# THE DIVERSIFICATION, BIOGEOGRAPHY, AND BODY SIZE EVOLUTION OF AUSTRALIAN HYLAEINE AND EURYGLOSSINE BEES

Pelin Kayaalp
B. Sc. Hons.

A THESIS SUBMITTED FOR THE DEGREE DOCTOR OF PHILOSOPHY

SCHOOL OF BIOLOGICAL SCIENCES
FLINDERS UNIVERSITY OF SOUTH AUSTRALIA

September 2011

A. *Hylaeus quadratus (*Hylaeinae) head, frontal view
B. *Hylaeus quadratus (*Hylaeinae) body, lateral view
C. *Callohesma* sp. (Euryglossinae) body, dorsal view
D. *Callohesma* sp. (Euryglossinae) head, frontal view
E. *Hylaeus (Rhodohylaeus)* sp. (Hylaeinae) head, frontal view
F. *Hylaeus (Rhodohylaeus)* sp. (Hylaeinae) body, lateral view
G. *Brachyhesma* sp. (Euryglossinae), dorsal view

# CONTENTS

# SUMMARY

Patterns in diversification, biogeographical histories, and the evolution of body size in euryglossine and hylaeine bees (Colletidae) are the topics of this thesis. Central to the work is the reconstruction of molecular phylogenies, which also enables the evaluation of current taxonomic arrangements within these groups.

*Hylaeus* (Hylaeinae) is the only globally distributed colletid bee genus, with subgeneric and species-level diversity highest in Australia. I use molecular phylogenetic analyses based on COI, EF-1$\alpha$, and 28S genes comprising 3,997 nucleotides and a total of 76 Hylaeinae species and five outgroup species to show that *Hylaeus* originated in Australia about 30 Mya. Log-lineages through time plots indicate high rates of cladogenesis early on in the history of Australian *Hylaeus*, and the phylogeny indicates only two dispersal events out of Australia, both shortly after its crown age. One of these dispersals was into New Zealand, with only a minor subsequent radiation, but a single dispersal event from Australia gave rise to the approximately 450 species of *Hylaeus* outside of Australasia with no evidence for subsequent dispersals from Australia into regions outside of Australasia, nor migration back into Australia. I discuss the possibility that despite a decreasing distance between Australia and Asia over the past 30 My, successful colonisation subsequent to the initial dispersal outside of Australasia would not entail the original benefits of new niche space, but would involve the costs of low genetic variation in colonizing lineages.

The subfamily Euryglossinae is endemic to Australasia, and comprises one of the most speciose bee groups in Australia. I use molecular phylogenetic analyses based on COI, EF-1$\alpha$, and 28S genes comprising 3,020 nucleotides and a total of 24 Euryglossinae species and 18 outgroup species to explore evolutionary relationships among the major euryglossine genera and the relationship of Euryglossinae to other colletid subfamilies. My analyses confirm the sister clade relationship between Euryglossinae and the southern African Scrapterinae, with a divergence date between the two clades of approximately 50 Mya. I argue that, based on that divergence date, the disjunct distribution of the subfamilies are unlikely to be a legacy of Gondwanan rifting, and I

discuss two alternative routes of dispersal between Africa and Australia. A northern Eurasian route would mean that extensive extinction has taken place throughout Eurasia and northern Africa, and a southern dispersal route is considered more likely, either between Africa and Australia via Antarctica, or with a most recent common ancestor in Antarctica with separate dispersals into Australia and Africa. Two clades within the Euryglossinae that had previously only been recognised from wing morphology are shown to have diverged ca. 45 Mya. Log-lineage through time plots show that diversification in the Australian Euryglossinae has been constant over time, and despite their abundance in Australia (approximately 400 species), only three species of Euryglossinae are known from outside Australia. I discuss the potential influence of nesting substrate (i.e., wood or soil) on dispersal ability and on the distributions and abundance of bees in tropical and Mediterranean climates as one possible explanation for this distribution.

As well as being two of the most abundant groups of bees in Australia, Euryglossinae and Hylaeinae also include some of the smallest species of bees in Australia. A generalised least- squares technique for inferring changes in body size in these two subfamilies found that phylogenetic signal is evident to only a small degree in the body size evolution within these two groups. There is some indication that evolution of body size might be adaptive in both groups, but the results are not strongly supported. It has been suggested that the relatively small size of species in both these subfamilies has led to the loss of pollen-collecting scopae. That Euryglossinae and Hylaeinae are not sister-clades infers that the loss of scopae and internal pollen transport has evolved twice at a higher taxonomic scale (i.e. subfamily) in bees and only within the Colletidae. The thesis discusses an alternative possibility for the loss of scopae, which also provides a possible clue to why these bees are so successful in Australia; that being that the nectar-rich larval provisions utilised by all colletids can easily be collected from the abundant and nectar-rich Myrtaceae species in Australia. It may be that euryglossines and hylaeines have adapted to swallowing both pollen and nectar with the high quantity of nectar used in provisioning allowing for a sufficient amount of liquid for regurgitating both materials back at the nest, thus negating the need for pollen-collecting scopae.

This thesis provides further insight into the evolutionary histories of two of Australia's most diverse and abundant bee groups by highlighting their similarities and differences in diversification and life history.

# DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Pelin Kayaalp

19 September, 2011

# GENERAL INTRODUCTION

Phylogenetic studies can now be undertaken with relative efficiency using DNA sequence data and, consequently, computational tools for exploring phylogenetic patterns have become widely available. Together these facets have accelerated the evaluation of underlying factors that might influence biological diversity. Answering these questions is becoming even more important with the increasing value that is placed on biodiversity and the growing concern that some of this biodiversity may be under threat.

Bees are one of the most important pollinators of angiosperms and their evolutionary histories are presumably intertwined (Michener 1979; Grimaldi 1999; Danforth et al. 2006). Along side this, recent concerns about pathogens affecting the honeybee *Apis mellifera* and the decline of honeybee populations in many parts of the world may require that humans rely more heavily on native bees to deliver agricultural pollination services in the future, and studies focusing on native bee groups will become ever more important. Despite the vast taxonomic and behavioural information that has been accumulated, an understanding of the origin and biogeography of many bee lineages is only now beginning to emerge through the use of molecular-based phylogenetics.

Bees, or super-family Apoidea, are part of the largest order of insects, Hymenoptera, and along with wasps and ants make up the sub-order Apocrita. The fossil record for bees is very patchy but Burmese amber has provided evidence for the existence of bee-like forms at least 90 Mya (Poinar and Danforth 2006). The only other Cretaceous-aged bee fossil is closely related to extant Meliponini clades and has been dated at approximately 65 Mya (Engel 2000). The majority of bee fossils are otherwise limited to Baltic and Dominican amber (55 - 10 Mya) (e.g., Michener and Poinar 1996; Engel 1999, 2001; Michez et al. 2007; Michener 2007). Using fossils from these strata as calibration points for minimum divergence times, exploration of patterns and periods of bee diversification has been made easier through methods for converting molecular-based phylogenies into chronograms where branch lengths correspond approximately to time. With the emergence of techniques of estimating divergence dates from molecular phylogenies,

evidence from different bee groups is accumulating and holds the promise of eventually creating a comprehensive history of bee evolution. Focus has been on evolutionary relationships, divergence times, diversification, and the origins and evolution of sociality.

These studies have altered our understanding of the higher-level relationships of bees and have inferred places and times of origins of many extant bee groups. Much of the recent research into the historical biogeography of bees has focused on groups of long tongued bees and an African origin has been inferred for the apid tribes Allodapini, Ceratinini and Ctenoplectrini, with the earliest dispersals out of that continent all estimated to have occurred around 50-20 mya. The Australian Allodapini diverged from an African+Malagasy clade at least 30 mya (Schwarz et al. 2006) and their sister tribe, the Ceratinini, also originated in Africa around 50 mya with dispersals into the Americas and Eurasia soon after (Rehan et al. 2010). Today, the allodapines are largely restricted to the southern Old World, whereas ceratinines spread to all regions of the world very rapidly (Rehan et al. 2010). Ctenoplectrini are a small tribe of 19 species, ten of which are found in Asia and represent two dispersals from Africa some 30-20 mya (Schaefer and Renner 2008). Bombini (Apidae) had an Old World origin near the Eocene-Oligocene boundary and showed multiple dispersal events from the Palearctic from about 20 Mya to the Nearctic and then two dispersals into the Neotropical region, both less than 8 Mya (Hines 2008). Meliponini (Apidae) are inferred to have an origin approximately 80 Mya, with major radiations across the tropics 60-30 Mya (Ramussen and Cameron 2010).

Results from family-level bee phylogenies using sequence data indicate a major change from the more traditional treatment of Colletidae as sister-group to all other extant bees (Danforth et al. 2006; Brady et al. 2009). Melittidae *sensu lato* (Melittidae *sensu stricto*, Dasypodaidae, Meganomiidae) are now thought to be sister group to the remaining bees (Danforth et al. 2006), which suggests an African origin of the bees. The long-tongued bees (i.e., Apidae and Megachilidae) are sister to the short-tongued bees (excluding Melittidae), with Colletidae relatively derived (Danforth et al. 2006; Brady et al. 2009). Those results are supported by more recent family level age estimates of Apidae at 105-

115 Mya (Cardinal et al. 2010), and of Colletidae approximately 70 Mya (Almeida et al. in press).

The colletid bees are far more abundant and diverse in the southern hemisphere than in the northern hemisphere (Michener 2007), and this led to a general conjecture that the group could be 'Gondwanan' in origin (Michener 1979). This presumption that they are a vicariant group and the apparent similarities in glossal morphology to the sphecoid wasps led to the inference that the colletids and the related family Stenotritidae may comprise a sister clade to all other bees. The Perkins-McGinley hypothesis constitutes a counter argument to this, whereby Perkins (1912) and McGinley (1980) have separately argued three main points: (i) closer examination of the colletid glossa reveals only a superficial similarity to the sphecoid wasps; (ii) the colletids use their glossa to apply a cellophane-like nest lining, whereas the wasps do not; and (iii) males of a few colletid genera have an acutely pointed glossa. These features, along with the presence of an intron in the nuclear gene EF-1α (Brady and Danforth 2004) support the monophyly of the Colletidae and the placement of the stenotritid bees as a separate family, Stenotritidae, sister to the colletids.

The family Colletidae contains around 2,000 bee species globally (Michener 2007). Australia is the centre of diversity and abundance, where they constitute approximately half of the continent's more than 2,000 described native bee species (Exley 2001; Michener 2007). Four of the eight colletid subfamilies occur in Australia, Callomelittinae, Eurglossinae, Hylaeinae, and Neopasiphaeinae, as well as a now unassigned genus *Paracolettes*. Callomelittinae is a monogeneric subfamily of 11 species restricted to Australia (Almeida 2008). Euryglossinae (~ 400 species) is nearly entirely restricted to Australia, with only one species in New Zealand and three species in New Caledonia (Michener 1965, 2007; Donovan 1983; 2007). Hylaeinae (~ 900 species) includes the only globally distributed colletid genus, *Hylaeus*, otherwise the six other hylaeine genera comprise endemic Australian/New Guinean clades (Michener 1965, 2007). Neopasiphaeinae (~ 400 species) is restricted to Australia and South America, but with only one genus, *Leioproctus*, found on both continents. Prior to Almeida et al. (in press) Paracolletinae Cockerell 1934 was used instead of

Neopasiphaeinae Cockerell 1930. However, molecular phylogenies placed *Paracolletes* as paraphyletic to the other Paracolletinae (Almeida and Danforth 2009; Almeida et al. in press), and the use of Neopasiphaeinae was resurrected (Almeida et al. in press), with *Paracolletes* deemed as *incertae sedis* by Almeida et al. (in press). *Paracolletes* (16 species) is endemic to Australia (Almeida et al. in press).

The other four colletid subfamilies are not found in Australia. Colletinae (~ 480 described species) is most diverse in South America, with a single genus, *Colletes*, found outside of South America, and the subfamily is entirely absent from Australia, Madagascar and Southeast Asia (Michener 2007). Scrapterinae (~ 42 species) is a monogeneric subfamily, which is largely restricted to southern Africa except for one species that has a distribution to Kenya (Davies et al. 2005; Davies and Brothers 2006; Almeida and Danforth 2009). Diphaglossinae (~ 130 species) and Xeromelissinae (~120 species) are both endemic to South America (although the range of some species does extend into central America and the southern USA) (Michener 2007).

This thesis extends the available sequence data of Colletidae, with focus on two colletid subfamilies, Euryglossinae and the Hylaeinae, and attempts to examine why they are most diverse in Australia, as well as exploring the biogeography of each group. Outside of Australia, geographical distributions of the two groups are very different with Euryglossinae restricted to Australasia (throughout the thesis, the term Australasia will refer to Australia, New Zealand, and New Caledonia, but not New Guinea) and Hylaeinae found globally. Both groups, however, have an Australia-wide distribution and this thesis explores and compares diversification patterns of the Australian species of each group. Euryglossinae and Hylaeinae are the only two non-parasitic bees to exhibit extreme reduction of pollen-collecting scopae at a broad taxonomic scale. Scopae are the generally dense and branched hairs on the hind legs and/or metasomal sternites of most bees that are specialised for carrying pollen. Given the importance of pollen as a food source for larvae it is significant that two Australian bee subfamilies have lost this characteristic. To address the suggestion by Pasteels and Pasteels (1976, in Davies and Brothers 2006, p. 160) that extreme reduction of scopae in hylaeines and euryglossines is a result of their small body size I look at phylogenetic signal in body

size within each group. The thesis ends by discussing the potential influence that climatic features and changes to the climate over the past 30 My might have had on the diversification, dispersal ability, and body size evolution of hylaeine and euryglossine bees in Australia.

The first data chapter, **Rapid diversification in Australia and two dispersals out of Australia in the globally distributed bee genus, *Hylaeus* (Colletidae: Hylaeinae)** focuses on the colletid subfamily Hylaeinae**.** *Hylaeus* is the largest genus in the Hylaeinae and comprises one of the most cosmopolitan bee genera, with 600 described species on all continents except Antarctica. Four of the other Hylaeinae genera are restricted to Australia, and a sixth is found mainly in New Guinea and northern Australia. Using molecular sequence data of *Hylaeus* species from around the world, in Chapter one I test the hypotheses that: (i) *Hylaeus* originated approximately 30 Mya, in line with the estimates of previous studies, with dispersals either via Antarctica and South America or south-east Asia, (ii) diversification in Australian *Hylaeus* has been rapid because of the large number of species relative to the estimated age of the clade , and (iii) *Hylaeus* species from outside Australia diverged soon after an *Hylaeus* origin in Australia due to available paths for dispersal being more restrictive closer to the present.

The second data chapter, **Phylogenetics and diversification of the Australian bee subfamily Euryglossinae (Colletidae), and support for the southern African Scrapterinae as sister clade,** focuses on the colletid subfamily Euryglossinae. This is a subfamily that is largely restricted to Australia, with only a few species having been recorded from New Zealand and New Caledonia. Using additional species to those in previous studies, this chapter examines the phylogeny of 24 species from eight euryglossine genera and asks whether current classifications reflect phylogeny. I test the hypotheses that: (i) Euryglossinae and the southern African Scrapterinae are sister clades using a more comprehensive taxon set than earlier studies, (ii) Scrapterinae/Euryglossinae had an ancestral range in Australia, and (ii) diversification patterns in Euryglossinae differ from that of Australian *Hylaeus*, with the possibility that this is related to differences in global abundance and clade age.

The third data chapter, **Trends in euryglossine and hylaeine (Hymenoptera: Colletidae) body size evolution**, explores a fundamental morphological trait – body size - and, examines whether there is phylogenetic signal in the evolution of body size in both euryglossines and hylaeines. It also explores the tempo of evolutionary change in body size, and estimates ancestral sizes at key nodes. It has been suggested that the relatively small body size in Hylaeinae and Euryglossinae is linked to an interesting morphological trait only found universally in the Euryglossinae and Hylaeinae: the lack of pollen-carrying scopae. Scopae are pollen-collecting hairs that most bees use to collect pollen and to transport it back to their nests. These two largely Australian groups have independently evolved to carry pollen internally in the crop, which they then regurgitate when back at the nest. I test the hypothesis that body size in euryglossines and hylaeines have both evolved at the same mode and tempo, and discuss the possibility that this shared trait and the abundance of both subfamilies in Australia might be linked to the prevalence of Myrtaceae in Australia.

## CHAPTER FORMATS

This thesis presents work in the form of multi-authored manuscripts using a format similar to that for papers submitted to peer-reviewed journals. The Candidate was responsible for all of the laboratory work and nearly all of the field work. Data analysis was guided by suggestions from supervisors Associate Professor Michael Schwarz and Dr. Mark Stevens.

## LITERATURE CITED

Almeida, E.A.B. 2008. Revised species checklist of the Paracolletinae (Hymenoptera, Colletidae) of the Australian Region, with the description of new taxa. Zootaxa. 1891:1-24.

Almeida, E.A.B., and B.N. Danforth. 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50:290-309.

Almeida, E.A.B., M.R. Pie, S.G. Brady, and B.N. Danforth. In press. Biogeography and diversification of Colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. J. Biogeogr.

Brady, S.G., and B.N. Danforth. 2004. Recent intron gain in elongation factor-1α of colletid bees (Hymenoptera: Colletidae). Mol. Biol. Evol. 21(4):691–696.

Brady, S.G., L. Larkin, and B.N. Danforth. 2009. Bees, ants, and stinging wasps (Aculeata), pp 264-269 in *The Timetree of Life* Hedges, S.B. and Kumar, S., eds. Oxford Univ. Press, U.K.

Cardinal, S., J. Straka, and B.N. Danforth. 2010. Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. P. Natl. Acad. Sci. USA. 107(37):16207-16211.

Danforth B.N., S. Sipes, J. Fang, and S.G. Brady. 2006. The history of early bee diversification based on five genes plus morphology. P. Natl. Acad. Sci. USA. 103:15118-15123.

Davies, G.B.P., C. D. Eardley, and D. J. Brothers. 2005. Eight new species of *Scrapter* (Hymenoptera: Apoidea: Colletidae), with descriptions of *S. albifumus* and *S. amplispinatus* females and a major range extension of the genus. Afr. Invertebr. 46:141–179.

Davies, G.B.P., and D.J. Brothers. 2006. Morphology of *Scrapter* (Hymenoptera: Anthophila: Colletidae), with description of three new species and taxonomic status of five Cockerell taxa. Afr. Invertebr. 47:135–183.

Donovan, B.J. 2007. Apoidea (Insecta: Hymenoptera) in Fauna of New Zealand (Ko te Aitanga Pepeke o Aotearoa). ManaakinWhenua Press. Canterbury, NZ.

Engel, M.S. 1999. A new Xeromelissine bee in Tertiary amber of the Dominican Republic. Entomol. Scand. 30(4):453-458.

Engel, M.S. 2000. A New Interpretation of the oldest fossil bee (Hymenoptera: Apidae). Am. Mus. Novit. 3296.

Engel, M.S. 2001. Monophyly and extensive extinction of advanced eusocial bees: insights from an unexpected Eocene diversity. P. Natl. Acad. Sci. USA. 98:1661–1664.

Grimaldi, D. 1999. The co-radiations of pollinating insects and angiosperms in the Cretaceous. Ann. Miss. Bot. Gdns. 86(2):373-406.

Hines, H.M. 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: Bombus). Syst. Biol. 57(1):58–75.

McGinley, R.J. 1980. Glossal morphology of the Colletidae and recognition of the Stenotritidae at the family level. J. Kansas Entomol. Soc. 53:539-552.

Michener, C.D. 1965. A classification of the bees of the Australian and South Pacific regions. B. Am. Mus. Nat. Hist. 130:1-362.

Michener, C.D. 1979. The biogeography of the bees. Ann. Mo. Bot. Gard. 66(3):277-347.

Michener, C.D. 2007. The Bees of the World, 2nd edn. Johns Hopkins University Press, Baltimore, MD.

Michener, C.D., and G. Poinar Jr. 1996. The known bee fauna of the Dominican amber. J. Kansas Entomol. Soc. 69(4):353-361.

Michez, D., A, Nel, J. Menier, and P. Rasmont. 2007. The oldest fossil of a melittid bee (Hymenoptera: Apiformes) from the early Eocene of Oise (France). Zool. J. Linn. Soc. 150:701-709.

Perkins, R.C.L. 1912. Notes, with descriptions of new species, on aculeate Hymenoptera of the Australian Region. Ann. Mag. Nat. Hist. 8(9):96-121.

Poinar Jr, J.O., and B.N. Danforth. 2006. A fossil bee from early Cretaceous Burmese amber. Science. 314:614.

Rasmussen, C., and S.A. Cameron. 2010. Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. Biol. J. Linn. Soc. 99:206–232.

Rehan, S.M., T.W. Chapman, A.I. Craigie, M.H. Richards, S.J.B. Cooper, and M.P. Schwarz. Molecular phylogeny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. Mol. Phylogen. Evol. 55(3):1042-1054.

Schwarz, M.P., S. Fuller, S.M. Tierney, and S.J.B. Cooper. 2006. Molecular phylogenetics of the exoneurine allodapine bees reveal an ancient and puzzling dispersal from Africa to Australia. Syst. Biol. 55(1):31–45.

Schaefer, H., and S. Renner. 2008. Phylogeny of the oil bee tribe Ctenoplectrini (Hymenoptera: Anthophila) based on mitochondrial and nuclear data: evidence for early Eocene divergence and repeated out-of-Africa dispersal. Mol. Phylogenet. Evol. 47(2):799-811.

**CHAPTER I.**

# Rapid diversification in Australia and two dispersals out of Australia in the globally distributed bee genus, *Hylaeus* (Colletidae: Hylaeinae).

Pelin Kayaalp[1,2], Michael P. Schwarz[1], Mark I. Stevens[3]

1. School of Biological Sciences, Flinders University, GPO Box 2100, SA 5001, Australia.
2. Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa.
3. South Australian Museum, North Terrace, GPO Box 234, SA 5000, Australia, and School of Earth and Environmental Sciences, University of Adelaide, SA 5005, Australia.

RUNNING TITLE: Radiation of the bee genus *Hylaeus*.

KEYWORDS: Niche space, radiation, historical biogeography, Southern Hemisphere, bee phylogeny, Miocene.

## *Abstract*

*Hylaeus* is the only globally distributed colletid bee genus, with subgeneric and species-level diversity highest in Australia. We used one mitochondrial and two nuclear genes to reconstruct a phylogeny using Bayesian analyses of *Hylaeus* based on species from Australia, Asia, Africa, Europe, Hawai'i, the New World and New Zealand. Pairwise genetic distance estimates of *Hylaeus* compared with other bee groups indicate that *Hylaeus* originated in Australia, approximately 30 Mya. Log-lineages through time plots indicate high rates of cladogenesis early on in *Hylaeus* history, and the phylogeny indicates only two dispersal events out of Australia, both shortly after its crown age. One of these dispersals was into New Zealand with only a minor subsequent radiation, but the second dispersal out of Australia resulted in a world-wide distribution. This second dispersal and radiation event, combined with early and rapid radiation of *Hyleaus* in Australia, poses a conundrum: what kinds of biogeographical and ecological factors could simultaneously drive global dispersal, yet strongly constrain successful migration out of Australia when physical barriers appear to be weak? We propose that movement into new niches and enemy-free spaces may favour initial dispersal events, but that subsequent dispersals would not entail the original benefits of new niche space.

## *Introduction*

The radiation of the bees has been linked with that of the angiosperms (Michener 1979; Grimaldi 1999; Danforth et al. 2006), which are believed to have first appeared in Gondwana at least 130 Mya (Soltis and Soltis 2004), although a late Triassic origin for angiosperms has recently been suggested (Smith et al. 2010). The oldest bee-like insect has been dated at more than 90 Mya from Burmese amber (Poinar and Danforth 2006), and the abundance and diversity of bees in the Southern Hemisphere suggests that bees may have been present on Gondwana at the same time or earlier, when the super-continent was in the midst of rifting.

The Australian bee fauna is highly endemic, with two short-tongued families, Colletidae and Halictidae together comprising approximately 80 per cent of the estimated 2,000 bee species in Australia (Michener 1965, 2007). Although there are a relatively large number of halictid species in Australia (~30% of the total bee fauna of Australia), their diversity and endemicity is limited above the subgeneric level (Michener 1965, 2007) when compared with the Colletidae (~50% of the total bee fauna) (Michener 1965, 2007). There are numerous colletid genera endemic to Australia, and one subfamily, Euryglossinae, is found largely in Australia with only three species in New Zealand and New Caledonia (Donovan 2007). Two families that are common elsewhere throughout the globe, Andrenidae and Melittidae, are completely absent from Australia (Michener 2007) and other globally distributed groups, Nomadinae, Anthidiini and Ceratinini, are only represented by a few species restricted to north-eastern Australia (Michener 2007).

Recent molecular phylogenetic studies (Danforth et al. 2006; Almeida and Danforth 2009; Almeida et al. in press), have refuted Michener's (1979) suggestion that because of similarities in glossal morphology to the sphecoid wasps the Colletidae are a biogeographically 'old' group and sister clade to all other bees. The same studies support appraisals made by Perkins (1912) and McGinley (1980) that the colletid glossa reveals only a superficial similarity to the sphecoid wasps, and that the glossae of the species that he examined exhibit what are likely to be derived traits – implying a more recent origin. The Colletidae are inferred to have a late-stage Gondwanan origin around

70 Mya (Almeida et al. in press), and the two largely Australian colletid subfamilies, Hylaeinae (most basal genera are restricted to Australasia) and Euryglossinae (restricted to Australia with a few species on surrounding islands) are no longer inferred to be sister clades. Instead, molecular phylogenies infer that Euryglossinae is sister to the sub-Saharan Scrapterinae (Almeida and Danforth 2009; Almeida et al. in press), with dispersal between Australia and Africa inferred at approximately 54 Mya (Almeida et al. in press). Almeida et al. (in press) also argued that for the sister subfamilies Xeromelissinae (restricted to the southern New World) and Hylaeinae (Almeida and Danforth 2009) a transantarctic interchange or South American/Antarctica/Australian vicariance event most likely explained their geographical centres of diversity, with a divergence around 58 Mya.

Major geological events that may be important when considering biogeographical scenarios involving the Australian biota include (i) Australia's terrestrial connection with Africa, via Antarctica and South America, which continued until approximately 100 Mya (McLoughlin 2001); (ii) the northward drift of the Australian plate towards Asia since the Miocene (McLoughlin 2001); (iii) rifting of Australia and South America from Antarctica during the Oligocene; and (iv) the isolation of Antarctica approximately 30 Mya (Veevers and Li 1991; Sanmartín and Ronquist 2004), which led to the formation of the Antarctic Circumpolar Current, and development of the first extensive ice sheets on Antarctica (Zachos et al. 1994). Importantly, when assessing biotic relationships between Australia and South America, it is believed that Antarctica had a temperate climate and was rich with angiosperms until the early Eocene (Sanmartín and Ronquist 2004), potentially providing few barriers for temperate-adapted species to move across those three landmasses up until about 20-30 Mya, and possibly up to 12-13 Mya (Lewis et al. 2008; Stevens et al. 2006). It seems likely that both late-stage Gondwanan vicariance (i.e., at the latter stages of Gondwanan rifting, around the end of the Cretaceous ) and long range dispersal may have played a role in the current geographic dispersal of colletid bees, but disentangling these two factors requires detailed phylogenetic and historical biogeographic analyses.

The only large-scale genus-level work that has examined the diversification and biogeography of the colletid bees has been for the genus *Colletes* (Colletinae) (Kuhlmann et al. 2009). However, that study did not employ dating techniques but referred to an unpublished source for a *Colletes* origin of 32-40 Mya. In a recent colletid-wide study Almeida et al. (in press) provided an estimated crown age for Colletinae at 33 Mya and 30 Mya for *Colletes*. After *Hylaeus*, *Colletes* (containing 469 described species) is the second-most widely distributed colletid genus, but is absent from Australia, Madagascar, and South-east Asia (Michener 2007). A South American origin for *Colletes* has been inferred, with subsequent dispersal into North America (Michener 1979, Kuhlmann et al. 2009), and with five dispersals from north America into Euarasia and two dispersals from Eurasia into sub-Saharan Africa (Kuhlmann et al. 2009).

The subfamily Hylaeinae is one of the most speciose and diverse groups in the family Colletidae, and contains the only globally distributed colletid genus, *Hylaeus*. *Hylaeus* currently occurs on all continents except Antarctica (Michener 2007), a distribution that provides an interesting contrast with that of *Colletes*. The diversity and endemism of *Hylaeus* in Australia, compared to other regions, is outlined in Table 1. Twenty-two of the 46 subgenera are found in Australasia, and 17 are endemic to the Australian continent. Some distributional overlap of subgenera occurs for other regions, but the five sub-Saharan subgenera are endemic to that region. There are more than 600 described species of *Hylaeus* (Michener 2007) revealing a species richness greater than other colletid genera and other colletid subfamilies. Nearly one third of the described *Hylaeus* species (~170) are endemic to the Australasian region (Houston 1975, 1981; Michener 2007) and the oldest lineages within Hylaeinae are found in Australia (Michener 1979; Almeida and Danforth 2009; Almeida et al. in press), suggesting that the origin and diversification of the group is associated with this region. While Colletidae abundance and diversity is greatest in Australia and South America, the representation of *Hylaeus* species in the Neotropics is depauperate, with an estimated 50 species (Michener 2007). Furthermore, the crown age of *Hylaeus* is estimated at 30 Mya (Almeida et al. in press), which means that *Hylaeus* originated in Australia at a time when dispersal to South America via Antarctica was possible but problematic because of increasing ice and/or

water barriers during this period.

Here we use a wide representation of Australian and non-Australian *Hylaeus* species to explore phylogenetic relationships and identify the patterns of diversification in this clade. Specifically, we investigate (i) whether *Hylaeus* originated in Australia, (ii) the patterns of diversification of *Hylaeus* in Australia; and (iii) if there was an Australian origin, the number of dispersal events from Australia and the biogeographic scenarios that may have contributed to those dispersals.

## *Methods*

### TAXON SAMPLING

The most recent colletid phylogeny included only eleven Hylaeinae species, all of which were Australian species (Almeida et al. in press). We have collected molecular data for an additional 51 Hylaeinae species, including 11 species from outside Australia, and included sequences from another 25 Australian and non-Australian Hylaeinae species from GenBank. We included one of the most divergent Hylaeinae taxa, *Hyleoides concinna*, where *Hyleoides* is thought to be sister clade to all other hylaeines (Almeida and Danforth 2009). Species from three other closely related lineages of colletid subfamilies, Euryglossinae, Scrapterinae, and Xeromelissinae (Almeida and Danfroth 2009) were used as outgroups. Taxa and sampling localities along with GenBank accession numbers are listed in Table 2. We refer to the two female *Hylaeus* species from sub-Saharan Africa collected by us as species from the *H. (Deranchylaeus)* group because only males of the three closely related African subgenera, *H. (Deranchylaeus)*, *H. (Cornhylaeus)*, *H. (Alfkenhylaeus)*, carry a diagnostic feature that distinguishes one subgenus from the other (Eardley and Urban 2010). Voucher specimens of species collected by the authors for this study are located at Flinders University, South Australia. Information regarding voucher specimens of sequence material taken from Genbank can be found using relevant accession numbers presented in Table 2.

