Balke *et al*. 2004; Watts and Humphreys 2009) and Tateidae (Gastropoda: Caenogastropoda: Truncatelloidea: Ponder and Humphreys, unpublished), the current study is the first to examine *Haloniscus* and the relationships amongst individuals across the Ngalia Basin in detail.

GAB springs

Groundwater-fed springs are found throughout inland Australia, but the largest and most well known are the clusters of groundwater discharge from the GAB. In the present study, we sampled individuals from two major clusters of GAB springs (Dalhousie and Lake Eyre). We do not know of any terrestrial *Haloniscus* being recorded from the wet edge of any other springs in Australia before Murphy *et al*. (2015*a*), making our findings especially significant. Known as spring 'supergroups', these clusters typically share hydrogeological characteristics that unite the springs. The Lake Eyre and Dalhousie supergroups are two of 13 supergroups (Fensham and Fairfax 2003). From these two supergroups we have delineated up to eight (bPTP) new putative species based on lineages of *Haloniscus*. The new putative species here match natural geographic patterns and most display clear morphological differences (unpubl. data). At a maximum, we observed one to two putative species from Dalhousie spring supergroup (based on the two models of species delimitation), and a further one to six putative species from Hermit Hills, Coward, Strangways, Neales, Freeling and FSA spring complexes. Dalhousie springs are known to be geographically isolated from other GAB spring up-wellings and to host other evolutionarily distinct spring-endemic taxa (Zeidler 1991; Kodric-Brown and Brown 1993; Ponder 1995; Ponder *et al*. 1996; Murphy *et al*. 2009, 2013). We observed 19–26% *CO1* sequence divergence between Dalhousie *Haloniscus* lineages and those from other spring complexes (Table 2), indicating long-term isolation of the Dalhousie species. Furthermore, most of the lineages within the Lake Eyre supergroup (Hermit Hills, Coward, Strangways, Neales and Freeling springs) correspond to the geographic distribution of springs at the spring complex level (Murphy *et al*. 2015*a*), as also found for phreatoicid isopods at GAB springs (Guzik *et al*. 2012). The Lake Eyre lineage revealed reciprocally monophyletic lineages of individuals associated with distinct spring groups and no shared haplotypes between them, except in very closely located springs (e.g. CSSG and CSSH). Little is known of the biology of these *Haloniscus* species and their dependence on the wet margins around springs. However, we have observed geographically structured lineages, allopatric and putative species that are made up of old lineages. Murphy *et al*. (2015*a*) have suggested that ~5 million years divergence has occurred between northern and southern lineages of the Lake Eyre group, indicating that dispersal and gene flow is limited and these taxa are likely to have an obligate association with the permanent flow of groundwater from GAB springs. In this study we have further demonstrated that there is no evidence of recent genetic mixing between spring groups. Adaptive life-history strategies and isolation are strong drivers of local adaptation in GAB springs (Murphy *et al*. 2010). We believe that even though *Haloniscus* has the physiological potential for dispersal between geographically proximate clusters of springs during periods of high flow from the GAB or flooding (Murphy *et al*. 2010) it does not disperse due to its strong adaptation to immediate spring conditions (i.e. physico-chemistry). Further, the strong phylogeographic structure we have observed is consistent with that reported for other GAB spring endemic fauna, including wolf spiders, amphipods, isopods and gastropods (Gotch *et al*. 2008; King 2009; Murphy *et al*. 2009, 2010, 2012, 2013, 2015*a*; Guzik *et al*. 2012). Surprisingly, though, the Francis Swamp spring group, which is geographically located within the Lake Eyre supergroup, was paraphyletic to all other Lake Eyre putative species and we discuss this occurrence below.

Groundwater relicts in Australia's arid zone

In contrast to many biodiverse habitats that maintain high levels of both taxonomic and functional diversity (i.e. tropical rainforests, reefs), groundwater ecosystems have a limited functional biodiversity (Gibert and Deharveng 2002) and comprise consistently similar and ancient taxon groups. For example, the most frequently observed taxon groups found in groundwater are: crustaceans (amphipods, isopods, copepods, ostracods, syncarids), which are by far the dominant groups; coleopterans (dytiscid diving beetles); and molluscs (hydrobiid snails). Their functional roles and range of tolerances within groundwater habitats appear to have rendered these groups adaptable and robust to environmental change in a variety of similar habitats that, at times, can be extreme in their physico-chemical composition (Brock 1986; Cognetti and Maltagliati 2000), especially for organisms that have limited dispersal opportunities (Bohonak and Jenkins 2003). Despite the narrow range of taxon groups found in groundwater, we continue to describe extremely high levels of intrataxon diversity that is represented by new uSRE lineages that have evolved *in situ* (King 2009; Guzik *et al*. 2012; Murphy *et al*. 2012,

Table 2. Genetic divergences (Kimura 2-Parameter corrected) for *CO1* sequence data within (Intra-lineage) and mean genetic divergences between Bayesian Poisson Tree Processes (bPTP)-delineated lineages of Great Artesian Basin (GAB) springs and Ngalia Basin (new to the present study) and one individual from Yilgarn calcrete aquifers, Western Australia (Jundee)

The ascending order of the lineages corresponds to the order of the delineated lineages in the *CO1* phylogeny (Fig. 3); NA, Napperby; NH, Newhaven

Lineage	Intra-lineage	Ngalia Basin, Northern Territory						GAB springs, South Australia									WA calcrete
		Newhaven (NH1)		(NH2)	(NA2)	Napperby (NA3)	(NA1)	Lake Eyre	Dalhousie		Lake Eyre		Northern springs				
		Gurner	Homestead	Camel	Herbert	New Well	Rabbit Hole	Sullivans	FSA	DAA009	DCA001a	McKewin	HOF090	EFS008	NHS6	NHS3	Jundee
		6667	9465.1	18773	8086	10054	10064.1	18754.1	764.5	1433.2	1459.1	669.4	795.3	1571.1	1008	1007	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Gurner	0–0.02																
2 Homestead	0–0.02	0.11															
3 Camel	0	0.16	0.14														
4 Herbert	0	0.15	0.16	0.15													
5 New Well	0	0.15	0.16	0.15	0.09												
6 Rabbit Hole	0	0.15	0.14	0.15	0.1	0.05											
7 Sullivans	0	0.15	0.16	0.14	0.1	0.06	0.03										
8 FSA	0	0.21	0.21	0.19	0.18	0.17	0.15	0.17									
9 DAA009	0–0.02	0.23	0.23	0.22	0.21	0.21	0.19	0.2	0.23								
10 DCA001a	0–0.04	0.22	0.23	0.2	0.2	0.19	0.18	0.18	0.21	0.05							
11 McKewin	0–0.03	0.21	0.21	0.22	0.22	0.21	0.21	0.23	0.22	0.24	0.22						
12 HOF090	0	0.2	0.21	0.22	0.21	0.21	0.2	0.21	0.22	0.23	0.2	0.03					
13 EFS008	0	0.2	0.21	0.22	0.19	0.2	0.2	0.19	0.21	0.24	0.22	0.11	0.09				
14 NHS6	0–0.01	0.22	0.21	0.19	0.21	0.22	0.21	0.21	0.21	0.25	0.22	0.08	0.08	0.1			
15 NHS3	0–0.02	0.22	0.22	0.21	0.2	0.2	0.22	0.22	0.21	0.26	0.23	0.08	0.09	0.12	0.07		
16 Jundee	–	0.19	0.19	0.22	0.19	0.17	0.17	0.19	0.2	0.24	0.22	0.21	0.21	0.23	0.24	0.21	

2013, 2015*b*). Much of the observed species diversity in the Ngalia Basin, GAB springs, Yilgarn aquifers and salt lake *Haloniscus* has possibly occurred since the Miocene, based on dated phylogenies for other groundwater-dependent taxon groups such as amphipods (Murphy *et al*. 2009, 2013) and dytiscid diving beetles (Leys *et al*. 2003; Cooper *et al*. 2007, 2008). However, future work will need to investigate dating of divergence times following the development of a robust phylogeny for the taxa. It should be noted, however, that for the GAB springs, most species diversification appears to have occurred more recently (Murphy *et al*. 2015*a*). The divergences observed here between arid-zone regions were relatively high compared with intraregion divergences (not all *CO1* divergences are shown in Table 2 but between the three major arid-zone regions, represented by Yilgarn calcrete Jundee, Ngalia and GAB springs, divergences were 17–24%). These divergences were higher than the crustacean species divergence threshold of Lefébure *et al*. (2006*a*) and were considered to have occurred as a result of colonisation events into new habitats by ancient lineages, followed by the isolation of these populations in groundwater refugia during aridification of the Australian continent after the late Miocene.

An unusual result was observed for individuals sampled from GAB springs at Francis Swamp spring group. Three major GAB spring clades of *Haloniscus* were consistently found in all trees, namely Dalhousie, Francis Swamp and Lake Eyre. These clades represent reciprocally monophyletic groups of individuals sampled at geographically proximate springs. However, these three groups did not together comprise a monophyletic group in any of the phylogenetic analyses. The Francis Swamp lineage, from a single group of springs in the Lake Eyre supergroup, comprised individuals that are highly divergent from all other spring individuals. Its next closest relatives were from calcrete aquifers in the Yilgarn in both *CO1* and the combined *CO1+18S* trees (Figs 3, 4), which was still a highly divergent relationship. PP support for the latter groupings was generally low, except in the *CO1+18S* tree, where PP was 97%. The paraphyly of the Francis Swamp lineage supports the hypothesis of an old divergence compared with those of the Dalhousie and Lake Eyre springs (Guzik *et al*. 2012; Murphy *et al*. 2013). Francis Swamp consistently shows unique lineages (i.e. *Phreatomerus* and chiltoniid amphipods: Guzik *et al*. 2012; Murphy *et al*. 2013). However, overall, our analyses showed low support at deep nodes, which could be explained by the limitations of the target genes, with highly variable *CO1* third codon positions and highly conserved *18S* sequences. This explanation is likely since we have observed other differences between the individual gene trees (i.e. the position of Ngalia Basin lineages and also *H. searlei* 2 varied between *CO1* and *18S* trees). Furthermore, it is also possible that our results reflect a hard polytomy that resulted from a short period when climate changed and ancestral species were isolated in each of the refugial locations. Extinct as well as unsampled lineages have also possibly affected our results. Sampling all representatives and closest relatives is difficult when much of the groundwater fauna is unexplored or extinct (Guzik *et al*. 2012; Murphy *et al*. 2012, 2013, 2015*a*). We explain the observation of Francis Swamp paraphyly by a scenario of three independent colonisation events by divergent *Haloniscus* ancestors to three areas of spring activity (Francis Swamp, Lake Eyre and Dalhousie). Consistent with other research (King *et al*. 2012), this scenario supports a relictualisation hypothesis of widespread populations across SA and WA with subsequent geographic isolation and diversification of new species within spring habitats (Murphy *et al*. 2015*a*). More data are needed to support this scenario, but we are not the first to recognise paraphyly between geographically disparate locations across the arid zone as a connection between *Austrochiltonia dalhousiensis* (Zeidler 1997) and chiltoniid species from calcrete aquifers in WA has also been demonstrated (King 2009; Murphy *et al*. 2009, 2013, 2015*b*; King *et al*. 2012).

Our results identify the Ngalia Basin in NT, calcrete aquifers in WA and the GAB springs in SA as significant locations for groundwater isopods and relictual fauna more generally. These patterns of extreme endemism seen in three refugial groundwater regions of the Australian arid zone are consistent with patterns for other extremophile isopods (Gouws and Stewart 2007, 2013; Trontelj and Fišer 2009; Wilson *et al*. 2009) and other arid-zone taxa. Our findings expand the geographic range of *Halonsicus* and identify at least six new species from Ngalia Basin and GAB springs and a possible total of 26 (and up to 44) species from the arid zone and salt lakes. We know very little about the basic biology of these species but they appear to be poor dispersers, based on the strong phylogeographic structure, a trait that would explain the high levels of putative allopatric diversity revealed in this study. These findings, in addition to a large body of evidence that has documented extreme endemism and species richness in refugial arid-zone groundwater habitats, highlights the importance of these ecological communities. Salt lakes and calcretes are the focus of considerable mining interest for water, carbonates, uranium and rare earths so that many sites can be inferred to be under threat. Combined with extent of occurrence being <100 km^2 and area of occupancy of <10 km^2, most of these putative taxa could be classed as critically endangered (either category B1a or B1b, or B2a or B2b: IUCN 2012). There is a clear need for preservation of groundwater habitats that support multiple refugial species. These communities represent an ancient snapshot of the Australian continent's prehistory and thus are deserving of high conservation status.

Conflicts ofinterest

The authors declare that they have no conflicts of interest.

Declaration of funding

Funding support was provided by a Bioplatforms Australia (supported by Australian Biological Resources Study (ABRS) and The University of Adelaide) grant: 'Framework DNA Barcode Datasets for Australia: Subterranean Invertebrate Biodiversity Data-set' to ADA and SJBC, Australian Research Council Linkage grants LP0669062, LP100200494 and LP140100555 to ADA, NPM, SJBC and WFH, and an ABRS Capacity Building grant (CT214-11) to DNS.

Acknowledgements

Thanks to Steve Delean (The University of Adelaide), and Lewis Harsch for field assistance, and Travis Gotch (Department for Environment and Water) for invaluable field advice at GAB springs. For collections in National Parks

in South Australia we obtained a permit to 'undertake scientific research' (Permit No. Z25519 to Dr Nicholas Murphy) using appropriate methods and ethics approval from The Government of South Australia, Department for Environment and Heritage, signed for the Minister for Environment and Conservation. For collection on private property we obtained general permission from Greg Campbell (Chief Executive Officer) of S. Kidman & Co Ltd and also directly from several station managers to carry out our field collections. The station managers were: Randall Crozier for Anna Creek Station, Peter Paisley for Stuart Creek Station, Bobby Hunter for The Peak Station. We appreciate the access given to us by the traditional owners of the GAB spring country, particularly Reg Dodd (Arabunna people) and Dean Ah Chee (Irrwanyere Ranger at Witjira National Park), which enabled us to undertake our field collection with permission and guidance to access culturally sensitive land. For sampling in the Northern Territory (Ngalia Basin) we obtained a Parks and Wildlife Commission, Northern Territory, permit to take wildlife for commercial purposes (Permit No. 54946). To collect at Newhaven Sanctuary, we conducted a 'Proposal to carry out research at an Australian Wildlife Conservancy sanctuary' (Australian Wildlife Conservancy) reviewed by Josef Schofield, Newhaven Sanctuary Manager, and the Regional Ecologist, David Roshier, and approved by John Kanowski, National Science and Conservation Manager. At Newhaven Sanctuary we liaised with Darcy Ginty, Land Management Officer, Newhaven Sanctuary. For sampling at Central Mt Wedge, we applied for an Aboriginal Land Special Purposes Permit (Central Land Council), approved by Jeff Hulcombe, the Ranger Coordinator for the Anangu Luritjiku Rangers. Jeff and a group of rangers also chaperoned and guided us during sampling. At Napperby Station, Roy and Janet Chisholm granted permission to sample on private land. Special thanks to Emma Matthews for figure design, Alana Delaine for imaging, and Steven Stringer for arrangement of Fig. 1.

References

Balke, M., Watts, C. H. S., Cooper, S. J. B., Humphreys, W. F., and Vogler, A. P. (2004). A highly modified stygobitic diving beetle of the genus *Copelatus* (Coleoptera, Dytiscidae): taxonomy and cladistic analysis based on mtDNA sequences. *Systematic Entomology* **29**, 59–67. doi:10.1111/j.1365-3113.2004.00229.x

Bayly, I. A. E., and Williams, W. D. (1966). Chemical and biological studies on some saline lakes of south-east Australia. *Marine and Freshwater Research* **17**, 177–228. doi:10.1071/MF9660177

Bohonak, A. J., and Jenkins, D. G. (2003). Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters* **6**, 783–796. doi:10.1046/j.1461-0248.2003.00486.x

Brock, M. A. (Ed.) (1986). 'Adaptation to Fluctuations Rather than to Extremes of Environmental Parameters.' Limnology in Australia, Monographiae Biologicae. (Springer: Dordrecht.)

Byrne, M., Yeates, D. K., Joseph, L., Kearny, M., Bowler, J., Williams, M. A. J., Cooper, S., Donnellan, S. C., Keogh, J. S., Leys, R., Melville, J., Murphy, D. J., Porch, N., and Wyroll, K.-H. (2008). Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology* **17**, 4398–4417. doi:10.1111/j.1365-294X.2008.03899.x

Chilton, C. (1920). On a new isopodan genus (Family Oniscidae) from Lake Corangamite, Victoria. *Proceedings of the Linnean Society of New South Wales* **44**, 723–734.

Cognetti, G., and Maltagliati, F. (2000). Biodiversity and adaptive mechanisms in brackish water fauna. *Marine Pollution Bulletin* **40**, 7–14. doi:10.1016/S0025-326X(99)00173-3

Cooper, S. J. B., Bradbury, J. H., Saint, K. M., Leys, R., Austin, A. D., and Humphreys, W. F. (2007). Subterranean archipelago in the Australian arid zone: mitochondrial DNA phylogeography of amphipods from central Western Australia. *Molecular Ecology* **16**, 1533–1544. doi:10.1111/j.1365-294X.2007.03261.x

Cooper, S. J. B., Saint, K. M., Taiti, S., Austin, A. D., and Humphreys, W. F. (2008). Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic isopods (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics* **22**, 195–203. doi:10.1071/IS07039

Davis, J., Pavlova, A., Thompson, R., and Sunnucks, P. (2013). Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Global Change Biology* **19**, 1970–1984. doi:10.1111/gcb.12203

de Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. In 'Endless Forms: Species and Speciation'. (Eds D. J. Howard, and S. H. Berlocher.) pp. 57–75. (Oxford University Press: Oxford.)

de Queiroz, K. (2005). Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6600–6607. doi:10.1073/pnas.0502030102

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* **56**, 879–886. doi:10.1080/10635150701701083

Edgar, R. C. (2004*a*). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113. doi:10.1186/1471-2105-5-113

Edgar, R. C. (2004*b*). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797. doi:10.1093/nar/gkh340

English, P., Spooner, N. A., Chappell, J., Questiaux, D. G., and Hill, N. G. (2001). Lake Lewis basin, central Australia: environmental evolution and OSL chronology. *Quaternary International* **83–85**, 81–101. doi:10.1016/S1040-6182(01)00032-5

Faille, A., Tänzler, R., and Toussaint, E. F. A. (2015). On the way to speciation: shedding light on the karstic phylogeography of the microendemic cave beetle *Aphaenops cerberus* in the Pyrenees. *The Journal of Heredity* **106**, 692–699.

Fensham, R. J., and Fairfax, R. J. (2003). Spring wetlands of the Great Artesian Basin, Queensland, Australia. *Wetlands Ecology and Management* **11**, 343–362. doi:10.1023/B:WETL.0000005532.95598.e4

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit 1 from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.

Gibert, J., and Deharveng, L. (2002). Subterranean ecosystems: a truncated functional biodiversity. *A.I.B.S. Bulletin* **52**, 473–481.

Gotch, T.B., Adams, M., Murphy, N.P., and Austin, A.D. (2008). A molecular systematic overview of wolf spiders associated with Great Artesian Basin springs in South Australia: evolutionary affinities and an assessment of metapopulation structure in two species. *Invertebrate Systematics* **22**, 151–165. doi:10.1071/IS07045

Gouws, G., and Stewart, B. A. (2007). From genetic structure to wetland conservation: a freshwater isopod *Paramphisopus palustris* (Phreatoicidea: Amphisopidae) from the Swan Coastal Plain, Western Australia. *Hydrobiologia* **589**, 249–263. doi:10.1007/s10750-007-0742-2

Gouws, G., and Stewart, B. A. (2013). Molecular species boundaries in the phreatoicidean genus *Amphisopus* (Isopoda: Amphisopidae) and evidence for a new freshwater isopod species from Western Australia. *Invertebrate Systematics* **27**, 173–185. doi:10.1071/IS12043

Greenslade, P., (1985). Terrestrial invertebrates of the mound spring bores, creek beds and other habitats. *South Australia's mound springs*. Nature Conservation Society of South Australia Inc, Adelaide, pp.64–77.

Guzik, M. T., Abrams, K. M., Cooper, S. J. B., Humphreys, W. F., and Cho, J.-L. (2008). Phylogeography of the ancient Parabathynellidae (Crustacea: Bathynellacea) from the Yilgarn region of Western Australia. *Invertebrate Systematics* **22**, 205–216. doi:10.1071/IS07040

Guzik, M. T., Austin, A. D., Cooper, S. J. B., Harvey, M. S., Humphreys, W. F., Bradford, T., Eberhard, S. M., King, R. A., Leys, R., Muirhead, K. A., and Tomlinson, M. (2011). Is the Australian subterranean fauna

uniquely diverse? *Invertebrate Systematics* **24**, 407–418. doi:10.1071/IS10038

Guzik, M. T., Adams, M., Murphy, N. P., Cooper, S. J. B., and Austin, A. D. (2012). Desert springs: deep phylogeographic structure in an ancient endemic crustacean (*Phreatomerus latipes*). *PLoS One* **7**, e37642.

Harvey, M. S. (2002). Short-range endemism among the Australian fauna: some examples from non-marine environments. *Invertebrate Systematics* **16**, 555–570. doi:10.1071/IS02009

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160–174. doi:10.1007/BF02101694

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913. doi:10.1038/35016000

Huelsenbeck, J. P., and Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. doi:10.1093/bioinformatics/17.8.754

Humphreys, W. (1993). The significance of the subterranean fauna in biogeographical reconstruction: examples from Cape Range peninsula, Western Australia. *Records of the Western Australian Museum* **45**, 165–192.

Humphreys, W. F. (1994). The subterranean fauna of the Cape Range coastal plain, northwestern Australia. Report to the Australian Heritage Commission and the Western Australian Heritage Committee. 202 pp. Western Australian Museum, unpublished report.

Humphreys, W. F. (1999). Relict stygofaunas living in sea salt, karst and calcrete habitats in arid northwestern Australia contain many ancient lineages. In 'The Other 99%. The Conservation and Biodiversity of Invertebrates'. (Eds W. Ponder, and D. Lunney.) Vol. 2088, pp. 219–227. (Transactions of the Royal Society of New South Wales: Sydney.)

Humphreys, W. F. (2000). Relict faunas and their derivation. In 'Ecosystems of the World. Subterranean Ecosystems'. (Eds H. Wilkens, D. C. Culver and W. F. Humphreys.) Vol. 30, pp. 417–432. (Elsevier: Amsterdam.)

Humphreys, W. F. (2001). Groundwater calcrete aquifers in the Australian arid zone: the context to an unfolding plethora of stygal biodiversity. *Records of the Western Australian Museum* **64**, 63–83. doi:10.18195/issn.0313-122x.64.2001.063-083

Humphreys, W. F. (2006). Aquifers: the ultimate groundwater-dependent ecosystems. *Australian Journal of Botany* **54**, 115–132. doi:10.1071/BT04151

Humphreys, W. F. (2008). Rising from down under: developments in subterranean biodiversity in Australia from a groundwater fauna perspective. *Invertebrate Systematics* **22**, 85–101. doi:10.1071/IS07016

Humphreys, W. F. (2009). Hydrogeology and groundwater ecology: does each inform the other? *Hydrogeology* **17**, 5–21.

Humphreys, W. F. (2017). 'Australasian Subterranean Biogeography.' (CRC Press: Boca Raton, FL.)

IUCN (2012). 'IUCN Red List Categories and Criteria: Version 3.1.' (IUCN: Gland.)

Javidkar, M., Cooper, S. J. B., Humphreys, W. F., King, R. A., Judd, S., and Austin, A. D. (2018). Biogeographic history of subterranean isopods from groundwater calcrete islands in Western Australia. *Zoologica Scripta* **47**, 206–220. doi:10.1111/zsc.12265

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., and Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* **33**, 1630–1638.

Kekkonen, M., and Hebert, P. D. N. (2014). DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Molecular Ecology Resources* **14**, 706–715. doi:10.1111/1755-0998.12233

Keppel, G., Van Niel, K. P., Wardell-Johnson, G. W., Yates, C. J., Byrne, M., Mucina, L., Schut, A. G. T., Hopper, S. D., and Franklin, S. E. (2012). Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography* **21**, 393–404. doi:10.1111/j.1466-8238.2011.00686.x

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120. doi:10.1007/BF01731581

King, R. K. (2009). Two new genera and species of chiltoniid amphipods (Crustacea: Amphipoda: Talitroidea) from freshwater mound springs in South Australia. *Zootaxa* **2293**, 35–52.

King, R. A., Bradford, T., Austin, A. D., Humphreys, W. F., and Cooper, S. J. B. (2012). Divergent molecular lineages and not-so-cryptic species: the first descriptions of stygobitic chiltoniid amphipods (Talitroidea: Chiltoniidae) from Western Australia. *Journal of Crustacean Biology* **32**, 465–488. doi:10.1163/193724012X626566

Kodric-Brown, A., and Brown, J. H. (1993). Highly structured fish communities in Australian desert springs. *Ecology* **74**, 1847–1855. doi:10.2307/1939942

Lefébure, T., Douady, C. J., Gouy, M., and Gibert, J. (2006*a*). Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* **40**, 435–447. doi:10.1016/j.ympev.2006.03.014

Lefébure, T., Douady, C. J., Gouy, M., Trontelj, P., Briolay, J., and Gibert, J. (2006*b*). Phylogeography of a subterranean amphipod reveals cryptic diversity and dynamic evolution in extreme environments. *Molecular Ecology* **15**, 1797–1806. doi:10.1111/j.1365-294X.2006.02888.x

Leys, R., and Watts, C. H. S. (2008). Systematics and evolution of the Australian subterranean hydroporine diving beetles (Dytiscidae), with notes on *Carabhydrus*. *Invertebrate Systematics* **22**, 217–225. doi:10.1071/IS07034

Leys, R., Watts, C. H. S., Cooper, S. J. B., and Humphreys, W. F. (2003). Evolution of subterranean diving beetles (Coleoptera: Dytiscidae: Hydroporini, Bidessini) in the arid zone of Australia. *Evolution* **57**, 2819–2834.

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 'Proceedings of the Gateway Computing Environments Workshop (GCE), 2010, New Orleans, Louisiana'. pp. 1–8. (Institute of Electrical and Electronics Engineers (IEEE): Louisana)

Miralles, A., and Vences, M. (2013). New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS One* **8**, –p.e68242. doi:10.1371/journal.pone.0068242

Moritz, C., and Agudo, R. (2013). The future of species under climate change: resilience or decline? *Science* **341**, 504–508. doi:10.1126/science.1237190

Murphy, N. P., Adams, M., and Austin, A. D. (2009). Independent colonization and extensive cryptic speciation of freshwater amphipods in the isolated groundwater springs of Australia's Great Artesian Basin. *Molecular Ecology* **18**, 109–122.

Murphy, N. P., Guzik, M. T., and Worthington Wilmer, J. (2010). The influence of landscape on population structure of four invertebrates in groundwater springs. *Freshwater Biology* **55**, 2499–2509. doi:10.1111/j.1365-2427.2010.02479.x

Murphy, N. P., Breed, M. F., Guzik, M. T., Cooper, S. J. B., and Austin, A. D. (2012). Trapped in desert springs: phylogeography of Australian desert spring snails. *Journal of Biogeography* **39**, 1573–1582. doi:10.1111/j.1365-2699.2012.02725.x

Murphy, N. P., Adams, M., Guzik, M. T., and Austin, A. D. (2013). Extraordinary micro-endemism in Australian desert spring amphipods. *Molecular Phylogenetics and Evolution* **66**, 645–653. doi:10.1016/j.ympev.2012.10.013

Murphy, N. P., Guzik, M. T., Cooper, S. J., and Austin, A. D. (2015*a*). Desert spring refugia: museums of diversity or evolutionary cradles? *Zoologica Scripta* **44**, 693–701. doi:10.1111/zsc.12129

Murphy, N. P., King, R. A., and Delean, S. (2015*b*). Species, ESUs or populations? Delimiting and describing morphologically cryptic diversity in Australian desert spring amphipods. *Invertebrate Systematics* **29**, 457–467. doi:10.1071/IS14036

Paz, A., and Crawford, A. J. (2012). Molecular-based rapid inventories of sympatric diversity: a comparison of DNA barcode clustering methods applied to geography-based vs clade-based sampling of amphibians. *Journal of Biosciences* **37**, 887–896. doi:10.1007/s12038-012-9255-x

Ponder, W. F. (1995). Mound spring snails of the Australian Great Artesian Basin. In 'The Conservation Biology of Molluscs'. (Ed. E. A. Kay.) pp. 13–18. (IUCN: Gland.)

Ponder, W. F., Colgan, D. J., Terzis, T., Clark, S. A., and Miller, A. (1996). Three new morphologically and genetically determined species of hydrobiid gastropods from Dalhousie Springs, northern South Australia, with the description of a new genus. *Molluscan Research* **17**, 49–109. doi:10.1080/13235818.1996.10673675

Posada, D., and Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. doi:10.1093/bioinformatics/14.9.817

Provan, J., and Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* **23**, 564–571. doi:10.1016/j.tree.2008.06.010

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012*a*). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology* **21**, 1864–1877. doi:10.1111/j.1365-294X.2011.05239.x

Puillandre, N., Modica, M. V., and Zhang, Y. (2012*b*). Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* **21**, 2671–2691. doi:10.1111/j.1365-294X.2012.05559.x

Rambaut, A., Drummond, A. J.,, Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**, 901–904. doi:10.1093/sysbio/syy032

Ratnasingham, S., and Hebert, P. D. N. (2013). A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS One* **8**, –p.e66213. doi:10.1371/journal.pone.0066213

Rivera, M. A., Howarth, F. G., Taiti, S., and Roderick, G. K. (2002). Evolution in Hawaiian cave-adapted isopods (Oniscidea: Philosciidae): vicariant speciation or adaptive shifts? *Molecular Phylogenetics and Evolution* **25**, 1–9. doi:10.1016/S1055-7903(02)00353-6

Sauer, J., and Hausdorf, B. (2012). A comparison of DNA-based methods for delimiting species in a Cretan land snail radiation reveals shortcomings of exclusively molecular taxonomy. *Cladistics* **28**, 300–316. doi:10.1111/j.1096-0031.2011.00382.x

Stamatakis, A. (2014). RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. doi:10.1093/bioinformatics/btu033

Taiti, S., and Humphreys, W. F. (2001). New aquatic Oniscidea (Crustacea, Isopoda) from groundwater calcretes of Western Australia. *Records of the Western Australian Museum* **64**, 133–151. doi:10.18195/issn.0313-122x.64.2001.133-151

Taiti, S., and Schotte, M. (Eds) (2016). *Haloniscus* Chilton, 1920. World Marine, Freshwater and Terrestrial Isopod Crustaceans Database (2008 onwards).

Taiti, S., and Xue, Z. (2012). The cavernicolous genus *Trogloniscus nomen novum*, with descriptions of four new species from southern China (Crustacea, Oniscidea, Styloniscidae). *Tropical Zoology* **25**, 183–209. doi:10.1080/03946975.2012.751240

Taiti, S., Ferrara, F., and Iliffe, T. M. (1995). A new species of *Haloniscus* Chilton, 1920 from New Caledonia (Isopoda, Oniscidea). *Crustaceana* **68**, 321–328. doi:10.1163/156854095X00502

Talavera, G., Dinč, V., and Vila, R. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution* **4**, 1101–1110. doi:10.1111/2041-210X.12107

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729. doi:10.1093/molbev/mst197

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* **17**, 57–86.

Trontelj, P., and Fišer, C. (2009). Cryptic species diversity should not be trivialised. *Systematics and Biodiversity* **7**, 1–3. doi:10.1017/S1477200008002909

Verovnik, R., Sket, B., and Trontelj, P. (2004). Phylogeography of subterranean and surface populations of water lice *Asellus aquaticus* (Crustacea: Isopoda). *Molecular Ecology* **13**, 1519–1532. doi:10.1111/j.1365-294X.2004.02171.x

Verovnik, R., Sket, B., and Trontelj, P. (2005). The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molecular Ecology* **14**, 4355–4369. doi:10.1111/j.1365-294X.2005.02745.x

Watts, C. H. S., and Humphreys, W. F. (2006). Twenty-six new Dytiscidae (Coleoptera) of the genera *Limbodessus* Guignot and *Nirripirti* Watts & Humphreys, from underground waters in Australia. *Transactions of the Royal Society of South Australia* **130**, 123–185. doi:10.1080/03721426.2006.10887055

Watts, C. H. S., and Humphreys, W. F. (2009). Fourteen new Dytiscidae (Coleoptera) of the genera *Limbodessus* Guignot, *Paroster* Sharp, and *Exocelina* Broun from underground waters in Australia. *Transactions of the Royal Society of South Australia* **133**, 62–107. doi:10.1080/03721426.2009.10887112

Wells, A. T., and Moss, F. J. (1983). The Ngalia Basin, Northern Territory: stratigraphy and structure. Bulletin 212. Bureau of Mineral Resources, Australia.

Whiting, M. F. (2002). Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta* **31**, 93–104. doi:10.1046/j.0300-3256.2001.00095.x

Whiting, M. F., Carpenter, J. C., Wheeler, Q. D., and Wheeler, W. C. (1997). The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* **46**, 1–68.

Williams, W. D. (1983). On the ecology of *Haloniscus searlei* (Isopoda, Oniscoidea), an inhabitant of Australian salt lakes. *Hydrobiologia* **105**, 137–142. doi:10.1007/BF00025183

Wilson, G. D. F., and Keable, S. J. (1999). A new genus of phreatoicidean isopod (Crustacea) from the north Kimberley region, Western Australia. *Zoological Journal of the Linnean Society* **126**, 51–79. doi:10.1111/j.1096-3642.1999.tb00607.x

Wilson, G. D., Humphrey, C. L., Colgan, D. J., Gray, K. A., and Johnson, R. N. (2009). Monsoon-influenced speciation patterns in a species flock of *Eophreatoicus* Nicholls (Isopoda; Crustacea). *Molecular Phylogenetics and Evolution* **51**, 349–364. doi:10.1016/j.ympev.2009.02.001

Zeidler, W. (1991). A new genus and species of phreatic amphipod (Crustacea: Amphipoda) belonging in the "Chiltonia" generic group, from Dalhousie Springs, South Australia. *Transactions of the Royal Society of South Australia* **115**, 177–187.

Zeidler, W. (1997). A new species of freshwater amphipod, *Austrochiltonia dalhousiensis* sp. nov. (Crustacea: Amphipoda: Hyalellidae) from Dalhousie Springs, South Australia. *Transactions of the Royal Society of South Australia* **121**, 29–42.

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29**, 2869–2876. doi:10.1093/bioinformatics/btt499

Handling editor: Shane Ahyong

SUPPLEMENTARY MATERIAL

Supplementary Material for Chapter 2

Supplementary File S1: Continued transcriptome editing and assembly methods (provided by TB and AZ).

For editing of raw transcript reads from the *Paraplatyarthrus* sp., *Pa. subterraneus* and *Po. pruinosus* samples, Tagcleaner v.0.12 (Schmieder, Lim, Rohwer, & Edwards, 2010) was used to trim SMARTer II adapters, and Trimmomatic v0.22 (Bolger, Lohse, & Usadel, 2012) was used to remove further adapters (e.g. Illumina sequencing adapters), together with long and short poly A and T tails, using the module, ILLUMINACLP. Reads <30 bp post-trimming were discarded, resulting in paired and unpaired fastq files. Transcripts were *de novo* assembled with Trinity v2012-06-18 (Grabherr et al., 2011; Haas et al., 2013) with default settings on a Dell PowerEdge R910 server using 512GB RAM.

For *Ceratothoa* sp. and *A. vulgare*, raw transcript reads were trimmed to remove low quality reads and adapters with Trimmomatic v0.32 (Bolger, Lohse, & Usadel, 2012) using ILLUMINACLIP. Trimmed reads were assembled with IDBA-Tran v1.1.1 (Peng et al., 2013) with --mink 20 --maxk 60 --step 5.