### DNA EXTRACTIONS AND SEQUENCING

DNA extractions were made from individual specimens using the QIAamp® DNA Micro Kit. PCR reactions were carried out in 25 µl volumes, each containing 2.0 µl of

DNA 2.5 µl 10x Buffer 2.5 µl $MgCl_2$ (25mM) 2.5 µl dNTP mix (10 mM of each dNTP) 1 µl of each primer (5 mM) 13.3 µl $H_2O$, and 0.1 µl SuperthermTM Taq DNA polymerase (5 U/µl). Primer specifications and PCR conditions are summarised in Tables 3a and 3b. The mitochondrial gene cytochrome *c* oxidase I (COI) is commonly used in phylogenetics for inferring more recent divergences. Two regions of COI (660 + 798 bp) were used in our study in order to maximise overlap with other bee phylogenies. Two nuclear genes were used, the F2 copy of elongation factor 1α (EF-1α; 1121 bp) and the ribosomal subunit 28S (1418 bp), both of which have also been used in insect phylogenetic studies and have demonstrated utility for recovering deeper divergences (Danforth et al. 2005). Primers used in 28S amplification covered the D1, D2 and D3 regions. A total of 3,997 nucleotides were included in the analyses.

Alignments were generated using the CLC Free Workbench package (CLC bio A/S version 4.6), this program uses a progressive alignment algorithm (Feng and Doolittle 1987) to create a multiple alignment. Adjustments to alignments were made by eye, using published sequence data for 28S (Almeida et al. 2009), and ambiguous alignments between loop regions in 28S and a large intron in EF-1α were excluded from the analyses.

**PHYLOGENETIC ANALYSIS**

Bayesian Inference (BI) was carried out using the Unix version of MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). We partitioned COI and EF-1α into two parts each of 1st and 2nd codons together and the 3rd separately because of third codon position saturation bias, giving rise to five partitions including that for 28S. A general time reversible (GTR - a 6-parameter rate transition matrix) model with gamma distributed (Γ) rates with a proportion of invariant sites (I) was used because this substitution model can incorporate simpler models, such as the K2P and HKY models, if these emerge from the MCMC parameter space. Two sets of four Monte Carlo Markov chains (MCMC) with Metropolis Coupling were run in parallel for 40 million generations, sampling every 1000th generation. Convergence between parallel runs was assessed using the average standard deviation of split frequencies, and LnL values were plotted using Tracer version 1.41 (Rambaut and Drummond 2007, http://beast.bio.ed.ac.uk/Tracer) to

determine when stationarity was achieved, and we used a burnin of 20,000 out of 40,001 trees, well after stationarity was reached.

**ESTIMATING DIVERGENCE TIMES**

The crown age of Hylaeinae has been estimated to be 45 Mya, and 30 Mya for *Hylaeus* (Almeida et al. in press). These estimates were part of a family-wide colletid study that included only eleven species of Hylaeinae, and our independent estimate using more species will provide a valuable comparison. No known Hylaeinae fossils exist, and so we used genetic distance measures as a way to make estimates and to compare nucleotide substitution rates of COI and EF-1α to infer approximate crown ages (e.g., Regier and Shultz 2001) for three key nodes in our phylogeny, relative to other bee groups, under the assumption that divergence age will generally correspond to genetic distance.

Pairwise genetic distance estimates using COI and EF-1α were calculated for sister clades at: (i) the root node of the Australian *Hylaeus* only (Fig. 3, node A), where *Hylaeus* diverged into two major clades (the poorly supported clade consisting of *H. (Xenohylaeus) leptospermi*, *H. (Planihylaeus trilobatus)*, and *H. (Prosopisteron) microphenax* was excluded); (ii) the node uniting two New Zealand species, *H. (Prosopisteron) matamako* and *H. (Prosopisteron)* NZsp1 with the Australian species (Fig.3, node B); and (iii) the node separating the non-Australasian (i.e., species from outside Australia and New Zealand) from the Australian *H. (Euprosopis)* + *H. (Prosopisteron) burnsi* clade (Fig. 3, node C).

We used a general-time reversible (GTR) distance model implemented in PAUP* 4.0b10 (Swofford 2002) to compare *Hylaeus* COI substitution rates with Colletinae (Colletidae) (Kuhlmann et al. 2009), and EF-1α substitution rates to that of Colletinae and the sub-tribe Caenohalictini (Halictidae: Halictinae) (Danforth et al. 2004). Although the use of models that are fit to each group separately might provide more accurate results for comparing distances within each group, the use of separate models would not allow for meaningful comparisons between groups. The 1$^{st}$ and 2$^{nd}$ codon positions of COI were used for estimating pair-wise distances because substitutions at

the third codon positions in bees can approach saturation for even moderately old divergences (e.g., Dowton and Austin 2002; Schwarz et al. 2004). Pair-wise genetic distances between all terminal taxa that defined the node of the most recent common ancestor (mrca) for the clade of interest were used to describe the relative 'age' of each clade.

We used closely related clades to estimate relative divergence times under the assumption that base frequencies and substitution rate dynamics are most similar between closely related clades. Boxplot graphs of the estimated distances were used to graphically compare and estimate approximate node ages across groups. Colletinae is the only colletid group for which there is a large-scale phylogeny using COI data (Kuhlmann et al. 2009). The Colletinae crown age has been estimated at 33 Mya (Almeida et al. in press), and the *Colletes* crown age has been estimated to be 30 Mya (Almeida et al. in press). We plotted COI pairwise distances at the node that separates the sister Colletinae genera *Colletes* and *Mourecotelles* (Kuhlmann et al. 2009), which we will refer to as the 'Colletinae node'. Because there are a number of polytomies in the Colletinae phylogeny, we also generated distance estimates at an internal node with two distinct clades, which we will refer to as the '*Colletes* node'. Our EF-1α genetic distance plots were calibrated using Colletinae as well as a node within the sister family to Colletidae, Halictidae. A subset of the species used in the COI estimates were available for Colletinae (Almeida et al. 2011), however the same *Mourecotelles* species enabled estimates of a comparable 'Colletinae node', and we again estimated an internal *Colletes* node. The node which forms the two main clades within the Halictid sub-tribe Caenohalictini (Danforth et al. 2004) were used, those being: *Agapostemon*/ *Rhinetula*/ *Dinagapostemon*, and *Habralictus*/ *Caenohalictus*/ *Ruizantheda*/ *Pseudagapostemon*. Within their larger Halictidae phylogeny Danforth et al. (2004) calibrated this node at a minimum age of 23 Mya, based on the Dominican amber fossil *Eickwortapis dominicana* from the same tribe, and retrieved a crown age for the Caeonohalictini of 36-66 Mya, with a mean age of 45-49 Mya (Danforth et al. 2004).

In addition to estimating node ages using distance measures, we calculated relative divergence times using penalized likelihood (PL) rate smoothing implemented in r8s

v1.71 (Sanderson 2003) with the smoothing parameter determined by a cross-validation method. Outgoup taxa were pruned using TreeEdit v1.0a10 so that the focal group could be presented in as large a space as possible. We compared this to a relaxed clock estimate using a Bayesian Markov chain Monte Carlo approach in BEAST v1.3 (Drummond and Rambaut 2003). We used the uncorrelated log-normal relaxed clock and a substitution model with gamma distributed and invariant site heterogeneity (GTR+I+Γ) model and constrained the topology at nodes where there was 100 posterior probability support inferred by the Bayesian consensus phylogeny. In response to the results of the genetic distance estimates, as well as the 30 Mya estimate of *Hylaeus* origin (Almeida et al. in press), we set the crown age of *Hylaeus* at 30 Mya.

## ESTIMATING PATTERNS OF DIVERSIFICATION IN AUSTRALIAN *HYLAEUS*

We plotted a log-lineages through time (log-LTT) graph to graphically explore lineage accumulation for the Australian *Hylaeus*, with the understanding that speciation and/or extinction rates can be underestimated if cryptic diversity is not recognised or if older lineages are not well sampled (Ricklefs 2007). In order to take phylogenetic uncertainty into account, we randomly selected 500 phylograms from a post-burnin sample of 20,000 trees sampled over 40 million generations, and subjected these to the same r8s PL transformation applied to the consensus phylogram. We then used the 500 resulting chronograms, and the consensus chronogram to generate log-LTT graphs, using the mltt.plot command in the APE module (Paradis et al. 2004) in the R package.

We also conducted constant-rates (CR) tests in the R package LASER v2.2. (Rabosky 2008) to derive γ (gamma) values (Pybus and Harvey 2000), used to assess the node or branching time distribution. The γ statistic as used in the CR test serves as an indication of whether a constant rate of diversification can be rejected. A γ value lower than -1.645 (one-tailed test) indicates a rejection of constant-rate diversification in favour of a temporal decrease in diversification, with lower values indicating greater diversification closer to the root node. Inferences of declining trends in diversification are known to sometimes reflect undersampling of taxa rather than a true trend (Pybus and Harvey 2000). By first calculating an observed γ value for the available data, the Monte Carlo

constant rates (MCCR) test (Pybus and Harvey 2000) accounts for incomplete sampling by generating random trees that take into account the number of missing taxa. We used the same 500 chronograms to calculate the observed γ of the 56 Australian species, and then the simulated γ using an estimated species number of 180 Australian species, which includes 19 *Meroglossa* species and one *Hemirhiza* species (Michener 2007) that we assign to *Hylaeus* as discussed below in the results. Our MCCR test used 10,000 replicate simulations four times. The type II error associated with the MCCR test means that failure to reject the null hypothesis is not proof of constant diversification (Pybus and Harvey 2000). We then plotted a histogram of the γ values derived from the CR analysis and a comparative histogram of simulated trees derived from the MCCR analysis. Identical distributions of the simulated and observed γ values indicate an effect of under-sampling. If the distributions are not the same then we cannot say that the observed diversification patterns are a reflection of our sampling regime alone.

## *Results*

### PHYLOGENETIC TOPOLOGY

There was strong support for the monophyly of the Australian, New Zealand, and non-Australasian *Hylaeus* (100 PP) in our BI consensus phylogeny (Fig. 1). Many of the internal nodes had low PP support (<50). There was however strong support (100 PP) for the monophyly for *Gnathoprosopis*, and for *Meroglossa*. There was also strong support (100 PP) for the monophyly of a clade including *Hylaeus (Hylaeteron)* and *H. (Hylaeorhiza)* with two groups that have traditionally been assigned generic status, *Meroglossa* and *Hemirhiza*. We also found that *H. (Prosopisteron)* species were recovered in a number of deeper clades, supporting the suspected polyphyly of the group (Magnacca and Danforth 2006). Furthermore, the New Zealand subgenera have been identified as *Prosopisteron* (Michener 2007; Donovan 2007) but our phylogeny indicates that these two New Zealand species, *Hylaeus (Prosopisteron) motamako* and a *H. (Prosopisteron)* species that we could not identify, together form a divergent clade to the Australian subgenera. A second *H. (Prosopisteron)* species from New Zealand, which we could not identify to species level, is sister to two Australian *H. (Prosopisteron)* species (100 PP), and probably represents an anthropogenic dispersal.

There was strong support (100 PP) for a clade that included three of the major Australian subgenera, *Rhodohylaeus*, *Euprosopoides* and *Euprosopis*, as well as *Euprosopellus pergibbosus* and *Macrohylaeus alcyoneus*. Sister to *H. (Euprosopis)* + *H. (Prosopisteron) burnsi* is a clade including all species from outside Australia and New Zealand (100 PP). The African *H. (Deranchylaeus)* group are monophyletic and sister to the rest of the non-Australasian clade (100 PP). The two other African subgenera that range into Madagascar are paraphyletic, with 100 PP support for the *H. (Nothylaeus)* species as sister to the Holarctic *H. (Paraprosopis) wootoni*, and the *H. (Metylaeus)* species as sister to the European *H. (Spatulariella) punctatus*. Our specimen of *H. (Spatulariella) punctatus* was collected from Chile, where it was introduced from Europe (Michener 2007). The two *H. (Nesoprosopis)* species are also paraphyletic. There is strong support (100 PP) for the monophyly of the South American *H. (Gongyloprosopis) preposterosus* and the North American *H. (Prosopella) hurdi*.

**ESTIMATING DIVERGENCE TIMES**

We found that mean COI genetic distances (Fig. 2a) for the root node of the Australian *Hylaeus* clades (Fig. 3, node A) were more than half the values of the distances of the node separating *Mourecotelles* from the other *Colletes*, and from the internal *Colletes* node, which have been estimated at approximately 30 Mya (Kuhlmann et al. 2009; Almeida et al. in press). The EF-1α estimates, however, suggest that node A is approximately the same age as Colletinae and younger than the Caeonohalictini node, which has been estimated to be approximately 45 Mya. The COI and EF-1α genetic distance estimates for node B (the node that separates the New Zealand clade from the Australian clade) and node C (the node that separates the non-Australian clade from the Australian *H.* (*Euprosopis*) + *H. (Prosopisteron) burnsi*) are of a similar age to node A (with node B slightly younger than nodes A and C in the EF-1α estimates). We suggest that the EF-1α estimates are more accurate due to saturation of COI at deep divergences, and this may be exentuated when comparing a widely distributed clade, Colletinae, to mainly Australian taxa.

The PL rate smoothing consensus chronogram (Fig. 3) shows that the divergence between the two New Zealand species of *Prosopisteron* (node B) and the other non-Australian species of *Hylaeus*, which eventually led to a global distribution (node C), occurred at about the same time (approximately 23-25 Mya) and this was at an early stage of diversification among the Australian clades, which we calibrated at 30 Mya. These relative divergence times correspond more closely with the COI pairwise genetic distance estimates than the EF-1α estimates.

## ESTIMATING PATTERNS OF DIVERSIFICATION IN AUSTRALIAN *HYLAEUS*

The log-lineages through time (LTT) plots for Australian *Hylaeus* species (Fig. 4) suggests a rapid accumulation of lineages early in their history and a very marked slowing down of diversification closer to the present. This kind of Moran-type LTT pattern is often interpreted as evidence for early adaptive evolution, but Moran-type curves can also be a reflection of under-sampling of taxa (Rabosky and Lovette 2008; Crisp and Cook 2009).

The observed constant rates γ value was -3.596. A value below zero suggests that internal nodes are closer to the root than expected by a pure birth model (Pybus and Harvey 2000). Because this can be an artefact of undersampling, we compared the observed γ statistic to that expected from a Monte Carlo constant rates (MCCR) simulation, where the number of missing taxa are randomly pruned from simulated trees based on random extinction and speciation events. We derived a distribution of the observed and simulated γ values, and these are summarised as histograms in Figure 5. The mode of the simulated values is less negative than that of the observed γ values, indicating that our observed negative values can not be explained purely asthe result of undersampling, suggesting early and rapid accumulation of lineages. Rehan et al. (2010) noted that outside of the visual comparison of the observed and simulated gamma distributions provided here, which suggests that cladogenesis has declined over time, any statistical comparison of the two gamma values would not be useful as the post-burnin trees are not derived from independent samples. As well as this, a sufficiently

large number of post-burnin generations or simulated trees could provide an arbitrarily significant difference between the two values (Rehan et al. 2010).

## *Discussion*

### PHYLOGENETIC RELATIONSHIPS

Our phylogeny had very weak support (<50PP) for a number of nodes, making interpretation of some subgeneric relationships unreliable. At the same time we were able to confirm the monophyly of *Gnathoprosopis*, and *Meroglossa*, as well as raise doubt about the taxonomic classification of other subgenera. Molecular phylogenies have grouped the hylaeine genus *Palaeorhiza* within *Hylaeus* (Almeida et al. in press) and our phylogeny supported the monophyly of the genera *Meroglossa* and *Hemirhiza*, with *Hylaeus* subgenera (*Hylaeteron* and *Hylaeorhiza*). This indicates that characters that have previously been used to justify generic and subgeneric status may not be robust indicators of deep phylogenetic history. This could be due to homoplasies (e.g. due to convergent evolution) but it could also be due to problematic treatment of character states.

Our phylogeny included 14 species from regions other than Australia and New Zealand and these species formed a well-supported monophyletic group (100PP), which we refer to as the 'non-Australasian group'. There is good support (100 PP) for a shared common ancestor between the non-Australasian group and one of the most wide-spread Australian *Hylaeus* subgenera, *Euprosopis* (Houston 1975, 1981) + *Prosopisteron burnsi*. Although the representation of non-Australasian subgenera in our phylogeny is incomplete, it includes all non-Australasian species that we were able to obtain sequences for, using GenBank and our own field-collected samples, and includes species from all major geographic regions. This geographically wide sampling increases our understanding of the historical biogeography of *Hylaeus*, and the following relationships from our phylogenetic reconstruction suggest that numerous long-range dispersal episodes underlie the biogeographical history of the non-Australasian *Hylaeus*, which we expand on later.

The two African subgenera, *Nothylaeus* and *Metylaeus*, have a sub-Saharan distribution

that (unlike the *Deranchylaeus* group) extends into Madagascar (Michener 2007). However, the two Malagasy species of the genera *Nothylaeus* and *Metylaeus* are paraphyletic, not part of the sub-Saharan clade, nor sister group to the mainland African clade. Our results do not support the proposal of assigning generic status to *Nothylaeus*, as suggested by Snelling (1985). The two *H. (Nesoprosopis)* species from Hawai'i and Malaysia respectively may be paraphyletic but PP values for critical nodes are low. This is a subgenus that is most common in Hawai'i and less so in Japan and tropical Asia (Magnacca and Danforth 2006). *Hylaeus (Nesoprosopis) inquilina* is one of the 62 Hawai'ian species and our results suggest it forms a monophyletic group with the South American *H. (Gongyloprosopis) preposterosus* and the North American *H. (Prosopella) hurdi*. The *H. (Nesoprosopis)* species collected from Malaysia also formed a monophyletic group with a *H. (Metylaeus)* species from Madagascar and the Holarctic *H. (Paraprosopis) wootoni*.

**DIVERGENCE TIMES**

Our EF-1α genetic distance estimates suggest that the root node of the Australian *Hylaeus* (Fig. 3, node A) is younger than the Caeonohalictini node (that separating *Agapostemon/ Rhinetula*/ *Dinagapostemon* from *Habralictus*/ *Caenohalictus*/ *Ruizantheda*/ *Pseudagapostemon*) (estimated to be 45-49 Mya, Danforth et al. 2004), and approximately the same age as Colletinae (at the node separating *Mourecotelles* from *Colletes*), (estimated mean age 33 Mya, Almeida et al. 2011). Our COI genetic distance estimates, however, suggest that the Australian *Hylaeus* root node is much younger than the Colletinae node and the *Colletes* node (estimated mean age 24 Mya, Almeida et al. 2011). This suggests that relative rates of COI and EF-1α evolution are not directly comparable with each other. Substitution rates in mitochondrial genes tend to be higher than in nuclear genes, and the reliability of COI, as an indicator of divergence age used alone, has been questioned for estimates of divergences older then 5-10 Mya (Lin and Danforth 2004). COI is best used for species delimitations as substitutions have been shown to be unlinked to population size, as opposed to nuclear genes (Bazin et al. 2006).

We therefore relied on the inference derived from the EF-1α estimates, which indicate that *Hylaeus* and Colletinae are of a simlar age, and applied a crown age calibration of 30 Mya for *Hylaeus* when constructing our relaxed clock chronogram. Relative divergence times in the chronogram suggest that the origin of the New Zealand lineage (Fig. 3, node B) was around the same time or earlier than that of the non-Australasian clade, which corresponds more with the COI distances. The approximate timeframe for the divergence of the non-Australasian *Hylaeus* is important for developing both dispersal and vicariance hypotheses and inferring broad historical biogeography scenarios. Although we are missing some non-Australasian subgenera from our dataset, our results do suggest that *Hylaeus* originated in Australia, and that there was at least one dispersal event out of Australia, not via New Zealand, that ultimately resulted in a worldwide distribution of divergent subgenera that are not found in Australia and surrounding islands.

**PATTERNS OF DIVERSIFICATION IN AUSTRALIAN *HYLAEUS***

Our log-LTT plot of Australian *Hylaeus* species from 500 randomly sampled post burnin trees suggests rapid early diversification of Australian *Hylaeus*, with slower cladogenesis in more recent times. This kind of Moran-type LTT pattern is often interpreted as evidence for early adaptive evolution, but Moran-type curves can also be a reflection of under-sampling of taxa (Rabosky and Lovette 2008; Crisp and Cook 2009). Our γ (gamma) simulations suggest that higher early rates of cladogenesis are not explainable purely in terms of under-sampling of taxa, suggesting that *Hylaeus* very likely did have higher rates of cladogenesis closer to its origin. Adaptive radiations can represent an initial rapid filling of available niche space over time (Schluter 2000; Phillimore and Price 2008), leaving fewer and fewer niches to be explored (Schluter 2000; Phillimore and Price 2008; Rabosky and Lovette 2008). A pattern of rapid diversification early on in *Hylaeus* evolution approximately 30 Mya is in line with the supposition made by Almeida et al. (in press) that otherwise constant diversification in Colletidae from 70 Mya increased suddenly at 30 Mya, as a result of the origin of *Hylaeus*, as well as a second widespread genus, *Colletes* (Colletinae).

Crisp and Cook (2009) observed an anti-sigmoidal curve in log-LTT plots for Australian

legume groups and an African legume clade. They suggest that mass extinctions can give rise to an anti-sigmoidal curve, which can misleadingly suggest an early rapid radiation. Ezrad et al. (2011) show this by using the rich fossil record of Cenozoic macroperforate planktonic foraminifera. Their log-LTT analyses, including only extant taxa, had a distinctly anti-sigmoidal curve, whereas the plot including both fossil and extant lineages was very different. However, our Moran-type LTT plot shows no evidence of an anti-sigmoidal shape that may indicate large extinction events (e.g., Crisp and Cook 2009; Ezrad et al. 2011). Additional evidence for changing speciation rather than changing extinction rates in *Hylaeus* comes from the negative $\gamma$ values of our MCCR test. Negative values are associated with increased speciation and constant extinction rates (Rabosky and Lovette 2008), whereas positive $\gamma$ values are associated with increased extinction rates and constant speciation (Rabosky and Lovette 2008). Over-dispersed sampling, where an effort is made to cover all representative groups or regions, is also likely to inflate evidence for early radiation pulses because it increases the amount of deeper divergences relative to crown clades (Cusimano and Renner 2010). Because we did not restrict sequencing to representatives of a few species per subgenus but obtained sequences from all material we collected, provided that species distinctness was evident from morphology, our experimental design is unlikely to lead to over-dispersed sampling of taxonomic groups.

The only other study to look at diversification patterns in Australian bees is on the long-tongued bee tribe Allodapini (Chenoweth and Schwarz 2011). The crown age of the Australian allodapines is about 35-30 Mya, which may be similar to the crown age for *Hylaeus*. LTT plots in that study were mostly linear, indicating a constant diversification rate, but with an increased rate of cladogenesis between 10 and 6 Mya, which coincides with the 'Hill Gap' (Chenoweth and Schwarz 2011). The Hill Gap corresponds to a major increase in aridification from a milder mid-Miocene Australian climate (Hill 1994; Byrne et al. 2008), and Chenoweth and Schwarz (2011) suggested that increasing cladogenesis of allodapines during that time may have been due to increased rates of allopatric speciation as temperate habitats became more subdivided. Support for this notion comes from the fact that the increase in cladogenesis was restricted to temperate-adapted allodapine genera, rather than the arid distributed

*Exoneurella*. The only way that our *Hylaeus* LTT plots could correspond to accelerated cladogenesis during the Hill Gap would be for the genus to have an origin close to 10 Mya, which is not concordant with estimated origins from Almeida et al. (in press), our genetic distance data, or with the greater species diversity of *Hylaeus*. Indeed, large clade size can be related to a higher than average diversification rate (Ricklefs 2007; Phillimore and Price 2008), or be a consequence of the longevity of an old clade (McPeek and Brown 2007). Given that a 30 Mya crown age makes *Hylaeus* one of the younger colletid clades (Almeida et al. in press) despite being one of the most speciose clades (Michener 2007), we suggest that high diversification rate has contributed more to species richness in *Hylaeus* than has age.

## DISPERSALS OUT OF (BUT NOT INTO) AUSTRALIA

Our phylogeny suggests that there has been at least one dispersal of *Hylaeus* from Australia to New Zealand and another dispersal from Australia that resulted in a very diverse and widespread non-Australasian *Hylaeus* group. *Hylaeus* species found in New Zealand, New Caledonia, and other south Pacific Islands (Australasia) have been placed in the otherwise Australian subgenus *Prosopisteron*. *Hylaeus* found outside of Australasia have been classified into subgenera that are not found in Australia (Michener 2007), and are probably more closely related to each other than to Australasian forms based on morphological characters (Michener 1979). Therefore, our phylogeny here is congruent with morphological characteristics that suggest forms of *Hylaeus* from outside Australasia are more closely related to each other than to Australasian forms (Michener 1979). We discuss two possible dispersal patterns that explain the current distribution of *Hylaeus* separately below. We then discuss the lack of evidence for any 'back-migration' into Australia.

### *From Australia to New Zealand*

New Zealand has been geologically isolated from Australia and from any other landmass for about 80 My (McLoughlin 2001). The origin of colletids has been estimated at approximately 70 Mya (Brady et al. 2009; Almeida et al. in press), with an estimated crown age of *Hylaeus* at 30 Mya (Almeida et al. in press). From our results, the crown age of the two New Zealand species is estimated at about 25 Mya, strongly favouring a

dispersal hypothesis for the presence of a *Hyleaus* lineage in New Zealand over a vicariance explanation. Only eight *Hylaeus* species are found in New Zealand, and the overall bee fauna is depauperate (Donovan 2007). When compared with the highly diverse bee fauna in Australia these observations suggest dispersal and colonisation of New Zealand by bees has been rare, there is relatively poor ecological opportunity to speciate once established, or extinction has played a major role there. The case for a complete biotic turnover caused by a major extinction event in New Zealand, "the Oligocene drowning" approximately 27 Mya (Cooper and Cooper 1995; Trewick and Morgan-Richards 2005; Stevens and Hicks 2009) has been refuted by evidence of lineages that have existed through the hypothesized drowning period (e.g., Knapp et al. 2007; Allwood et al. 2010; Giribet and Boyer 2010). Combined, this evidence indicates that a relatively large proportion of once-exposed land surface became submerged, which would have no doubt led to some extinction events, as would tectonic events 5-2 Mya, and glacial cycles over the last 2.5 My (Winkworth et al. 2002). These events could have been of particular significance to bees that forage at lower altitudes and around the coastline where flooding would have been most apparent.

### *Only once out of Australasia*

The other dispersal event out of Australia ultimately resulted in the world-wide distribution of *Hylaeus* subgenera divergent from Australasian forms. Despite being drawn from a very wide geographic range, including species from southern Africa, Madagascar, Hawai'i, Malaysia, North America, South America, and Eurasia, lack of support for some nodes in the non-Australasian clade only allows us to cautiously discuss the biogeographical patterns that led to the current distribution of *Hylaeus* outside of Australasia.

The crown age of the non-Australasian *Hylaeus* was estimated at between 25-22 Mya, and this clade is sister to the Australian *Euprosopis* + *Prosopisteron burnsi*, with the sub-Saharan *Deranchylaeus* group sister to the other non-Australasian taxa (Fig. 3). Michener (1979) suggested that a dispersal route out of Australia via Africa (or via India across a reduced Indian Ocean, or both) would explain the high diversity of *Hylaeus* in the Palearctic (second to Australia), rather than a route through Asia where

representation of the group is low (Michener 1979), but largely understudied (Snelling 1985). However, these scenarios were suggested under the assumption that the subfamily Hylaeinae comprises a much more ancient lineage that possibly diverged 70 Mya when distances between Australia and Asia were vast (Michener 1979). Given the results of Almeida et al. (in press) that support a much more recent crown age of Hylaeinae approximately 45 Mya, and a crown age of *Hylaeus* approximately 30 Mya, we cannot discount the possibility of dispersal out of Australia through Asia. The northward drifting of the Australian plate reached tropical latitudes during the Miocene (McLoughlin 2001) and potentially provided easier access to the northern hemisphere via Asia.

We propose that the dearth of dispersal events out of Australia is not a result of unsuccessful dispersal *per se*, but unsuccessful (re)establishment (e.g. Silvertown 2004). Winkworth et al. (2002) made two speculations about why plant distributions in the Southern Hemisphere have largely been inferred to have arisen in the last 10 My when opportunities for dispersal have existed for 65 My: (i) the extinction of taxa; and (ii) the limitation of suitable and/or available habitats. Extinction of *Hylaeus* lineages that hypothetically diversified after dispersals from Australia might have contributed to our observations, but this is difficult to test. The second point implicates unsuitable habitat or unavailable niche-space for dispersers to colonise, where an earlier lineage has already established, because of competition for shared resources (Winkworth et al. 2002; Silvertown 2004).

### *No Hylaeus re-colonisation of Australia*

The global radiation of *Hylaeus* following its dispersal out of Australia is remarkable for the lack of evidence for any backwards migration into Australia, despite the increasing geographical closeness of Laurasia to Australia since the late Cretaceous (McLoughlin 2001). Molecular phylogenetic evidence for the dispersal of multiple bee lineages into Australia (all outside Colletidae) (e.g., Danforth and Ji 2001; Leys et al. 2002; Fuller et al 2005; Schaefer and Renner 2008; Rasmussen and Cameron 2010; Rehan et al. 2010; Chenoweth and Schwarz 2011) further highlights the availability of dispersal opportunities since the late Eocene. One possibility for our finding involves the same

arguments as above for why only a single successful dispersal out of Australia occurred, viz. that opportunities to exploit new niches in a relatively enemy-free space by early colonizers are less likely to arise for subsequent dispersals (Winkworth et al. 2002; Silvertown 2004). This might be especially relevant when considering migration into Australia where *Hylaeus* diversification and thus filling of niche-space has been highest.

## *Conclusions*

Our results suggest that successful dispersals, and subsequent radiations, cannot be explained in terms of just overcoming geographic barriers. Instead, they suggest that early dispersals inhibit subsequent successful dispersals, or at least that factors enabling early dispersals do not operate at later times when geographic barriers to dispersal are lower. Determining whether this pattern is widespread is important. If it reflects a general pattern, then it suggests that biotic composition of a region may be determined more by the early appearance of dispersal opportunities than by later relaxation of geographic dispersal barriers.

**ACKNOWLEDGEMENTS**

Terry Houston of the Western Australian Museum (WAM) identified many of the Australian Hylaeinae species, and WAM, and Kirstenbosch Research Centre in South Africa hosted PK. Luke Chenoweth, Scott Groom, Sally Harradine, Michael McLeish, and Sandra Rehan helped with fieldwork. This work was funded by ARC grants awarded to MPS, as well as Holsworth Wildlife Fund, and Sir Mark Mitchell awards to PK.