References

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *15*, 644–652. https://doi.org/10.1038/nbt.1883

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., … Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*, 1494–1512. https://doi.org/10.1038/nprot.2013.084

Peng, Y., Leung, H. C. M., Yiu, S-M., Lv, M-J., Zhu, X-G., & Chin, F. Y. L. (2013). IDBA-tran: a more robust de novo de Bruihn graph assembler for transcriptomes with uneven expression levels. *Bioinformatics*, *29*, i326–i334. https://doi.org/10.1093/bioinformatics/btt219

Schmieder, R., Lim, Y. W., Rohwer, F., & Edwards, R. (2010). TagCleaner: Identification and removal of tag sequences from genomic and metagenomics datasets. *BMC Bioinformatics*, *11*, 341. https://doi.org/10.1186/1471-2105-11-341

Supplementary File S2: Automated script for filtering reads, mapping reads to a reference, removing duplicate reads, and estimating coverage (or sequencing depth) during post-processing of the capture data.

```
#!/bin/bash
#usage: program.sh
#Automated exon capture data filtering and mapping
#Danielle Stringer
#November 2018

function error_exit
{
   # Exit function due to fatal error
   # Accepts 1 arg:
   # string - descriptive error message

   echo "${PROGNAME}: ${1:-"Unknown error"}" 1>&2
   exit 1
}

#------adjust these for your run-----
THREADS=10
#-----------------------------------------
for file in *_R1.fastq.gz
do
FILESTEM=${file%_*}

#run all paired data through bbmap
sh bbduk.sh in=$file in2=$FILESTEM"_R2.fastq.gz" out=../clean/$FILESTEM"_R1_clean.fq.gz"
out2=../clean/$FILESTEM"_R2_clean.fq.gz" outs=../clean/$FILESTEM"_singletons.fq.gz"
literal=AGATCGGAAGAGCAC,AGATCGGAAGAGCGT ktrim=r k=15 mink=15 hdist=0 tbo qtrim=rl
trimq=20 minlength=30 threads=$THREADS || error_exit "$LINENO: Error cleaning R1 or R2"

#run cleaned data through fastqc
mkdir -p ../clean/fastqc
fastqc --noextract --threads $THREADS -o ../clean/fastqc ../clean/$FILESTEM"_R1_clean.fq.gz"
../clean/$FILESTEM"_R2_clean.fq.gz" ../clean/$FILESTEM"_singletons.fq.gz"

#collapse reads with PEAR
pear -f ../clean/$FILESTEM"_R1_clean.fq.gz" -r ../clean/$FILESTEM"_R2_clean.fq.gz" -o
../clean/$FILESTEM

#zip output
gzip ../clean/$FILESTEM".assembled.fastq" ../clean/$FILESTEM".unassembled.forward.fastq"
../clean/$FILESTEM".unassembled.reverse.fastq"

#index reference
#if <*.amb> file exists move on
bwa index ../refs/469HaloniscusGenes_Fixed6_concat.fasta

#map collapsed, unassembled and singletons to reference
mkdir -p ../clean/files
```

```
bwa mem -A 1 -B 2 -t $THREADS -R "@RG\tID:Collapsed\tSM:"$FILESTEM
../refs/469HaloniscusGenes_Fixed6_concat.fasta <(zcat ../clean/$FILESTEM".assembled.fastq.gz") |
samtools view -hu -q 1 -@ $THREADS - | samtools sort -o
../clean/files/$FILESTEM"_collapsed.sorted.bam" -T temp.sort -@ $THREADS - || error_exit
"$LINENO: Error mapping collapsed"

bwa mem -A 1 -B 2 -t $THREADS -R "@RG\tID:PE\tSM:"$FILESTEM
../refs/469HaloniscusGenes_Fixed6_concat.fasta <(zcat
../clean/$FILESTEM".unassembled.forward.fastq.gz") <(zcat
../clean/$FILESTEM".unassembled.reverse.fastq.gz") | samtools view -hu -q 1 -@ $THREADS - |
samtools sort -o ../clean/files/$FILESTEM"_PE.sorted.bam" -T temp.sort -@ $THREADS - || error_exit
"$LINENO: Error mapping PE"

bwa mem -A 1 -B 2 -t $THREADS -R "@RG\tID:Singleton\tSM:"$FILESTEM
../refs/469HaloniscusGenes_Fixed6_concat.fasta <(zcat ../clean/$FILESTEM"_singletons.fq.gz") |
samtools view -hu -q 1 -@ $THREADS - | samtools sort -o
../clean/files/$FILESTEM"_singletons.sorted.bam" -T temp.sort -@ $THREADS - || error_exit
"$LINENO: Error mapping singletons"

#merge collapsed, unassembled and singletons sorted.bam files together
samtools merge ../clean/files/$FILESTEM"_merged.bam"
../clean/files/$FILESTEM"_collapsed.sorted.bam" ../clean/files/$FILESTEM"_PE.sorted.bam"
../clean/files/$FILESTEM"_singletons.sorted.bam"

#remove duplicates
java -jar $EBROOTPICARD/picard.jar MarkDuplicates  I=../clean/files/$FILESTEM"_merged.bam"
O=../clean/files/$FILESTEM"_merged_nodup.bam" AS=TRUE M=/dev/null
REMOVE_DUPLICATES=TRUE VALIDATION_STRINGENCY=LENIENT

#index PEandSingletons.bam
samtools index ../clean/files/$FILESTEM"_merged_nodup.bam"

#estimate coverage
mkdir -p ../clean/coverage
bedtools coverage -a ../refs/469HaloniscusGenes_Fixed6_concat.bed -b
../clean/files/$FILESTEM"_merged_nodup.bam" > ../clean/coverage/$FILESTEM"_coverage.txt"

bedtools coverage -a ../refs/469HaloniscusGenes_Fixed6_concat.bed -b
../clean/files/$FILESTEM"_merged_nodup.bam" -d > ../clean/coverage/$FILESTEM"_bases.txt"
```

done

Supplementary File S3: Automated script for variant calling, variant filtering, and creating consensus sequences during post-processing of the capture data.

```
#!/bin/bash
#usage: program.sh
#Automated variant calling, filtering and consensus sequence production
#Danielle Stringer
#November 2018
#For included Perl scripts, see https://bitbucket.org/tbertozzi/scripts/src/master/

function error_exit
{
   # Exit function due to fatal error
   # Accepts 1 arg:
   # string - descriptive error message

   echo "${PROGNAME}: ${1:-"Unknown error"}" 1>&2
   exit 1
}

#------adjust these for your run-----
THREADS=10
#-----------------------------------------
for file in *_merged_nodup.bam
do
FILESTEM=${file%_merged*}

#variant calling using freebayes
freebayes --min-base-quality 20 --min-alternate-fraction 0.2 -f
../refs/469HaloniscusGenes_Fixed6_concat.fasta $file > ../clean/variants/$FILESTEM"_freebayes.vcf"

#separate complexes/mnps into single snps using vcflib
vcfallelicprimitives -kg ../clean/variants/$FILESTEM"_freebayes.vcf" >
../clean/variants/$FILESTEM"_vcflib_primitives.vcf"

#intersect file using bedtools
intersectBed -header -a ../clean/variants/$FILESTEM"_vcflib_primitives.vcf" -b
../refs/469HaloniscusGenes_Fixed6_concat.bed >
../clean/variants/$FILESTEM"_vcf_intersectBed.vcf"

#filter vcf for depth using vcflib
vcffilter -f "DP > 9" ../clean/variants/$FILESTEM"_vcf_intersectBed.vcf" >
../clean/variants/$FILESTEM"_vcffilter_depth.vcf"

#filter heterozygous alleles (perl script: filterVCF.pl)
perl filterVCF.pl --vcf ../clean/variants/$FILESTEM"_vcffilter_depth.vcf" --freq 0.2

# create consensus sequences (perl script: applyVariants.pl)
perl applyVariants_11iv18.pl --bed ../refs/469HaloniscusGenes_Fixed6_concat.bed --fasta
../refs/469HaloniscusGenes_Fixed6.fasta --vcf $FILESTEM"_vcffilter_depth_mod_hets.vcf" --cover
../clean/coverage/$FILESTEM"_bases.txt" --min 10

done
```

Supplementary File S4: R script for exon capture evaluation.

```
## Load libraries
library(tidyverse)
library(modeest)
library(lme4)
library(car)
library(emmeans)

##------------------------------------------------------------------------------------------------
## Read in bases data from all files
##------------------------------------------------------------------------------------------------

listFiles <- list.files(path = "", pattern = "*.txt", full.names = TRUE)
bases_allCombined <- NULL
for (i in listFiles) {
   df <- read.csv(i, sep = "\t", header = FALSE, stringsAsFactors = FALSE)
   bases_allCombined <- bind_rows(bases_allCombined, df)
}

## We then added variables to calculate the start and end positions of each exon, converting values
to be sequential starting at zero for each orthologue.
## We also set the order of samples, classifying specimens in groups for plotting, and specified the
different sequencing runs (1-3).

## See reduced example dataset for bases_allCombined on Figshare (doi:10.25909/5d3ba3b7b1b2d)
## Load example dataset
loadRDS("bases_allCombined_exampledataset.rds")

##------------------------------------------------------------------------------------------------
#' ### Number and position of exons
##------------------------------------------------------------------------------------------------

## How many exons per gene - use one specimen to determine as same for all
noExons.df <- bases_allCombined %>%
   filter(specimenID == "27810") %>%
   group_by(geneName) %>%
   summarize(no.exons.per.gene = n_distinct(endPos)) %>%
   arrange(match(geneName, as.character(unique(bases_allCombined$geneName))))
noExons.df <- as.data.frame(noExons.df)
noExons.df

## Calculate positions of of exons, also calculate length of exons
exonPos.df <- bases_allCombined %>%
   filter(specimenID == "27810") %>%
   group_by(geneName) %>%
   filter(!duplicated(endPos)) %>%
   select(-c(1, 4))
exonPos.df
exonPos.df$nextStartPos <- c(exonPos.df$startPos[-1], 0)
exonPos.df <- as.data.frame(exonPos.df)
exonPos.df$exonLength <- with(exonPos.df, endPos-startPos)
exonPos.df
```

```
##-------------------------------------------------------------------------------------------------
#' exonLengths for all exons with Nreads > 9 (coverage threshold used for inclusion in consensus
sequences) across individuals
##-------------------------------------------------------------------------------------------------

## First calculate maximum number of reads per gene/exon/specimen
## Then filter data for gene/exons with > 9 reads
## Lastly, calculate exonLength
exonPos.allExons.df <- bases_allCombined %>%
   group_by(geneName, exonNo, specimenID) %>%
   mutate(maxReads = max(noReadsMappedAtBase, na.rm = TRUE)) %>%
   ungroup() %>%
   group_by(geneName, specimenID) %>%
   filter(!duplicated(endPos)) %>%
   filter(maxReads > 9) %>%
   select(-c(1, 4)) %>%
   mutate(exonLength = endPos-startPos)
exonPos.allExons.df

## Calculate median exon length across all genes/exons with >9 reads for each run (summary
medians to use in plots below)
medianLength.run <- exonPos.allExons.df %>%
   ungroup() %>%
   #group_by(specID.run) %>%
   summarise(medianLength = median(exonLength), rangeMin = min(exonLength), rangeMax =
max(exonLength), n = n())
medianLength.run

## Total exons captured
noExons.df.readsGT9 <- exonPos.allExons.df %>%
   group_by(specimenID) %>%
   summarise(n = n())
noExons.df.readsGT9

## Minimum, median and maximum number of exons captured across specimens
min(noExons.df.readsGT9$n)
median(noExons.df.readsGT9$n)
max(noExons.df.readsGT9$n)

##-------------------------------------------------------------------------------------------------
#' ### Plot exon sequencing depth by gene and species (specimenID)
##-------------------------------------------------------------------------------------------------

## Example plot of coverage profiles
f2plotDataAB <- droplevels(subset(bases_allCombined, geneName == "EOG54MW8B" & specimenID
%in% unique(bases_allCombined$specimenID)[c(2:6, 22:26)]))

## Load libraries
library(ggplot2)
library(RColorBrewer)

my_cols <- brewer.pal(6, "Dark2")
intr.pos <- exonPos.df[exonPos.df$geneName == "EOG54MW8B", ]
```

```
max.coverage <- max(f2plotDataAB$noReadsMappedAtBase)
f2.plotA <- ggplot(data = subset(f2plotDataAB, specID.grp == 2)) +
  aes(startPos + baseNumberPos, noReadsMappedAtBase, col = specimenID) +
  geom_line(size = 0.6) + #geom_point(size=0, shape = ".") +
  guides(colour = guide_legend(override.aes = list(size = 0.2, linetype = 1))) +
  theme_bw(base_size = 18) + labs(x = "", y = "Sequencing depth") +
  geom_vline(data = intr.pos, aes(xintercept = endPos), col = "grey80", size = 0.55) +
  geom_vline(data = intr.pos, aes(xintercept = nextStartPos), col = "grey80", size = 0.55) +
  ylim(c(0, max.coverage)) +
  scale_size_manual(values = 0.2) + scale_colour_manual(values = my_cols[c(1,1,1,2,2)]) +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), axis.title.x =
element_text(size = 0),
      legend.title = element_blank(), legend.position = "bottom", legend.text =
element_blank(),#text(size = 14),
      legend.spacing.y = unit(0.25, 'cm'), legend.margin = margin(0, 0, 0, 0),
legend.box.margin=margin(0,0,0,0),
      legend.key.size = unit(0, "line"))
f2.plotA

## Iterate over genes, take median of number of reads for each specimen at each exon, then plot
distribution of medians for each exon
## Also below we calculate the median absolute deviations (mad) - a robust measure of dispersion to
use in calculating a coefficient of dispersion (which is a robust version of the coefficient of variation)
boxPlotsMed <- as.data.frame(bases_allCombined) %>%
  mutate(geneName = factor(geneName, levels = noExons.df$geneName)) %>%
  arrange(geneName) %>%
  group_by(geneName, exonNo, specimenID, specID.run) %>%
  summarize(medianReads = median(noReadsMappedAtBase), madReads =
mad(noReadsMappedAtBase))

##-------------------------------------------------------------------------------------------
#' ### Identify targeted exons
##-------------------------------------------------------------------------------------------

## Choose "best exon in each gene"
## For each gene take highest median in each run * exon and modal category across exons chooses
"targeted" exon
exonMode <- boxPlotsMed %>%
  group_by(geneName, exonNo, specID.run) %>%
  summarize(medianSpecID = median(medianReads)) %>%
  filter(medianSpecID > 0) %>%
  ungroup() %>%
  group_by(geneName, specID.run) %>%
  summarize(maxExonName = exonNo[which.max(medianSpecID)]) %>%
  ungroup() %>%
  group_by(geneName) %>%
  summarize(modalExon = mlv(maxExonName, method='mfv')[[1]])
exonMode

##-------------------------------------------------------------------------------------------
#' ### Extract 50 randomly-selected genes and summarise these
##-------------------------------------------------------------------------------------------
```

```
## From here summaries are by geneName and specID.run
## Create random sample index of geneName and use this to filter out the 50 randomly selected
genes
set.seed(562)
rand50genes <- sample(noExons.df$geneName, 50)
boxPlotsMed_target_rand50 <- boxPlotsMed_target %>%
   ungroup() %>%
   mutate(specID.run = factor(specID.run, levels = c("1", "2", "3"))) %>%
   filter(geneName %in% rand50genes)
boxPlotsMed_target_rand50

##------------------------------------------------------------------------------------------------
#' ### Analysis of differences in sequencing depth between runs
##------------------------------------------------------------------------------------------------

## Fit generalised linear mixed model to estimate average differences in median coverage between
runs.
## Gene and specimenID are random effects to account for average differences in coverage between
each factor.
## Model is fitted on scale of log link
## Model assumes variances are distributed as a negative binomial (an alternative to poisson where
the variation is overdispersed)

## Round medians up to remove effects of ties
boxPlotsMed_target_rand50$medianReads.rnd <-
round(boxPlotsMed_target_rand50$medianReads, 0)
## Fit model
summary(zmod <- glmer.nb(medianReads.rnd ~ specID.run + (1|geneName) + (1|specimenID), data
= boxPlotsMed_target_rand50))
Anova(zmod)

## Intra-class coefficients - variance explained by each level of random variance
## Shows that most variability is among specimen, rather than among genes (in keeping with plots
above showing specimenID variation)
## geneName
0.199/(0.199 + 0.626 + 0.2)
## specimenID
0.626/(0.199 + 0.626 + 0.2)

## Estimate marginal mean coverage values (and CIs) for each run - could plot these average values
emmeans(zmod, specs = "specID.run")

## Contrasts show that first two runs yield similar median coverage, but that run three is markedly
lower.
contrast(emmeans(zmod, specs = "specID.run"), method = "consec", reverse  = TRUE, type = "link")

## Same contrasts but transformed to ratios of effects
## Ratio of 1 indicates equivalence, whilst ratio of 10 says median coverage for third run is lower by a
factor of 10
contrast(emmeans(zmod, specs = "specID.run"), method = "consec", reverse  = TRUE, type =
"response")

## Confidence intervals on ratio of effects
```

```
confint(contrast(emmeans(zmod, specs = "specID.run"), method = "consec", reverse  = TRUE, type =
"response"))

## Alternative approach where planned comparisons are run 1 versus run 2, and the average of first
two runs versus run 3
(custContrasts <- contrast(estMeans, list(run1Vs2 = c(1, -1, 0), run12Vs3 = c(0.5, 0.5, -1)), type =
"response"))
confint(custContrasts)
```

Table S1: Taxon sampling for transcriptome sequencing as well as the number of individuals pooled for each sample.

Higher taxa	Family	Genus	Species	Sequencing ID	Country	Locality	# Individuals pooled
Oniscidea	Philosciidae	*Haloniscus*	sp.	Halonisc	Australia	Quandong Bore, Laverton Downs, Yilgarn, WA	9
Oniscidea	Paraplatyarthridae	*Paraplatyarthrus*	sp.	G1	Australia	Laverton Downs, Yilgarn, WA	12
Oniscidea	Paraplatyarthridae	*Paraplatyarthrus*	*subterraneus*	G2	Australia	Laverton Downs, Yilgarn, WA	7
Oniscidea	Porcellionidae	*Porcellionides*	*pruinosus*	G3	Australia	Mt Windarra, Laverton Downs, Yilgarn, WA	2
Oniscidea	Armadillidiidae	*Armadillidium*	*vulgare*	lib35249	Germany	Stuttgart	1
Cymothoida	Cymothoidae	*Ceratothoa*	sp.	lib37015	Croatia	Osor	1

Table S2: Taxon sampling for exon capture with detailed collection data. Note: the sample from Windimurra (Yilgarn, WA) consists of pooled DNA extracts from three individuals (BES identifiers specified).

Superfamily	Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18775	27809	Robb's Bore, Newhaven Sanctuary, Ngalia Basin, NT	-22.7314	131.0867	13/09/2015	W. Humphreys, S. Cooper & D. Stringer
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18774	27810	Homestead Bore, Newhaven Sanctuary, Ngalia Basin, NT	-22.7252	131.1664	14/09/2015	W. Humphreys, S. Cooper & D. Stringer
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18754	27817	Sullivan's Well, Napperby Station, Ngalia Basin, NT	-22.7361	132.4610	16/09/2015	W. Humphreys, S. Cooper & D. Stringer
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18773	1	Camel Bore, Newhaven Sanctuary, Ngalia Basin, NT	-22.9344	131.2397	14/09/2015	W. Humphreys, S. Cooper & D. Stringer
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18759.3	2	Rabbit Hole Well, Central Mt Wedge, Ngalia Basin, NT	-22.7178	132.3239	16/09/2015	W. Humphreys, S. Cooper & D. Stringer
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES6667.2	17	Gurner Bore, Newhaven Sanctuary, Ngalia Basin, NT	-22.7160	130.9842	14/06/2001	W. Humphreys & A. Russ
Oniscoidea	Philosciidae	*Haloniscus*	*rotundus*	GAB01433	28077	Kingfisher, Dalhousie Springs, SA	-26.4083	135.5216	7/07/2009	M. Guzik, R. King & L. Harsche
Oniscoidea	Philosciidae	*Haloniscus*	*yardiyaensis*	GAB01616	28078	Freeling South Springs, Mount Dennison, SA	-28.0733	135.9036	3/07/2009	M. Guzik, R. King & L. Harsche
Oniscoidea	Philosciidae	*Haloniscus*	*fontanus*	GAB00736	28079	Strangways Springs, SA	-29.1622	136.5517	1/11/2007	M. Guzik & N. Murphy
Oniscoidea	Philosciidae	*Haloniscus*	*fontanus*	GAB00795	11	Old Finnis Springs, Hermit Hills, SA	-29.5832	137.4408	4/11/2007	M. Guzik & N. Murphy
Oniscoidea	Philosciidae	*Haloniscus*	*fontanus*	GAB00765	12	Bubbler Spring, SA	-29.4464	136.8580	3/11/2007	M. Guzik & N. Murphy
Oniscoidea	Philosciidae	*Haloniscus*	*fontanus*	GAB01007.1	28082	Hawker Springs, Neales, SA	-28.4251	136.1861	27/08/2008	M. Guzik & N. Murphy
Oniscoidea	Philosciidae	*Haloniscus*	*microphthalmus*	GAB00764.1	28080, 28083	Francis Swamp Springs, SA	-29.0797	136.2769	3/11/2007	M. Guzik & N. Murphy
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18659	27813, 27814, 27815	Shady Well, Laverton Downs, Yilgarn, WA	-28.4074	122.2038	21/04/2015	W. Humphreys, S. Cooper & J. Hyde
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18601	27816	Quandong Bore, Laverton Downs, Yilgarn, WA	-28.3393	122.2097	20/04/2015	W. Humphreys, S. Cooper & J. Hyde
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18645	27812	Laverton South, Laverton Downs, Yilgarn, WA	-28.5161	122.1833	21/04/2015	W. Humphreys, S. Cooper & J. Hyde

Superfamily	Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES18644	27818	Laverton South, Laverton Downs, Yilgarn, WA	-28.5170	122.1813	21/04/2015	W. Humphreys, S. Cooper & J. Hyde
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES16434	28076	Lake Miranda East, Yilgarn, WA	-27.6792	120.6022	23/10/2011	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES17062	28081	Bubble Well, Millbillillie, Yilgarn, WA	-26.5607	120.0408	15/05/2012	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES16348	4	Bubble Well, Millbillillie, Yilgarn, WA	-26.5607	120.0409	21/10/2011	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES6655	3	Yuinmery South, Yilgarn, WA	-28.5486	119.0911	15/05/2001	W. Humphreys, C. Watts & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES8623.1	5	Three Rivers Plutonic, Yilgarn, WA	-25.2831	119.1757	26/08/2001	W. Humphreys, T. Karanovic & J. Waldock
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES13246	7	Mt Morgans, Yilgarn, WA	-28.7318	122.1569	10/05/2007	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES13396	8	Lake Uramurdah, Yilgarn, WA	-26.6877	120.3528	16/05/2007	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES13314	10	Lake Uramurdah, Yilgarn, WA	-26.6876	120.3027	16/05/2007	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES14385	9	Gum Well, Perrinvale, Yilgarn, WA	-28.7750	120.4170	8/05/2007	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES6601.2	15	Jundee South Hill, Yilgarn, WA	-26.2688	120.6809	11/05/2001	W. Humphreys, C. Watts & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES10410	16	Lake Mason, Yilgarn, WA	-27.5400	119.6243	30/05/2004	W. Humphreys, C. Watts & C. Clay
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES13452	18	Lake Violet, Yilgarn, WA	-26.6876	120.2866	16/05/2007	W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	sp.	BES8956, BES13133.1, BES13133.2	19	Windimurra, Yilgarn, WA	-28.2860	118.5754	31/08/2001, 24/10/2004	R. Leijs/W. Humphreys & S. Cooper
Oniscoidea	Philosciidae	*Haloniscus*	*searlei*	BES6573	27811	Lighthouse Swamp, Rottnest Island, WA	-32.0000	115.5000	1/04/2001	W. Humphreys

Superfamily	Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by
Outgroup: Oniscoidea	Philosciidae	*Haloniscus*	*anophthalmus*	BES10201	13	Isle of Pines, Grotte de Ouiudea, New Caledonia	-22.6000	167.4300	19/05/2003	F. Bzescia
Outgroup: Armadilloidea	Armadillidae	*Buddelundia*	sp.	B002	21	Andamooka, SA	-30.7564	137.1772	00/08/2016	M. Javidkar
Outgroup: Armadilloidea	Armadillidae	*Troglarmadillo*	sp.	BES15537.2	23	Lake Miranda West, Yilgarn, WA	-27.7467	120.5266	00/07/2010	W. Humphreys & S. Cooper
Outgroup: Oniscoidea	Paraplatyarthridae	*Paraplatyarthrus*	*subterraneus*	BES15525.10	22	Laverton South, Laverton Downs, Yilgarn, WA	-28.5028	122.1773	13/07/2010	W. Humphreys & S. Cooper
Outgroup: Oniscoidea	Paraplatyarthridae	*Paraplatyarthrus*	sp.	BES16400.2	6	Halfpenny Well, Yilgarn, WA	-27.6966	121.3395	21/10/2011	W. Humphreys & S. Cooper
Outgroup: Oniscoidea	Paraplatyarthridae	Gen.	sp.	Ja243	20	Porto Alegre, Belém Novo, Rio Grande do Sul, Brazil	-30.2086	-51.1697	19/03/2011	D. Kenne & I. Campos Filho

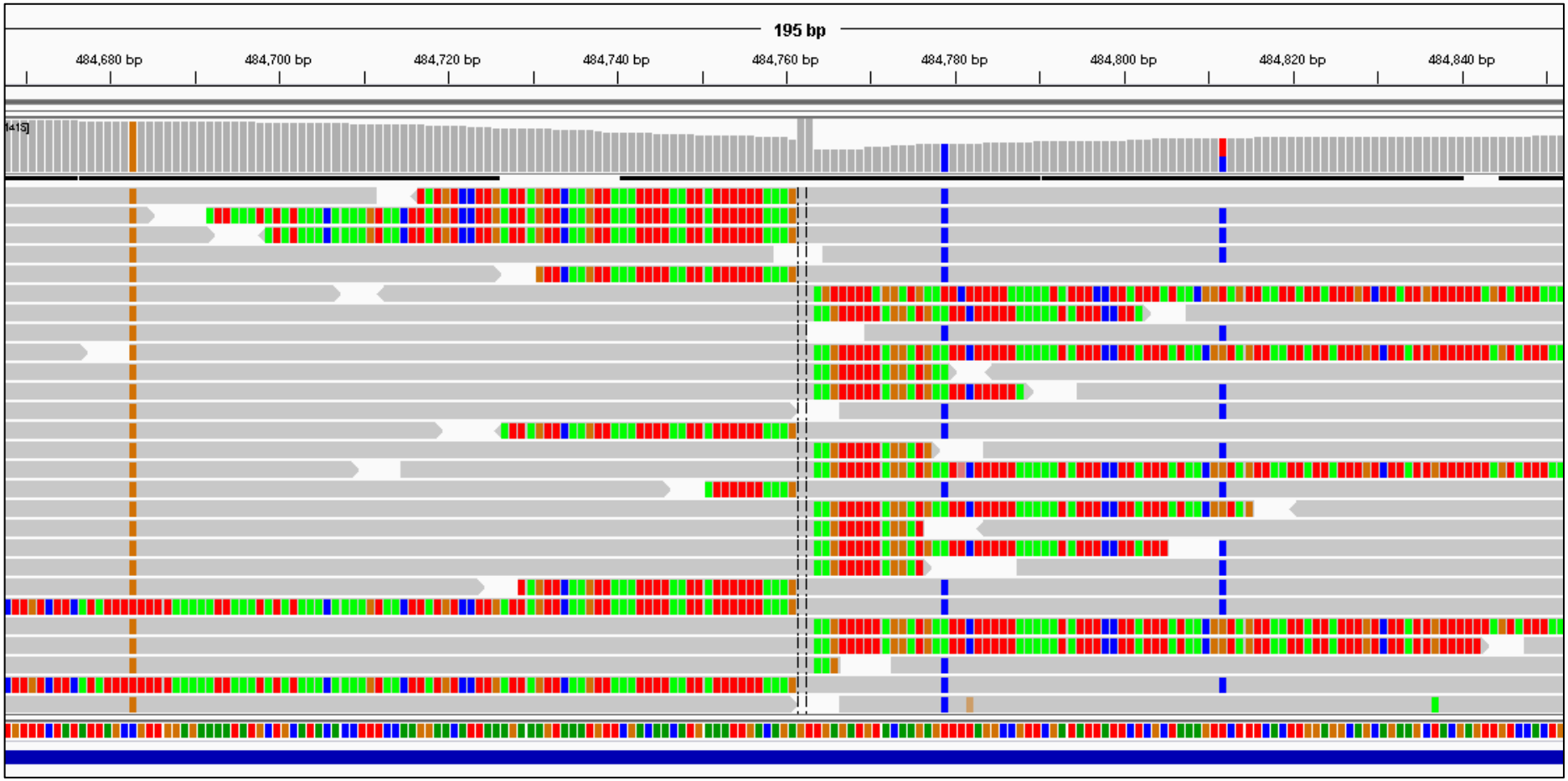


Figure S1. A short read alignment (in Integrative Genomics Viewer) highlighting the position of an intron-exon boundary. Each bar represents a single read. Bases matching the reference sequence are shown in grey, while soft-clipped (mismatched) bases are coloured. The *Haloniscus* reference sequence is indicted by the blue bar at the bottom of the figure.

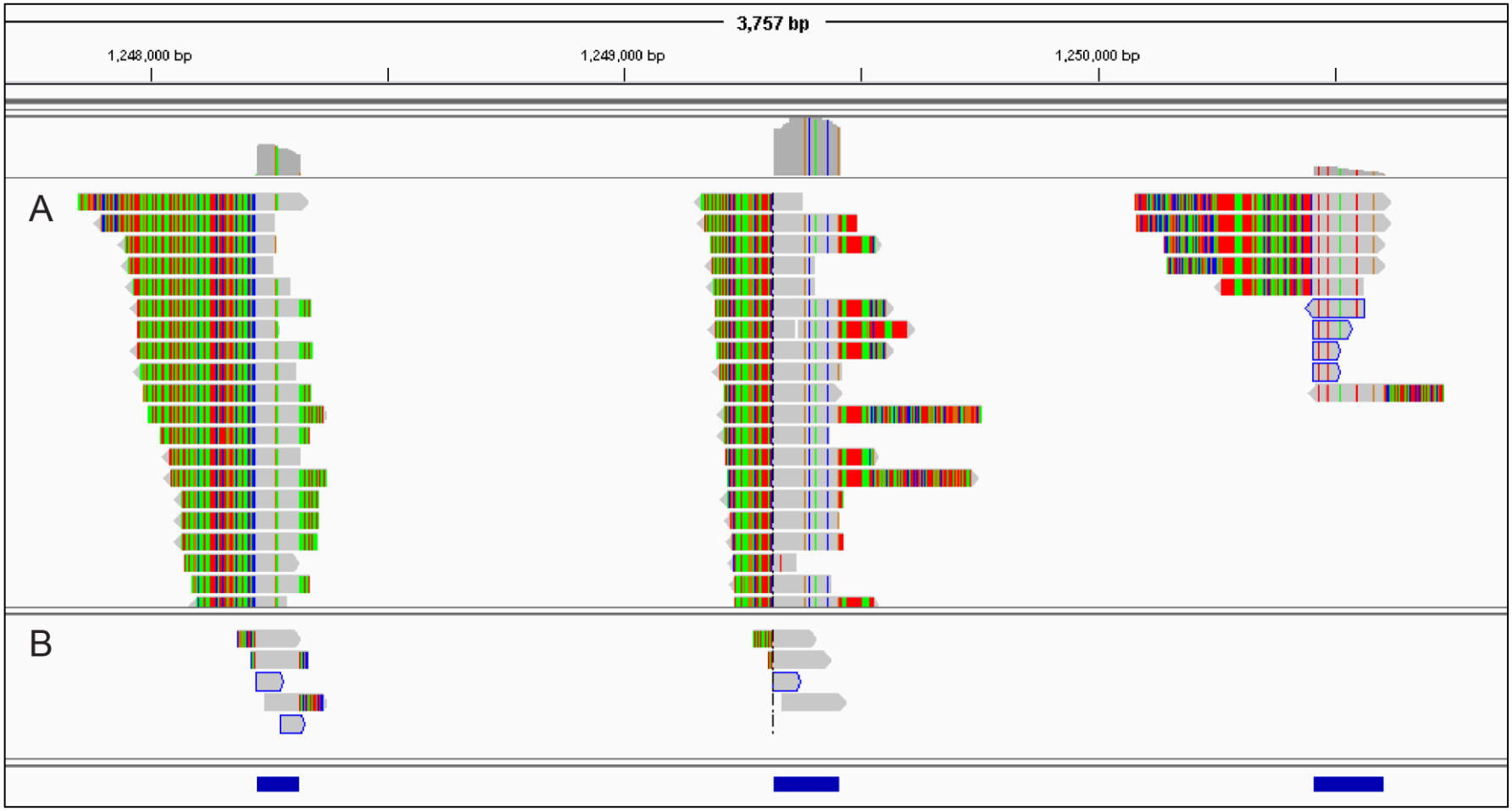


Figure S2. Alignments of short reads (in Integrative Genomics Viewer) to three putative exons (blue bars) after inferring intron-exon boundaries from the *Haloniscus* reference sequence used in bait design. (A) Bases matching the reference sequence are shown in grey, while mismatched bases representing introns and single nucleotide polymorphisms are coloured, and (B) position of baits. The reads and the reference are from different species, but both are from the *Haloniscus* genus.

Figure S3: Sequencing depth summaries for all 469 targeted orthologues and isopod samples. Each page shows one orthologue and the sequencing depth results for all taxa. Vertical lines indicate intron positions.

This figure is publically available on Figshare at doi:10.25909/5d3b90d1b424f

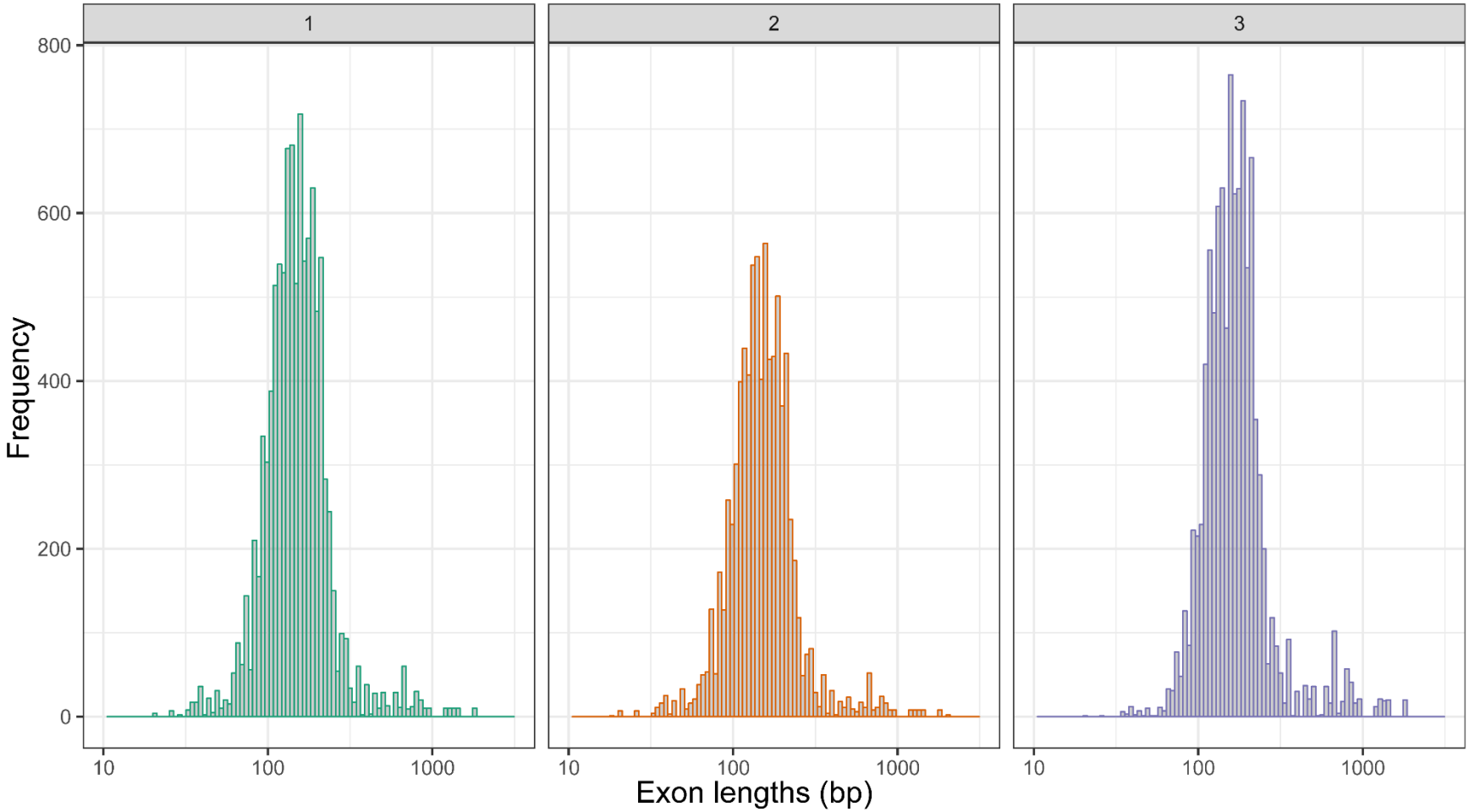
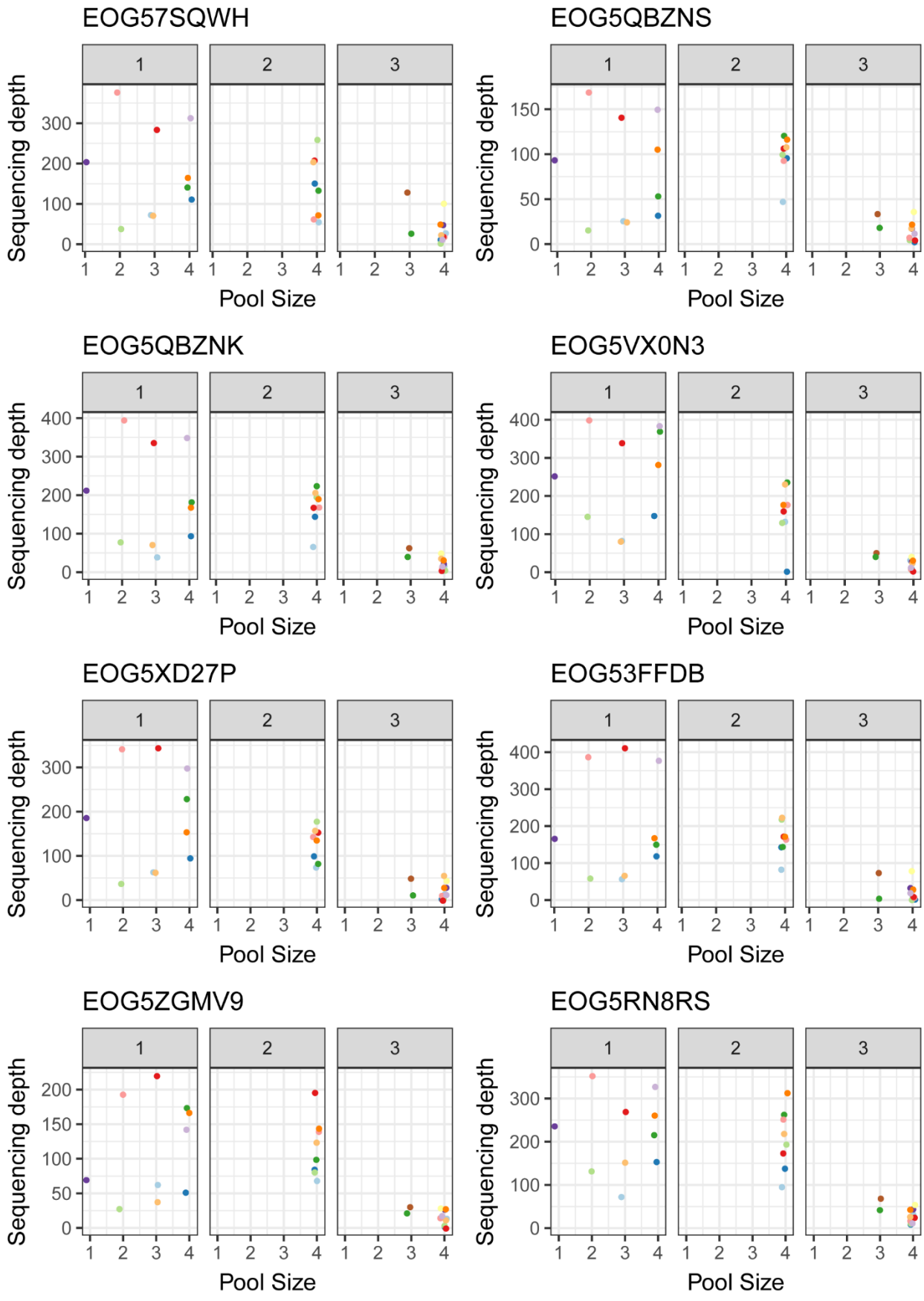


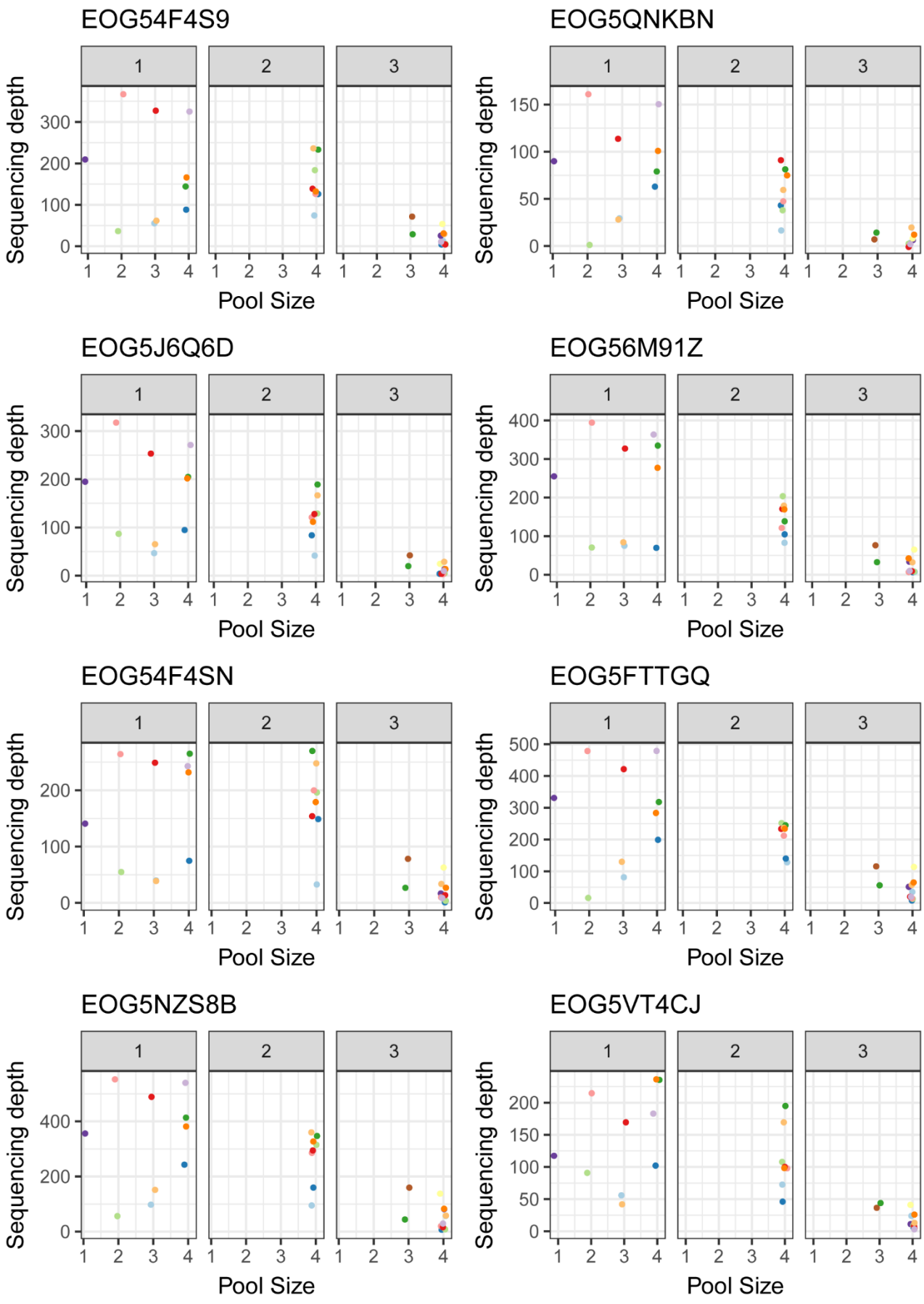
Figure S4: A frequency distribution for the length of exons (bp) captured across the three sequencing runs (run 1 is indicated in green, run 2 in orange, and run 3 in purple).

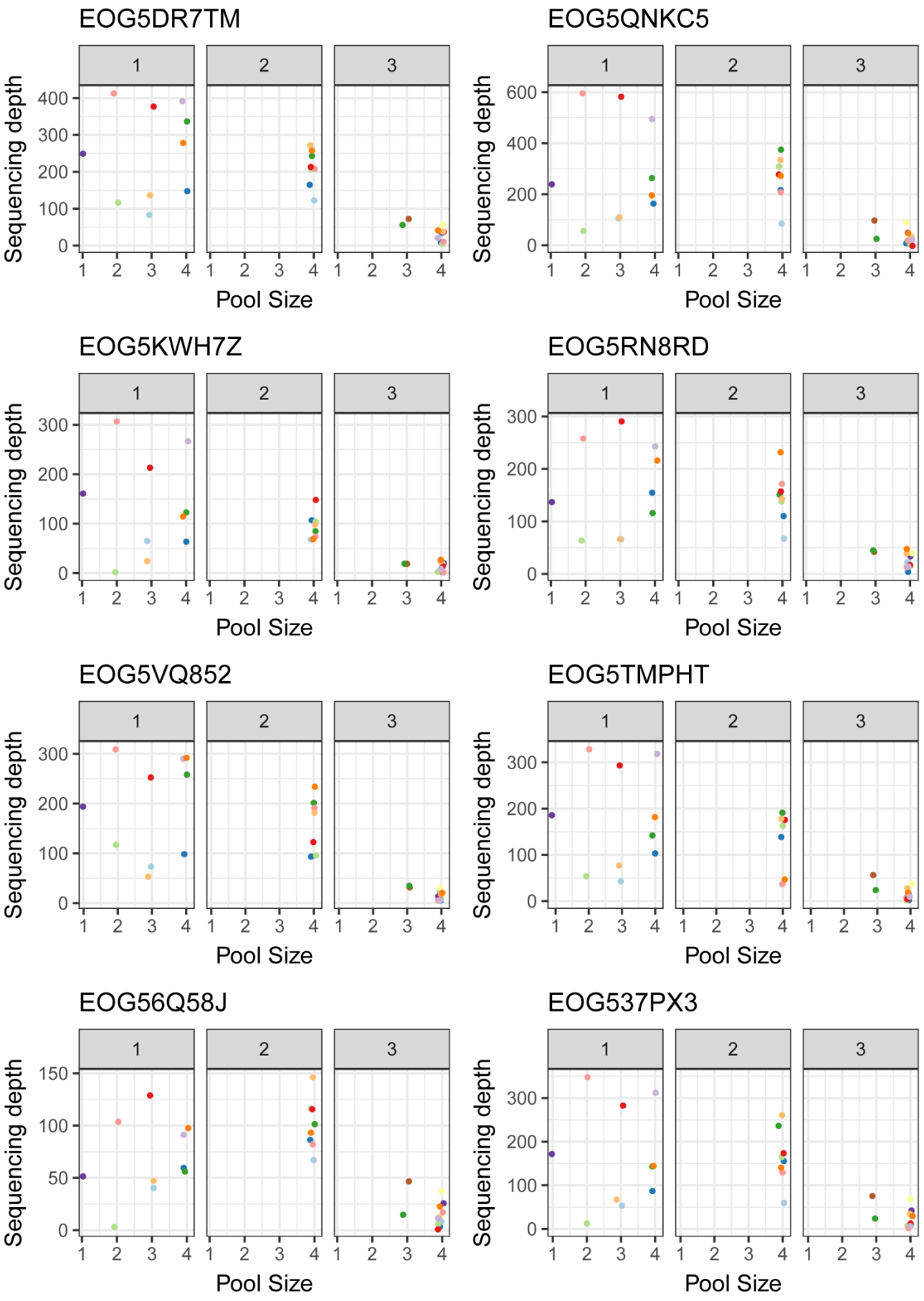
Figure S5: Distribution of sequencing depth for samples at each exon within all orthologues, grouped by the three sequencing runs (run 1 depicted in green, run 2 in orange, and run 3 in purple). Horizontal lines are median sequencing depths, vertical lines indicate boxplot whiskers, and solid points represent outliers.

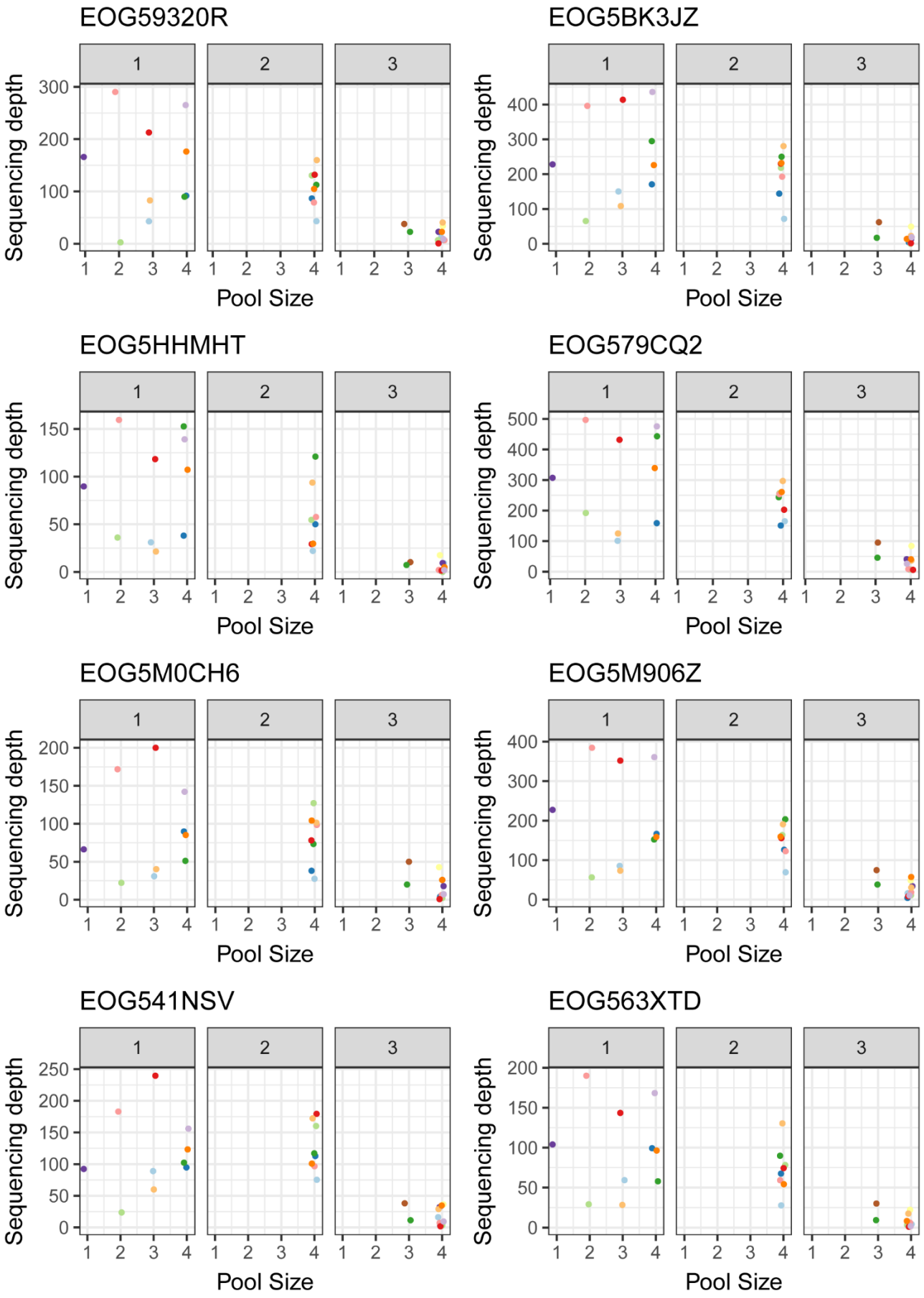
This figure is publically available on Figshare at doi:10.25909/5d3b94f9d4227

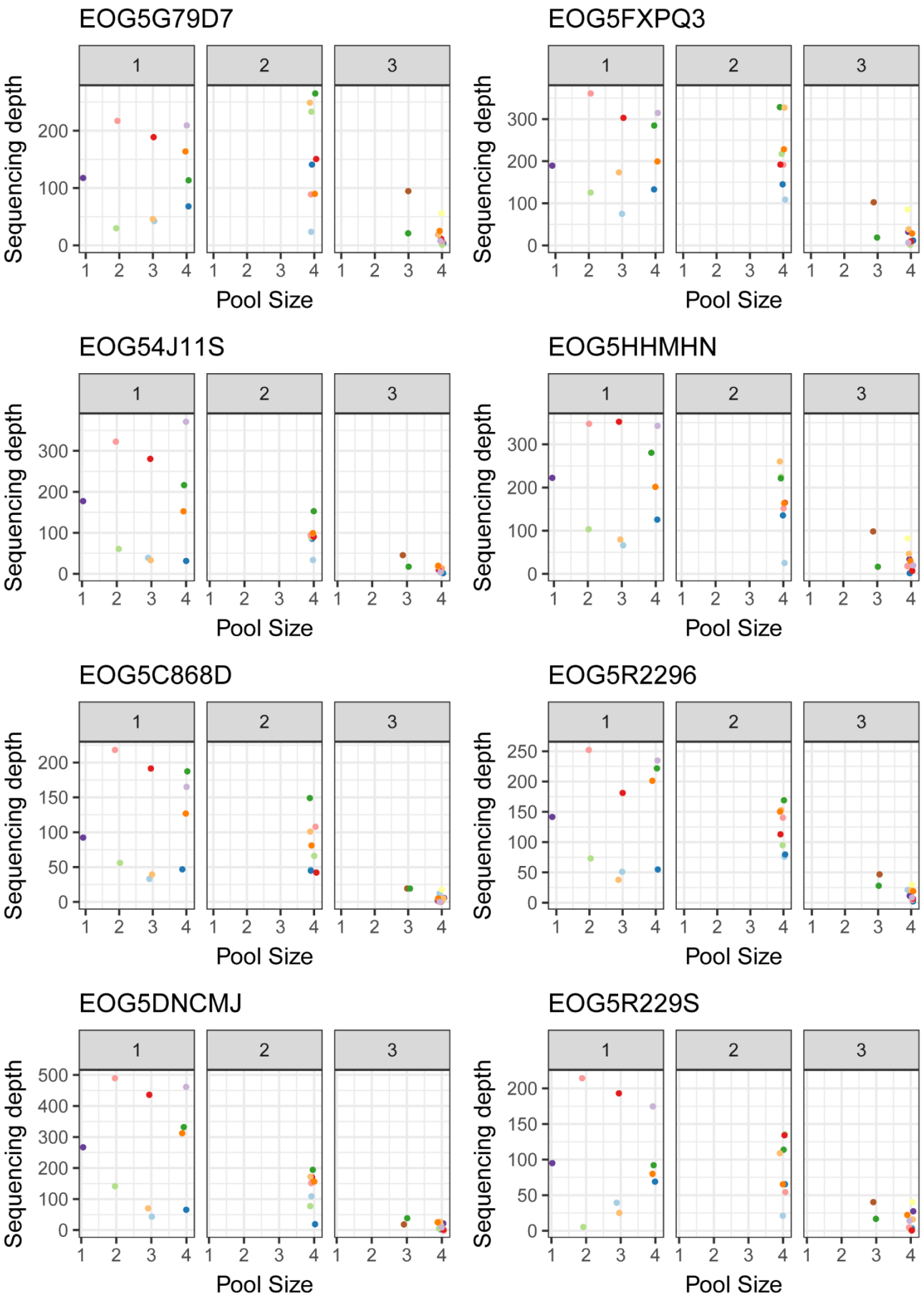
Figure S6: Plots of pooling sizes (1–4) prior to capture against median sequencing depth across samples for exons of 50 randomly targeted orthologues, separated by sequencing run (1–3). Orthologue code is specified above plots, and points are coloured according to specimen ID.

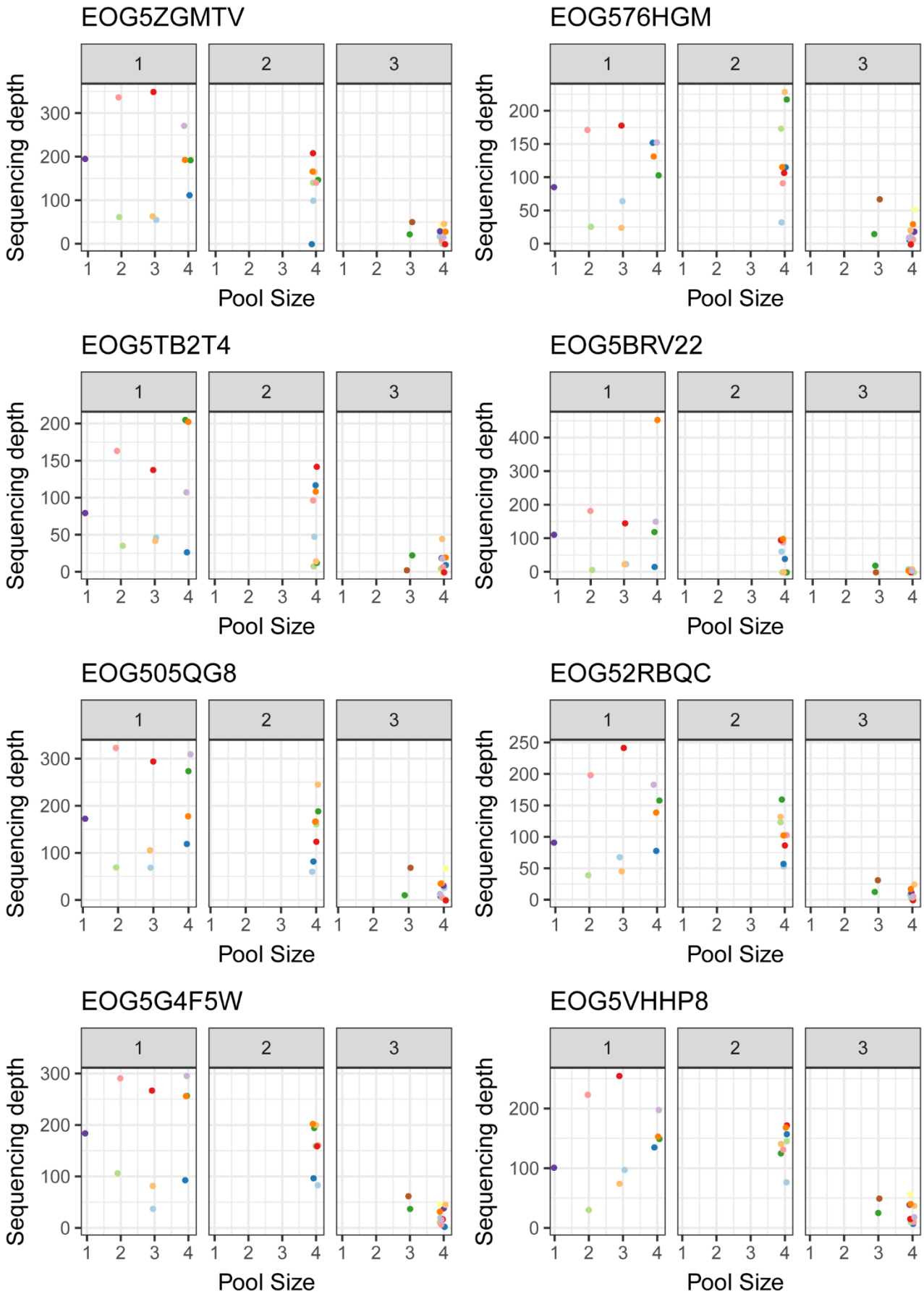



EOG54F4S9
EOG5QNKBN
EOG5J6Q6D
EOG56M91Z
EOG54F4SN
EOG5FTTGQ
EOG5NZS8B
EOG5VT4CJ
1
2
3
Sequencing depth
Pool Size


EOG5DR7TM
EOG5QNKC5
EOG5KWH7Z
EOG5RN8RD
EOG5VQ852
EOG5TMPHT
EOG56Q58J
EOG537PX3
1
2
3
Sequencing depth
Pool Size


EOG59320R
EOG5BK3JZ
EOG5HHMHT
EOG579CQ2
EOG5M0CH6
EOG5M906Z
EOG541NSV
EOG563XTD
1
2
3
Sequencing depth
Pool Size


EOG5G79D7
EOG5FXPQ3
EOG54J11S
EOG5HHMHN
EOG5C868D
EOG5R2296
EOG5DNCMJ
EOG5R229S
Sequencing depth
Pool Size
1
2
3


EOG5ZGMTV
EOG576HGM
EOG5TB2T4
EOG5BRV22
EOG505QG8
EOG52RBQC
EOG5G4F5W
EOG5VHHP8
Sequencing depth
Pool Size
1
2
3

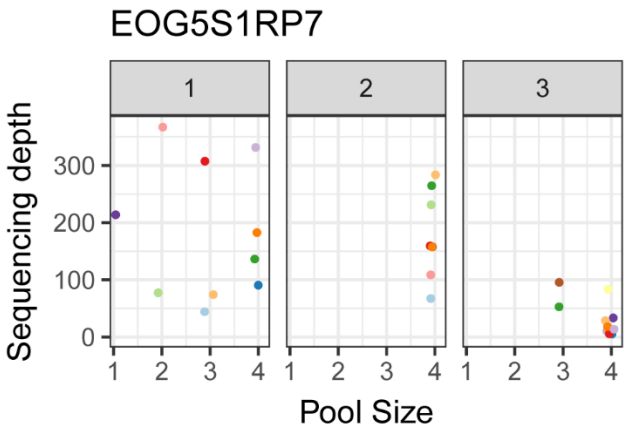
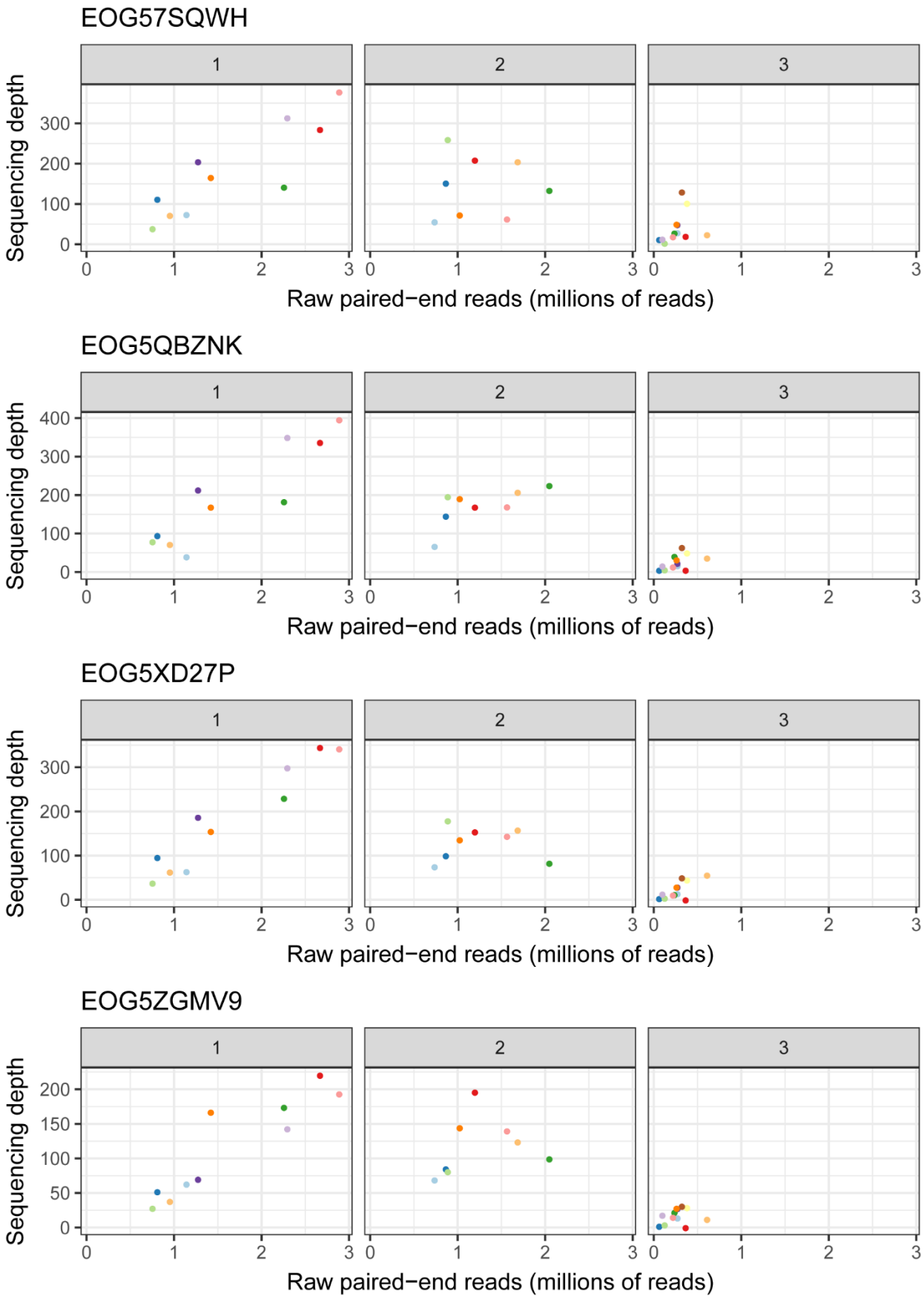
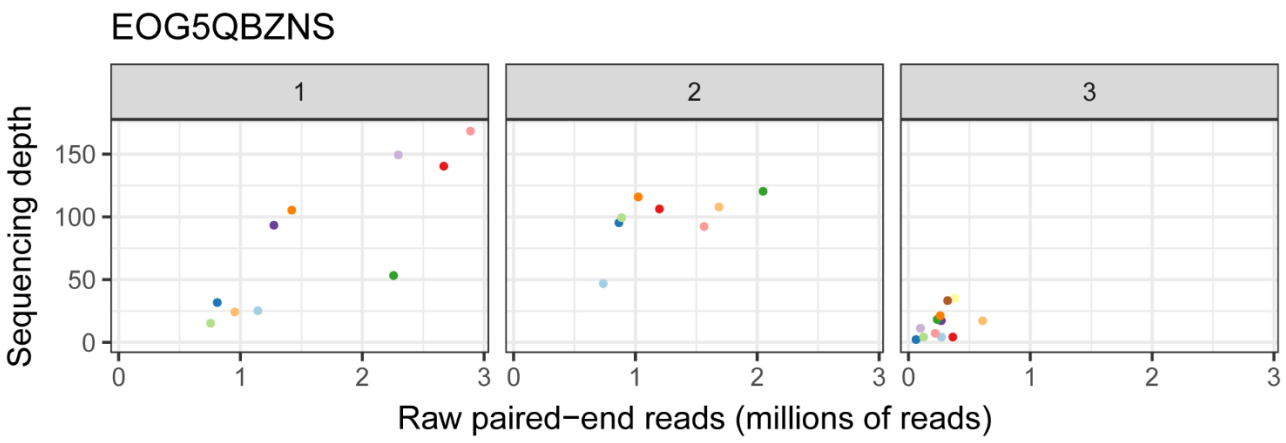
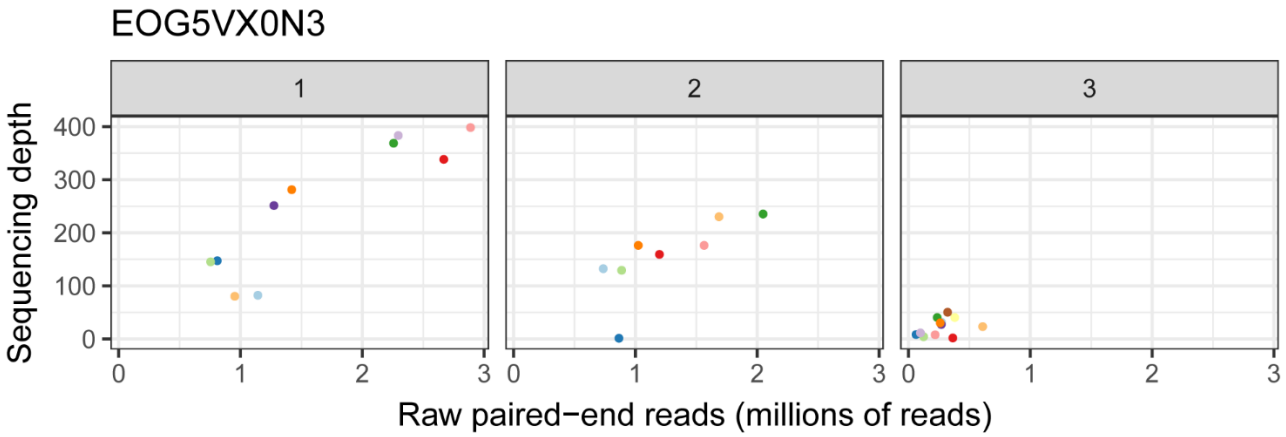
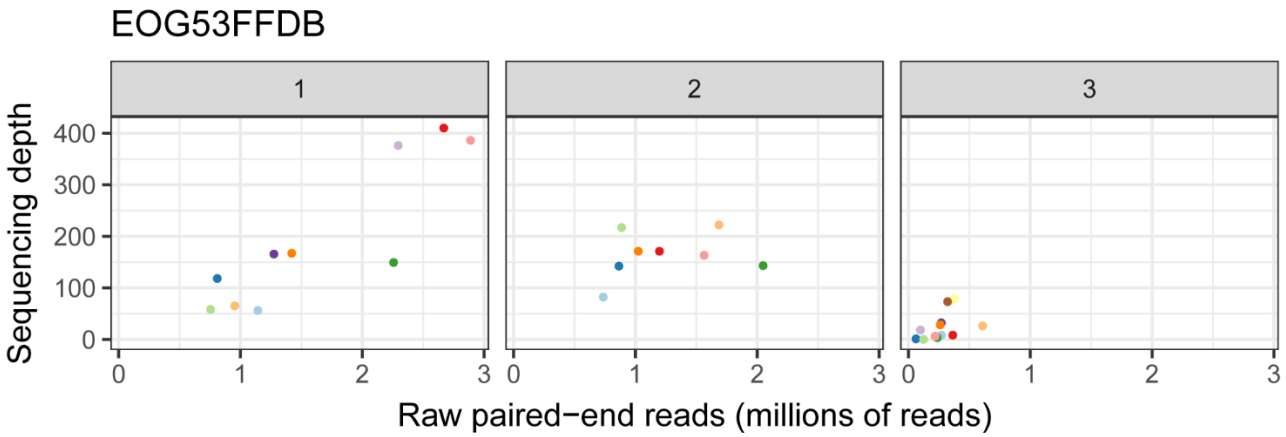

EOG5S1RP7
1
2
3
Sequencing depth
300
200
100
0
1 2 3 4 1 2 3 4 1 2 3 4
Pool Size

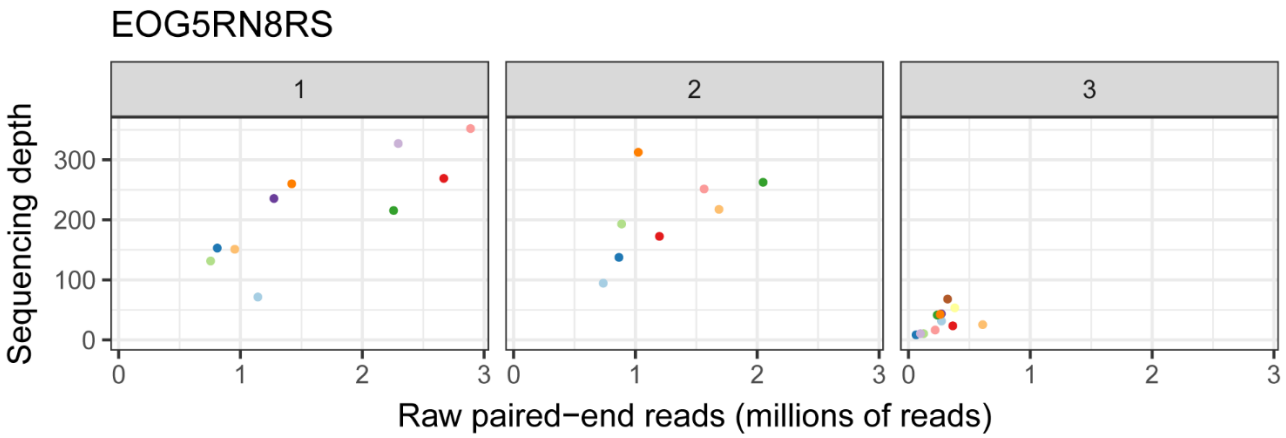
Figure S7: Plots of raw paired-end reads against median sequencing depth across samples for exons of 50 randomly targeted orthologues, separated by sequencing run (1–3). Orthologue code is specified above plots, and points are coloured according to specimen ID.