**LITERATURE CITED**

Allwood, J., D. Gleeson, G. Mayer, S. Daniels, J.R. Beggs, and T.R. Buckle. 2010. Support for vicariant origins of the New Zealand Onychophora. J. Biogeogr. 37:669–68.

Almeida, E.A.B., and B.N. Danforth. 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50:290-309.

Almeida, E.A.B., M.R. Pie, S.G. Brady, and B.N. Danforth. In press. Biogeography and diversification of Colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. J. Biogeogr.

Bazin, E., S. Glemin, and N. Galtier. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science. 312:570-572.

Brady, S.G., L. Larkin, and B.N. Danforth. 2009. Bees, ants, and stinging wasps (Aculeata), pp 264-269 in *The Timetree of Life* Hedges, S.B. and Kumar, S., eds. Oxford Univ. Press, U.K.

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

Chenoweth, L.B., and M.P. Schwarz. 2011. Biogeographical origins and diversification of the exoneurine allodapine bees of Australia (Hymenoptera, Apidae). J. Biogeogr. 38:no. doi: 10.1111/j.1365-2699.2011.02488.x.

Cooper, A., and R.A. Cooper. 1995. The Oligocene bottleneck and New Zealand biota: genetic record of a past environmental crisis. P. Roy. Soc. Lond. B. Bio. 261:293-302.

Crisp, M.D., and L.G. Cook. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. Evolution. 63(9):2257-2265.

Cusimano, N., and S.S. Renner. 2010. Slowdowns in diversification rates from real phylogenies may not be real. Syst. Biol. 59:1-7.

Danforth, B.N., H. Sauquet, L. Packer. 1999. Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae), based on parsimony and likelihood analyses of nuclear EF-1$\alpha$ sequence data. Mol. Phylogenet. Evol. 13(3):605–618.

Danforth B.N., and S. Ji. 2001. Australian *Lasioglossum +Homalictus* form a monophyletic group: resolving the "Australian enigma". Syst. Biol. 50:268–283.

Danforth, B.N., S.G. Brady, S.D. Sipes, and A. Pearson. 2004. Single-copy nuclear genes recover cretaceous-age divergences in bees. Syst. Biol. 53(2):309-326.

Danforth, B.N., C. Lin, and J. Fang. 2005. How do insect nuclear ribosomal genes compare to protein-coding genes in phylogenetic utility and nucleotide substitution patterns. Syst. Biol. 30:549-562.

Danforth B.N., S. Sipes, J. Fang, and S.G. Brady. 2006. The history of early bee diversification based on five genes plus morphology. P. Natl. Acad. Sci. USA. 103(41):15118–15123.

Donovan, B.J. 2007. Apoidea (Insecta: Hymenoptera) in Fauna of New Zealand (Ko te Aitanga Pepeke o Aotearoa). ManaakinWhenua Press. Canterbury, NZ.

Dowton, M., and A.D. Austin. 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy – the behaviour of the incongruence length difference test in mixed-model analyses. Syst. Biol. 51(1):19-31.

Drummond, A.J., and A. Rambaut. 2003. BEAST. Version 1.3 [computer program]. Available: http://evolve.zoo.ox.ac.uk/. Accessed 22 April 2009.

Eardley, C., and R. Urban. 2010. Catalogue of Afrotropical bees (Hymenoptera: Apoidea: Apiformes). Zootaxa. 2455:1-548.

Ezard, T.H.G., T. Aze, P.N. Pearson, and A. Purvis. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. Science. 332:349-351.

Feng, D., and R.F. Doolittle. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. 25(4):351-360.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3:294–299.

Fuller, S., M. Schwarz, and S. Tierney. 2005. Phylogenetics of the allodapine bee genus *Braunsapis*: historical biogeography and long-range dispersal over water. J. Biogeogr. 32:2135–2144.

Giribet, G., and S.L. Boyer. 2010. 'Moa's Ark' or 'Goodbye Gondwana': is the origin of New Zealand's terrestrial invertebrate fauna ancient, recent, or both? Invertebr. Syst. 24:1–8.

Grimaldi, D. 1999. The co-radiations of pollinating insects and angiosperms in the Cretaceous. Ann. Mo. Bot. Gard.86(2):373-406.

Hill, R.S. (ed.) 1994. The history of selected Australian taxa in *History of Australian vegetation: Cretaceous to Recent*. Cambridge Univ. Press, U.K.

Houston, T.F. 1975. A revision of the Australian hylaeine bees (Hymenoptera: Colletidae). Aust. J. Zool. Supp. Ser. No. 36:1-135.

Houston, T.F. 1981. A revision of the hylaeine bees (Hymenopter: Colletidae). II. Aust. J. Zool. Supp. Ser. No. 80:1-128.

Huelsenbeck, J.P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics. 17:754-755.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-120.

Knapp, M., R. Mudaliar, D. Havell, S.J. Wagstaff, and P.J. Lockhart. 2007. The drowning of New Zealand and the problem of *Agathis*. Syst. Biol. 56(5):862-870.

Kuhlmann, M., A.E.B. Almeida, N. Laurenne, and D.L.J. Quicke. 2009. Molecular phylogeny and historical biogeography of the bee genus *Colletes* Latreille, 1802 (Hymenoptera: Apiformes: Colletidae), based on mitochondrial COI and nuclear 28S sequence data. Insect Syst. Evol. 40:291-318.

Lewis, R., D.R. Marchant, A.C. Ashworth, L. Hedenas, S.R. Hemming, J.V. Johnson, M.J. Leng, M.L. Machlus, A.E. Newton, and J.I. Raine. 2008. Mid-Miocene cooling and the extinction of tundra in continental Antarctica. P. Natl. Acad. Sci. USA. 105(31):10676-10680.

Lin, C., and B.N. Danforth. 2004. How do insect nuclear and mitochondrial gene substitution
patterns differ? Insights from Bayesian analyses of combined datasets. Mol. Phylogen. Evol. 30:686–702.

Leys, R., S.J.B. Cooper, and M.P. Schwarz. 2002. Molecular phylogeny and historical biogeography of the large carpenter bees, genus *Xylocopa* (Hymenoptera: Apidae). Biol. J. Linn. Soc. 77:249-266.

McGinley, R.J. 1980. Glossal morphology of the Colletidae and recognition of the Stenotritidae at the family level. J. Kansas Entomol. Soc. 53:539-552.

McLoughlin, S. 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. 49:271-300.

McPeek, M.A., and J.M. Brown. 2007. Clade age and not diversification rate explain species richness among animal taxa. Am. Nat. 169:E97-106.

Magnacca, K.N., and B.N. Danforth. 2006. Evolution and biogeography of native Hawaiian *Hylaeus* bees (Hymenoptera: Colletidae). Cladistics. 22:393–411.

Mardulyn, P., and J.B Whitfield. 1999. Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of microgastrinae (Hymenoptera: Braconidae): evidence of a high diversification rate in this group of parasitoids. Mol. Phylogenet. Evol. 12, 282–294.

Michener, C.D. 1965. A classification of the bees of the Australian and South Pacific regions. B. Am. Mus. Nat. Hist. 130:1-362.

Michener, C.D. 1979. The biogeography of the bees. Ann. Mo. Bot. Gard. 66(3):277-347.

Michener, C.D. 2007. The Bees of the World, 2nd edn. Johns Hopkins University Press, Baltimore, MD.

Phillimore A.B., and T.D. Price. 2008. Density dependent cladogenesis in birds. PLoS Biol. 6:e71.

Poinar Jr, J.O., and B.N. Danforth. 2006. A fossil bee from early Cretaceous Burmese amber. Science. 314: 614.

Pybus, O.G., and P.H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. P. Roy. Soc. Lond. B. Bio. 267:2267-2272.

Rabosky D.L. 2008. LASER: Likelihood Analysis of Speciation/Extinction Rates from Phylogenies. R package. version 2.2.

Rabosky, D.L., and I.J. Lovette. 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution. 62(8):1866-1875.

Rasmussen, C., and S.A. Cameron. 2010. Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. Biol. J. Linn. Soc. 99:206–232.

Regier, J.C., and J.W. Shultz. 2001. Elongation Factor-2: A Useful Gene for Arthropod Phylogenetics. Mol. Phylogenet. Evol. 20(1):136–148.

Rehan, S.M., T.W. Chapman, A.I. Craigie, M.H. Richards, S.J.B. Cooper, and M.P. Schwarz. 2010. Molecular phylogeny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. Mol. Phylogen. Evol. 55(3):1042-1054.

Ricklefs, R.E. 2007. Estimating diversification rates from phylogenetic information. Trends Ecol. Evol. 22(11):601-609.

Sanderson, M.J. 2003. r8s: inferring absolute rates of evolution and divergence times in the absence of a molecular clock. Bioinformatics. 19:301–301.

Sanmartín, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53(2):216-243.

Schaefer, H., and S. Renner. 2008. Phylogeny of the oil bee tribe Ctenoplectrini (Hymenoptera: Anthophila) based on mitochondrial and nuclear data: evidence for early Eocene divergence and repeated out-of-Africa dispersal. Mol. Phylogenet. Evol. 47(2):799-811.

Schluter, D. 2000. Ecological Character Displacement in Adaptive Radiation. Am. Nat. 156:S4-S16.

Schwarz, M.P., S.M. Tierney, S.J.B. Cooper, and N.J. Bull. 2004. Molecular phylogenetics of the allodapine bee genus Braunsapis: A–T bias and heterogeneous substitution parameters. Mol. Phylogenet. Evol. 32:110–122.

Silvertown, J. 2004. The ghost of competition past in the phylogeny of island endemic plants. J. Ecol. 92:168-173.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87:651–701.

Smith, S.A., J.M. Beaulieu, and M.J. Donoghue. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. P. Natl. Acad. Sci. USA. 107(13):5897-5902.

Snelling, R.R. 1985. The systematics of the Hylaeine bees (Hymenoptera: Colletidae) of the Ethiopian zoogeographical region: the genera and subgenera with revisions of the smaller groups. Contr. Sci., Nat. Hist Mus. LA. Co. 361:1-33.

Soltis, P.S., and D.E. Soltis. 2004. The origin and diversification of angiosperms. Am. J. Bot. 91(10):1614-1626.

Stevens, M.I., P. Greenslade, I.D. Hogg, and P. Sunnucks. 2006. Examining Southern Hemisphere springtails: could any have survived glaciation of Antarctica? Mol. Biol. Evol. 23:874-882.

Stevens, M.I., and B.J. Hicks. 2009. Mitochondrial DNA reveals monophyly of New Zealand's *Gobiomorphus* (Teleostei: Eleotridae) among a morphological complex. Evol. Ecol. Res. 11:109-123.

Swofford, D.L. 2002. PAUP. Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10 for Macintosh. Sinauer Associates, Sunderland, MA.

Trewick, S.A., and M. Morgan-Richards. 2005. After the deluge: mitochondrial DNA indicates Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera: Anostostomatidae). J. Biogeogr. 32:295–309.

Veevers, J.J., and Z.X. Li. 1991. Review of seafloor spreading around Australia. II. Marine magnetic anomaly modelling. Aust. J. Earth Sci. 38:391-408.

Ward, P.S., and S.G Brady. 2003. Phylogeny and biogeography of the ant subfamily Myrmeciinae (Hymenoptera: Formicidae). Invertebr. Syst. 17:361–386.

Winkworth, R.C., S. J. Wagstaff, D. Glenny, and P.J. Lockhart. 2002. Plant dispersal N.E.W.S from New Zealand. Trends Ecol. Evol. 17 (11): 14-520.

Zachos, J.C., L.D. Stott, and K.C. Lochmann. 1994. Evolution of early Cenozoic marine temperatures. Paleoceanography. 9:353-387.

Table 1: Subgeneric distributions of *Hylaeus* modified from Michener (2007), and number of total species from each region.

| | Number of subgenera | Number of species |
|---|---|---|
| Australasia | 22 | 173 |
| Nearctic | 7 | 35 |
| Neotropical | 4 | 50 |
| Oriental | 8 | ? |
| Palearctic | 10 | 186 |
| Sub-Saharan | 5 | 95 |

Table 2: Collection localities, and GenBank accession numbers for 28S, EF-1α and two regions of COI. Abbreviated collection localities refer to states in Australia: NT=Northern Territory; Qld=Queensland; SA=South Australia; Vic=Victoria, WA=Western Australia. Distributions of non-Australasian species where sequence material was obtained from GenBank are bracketed. GenBank accession numbers that begin with JN are from this study; those beginning with AY or DQ are from two other studies.

| Species | Collection locality | 28S | EF-1α | COI | |
|---|---|---|---|---|---|
| **Australasian species** | | | | | |
| *Hemirhiza* group | Cardwell, Qld | JN566243 | JN566289 | - | JN578924 |
| Hylaeinae sp. | North of Highway Inn Road House, NT | JN566244 | JN566290 | - | JN578925 |
| *Hylaeus* sp. 3 | Lake Gilles, SA | JN566278 | JN566329 | JN578912 | JN578962 |
| *Hylaeus (Edriohylaeus) ofarrelli* | GenBank | DQ768585 | DQ884676 | - | - |
| *Hylaeus (Euprosopellus) pergibbosus* | Ngaraket, SA | JN566247 | JN566293 | JN578891 | JN578928 |
| *Hylaeus (Euprosopis)* sp. 7 | Glen Eagle, WA | - | JN566296 | JN578892 | JN578931 |
| *Hylaeus (Euprosopis) disjunctus* | GenBank | DQ768586 | DQ884677 | - | - |
| *Hylaeus (Euprosopis) elegans* | Quorn, SA | JN566248 | JN566294 | - | JN578929 |
| *Hylaeus (Euprosopis) husela* | Larrimah, NT | JN566249 | JN566295 | - | JN578930 |
| *Hylaeus (Euprosopis) violaceus* | Crusoe Beach, WA | JN566250 | JN566297 | - | JN578932 |
| *Hylaeus (Euprosopoides)* sp. | Augusta, WA | JN566252 | JN566299 | JN578894 | JN578934 |

| | | | | | |
|---|---|---|---|---|---|
| *Hylaeus (Euprosopoides)* sp. 1 | Augusta, WA | JN566253 | JN566300 | JN578895 | JN578935 |
| *Hylaeus (Euprosopoides) cyanurus* | GenBank | DQ768587 | DQ884678 | - | - |
| *Hylaeus (Euprosopoides) ruficeps* | Mataranka, NT | JN566251 | JN566298 | JN578893 | JN578933 |
| *Hylaeus (Gephrohylaeus) sculptus* | GenBank | DQ768598 | DQ884688 | - | - |
| *Hylaeus (Gnathoprosopis) albonitens* | Larrimah, NT | JN566254 | JN566301 | JN578896 | JN578936 |
| *Hylaeus (Gnathoprosopis) amiculus* | GenBank | DQ872777 | DQ212154 | - | AY913955 |
| *Hylaeus (Gnathoprosopis) euxanthus* | Port Lincoln, SA | - | JN566304 | JN578898 | JN578939 |
| *Hylaeus (Gnathoprosopoides) bituberculatus* | GenBank | DQ768610 | DQ884699 | - | - |
| *Hylaeus (Heterapoides) exleyae* | GenBank | DQ768599 | DQ884689 | - | - |
| *Hylaeus (Heterapoides) extensus* | GenBank | DQ768600 | DQ884690 | - | - |
| *Hylaeus (Hylaeorhiza) nubilosus* | GenBank | DQ768591 | DQ884681 | - | - |
| *Hylaeus (Hylaeteron) douglasi* | South-west WA | JN566255 | JN566306 | - | - |
| *Hylaeus (Macrohylaeus) alcyoneus* | GenBank | DQ768577 | DQ884668 | - | - |
| *Hylaeus (Planihylaeus) trilobatus* | GenBank | DQ768611 | DQ884700 | - | - |
| *Hylaeus (Prosopisteron)* sp. 1 | Charters Towers, Qld | JN566272 | JN566323 | - | JN578956 |

| | | | | | |
|---|---|---|---|---|---|
| *Hylaeus (Prosopisteron)* sp. 5 | Daintree, Qld | JN566269 | JN566320 | JN578905 | JN578953 |
| *Hylaeus (Prosopisteron)* sp. 2 | 94km south of Williams, WA | JN566271 | JN566322 | JN578907 | JN578955 |
| *Hylaeus (Prosopisteron)* sp. | Crusoe Beach, WA | JN566262 | JN566313 | JN578902 | JN578947 |
| *Hylaeus (Prosopisteron)* sp. 3 | Geraldton, WA | JN566270 | JN566321 | JN578906 | JN578954 |
| *Hylaeus (Prosopisteron) aralis* | GenBank | DQ768590 | DQ884680 | - | - |
| *Hylaeus (Prosopisteron) bidentatus* | GenBank | DQ768614 | DQ884702 | - | - |
| *Hylaeus (Prosopisteron) burnsi* | GenBank | DQ768620 | DQ884708 | - | - |
| *Hylaeus (Prosopisteron) chlorosomus* | Whyalla, SA | JN566263 | JN566314 | JN578903 | JN578948 |
| *Hylaeus (Prosopisteron) cyaneomicans* | GenBank | DQ768616 | DQ884704 | - | - |
| *Hylaeus (Prosopisteron) eugeniellus* | GenBank | DQ768594 | DQ884684 | - | - |
| *Hylaeus (Prosopisteron) matamoko* | New Zealand | JN566264 | JN566315 | - | JN578949 |
| *Hylaeus (Prosopisteron) microphenax* | GenBank | DQ768595 | DQ884685 | - | - |
| *Hylaeus (Prosopisteron)* NZ sp. 1 | New Zealand | JN566267 | JN566318 | - | JN578951 |
| *Hylaeus (Prosopisteron)* NZ sp. 4 | New Zealand | JN566268 | JN566319 | - | JN578952 |
| *Hylaeus (Prosopisteron)* | Port Lincoln, SA | JN566259 | JN566310 | JN578901 | JN578944 |

| | | | | | |
|---|---|---|---|---|---|
| *perhumilis* | | | | | |
| *Hylaeus (Prosopisteron) primulipictus* | GenBank | DQ768593 | DQ884683 | - | - |
| *Hylaeus (Prosopisteron) quadratus* | Augusta, WA | JN566265 | JN566316 | - | - |
| *Hylaeus (Prosopisteron) simplus* | GenBank | DQ768573 | DQ884664 | - | - |
| *Hylaeus (Prosopisteron) vittatifrons* | Alawoona, SA | JN566266 | JN566317 | JN578904 | JN578950 |
| *Hylaeus (Rhodohylaeus)* sp. 4 | Perth, WA | JN566273 | JN566324 | - | JN578957 |
| *Hylaeus (Rhodohylaeus)* sp. 3 | Nelson, Vic | JN566275 | - | JN578909 | JN578959 |
| *Hylaeus (Rhodohylaeus)* sp. | Whyalla, SA | JN566274 | JN566326 | JN578908 | JN578958 |
| *Hylaeus (Rhodohylaeus)* sp. 1 | Halls Creek, WA | JN566276 | JN566327 | JN578910 | JN578960 |
| *Hylaeus (Rhodohylaeus)* sp. 2 | Kojunup, WA | JN566277 | JN566328 | JN578911 | JN578961 |
| *Hylaeus (Rhodohylaeus) constrictiformis* | GenBank | DQ768596 | DQ884686 | DQ872733 | - |
| *Hylaeus (Rhodohylaeus)* near *lateralis* | Kojunup, WA | - | JN566325 | - | - |
| *Hylaeus (Rhodohylaeus) proximus* | GenBank | AY654493 | AY585130 | DQ87273 | - |
| *Hylaeus (Sphaerhylaeus)* sp. 4 | Ngarakat, SA | JN566280 | JN566331 | JN578913 | JN578964 |
| *Hylaeus (Xenohylaeus) leptospermi* | GenBank | DQ768612 | DQ884701 | - | - |

| | | | | | |
|---|---|---|---|---|---|
| *Meroglossa striaticeps* | Qld | JN566283 | - | JN578917 | JN578966 |
| *Meroglossa* near *eucalypti* | ? | JN566281 | JN566333 | JN578915 | JN578965 |
| *Meroglossa eucalypti* group | 90 km north of Adelaide River, NT | - | JN566332 | JN578914 | - |
| *Meroglossa torrida* group | 30 km north of Katherine, NT | - | JN566335 | JN578918 | - |
| *Meroglossa rubricata* | Gerladton, WA | JN566282 | JN566334 | JN578916 | - |
| **Non-Australasian species** | | | | | |
| *Hylaeus (Deranchylaeus)* sp. | GenBank (Sub-Saharan Africa) | DQ768575 | DQ884666 | - | - |
| *Hylaeus (Deranchylaeus)* group Swazi sp. | Hawane, Swaziland | JN566245 | JN566291 | JN578889 | JN578926 |
| *Hylaeus (Deranchylaeus)* group Zambia sp. | Zambia | JN566246 | JN566292 | JN578890 | JN578927 |
| *Hylaeus (Gongyloprosopis) preposterosus* | GenBank (Bolivia) | DQ768576 | DQ884667 | - | - |
| *Hylaeus (Hylaeus) leptocephalus* | GenBank (Holarctic) | DQ768580 | DQ212156 | - | AY913959 |
| *Hylaeus (Metylaeus)* Madagascar sp. | Ifaty, Madagascar | JN566256 | JN566307 | JN578899 | JN578941 |
| *Hylaeus (Nesoprosopis) inquilina* | GenBank (Hawai'i) | DQ768578 | DQ212162 | - | AY914012 |
| *Hylaeus (Nesoprosopis)* Malaysia sp. | Sarawak, Malaysia | JN566257 | JN566308 | - | JN578942 |

| | | | | | |
|---|---|---|---|---|---|
| *Hylaeus (Nothylaeus)* Madagascar sp. | Ifaty, Madagascar | JN566258 | JN566309 | JN578900 | JN578943 |
| *Hylaeus (Paraprosopis) wootoni* | GenBank (North and central America) | DQ768582 | DQ884673 | - | AY913963 |
| *Hylaeus (Prosopella) hurdi* | GenBank (North and central America) | DQ768584 | DQ884675 | - | AY913964 |
| *Hylaeus (Prosopis) affinis* | St. Catharines, Canada | JN566260 | JN566311 | - | JN578945 |
| *Hylaeus (Prosopis)* Malaysia sp. | Sarawak, Malaysia | JN566261 | JN566312 | - | JN578946 |
| *Hylaeus (Spatulariella) punctatus* Chile sp. | West Conquimbo, Chile (Western Palearctic) | JN566279 | JN566330 | - | JN578963 |
| **Outgroup species** | | | | | |
| *Hyleoides concinna* (Hylaeinae) | GenBank | DQ884691 | DQ884691 | - | - |
| *Callohesma sinapipes* (Euryglossinae) | Whyalla, SA | JN566238 | JN566284 | JN57886 | JN578919 |
| *Chilicola* sp. 13 (Xeromelissinae) | West Conquimbo, Chile | JN566240 | JN566286 | - | JN578921 |
| *Pachyprosopis* sp. A (Euryglossinae) | Larrimah, NT | JN566239 | JN566285 | JN578887 | JN578920 |
| *Scrapter bicolor* (Scrapterinae) | Pakhuis Pass, South Africa | JN566241 | JN566287 | JN578888 | JN578922 |
| Xeromelissinae sp. 2 (Xeromelissinae) | West Conquimbo, Chile | JN566242 | JN566288 | - | JN578923 |

Table 3a: Primer specifications.

| Locus | Primer name | Sequence 5'-3' | Reference |
|---|---|---|---|
| EF-1α | HaF2For1 | GGG YAA AGG WTC CTT CAA RTA TGC | Danforth et al. 1999 |
| | F2-rev1 | A ATC AGC AGC ACC TTT AGG TGG | Danforth et al. 1999 |
| 28S rRNA | A-28S-For | CCC CCT GAA TTT AAG CAT AT | Ward and Brady 2003 |
| | Mar28S-Rev | TAG TTC ACC ATC TTT CGG GTC CC | Mardulyn and Whitfield 1999 |
| COI | Jerry (C1-J-2183) | CAA CAT TTA TTT TGA TTT TTT GG | Simon et al. 1994 |
| | Pat (TL2-N-3014) | TCC AAT GCA GTA ATC TGC CAT ATT A | Simon et al. 1994 |
| | LCOI490 | CCT TTT ATA ATT GGA GGA TTT GG | Folmer et al. 1994 |
| | M399 | TCA TCT AAA AAC TTT AAT TCC TG | Schwarz et al. 2004 |

Table 3b: PCR conditions for primers used showing temperature and time of: initial denaturation; and 35 cycles of denaturation, annealing, extension.

| Primer | PCR Conditions |
|---|---|
| HaF2For1/F2-rev1 | 94˚C 3min; 94˚C 60sec, 52˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| A-28S-For/Mar28S-Rev | 94˚C 3min; 94˚C 60sec, 58˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| Jerry/Pat/ LC01490/M399 | 94˚C 3min; 92˚C 30sec, 52˚C 90sec, 72˚C 2.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |

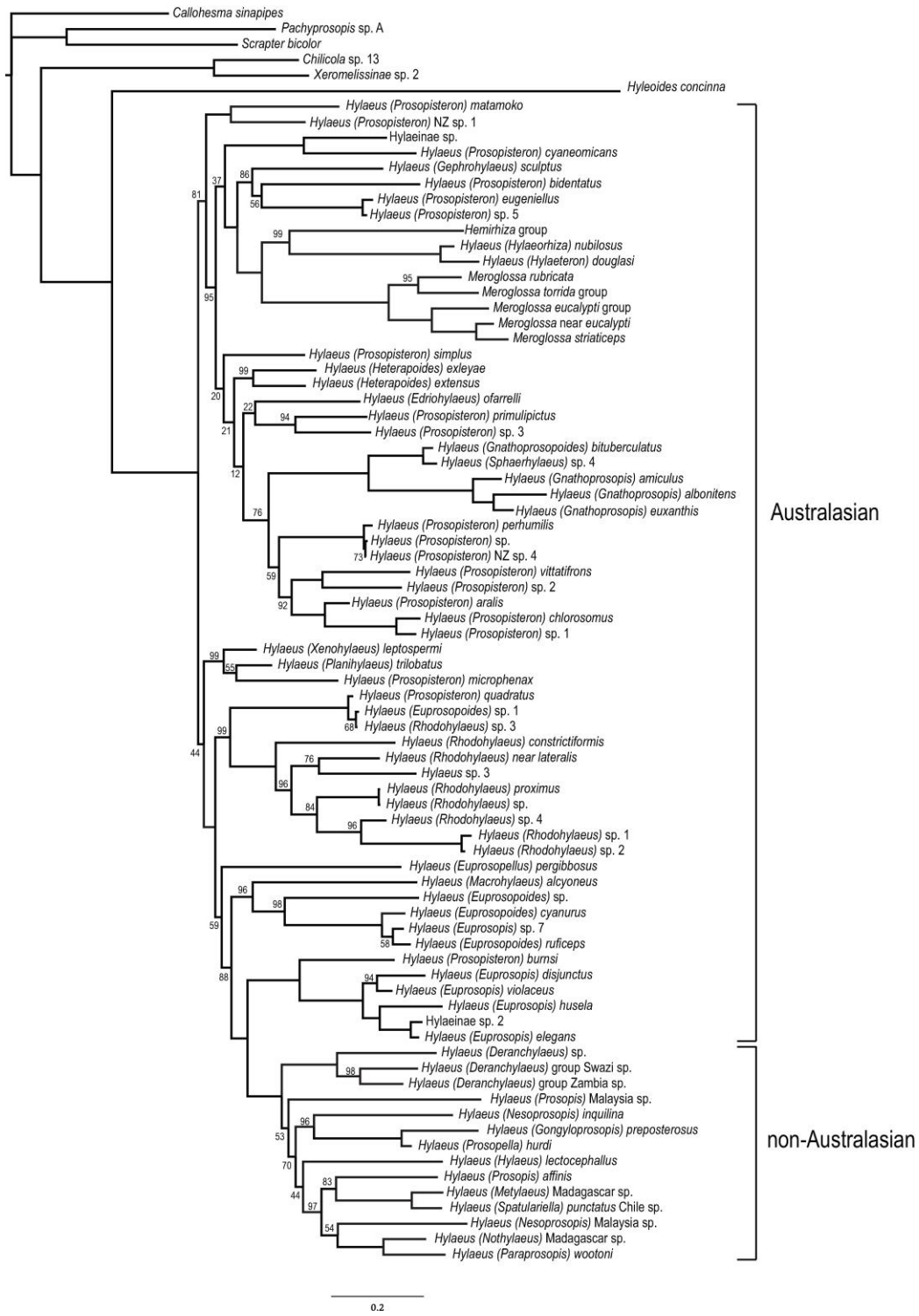



Figure 1. Bayesian consensus phylogram of Hylaeinae and outgroups. Posterior probabilities are indicated for each node under 100. Square brackets indicate Australasian (Australia and New Zealand) and non-Australasian Hylaeinae species used in this study. Other species are outgroups.

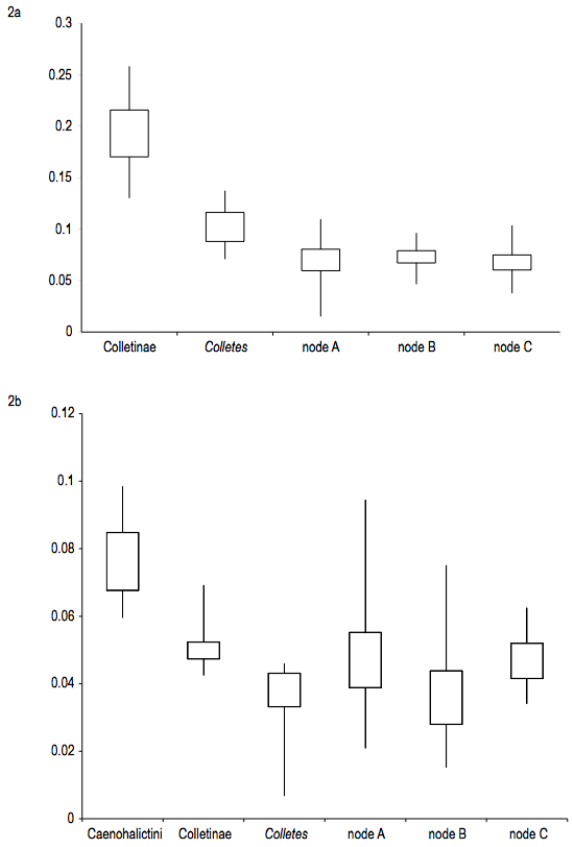



Figure 2. (a) COI pairwise genetic distance estimates from the GTR parameter for three Hylaeinae nodes compared with the divergence of *Mourocolletes* from *Colletes* (Colletinae; Colletidae) and an internal Colletinae clade, '*Colletes*' (Kuhlmann et al. 2009). (b) EF-1α pairwise genetic distance estimates from the GTR parameter for three Hylaeinae nodes compared with the divergence of the subtribe Caenohalictini (Halictini; Halictidae): *Agapostemon*/ *Rhinetula*/ *Dinagapostemon*/ from *Habralictus*/ *Caenohalictus*/ *Ruizantheda*/ *Pseudagapostemon* (Danforth et al. 2004), as well as Colletinae and *Colletes* clades (Almeida et al. 2011) comparable to those used in COI estimates. node A: all Australian species except for the poorly supported clade including *H. (Xenohylaeus) leptospermi, H. (Planihylaeus) trilobatus, H. (Prosopisteron) microphenax*; node B: the divergence of the two New Zealand species from Australian species in the same clade, A1 (see Figure 3 for more detail); node C: divergence of non-Australasian species from *H. (Euprosopis)* species + *H. (Prosopisteron) burnsi*.