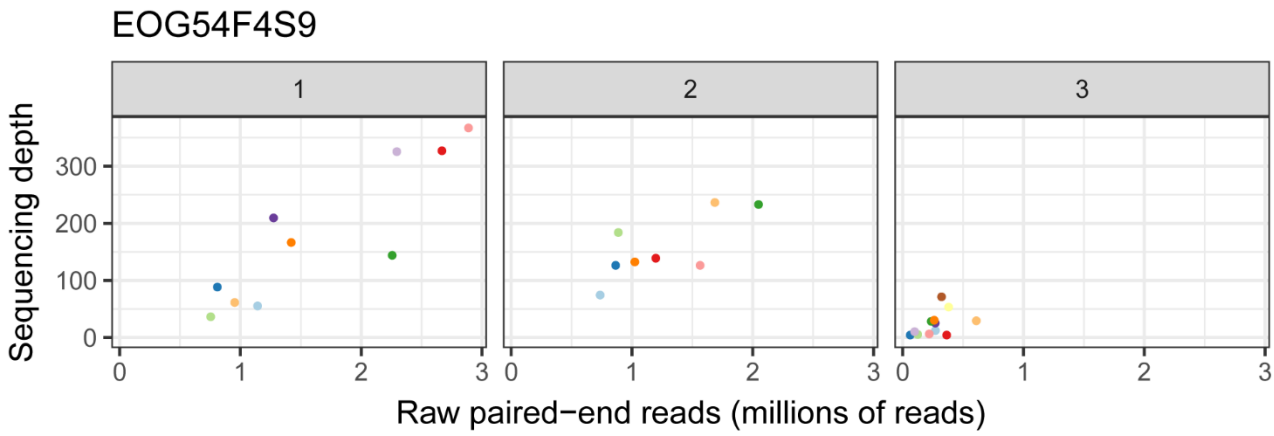


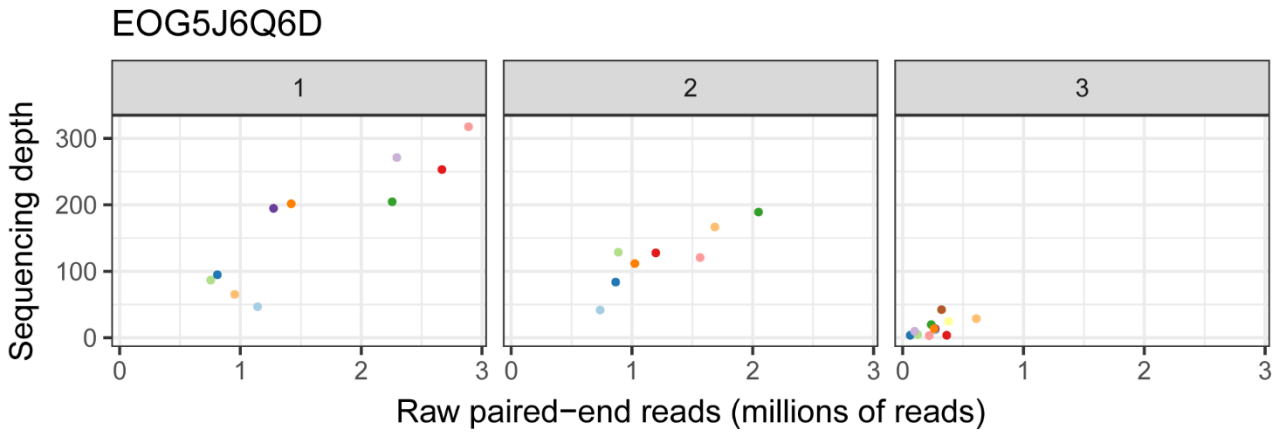

EOG5QBZNS
1
2
3
Sequencing depth
150
100
50
0
0
1
2
3
Raw paired−end reads (millions of reads)

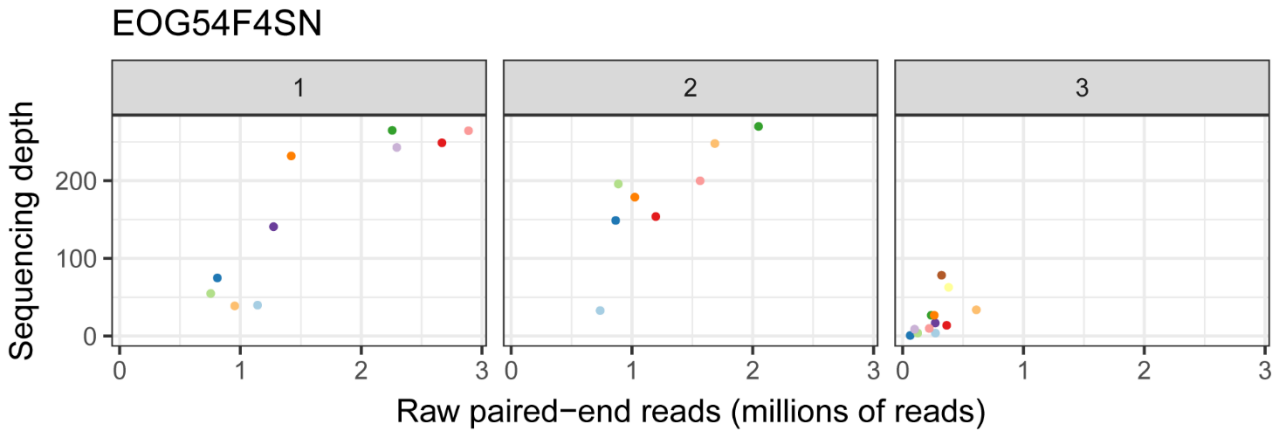

EOG5VX0N3
1
2
3
Sequencing depth
400
300
200
100
0
0
1
2
3
Raw paired−end reads (millions of reads)

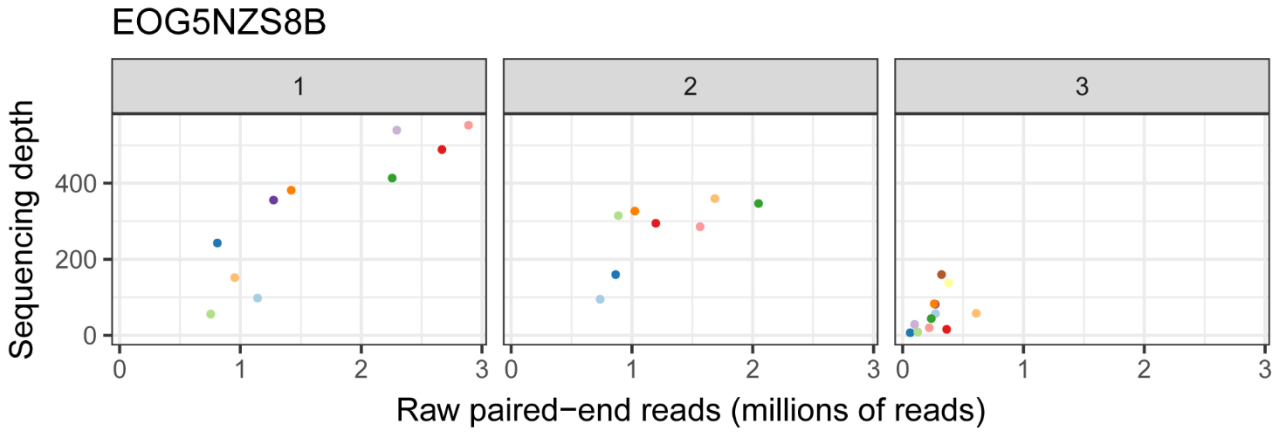

EOG53FFDB
1
2
3
Sequencing depth
400
300
200
100
0
0
1
2
3
Raw paired−end reads (millions of reads)

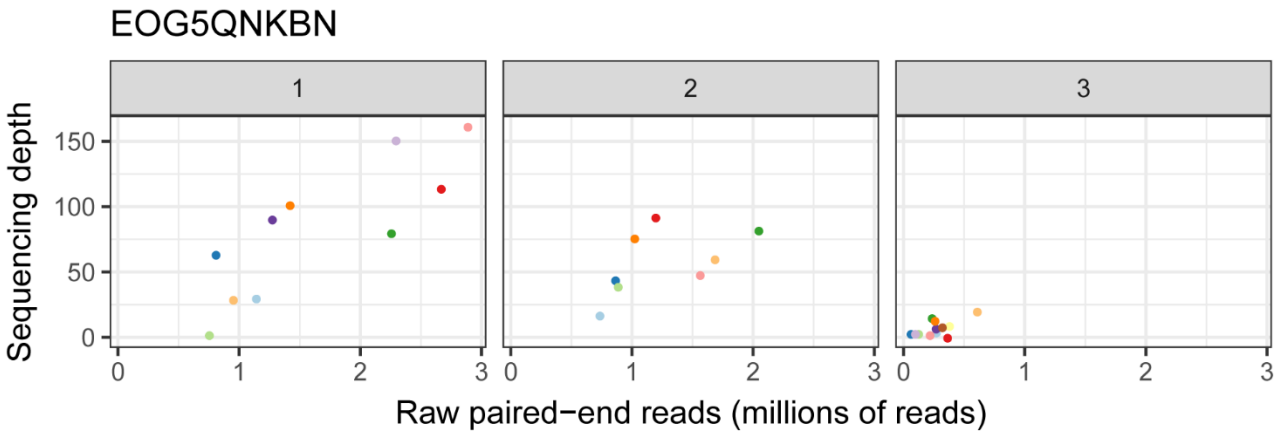

EOG5RN8RS
1
2
3
Sequencing depth
300
200
100
0
0
1
2
3
Raw paired−end reads (millions of reads)

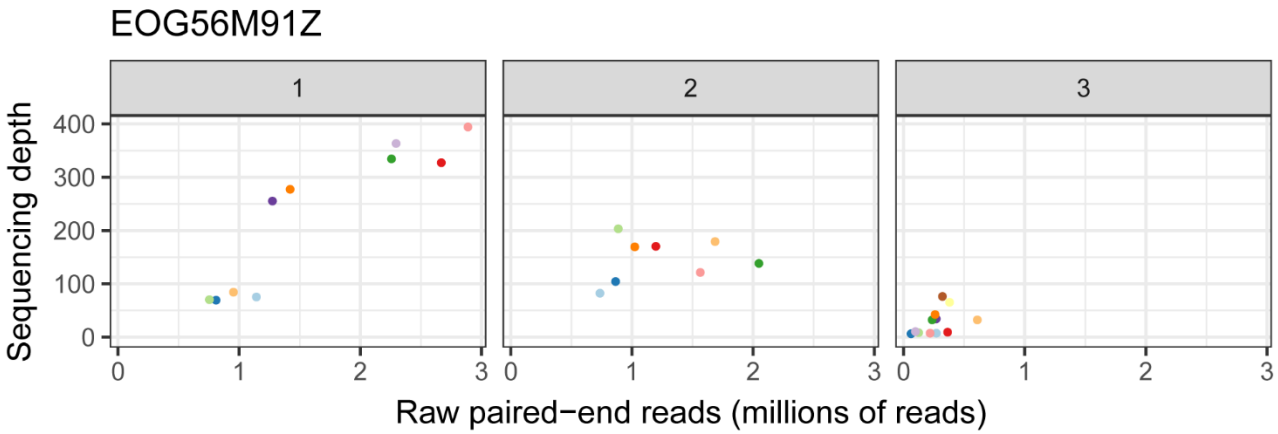

EOG54F4S9
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

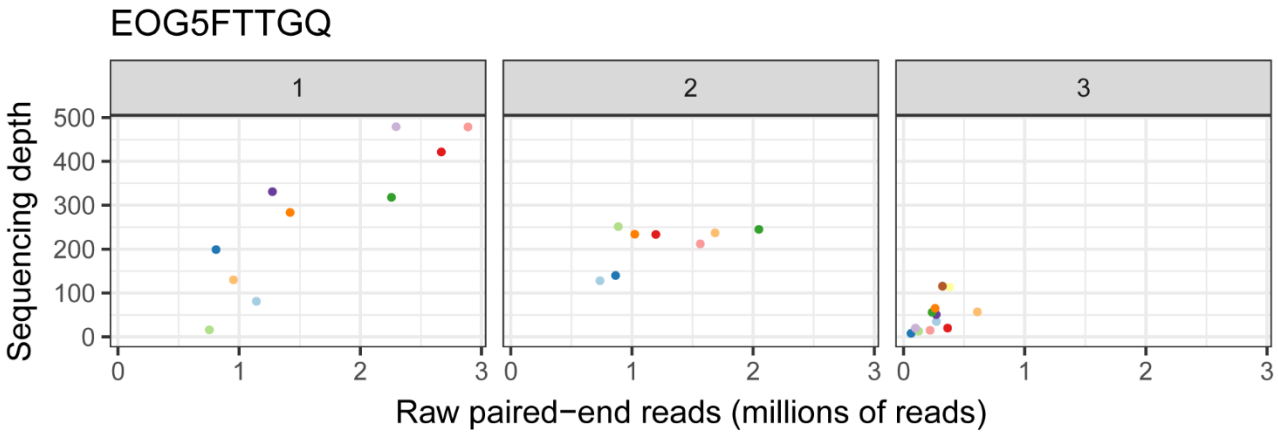

EOG5J6Q6D
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

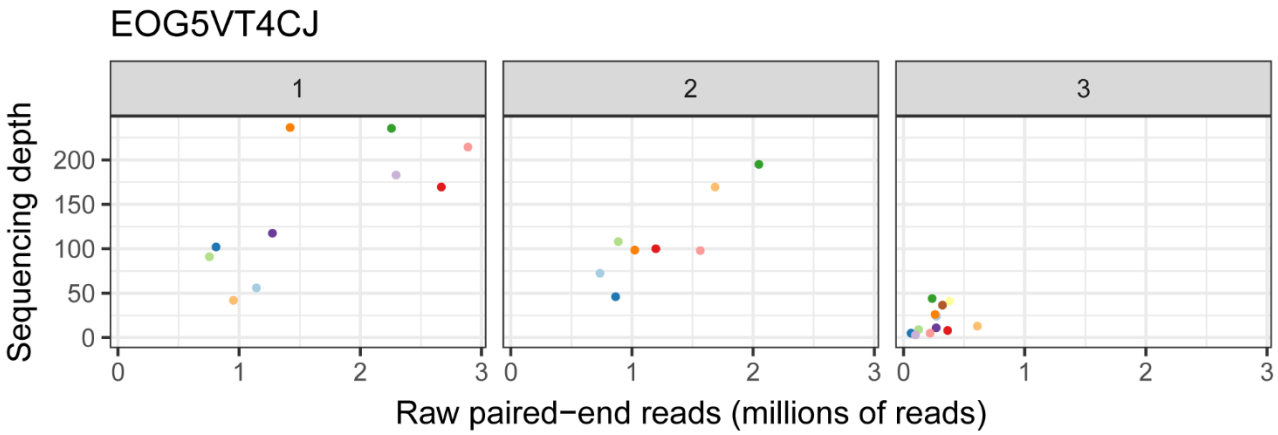

EOG54F4SN
1
2
3
Sequencing depth
0
100
200
0
1
2
3
Raw paired−end reads (millions of reads)

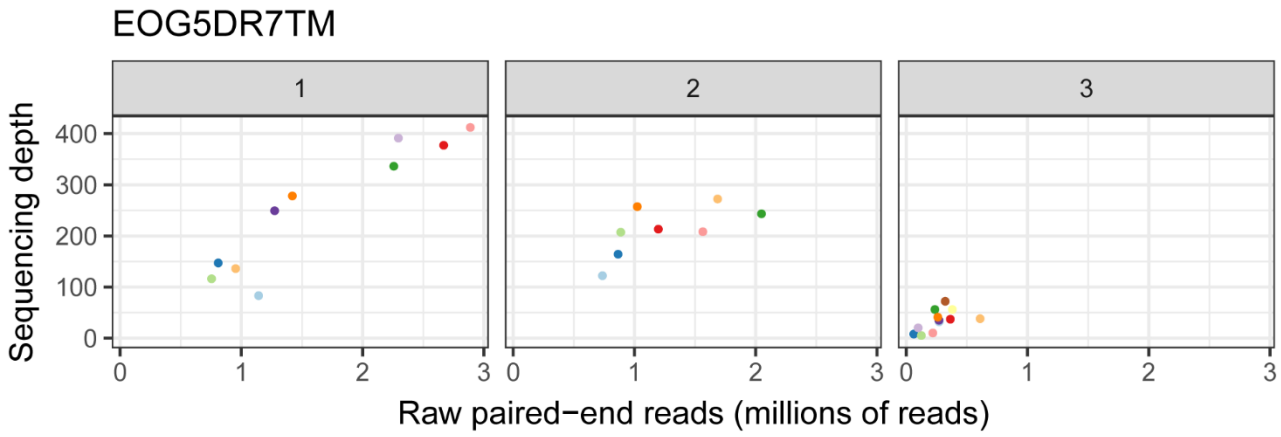

EOG5NZS8B
1
2
3
Sequencing depth
0
200
400
0
1
2
3
Raw paired−end reads (millions of reads)

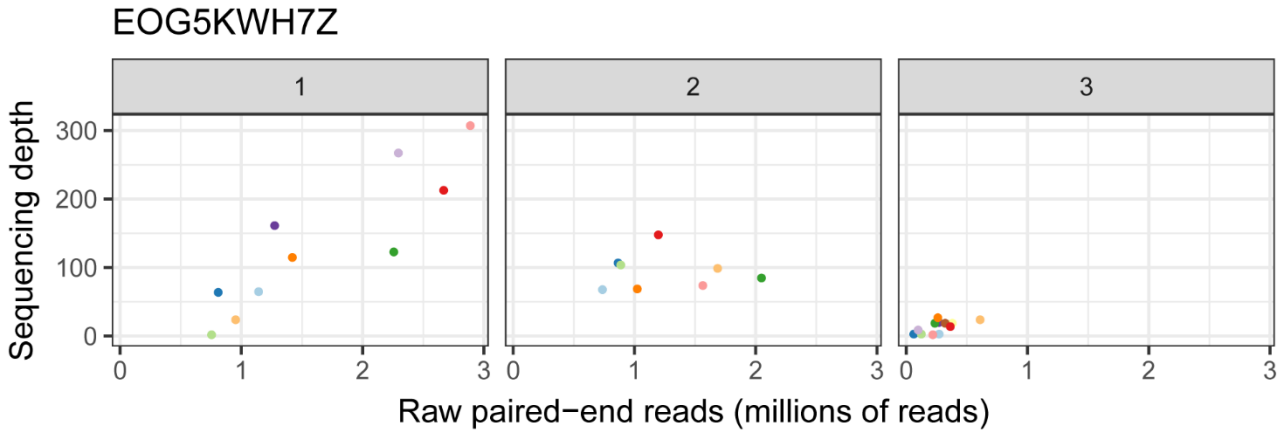

EOG5QNKBN
1
2
3
Sequencing depth
150
100
50
0
0
1
2
3
Raw paired−end reads (millions of reads)

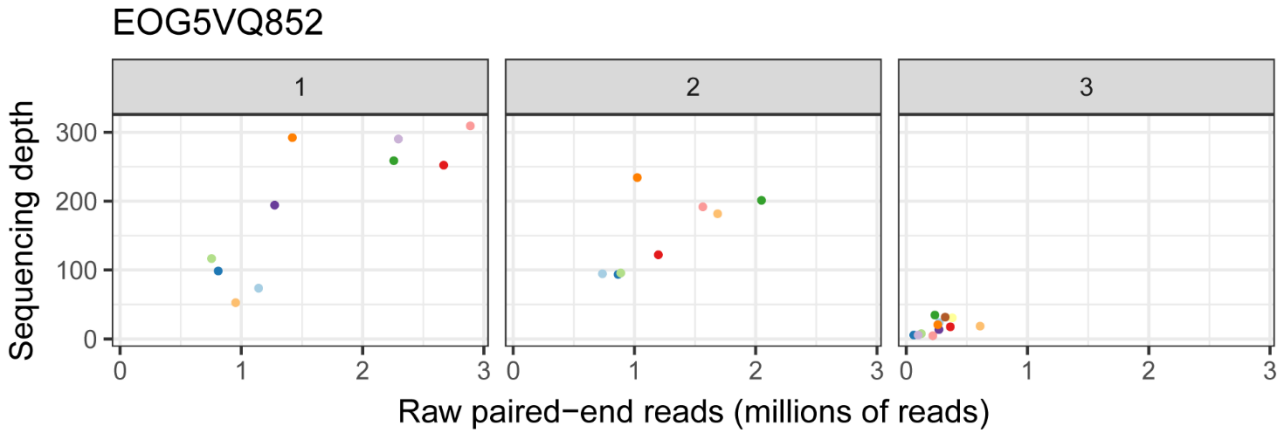

EOG56M91Z
1
2
3
Sequencing depth
400
300
200
100
0
0
1
2
3
Raw paired−end reads (millions of reads)

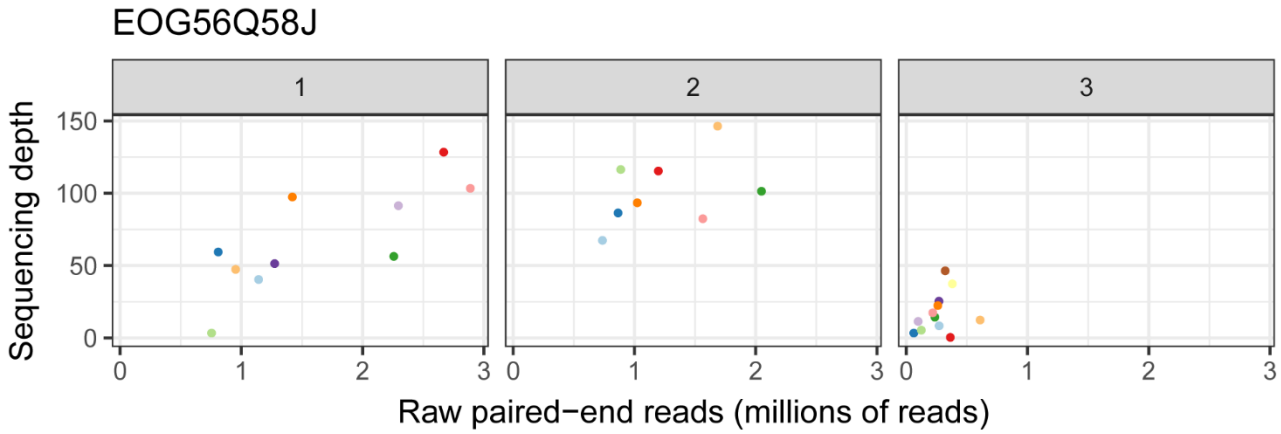

EOG5FTTGQ
1
2
3
Sequencing depth
500
400
300
200
100
0
0
1
2
3
Raw paired−end reads (millions of reads)

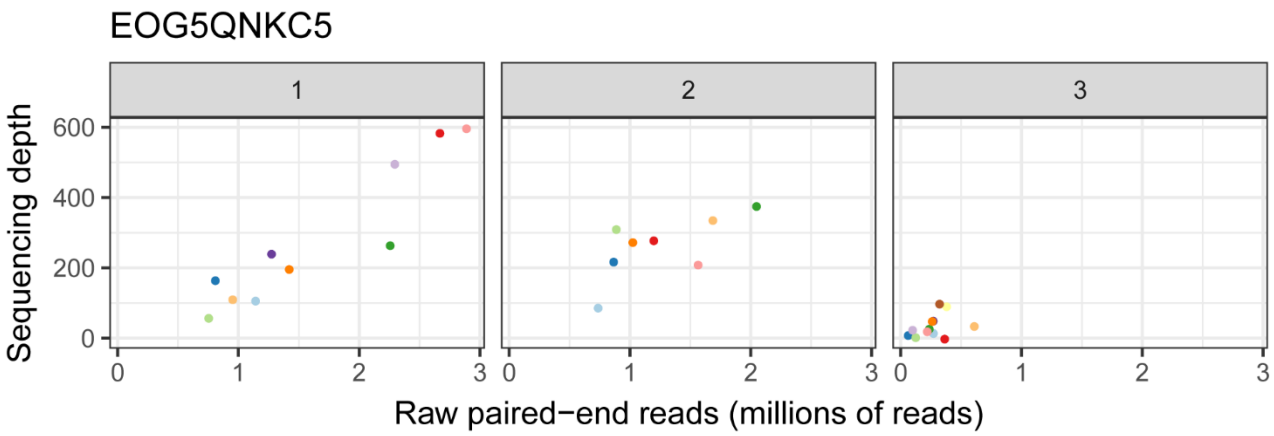

EOG5VT4CJ
1
2
3
Sequencing depth
200
150
100
50
0
0
1
2
3
Raw paired−end reads (millions of reads)

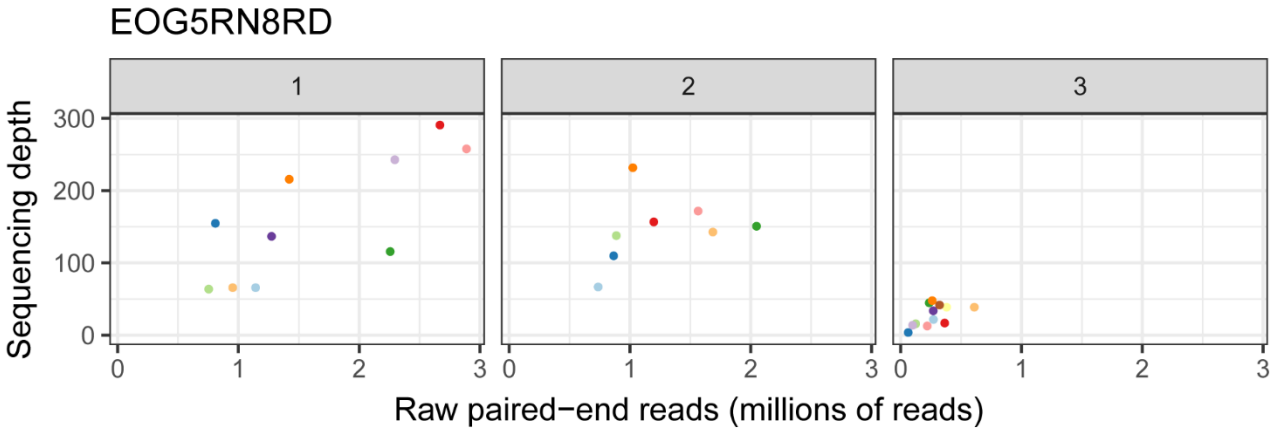

EOG5DR7TM
1
2
3
Sequencing depth
0
100
200
300
400
0
1
2
3
Raw paired−end reads (millions of reads)

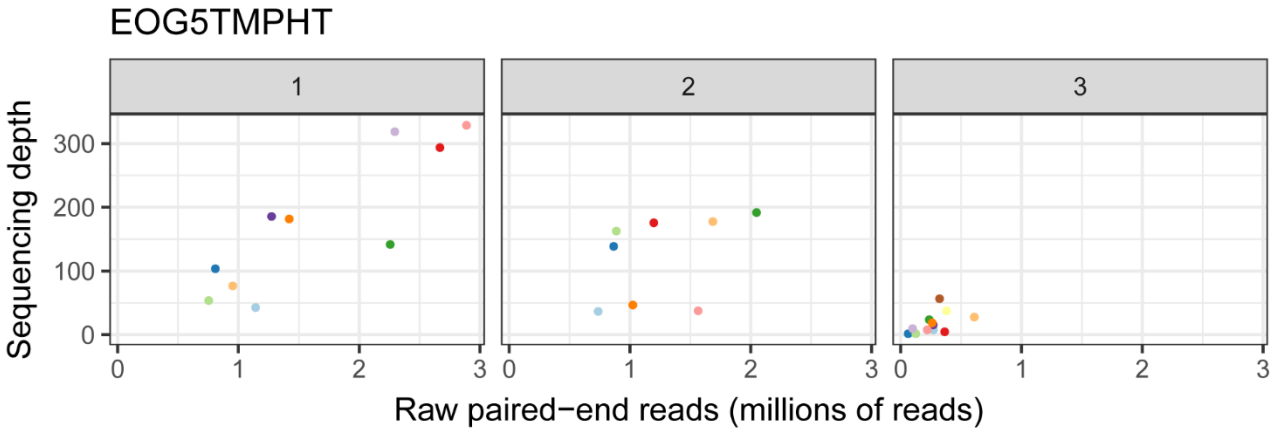

EOG5KWH7Z
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

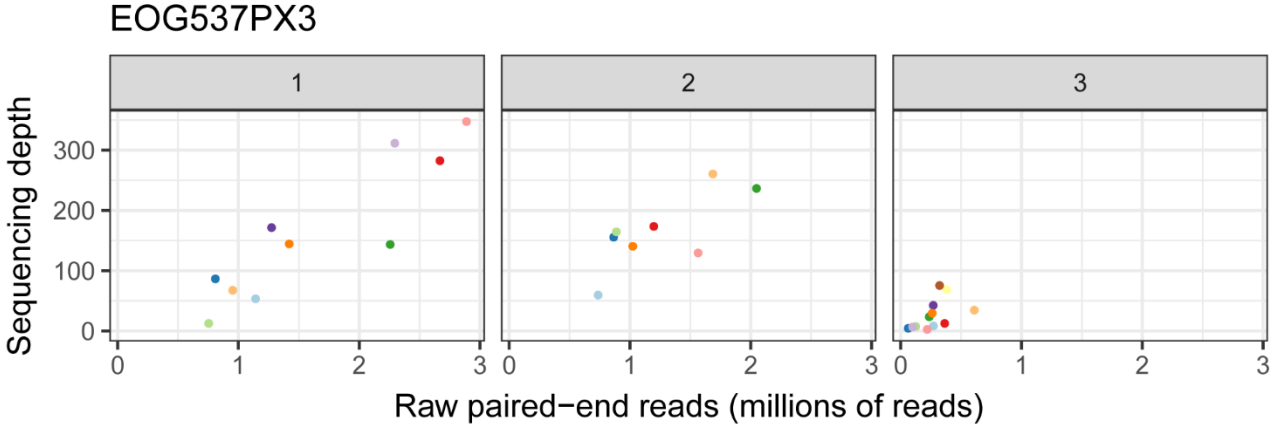

EOG5VQ852
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

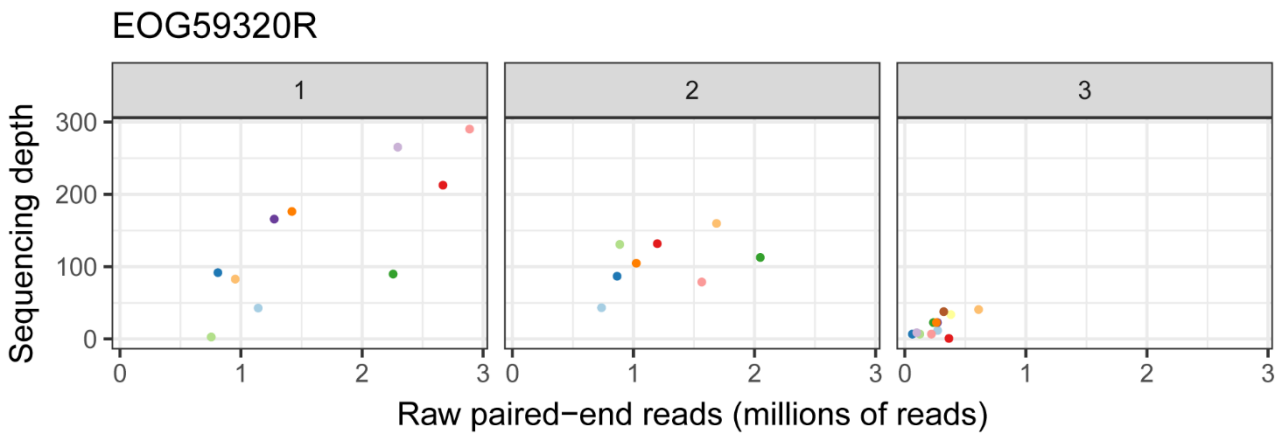

EOG56Q58J
1
2
3
Sequencing depth
0
50
100
150
0
1
2
3
Raw paired−end reads (millions of reads)

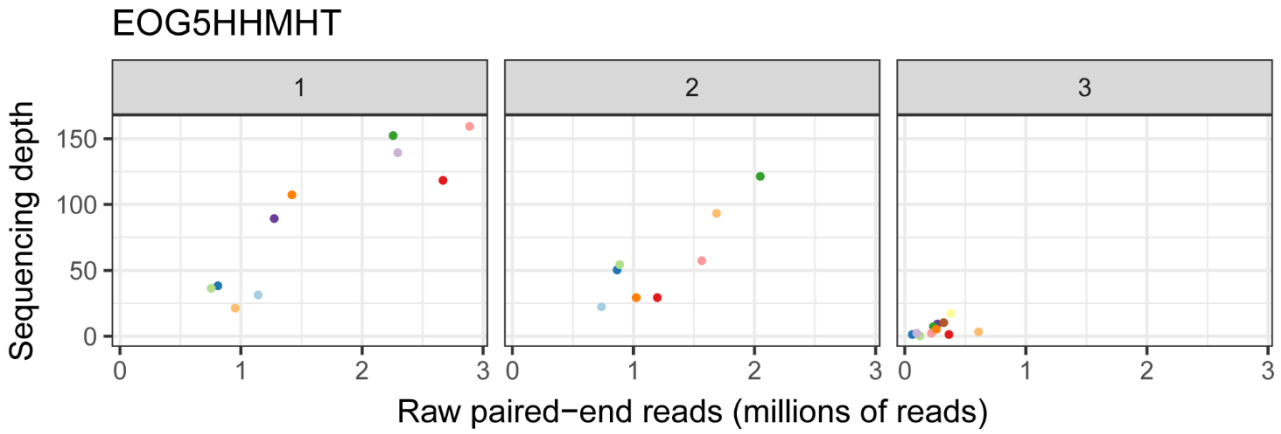

EOG5QNKC5
1
2
3
Sequencing depth
0
200
400
600
0
1
2
3
Raw paired−end reads (millions of reads)

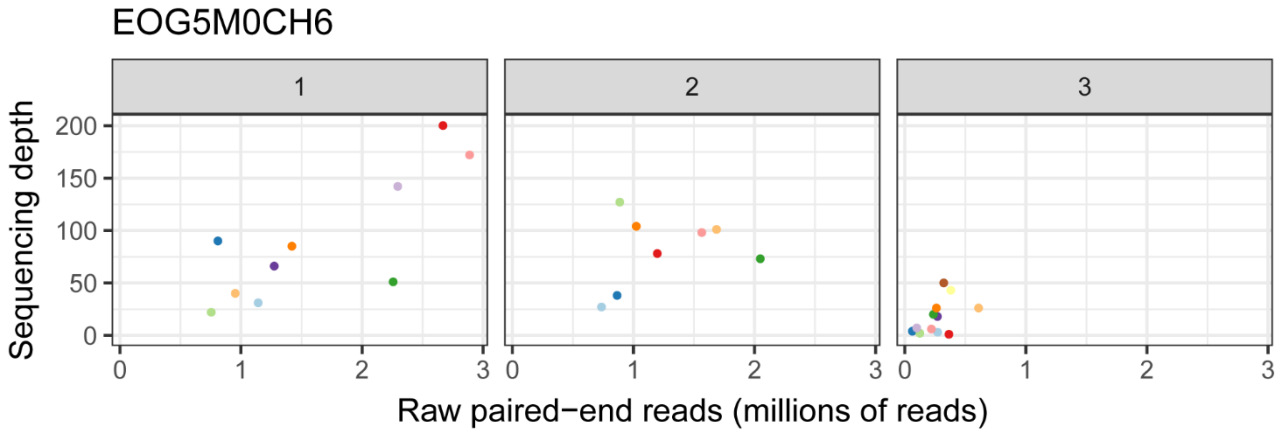

EOG5RN8RD
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

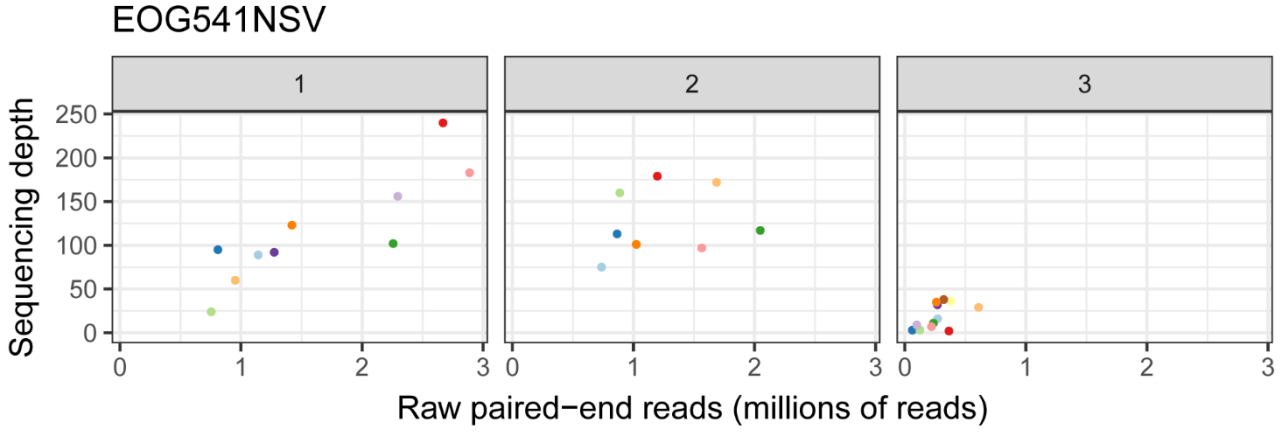

EOG5TMPHT
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

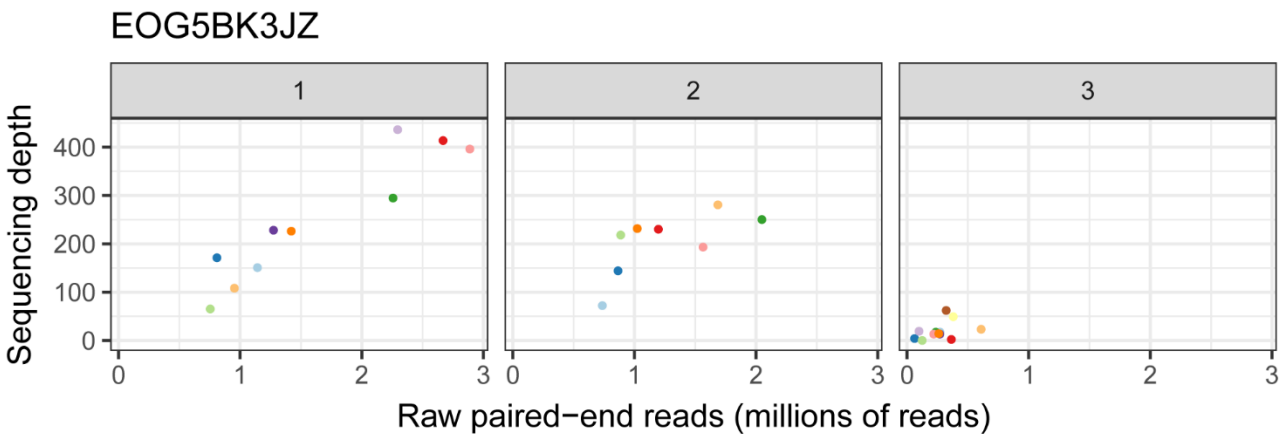

EOG537PX3
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

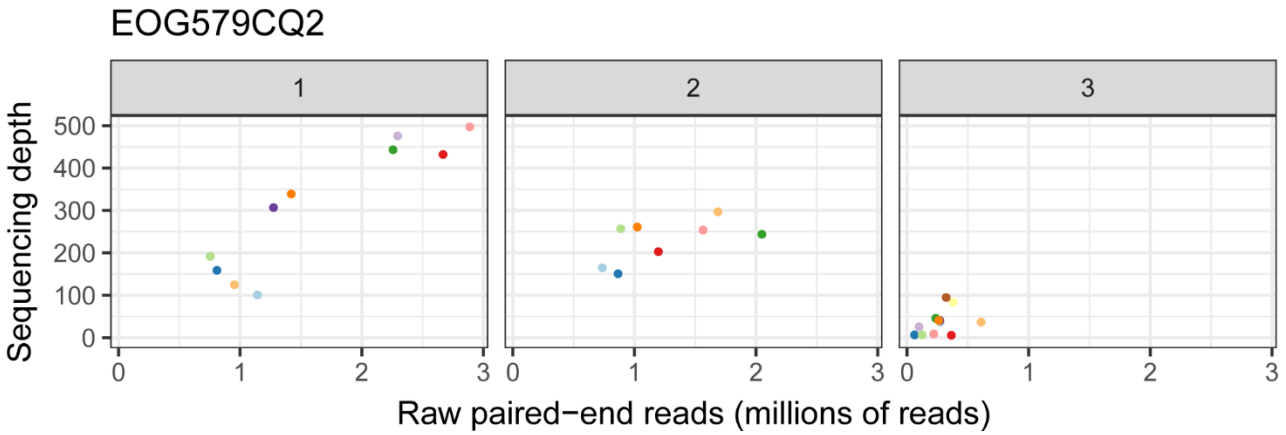

EOG59320R
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

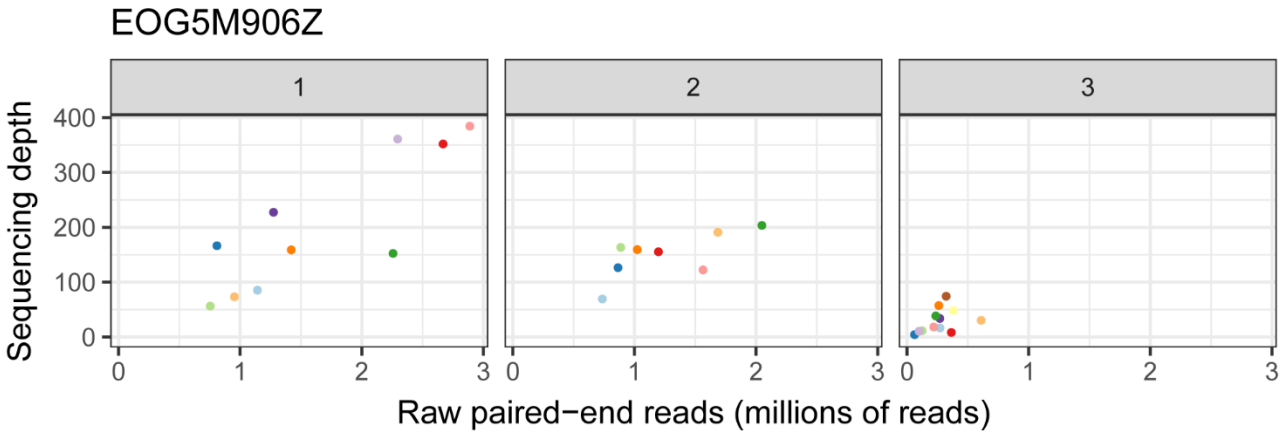

EOG5HHMHT
1
2
3
Sequencing depth
0
50
100
150
0
1
2
3
Raw paired−end reads (millions of reads)

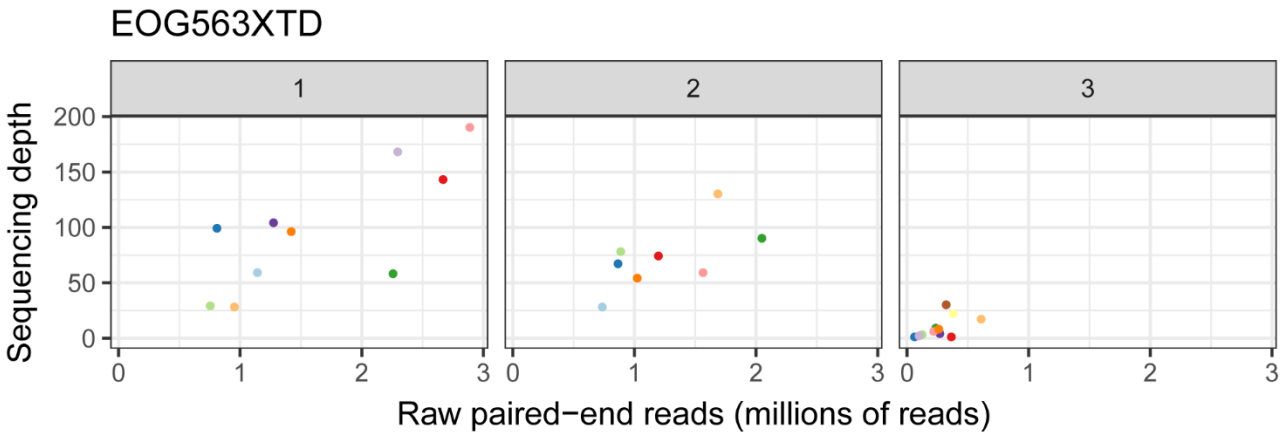

EOG5M0CH6
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)

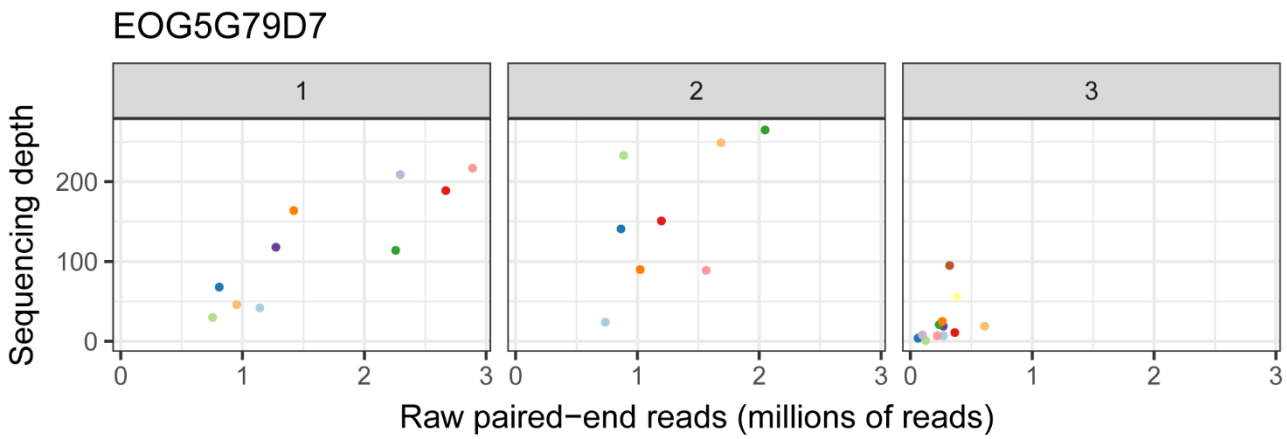

EOG541NSV
1
2
3
Sequencing depth
0
50
100
150
200
250
0
1
2
3
Raw paired−end reads (millions of reads)

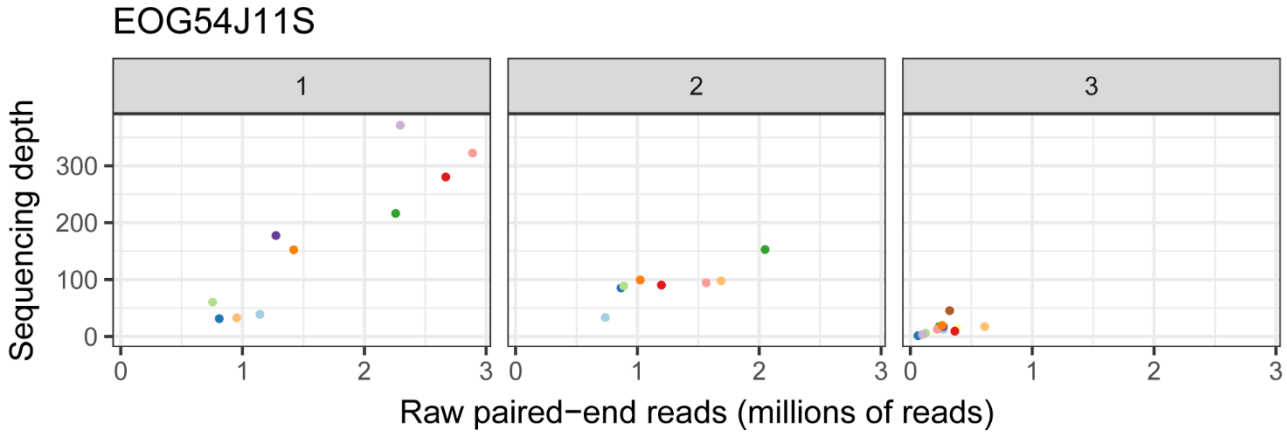

EOG5BK3JZ
1
2
3
Sequencing depth
0
100
200
300
400
0
1
2
3
Raw paired−end reads (millions of reads)

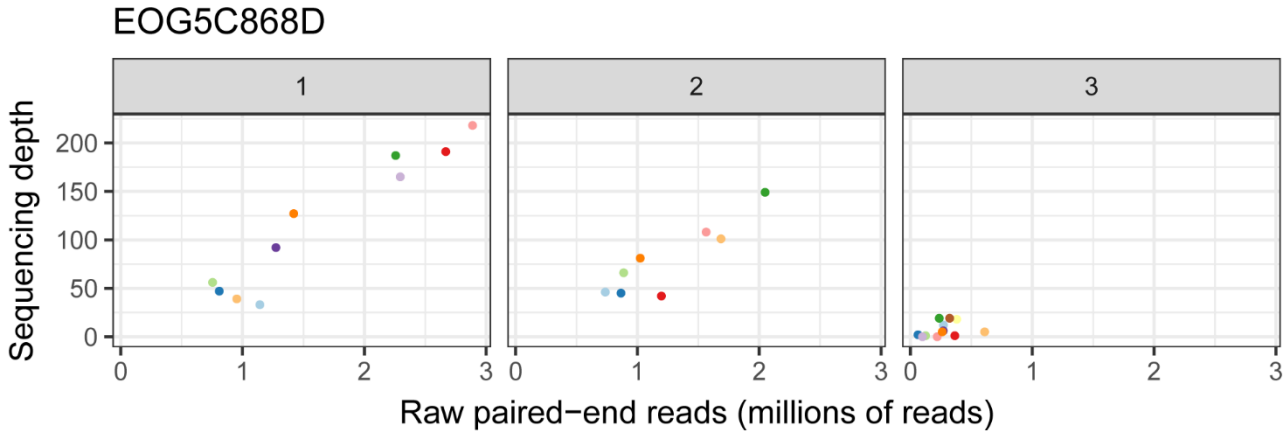

EOG579CQ2
1
2
3
Sequencing depth
0
100
200
300
400
500
0
1
2
3
Raw paired−end reads (millions of reads)

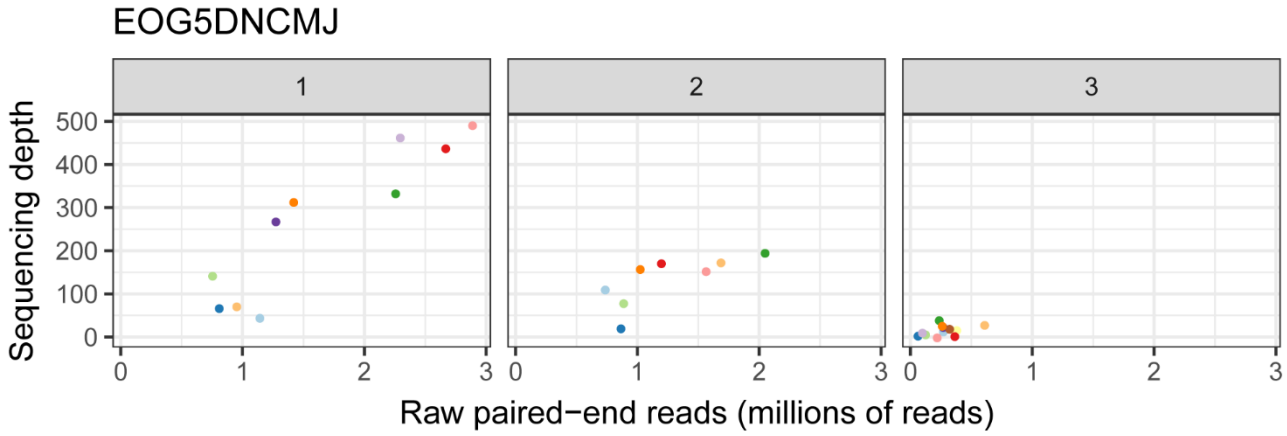

EOG5M906Z
1
2
3
Sequencing depth
0
100
200
300
400
0
1
2
3
Raw paired−end reads (millions of reads)

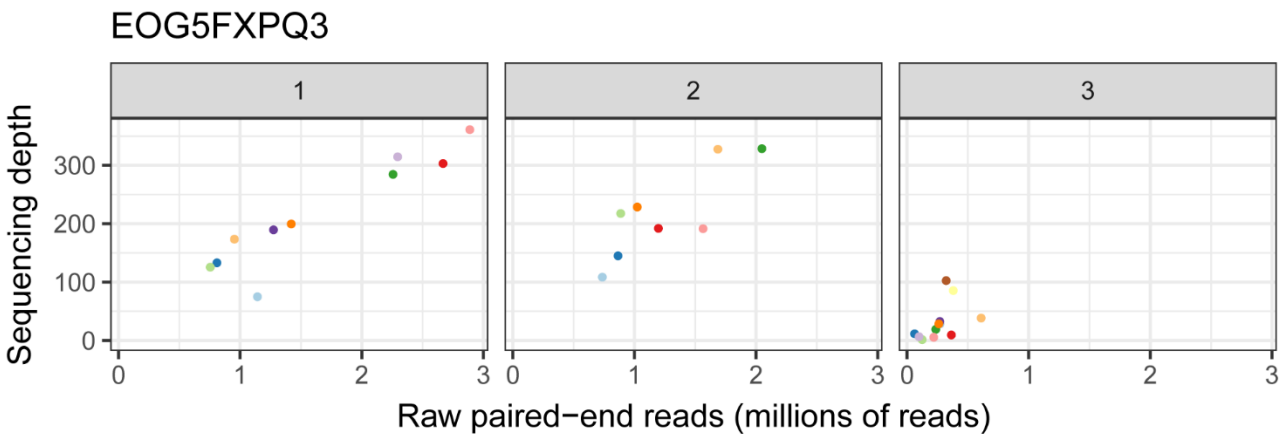

EOG563XTD
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)

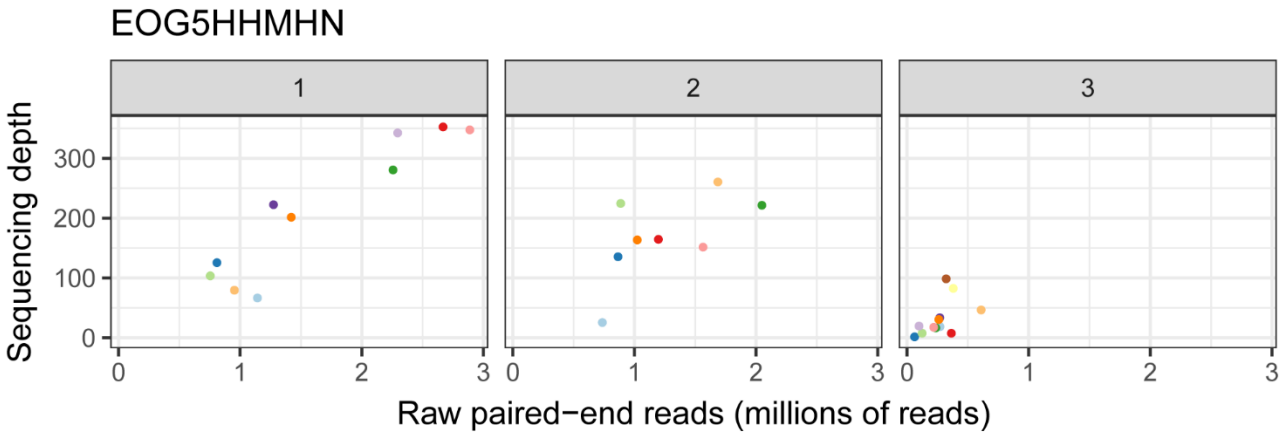

EOG5G79D7
1
2
3
Sequencing depth
0
100
200
0
1
2
3
Raw paired−end reads (millions of reads)

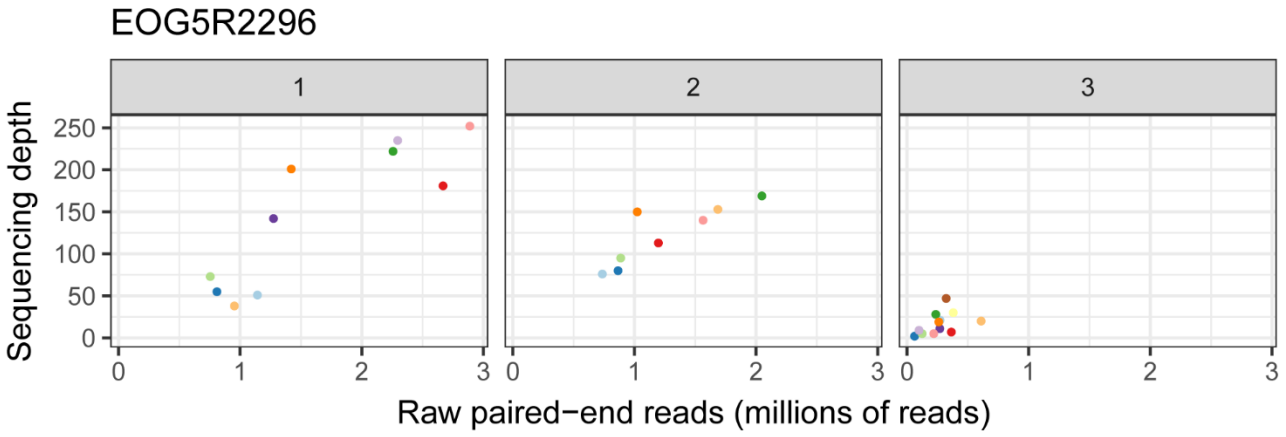

EOG54J11S
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

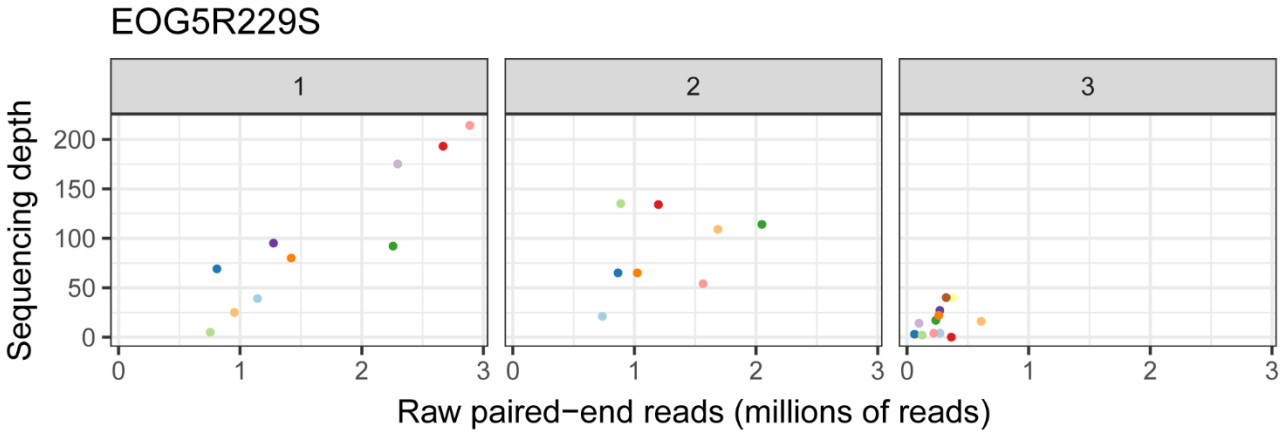

EOG5C868D
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)

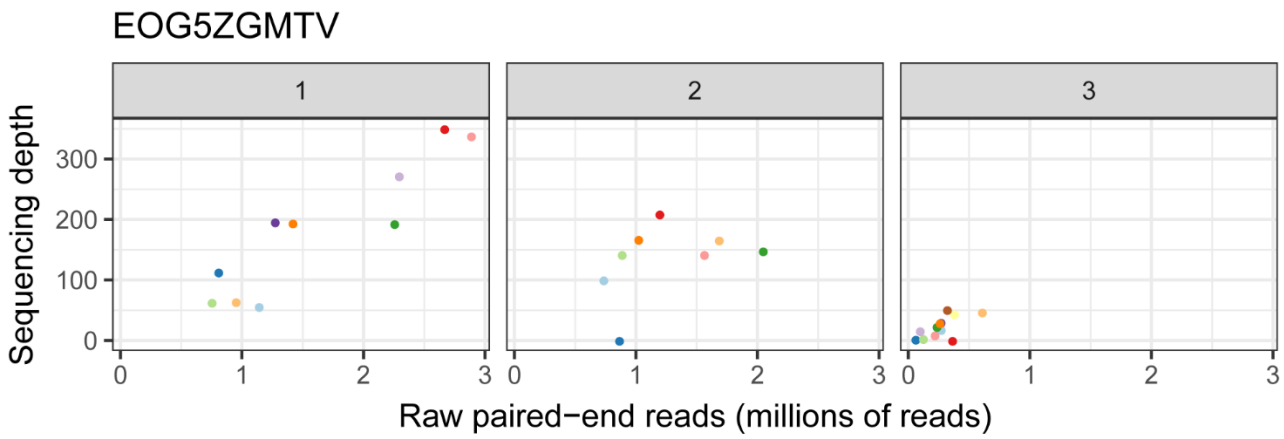

EOG5DNCMJ
1
2
3
Sequencing depth
0
100
200
300
400
500
0
1
2
3
Raw paired−end reads (millions of reads)

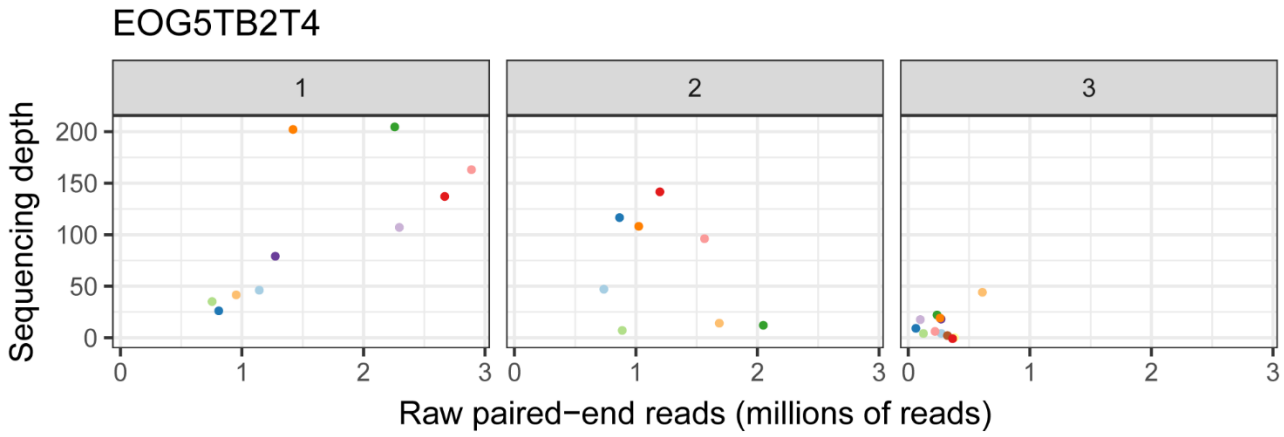

EOG5FXPQ3
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

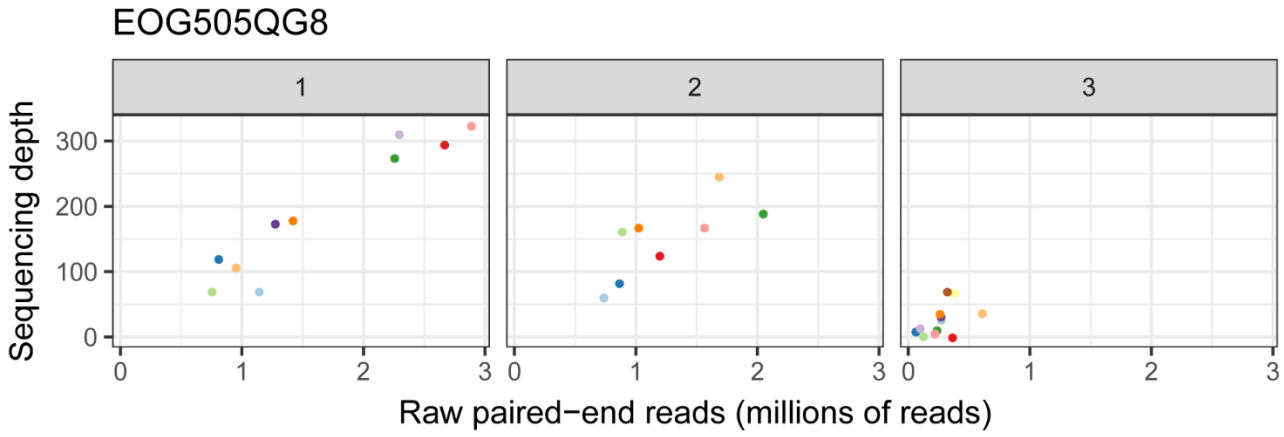

EOG5HHMHN
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

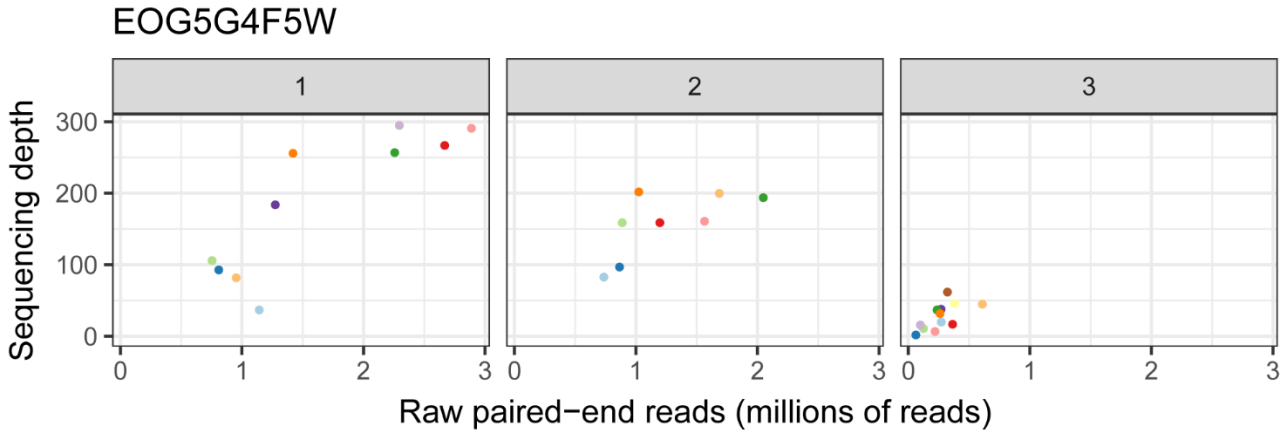

EOG5R2296
1
2
3
Sequencing depth
0
50
100
150
200
250
0
1
2
3
Raw paired−end reads (millions of reads)

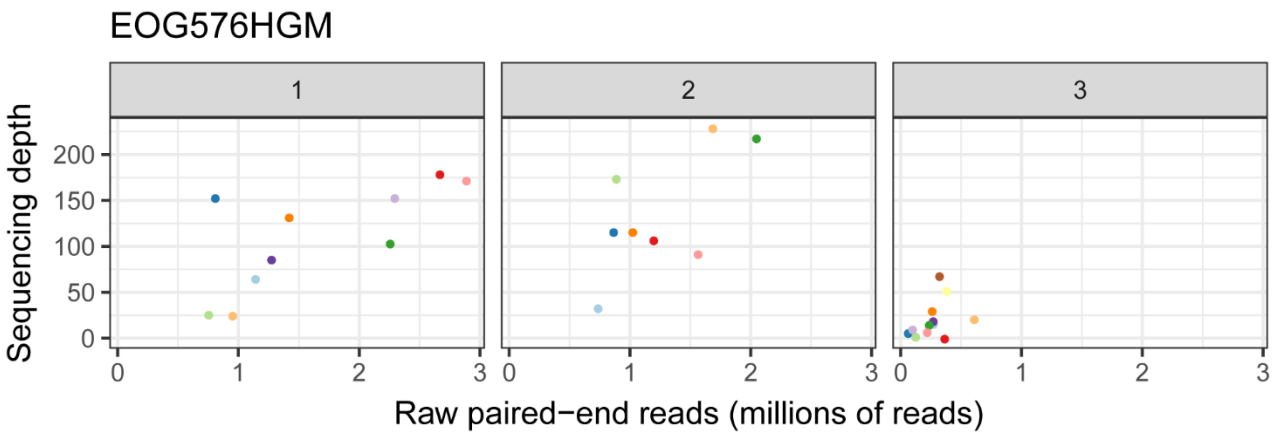

EOG5R229S
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)

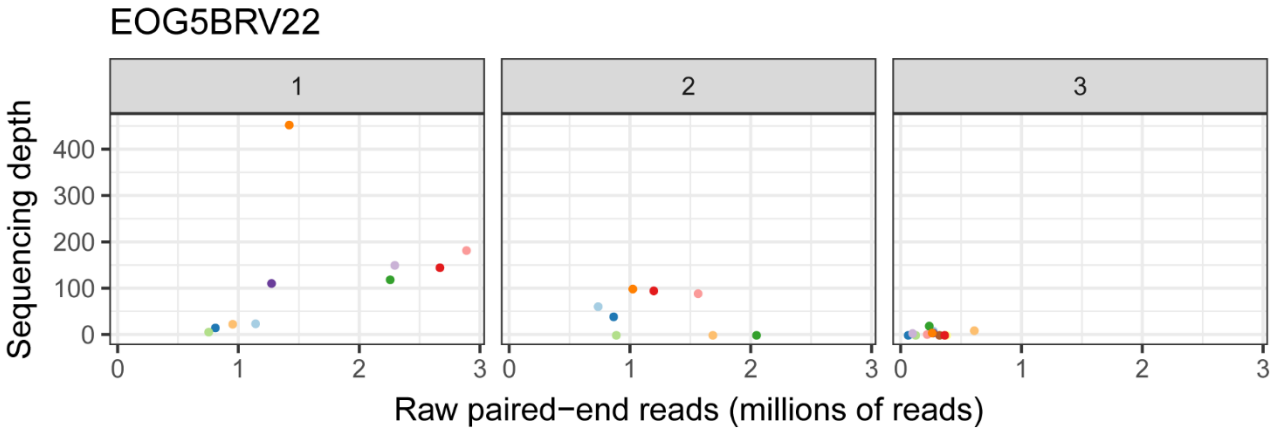

EOG5ZGMTV
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

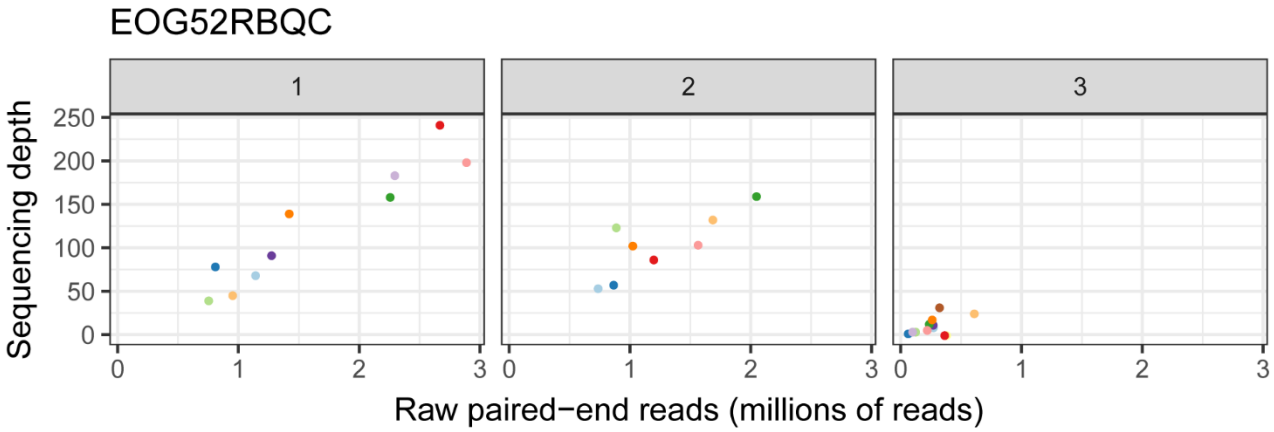

EOG5TB2T4
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)

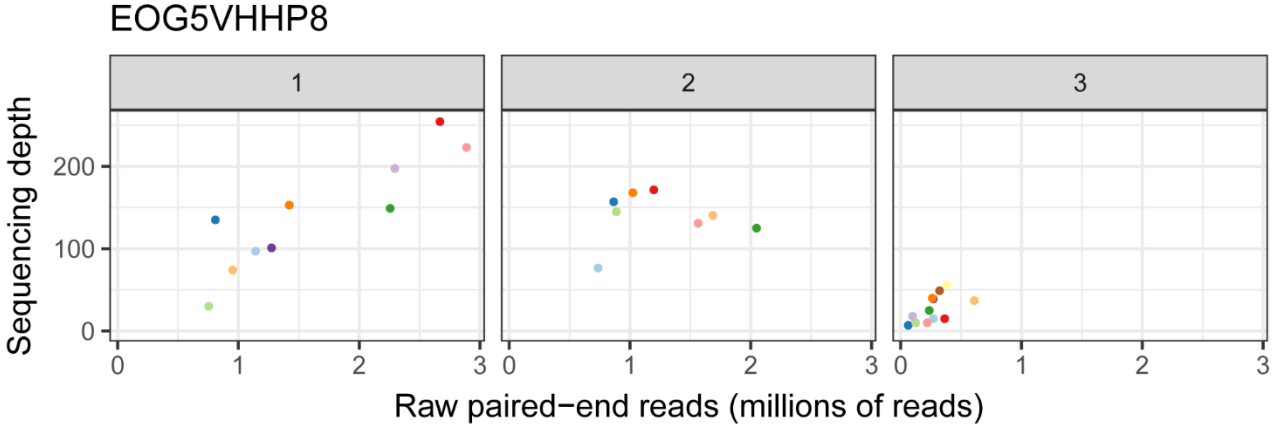

EOG505QG8
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)


EOG5G4F5W
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)


EOG576HGM
1
2
3
Sequencing depth
0
50
100
150
200
0
1
2
3
Raw paired−end reads (millions of reads)


EOG5BRV22
1
2
3
Sequencing depth
0
100
200
300
400
0
1
2
3
Raw paired−end reads (millions of reads)


EOG52RBQC
1
2
3
Sequencing depth
0
50
100
150
200
250
0
1
2
3
Raw paired−end reads (millions of reads)


EOG5VHHP8
1
2
3
Sequencing depth
0
100
200
0
1
2
3
Raw paired−end reads (millions of reads)