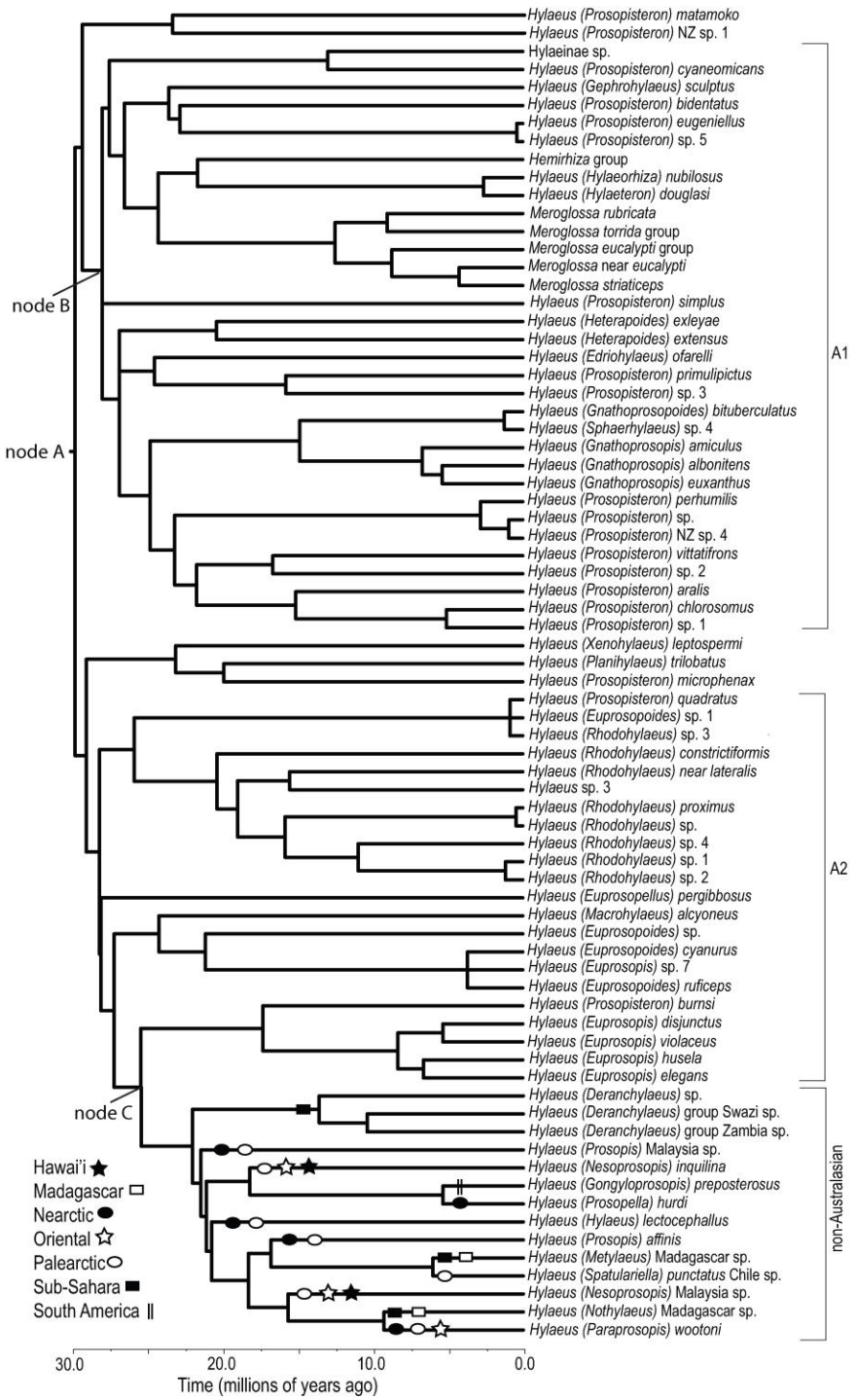



Figure 3. Chronogram of Hylaeinae derived from penalized likelihood transformation of the consensus phylogram. A, B, and C refer to three major nodes that were used to estimate genetic distances (see Fig. 2). A1 and A2 identify the clades used for genetic distance estimates of Node A, as well as Node B in the case of A1 (see Fig. 2). Symbols on branches represent subgenus distributions of non-Australasian *Hylaeus* lineages.

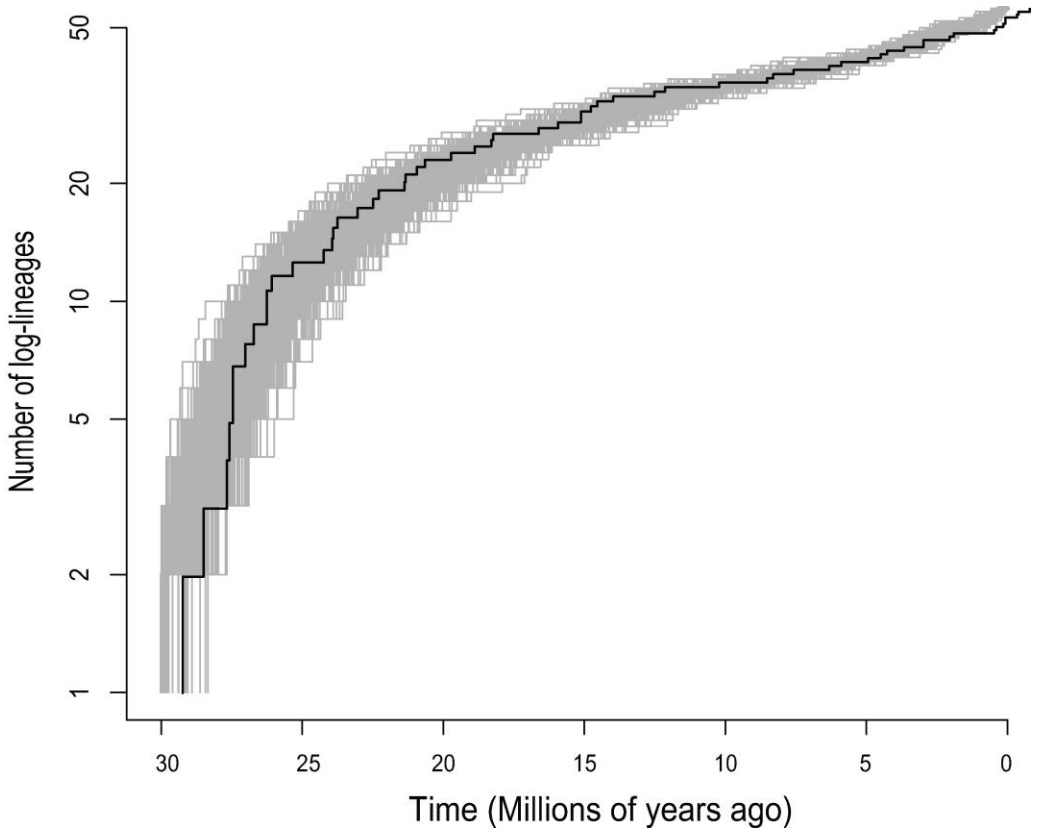



Figure 4. Log-lineages through time plots of Australian *Hylaeus* cladogenesis over time. Grey lines represent 500 randomly selected post burnin chronograms and the black line represents the log-LTT plot from the consensus chronogram.

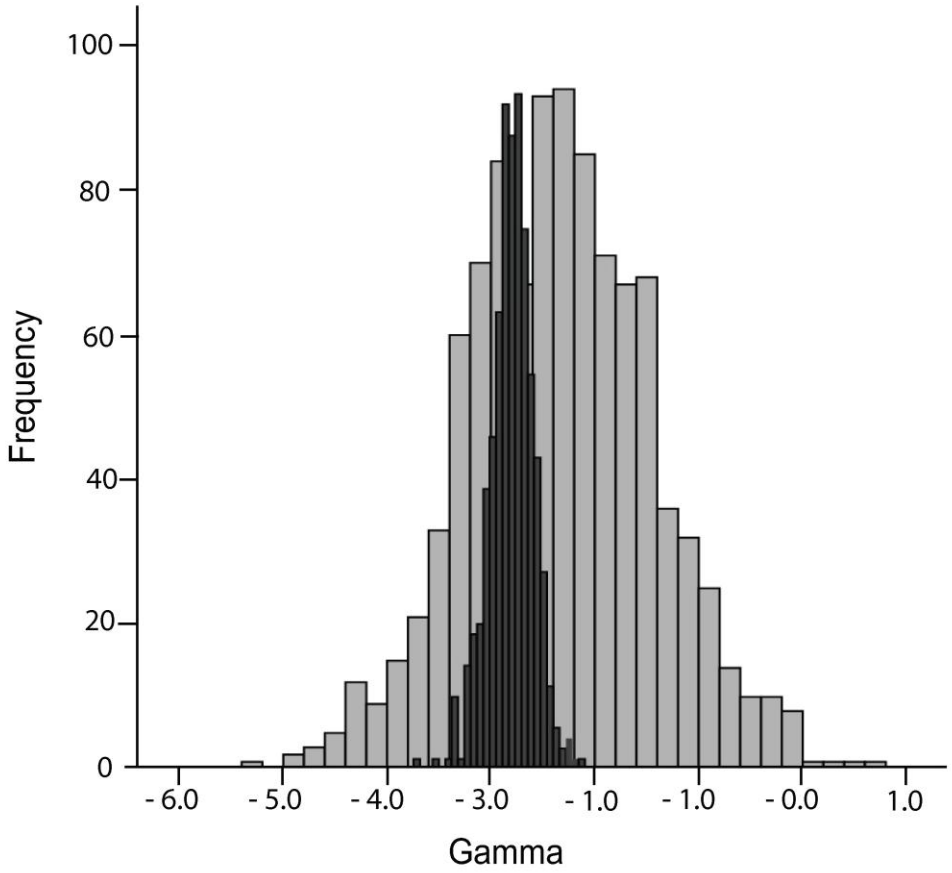



Figure 5. γ (gamma) distributions of 500 randomly-selected phylogenies including the 56 sampled Australian *Hylaeus* species (dark grey), and simulated (light grey) phylogenies based on an estimated 180 species Australian *Hylaeus* species.

**CHAPTER II.**

# Phylogenetics and diversification of the Australian bee subfamily Euryglossinae, and support for the southern African Scrapterinae as sister clade (Colletidae).

Pelin Kayaalp[1,2], Michael P. Schwarz[1], Mark I. Stevens[3]

1. School of Biological Sciences, Flinders University, GPO Box 2100, Adelaide, S.A, 5001.
2. Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa.
3. South Australian Museum, North Terrace, GPO Box 234, SA 5000, and School of Earth and Environmental Sciences, University of Adelaide, SA 5005, Adelaide, Australia.

RUNNING TITLE: Phylogeny of the bee subfamily Euryglossinae.

KEYWORDS: diversification, historical biogeography, dispersal, nesting habit, morphology

## *Abstract*

The subfamily Euryglossinae (Colletidae) is endemic to Australasia, and comprises one of the most speciose bee groups in Australia. Although the taxonomy of this subfamily is well developed, past molecular phylogenetic studies have included only upto six species of Euryglossinae. Using sequences from three genes, COI, EF-1α and 28S, we reconstruct a phylogeny of the Euryglossinae based on 24 species from eight genera, and use this to explore timings and patterns of diversification. We also included 10 species from the southern African subfamily Scrapterinae to further explore a recent analysis indicating that they form the sister clade to the Euryglossinae. Our analyses confirm a sister-clade relationship for these two subfamilies, with a divergence date of approximately 54 - 49 Mya. Analyses on ancestral distributions indicate that an Australian Scrapterinae/Euryglossinae ancestor is more likely than an African or South American ancestor. We explore various biogeographical scenarios for the current distribution of these two subfamilies and argue that a southern dispersal route is more likely, either directly between Africa and Australia, via Antarctica, or with a most recent common ancestor in Antarctica with separate dispersals into Australia and Africa. We also show that the Euryglossinae diverged into two major clades ca. 45 Mya and that these clades differ from each other in wing venation and nesting habit. The ancestral states at the Euryglossinae root node for the three wing characteristics are aligned with one clade, but parsimony and Bayesian approaches differed in their estimation of whether ancestral nesting substrate was wood or soil. We infer that diversification rates in the Australian Euryglossinae have been constant over time, and discuss the potential influence of nesting substrate on dispersal ability, and on the distributions and abundance of bees in tropical and Mediterranean climates.

## *Introduction*

Australia has a highly unusual bee fauna, perhaps the most distinctive in the World (Michener 1965) with around eighty percent of species from two short-tongued families, Colletidae and Halictidae (Michener 2007). Approximately fifty per cent of the Australian bee fauna belong to the family Colletidae (Exley 2001; Michener 2007). The colletids have a predominantly southern hemisphere distribution and most of the older clades occur in South America and Australia (Michener 1965, 1979, 2007; Almeida and Danforth 2009).

Four of the seven colletid subfamilies occur in Australia, Callomelittinae, Eurglossinae, Hylaeinae, and Neopasiphaeinae, as well as a now unassigned genus *Paracolettes*. Callomelittinae is a monogeneric subfamily of 11 species restricted to Australia (Almeida 2008a). Euryglossinae are the most speciose of the four colletid subfamilies found in Australia, with nearly 400 described species (Almeida and Danforth 2009). The subfamily is entirely restricted to the Australasian region, and outside of Australia *per se* it has only minimal representation in New Zealand and New Caledonia (Michener 1965; Donovan 2007). Hylaeinae includes the only globally distributed colletid genus, *Hylaeus*, and the other six hylaeine lineages comprise endemic clades restricted to Australia and New Guinea (Michener 2007; Almeida and Danforth 2009) suggesting an origin of this subfamily in Australia, or perhaps Australia + Antarctica (Almeida et al. in press). Nearly 200 Australian Hylaeinae species have been described (Michener 2007). The low level of *Hylaeus* diversity in South America (Michener 2007) suggests an origin at a stage when interchange between Australia and South America was impeded (Chapter 1). *Hylaeus* originated ca. 30 Mya in Australia (Almeida et al. in press; Chapter 1), shortly after separation from Antarctica, and diversified rapidly in Australia (Chapter 1). The global *Hylaeus* distribution is suggested to have arisen from two events (Chapter 1). The first involved dispersal to New Zealand followed by minimal diversification, and a second dispersal event out of Australia eventually led to an enormous radiation with extant *Hylaeus* lineages occurring in all other regions of the world except Antarctica (Chapter 1). Neopasiphaeinae are restricted to Australia and South America, but with only one genus, *Leioproctus*, found on both continents (Michener 2007; Almeida et al. in press). However, the South American and Australian

*Leioproctus* belong to separate clades and the genus itself is paraphyletic (Almeida and Danforth 2009). There are nearly 300 Australian Neopasiphaeinae species (Almeida 2008a; Almeida et al. in press).

Interestingly, all species of the two largely Australian subfamilies, Hylaeinae and Euryglossinae transport pollen in the crop, rather than on external scopae (Michener 1965, 2007; Houston 1975, 1981). Some earlier studies suggested that this mode of pollen transport may reflect a synapomorphy between Euryglossinae and Hylaeinae and that they were sister lineages (Michener 1965, 1979; Houston 1969), but more recent studies indicate that this is not the case (McGinley 1980; Almeida and Danforth 2009). There is now strong evidence from molecular phylogenies that the South American subfamily Xeromelissinae is sister clade to Hylaeinae, and that the southern African Scrapterinae is sister to Euryglossinae (Almeida and Danforth 2009; Almeida et al. in press). The Australian Callomelittinae is sister to the Australian + South American Neopasiphaeinae (Almeida and Danforth 2009; Almeida et al. in press), and the Australian genus *Paracolletes* (*incertae sedis*, Almeida et al. in press) is sister to the South American subfamily Diphaglossinae (Almeida and Danforth 2009; Almeida et al. in press). That each subfamily shares a common ancestor with another Southern Hemisphere colletid subfamily that has a current distribution disjunct from its own, all with an estimated crown age of at least 54 My (Almeida et al. in press), presents a classic Gondwanan hypothesis of distribution. However, an origin of the colletids at 70 Mya (± 15 Mya) (Almeida et al. in press) post-dates the separation of South America from Africa, and indicates that long-range dispersals may have played an important role in the present-day distributions of colletids. Although South America and Africa were perhaps only separated by a few hundred kilometers at the end of the Cretaceous (Ali and Krause 2011), the only present day colletid affinities between these two continents are the two widely distributed *Colletes* and *Hylaeus* (Michener 2007).

Here we focus on the relationship between Euryglossinae and Scrapterinae (Almeida and Danforth 2009). Euryglossinae have been considered to have subfamily status for a long time, however, Scrapterinae has traditionally been placed as a genus, *Scrapter*, within Colletinae (Michener 2007). Recently, Scrapterinae has been treated as a tribe of Colletinae (Scrapterini, Melo and Goncalves 2005), and as a monogeneric subfamily,

either as Scraptrinae (Ascher and Engel 2006) or most recently as Scrapterinae (Almeida et al. 2008; Almeida and Danforth 2009). The fact that the most abundant colletid group in southern Africa is the Colletinae genus *Colletes* (ca. 100 species) (Kuhlmann 2005), and the similarity in nesting habit of Scrapterinae to other Colletinae, along with the method of larval provisioning (Michener 1965), supports the inclusion of *Scrapter* in Colletinae. However, morphologically Scrapterinae are unlike Colletinae in that they have two sub-marginal cells rather than three, are elongate, and the seventh sternite of the male protrudes posteriorly (Eardley 1996; Michener 2007). Furthermore, there are similarities in larval morphology between Scrapterinae, Euryglossinae, Hylaeinae, and Xeromelissinae, (McGinley 1980), and this relationship has since been supported by molecular phylogenetic results (Almeida and Danforth, 2009). Scrapterinae and Euryglossinae also have similarities in the tubercles at the margins of the female basitibial plate (Michener 2007).

Scrapterinae contains 42 described species (Davies and Brothers 2006) endemic to southern Africa with diversity greatest in the winter rainfall Succulent Karoo biome of the western and northern Cape in South Africa (Davies et al. 2005), but with one known species that ranges into Kenya (Michener 2007). Euryglossinae is endemic to Australia and surrounding islands (Australasia), with one species, *Euryglossina (Euryglossina) proctotrypoides* also found in New Zealand and three species in New Caledonia - one *Euryglossa* and two *Euryglossina* species (Donovan 2007). A single specimen of a *Euryglossina (Euryglossina)* species has been recorded from South Africa, but is considered to be an anthropogenic introduction (Michener 2007). The common ancestor of Scrapaterinae and Euryglossinae is likely to be younger than the rifting of Africa from the rest of Gondwana (Almeida et al. in press), thus implying a long-distance dispersal event linking these two geographically disjunct groups.

Here, we use mitochondrial and nuclear DNA sequences to reconstruct a phylogeny of the Euryglossinae and Scrapterinae in order to: (i) examine the robustness of the proposed sister-relationship between the two subfamilies using more extensive taxon sampling; (ii) provide an estimate of the divergence of the most recent common ancestor between these two groups; and (ii) contrast diversification patterns of Euryglossinae with other Australian bees.

## *Methods*

### TAXON SAMPLING

Euryglossinae were sampled from around Australia, covering tropical, sub-tropical, temperate, and arid regions. A total of 24 species from eight of the 13 described genera were included in our analyses, including five species from GenBank. We also included ten southern African Scrapterinae species in our study, including five species from GenBank. Collection locations and GenBank accession numbers are summarised in Table 1. In order to confirm the sister relationships inferred by Almeida and Danforth (2009) we included four Hylaeinae and two Xeromelissinae species. We also included *Callomelitta antipodes* (Callomellitinae) and a *Trichocolletes* species (Neopasiphaeinae) as outgroups. Hylaeinae + Xeromelissinae represents a sister group to the Euryglossinae + Scrapterinae, and *Callomelitta* and *Trichocolletes* comprise more basal colletid groups (Almeida and Danforth 2009). Voucher specimens sequenced as part of this study are located at Flinders University, South Australia. Details of the voucher specimens for those species obtained from Genbank can be found using the relevant accession number (Table 1).

### DNA EXTRACTIONS AND SEQUENCING

DNA extractions were made from individual specimens using the QIAamp® DNA Micro Kit. PCR reactions were carried out in 25 µl volumes, each containing 2.0 µl of DNA, 2.5 µl 10x Buffer, 2.5 µl MgCl2 (25mM), 2.5 µl dNTP mix (10 mM of each dNTP), 1 µl of each primer (5 mM), 13.3 µl H2O, and 0.1 µl SuperthermTM Taq DNA polymerase (5 U/µl). Primer specifications and PCR conditions are summarised in Tables 2a and 2b. The mitochondrial gene cytochrome *c* oxidase I **(**COI) is commonly used in phylogenetics for inferring more recent divergences. Two regions of COI were used in our study in order to maximise overlap with other bee phylogenies (1425 bp). Two nuclear genes were used, the F2 copy of elongation factor 1$\alpha$ (EF-1$\alpha$, 747 bp) and 28S (858 bp), both of which have also been used in insect phylogenetic studies and have demonstrated utility for recovering deeper divergences (Danforth et al. 2005). Primers used in 28S amplification covered the D1, D2 and D3 regions. A total of 3030 nucleotides were included in the analyses covering all three gene regions. Alignments

were generated using the CLC Free Workbench package (CLC bio A/S version 4.6). The CLC workbench uses a progressive alignment algorithm (Feng and Doolittle 1987) in order to create a multiple alignment. Adjustments to ambiguous alignments were made by eye and unalignable loops in 28S were excluded from the analyses, as was a large unalignable intron in EF-1α.

**PHYLOGENETIC ANALYSIS**

Bayesian Inference (BI) was carried out using the Unix version of MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). Because of third codon saturation problems in Hymenoptera, including bees (e.g., Dowton and Austin 2002; Schwarz et al. 2004), we partitioned CO1 and EF-1α (F2) into two parts each, comprising 1st and 2nd codons together and the 3rd separately, giving rise to four partitions, with 28s treated as a separate fifth partition. We used a general time reversible (GTR) model with gamma distributed (Γ) rates with a proportion of invariant sites (I) assumed. We used GTR (a 6-parameter rate transition matrix) because this is the most general substitutional model and can incorporate simpler models, such as the K2P and HKY models if these emerge from the MCMC parameter space. Two sets of four Monte Carlo Markov chains with Metropolis Coupling were run in parallel for 40 million generations, sampling every 1000th generation and retaining node support for compatible groups with less than 50% bootstrap support. Convergence between parallel runs was assessed using the average standard deviation of split frequencies, and log likelihood (LnL) values were plotted using Tracer version 1.41 (Rambaut and Drummond 2007, http://beast.bio.ed.ac.uk/Tracer) to determine when stationarity was achieved.

**ESTIMATING DIVERGENCE TIMES**

No known fossils of Euryglossinae or Scrapterinae exist, and we therefore used genetic distance measures to estimate and compare nucleotide substitution rates of COI and EF-1α to infer an approximate age (e.g., Regier and Shultz 2001) for two key nodes in our phylogeny under the assumption that divergence age will approximately correspond to genetic distance. Genetic distance estimates were applied to: (i) the node uniting Scrapterinae and Euryglossinae, and (ii) the node uniting Euryglossinae Group A (*Pachyprosopis, Euryglossina*, and *Euryglossula*) with Group B (*Brachyhesma*,

*Callohesma*, *Euhesma*, *Hyphesma*, *Xanthesma*) (see *Phylogenetic Results* and Fig. 3 for descriptions of Groups A and B).

We used closely related clades to estimate relative divergence times under the assumption that base frequencies and substitution rate dynamics are most similar between closely related clades. Boxplot graphs of the estimated distances were used to graphically compare and estimate approximate node ages across groups. We used a general-time reversible (GTR) model of evolution calculated in PAUP* to estimate and compare nucleotide substitution rates of COI and EF-1α. Although the use of models that are fit to each group separately might provide more accurate results for comparing distances within each group, the use of separate models would not allow for meaningful comparisons between groups.

*Hylaeus* and Colletinae are the only two colletid groups for which there are large-scale dated phylogenies using COI data. Euryglossinae/ Scrapterinae COI divergence estimates were compared with the root node of *Colletes* and *Mourecotelles* (Colletinae: Colletidae) (Kuhlmann et al. 2009), which we will refer to as the 'Colletinae node'. Because there are a number of polytomies in the phylogeny, we also generated distance estimates at an internal node within the Colletinae phylogeny (Kuhlmann et al. 2009), which we will refer to as the '*Colletes* node'. We also used the root node of the Australian *Hylaeus* (Hylaeinae: Colletidae) where it diverged into two major clades (the poorly supported clade consisting of *H. (Xenohylaeus) leptospermi*, *H. (Planihylaeus trilobatus)*, and *H. (Prosopisteron) microphenax* was excluded) (Chapter 1) to compare COI distance, which we will refer to as the '*Hylaeus* node'. We fitted only the 1$^{st}$ and 2$^{nd}$ codon positions of COI because the third codon positions in bees can approach saturation for even moderately old divergences (e.g., Dowton and Austin 2002; Schwarz et al. 2004).

Our EF-1α genetic distance plots were calibrated using Colletinae and *Hylaeus*, as well as a node within the sister family to Colletidae, Halictidae. Colletinae EF-1α sequences were available for a subset of the species used in the COI estimates (Almeida et al. 2011), however the same *Mourecotelles* species enabled estimates of a comparable

'Colletinae node', and we again estimated an internal *Colletes* node. We also estimated distances at the root node of the sub-tribe Caenohalictini (Halictini: Halictidae) (Danforth et al. 2004): *Agapostemon/ Rhinetula/ Dinagapostemon* and *Habralictus/ Caenohalictus/ Ruizantheda/ Pseudagapostemon*, as well as for the same *Hylaeus* node used for COI estimates.

The crown age for the subfamily Colletinae has been estimated to be 33 Mya (Almeida et al. in press) and the *Colletes* crown age has been estimated to be 30 Mya (Almeida et al. in press). There are a number of internal fossil calibrations for the halictines that provide conservative estimates of divergences within that clade (Danforth et al. 2004; Brady et al. 2006). Within their larger Halictidae phylogeny Danforth et al. (2004) calibrated Caenohalictini with a minimum crown age of 23 Mya, based on the Dominican amber fossil *Eickwortapis dominicana* from the same tribe, and retrieved a crown age for Caeonohalictini of 36-66 Mya, with a mean age of 45-49 Mya (Danforth et al. 2004). The crown age of *Hylaeus* has been estimated to be approximately 30 Mya (Almeida et al. in press; Chapter 1).

In addition to estimating node ages using distance measures, we calculated relative divergence times using penalized likelihood (PL) rate smoothing implemented in r8s v1.71 (Sanderson 2003) with the smoothing parameter determined by a cross-validation method. Outgroup taxa with a younger estimated crown age were pruned using TreeEdit v1.0a10 so that the focal group could be presented in as large a space as possible. *Callomelitta antipodes* was retained in order to utilise and 'estimate' a 'root' age. We used 500 post-burnin phylograms, obtained by filtering the post-burnin trees to the consensus tree topology. We compared this to a relaxed clock estimate using a Bayesian Markov chain Monte Carlo approach in BEAST v1.3 (Drummond and Rambaut 2003). We used the uncorrelated log-normal relaxed clock and a substitution model with gamma distributed and invariant site heterogeneity (GTR+I+Γ) model and constrained the topology at nodes where there was 100 posterior probability support inferred by the Bayesian consensus phylogeny obtained from the MrBayes analysis. In response to the results of the genetic distance estimates (see *Estimation of divergence times*), we set a crown age for the divergence between Scrapterinae and Euryglossinae at 49 Mya. This is at the upper limit of the inferred estimate for Caeonohalictini (45-49 Mya, Danforth et

al. 20004), and at the lower limit of the 54 (± 12) Mya estimate for the Scrapaterinae + Euryglossinae crown age made by Almeida et al. (in press). In order to increase confidence in these age estimates we also constrained the Scrapterinae + Euryglossinae node to 54 Mya (Almeida et al. in press), and the root node of the outgroup to 60 Mya (Almeida et al. in press).

**ESTIMATING HISTORICAL BIOGEOGRAPHY**

Ancestral distributions were inferred using a maximum likelihood approach implemented in Lagrange build 20110117 (Ree and Smith 2008). Species were coded as having their distribution as Australia (A), Africa (B) or South America (C) (see Table 1), and no species occurred in more than one region. Because we did not want to introduce any preconceptions into possible historical scenarios, historical ranges were not restricted, nor were dispersal possibilities. These settings for Lagrange therefore mean that historically, a species could have conceivably occurred in South America at the same time as it occurred in Australia, and that dispersals between Australia, South America and Africa were not constrained. Lagrange estimates a probability for the ancestral region of each branch leading from a bifurcating node. Lagrange sometimes reports the probabilities for more than two possible historical scenarios. In such cases alternative scenarios were only reported if their log likelihood value was less than 1.96 units lower than the most likely scenario, since differences greater than this represent outcomes that are not significantly supported at an alpha level of 0.05. When there were two or more alternative outcomes that differed by less than 1.96 LnL units, these outcomes are reported and discussed.

**ESTIMATING DIVERSIFICATION PATTERNS IN AUSTRALIAN EURYGLOSSINAE**

We used log-lineage through time (LTT) plots to graphically explore patterns of diversification for Euryglossinae, with the understanding that speciation and/or extinction rates can be underestimated if cryptic diversity is not recognised or if older lineages are not well sampled (Ricklefs 2007). In order to take phylogenetic uncertainty into account, we randomly selected 500 phylograms from a post-burnin sample of 20,000 trees sampled over 40 million generations, and subjected these to the same r8s

PL transformation applied to the consensus phylogram. We then used the 500 resulting chronograms and the consensus chronogram to generate the log-LTT graphs using the mltt.plot command in the APE module (Paradis et al. 2004) in the R package.

We used the same 500 randomly-sampled trees to derive a γ (gamma) statistic (Pybus and Harvey 2000), as used in the constant-rates test (CR) in LASER v2.2 (Rabosky 2008), which serves as an indication of whether a constant rate of diversification can be rejected, and whether diversification rates in a clade have changed over time. A significant γ value lower than -1.645 (one-tailed test) indicates a rejection of constant-rate diversification in favour of a temporal decrease in diversification, with lower values indicating greater diversification closer to the root node. A pure-birth model can be rejected at the 95% level if $\gamma < -1.96$ or $\gamma > 1.96$. (Pybus and Harvey 2000). Incomplete sampling will tend to produce γ values that will give the appearance of higher rates of cladogenesis closer to the root (Pybus and Harvey 2000), a confounding factor from which our sampling regime suffers.