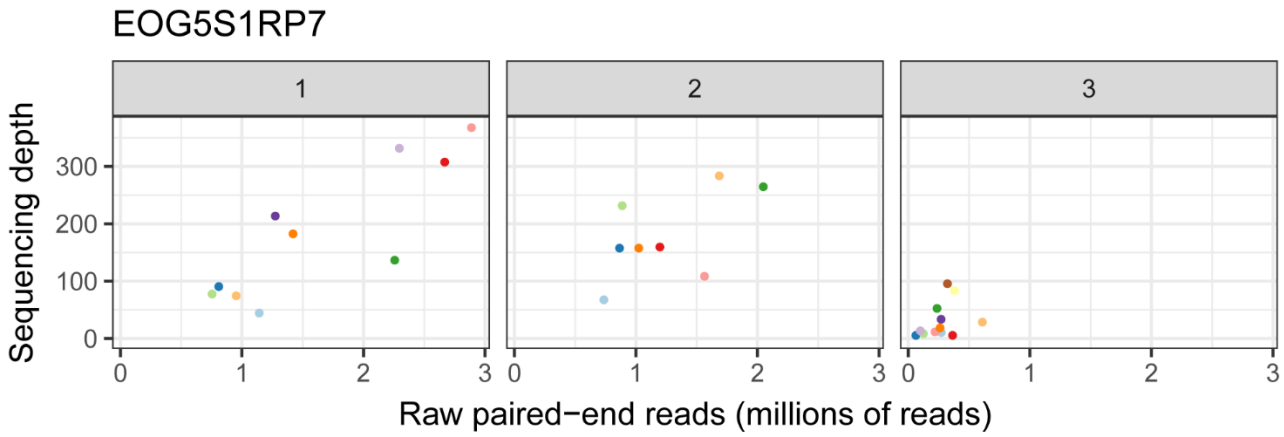
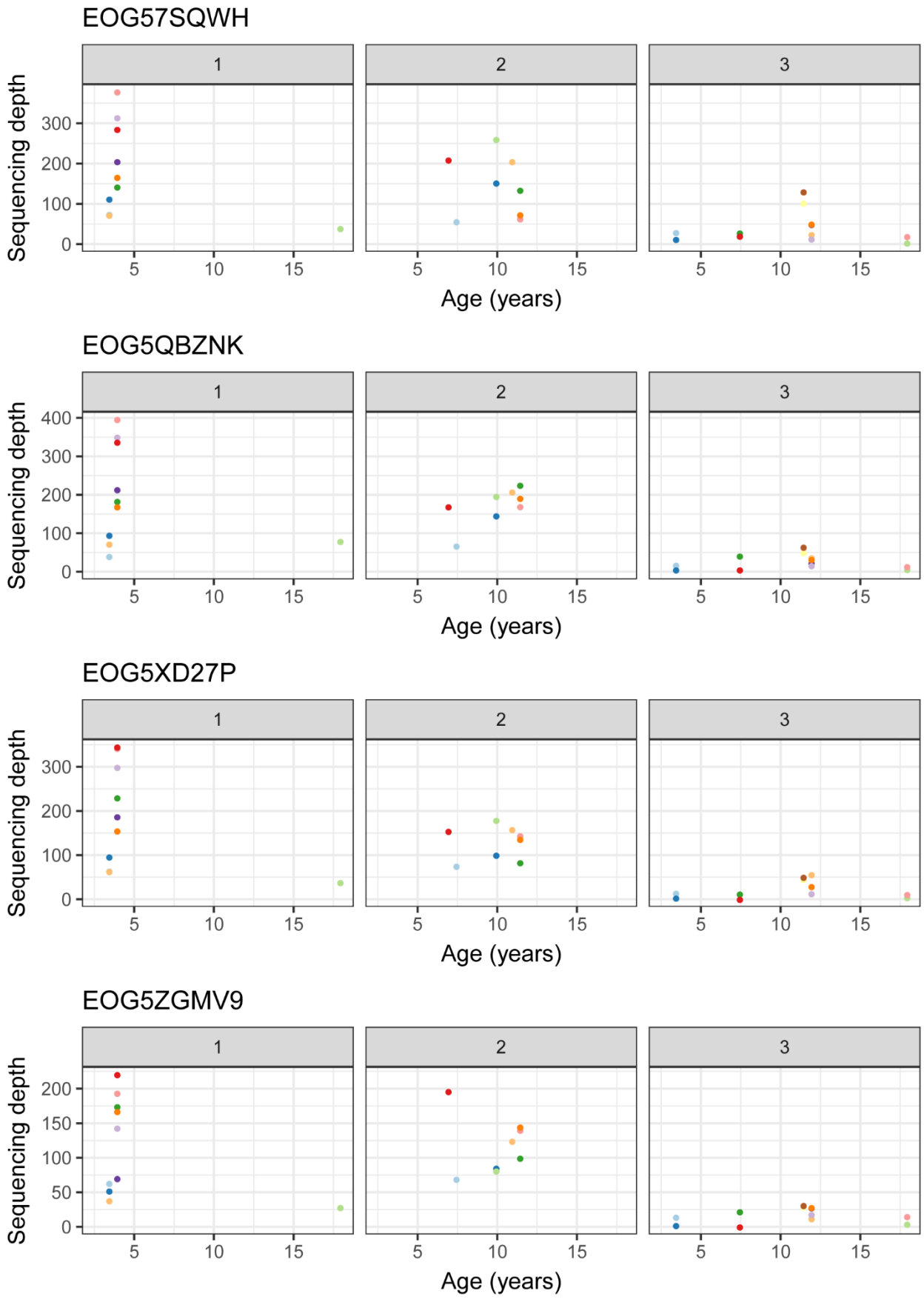
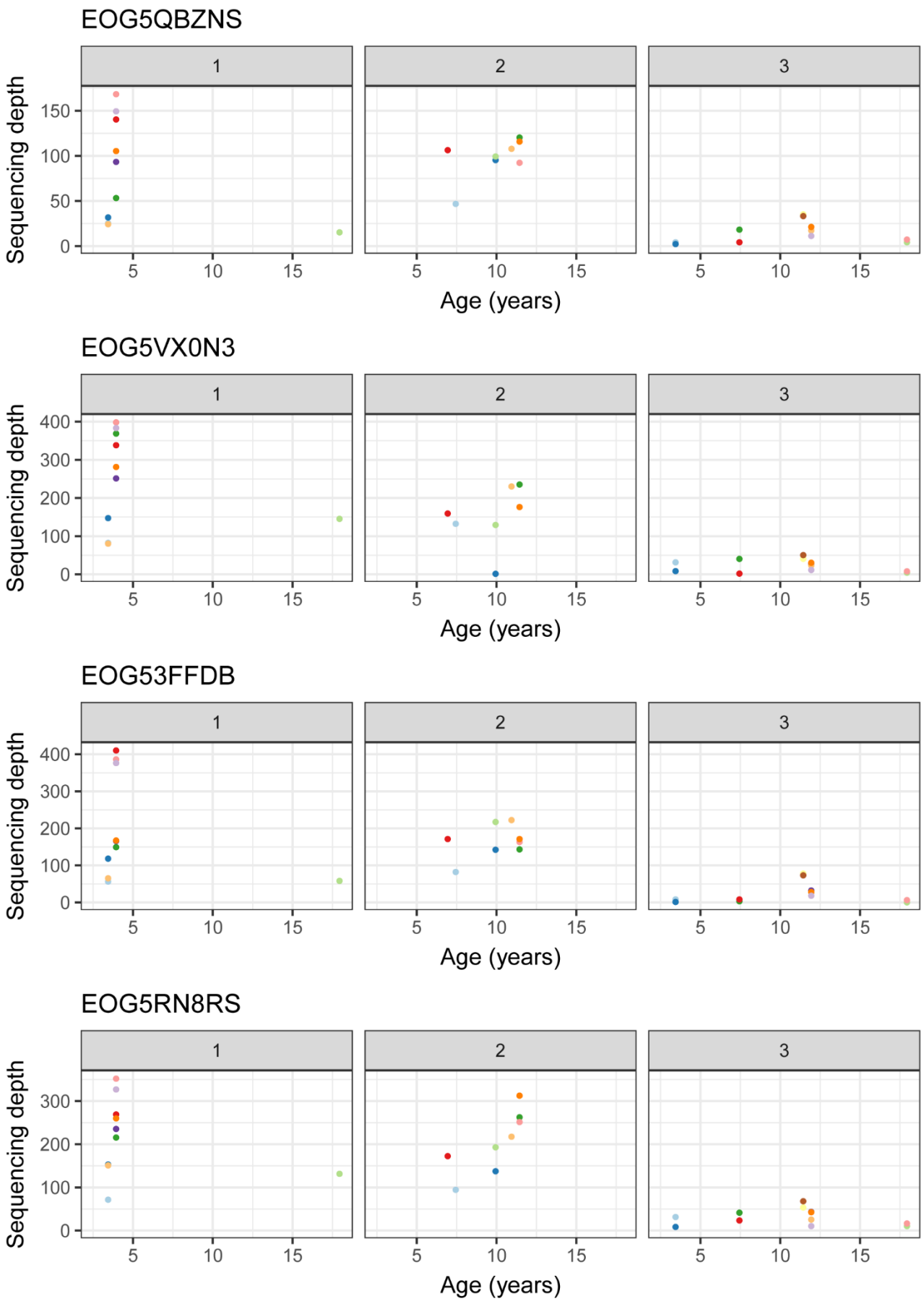
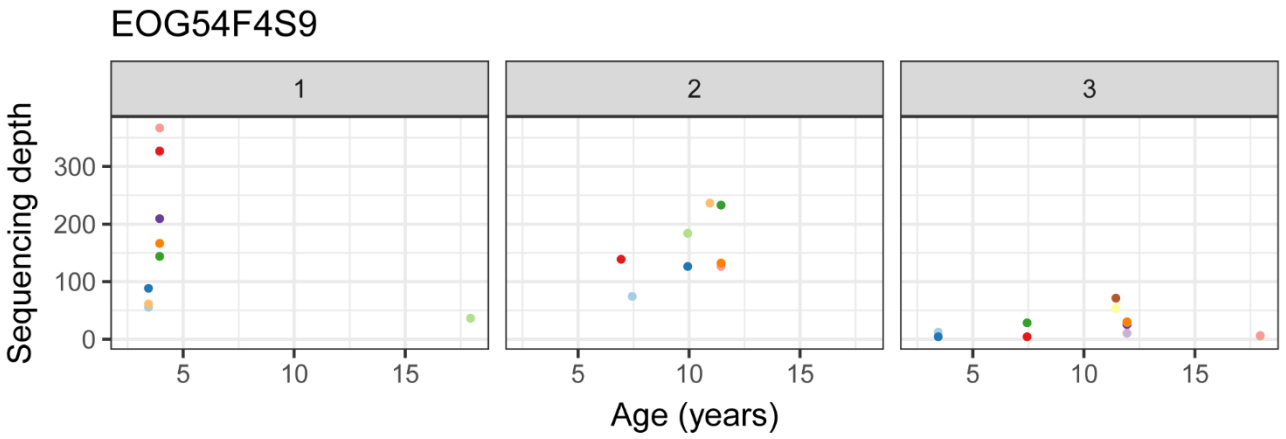
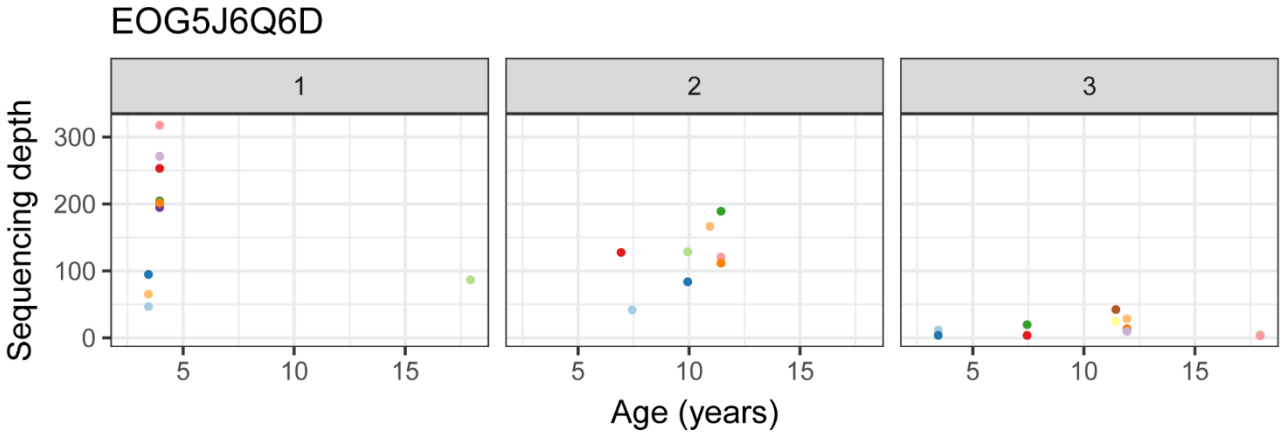

EOG5S1RP7
1
2
3
Sequencing depth
0
100
200
300
0
1
2
3
Raw paired−end reads (millions of reads)

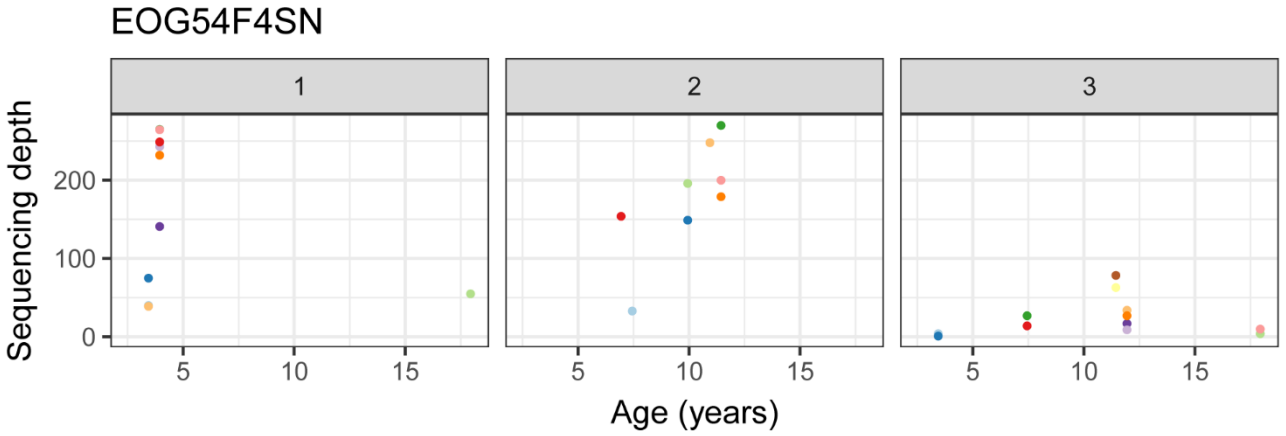
Figure S8: Plots for specimen preservation age (years since collected) against median sequencing depth across samples for exons of 50 randomly targeted orthologues, separated by sequencing run (1–3). Orthologue code is specified above plots, and points are coloured according to specimen ID.

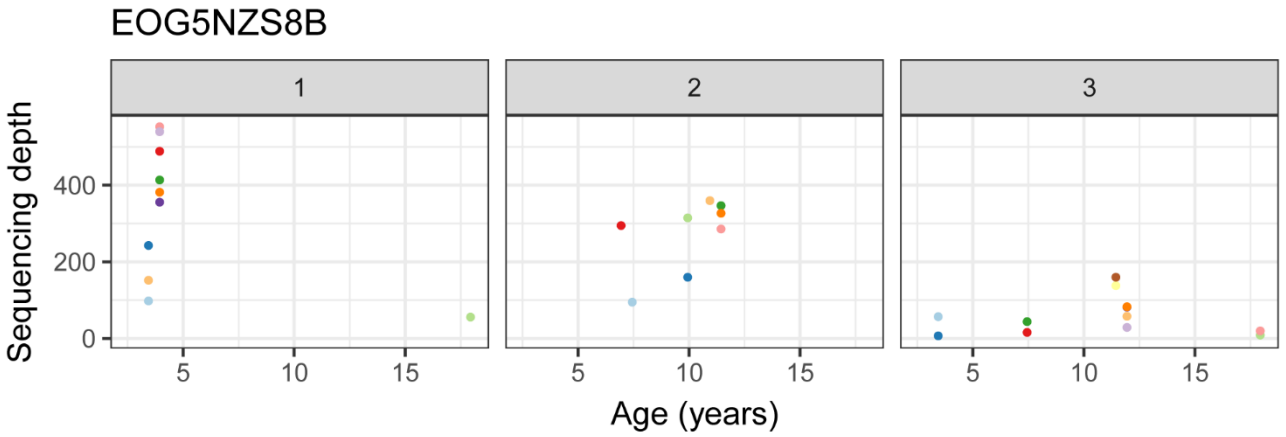


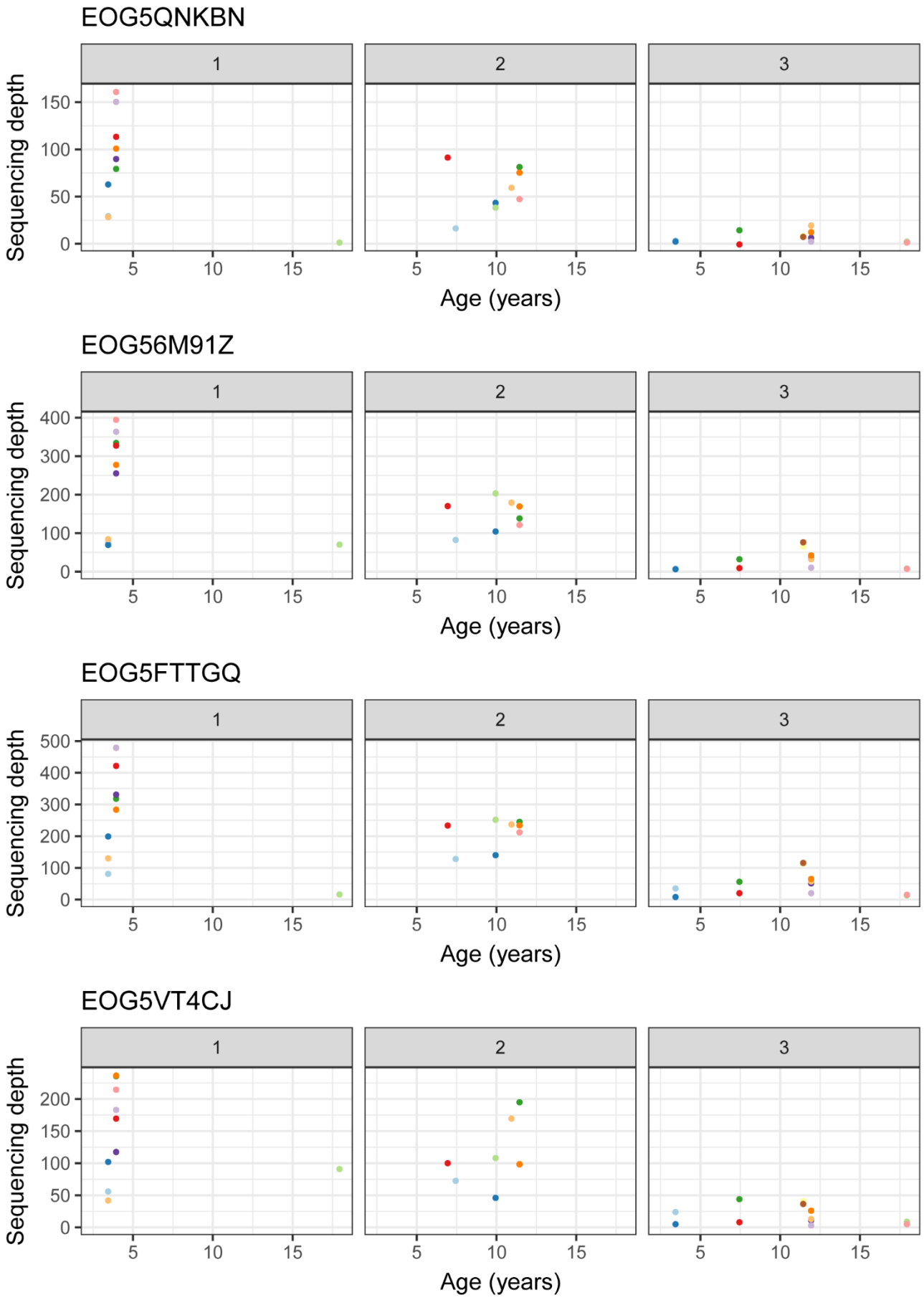

EOG5QBZNS
1
2
3
Sequencing depth
0
50
100
150
5
10
15
Age (years)
EOG5VX0N3
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)
EOG53FFDB
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)
EOG5RN8RS
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)

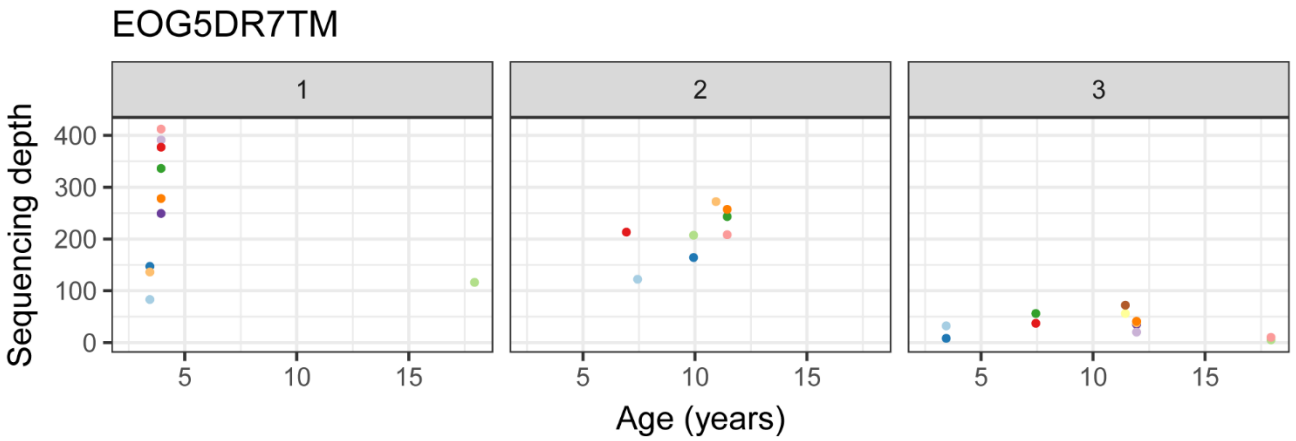

EOG54F4S9
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

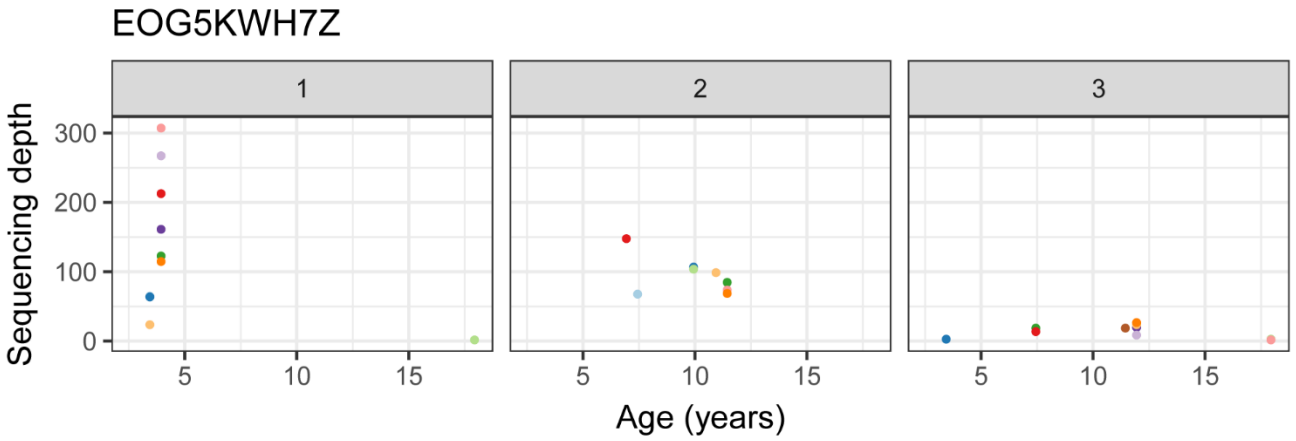

EOG5J6Q6D
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

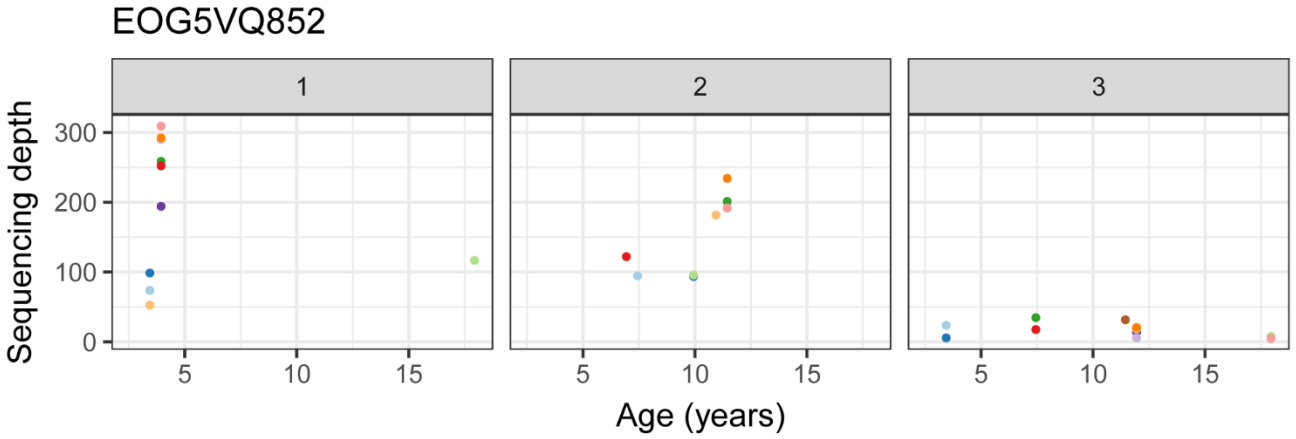

EOG54F4SN
1
2
3
Sequencing depth
200
100
0
5
10
15
Age (years)

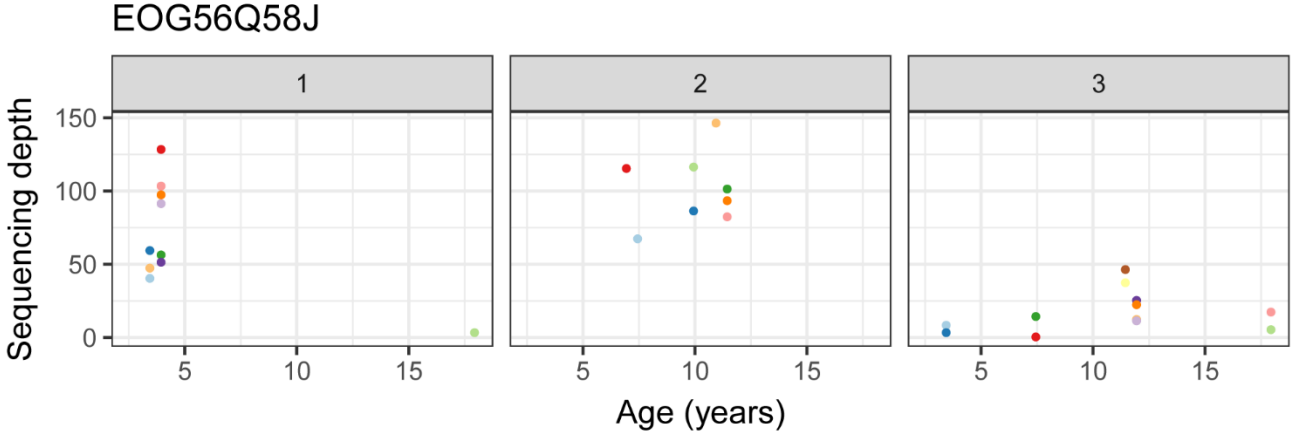

EOG5NZS8B
1
2
3
Sequencing depth
400
200
0
5
10
15
Age (years)

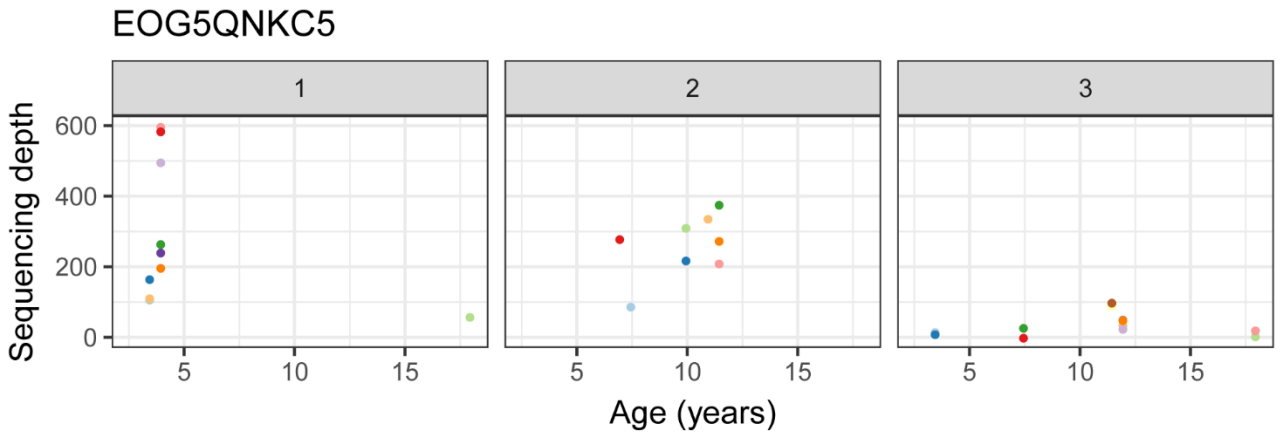

EOG5QNKBN
1
2
3
Sequencing depth
0
50
100
150
5
10
15
Age (years)
EOG56M91Z
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)
EOG5FTTGQ
1
2
3
Sequencing depth
0
100
200
300
400
500
5
10
15
Age (years)
EOG5VT4CJ
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)

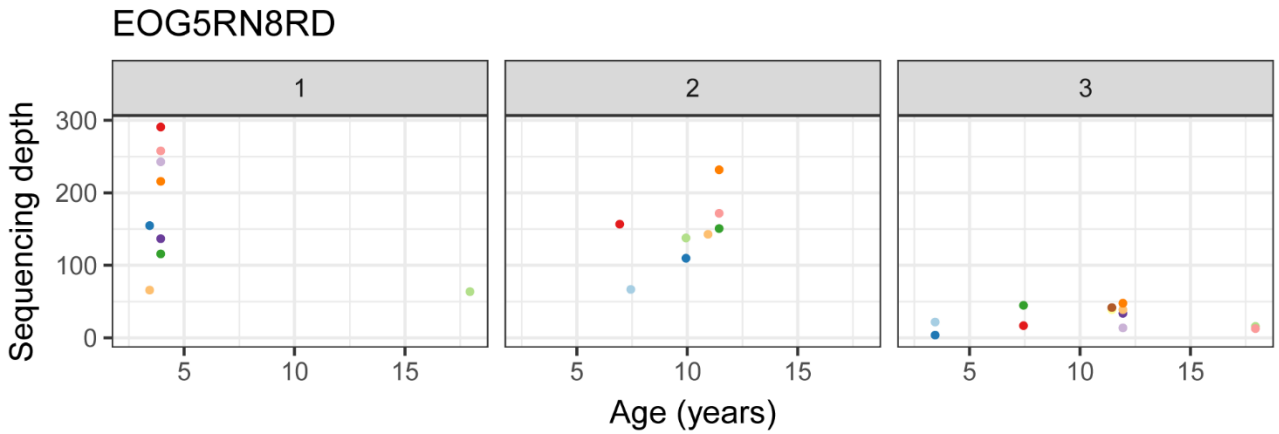

EOG5DR7TM
1
2
3
Sequencing depth
400
300
200
100
0
5
10
15
Age (years)

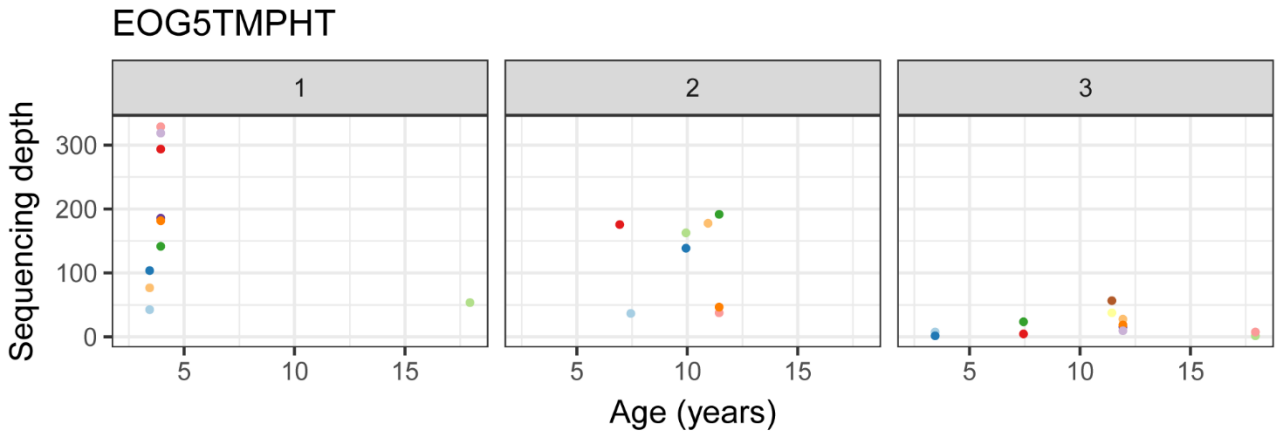

EOG5KWH7Z
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

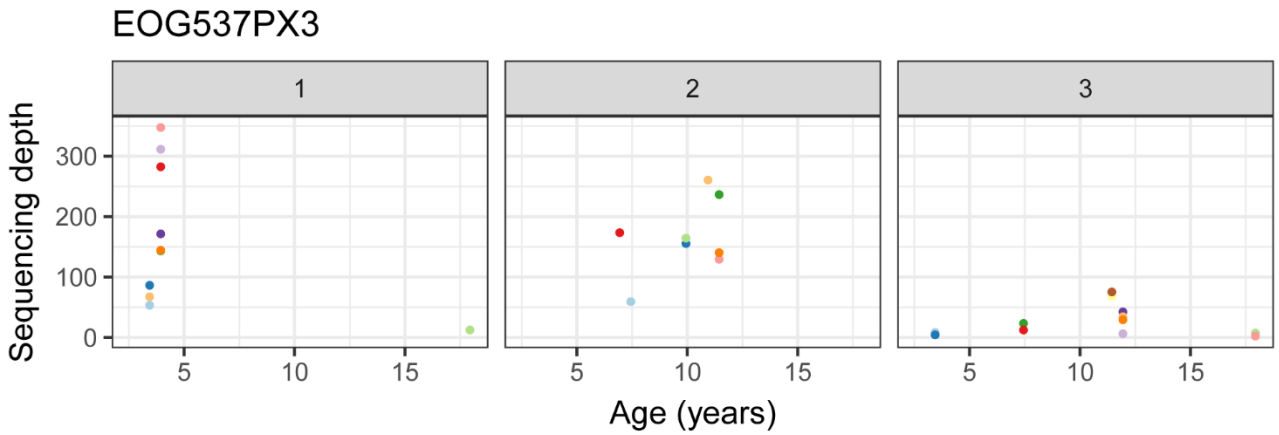

EOG5VQ852
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

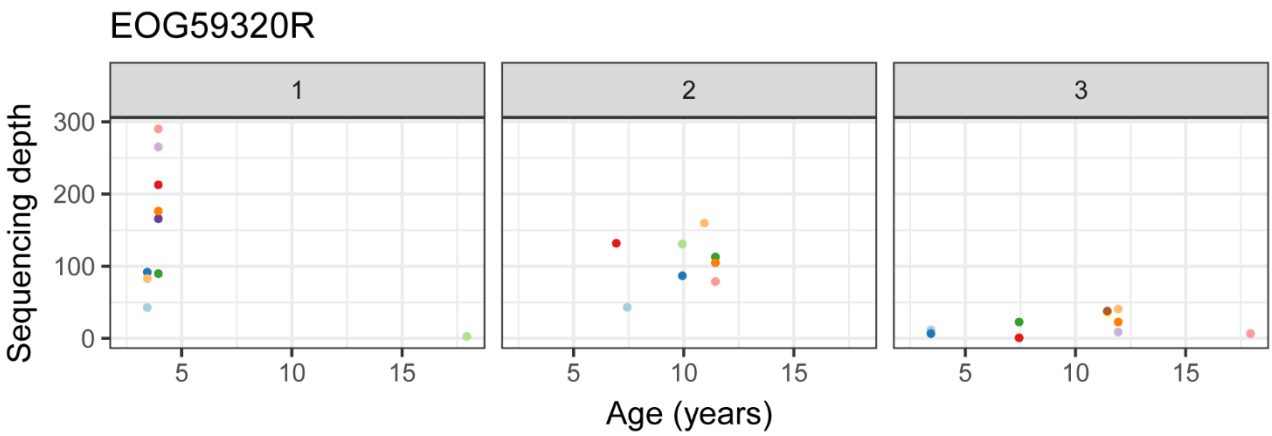

EOG56Q58J
1
2
3
Sequencing depth
150
100
50
0
5
10
15
Age (years)

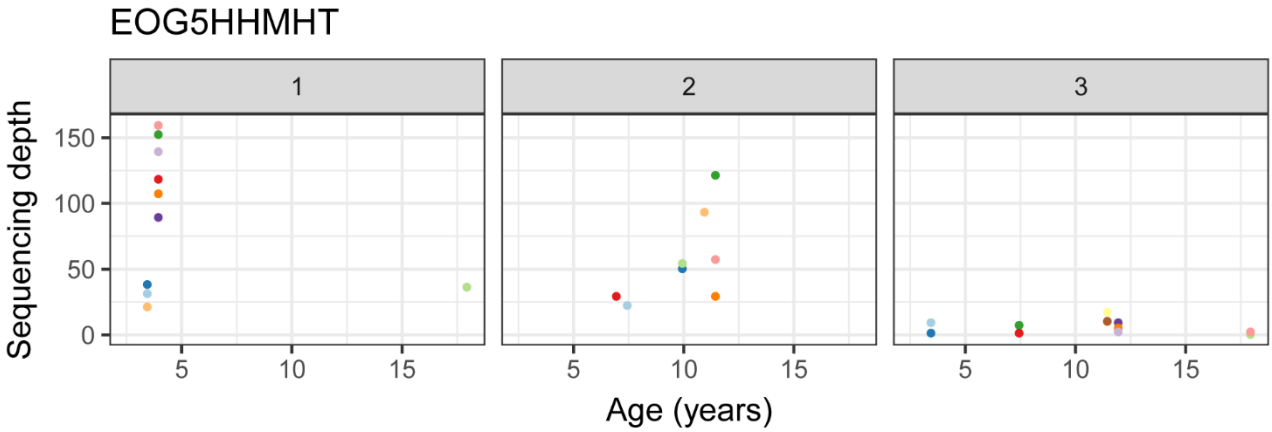

EOG5QNKC5
1
2
3
Sequencing depth
600
400
200
0
5
10
15
Age (years)

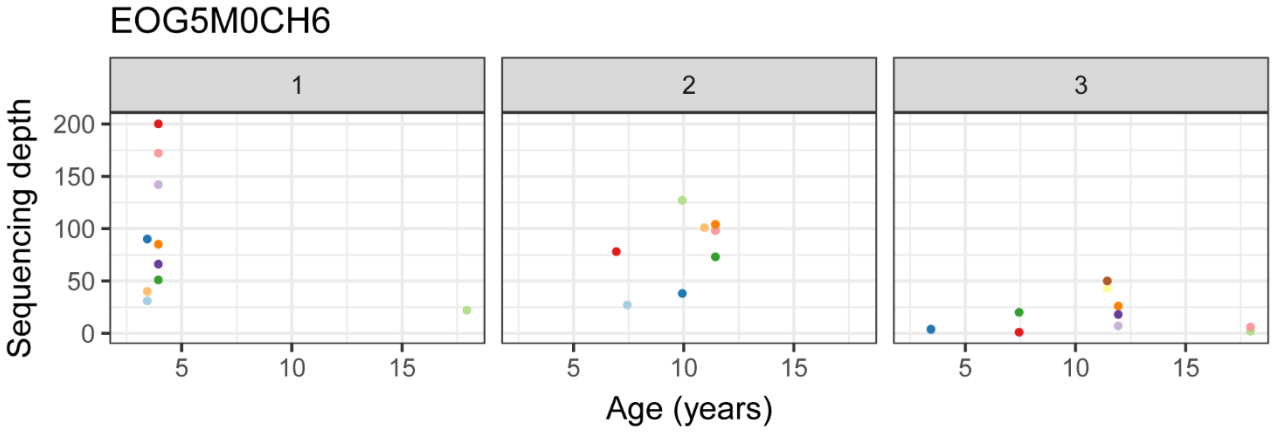

EOG5RN8RD
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

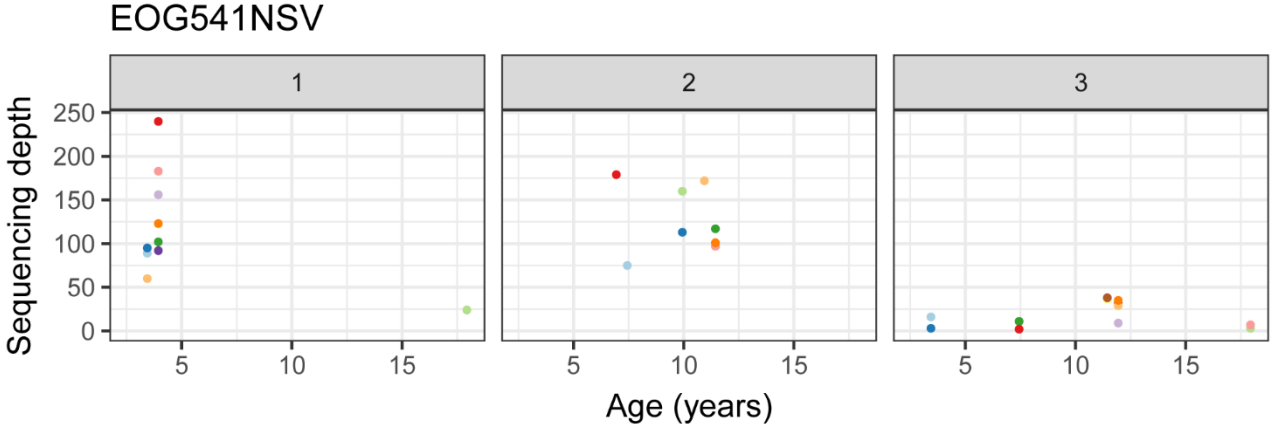

EOG5TMPHT
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

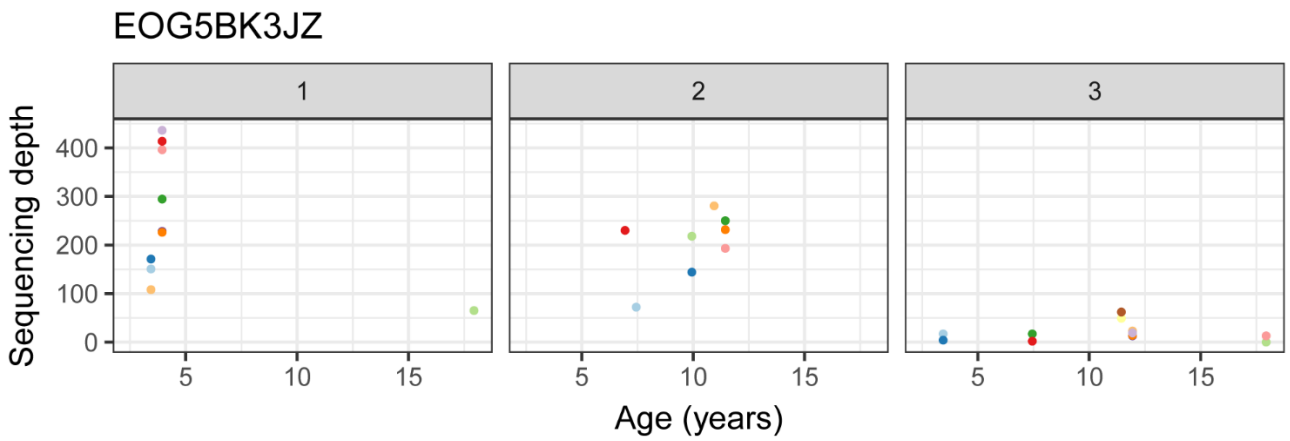

EOG537PX3
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

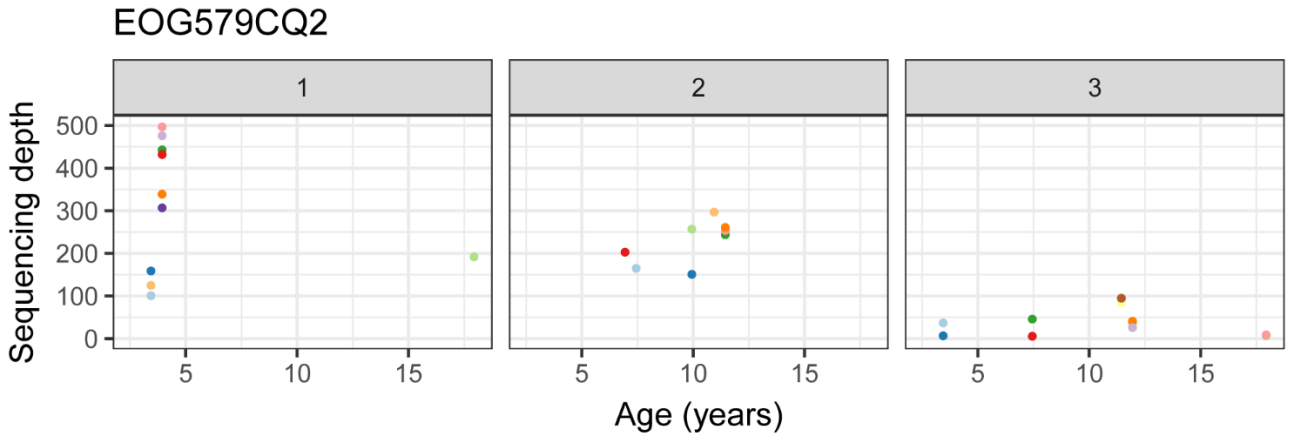

EOG59320R
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

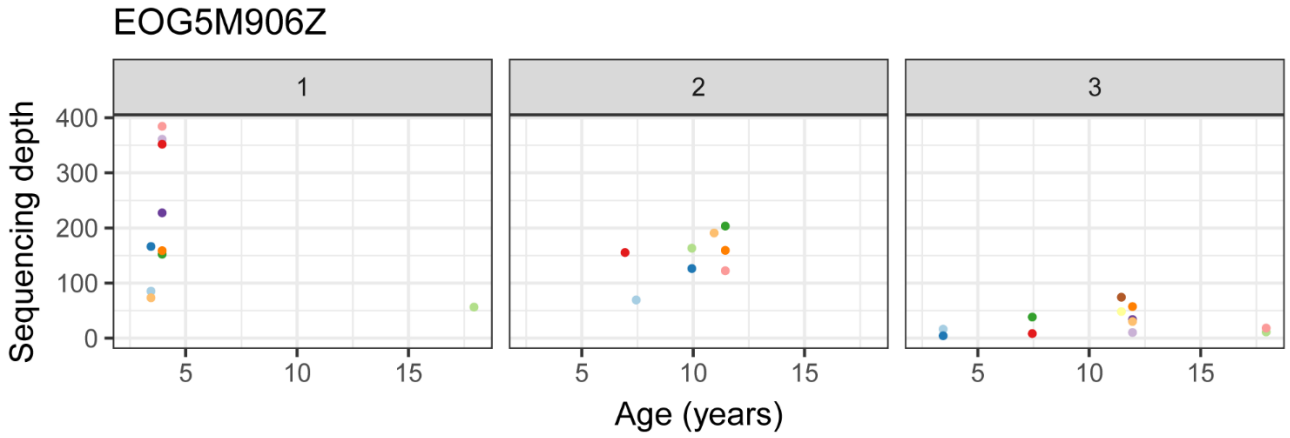

EOG5HHMHT
1
2
3
Sequencing depth
150
100
50
0
5
10
15
Age (years)

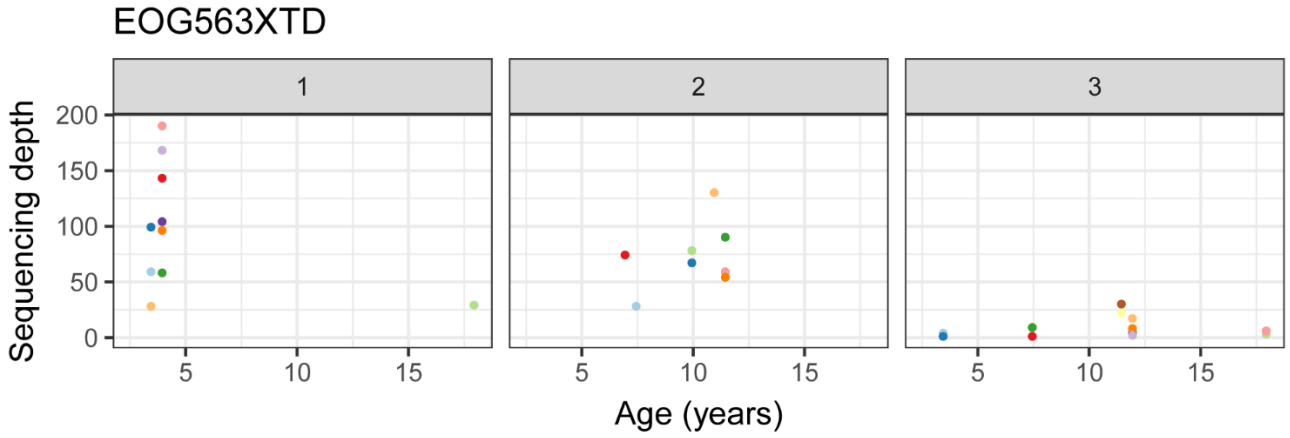

EOG5M0CH6
1
2
3
Sequencing depth
200
150
100
50
0
5
10
15
Age (years)

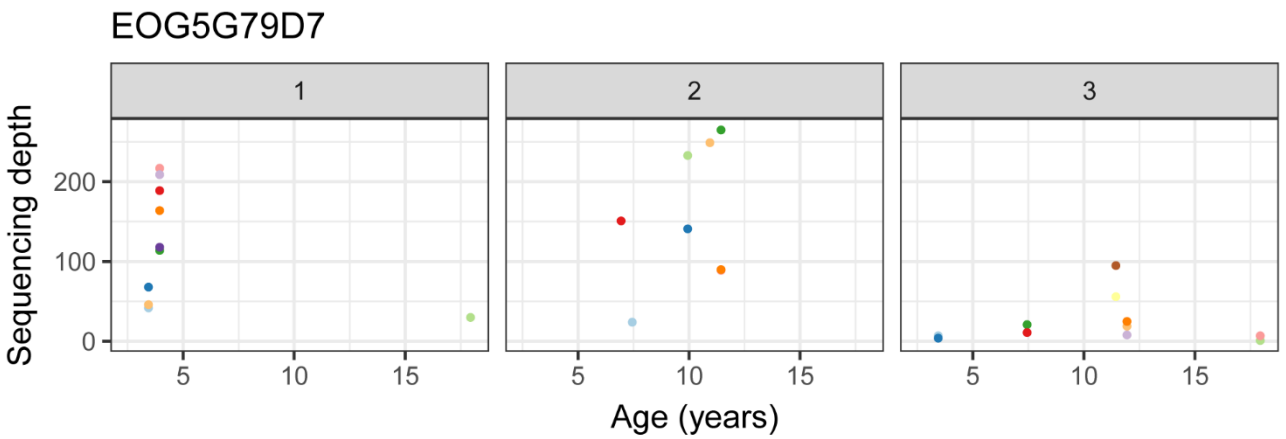

EOG541NSV
1
2
3
Sequencing depth
250
200
150
100
50
0
5
10
15
Age (years)

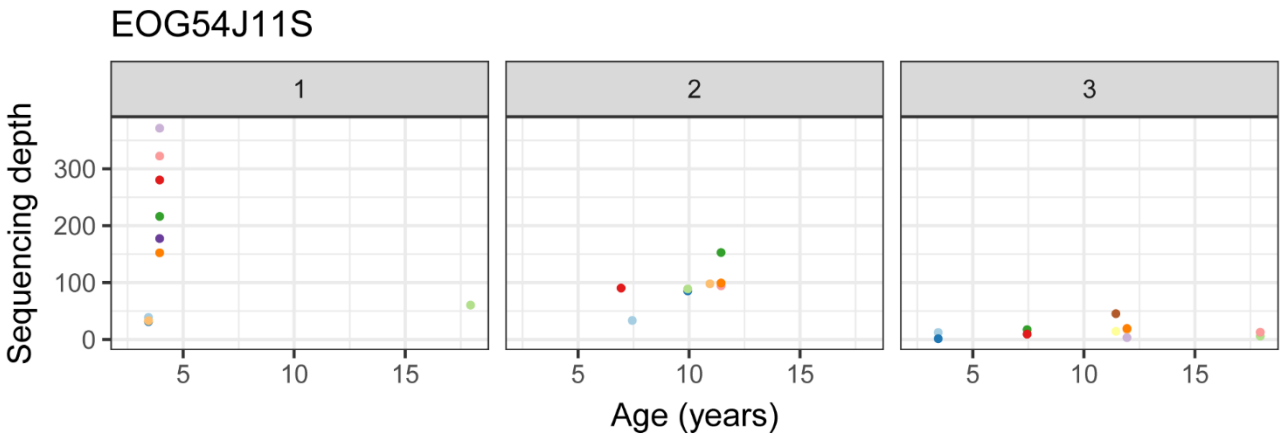

EOG5BK3JZ
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)

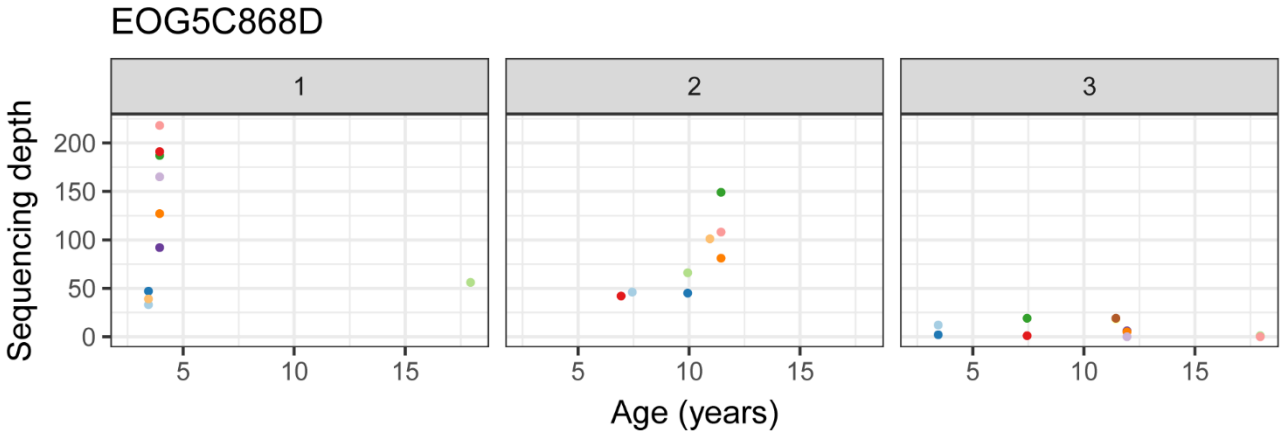

EOG579CQ2
1
2
3
Sequencing depth
0
100
200
300
400
500
5
10
15
Age (years)

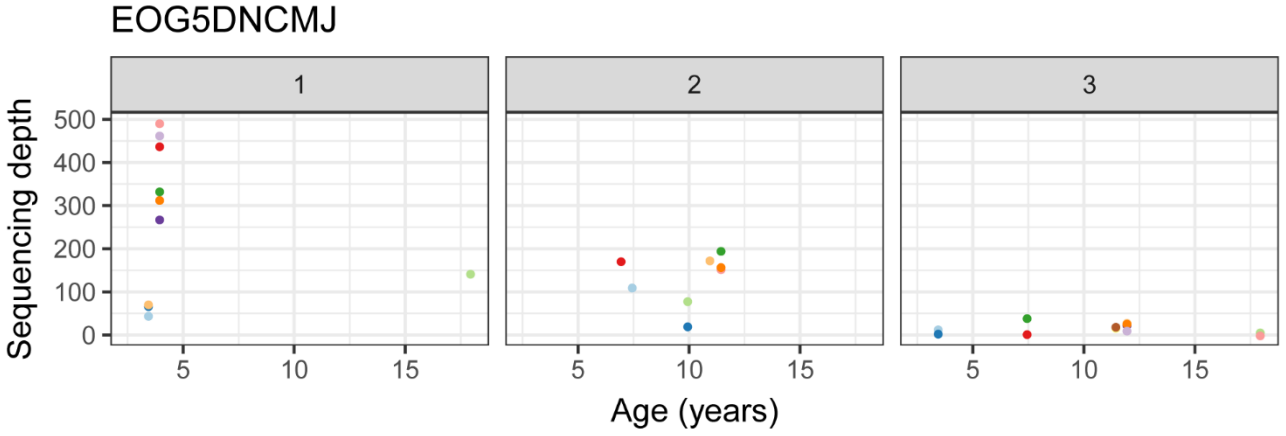

EOG5M906Z
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)

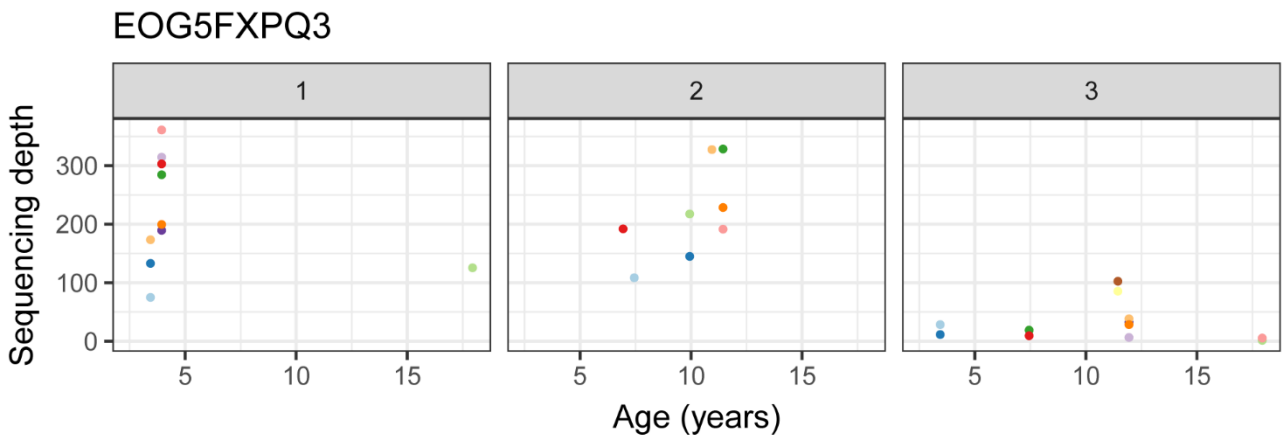

EOG563XTD
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)

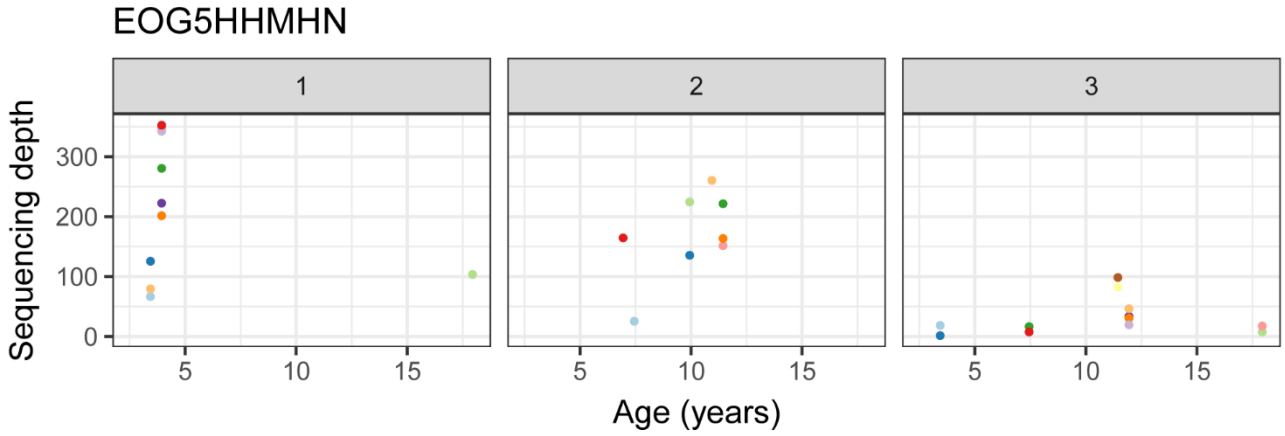

EOG5G79D7
1
2
3
Sequencing depth
0
100
200
5
10
15
Age (years)

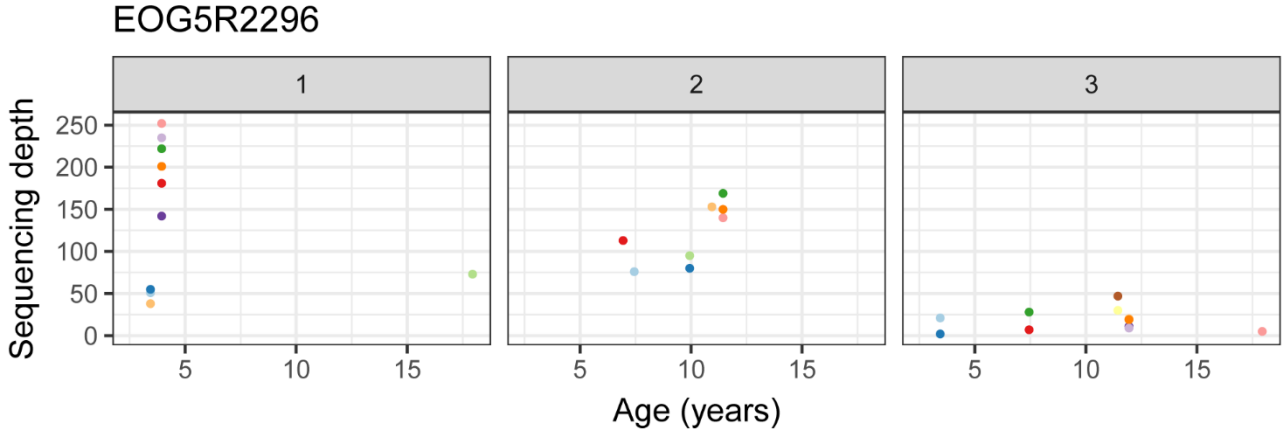

EOG54J11S
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)

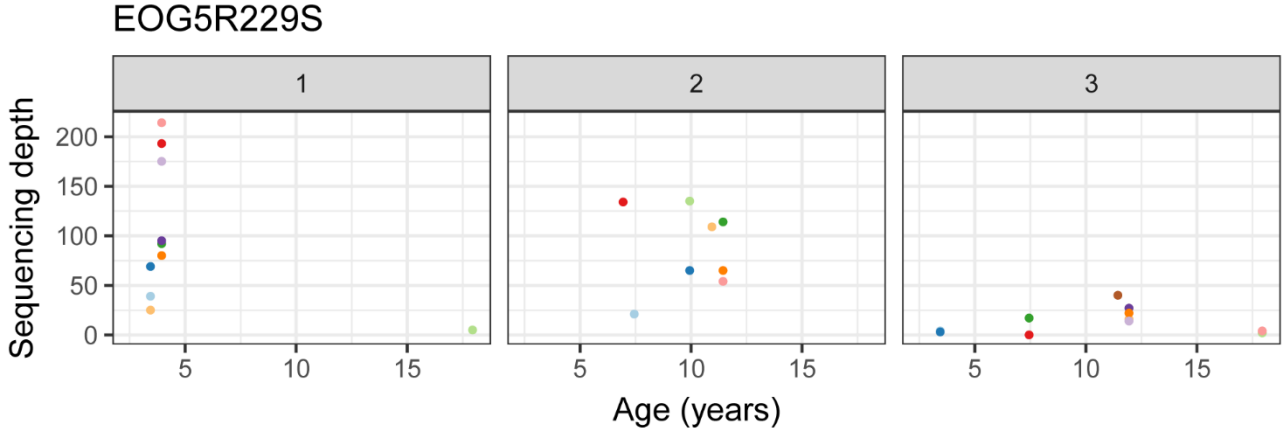

EOG5C868D
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)

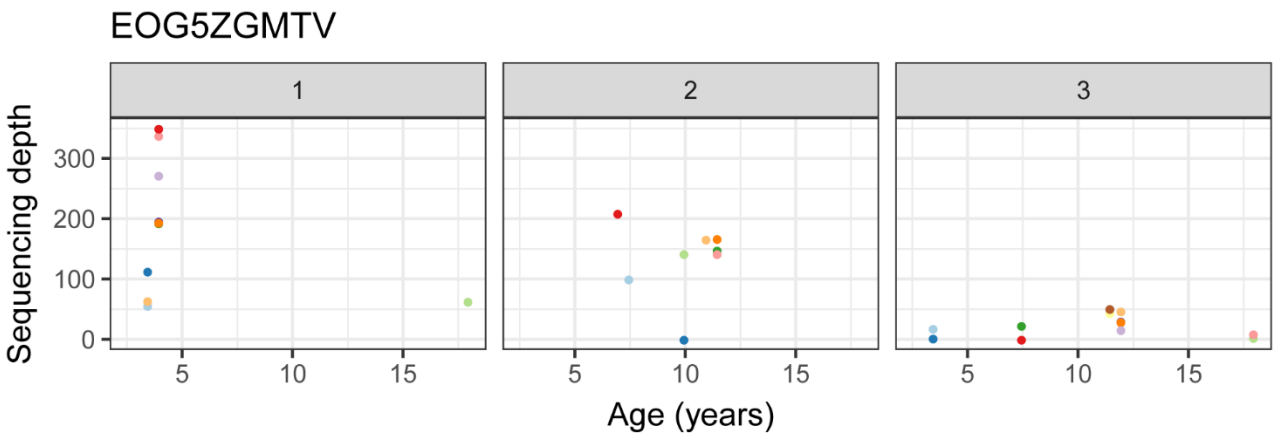

EOG5DNCMJ
1
2
3
Sequencing depth
0
100
200
300
400
500
5
10
15
Age (years)

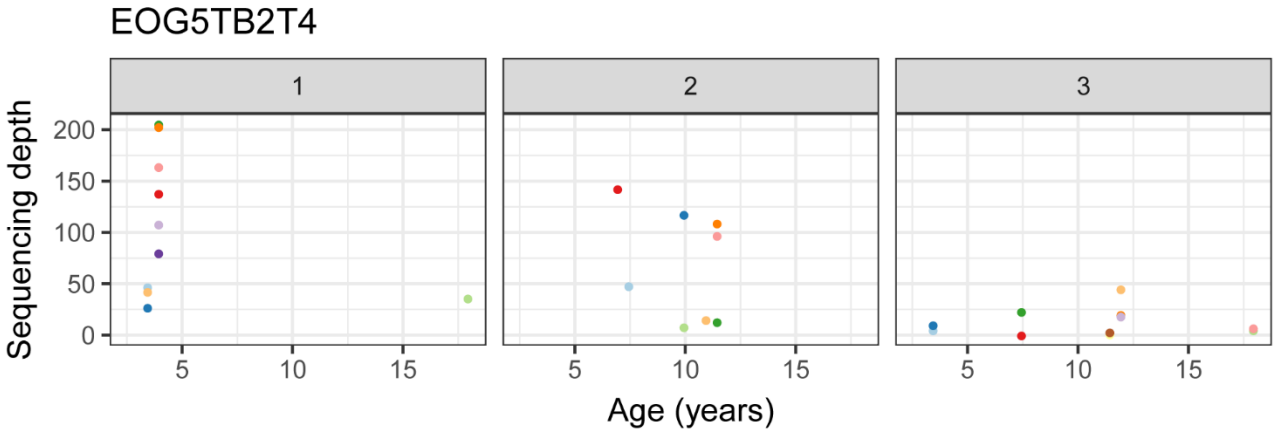

EOG5FXPQ3
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)

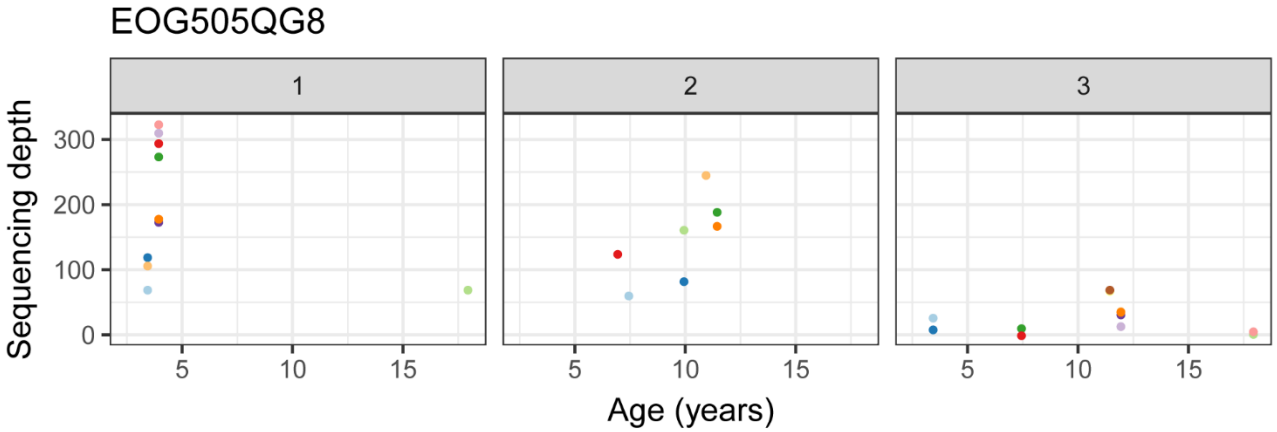

EOG5HHMHN
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)

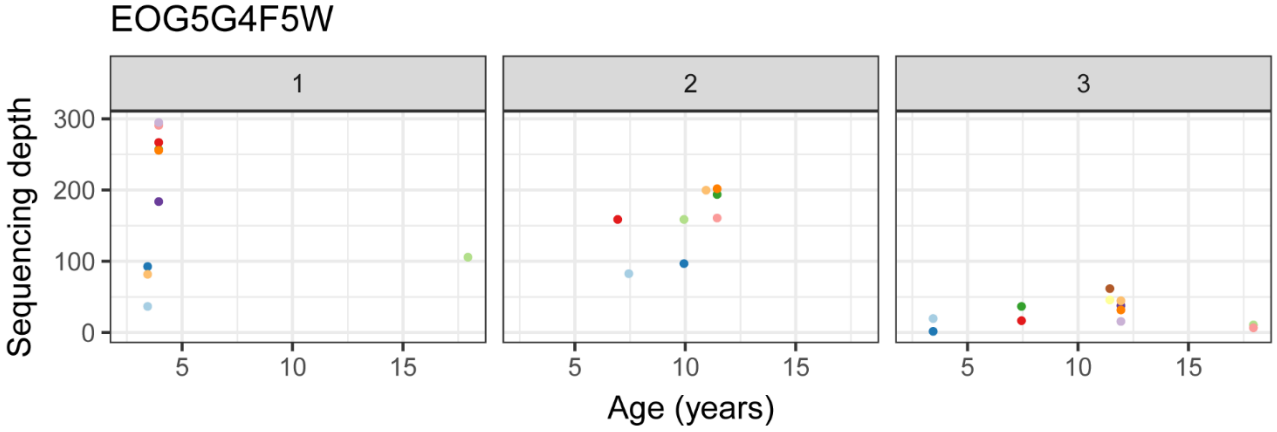

EOG5R2296
1
2
3
Sequencing depth
0
50
100
150
200
250
5
10
15
Age (years)

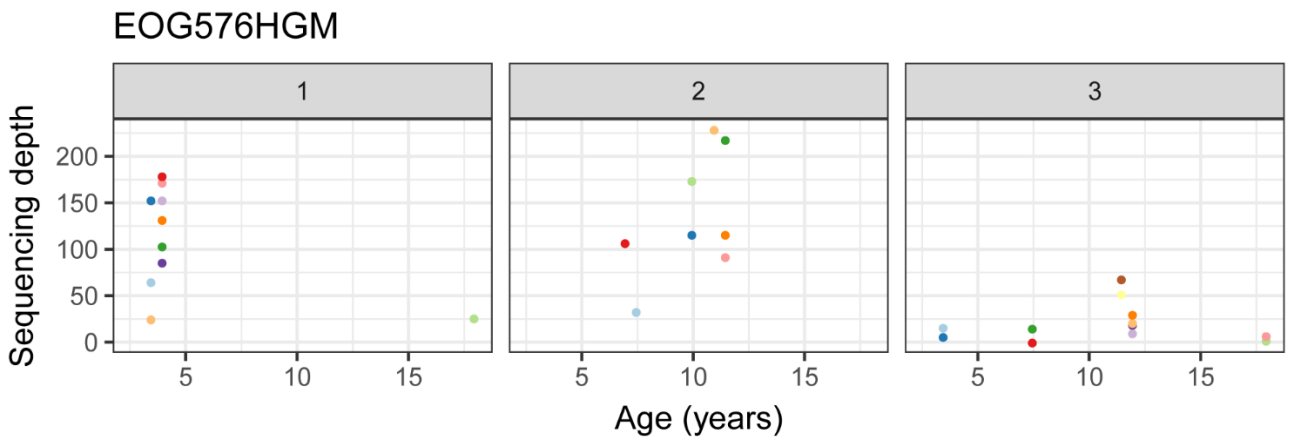

EOG5R229S
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)

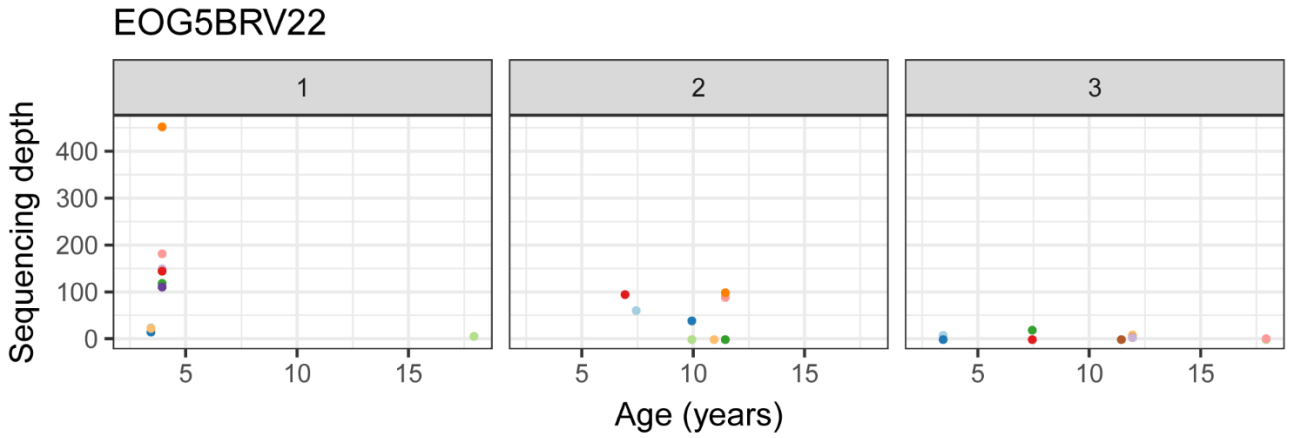

EOG5ZGMTV
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)