We first calculated observed γ values using the 500 random trees for the available 24 species. We then subjected the same 500 trees to a Monte Carlo constant rates (MCCR) test (Pybus and Harvey 2000). This method provides a simulated distribution of γ by accounting for incomplete sampling by generating random trees and randomly pruning the number of known species (in this case 400) to the number in the given phylogeny (in this case 24). The type II error associated with the MCCR test means that failure to reject the null is not proof of constant diversification (Pybus and Harvey 2000). We plotted a histogram of the γ values derived from the 500 randomly-sampled trees and a comparative histogram of γ values from simulated trees. If the distribution of values of the simulated trees is the same as our random chronograms, then we cannot say that the observed diversification patterns are a reflection of our sampling regime – specifically undersampling.

We used MEDUSA (Alfaro et al. 2009) to detect any shifts in the rates of diversification in the Euryglossinae. By pruning taxa from the chronogram so that a single representative from each clade is left and by creating a matrix with the total species

number for each clade, MEDUSA can detect whether there have been changes in the rate of diversification at any stage in a pruned phylogeny. Based on our phylogeny (Fig. 1), and also on taxonomic assignments (Michener 2007), we treated each of the eight genera in our tree as single clades. Species numbers for each genus were taken from Michener (2007), and were assigned as follows: *Brachyhesma*: 22, *Callohesma*: 34, *Euhesma*: 60, *Euryglossina*: 75, *Euryglossula*: 7, *Hyphesma*: 7, *Pachyprosopis*: 23, *Xanthesma*: 48.

### ESTIMATING ANCESTRAL CHARACTER STATES

Michener (2007) suggested that there are two distinct groups within Euryglossinae (Group A and Group B) based on wing venation, and our results (see *Phylogenetic results*) somewhat support this. We used *BayesMultiState* (Pagel et al. 2004; available from www.evolution.rdg.ac.uk) to estimate ancestral states for three wing characteristics within Euryglossinae, as well as nesting substrate within Scrapterinae/Euryglossinae, using 500 randomly selected post-burnin trees. The inclusion of the sister group to Euryglossinae, Scrapterinae (see *Phylogenetic results*), when considering life-history traits, such as nesting substrate, is important for understanding the evolutionary history of the group. In addition to the model-based analysis, for visualisation of character state distributions over the phylogeny we used MacClade's (Maddison and Maddison 2000) parsimony reconstruction tool that assumes unordered states. The three wing traits for each species were coded according to: (i) whether the posterior margin of the first submarginal cell is straight, or, sinuate, (ii) whether the second submarginal cell is much less than half as long as the first submarginal cell, or, nearly half as long to more than half as long, (iii) whether the first abscissa of vein Rs is transverse, or, oblique (Michener 2007). Species were assigned as being either soil or wood nesters, based on available genus-level data (Houston 1969; Exley 1968, 1972; Michener 2007), except for *Callohemsa* for which nesting habit is unknown.

## *Results*

### PHYLOGENETIC RESULTS

Although we have included very few species from other colletid groups, our Bayesian inference consensus phylogeny (Fig. 1) supports the finding of Almeida et al. (in press) that Scrapterinae and Euryglossinae are sister groups (99 PP). Monophyly of

Scrapterinae is well supported (100 PP), as is the monophyly of Euryglossinae (100 PP). We found strong support (100 PP) for the monophyly of *Pachyprosopis*/ *Euryglossina*/ *Euryglossula*. We partly follow Michener's (2007, pp 210-211) terminology by referring to these three genera collectively as 'Group A' and all other Euryglossinae as 'Group B', groupings originally based on wing venation (see Discussion, Fig. 3, and Fig. 6). The monophyly of Group B is less well supported (92 PP). The monophyly of each of the genera *Callohesma*, *Euryglossula*, and *Hyphesma* are all well supported (100 PP), while the monophyly of *Euhesma* has very low support (65 PP).

**ESTIMATION OF DIVERGENCE TIMES**

We used four bee groups that have been subjected to dating analyses to estimate and compare genetic distance estimates with: (i) the Euryglossinae/ Scrapterinae node; and (ii) the node where Euryglossinae Group A clade diverged from Group B clade. COI estimates were compared with the root node of the Australian *Hylaeus* (Chapter 1), the root node of Colletinae (Colletidae) (Kuhlmann et al. 2009), and an internal *Colletes* (Colletinae) node (Kuhlmann et al. 2009). EF-1α estimates were compared with the same three nodes, but with Colletinae and *Colletes* data taken from Almeida et al. (2011), as well as with the root node of Caeonohalictini (Halictidae) (Danforth et al. 2004). Our COI estimates (Fig. 2a) suggest that the origin of the most recent common ancestor of Euryglossinae and Scrapterinae is much younger than the Colletinae root node (ca. 33 Mya, Almeida et al. in press), slightly younger than the *Colletes* node (ca. 24 Mya, Almeida et al. 2011), and slightly older than the *Hylaeus* root node (ca. 30 Mya, Almeida et al. in press; Chapter 1), and that the divergence of Euryglossinae Group A from Group B is similar to that of the *Hylaeus* root node. The EF-1α genetic distance estimates (Fig. 2b) suggest that the Scrapterinae/ Euryglossinae and Euryglossinae Group A/Group B node both have a crown age slightly older than Caeonohalictini (45-49 Mya, Danforth et al. 2004), and much older than the Colletinae, *Colletes*, and Hylaeus nodes.

Our three penalized likelihood chronograms all gave similar results (summarised in Table 4) and are more congruent with the EF-1α genetic distance estimates than with the COI estimates, where the divergence between Euryglossinae A and B is older than 30

Mya. The consensus chronogram using an estimated divergence between Scrapterinae and Euryglossinae, based on our genetic distance results is shown in Figure 3a. The two chronograms based on dates inferred by Almeida et al. (in press), using our consensus phylogram, are shown in Figure 3b and 3c.

**ESTIMATION OF HISTORICAL BIOGEOGRAPHY**

In order to infer the ancestral distribution of the Scrapterinae/Euryglossinae clade we considered the outcomes of two major nodes (Fig. 1) in the Maximum likelihood analysis of historical biogeography in Lagrange. Lagrange provides probabilities of the ancestral ranges of the two branches leading from each node. At the root node (Fig. 1, N82), there is (i) a 61.38% probability (lnL -21.56) that the branch leading to *Trichocolletes* (Neopasiphaeinae: Colletidae) had an Australian distribution, and that the branch leading to the other clade, comprised of all other species in the phylogeny (including Scrapterinae and Euryglossinae) had an Australian and/or African distribution, and (ii) a 28.75% probability (lnL -22.32) that both branches had an Australian distribution. At the node leading to the Scrapterinae/Euryglossinae clade (Fig. 1, N81), there is (i) a 71.86% probability (lnL – 21.41) that the branch leading to Hylaeinae/Xeromelissinae had an Australian distribution, and that the branche leading to Scrpaterinae/Euryglossinae had an Australian and/or African distribution, and (ii) a 18.31% probability (lnL – 22.77) that both branches had an Australian distribution. At both nodes the third alternative was statistically unlikely compared with the most likely possibilities and are not reported here.

**ESTIMATION OF DIVERSIFICATION PATTERNS IN EURYGLOSSINAE**

The log-lineages through time (LTT) plot (Fig. 4) of the consensus chronogram and the 500 randomly sampled Euryglossinae chronograms is approximately linear, which can be an indication of a constant birth-death rate (Crisp and Cook 2009). Although we have included only 24 of the estimated 400 Euryglossinae species, the log-LTT plot does not have the Moran-type curve that is often an artefact of undersampling and consequently a false indication of early rapid diversification (Rabosky and Lovette 2008; Crisp and Cook 2009).

The mean distribution of our observed γ (gamma) values from 500 randomly chosen post-burnin chronograms is 0.4 (Fig. 5). Although the MCCR test is used to detect the effect of undersampling (which is inferred by a negative γ value) (Pybus and Harvey 2000), we compared our observed γ distribution to a simulated distribution of 500 randomly chosen trees. The MCCR method simulates trees based on the described number of taxa (i.e. 400 Euryglossinae species) and randomly prunes branches from these trees to match the actual sample size (in our case 24 species), providing a distribution of γ values that would be expected. The mean distribution of the simulated γ is -3.5 (Fig. 5), and reflects the large amount of taxa that we are missing, and the observed value falls outside the 95 % confidence interval. This means we can be confident that the observed mean γ value of 0.4 is not due solely to undersampling. This pattern indicates that there has not been a major shift in euryglossine diversification, however the type II error associated with the MCCR test means that failure to reject the null is not proof of constant diversification (Pybus and Harvey 2000). Estimates from MEDUSA indicate a net diversification rate (b-d) for the Euryglossinae of 1.51, and an extinction fraction (d/b) of 0.986.

**ANCESTRAL CHARACTER STATES**

Under MacClade the ancestral states, at the Euryglossinae root node, of the three wing characteristics could not be determined (Figs. 6a-6c) and the ancestral nesting substrate was inferred to be soil (Fig. 6d). *BayesMultiState* was used to estimate the ancestral states of three wing characteristics at the Euryglossinae root node, as well as for the two major euryglossine groups, Group A and Group B. Ancestral nesting habit was estimated for the same three nodes, as well as for the Scrapterinae/Euryglossinae node (Fig. 3a).

In Group A species the posterior margin of the first submarginal cell is straight, and in Group B it is sinuate (except for *Brachyhesma* sp. where the margin is straight). The inferred ancestral state at the Euryglossinae root node is slightly sinuate (98%), for Group A, straight (>99%), and for Group B, slightly sinuate (98%) (Fig. 3a). In Group A species the length of the second submarginal cell is much less than half as long as the first submarginal cell, and for Group B it is at least nearly half as long (except for

*Brachyhesma* sp. where it is much less than half as long). The inferred ancestral state at the root node is nearly half as long (98%), for Group A, much less than half as long (>99%), and for Group B, nearly half as long (98%) (Fig. 3a). In Group A species the first abscissa of vein Rs is transverse and in Group B it is oblique. The inferred ancestral state at the root node is oblique (89%), for Group A, transverse (>99%), and for Group B, oblique (>99%) (Fig. 3a).

Scrapterinae species were included in the estimation of ancestral nesting habits. All Scrapterinae are soil nesters, the Euryglossinae clade, Group A, consists of both soil and wood nesting species, and Group B has only soil nesting species. A wood nesting ancestor is inferred at the Scrapterinae/Euryglossinae root node (62%), the Euryglossinae root node (79%), and for Euryglossinae Group A (98%), and a soil nesting ancestor is inferred at the Scrapterinae root node (92%), and Euryglossinae Group B (86%) (Fig. 3a).

## *Discussion*

### PHYLOGENETIC RELATIONSHIPS

Despite undersampling, this study provides further support for Almeida and Danforth's (2009) finding that Scrapterinae and Euryglossinae are sister clades, and that Scrapterinae is best placed as a monogeneric subfamily. Although Scrapterinae was originally placed in Colletinae as the genus *Scrapter*, the similarities between Scrapterinae and Euryglossinae were noted at least as far back as McGinley (1980) who used larval characteristics to examine colletid relationships.

Our analyses indicate that the eight euryglossine genera in our phylogeny make up two distinct clades: *Euryglossina*/ *Euryglossula/ Pachyprosopis*, and *Brachyhesma*/ *Callohesma*/ *Euhesma*/ *Hyphesma*/ *Xanthesma*. These groups correspond largely to earlier proposed groupings based on venation of the forewing (Michener 2007; see Figs 6a-6c), except that Michener (2007) groups *Brachyhesma* together with *Euryglossina, Euryglossula* and *Pachyprosopis* because (i) "the posterior margin of the first submarginal cell is straight in those genera, whereas it is typically gently sinuate in others" (Michener 2007: 211), and (ii) the second submarginal cell "is much less than

half as long as the first (or absent)" (Michener 2007: 210) whereas it is "nearly half as long as the first to more than half as long as the first" in all other Euryglossinae genera (Michener 2007: 210). Michener (2007) suggests that the latter characteristic might have resulted from convergent evolution due to small body size. However, he also suggests that *Brachyhesma*, along with *Xanthesma*, is minute and that those two genera "have plesiomorphic characters associated with the base of the first sub-marginal cell [and that] these genera have become minute independently of *Pachyprosopis*, *Euryglossina*, and *Euryglossula*" (Michener 2007: 210). Our phylogeny is concordant with this notion, where *Pachyprosopis*, *Euryglossina*, and *Euryglossula* form a clade separate from the other genera in our phylogeny, including *Brachyhesma*. Together, our molecular phylogeny and Michener's groupings based on wing venation provide good support that there Euryglossinae is made up of two distinct clades.

The main wing feature that unites *Pachyprosopis*, *Euryglossina*, and *Euryglossula* and separates them from other Euryglossinae, including *Brachyhesma*, is that "the first abscissa of vein Rs of the forewing is transverse, not oblique as in other bees, with resultant effects on the angles at the ends of this abscissa" (Michener 2007: 211). Therefore, based on the two Euryglossinae clades inferred by our phylogeny and the general congruence between these clades and wing characteristics, we will refer to *Pachyprosopis*, *Euryglossina*, and *Euryglossula* as Group A and to *Brachyhesma, Callohesma, Euhesma, Hyphesma,* and *Xanthesma* as Group B. Michener (2007) suggests that *Eurylgossula* may be sister to *Euryglossina-Pachyprosopis*, but our phylogeny suggests that *Pachyprosopis* is sister to the others. *Euhesma* has been described as a 'dumping ground' (Exley 2001) for species that are otherwise unclear in their taxonomic position, and the low support (65 PP) we found for the monophyly of *Euhesma* likely reflects this.

**DIVERGENCE TIMES**

Genetic distance estimates were compared to key dated nodes within four other bee groups, *Hylaeus*, Colletinae, *Colletes*, and Caenohalictini. Estimates derived from COI distances suggested that Colletinae was perhaps twice as old as any of the other groups, considering the Colletinae root node has been estimated to be 33 Ma, this would mean then the Scrapterinae/Euryglossinae node would be 15 Ma. Such a young estimation

seems highly unlikely given previous estimates of the same node of approximately 50 Mya (e.g. Brady et al. 2009; Almeida et al. 2011). Estimates derived from EF-1α estimates were substantially different, and relative positions of the boxplots agreed more with published dates. The plotted distributions of the EF-1α estimates suggest that the crown age of Scrapterinae/ Euryglossinae is similar to that of Caenohalictini, which has been dated at 45-49 Mya. This estimate corresponds to the younger end of the estimate by Almeida et al. (in press) of a crown Scrapterinae + Euryglossinae age of 54 Mya (±12 Mya).

Our results reveal that relative rates of COI and EF-1α evolution are not directly comparable with each other. We suggest that the EF-1α estimates are more accurate due to saturation of COI at deep divergences, and this may be exentuated when comparing a widely distributed clade, Colletinae, to mainly Australian taxa. Substitution rates in mitochondrial genes tend to be higher than in nuclear genes, and the reliability of COI, as an indicator of divergence age used alone, has been questioned for estimates of divergences older then 5-10 Mya (Lin and Danforth 2004). COI is best used for species delimitations as substitutions have been shown to be unlinked to population size, as opposed to nuclear genes (Bazin et al. 2006). We therefore relied on the EF-1α estimates use as a calibration point in further analyses.

We applied three separate calibration points, one based on our genetic distance results and two based on the results of Almeida et al. (in press). This enabled us to compare the estimated crown ages of Scrapterinae and Euryglossinae. The crown age of Euryglossinae in all three of our chronograms fell within the 38-60 My age estimated by Almeida et al. (in press). The crown age of the ten Scrapterinae species estimated by our chronograms was around 35-40 My compared with the 17-31 Mya estimate of Almeida et al. (in press). However, *Scrapter* sp. 1 and the five species that were part of Almeida et al. (2011) study (sequences from GenBank) form a monophyletic group and that clade has a crown age of approximately 28 My (17-31 Mya estimated by Almeida et al. in press), whereas the crown age of the clade including *Scrapter nitidus*, *S. striatus*, *S. bicolor*, and *S. chloris* is older. The crown age of Euryglossinae Group A is

approximately 25-15 My younger than the crown age of Group B. As with Scrapterinae above this may be an effect of undersampling older lineages.

## BIOGEOGRAPHICAL IMPLICATIONS OF DIVERGENCE TIMES AND HISTORICAL RANGE

Almeida et al. (in press) suggest an Australia to Africa direction of dispersal for a common Euryglossinae/Scrpaterinae ancestor, based largely on the fact that Euryglossinae abundance is higher than Scrapterinae abundance, and that colletid abundance is higher in Australia. Our results from Lagrange support that scenario, where it is very likely that the ancestral region of Euryglossinae/Scrapterinae and Hylaeinae/Xeromelissinae (Fig. 1, Node 81) included Australia, since a presence in Australia is covered by all most statistically likely possibilities. Inference of an ancestral presence in Australia and/or Africa is indicative of a dispersal event between the two disjunct continents, rather than vicariance because inferred divergence times post-date the separation of Australia and Africa by at least 45 My.

Southern Hemisphere vicariance events involving Africa rely on a stem age predating the separation of Africa from other landmasses, which began 165 Mya (Sanmartín and Ronquist 2004). Gondwanan rifting events that are able to explain a disjunction between African and Australian clades (perhaps via Antarctica) require the African and Australian clades to have diverged at least 105 Mya (McLoughlin 2001), making a vicariance event involving a common Euryglossinae and Scrapterinae ancestor unlikely. The origin of the colletids has been dated at 71(±15) Mya (Almeida et al. in press) and the stem age of Scrapterinae + Euryglossinae clade is estimated at approximately 61 Mya and both dates significantly post-date the rifting of Africa from the other Gondwanan landmasses.

Although long-distance dispersal events that involve extensive barriers can seem improbable, such dispersal in the Southern Hemisphere may have been more important during the Cretaceous and Paleogene when oceanic barriers in the Southern Hemisphere were less extensive than at present (McLoughlin 2001; Sanmartín and Ronquist 2004). The geological connection between Australia, South America, and Antarctica up until about 40 Mya (McLoughlin 2001) allows for the possibility that some taxa had either a

continuous range across these now disjunct continents, or were restricted to any section within the expanse. In that case, we cannot know without further evidence whether the most recent common ancestor of Scrapterinae and Euryglossinae was present across the whole geographic range or a section of what is now Australia, Antarctica, or South America (or a combination of these).

With this in mind, we suggest two possible dispersal routes to explain the disjunct distribution of the southern African Scrapterinae and Australian Euryglossinae: (1) a trans-oceanic dispersal route between Africa and Australia/Antarctica/South America; or (2) a trans-oceanic and over-land route covering south-east Asia, Eurasia and northern Africa.

### *The trans-oceanic route*

Many authors have proposed a direct trans-oceanic dispersal between Australia and Africa, or via Antarctica and now submerged islands (e.g. Baum et al. 1998; Barker et al. 2007; Schweizer et al. 2010; Chenoweth and Schwarz 2011). Almeida et al. (in press) provide insightful discussion on the potential dispersal scenarios from Antarctica or Australia into Africa, as well as an island-hopping route over the Kerguelen Plateau and/or the Crozet Plateau, which are hypothesized to have connected Antarctica, Madagascar, India, and perhaps continental Africa, from the late or mid-Cretaceous until the mid-Tertiary (McLoughlin 2001; Ali and Aitchison 2009). Likewise, there is evidence for an eastern Antarctic + Australian origin for two clades of Parrots, each with a dispersal into Madagascar and then Africa (Schweizer et al. 2010). Like Almeida et al. (in press), Schweizer et al. (2010) argued that the presence of an emergent subaerial Kerguelen Plateau could have been part of the dispersal route. However, the roles of those Indian Ocean features as part of a potential dispersal route have recently been strongly questioned (Ali and Aitchinson 2009; Ali and Krause 2011). Apparently only small portions of the present-day Kerguelen Plateau were sub-aerial with large traverses between the sub-aerial sections (Ali and Aitchison 2009). Almeida et al. (in press) do suggest that even so, any emergent areas could have provided potential stepping-stones. Eastward dispersals from Africa to Australia have been inferred for the Baobabs (Baum et al. 1998), and for the bee tribe Allodapini (Apidae) (Schwarz et al. 2006; Chenoweth and Schwarz 2011).

Two independent divergences between sister lineages of Australian and South African Proteaceae have also been inferred to post-date a vicariance explanation (Barker et al. 2007). One divergence has been dated at approximately 65 Mya, and the second approximately 40 Mya (Barker et al. 2007). The first date is close to the inferred divergence date for Scrapterinae + Euryglossinae, and the latter is close to that of the allodapine lineage that dispersed from Africa to Australia, suggesting that there may have been increased opportunities for dispersal during the Paleocene-Eocene. At that stage, Australia and Antarctica were connected, and presumably the circum-polar currents that were initiated by the movement of Australia away from Antarctica ca. 30 Mya (McLoughlin 2001) were not present. This is a strong indication that the ocean and wind currents in the southern hemisphere prior to the separation of Australia and Antarctica were likely to have been different from what we see today.

The proposal of Antarctica as an ancient biotic gateway, and a possible ancestral ground is plausible given that it was connected to the southern extremes of Australia and South America. However, although Antarctica remained largely ice free during the Eocene (Thorn and DeConto 2006), it was occupied by a cold climate flora dominated by Proteaceae, Podocarpaceae and *Nothofagus* (Truswell and Macphail 2009). The existence of bees, which are generally more adapted to Mediterranean climates (Wcislo and Cane 1996), on Antarctica during the Miocene with dark winters and mean annual temperatures of 10–15° C (Francis and Poole 2002), although possible, should be considered carefully. Furthermore, many organisms with post-Gondwanan affinities between South America and Africa are tropically distributed (e.g. Raven and Axelrod 1974; Sanmartín et al. 2010), although, the temperately distributed Southern Hemisphere Proteaceae has been inferred to have dispersed both between Australia and South Africa, as well as between South America and Africa (Barker et al. 2007). We agree that dispersal between Australia (and possibly Antarctica) and Africa is most plausible (Almeida et al. in press), however we will briefly discuss the possibility of a Eurasian route.

### *The Eurasian route*

Dispersal of an African ancestor via Eurasia into Australia has been inferred for the bee genus *Braunsapis* (Allodapini) (Fuller et al. 2005) some 8 Mya, well after the beginning of the Australian/Laurasian interchange that began with the collision of these two plates (Lee and Lawver 1995). At the time of inferred divergence between Scrapterinae and Euryglossinae however (ca. 50-60 Mya), the distance between the northern tip of Australia and the southernmost island of Asia was approximately 2,100 km (Lee and Lawver 1995). In comparison, the distance between the southern tip of Africa and the western tip of Antarctica 40 Mya was 2,600 km (Houle 1999) and was probably less at 50 Mya. This means that around 50 Mya the stretch of water between Africa and Antarctica was a similar distance as the oceanic barrier between and Australia and Asia - enough to discount one or the other as impassable. Furthermore, it has been shown that at least two other colletid groups, *Hylaeus* (Hylaeinae) and Neopasiphaeinae have dispersed across the Tasman Sea from Australia to New Zealand over the past 30 My (Chapter 1; Almeida et al. in press), and this would entail a trans-oceanic dispersal of at least 1,500 km (Wallis and Trewick 2009). Only since the Miocene has there been a high frequency of exchange between Africa and Laurasia (Gheerbrant and Rage 2006), largely associated with the closing of the Tethy's Sea from approximately 25 Mya (Hrbek and Meyer 2003; Gheerbrant and Rage 2006).

The Eurasian route seems less likely because of the absence of any related extant (or known fossilised) lineages to Euryglossinae or Scrapterinae in the Palearctic or Asia (Michener 2007). If considered, this route would necessitate extensive extinction throughout south-east Asia, Asia, Europe, and northern Africa. Furthermore, a Eurasian/northern African dispersal route would have required tropically adapted species, and the absence of Scrapterinae from tropical regions (Eardely 1996), and the generally more temperate and arid distributed Euryglossinae (Exley 2001) suggests that this is unlikely.

**DIVERSIFICATION PATTERNS**

The log-LTT plot suggests that diversification in Euryglossinae has occurred at a constant rate, and this inference is not rejected by our MCCR $\gamma$ (gamma) simulations. The positive $\gamma$ value inferred for Euryglossinae ($\gamma = 0.4$) reflects the occurrence of

speciation events closer to the present, and also contrasts with the negative value that is expected from undersampling (Pybus and Harvey 2000). A high extinction ratio, which is inferred from our MEDUSA results (d/b = 0.986), however, might result in a positive γ distribution (Rabosky and Lovette 2008). Although the assumption of constant diversification through time allows for a more confident interpretation of the results from MEDUSA (Pybus and Harvey 2000; Rabosky 2006), without a fossil record there is little room for enhancing the reliability of these results (Rabosky 2010).

The more-or-less constant rate of diversification inferred for Euryglossinae (with higher speciation closer to the present) contrasts with the pattern of diversification of another common Australian colletid, *Hylaeus* (Hylaeinae), which showed rapid cladogenesis early on and subsequent slowing down in diversification (Chapter 1). A constant diversification rate was also inferred for the Australian allodapine bees (Chenoweth and Schwarz 2011). Although the Euryglossinae log-LTT plot is more convex in shape and does not indicate the increased diversification rate of the Australian allodapines from 10-6 My (which Chenoweth and Schwarz 2011 attribute to an increase in speciation at the 'Hill Gap', Hill 1994) both these groups show a very clear difference from the pattern of diversification in *Hylaeus* (Chapter 1).

It has been suggested that higher diversification rates in some clades might be positively linked with clade age (McPeek and Brown 2007), the number of species in a clade (Ricklefs 2006; e.g., Mullen et al. 2011), or the size of the area a clade inhabits (Ricklefs 2006). In comparing the Australian *Hylaeus*, Euryglossinae, and Allodapini, we might expect that the higher rate of diversification inferred from the Australian *Hylaeus* might mean that it is an older clade, it has the highest number of species, and that it has the widest distribution. Interestingly, the number of Euryglossinae in Australia is nearly twice that of *Hylaeus*, despite the two groups having comparable ranges in Australia, and has an older crown age (*Hylaeus* 30 Mya, Euryglossinae 48 Mya, Almeida et al. in press).

**ANCESTRAL CHARACTER STATES**

Results from *BayesMultiStates* suggest that of the wing traits used to differentiate between Group A and Group B species, the characters in the Group B species are most

likely to be the ancestral states. This means that the wings of the most recent common ancestor of Group A and Group B most likely had: (i) a 2nd submarginal cell nearly half as long to more than half as long as the first, (ii) a slightly sinuate posterior margin of the 1st submarginal cell, (iii) and an oblique first abscissa of vein Rs. Evolution of the two traits related to the 1st and 2nd submarginal cells in *Brachyhesma* (Group B) and Group A appear to be convergent.

The parsimony analysis in MacClade and the Maximum Likelihood analysis in *BayesMultiStates* differed in their estimation of inferred ancestral states for nesting habit, with a soil nesting ancestor for Euryglossinae estimated under parsimony and a wood nesting ancestor estimated under Maximum Likelihood. Despite all species outside of Group A in the phylogeny being soil nesters, the ancestral state at all estimated nodes, apart from that of Scrapterinae and Group B, was inferred to be wood nesting. This might be an effect of the long branch leading to Group A, however, must be considered as an alternative possibility, albeit cautiously. The sister clades to Scrapterinae/Euryglossinae, Hylaeinae/Xeromelissinae are also wood nesting species (Michener 2000), and if results from our *BayesMultiState* character state reconstructions are correct, this implies there have been multiple reversions to soil nesting within Scrapterinae/Euryglossinae.

Life history data for Euryglossinae are scarce with nesting habit described from only eight of the 13 genera (Houston 1969; Exley 1972; Michener 2007). Nesting habit is known from at least a single species from all genera in our phylogeny except *Callohesma*. Documented nesting habits suggest that most Euryglossinae species nest in soil (Exley 1968; Houston 1969; Michener 2007), and the only known exceptions are *Euryglossina* and *Pachyprosopis*, two of the three Group A genera in our phylogeny. *Euryglossina* nests have been recorded from preformed cavities in hardwood as well as abandoned beetle holes in dead wood (Exley 1968; Houston 1969). *Pachyprosopis* species have been reported to nest in both wood and soil. *Pachyprosopis haematostoma* has been observed occupying the vacated burrows of beetle larvae (Houston 1969), and two other *Pachyprosopis* species nests have been described from ‘termite soil’ at the bases of or in hollow trees (Houston 1969; Exley; 1972; Michener 2007). Nests are only known from one species each of *Brachyhesma, Hyphesma, and Xanthesma* and for other

genera the number of species with nest habit recorded are also few, so these delimitations will need to be re-examined once more nests from more species are studied.

**NESTING HABIT AND DISPERSAL SUCCESS**

It is surprising that a wood-nesting ancestor was inferred for three of the four major nodes considered in the phylogeny, especially at the Scrpaterinae/Euryglossinae node, considering all Scrapterinae and most Euryglossinae are soil nesting species. It is obvious, however, that extant lineages in both families are largley soil nesters and this may be relevant when considering the lack of dispersal events in both Euryglossinae and Scrapterinae, where the higher species richness of euryglossines in Australia but their limited success in over-water dispersal might be linked to their largely soil nesting habit. Evidence from extant species provides information for only a single over-water dispersal of a Scrapterinae/Euryglossinae ancestor, however, if we consider the evidence for a wood-nesting ancestor, then we must also consider extinction events of potentially a number of those lineages that may have dispersed between continents. Stem-nesting (as in hylaeines and allodapines) might better facilitate successful over-water dispersal, providing a substrate better suited to successfully being carried by wind or vegetation rafts, as well as comprising a nesting habitat that is less vulnerable to inundation in tropical climates.

Where soil nesting habitat is available and less prone to disturbance, i.e., in drier climates, soil nesters might flourish because nesting opportunities are not as restrictive as for stem-nesters. The prevelance of bees in Mediterranean climates, where dry substrate is more readily available when compared with humid environments, is assumed to be for this reason (Michener 2007). As well as Euryglossinae, Neopasiphaeinae (Colletidae) and Halictinae (Halictidae) species that are endemic to and common in Australia are also largely soil nesters (Michener 2007). The relative abundance of stem-nesting lineages in humid regions might in part be because it provides better defence against fungal attack on larval provisions (Michener 2007).

The inclusion of *Euryglossina* (Group A, stem nesting) species from New Zealand and Caledonia, and *Euryglossa* (Group B, soil nesting) species from New Caledonia in

future dating studies might provide further insight into whether nesting substrate contributes to relative dispersal success. With the evidence at hand, we might expect that *Euryglossa* is more likely to have arrived in New Caledonia via anthropogenic dispersal and that *Euryglossina* dispersed naturally.

## *Conclusions*

These results confirm the sister relationship between the Australasian Euryglossinae and the southern African Scrapterinae, and that they split at a time when Africa and Australia were no longer part of Gondwana. This provides further support for the role of long distance dispersal in the present-day distributions of southern hemisphere biota. We propose that the limited dispersal of either group subsequent to their split might partly be an artefact of a now largely soil nesting habit, but that unknown wood-nesting lineages might have been prevelant and enabled a wider distribution than that of extant soil nesting species.

**ACKNOWLEDGEMENTS**

Terry Houston of the Western Australian Museum (WAM) identified many Euryglossinae species, and Connal Eardley of the Agricultural Research Council in Pretoria identified Scrapterinae species. Simon van Noort's (Iziko Museum, Cape Town) assistance in the field in South Africa, and that of Luke Chenoweth, Scott Groom, Sally Harradine, Michael McLeish, and Sandra Rehan in Australia is very much appreciated. WAM, and Kirstenbosch Research Centre in South Africa hosted PK. Michael McLeish assisted with graphics. This work was funded by ARC grants awarded to MPS, as well as Holsworth Wildlife Fund, and Sir Mark Mitchell awards to PK.

**LITERATURE CITED**

Alfaro, A.E., F Santini, C. Brock, H. Alamillo, A. Dornburg, D.L. Rabosky, G. Carnevale, and L.J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. P. Natl. Acad. Sci. USA. 106: 13410–13414.

Ali, J.R., and J.C. Aitchison. 2009. Kerguelen Plateau and the Late Cretaceous southern-continent bioconnection hypothesis: tales from a topographical ocean. J. Biogeogr. 36: 1778–1784.

Ali, J.R., and D.W. Krause. 2011. Late Cretaceous bioconnections between Indo-Madagascar and Antarctica: refutation of the Gunnerus Ridge causeway hypothesis. J. Biogeogr. doi:10.1111/j.1365-2699.2011.02546.x

Almeida, E.A.B. 2008a. Revised species checklist of the Paracolletinae (Hymenoptera, Colletidae) of the Australian Region, with the description of new taxa. Zootaxa. 1891:1-24.

Almeida, E.A.B. 2008b. Colletidae nesting biology (Hymenoptera: Apoidea). Apidologie. 39:16–29.

Almeida, E.A.B., L. Packer, and B.N. Danforth. 2008. Phylogeny of the Xeromelissinae (Hymenoptera: Colletidae) based upon morphology and molecules. Apidologie 39:75–85.

Almeida, E.A.B., and B.N. Danforth. 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50:290–309.

Almeida, E.A.B., M.R. Pie, S.G. Brady, and B.N. Danforth. in press. Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the World. J. Biogeogr.

Ascher, J.S., and M.S. Engel. 2006. On the availability of family-group names based on *Scrapter* (Hymenoptera: Colletidae). Entomol. News. 117:117-119.

Baum, D.A., R.L. Small, and J.F. Wendel. 1998. Biogeography and floral evolution of baobabs (Adansonia, Bombacaceae) as infer red from multiple data sets. Syst. Biol. 47(2):181-207.

Barker, N.P., P.H. Weston, F. Rutschmann, and H. Sauquet. 2007. Molecular dating of the 'Gondwanan' plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana. J. Biogeogr. doi:10.1111/j.1365-2699.2007.01749.x

Brady, S.G., S. Sipes, A. Pearson, and B.N. Danforth. 2006. Recent and simultaneous origins of eusociality in halictid bees. Proc. R. Soc. B. 273:1643–1649.

Brady, S.G., Larkin, L., Danforth, B.N., 2009. Bees, ants, and stinging wasps (Aculeata), pp 264-269 in *The Timetree of Life* Hedges, S.B. and Kumar, S., eds. Oxford Univ. Press, U.K.

Bazin, E., Glemin, S., Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. Science. 312, 570-572.

Chenoweth, L.B., and M.P. Schwarz. 2011. Biogeographical origins and diversification of the exoneurine allodapine bees of Australia (Hymenoptera, Apidae). J. Biogeogr. 38:no. doi: 10.1111/j.1365-2699.2011.02488.x.

Crisp MD, and L.G. Cook. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. Evolution. 63(9):2257-2265.

Danforth, B.N., H. Sauquet, and L. Packer. 1999. Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae), based on parsimony and likelihood analyses of nuclear EF-1$\alpha$ sequence data. Mol. Phylogenet. Evol. 13(3):605–618.

Danforth, B.N., S.G. Brady, S.D. Sipes, and A. Pearson. 2004. Single-copy nuclear genes recover cretaceous-age divergences in bees. Syst. Biol. 53(2):309-326.

Danforth, B.N., C. Lin, and J. Fang. 2005. How do insect nuclear ribosomal genes compare to protein-coding genes in phylogenetic utility and nucleotide substitution patterns. Syst. Biol. 30:549-562.

Davies, G.B.P., C. D. Eardley, and D. J. Brothers. 2005. Eight new species of *Scrapter* (Hymenoptera: Apoidea: Colletidae), with descriptions of *S. albifumus* and *S. amplispinatus* females and a major range extension of the genus. Afr. Invertebr. 46:141–179.

Davies, G.B.P., and D. J. Brothers. 2006. Morphology of Scrapter (Hymenoptera: Anthophila: Colletidae), with description of three new species and taxonomic status of five Cockerell taxa. Afr. Invertebr. 47:135–183.

Donovan, B.J. 2007. Apoidea (Insecta: Hymenoptera) in Fauna of New Zealand (Ko te Aitanga Pepeke o Aotearoa). ManaakinWhenua Press. Canterbury, NZ.

Dowton, M., and A.D. Austin. 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy – the behaviour of the incongruence length difference test in mixed-model analyses. Syst. Biol. 51(1):19-31.

Drummond, A.J., and A. Rambaut. 2003. BEAST. Version 1.3 [computer program]. Available: http://evolve.zoo.ox.ac.uk/. Accessed 22 April 2009.

Eardley, C.D. 1996. The genus *Scrapter* Lepeletier & Serville (Hymenoptera: Colletidae). Afr. Entomol. 4:37-92.

Exley, E.M. 1968. Revision of the genus *Brachyhesma* Michener. (Apoidea : Colletidae). Aust. J. Zool. 16:167-201.

Exley, E.M. 1972. Revision of the genus *Pachyprosopis* Perkins (Apoidea : Colletidae). Aust. J. Zool. Suppl. Ser. No. 10:1-43.

Exley, E.M. 2001. The *walkeriana* species-group of *Euhesma* Michener (Hymenoptera: Colletidae: Euryglossinae). Aus. J. Entomol. 40(2):102-112.

Feng, D., and R.F. Doolittle. 1987. Progressive sequence alignment as a prerequisitet to correct phylogenetic trees. J. Mol. Evol. 25(4):351-360.

Francis, J.E., and I. Poole. 2002. Cretaceous and early Tertiary climates of Antarctica: evidence from fossil wood. Palaeogeogr. Palaeocl. 182: 47-64.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3:294–299.

Fuller, S., M. Schwarz, and S. Tierney. 2005. Phylogenetics of the allodapine bee genus *Braunsapis*: historical biogeography and long-range dispersal over water. J. Biogeogr. 32:2135–2144.

Gheerbrant, E., and J. Rage. 2006. Paleobiogeography of Africa: How distinct from Gondwana and Laurasia? Palaeogeogr. Palaeocl. 241:224–246.

Hill, R.S. (ed.) 1994. The history of selected Australian taxa in *History of Australian vegetation: Cretaceous to Recent*. Cambridge Univ. Press, U.K.

Houle, A. 1999. The origin of platyrrhines: an evaluation of the antarctic scenario and the floating island model. Am. J. Phys. Anthropol. 109:541–559.

Houston, T.F. 1969. Observations on the nests and behaviour of some euryglossine bees (Hymenoptera : Colletidae). J. Aust. Entomol. Soc. 8:1-10.

Houston, T.F. 1975. A Revision of the Australian Hylaeine Bees (Hymenoptera: Colletidae). Aust. J. Zool. Supp. Ser. No. 36:1-135.
Houston, T.F. 1981. A revision of the Hylaeinae Bees (Hymenoptera: Colletidae). II. Aust. J. Zool. Supp. Ser. No. 80:1-128.

Hrbek, T., and A. Meyer. 2003. Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). J. Evol. Biol. 16:17-36.

Huelsenbeck, J.P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics. 17:754-755.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-120.

Kuhlmann, M. 2005. Diversity, distribution patterns and endemism of southern African bees, pp 167-172. In African Biodiversity: molecules, organisms, ecosystems, eds, B.A., Huber, B.J. Sinclair, K.H. Lamp.

Kuhlmann, M., A.E.B. Almeida, N. Laurenne, and D.L.J. Quicke. 2009. Molecular phylogeny and historical biogeography of the bee genus *Colletes* Latreille, 1802 (Hymenoptera: Apiformes: Colletidae), based on mitochondrial COI and nuclear 28S sequence data. Insect Syst. Evol. 40:291-318.

Lee, T, and L.A. Lawver. 1995. Cenozoic plate reconstruction of Southeast Asia. Tectonophysics. 251:85-138.

McGinley, R.J. 1980. Glossal morphology of the Colletidae and recognition of the Stenotritidae at the family level. J. Kansas Entomol. Soc. 53:539-552.

McLoughlin, S. 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. 49:271-300.

McPeek, M.A., and J.M. Brown. 2007. Clade age and not diversification rate explains species richness among animal taxa. Am. Nat. 169:E97-106.

Maddison, D. R. and W. P. Maddison, 2000. MacClade 4: Analysis of phylogeny and character evolution. Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Mardulyn, P., and J.B Whitfield. 1999. Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of microgastrinae (Hymenoptera: Braconidae): evidence of a high diversification rate in this group of parasitoids. Mol. Phylogenet. Evol. 12, 282–294.

Melo, G.A.R., and R.B. Goncalves. 2005. Higher-level bee classifications (Hymenoptera, Apoidea, Apidae *sensu lato*). Rev. Bras. Zool. 22(1):153–159.

Michener, C.D. 1965. A classification of the bees of the Australian and South Pacific regions. B. Am. Mus. Nat. Hist. 130:1-362.

Michener, C.D. 1979. The biogeography of the bees. Ann. Mo. Bot. Gard. 66(3):277-347.

Michener, C.D. 2007. The Bees of the World, 2nd edn. Johns Hopkins University Press, Baltimore, MD.

Mullen, S.P., W.K. Savage, N. Wahlberg, and K.R. Willmott. 2011. Rapid diversification and not clade age explains high diversity in neotropical Adelpha butterflies. Proc. R. Soc. B. 278:1777-1785.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53:673-684.

Paradis, E., J. Claude, and K. Strimmer. 2004 APE: Analyses of Phylogenetics and Evolution
in R language. Bioinformatics. 20(2): 289–290.

Pybus, O.G., and P.H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. P. Roy. Soc. Lond. B. Bio. 267:2267-2272.

Rabosky, D.L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. Evolution. 60(6):1152–1164.

Rabosky D.L. 2008. LASER: Likelihood Analysis of Speciation/Extinction Rates from Phylogenies. R package. version 2.2.

Rabosky, D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution. 64(6): 1816–1824.

Rabosky, D.L., and I.J. Lovette. 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution. 62(8):1866-1875.

Rambaut, A., and A.J. Drummond. 2007. [computer program] Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Raven, P.H., and D.I. Axelrod. 1974. Angiosperm biogeography and past continental movements. Ann. Mo. Not. Gard.61(3):539-673.

Regier, J.C., and J.W. Shultz. 2001. Elongation Factor-2: A Useful Gene for Arthropod Phylogenetics. Mol. Phylogenet. Evol. 20(1):136–148.

Ricklefs, R.E. 2006. Evolutionary diversification and the origin of the diversity–environment relationship. Ecology. Suppl. 87(7): S3–S13.

Ricklefs, R.E. 2007. Estimating diversification rates from phylogenetic information. Trends Ecol. Evol.. 22(11):601-609.

Rozen J.G. Jr., and C.D. Michener. 1968. The biology of *Scrapter* and its cuckoo bee, *Pseudodichroa*. (Hymenoptera: Colletidae and Anthophoridae). Am. Mus. Novit. 2335:1–13.

Sanderson, M.J. 2003. r8s: inferring absolute rates of evolution and divergence times in the absence of a molecular clock. Bioinformatics. 19:301–301.

Sanmartín, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53(2):216-243.

Sanmartín, I., C.L. Anderson, A. Alarcon, F. Ronquist, and J.J. Aldasoro. 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. Biol. Lett.

doi: 10.1098/rsbl.2010.0095

Schwarz, M.P., S.M. Tierney, S.J.B. Cooper, and N.J. Bull. 2004. Molecular phylogenetics of the allodapine bee genus Braunsapis: A–T bias and heterogeneous substitution parameters. Mol. Phylogenet. Evol. 32:110–122.

Schwarz, M.P., S. Fuller, S.M. Tierney, and S.J.B. Cooper. 2006. Molecular Phylogenetics of the Exoneurine Allodapine Bees Reveal an Ancient and Puzzling Dispersal from Africa to Australia. Syst. Biol. 55(1):31–45.

Schweizer, M., M. Güntert, O. Seehausen, and S.T. Hertwig. 2010. The evolutionary diversification of parrots supports a taxon pulse model with multiple trans-oceanic dispersal events and local radiations. Mol. Phylogenet. Evol. 54(3):984-994.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87:651–701.

Thorn, V.C., and R. DeConto. 2006. Antarctic climate at the Eocene/Oligocene boundary-climate model sensitivity to high latitude vegetation type and comparisons with the palaeobotanical record. Palaeogeogr. Palaeocl. 231:134–157.

Truswell, E.M., and K. Macphail 2009. Polar forests on the edge of extinction: what does the fossil spore and pollen evidence from East Antarctica say? Aust. Syst. Bot. 22: 57–106.

Wallis, G.P., and S.A. Trewick. 2009. New Zealand phylogeography: evolution on a small continent. Mol. Ecol. 8:3548–3580.

Wcislo, W.T. and J.H. Cane. 1996. Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies. Annu. Rev. Entomol. 41:257-286.

Table 1. Collection localities, and GenBank accession numbers for 28S, EF-1α and two regions of COI. All Euryglossinae species are from around Australia, and all Scrapterinae species are from the Western Cape province in South Africa. Abbreviated collection localities refer to states in Australia: NT=Northern Territory; Qld=Queensland; SA=South Australia; Vic=Victoria, WA=Western Australia. 'Grp' refers to whether Euryglossinae are part of Group A or Group B, as distinguished by wing venation and our BI phylogeny. (A) refers to Michener's (2007) grouping, which was based on additional characters. GenBank accession numbers that begin with JN are from this study, those beginning with AY or DQ are from two other studies. 'Lagrange' refers to the coded distributions used for inferring historical biogeography, where 'A' = Australia, 'B' = Africa, and 'C' = South America.

| Species | Collection Location | Grp | Lagrange | 28S | EF1 | COI | |
|---|---|---|---|---|---|---|---|
| **Euryglossinae, Australia** | | | | | | | |
| *Brachyhesma* sp. | Adelaide, SA | B(A) | A | JN603373 | JN603438 | - | JN603396 |
| *Callohesma* sp. | NT | B | A | JN603376 | JN603441 | JN603421 | JN603399 |
| *Callohesma* sp. 1 | Adelaide, SA | B | A | JN603377 | JN603442 | JN603422 | JN603400 |
| *Callohesma* sp. 2 | Ngaraket, SA | B | A | JN603378 | JN603443 | JN603423 | JN603401 |
| *Callohesma calliopsella* | GenBank | B | A | DQ872768 | AY585126 | DQ872696 | - |
| *Callohesma flavopicta* | Vic | B | A | JN603374 | JN603439 | JN603419 | JN603397 |
| *Callohesma lacteipennis* | Scott Creek, SA | B | A | JN603375 | JN603440 | JN603420 | JN603398 |
| *Callohesma sinapipes* | Whyalla, SA | B | A | JN566238 | JN566284 | JN57886 | JN578919 |
| *Euhesma* sp. 1 | Cooee Bay, Qld | B | A | JN603379 | JN603444 | JN603424 | JN603402 |
| *Euhesma* sp. 2 | Cooee Bay, Qld | B | A | JN603380 | JN603445 | JN603425 | JN603403 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Euhesma* sp. 3 | 94 km south of Williams, WA | B | A | JN603381 | JN603446 | JN603426 | JN603404 |
| *Euhesma crabronica* | GenBank | B | A | DQ768543 | DQ884654 | - | - |
| *Euhesma platyrhina* | GenBank | A | A | DQ768541 | DQ884652 | DQ872694 | - |
| *Euryglossina* sp. 1 | Broome, WA | A | A | JN603382 | JN603447 | - | JN603405 |
| *Euryglossina* sp. 2 | Qld | A | A | JN603383 | JN603448 | JN603427 | JN603406 |
| *Euryglossina* sp. 3 | Whyalla, SA | A | A | JN603384 | JN603449 | JN603428 | JN603407 |
| *Euryglossina globuliceps* | GenBank | A | A | AY654490 | AY585127 | DQ115551 | - |
| *Euryglossula* sp. 1 | Larrimah, NT | A | A | JN603385 | JN603450 | JN603429 | JN603408 |
| *Euryglossula* sp. 2 | Jurien Bay, WA | A | A | JN603386 | JN603451 | JN603430 | JN603409 |
| *Euryglossula* sp. 3 | Charters Towers, Qld | A | A | JN603387 | JN603452 | JN603431 | JN603410 |
| *Hyphesma* sp. 1 | Kojonup, WA | B | A | JN603388 | JN603453 | JN603432 | JN603411 |
| *Hyphesma* sp. 2 | Scott Creek, SA | B | A | JN603389 | JN603454 | JN603433 | JN603412 |
| *Pachyprosopis* (*Pachyprosopula*) sp. | Warwick, Qld | A | A | JN603389 | JN603455 | JN603434 | JN603413 |
| *Xanthesma furcifera* | GenBank | B | A | DQ872770 | AY585140 | - | - |
| **Scrapterinae, South Africa** | | | | | | | |
| *Scrapter* sp. 1 | Constantiaberg, Cape Town | | B | JN603394 | JN603458 | - | JN603416 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Scrapter algoensis* | GenBank | B | DQ872771 | EF032901 | - | - |
| *Scrapter bicolor* | Pakhuis Pass, Clanwilliam | B | JN566241 | JN566287 | JN578888 | JN578922 |
| *Scrapter chloris* | Pakhuis Pass, Clanwilliam | B | JN603391 | JN603456 | JN603435 | JN603414 |
| *Scrapter erubescens* | GenBank | B | DQ872772 | AY585135 | - | - |
| *Scrapter heterodoxus* | GenBank | B | DQ872773 | AY585136 | - | - |
| *Scrapter niger* | GenBank | B | DQ872774 | AY585137 | - | - |
| *Scrapter nitidus* | Pakhuis Pass, Clanwilliam | B | JN603392 | JN603457 | JN603436 | JN603415 |
| *Scrapter ruficornis* | GenBank | B | DQ872775 | AY585138 | - | - |
| *Scrapter striatus* | Pakhuis Pass, Clanwilliam | B | JN603393 | JN603459 | JN603437 | JN603417 |
| **Outgroups** | | | | | | |
| *Callomelitta antipodes* (Neopasiphaeinae) | Cape Bridgewater, Vic | A | DQ872767 | AY585122 | JN603418 | JN603395 |
| *Chilicola sp. 13* (Xeromelissinae) | West Conquimbo, Chile | C | JN566240 | JN566286 | - | JN578921 |
| *Deranchylaeus sp.* | Zambia | B | JN566246 | JN566292 | JN578890 | JN578927 |

| | | | | | | |
|---|---|---|---|---|---|---|
| (Hylaeinae) | | | | | | |
| *Hylaeus* near *lateralis* spB (Hylaeinae) | Kojunup, WA | A | JN566277 | JN566328 | JN578911 | JN578961 |
| *Hylaeus perhumilis* (Hylaeinae) | Port Lincoln, SA | A | JN566259 | JN566310 | JN578901 | JN578944 |
| *Meroglossa* near *eucalypti* (Hylaeinae) | Australia | A | JN566281 | JN566333 | JN578915 | JN578965 |
| *Trichocolletes* sp. (Neopasiphaeinae) | GenBank | A | DQ872760 | AY585139 | DQ872677 | - |
| Xeromelissinae sp2 (Xeromelissinae) | West Conquimbo, Chile | C | JN566242 | JN566288 | - | JN578923 |

Table 2a: Primer specifications.

| Locus | Primer name | Sequence 5'-3' | Reference |
|---|---|---|---|
| EF-1α | HaF2For1 | GGG YAA AGG WTC CTT CAA RTA TGC | Danforth et al. 1999 |
| | F2-rev1 | A ATC AGC AGC ACC TTT AGG TGG | Danforth et al. 1999 |
| 28S rRNA | A-28S-For | CCC CCT GAA TTT AAG CAT AT | Ward and Brady 2003 |
| | Mar28S-Rev | TAG TTC ACC ATC TTT CGG GTC CC | Mardulyn and Whitfield 1999 |
| COI | Jerry (C1-J-2183) | CAA CAT TTA TTT TGA TTT TTT GG | Simon et al. 1994 |
| | Pat (TL2-N-3014 | TCC AAT GCA GTA ATC TGC CAT ATT A | Simon et al. 1994 |
| | LCO1490 | CCT TTT ATA ATT GGA GGA TTT GG | Folmer et al. 1994 |
| | M399 | TCA TCT AAA AAC TTT AAT TCC TG | Schwarz et al. 2004 |

Table 2b: PCR conditions for primers used showing temperature and time of: initial denaturation; and 35 cycles of denaturation, annealing, extension.

| Primer | PCR Conditions |
|---|---|
| HaF2For1/F2-rev1 | 94˚C 3min; 94˚C 60sec, 52˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| A-28S-For/Mar28S-Rev | 94˚C 3min; 94˚C 60sec, 58˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| Jerry/Pat/ LC01490/M399 | 94˚C 3min; 92˚C 30sec, 52˚C 90sec, 72˚C 2.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |

Table 3: Results of three Penalized Likelihood chronograms calibrated at three different dates on two different nodes (Euryglossinae + Scrapterinae, and the Root node) followed by the resulting crown age estimates of six major nodes. See Figure 3 for descriptions of the six nodes. Crown ages are given from the results of Almeida et al. (in press) for comparison.

| **Calibration point**<br>*** based on this study, derived from genetic distance estimates**<br>**** based on Almeida et al. (in press)** | **Root** | **Scrapterinae / Euryglossinae** | **Scrapterinae** | **Euryglossinae** | **Group A** | **Group B** |
|---|---|---|---|---|---|---|
| **Euryglossinae + Scrapterinae 49 Mya*** (see Figure 3a) | 54 Mya | - | 35 Mya | 42 Mya | 19 Mya | 34 Mya |
| **Euryglossinae + Scrapterinae 54 Mya**** (see Figure 3b) | 65 Mya | - | 35 Mya | 47 Mya | 19 Mya | 41 Mya |
| **Root 60 mya**** (see Figure 3c) | - | 56 Mya | 38 Mya | 50 Mya | 20 Mya | 46 Mya |
| **Almeida et al. (in press) crown age estimate** | 60 (54-65) Mya | 54 (42-66) Mya | 24 (17-31) Mya | 48 (38-60) Mya | n/a | n/a |

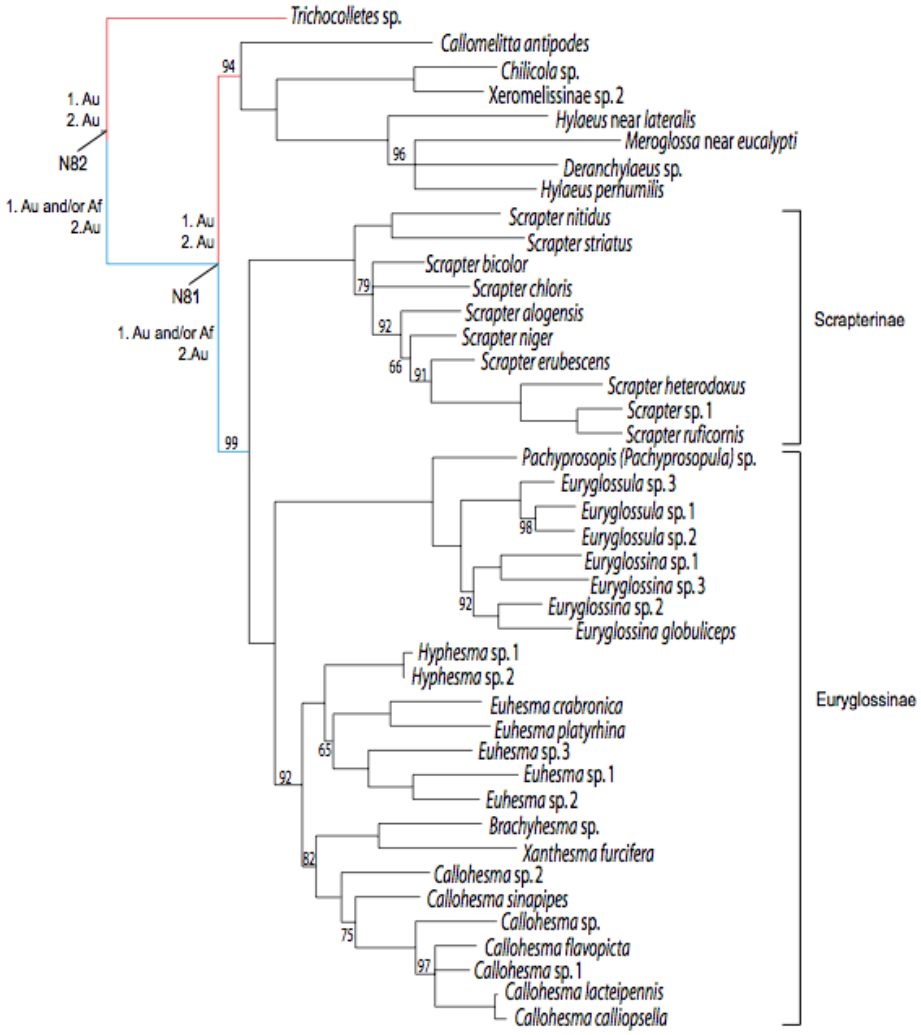



Figure 1. Bayesian consensus phylogram of Euryglossinae and outgroups. Posterior probabilities are indicated for each node under 100. Square brackets indicate Scrapterinae and Euryglossinae, other species are outgroup species. N82 and N81 refer to nodes for which ancestral ranges were inferred for the upper (red) and lower (blue) branches leading from each node under Maximum Likelihood in Lagrange. The two lower (blue) branches are most significant when considering the ancestral range of Scrapterinae/Euryglossinae. The two most likely ancestral ranges for each branch are listed as 1. and 2. – the third possibilities were not statistically significant. Au=Australia, Af=Africa. For the branches leading from N82, the probability of possibility 1. = 61.38% with log-likelihhod (lnLh) -21.56. The probability of possibility 2. = 28.75%, lnL -22.32. For the branches leading from N81, the probability of possibility 1. = 71.86%, lnL – 21.41. The probability of possibility 2. is 18.31% , lnL – 22.77.

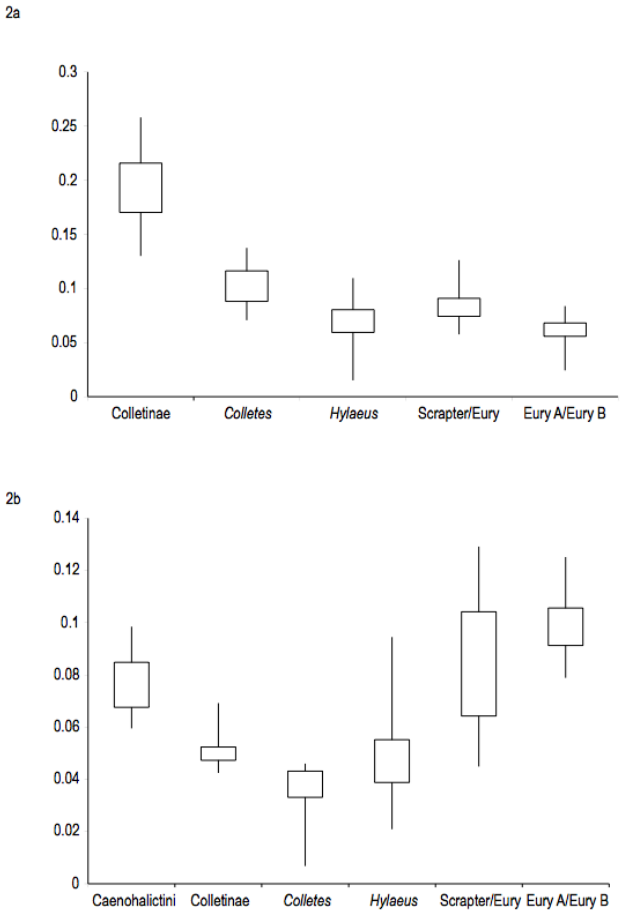



Figure 2. Pairwise genetic estimates from the GTR parameter from two nodes in the phylogeny compared with other groups. Scrapter/Eury is the node separating Scrapterinae and Euryglossinae, and Eury A/Eury B is the node separating Euryglossinae Group A from Euryglossinae Group B. (a) COI pairwise genetic distance compared with the divergence of Colletinae (i.e., *Mourocolletes* from *Colletes*), and a clade of *Colletes* within Colletinae (Kuhlmann et al. 2009), as well as Australian *Hylaeus* species except for the poorly supported clade including *H. (Xenohylaeus) leptospermi, H. (Planihylaeus) trilobatus, H. (Prosopisteron) microphenax* (Chapter 1). (b) EF-1α pairwise genetic distance estimates compared with the divergence of the same three groups as for COI, as well as for the subtribe Caenohalictini (Halictini; Halictidae): *Agapostemon*/ *Rhinetula*/ *Dinagapostemon* from *Habralictus*/ *Caenohalictus*/ *Ruizantheda*/ *Pseudagapostemon* (Danforth et al. 2004).