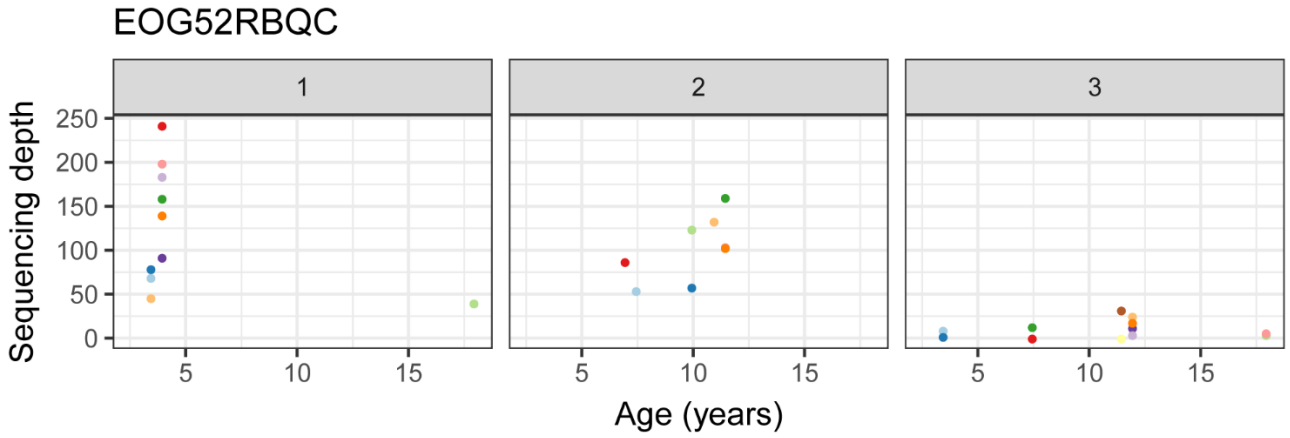

EOG5TB2T4
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)

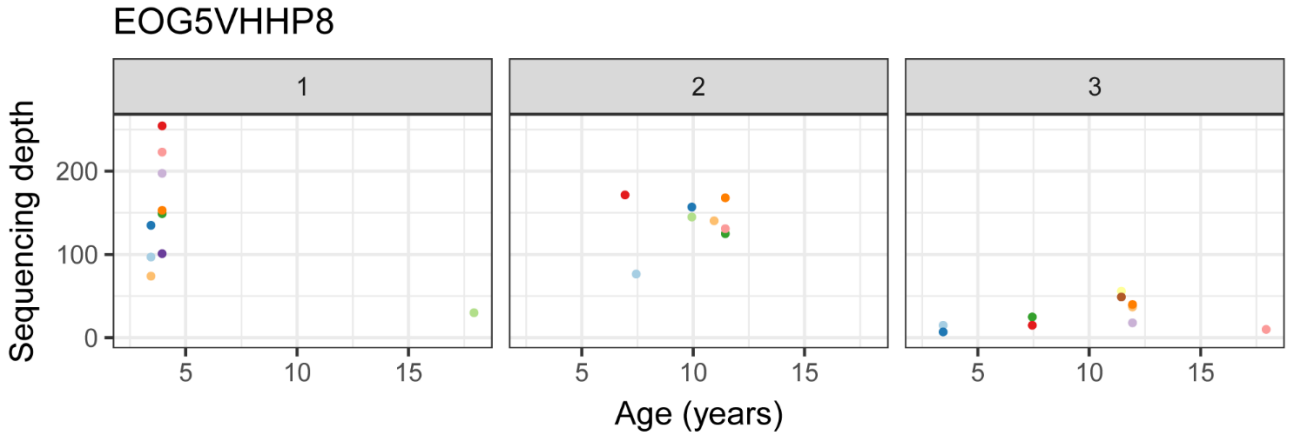

EOG505QG8
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)


EOG5G4F5W
1
2
3
Sequencing depth
0
100
200
300
5
10
15
Age (years)


EOG576HGM
1
2
3
Sequencing depth
0
50
100
150
200
5
10
15
Age (years)


EOG5BRV22
1
2
3
Sequencing depth
0
100
200
300
400
5
10
15
Age (years)


EOG52RBQC
1
2
3
Sequencing depth
0
50
100
150
200
250
5
10
15
Age (years)


EOG5VHHP8
1
2
3
Sequencing depth
0
100
200
5
10
15
Age (years)

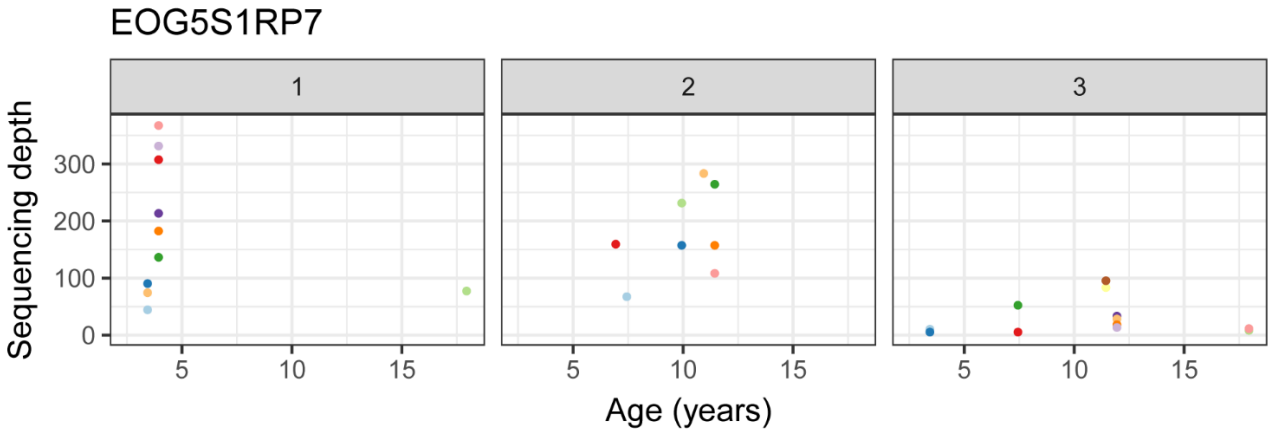
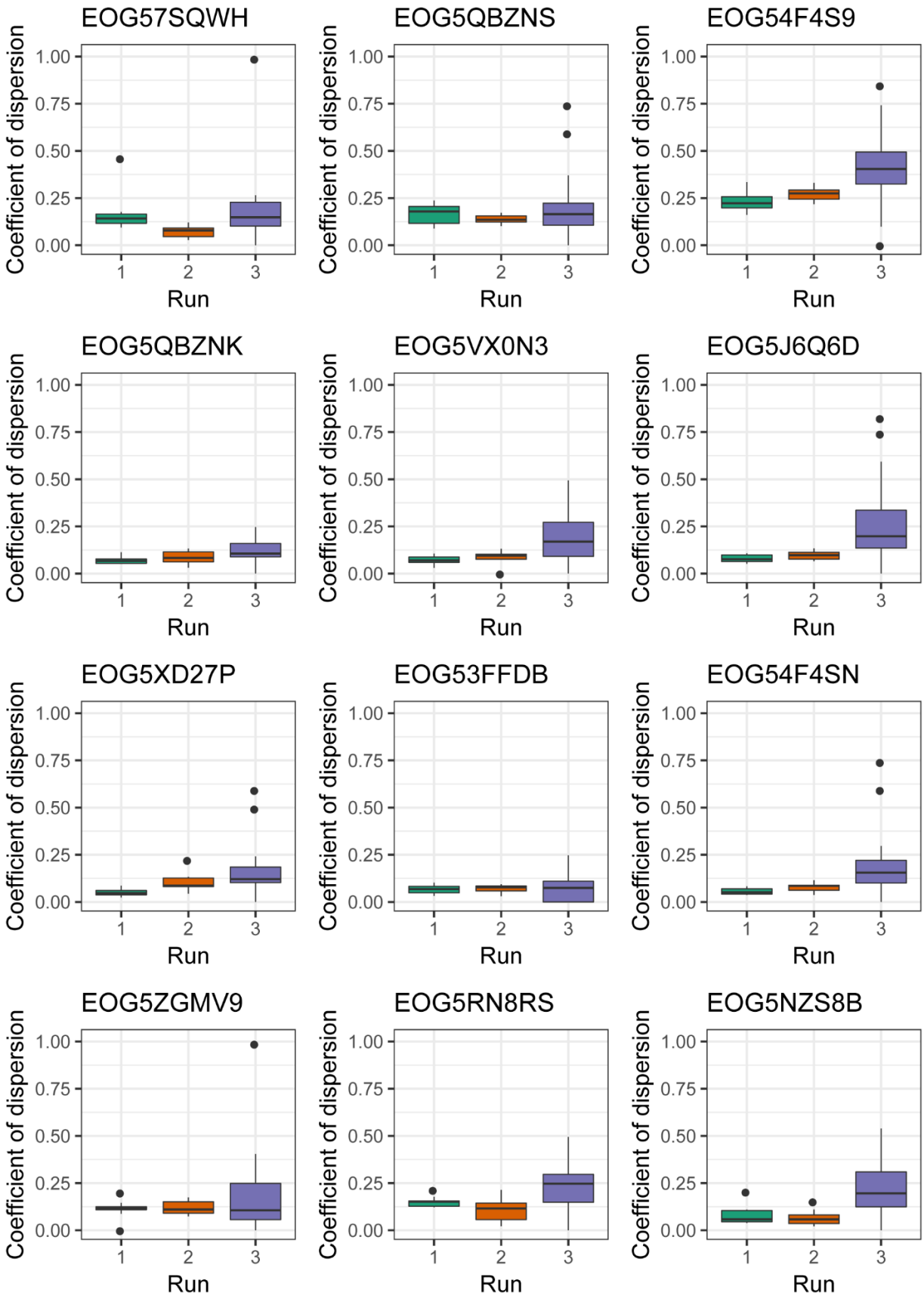
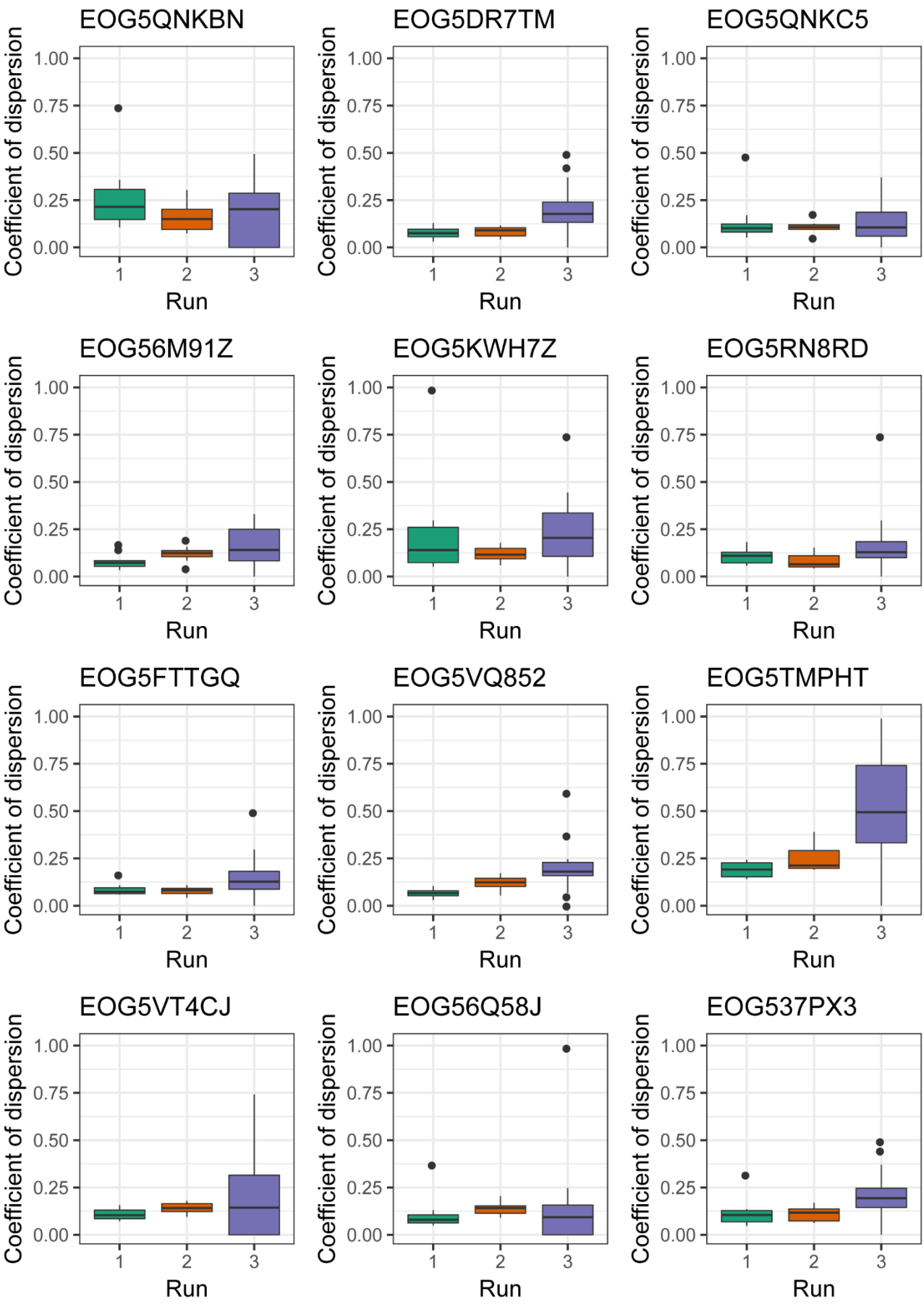
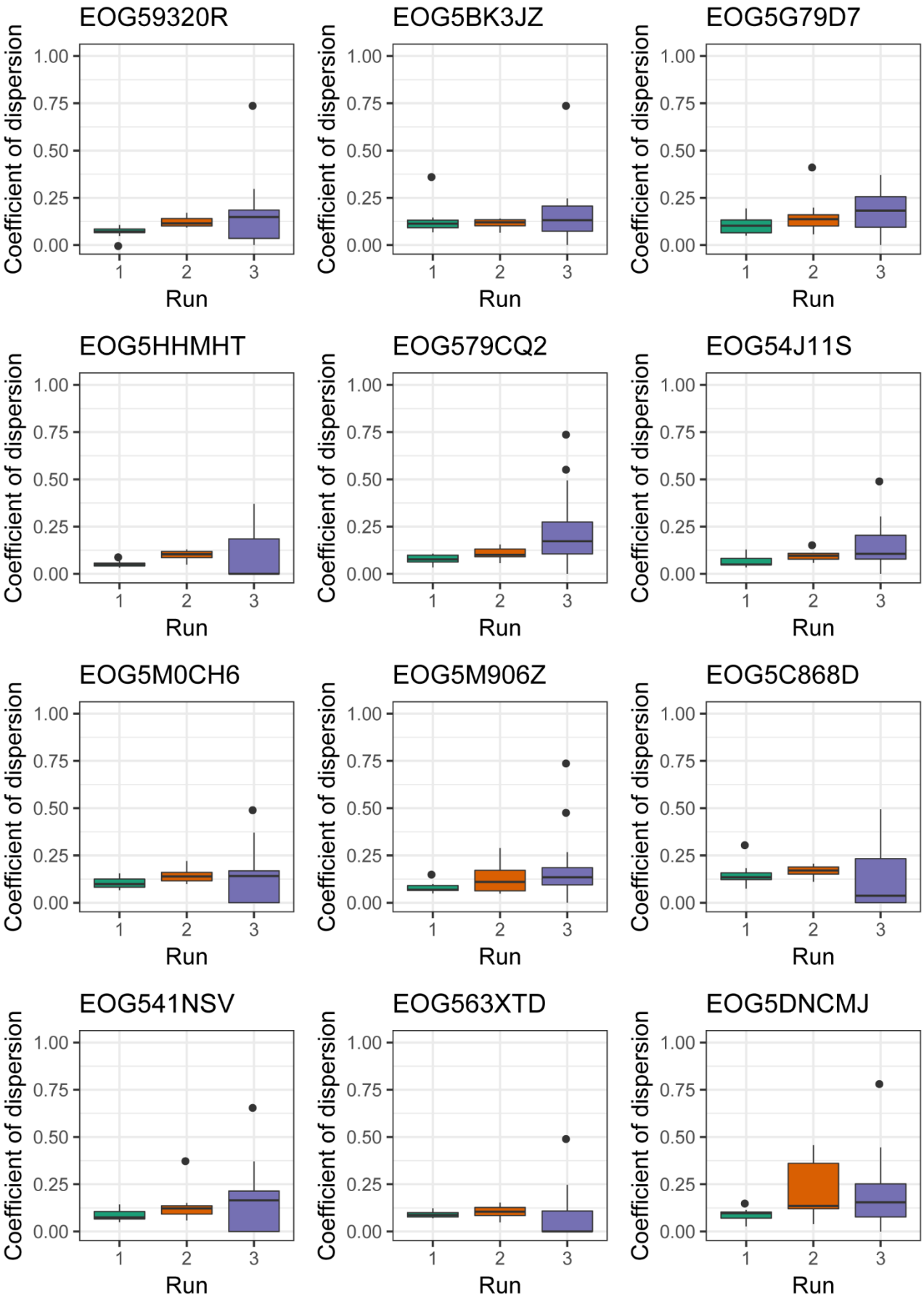
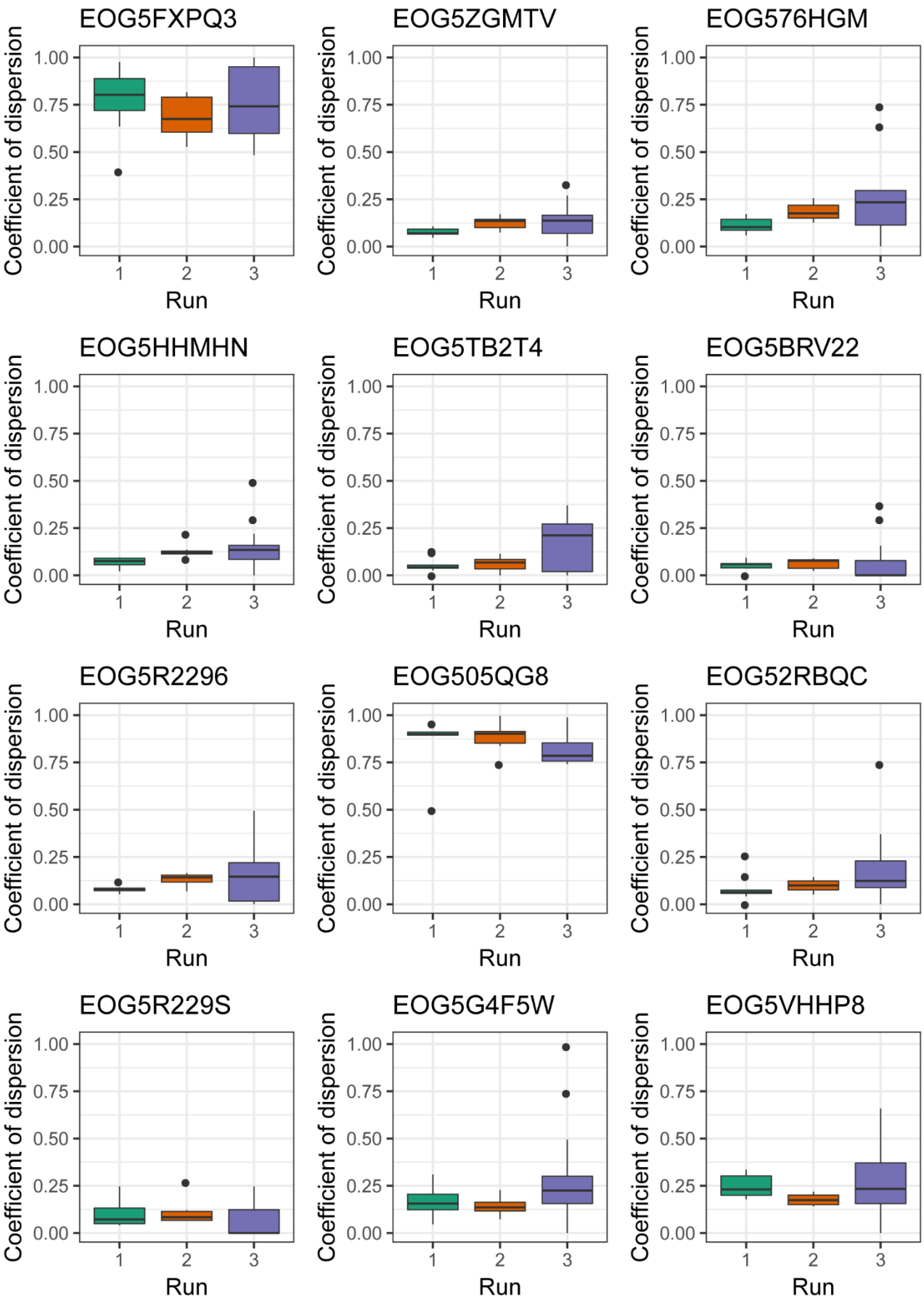

EOG5S1RP7
1
2
3
Sequencing depth
300
200
100
0
5
10
15
Age (years)

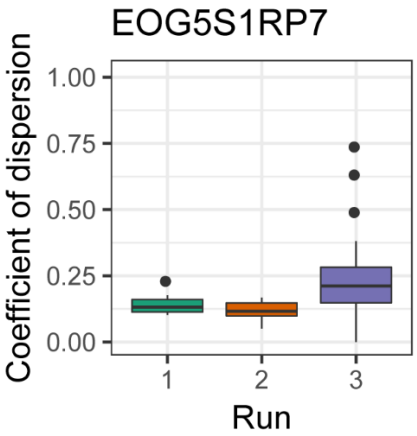
Figure S9: Distribution of the robust coefficient of dispersion (CD) across all samples at each exon for 50 randomly targeted orthologues, separated by sequencing run (run 1 depicted in green, run 2 in orange, and run 3 in purple). Orthologue code is specified above plots. Horizontal lines show median CD, vertical lines depict boxplot whiskers, and solid points represent outliers.




EOG5QNKBN
EOG5DR7TM
EOG5QNKC5
EOG56M91Z
EOG5KWH7Z
EOG5RN8RD
EOG5FTTGQ
EOG5VQ852
EOG5TMPHT
EOG5VT4CJ
EOG56Q58J
EOG537PX3
Coefficient of dispersion
0.00
0.25
0.50
0.75
1.00
1
2
3
Run


EOG59320R
EOG5BK3JZ
EOG5G79D7
EOG5HHMHT
EOG579CQ2
EOG54J11S
EOG5M0CH6
EOG5M906Z
EOG5C868D
EOG541NSV
EOG563XTD
EOG5DNCMJ
Coefficient of dispersion
0.00
0.25
0.50
0.75
1.00
Run
1
2
3


EOG5FXPQ3
EOG5ZGMTV
EOG576HGM
EOG5HHMHN
EOG5TB2T4
EOG5BRV22
EOG5R2296
EOG505QG8
EOG52RBQC
EOG5R229S
EOG5G4F5W
EOG5VHHP8
Coefficient of dispersion
Run
0.00
0.25
0.50
0.75
1.00
1
2
3


EOG5S1RP7
Coefficient of dispersion
1.00
0.75
0.50
0.25
0.00
1
2
3
Run

Figshare files for Chapter 2

These additional files are publicly available on the Figshare online repository.

Six assembled isopod transcriptomes (doi:10.25909/5d3674926d717)

- Armadillidium_vulgare.IDBA_tran.transcript-60.fa
- Ceratothoa_sp.IDBA_tran.transcript-60.fa
- Halonisc.fas
- Paraplatyarthrus_spG1_EST.fas
- Paraplatyarthrus_subterraneusG2_EST.fas
- Porcellionides_pruinosusG3_EST.fas

MARE information content results (doi:10.25909/5d3678273b4f0)

- matrix_legend.png
- MARE_OGsIC.xls
- matrix_unred_sort.png

Selected 469 orthologous groups (doi:10.25909/5d3672cf76c28)

- 469 orthologous groups (nucleotide alignments)

Bait design (doi:10.25909/5d3548f059aed)

- coverted_baits_maxseq

Alignments and partition files (doi:10.25909/5d354467c921e)

- DatasetA_concatenated_alignment_fullexons25.fasta
- DatasetA_alignment_partitions_ fullexons25.txt
- DatasetB_concatenated_alignment_fullexons50.fasta
- DatasetB_alignment_partitions_ fullexons50.txt
- DatasetC_concatenated_alignment_fullexons75.fasta
- DatasetC_alignment_partitions_ fullexons75.txt

Example dataset for use with R script (doi:10.25909/5d3ba3b7b1b2d)

- bases_allCombined_exampledataset.rds

Supplementary Material for Chapter 3

Table S1: Taxon sampling for exon capture with detailed collection data. Specimens included in the StarBEAST2 species tree analysis (SC1) are also listed here. Note: the sample from Windimurra (Yilgarn, WA) consists of pooled DNA extracts from three individuals (BES identifiers specified).

Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by	StarBEAST2
Ngalia Basin, Northern Territory										
Philosciidae	*Haloniscus*	sp.	BES18775	27809	Robb's Bore, Newhaven Sanctuary	-22.7314	131.0867	13.09.2015	W. Humphreys, S. Cooper & D. Stringer	Y
Philosciidae	*Haloniscus*	sp.	BES18774	27810	Homestead Bore, Newhaven Sanctuary	-22.7252	131.1664	14.09.2015	W. Humphreys, S. Cooper & D. Stringer	N
Philosciidae	*Haloniscus*	sp.	BES18754	27817	Sullivan's Well, Napperby Station	-22.7361	132.4610	16.09.2015	W. Humphreys, S. Cooper & D. Stringer	Y
Philosciidae	*Haloniscus*	sp.	BES18773	1	Camel Bore, Newhaven Sanctuary	-22.9344	131.2397	14.09.2015	W. Humphreys, S. Cooper & D. Stringer	Y
Philosciidae	*Haloniscus*	sp.	BES18759.3	2	Rabbit Hole Well, Central Mt Wedge	-22.7178	132.3239	16.09.2015	W. Humphreys, S. Cooper & D. Stringer	N
Philosciidae	*Haloniscus*	sp.	BES6667.2	17	Gurner Bore, Newhaven Sanctuary	-22.7160	130.9842	14.06.2001	W. Humphreys & A. Russ	N
Great Artesian Basin Springs, South Australia										
Philosciidae	*Haloniscus*	*rotundus*	GAB01433	28077	Kingfisher, Dalhousie Springs	-26.4083	135.5216	7.07.2009	M. Guzik, R. King & L. Harsche	Y
Philosciidae	*Haloniscus*	*yardiyaensis*	GAB01616	28078	Freeling South Springs, Lake Eyre	-28.0733	135.9036	3.07.2009	M. Guzik, R. King & L. Harsche	Y
Philosciidae	*Haloniscus*	*fontanus*	GAB00736	28079	Strangways Springs, Lake Eyre	-29.1622	136.5517	1.11.2007	M. Guzik & N. Murphy	N
Philosciidae	*Haloniscus*	*fontanus*	GAB00795	11	Old Finnis Springs, Hermit Hills, Lake Eyre	-29.5832	137.4408	4.11.2007	M. Guzik & N. Murphy	N
Philosciidae	*Haloniscus*	*fontanus*	GAB00765	12	Bubbler Spring, Coward, Lake Eyre	-29.4464	136.8580	3.11.2007	M. Guzik & N. Murphy	N
Philosciidae	*Haloniscus*	*fontanus*	GAB01007.1	28082	Hawker Springs, Neales, Lake Eyre	-28.4251	136.1861	27.08.2008	M. Guzik & N. Murphy	Y
Philosciidae	*Haloniscus*	*microphthalmus*	GAB00764.1	28080	Francis Swamp Springs, Lake Eyre	-29.0797	136.2769	3.11.2007	M. Guzik & N. Murphy	Y

Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by	StarBEAST2
Yilgarn, Western Australia										
Philosciidae	*Haloniscus*	sp.	BES18659	27814	Shady Well, Laverton Downs, Carey Drainage	-28.4074	122.2038	21.04.2015	W. Humphreys, S. Cooper & J. Hyde	Y
Philosciidae	*Haloniscus*	sp.	BES18601	27816	Quandong Bore, Laverton Downs, Carey Drainage	-28.3393	122.2097	20.04.2015	W. Humphreys, S. Cooper & J. Hyde	N
Philosciidae	*Haloniscus*	sp.	BES18645	27812	Laverton South, Laverton Downs, Carey Drainage	-28.5161	122.1833	21.04.2015	W. Humphreys, S. Cooper & J. Hyde	Y
Philosciidae	*Haloniscus*	sp.	BES18644	27818	Laverton South, Laverton Downs, Carey Drainage	-28.5170	122.1813	21.04.2015	W. Humphreys, S. Cooper & J. Hyde	N
Philosciidae	*Haloniscus*	sp.	BES16434	28076	Lake Miranda East, Carey Drainage	-27.6792	120.6022	23.10.2011	W. Humphreys & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES17062	28081	Bubble Well, Millbillillie, Carey Drainage	-26.5607	120.0408	15.05.2012	W. Humphreys & S. Cooper	N
Philosciidae	*Haloniscus*	sp.	BES16348	4	Bubble Well, Millbillillie, Carey Drainage	-26.5607	120.0409	21.10.2011	W. Humphreys & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES6655	3	Yuinmery South, Raeside Drainage	-28.5486	119.0911	15.05.2001	W. Humphreys, C. Watts & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES8623.1	5	Three Rivers Plutonic, Gascoyne Drainage	-25.2831	119.1757	26.08.2001	W. Humphreys, T. Karanovic & J. Waldock	N
Philosciidae	*Haloniscus*	sp.	BES13246	7	Mt Morgans, Carey Drainage	-28.7318	122.1569	10.05.2007	W. Humphreys & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES13396	8	Lake Uramurdah, Carey Drainage	-26.6877	120.3528	16.05.2007	W. Humphreys & S. Cooper	N
Philosciidae	*Haloniscus*	sp.	BES14385	9	Gum Well, Perrinvale, Raeside Drainage	-28.7750	120.4170	8.05.2007	W. Humphreys & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES13314	10	Lake Uramurdah, Carey Drainage	-26.6876	120.3027	16.05.2007	W. Humphreys & S. Cooper	Y
Philosciidae	*Haloniscus*	sp.	BES6601.2	15	Jundee South Hill, Gascoyne Drainage	-26.2688	120.6809	11.05.2001	W. Humphreys, C. Watts & S. Cooper	Y

Family	Genus	Species	Specimen ID	Sequencing ID	Locality	Latitude	Longitude	Collection date	Collected by	StarBEAST2
Philosciidae	*Haloniscus*	sp.	BES10410	16	Lake Mason, Raeside Drainage	-27.5400	119.6243	30.05.2004	W. Humphreys, C. Watts & C. Clay	Y
Philosciidae	*Haloniscus*	sp.	BES13452	18	Lake Violet, Carey Drainage	-26.6876	120.2866	16.05.2007	W. Humphreys & S. Cooper	N
Philosciidae	*Haloniscus*	sp.	BES8956/BES 13133.1, BES13133.2	19	Windimurra, Murchison Drainage	-28.2860	118.5754	31.08.2001/ 24.10.2004	R. Leijs/W. Humphreys & S. Cooper	Y
Salt lake species, Western Australia										
Philosciidae	*Haloniscus*	*searlei*	BES6573	27811	Lighthouse Swamp, Rottnest Island	-32.0000	115.5000	1.04.2001	W. Humphreys	Y
Outgroups										
Armadillidae	*Troglarmadillo*	sp.	BES15537.2	23	Lake Miranda West, Yilgarn, WA	-27.7467	120.5266	00.07.2010	W. Humphreys & S. Cooper	N
Paraplatyarthridae	*Paraplatyarthrus*	*subterraneus*	BES15525.10	22	Laverton South, Laverton Downs, Yilgarn, WA	-28.5028	122.1773	13.07.2010	W. Humphreys & S. Cooper	Y
Paraplatyarthridae	*Paraplatyarthrus*	sp.	BES16400.2	6	Halfpenny Well, Yilgarn, WA	-27.6966	121.3395	21.10.2011	W. Humphreys & S. Cooper	N
Paraplatyarthridae	Gen.	sp.	Ja243	20	Porto Alegre, Belém Novo, Rio Grande do Sul, Brazil	-30.2086	-51.1697	19.03.2011	D. Kenne & I. Campos Filho	Y

Table S2: Details for all models tested in independent StarBEAST2 analyses. The model in grey (SC1) was used for the analysis depicted in Figure 3.

Run	Exons	Taxa	*COI*?	Clock model	Substitution model	Process	Chains	ESS (convergence?)	Calibration	Support values	Important nodes (Ma)
SC1	90	21	N	Strict	HKY	Yule	1 billion	Yes, all parameters >200	Gondwana split	Majority PP = 1, except some recent branches (4)	A: 8.98, B: 8.23, C: 5.05, F: 3.81, G: 5.00, J: 25.41, L: 1.20
SC2	90	21	N	Strict	HKY	Birth-death	1 billion	Yes, all parameters >200	Gondwana split	Majority PP = 1, except some recent branches (4)	A: 8.93, B: 8.19, C: 5.29; F: 3.83, G: 4.99, J: 25.37, L:1.22
SC3	90	21	N	Strict	HKY	Calibrated Yule	1 billion	All >200 except Prior, cySpeciationRate and CalibratedYuleModel	Gondwana split	Majority PP = 1, except some recent branches (4)	A: 8.80, B: 8.19; C: 5.21, F: 3.73: G: 4.90; J: 24.88; L: 1.20
SC4	45	21	N	Strict	HKY	Birth-death	500 million	Yes, all parameters >200	Gondwana split	Majority PP = 1, except some recent branches (5)	A: 8.74; B: 8.00; C: 5.64; F: 3.32; G: 4.85; J: 24.91; L :1.07
SC5	45	21	Y	Strict	HKY	Birth-death	500 million	Yes, all parameters >200	Gondwana split	8 unresolved branches, position of Windimurra changes	Dates much older than SC1/SC4, A: 14.22; B: 12.99; C: 9.42; F: 4.79; G: 7.68; J: 28.63; L: 1.14
UCLN1	90	21	N	Uncorrelated Lognormal	HKY	Birth-death	600 million	No (parameters <100)	Gondwana split	Failed: topology inconsistent	Failed
UCLN2	5	21	N	Uncorrelated Lognormal	HKY	Birth-death	1 billion	No (parameters <100)	Gondwana split	Majority PP > 0.9, except some recent branches (5)	A: 7.63; B: 6.48; C: 4.52; F:3.98; G: 3.83; J: 22.66; L: 1.11
UCLN3	15	11	N	Uncorrelated Lognormal	HKY	Yule	500 million	No (parameters <100)	Gondwana split	All branches PP = 1	A: 8.84, B: 5.72, E: 4.75, J: 22.97
UCLN4	60	21	Y	Uncorrelated Lognormal	HKY	Birth-death	500 million	No (parameters <100)	*COI* substitution rate 0.0125	Majority PP = 1, except some recent branches (4)	Dates younger than SC3, large HPD intervals, A: 4.42, B: 4.27; C: 2.72; F: 1.64; G: 2.46; J: 11.4; L: 0.36
UCED1	30	21	N	Uncorrelated Exponential	HKY	Birth-death	500 million	No (parameters <100)	Gondwana split	Majority of branches unresolved	Large HPD intervals, A: 8.43; B: 7.24, C: 5.03; F: 1.88; G: 3.12; L: 1.36

Run	Exons	Taxa	*COI*?	Clock model	Substitution model	Process	Chains	ESS (convergence?)	Calibration	Support values	Important nodes (Ma)
UCED2	5	21	N	Uncorrelated Exponential	HKY	Birth-death	500 million	No (parameters <100)	Gondwana split	Better resolution than UCED1; majority PP > 0.9, except some recent branches (6)	Large HPD intervals, A: 7.96; B: 8.87; C: 4.80; F: 3.56; G: 4.79; J: 27.13; L: 1.29
RLC1	10	21	N	Random Local Clock	HKY	Yule	500 million	No (parameters <100)	Gondwana split	Majority PP > 0.9, except some recentbranches (5)	A: 7.93; B: 6.46; C: 5.95; F: 3.22; G: 4.43; J: 22.42; L: 0.88

Figshare files for Chapter 3

These additional files are publicly available on the Figshare online repository.

Final concatenated exon capture alignment and exon partition files (doi:10.25909/5d3512c0a42be)

- concatenated_alignment_thresh50_exons.fasta
- alignment_partitions_thresh50_exons.txt