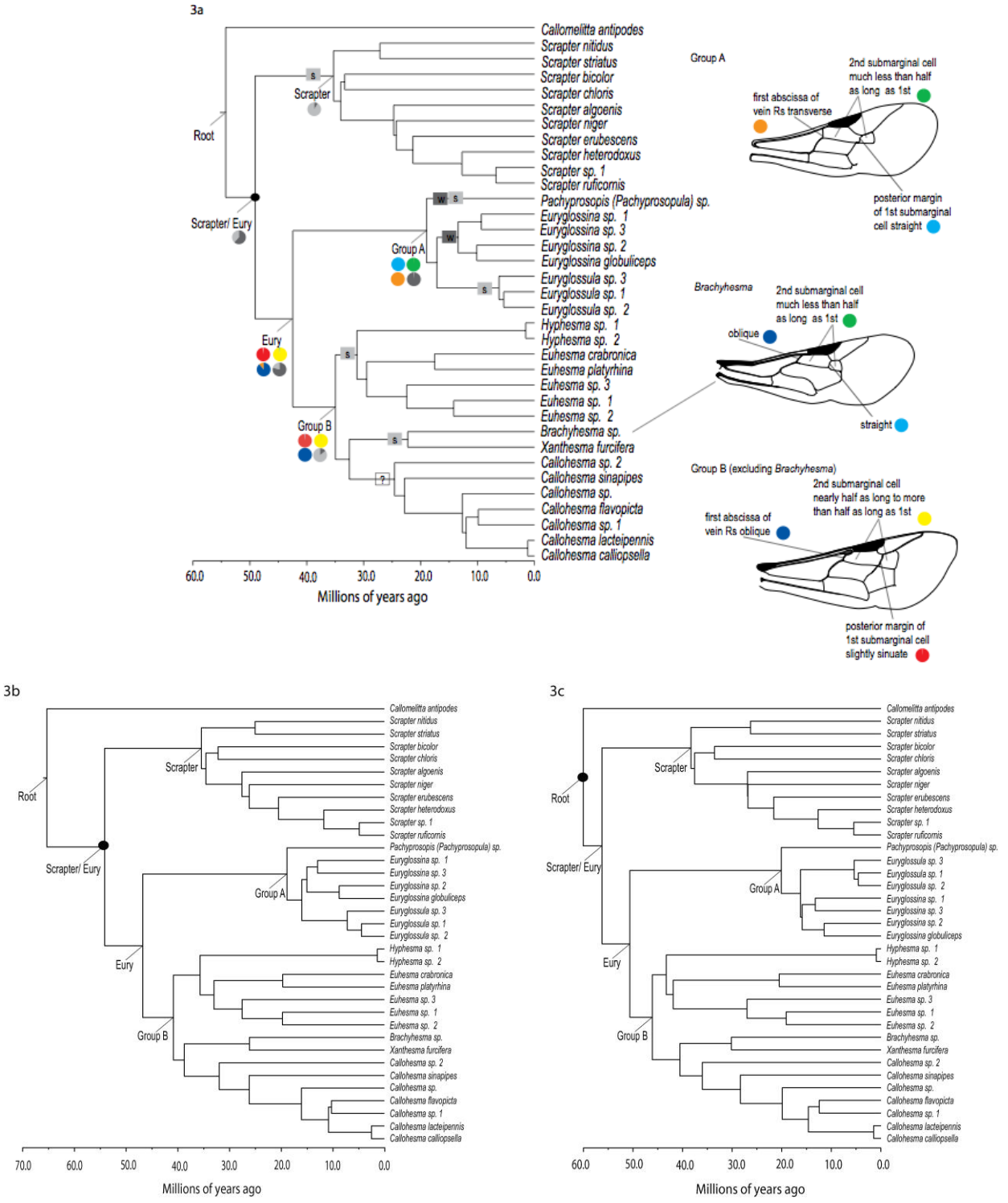



Figure 3. Consensus chronograms derived from penalized likelihood transformation of the consensus phylogram using three different age constraints at two different nodes, and indicated by black circles. (a) Scrapterinae + Euryglsossinae node fixed at 49 Mya based on genetic distance results. Euryglossinae Groups A and B are defined in congruence

with three main characteristics in wing venation. Wing diagrams refer to Group A, *Brachyhesma* (Group B), and all Group B excluding *Brachyhesma*. Pie charts at nodes present proportion of likelihood of the ancestral state for three wing characteristics and nesting substrate – colours of each trait are defined in the wing diagrams. Note that *Brachyhesma* is similar to Group A in length of the second submarginal cell in relation to the first (green/yellow), and in the straight posterior margin of the 1st submarginal cell (light blue/red), whereas it is similar to the other Group B species in the oblique angle of the first abscissa in vein Rs (dark blue/orange). W (dark grey) = lineages for which only wood nesting species are known, S (light grey) = lineages for which only soil nesting species are known, ? = no known records of nesting habit. (b) Scrapterinae + Euryglossinae node fixed at 54 Mya (Almeida et al. in press). (c) Root node fixed at 60 Mya (Almeida et al. in press). The six nodes for which age stimates are defined in Table 4 are indicated as: Root (including outgroup taxa), Scrapter/Eury (Scrapterinae/ Euryglossinae), Scrapter (Scrpaterinae), Group A (Euryglossinae), Group B (Euryglossinae).

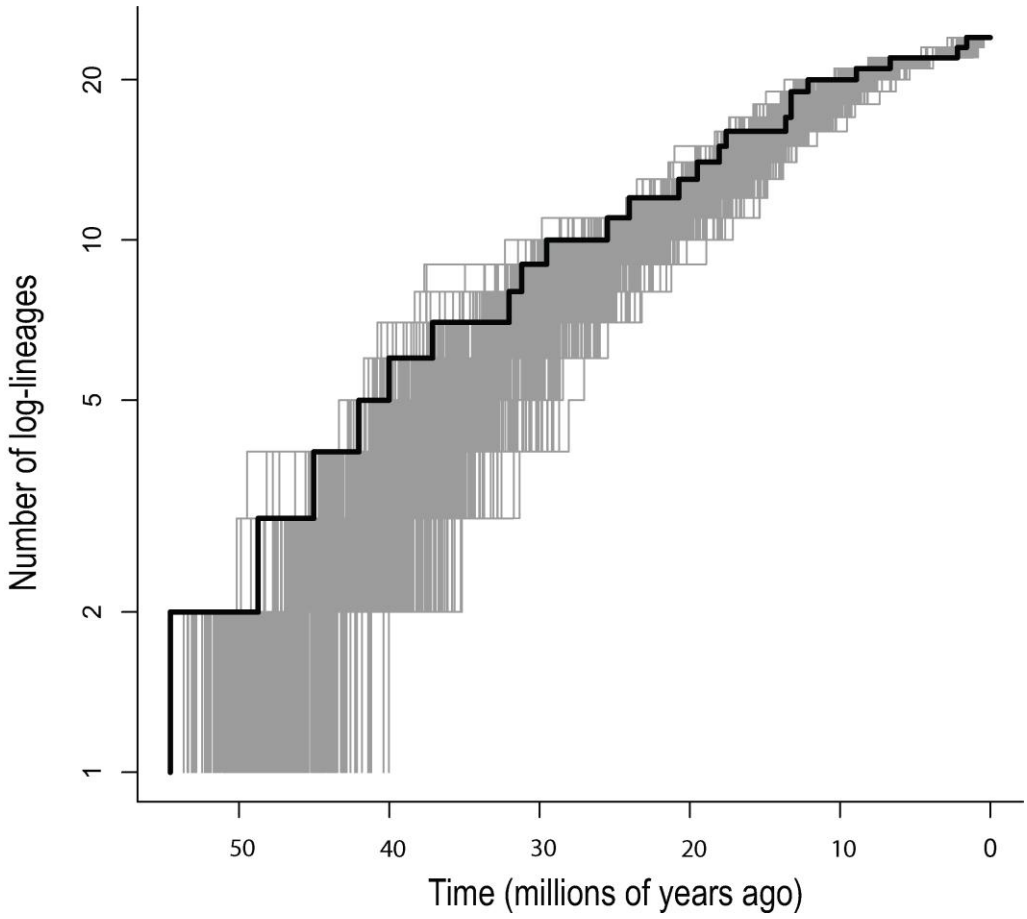



Figure 4. Log-lineages through time plot of Euryglossinae cladogenesis over time. Grey lines represent 500 randomly selected post burnin chronograms and the black line represents the log-LTT plot from the consensus chronogram.

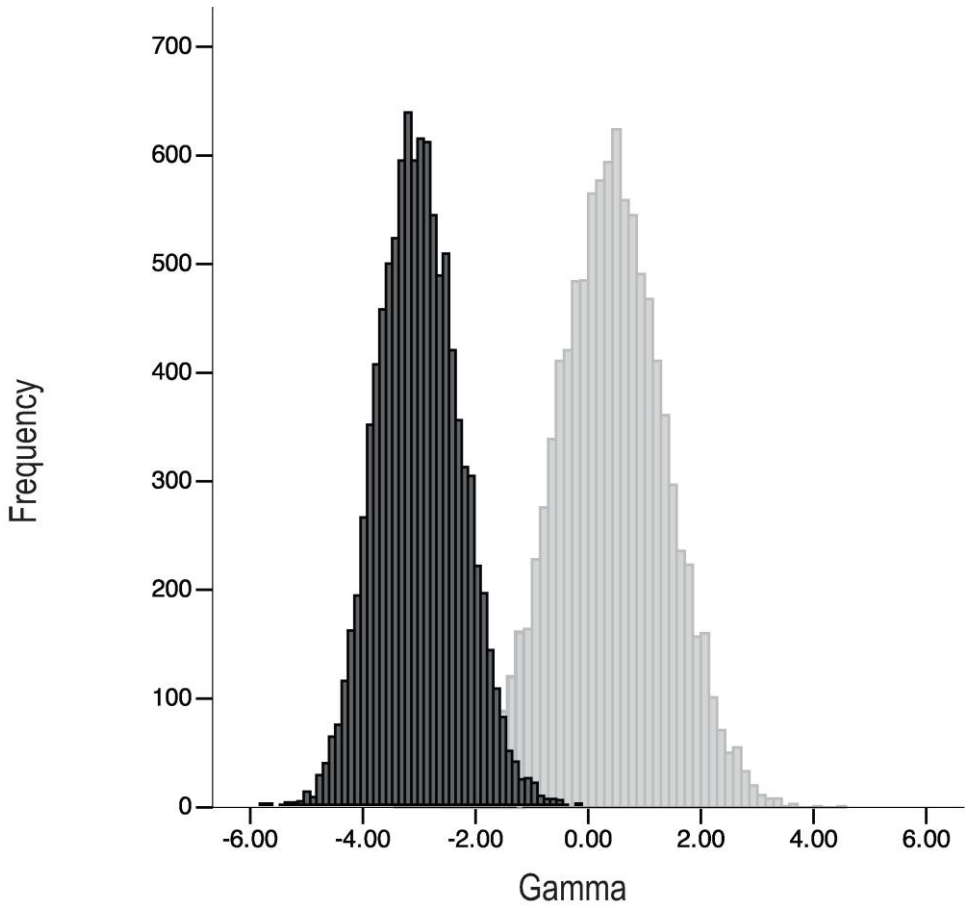



Figure 5. γ (gamma) distributions of 500 randomly-selected phylogenies including the 24 sampled Euryglossinae species (dark grey), and simulated (light grey) phylogenies based on an estimated 400 species.

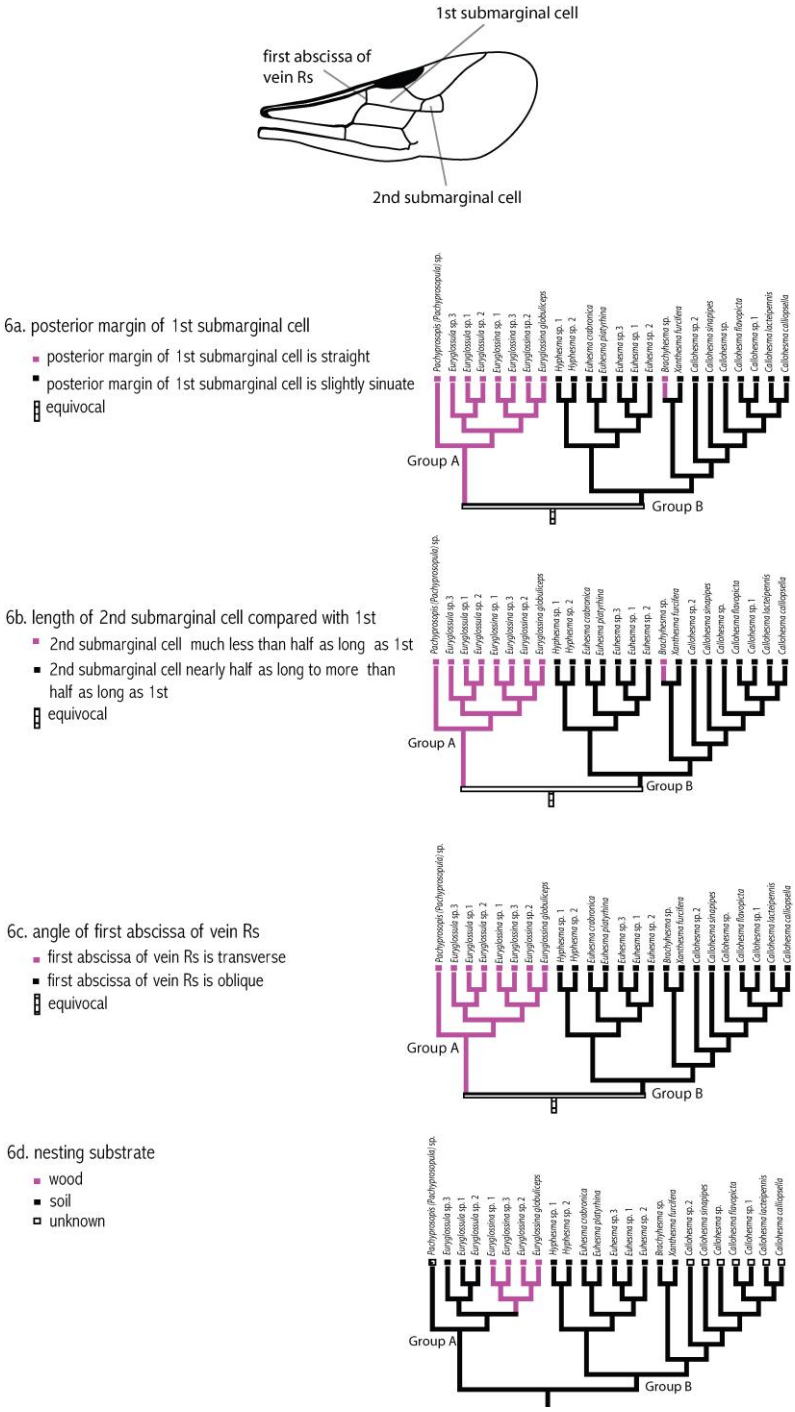



Figure 6: Character tracing in MacClade of three wing characteristics: (a) whether the posterior margin of the first submargical cell is straight or sinuate, (b) length of the 2nd submarginal cell compared with the first submarginal cell, (c) whether the angle of the first abscissa of vein Rs is transverse or oblique, and of (d) nesting substrate: soil, wood, or unknown in the case of *Callohesma*.

**CHAPTER III.**

# Trends in euryglossine and hylaeine (Hymenoptera: Colletidae) body size evolution.

Pelin Kayaalp[1,2] and Michael P. Schwarz[1]

1. School of Biological Sciences, Flinders University, GPO Box 2100, Adelaide, S.A, 5001.
2. Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa.

RUNNING TITLE: Australian bee body size evolution

KEYWORDS: body size, bee phylogeny, scopae, pollen

## *Abstract*

Morphological evolution, particularly body size, is thought to occur rapidly when organisms enter new ecological zones. Body size in animals can strongly influence energetic requirements, potential resource exploitation, and susceptibility to predation. An increase in body size over evolutionary time is often predicted for vertebrates, but the opposite is frequently inferred for invertebrates. Some of the smallest species of bees in Australia, Euryglossinae and Hylaeinae (Colletidae) are also the most abundant, and it has been proposed that their small size has led to the loss of pollen-collecting scopae. We applied a generalised least-squares technique for inferring trends in body size evolution, and to estimate ancestral size for both the hylaeines and euryglossines as well as for two internal nodes within each group. We found that phylogenetic signal is evident to only a small degree in the body size evolution of these two groups. The variation in the tempo of trait evolution as a function of branch length, and whether trait evolution is linked to speciation differed in the hylaeines depending on whether a Maximum Likelihood approach or an MCMC approach was used. Both models indicated that body size in the euryglossines has changed more in longer brances. There is some indication that evolution of body size might be adaptive in both groups, but the results are not strongly supported. Estimation of ancestral sizes suggest that within both the euryglossines and hylaeines there is one clade that has retained the ancestral size of the root node. In the euryglossines, a second clade has become smaller, while in the hylaeines a second clade has become larger. We suggest that more so than small body size, the abundance of nectar-rich Myrtaceae species in Australia might be linked to the lack of scopae and internal pollen transport in hylaeines and euryglossines, as well as to their diversity in Australia compared to the rest of the World.

## *Introduction*

Body size in animals can strongly influence energetic requirements, potential resource exploitation, and susceptibility to predation (Tokeshi and Schmid 2002). As such, both extremes of body size entail adaptations associated with morphology, physiology, locomotion, and reproduction (Hanken and Wake 1993; Mayhew 2007; Chown and Gaston 2010), and body size in animals tends to show a strong signal associated with phylogeny and/or taxonomic grouping above the species level (e.g. among genera, families and orders) (Gaston and Blackburn 2000). Cope's Rule predicts evolution toward a larger body size (as far as morphological and other constraints will allow, Hone and Benton 2005); possibly as an artefact of the small-bodied ancestors of most lineages (Stanley 1973), but has not been well tested (Hone and Benton 2005). In many insects, however, a decrease from the ancestral body size is inferred (Peters 1983; Dudley 2002; Chown and Gaston 2010) (outside of island systems where the opposite is frequently inferred, arguably due to fewer predation pressures), possibly enabled as a result of the adaptation of asynchronous flight muscles and high wingbeat frequencies (Dudley 2002).

Of the approximately 17,500 valid bee species (Michener 2007: pp. 1, 75), the largest is *Megachile pluto* (Megachilidae), which can be as large as 39.0 mm (Michener 2007), and some of the smallest species are euryglossine (Colletidae) and meliponine (Apidae) bees at around 2.0 mm long (Exley 1974; Michener 2007). This implies that within the bees there exists a wide range of ecological, behavioural, physiological and morphological differences, such as habitat preference, and flight speed (Dudley 2002). In bees, larger males have been found to be more successful in establishing mating territories (e.g., Eickwort and Ginsberg 1980; Leys 2000), and larger females tend to have higher fecundity and larger offspring (e.g., Giovanetti and Lasso 2005). Many of the studies on variation in bee body size have investigated the effect of body size on reproductive dominance in social species (e.g. Michener 1974; Breed et al. 1978; Wheeler 1986; Tierney et al. 2002). For solitary bees, on the other hand, there is no within-colony competition and thus the advantages of small size, such as fewer resource requirements, and physiological benefits might enable a decrease in body size that would be detrimental to some social species.

Body size can have a great influence on the amount of resources needed and how they are acquired. For most bees this specifically involves flower visitation and the ability to transport pollen and nectar loads, and in some cases oil, as primary larval provisions (Michener 2007). Exceptions include the use of extra-floral nectaries as a nectar source (e.g. O'Brien 1995), and the collection of fungal spores by Honeybees (Shaw and Robertson 1980; Shaw 1998). One subgroup of stingless bees, including *Trigona necrophaga*, *Trigona crassipes*, *Trigona hypogea*, are specialised to feed on carrion (Roubik 1982; Camargo and Roubik 1991), but otherwise non-parasitic bees rely largely on floral provisions.

Studies that look at body size correlates in bees have found both negative and positive associations between size and pollen load (Ramalho et al. 1994, 1998), foraging distance (Greenleaf et al. 2007), and foraging time (Wetterer 1989), whereas size is not associated with oligolecty or polylecty in bees (Wcislo and Cane 1996). Transfer of collected pollen back to the nest is generally a function carried out by specialised hairs on the hind legs and/or metasomal sternites of bees, the scopae, and most bees carry some pollen mixed with nectar in the crop (Thorp 1979; 2000). However, in Euryglossinae and Hylaeinae (Colletidae) the scopae are completely absent or so reduced that alimentary transport of pollen and nectar in the crop is always utilised (Michener 1964, 1965, 2007). It has been suggested that the lack of scopae in these subfamilies is a result of small body size: '...rappelons que ce sont des insectes de petite taille, chez lesquels une scopa formée d'une couverture de soies ne serait probablement pas efficace' ('…nudity in euryglossines and hylaeines is secondarily derived because scopa would be ineffectual in tiny bees', Pasteels and Pasteels 1976 in Davies and Brothers 2006, p. 160). The topic of this study is to detect any trends in the evolution of body size in the euryglossines and hylaeines, and discuss the role of reduced scopae and internal pollen transport.

Depending on the classification system used, Colletidae are divided into five to seven subfamilies (Michener 2007; Almeida and Danforth 2009) and are found on all continents (excluding Antarctica), but are substantially more abundant and diverse in Australia and South America compared with the rest of the world (Michener 2007).

Colletidae are almost completely solitary (but see Houston 1969 and Spessa et al. 2000), short-tongued, with a synapomorphic blunt, truncate, or bifid glossa in all females and most males (Michener 1965). The shape of the female glossa is associated with the application of cellophane-like lining to the walls of the nest (Michener 1965). The short glossa of Colletidae restricts their access to some nectar sources and so a smaller size would enable access to a larger variety of flowers, i.e, a minute body size might allow them to enter into the corolla tube to collect pollen and/or nectar (Wcislo 1989). Euryglossines and hylaeines in Australia, for example, are known to largely visit flowers of shallow-cupped Myrtaceae species (Michener 1965; Exley 1967, 1968, 1972, 1974, 1975, 1976, 1978, 2001; Houston 1975, 1981), which are abundant throughout Australia. Easy access to nectar and pollen might have contributed to the diversity and abundance of these short-tongued groups in Australia.

Alhough the assumed phylognetic relationship between Hylaeinae and Euryglossinae was based on their abundance and diversity in Australia (only two genera of Euryglossinae and one genus of Hylaeinae are found outside of Australia), and lack of scopae (Michener 1979), molecular data do not support a direct sister relationship between the two groups (Almeida and Danforth 2009; Chapter 2). This prompts the question as to why the lack of scopae has evolved twice in two groups that are most dominant in Australia. Although they are amongst the smallest bees found in Australia (particularly the Euryglossinae where some species from the genera *Euryglossina*, *Euryglosella*, *Quasihesma* are less than 3mm long, Exley 1974), there are bees from other groups around the world that are the same size or smaller than some Euryglossinae and Hylaeinae (Michener 2007).

In summary, Euryglossinae and Hylaeinae share three characteristics: (i) they have relatively small body size, both within the Colletidae and across all bee groups; (ii) they are the only two non-parasitic bee groups that lack scopae and transport pollen entirely in the crop; and (iii) they are most diverse in Australia. Here we suggest that these three factors may be inter-related. Firstly, we explore whether there is phylogenetic signal in body size species of Australian euryglossines and hylaeines. We do this by applying a phylogenetic Bayesian inference procedure to inferring evolutionary patterns in body size within the hylaeines and the euryglossines. We then discuss whether internal pollen

transport in Hylaeinae and Euryglossinae might be associated with the abundance of Myrtaceae in Australia, and whether this in turn might be associated with the reduction of scopae in these two groups, their small size, and relative abundance in Australia. From here on, reference to hylaeines or euryglossines will be restricted to Australian species of these groups.

## *Methods*

### PHYLOGENETIC RECONSTRUCTION

Extractions, PCR protocols, and methods of phyloegentic construction were carried out using the same methods as for previous work on hylaeines (Chapter 1) and euryglossines (Chapter 2). Voucher specimens sequenced for this study are located at Flinders University, South Australia. Details of the voucher specimens for sequence material obtained from Genbank can be found using the relevant accession number (Tables 3a and 3b). Three genes, cytochrome *c* oxidase I (COI), the F2 copy of elongation factor-1$\alpha$ (EF-1$\alpha$), and the ribosomal subunit 28S, as per Chapter 1 and Chapter 2, were used. A euryglossine phylogeny was reconstructed using a Bayesian Inference method carried out using the Unix version of MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). Because of known third codon saturation problems in hymenoptera and bees (e.g., Dowton and Austin 2002; Schwarz et al. 2004) we partitioned CO1 and EF-1$\alpha$ (F2) into two parts each of 1$^{st}$ and 2$^{nd}$ codons together and the 3$^{rd}$ separately, giving rise to five partitions, with 28s treated as a separate but single partition.

We used a general time reversible (GTR) model with gamma distributed ($\Gamma$) rates with a proportion of invariant sites (I). We used GTR (a 6-parameter rate transition matrix) because this is the most general substitutional model and can incorporate simpler models, such as the K2P and HKY models if these emerge from the MCMC parameter space. Two sets of four Monte Carlo Markov chains with Metropolis Coupling were run in parallel for 40 million generations, sampling every 1000$^{th}$ generation. Convergence between parallel runs was assessed using the average standard deviation of split frequencies, and log likelihood (LnL) values were plotted using Tracer version 1.41 (Rambaut and Drummond, 2007, http://beast.bio.ed.ac.uk/Tracer) to determine when stationarity was achieved. We used a Bayesian Inference (BI) tree from a previous study

of hylaeine species that was reconstructed in the same way as described above (Chapter 1). That study included species from outside of Australia, and we excluded those species for the purposes of this study. Primers and PCR conditions are summarised in Tables 1a and 1b.

**CHRONOGRAMS**

Our method for tracing character evolution requires a chronogram. Using the Bayesian consensus tree of each group we calculated relative divergence times using penalized likelihood (PL) rate smoothing implemented in r8s v1.71 (Sanderson 2003) with the smoothing parameter determined by a cross-validation method. The crown age of the hylaeines was set to 30 My (Chapter 1), and the crown age of euryglossines was set to 45 My (Chapter 2).

Because of poor support at some nodes (especially in the hylaeine phylogeny) we also analysed trends in body size evolution and ancestral body sizes using multiple chronograms. To do this, we filtered the post-burnin trees to match the topology of the consensus trees, and randomly chose 500 of these trees to convert into chronograms. Because *Continuous* does not handle polytomies, randomly selected trees with polytomies were excluded by eye. We pruned the polytomies in the hylaeine consensus tree, which led to the exclusion of six species. Body size measurements were available for two of those species, *Hylaeus (Prosopisteron)* sp. 3, and *Hylaeus (Euprosopis)* sp. 7, and inter-tegular data for the other four species, *Hylaeus (Prosopisteron) simplus*, *Hylaeus (Edriohylaeus) ofarrelli*, *Hylaeus (Prosopisteron) primulipictus*, and *Hylaeus (Euprosopoides)* sp. 1, was not available.

**BODY SIZE EVOLUTION AND ANCESTRAL STATE RECONSTRUCTION**

We measured the inter-tegular width (see Fig. 1) of 12 female euryglossine and 36 female hylaeine bee species in our trees using an eyepiece micrometer with a dissecting microscope and used this as a proxy for body size (Cane 1987). Where more than one specimen of a species was available, we measured the inter-tegular width of two or three individuals and took the mean. In order to explore phylogenetic patterns in the evolution of body size of euryglossines and hylaeines we used the *Continuous* (Pagel 1997, 1999) package in BayesTraits (www.evolution.rdg.ac.uk) (Pagel and Meade 2004), which

implements a generalised least-squares model for across-species analysis of comparative data (Pagel 1997, 1999).

Three parameters, λ (lambda), δ (delta), and κ (kappa) estimate key features in the evolution of traits within a phylogeny: λ is a measure of how well a phylogeny explains variation in a trait, δ is an indication of variation in the tempo of trait evolution as a function of branch length, and κ is an indication of whether trait evolution is linked to speciation *per se*. We estimated each parameter using the standard random walk model (Model A) in the Maximum Likelihood (ML) mode of *Continuous* using the consensus trees as well as the randomly selected trees. The Likelihood Ratio Test (LRT), 2(log-Likelihood [better fitting model] – (log-Likelihood [worse fitting model]), was used to determine whether each parameter was different from either zero or one (see Table 2 for interpretations of default and estimated values). We then used the significant λ, δ, and κ values (see Results below) to estimate ancestral sizes of three major nodes within hylaeines and euryglossines using the consensus trees as well as the sets of multiple chronograms.

Estimation of ancestral states for internal nodes in *Continuous* is only possible using the MCMC function. First, in order to determine whether the estimates from MCMC would be appropriate, we estimated the parameters and the ancestral node using the MCMC function in the same way as for ML. Given that the results were similar to those estimated using ML (see Results) we proceeded to estimate internal nodes. We used the Alpha values derived from the MCMC run along with the estimated parameter values in order to estimate internal nodes.

## *Results*

### PHYLOGENETIC RECONSTRUCTION AND CHRONOGRAMS

Tables 3a and 3b outline the collection locations and GenBank accession numbers of hylaeines and euryglossines from previous studies. A previous study has reconstructed the hylaeine phylogram (Chapter 1) used here, except that we pruned the species that were from outside Australia. That study and a colletid-wide molecular phylogeny (Almeida and Danforth 2009) placed two Hylaeinae species that have been described as

genera, based on morphology, within the genus *Hylaeus* – specifically *Paleorhiza* (Almeida and Danforth 2009), and *Meroglossa* (Almedia and Danforth 2009; Chapter 1). Further taxonomic revision is needed in order to determine whether these genera should be recognised as subgenera, or alternatively whether some *Hylaeus* subgenera should be given generic rank (Chapter 1). We have therefore opted to use a less defining terminology by referring to species in this study collectively as hylaeines.

Although the hylaeine consensus chronogram (Fig. 2) is presented with polytomies, those branches were pruned, as required for *Continuous*. Of the randomly selected trees with identical topologies, those with polytomies were excluded. Two distinct clades are recognised, Clade 1 and Clade 2 (see Fig. 2), and ancestral body size was estimated for each of those nodes as well as for the hylaeine root node.

The Euryglossinae phylogeny reconstructed here (Fig. 3) also resembles a phylogeny reconstructed in a previous study (Chapter 2). In particular, two distinct euryglossine clades are defined, Group A and Group B, and concur with differences in wing venation (Michener 2007; Chapter 2), as well as with nesting habit (Chapter 2). Ancestral body size was estimated for each of Group A and Group B, as well as for the root node.

**BODY SIZE MEASUREMENTS**

Figures 2 and 3 summarise the inter-tegular widths of the hylaeine and euryglossine female bee species used in our analyses. Of the euryglossine species measured by us, *Euryglossina* sp. 1 and *Euryglossina* sp. 2 had the shortest inter-tegular widths (0.5 mm), and *Callohesma sinapipes* had the longest (1.7 mm). The hylaeine species from our data set with the shortest inter-tegular width was *Hylaeus (Prosopisteron) perhumilis* (0.5 mm), and *Hylaeus (Euprosopellus) pergibbosus* and *Hylaeus (Hylaeteron) douglasi* (2.2 mm) both had the longest inter-tegular width.

**PHYLOGENETIC SIGNAL, MODE, AND TEMPO OF BODY SIZE EVOLUTION**

Tables 4a and 4b summarise the likelihood ratio tests (LRT) of estimated $\lambda$, $\delta$, and $\kappa$ values against 0.0 and 1.0 under Maximum Likelihood (ML) for the randomly selected chronograms and for the consensus chronograms of hylaeines (Table 4a) and

euryglossines (Table 4b). Estimates for the multiple chronograms gave very similar results to those of the consensus chronograms for both groups. Estimated λ for hylaeines was 0.892 for the multiple chronograms, and 0.890 for the consensus chronogram. Estimated λ for euryglossines was 0.834 for the multiple chronograms and 0.835 for the consensus chronogram. LRT > 9.0 indicate that in both cases these values were highly different from 0.0 and 1.0. Estimated λ values were therefore used in the estimation of δ and κ. Again the multiple chronograms and the consensus chronogram produced very similar results, with δ values not significantly different from 1.0 (LRT 0.779 – 1.783), and κ values very different from 0.0 (LRT 7.06 - 10.126) for hylaeines and euryglossines under both ML and MCMC.

**ANCESTRAL BODY SIZE**

Ancestral body size under Maximum Likelihhod (ML) was estimated using the estimated parameter values from both the consensus chronogram and the multiple randomly selected chronograms. The distribution of ancestral body size based on the randomly chosen post-burnin chronograms of hylaeines and euryglossines are summarised as histograms in Figures 4a and 4b, respectively. Both groups show a normal distribution of values (1.04 – 1.19 mm), centred at 1.16 mm for Euryglossinae, and (1.29 – 1.44) centred at 1.37 mm for the hylaeines. These central values correspond roughly with the ancestral state estimates inferred from the consensus chronograms, which were 1.16 mm for euryglossines and 1.39 mm for the hylaeines.

Ancestral body sizes of the internal hylaeine and euryglossine nodes were estimated under MCMC using the sets of multiple chronograms. In order to determine whether the internal node estimates would be comparable to the ML root estimates, we first estimated the ancestral inter-tegular width of the root node under MCMC. The parameter estimates and LRT are summarised in Tables 5a and 5b. Again, a λ significantly different from 0.0 for both groups implies that phylogenetic correction is needed when estimating δ and κ values. λ was very close to 1.0 for hylaeines, which is a strong indication that phylogeentic correction is required. δ was significantly less than 1.0 (LRT = 4.35), and κ was close to 0.0 (LRT = 1.79) for hylaeines. δ was close to 1.0 (LRT = 0.33), and κ significantly greater than 1.0 (LRT = 6.42) for euryglossines.

The range of MCMC ancestral values were much wider but the average of these were similar to the ML estimates. MCMC estimates for Euryglossinae centred at 1.26 mm (0.64 – 1.75), and for hylaeine 1.38 mm (-1.4 – 4.46 mm). The similarities in the estimated root nodes in both ML and MCMC modes increased our confidence in the MCMC estimates of the internal nodes. The estimate for hylaeine Clade 1 was similar to that of the root node, 1.44 mm, and larger for Clade 2, 1.66 mm. For Euryglossinae, the estimate for Group A was smaller than for the root node at 0.7 mm (-0.38 – 1.5 mm), and for Group B was roughly the same at 1.27 mm (1.04 – 1.55 mm). The mean value for each estimated node is shown on the chronograms (Figs. 2 and 3).

## *Discussion*

### TRENDS IN BODY SIZE EVOLUTION IN EURYGLOSSINES AND HYLAEINES

As predicted, inter-tegular width (and hence body size) was found to deviate significantly from a species-specific model of evolution (Gaston and Blackburn 2000), in both hylaeines and euryglossines as evidenced by a $\lambda$ significantly different from 0.0. This suggests that phylogenetic history has had an effect on the evolution of body size, and phylogenetic correction is required, and so estimated $\lambda$ values were used to examine $\kappa$ and $\delta$. However, a $\lambda$ value significantly different from 1.0 also implies that phylogenetic history is not the only determinant of trait variation (Freckleton et al. 2002; Moen 2006). Only in the case of the hylaeine MCMC results was there an indication of strong phylogenetic influence, with $\lambda$ very close to 1.0 (estimated $\lambda$ = 0.99, LRT = 0.07). The accuracy of our results may also be weakened by undersampling, although, in trees with more than 20 tips, $\lambda$ can detect the phylogenetic dependence of data with more than 90% accuracy and with nearly 100% accuracy for trees with more than 40 tips (Freckleton et al. 2002).

For all parameter estimates in *Continuous* a value of 1.0 indicates default gradualism. Although any positive value calculated from the likelihood ratio test (LRT) favours the parameter with a larger log-likelihood (Lh), LRT values greater then 2.0 are considered to be significant. Estimated $\delta$ values were less than 1.0 but not significantly so (LRT <

2.0) in all cases except the MCMC hylaeine results where there was good support for $\delta < 1.0$ (LRT = 4.35). $\delta$ less than 1.0 suggests that evolution in body size might be adaptive in the hylaeines, however little support from the ML analysis makes this result questionable. Under Maximum Likelihood (ML) the $\kappa$ results for the hylaeines suggest that there was more change in longer branches, but the result was not significantly different from 1.0 (LRT = 1.5). Under the MCMC analysis hylaeine $\kappa$ was not different from 0.0 (LRT = 1.79) indicating that body size might follow a punctuational model of evolution. For the euryglossines $\kappa$ was greater than 1.0 under both ML and MCMC (LRT > 6) suggesting that there is disproportionately more change in body size in longer branches. None of these results provide conclusive evidence for body size in the hylaeines and euryglossines to evolve outside of the default gradualism model.

Based on Simpson's (1949, 1953) model, the evolution of morphological adaptations in a new environment should follow an adaptive burst, as is commonly seen in patterns of lineage diversification. Harmon et al. (2010) looked at patterns of evolution in both body size and body shape in a variety of animal clades (squamates, birds, fish, arthropods, mammals, amphibians) that show adaptive radiations in their diversification. Harmon et al. (2010) conclude that the evolution of body size and shape within a clade does not follow patterns of lineage diversification through time. Rather, they found that evolution of body size and shape in all groups, except birds, follows a random walk model with a single-stationary peak (SSP), even in groups that are classic examples of adaptive radiation. Log-lineages through time (LTT) plots and $\gamma$ (gamma) simulations of Australian hylaeines suggested that early diversification occurred rapidly with a slowing down of diversification later on, suggesting adaptive radiation (Chapter 1), but this is not mirrored in the estimation of body size evolution. In the euryglossines, however, log-LTT plots and $\gamma$ (gamma) simulations inferred a constant rate of diversification (Chapter 2), which does concord with the gradual evolution of body size inferred from our *Continuous* results presented here. However, Harmon et al. (2010) further propose that a constant Brownian motion model is also less common than the SSP model they have tested, and indicate that future studies should incorporate this more complex model.

**ANCESTRAL SIZE**

The ancestral body size of hylaeines is larger than estimated for the euryglossine ancestral size, and this supports the overall size differences in extant hylaeines and euryglossines (Michener 2007). The inferred ancestral inter-tegular widths of both the hylaeines and Euryglossinae in our phylogenies are in the middle range of the measured species in our phylogenies. This is not consistent with Cope's Rule of increasing body size over time. This might be an artefact of the random walk method used (Moen 2006), however, the generalised-least squares platform of *Continuous* should counter this (Pagel 1997, 1999; Moen 2006). Of the two internal nodes estimated in the euryglossine phylogeny, the estimated ancestral size of Group B is of similar size to the root node, whereas the ancestral size of Group A is smaller. Of the two internal nodes estimated in the hylaeine phylogeny, the estimated ancestral size of Clade 1 is similar to the root node, and that of Clade 2 is larger.

Euryglossine species from Group A include the only known wood nesting species from the subfamily, and the ability to utilise nesting substrates with small preformed entrance holes might be linked to an evolutionary decrease in body size. However, Hylaeinae species are generally the same size and often larger than Euryglossinae (see Figures 2 and 3) and have been successful in retaining a wood nesting habit. Furthermore, there is some evidence supporting a wood nesting ancestor of Euryglossinae, as well as of Scrapterinae/Euryglossinae (see Chapter II), and this is perhaps an indication that larger species have reverted to a soil nesting habit. However, because nesting habit has been described from only eight of 13 Euryglossinae genera (and in most cases these descriptions are limited to only one to three species from each genus) speculation should be reserved for when more data is available.

It is possible that hairs can function as an insulator in bees, with the cold-climate adapted bumble bee, for example being very hairy (Church 1959). In this way, Australia's relatively warm climate might have allowed for euryglossines and hylaeines to develop a successful hairless alternative to pollen collection. For small bees, being hairless and without scopae might also be beneficial in flight. Most large bees (bumblebees, honey bees, Megachilidae) are quite hairy (even without scopae), but virtually all minute bees have few body hairs. Energy costs per gram of body mass increase dramatically as body size becomes minute (Schmidt-Nielsen 1972), and the

impact of air resistance (drag) might change with body size and shape with hair and scopae potentially increasing drag quite strongly. This would mean that for small bees, where the energy costs of flight are already greater than for large bees (Schmidt-Nielsen 1972), drag would make this even more costly. Reducing hair should reduce drag, and reducing scopae should have the same affect. Although some species of euryglossine and hylaeine bees are very small (Exley 1996; Michener 2007), considering the small size of Xeromelissinae (Colletidae), as well as other non-colletid bees it would seem that there is a reason other than small size for the extreme reduction of scopae in euryglossines and hylaeines.

**SCOPAE LOSS, INTERNAL POLLEN TRANSPORT, AND RICH NECTARIES**

Across all species of hylaeines and euryglossines, large and small, along with the absence of pollen-collecting scopae, is the feature of internal transport of pollen. All observed species of colletids provide offspring provisions that consist of relatively more nectar than pollen, whereas in most other bees the mix is reversed (Michener 1964). We suggest that transport of pollen in the crop is only viable if bees have collected enough nectar to make the crop contents a liquid or semi-liquid in order to regurgitate the mixture at the nest. This would require that abundant nectar was available close to the pollen source, especially in small bees, given that in bee species from six families, those with an inter-tegular width smaller than one millimetre had a foraging distance of less than 100 metres (Greenleaf et al. 2007). Around a third of the species we measured had an inter-tegular width smaller than 1.0 milimetre, and most had a width of between 1.0 and 2.0 mm.

Australian hylaeine and euryglossine species most commonly visit Myrtaceae species (Michener 1965; Exley 1967, 1968, 1972, 1974, 1975, 1976, 1978, 2001; Houston 1975, 1981). Myrtaceae flowers generally provide large amounts of nectar in shallow cups that do not require long tongues to access, possibly because of the importane of birds as pollinators for this family (Ford et al. 1979). Their lack of scopae means that euryglossines and hylaeines are unlikely to assist in pollination of Myrtaceae (as was suggested by Beardsell et al. 1993). However, because of the birds and other animals associated with Myrtaceae (Ford et al. 1979; Beardsell et al. 1993) the flowers probably provide an abundant and consistent pollen/nectar supply to those bees, and the small size

of Myrtaceae pollen grains (10-15 μm) (Beardsell et al. 1993) are likely to further allow for efficient resource transportation in the crop. Small bees prefer flowers in which they can forage with higher efficiency (Inouye 1980) and utilising flowers that rely on vertebrate pollinators means that pollen and nectar yields would be more than sufficient for the daily requirements of solitary bees. Furthermore, where pollen is abundantly available in a small area transport in the crop may become a more efficient method of transfer (Kuhlmann 2006), and especially, if, as in the orchid bees, the volume of nectar collected is not affected by body size (Borrell 2007).

Two other known examples of alimentary transport of pollen and nectar in bees are also colletids: the *Colletes fasciatus*-group (Colletinae) where eleven species have partly-reduced or reduced scopa (Kuhlmann 2006), and *Leioproctus (Euryglossidia) cyanescens* (Paracolletinae) (Houston 1981). Both authors suggest that the greater availability of flowers with small (< 40 μm) pollen grains, where scopa and hairs on the bee are not sufficiently dense to carry fine pollen, might lead to pollen collection in the crop. Evidence for increased efficiency of pollen and/or nectar transport with body size is not strong in non-social bees. For example, no correlation was found between female size and transported load in the aggregate nesting species, *Andrena agilissima* (Giovanetti and Lasso 2005), and pollen load does not increase proportionally with body size in scopae-using solitary bees (Neff 2008). In the solitary bee, *Megachile rotundata*, however, pollen load did increase with size (Klostermeyer et al. 1973).

Considering all colletids utilise at least as much nectar as they do pollen by volume in larval provisioning (as opposed to most other bees who use more pollen) (Michener 1964), we suggest that euryglossines and hylaeines have adapted to utilising the rich nectar source provided by Myrtaceae in Australia by swallowing both pollen and nectar (Michener 1964) (the abundant liquid nectar presumably enabling them to regurgitate the pollen as well as the nectar), thus inhibiting the use of scopae. A family-level body size study will best determine whether Euryglossinae and Hylaeine are phylogenetically smaller, and thus provide more conclusive evidence as to whether scopae loss at the sub-family level in these two groups is linked to small size.

## *Conclusions*

These results provide little evidence for phylogenetic signal in body size evolution, or that it occurs in an adaptive manner, and further investigation is needed with a larger dataset in order to determine whether the slightly positive signals are representative of any real patterns. Ancestral sizes in both the eurygliossines and hylaeines are of an intermediate size to the species in our phylogenies. The estimates derived from the internal nodes suggest that within the euryglossines there is a clade that has a decreased size from that of the root size. In the hylaeines, however, one clade has an increased size to that of the root estimate. It has been suggested that the lack of pollen collecting scopae in these two bee groups is due to their small size. We argue that the abundance of these groups in Australia and their predominance in foraging at nectar rich Myrtaceae might be an additional reason that they have adapted to transporting pollen internally.

**ACKNOWLEDGEMENTS**

Terry Houston of the Western Australian Museum (WAM) identified many Euryglossinae species, and Connal Eardley of the Agricultural Research Council in Pretoria identified Scrapterinae species used in this study. Simon van Noort's (Iziko Museum, Cape Town) assistance in the field in South Africa and that of Luke Chenoweth, Scott Groom, Sally Harradine, Michael McLeish, and Sandra Rehan in Australia is appreciated. WAM, and Kirstenbosch Research Centre in South Africa hosted PK. Michael McLeish assisted with graphics. This work was funded by ARC grants awarded to MPS, as well as Holsworth Wildlife Fund, and Sir Mark Mitchell awards to PK.

**LITERATURE CITED**

Almeida, E.A.B., and B.N. Danforth. 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50:290-309.

Beardsell, D.V., S.P. O'Brien, E.G. Williams, R.B. Knox, and D.M. Calder. 1993. Reproductive biology of Australian Myrtaceae. Aust. J. Bot. 41:511-26.

Borrell, B.J. 2007. Scaling of nectar foraging in orchid bees. Am. Nat. 169(5): 569-580.

Breed, M.D., J.M. Silverman, and W.J. Bell. 1978. Agonistic behavior, social interactions, and behavioral specialization in a primitively eusocial bee. Insect. Soc. 25(4):351-364.

Camargo, J.M.F., and D.W. Roubik. 1991. Systematics and bionomics of the apoid obligate necrophages: the *Trigona hypogea* group. Biol. J. Linn. Soc. 44: 13-39.

Cane, J.H. 1987. Estimation of bee size using intertegular span (Apoidea). J. Kansas Entomol. Soc. 60(1): 145-147.

Chown, S.L., and K.J. Gaston. 2010. Body size variation in insects: a macroecological perspective. Biol. Rev. 85:139-169.

Church, N.S. 1959. Heat loss and the body temperatures of flying insects ii. Heat conduction within the body and its loss by radiation and convection. J. Experimental Biology. 37(1): 186-213.

Davies, G.B.P., and D.J. Brothers. 2006. Morphology of *Scrapter* (Hymenoptera: Anthophila: Colletidae), with description of three new species and taxonomic status of five Cockerell taxa. Afr. Invertebr. 47:135-183.

Dowton, M., and A.D. Austin. 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy – the behaviour of the incongruence length difference test in mixed-model analyses. Syst. Biol. 51(1):19-31.

Dudley R. 2002. The biomechanics of insect flight: form, function, evolution. Princeton Univ. Press. USA.

Eickwort, G.C., and H.S. Ginsberg. 1980. Foraging and mating behavior in Apoidea. Ann. Rev. Entomol. 25:421-46

Exley, E.M. 1967. Revision of the genus *Euryglossula* Michener (Apoidea : Colletidae). Aust. J. Zool. 16:203-217.

Exley, E.M. 1968. Revision of the genus *Brachyhesma* Michener. (Apoidea : Colletidae). Aust. J. Zool. 16:167-201.

Exley, E.M. 1972. Revision of the genus *Pachyprosopis* Perkins (Apoidea : Colletidae). Aust. J. Zool.

Exley, E.M. 1974. A contribution to our knowledge of Australia's smallest bees with descriptions of new species (Hymenoptera: Colletidae: Euryglossinae). J. Aust. Entomol. Soc. 13:1-9.

Exley, E.M. 1975. Revision of the genus *Hyphesma* Michener (Apoidea : Colletidae). Aust. J. Zool. 23: 277-291.

Exley, E.M. 1976. Revision of the subgenus *Euryglossa* Smith (Apoidea : Colletidae : Euryglossinae). Aust. J. Zool. Suppl. Ser. No. 41: 1-72.

Exley, E.M. 1978. *Chaetohesma* - a new genus of Australian bees (Apoidea : Colletidae : Euryglossinae) Aust. J. Zool. 26:373-97.

Exley, E.M. 2001. The *walkeriana* species-group of *Euhesma* Michener (Hymenoptera: Colletidae: Euryglossinae). Aust. J. Entomol. 40(2):102–112.

Ford, H.A., D.C. Paton, and N. Forde. 1979. Birds as pollinators of Australian plants. New Zeal. J. Bot. 17:509-519.

Freckleton, R.P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. Am. Nat. 160(6): 712-726.

Gaston, K.J., and T.M. Blackburn. 2000. Pattern and process in macroecology. Blackwell Science, Oxford.

Giovanetti M., and E. Lasso. 2005. Body size, loading capacity and rate of reproduction in the communal bee *Andrena agilissima* (Hymenoptera; Andrenidae). Apidologie. 36:439–447.

Greenleaf, S.S., N.M. Williams, R. Winfree, and C. Kremen. 2007. Bee foraging ranges and their relationship to body size. Oecologia. 153:589-596.

Hanken J., and D.B. Wake. 1983. Miniaturization of body size: organismal consequences and evolutionary significance. Ann. Rev. Ecol. Syst. 24:501-519.

Harmon, L.J., J.B. Losos, T.J.Davies, R.G. Gillespie, J.L. Gittleman, W.B. Jennings, K.H. Kozak, M.A. McPeek, F. Moreno-Roark, T.J. Near, A. Purvis, R.E. Ricklefs, D.Schluter, J.A. Schulte II, O. Seehausen, B.L. Sidlauskas, O. Torres-Carvajal, J.T. Weir, and A.Ø. Mooers. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64(8):2385–2396.

Hone, D.W.E., and M.J. Benton. 2005. The evolution of large size: how does Cope's Rule work? Trends Ecol. Evol. 20(1):4-6.

Houston, T.F. 1969. Observations on the nests and behaviour of some euryglossine bees (Hymenoptera: Colletidae). J. Aust. Ent. Soc. 8:1-10.

Houston, T.F. 1975. A revision of the Australian Hylaeine bees (Hymenoptera: Colletidae). Aust. J. Zool. Suppl. Ser. No. 36:1-135.

Houston, T.F. 1981. A revision of the Australian Hylaeine bees (Hymenoptera: Colletidae) II. Aust. J. Zool. Suppl. Ser. No. 80:1-128.

Houston 1981. Leioproctus (euryglossidia) cyanscens

Huelsenbeck, J.P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics. 17:754-755.

Inouye, D.W. 1980. The effect of proboscis and corolla tube lengths on patterns and rates of flower visitation by bumblebees. Oecologica. 45:197-201.

Klostermeyer, E. C., S. J. Mech, Jr., and W. B. Rasmussen. 1973. Sex and weight of *Megachile rotundata* (Hymenoptera: Megachilidae) progeny associated with provision weights. J. Kansas Entomol. Soc.. 46: 536-548.

Kuhlmann, M. 2006. Scopa reduction and pollen collecting of bees of the *Colletes fasciatus*-group in the winter rainfall area of South Africa (Hymneoptera: Colletidae). J. Kansas Entomol. Soc.. 79(2):165-175.

Leys, R. 2000. Mate locating strategies of the green carpenter bees *Xylocopa* (Lestis) *aeratus* and *X*. (*L*.) *bombylans*. J. Zool. 252: 453-462.

Mayhew, P.J. 2007. Why are there so many insect species? Perspectives from fossils and phylogenies. Biol. Rev. 82:425–454.

Michener, C.D. 1964. Evolution of the nests of bees. American Zoologist. 4(2):227-239.

Michener, C.D. 1965. A classification of the bees of the Australian and South Pacific regions. B. Am. Mus. Nat. Hist. 130:1-362.

Michener, C.D. 1974. The Social Behavior of the Bees. Harvard University Press, Baltimore.

Michener, C.D. 2007. The Bees of the World, 2nd edn. Johns Hopkins University Press, Baltimore, MD.

Moen, D.S. 2006. Cope's rule in cryptodiran turtles: do the body sizes of extant species reflect a trend of phyletic size increase? Euro. Soc. Evol. Biol. 19:1210–1221.

Neff, J.L. 2008. Components of nest provisioning behavior in solitary bees (Hymenoptera: Apoidea). Apidologie. 39:30–45.

O'Brien, S.P. 1995. Extrafloral nectaries in *Chamelaucium uncinatum*: a first record in the Myrtaceae. Aust. J. Bot. 43:407-413.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. Zool. Scripta. 26:331-348.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature. 401 877-884.

Peters, R.H. 1983. The ecological implications of body size. Cambridge University Press, USA. 329 pp.

Ramalho, M., T.C. Giannini, K.S. Malagodi-Braga, and V.L. Imperatriz-Fonseca. 1994. Pollen harvest by stingless bee foragers (Hymenoptera, Apidae, Meliponinae). Grana. 33:239-244.

Ramalho, M., V.L. Imperatriz-Fonseca, and T.C. Giannini. 1998. Within-colony variation of foragers and pollen load capacity in the stingless bee *Melipona quadrifasciata anthidioides* Lepeletier (Apidae, Hymenoptera). Apidologie. 29:221-228.

Roubik, D.W. 1982. Obligate necrophagy in a social bee. Science 217: 1059-1060.

Sanderson, M.J. 2003. r8s: inferring absolute rates of evolution and divergence times in the absence of a molecular clock. Bioinformatics. 19:301–301.

Schmidt-Nielsen, K. 1972. Locomotion: energy cost of swimming, flying, and running. Science. 177:222-228.

Schoener, T.W., and D.H. Janzen. 1968. Notes on environmental determinants of tropical versus temperate insect size patterns. Am. Nat. 102(925):207-224.

Schwarz, M.P., S.M. Tierney, S.J.B. Cooper, and N.J. Bull. 2004. Molecular phylogenetics of the allodapine bee genus Braunsapis: A–T bias and heterogeneous substitution parameters. Mol. Phylogenet. Evol. 32:110–122.

Shaw, D.E., and D.F. Robertson. 1980. Collection of *Neurospora* by honeybees. Trans. Brit. Mycolog. Soc. 74(3):459-464

Shaw, D.E. 1998. Species of *Neurospora* recorded in Australia, and the collection of *Neurospora* conidia by honey bees in lieu of pollen. Mycologist. 12(4):154-158.

Simpson, G.G. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.

Simpson, G. G. 1953. The major features of evolution. Columbia Univ. Press, New York.

Spessa, A., M.P. Schwarz, and M. Adams. 2000. Sociality in *AmpHylaeine morosus* (Hymenoptera: Colletidae: Hylaeinae). Behaviour. 93(3):685-692.

Stanley, S.M. 1973. An explanation for Cope's rule. Evolution. 27:1-26.

Thorp, R.W. 1979. Structural, behavioral, and physiological adaptations of bees (Apoidea) for collecting pollen. Ann. Mo. Bot. Gard.66:788-812.

Thorp, R.W. 2000. The collection of pollen by bees. Plant Syst. Evol. 222:211-223.

Tierney, S., M.P. Schwarz, T. Neville, and P.M. Schwarz. 2002. Sociality in the phylogenetically basal allodapine bee genus *Macrogalea* (Apidae: Xylocopinae): implications for social evolution in the tribe Allodapini. Biol. J. Linn. Soc. 76:211-224.

Tokeshi, M., and P.E. Schmid. 2002. Niche division and abundance: an evolutionary perspective. Popul Ecol. 44:189–200.

Wcislo, W.T. 1989. Behavioral Environments and Evolutionary Change. Annu. Rev. Ecol. Syst. 20:137-169.

Wcislo, W.T., and J.H. Cane. 1996. Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies. Annu. Rev. Entomo. 41:257-286.

Wetterer, J.K. 1989. Central place foraging theory: when load size affects travel time. Theor. Pop. Biol. 36:267-280.

Wheeler, D.E. 1986. Developmental and physiological determinants of caste in social hymenoptera: evolutionary implications. Am. Nat. 128(1):13-34.

Table 1a: Primer specifications.

| Locus | Primer name | Sequence 5'-3' | Reference |
|---|---|---|---|
| EF-1α | HaF2For1 | GGG YAA AGG WTC CTT CAA RTA TGC | Danforth et al. 1999 |
| | F2-rev1 | A ATC AGC AGC ACC TTT AGG TGG | Danforth et al. 1999 |
| 28S rRNA | A-28S-For | CCC CCT GAA TTT AAG CAT AT | Ward and Brady 2003 |
| | Mar28S-Rev | TAG TTC ACC ATC TTT CGG GTC CC | Mardulyn and Whitfield 1999 |
| COI | Jerry (C1-J-2183) | CAA CAT TTA TTT TGA TTT TTT GG | Simon et al. 1994 |
| | Pat (TL2-N-3014 | TCC AAT GCA GTA ATC TGC CAT ATT A | Simon et al. 1994 |
| | LCO1490 | CCT TTT ATA ATT GGA GGA TTT GG | Folmer et al. 1994 |
| | M399 | TCA TCT AAA AAC TTT AAT TCC TG | Schwarz et al. 2004 |

Table 1b: PCR conditions for primers used showing temperature and time of: initial denaturation; and 35 cycles of denaturation, annealing, extension.

| Primer | PCR Conditions |
|---|---|
| HaF2For1/F2-rev1 | 94˚C 3min; 94˚C 60sec, 52˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| A-28S-For/Mar28S-Rev | 94˚C 3min; 94˚C 60sec, 58˚C 60sec, 72˚C 1.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |
| Jerry/Pat/ LC01490/M399 | 94˚C 3min; 92˚C 30sec, 52˚C 90sec, 72˚C 2.5min |
| | 94˚C 60sec; 94˚C 30sec, 52˚C 60sec, 72˚C 60sec |

Table 2: Scaling parameters lambda, kappa, and delta and the significance of values, taken from M. Pagel *Continuous* manual.

| **Parameter** | **Action** | **0** | **<1** | **1** | **>1** |
|---|---|---|---|---|---|
| λ lambda | Assess contribution of phylogeny | Star phylogeny (species independent) | Phylogenetic history has minimal effect | Default phylogeny | Not defined |
| κ kappa | Scale branch lengths in tree | Punctuational evolution | Stasis in longer branches | Default gradualism | Longer branches more change |
| δ delta | Scale total path (root to tip) in tree | Not defined | Temporally early change important (adaptive radiation) | Default gradualism | Temporally later change important (species-specific adaptation) |

Table 3a: Species names, collection locations, and GenBank accession numbers of hylaeine species used in this study. 'Clade' refers to which of the two major clades in the phylogeny the species belongs to (see Fig. 2). Abbreviated collection localities refer to states in Australia: NT=Northern Territory; Qld=Queensland; SA=South Australia; Vic=Victoria, WA=Western Australia. GenBank accession numbers that begin with JN are from a previous study of the authors', those beginning with AY or DQ are from two other studies.

| Species | Clade | Collection locality | 28S | EF-1 | COI | |
|---|---|---|---|---|---|---|
| *Hemirhiza* group | 1 | Cardwell, Qld | JN566243 | JN566289 | - | JN578924 |
| Hylaeinae sp. | 1 | North of Highway Inn Road House, NT | JN566244 | JN566290 | - | JN578925 |
| *Hylaeus* sp. 3 | 2 | Lake Gilles, SA | JN566278 | JN566329 | JN578912 | JN578962 |
| *Hylaeus (Edriohylaeus) ofarrelli* | 1 | GenBank | DQ768585 | DQ884676 | - | - |
| *Hylaeus (Euprosopellus) pergibbosus* | 2 | Ngaraket, SA | JN566247 | JN566293 | JN578891 | JN578928 |
| *Hylaeus (Euprosopis)* sp. 7 | 2 | Glen Eagle, WA | - | JN566296 | JN578892 | JN578931 |
| *Hylaeus (Euprosopis) disjunctus* | 2 | GenBank | DQ768586 | DQ884677 | - | - |
| *Hylaeus (Euprosopis) elegans* | 2 | Quorn, SA | JN566248 | JN566294 | - | JN578929 |
| *Hylaeus (Euprosopis) husela* | 2 | Larrimah, NT | JN566249 | JN566295 | - | JN578930 |
| *Hylaeus (Euprosopis) violaceus* | 2 | Crusoe Beach, WA | JN566250 | JN566297 | - | JN578932 |
| *Hylaeus (Euprosopoides)* sp. | 2 | Augusta, WA | JN566252 | JN566299 | JN578894 | JN578934 |
| *Hylaeus (Euprosopoides)* sp. 1 | 2 | Augusta, WA | JN566253 | JN566300 | JN578895 | JN578935 |
| *Hylaeus (Euprosopoides)* | 2 | GenBank | DQ768587 | DQ884678 | - | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| *cyanurus* | | | | | | |
| *Hylaeus (Euprosopoides) ruficeps* | 2 | Mataranka, NT | JN566251 | JN566298 | JN578893 | JN578933 |
| *Hylaeus (Gephrohylaeus) sculptus* | 1 | GenBank | DQ768598 | DQ884688 | - | - |
| *Hylaeus (Gnathoprosopis) albonitens* | 1 | Larrimah, NT | JN566254 | JN566301 | JN578896 | JN578936 |
| *Hylaeus (Gnathoprosopis) amiculus* | 1 | GenBank | DQ872777 | DQ212154 | - | AY913955 |
| *Hylaeus (Gnathoprosopis) euxanthus* | 1 | Port Lincoln, SA | - | JN566304 | JN578898 | JN578939 |
| *Hylaeus (Gnathoprosopoides) bituberculatus* | 1 | GenBank | DQ768610 | DQ884699 | - | - |
| *Hylaeus (Heterapoides) exleyae* | 1 | GenBank | DQ768599 | DQ884689 | - | - |
| *Hylaeus (Heterapoides) extensus* | 1 | GenBank | DQ768600 | DQ884690 | - | - |
| *Hylaeus (Hylaeorhiza) nubilosus* | 1 | GenBank | DQ768591 | DQ884681 | - | - |
| *Hylaeus (Hylaeteron) douglasi* | 1 | South-west WA | JN566255 | JN566306 | - | - |
| *Hylaeus (Macrohylaeus) alcyoneus* | 1 | GenBank | DQ768577 | DQ884668 | - | - |
| *Hylaeus (Planihylaeus) trilobatus* | 2 | GenBank | DQ768611 | DQ884700 | - | - |
| *Hylaeus (Prosopisteron)* sp. 2 | 1 | Charters Towers, Qld | JN566272 | JN566323 | - | JN578956 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Hylaeus (Prosopisteron)* sp. 5 | 1 | Daintree, Qld | JN566269 | JN566320 | JN578905 | JN578953 |
| *Hylaeus (Prosopisteron)* sp. 2 | 1 | 94km south of Williams, WA | JN566271 | JN566322 | JN578907 | JN578955 |
| *Hylaeus (Prosopisteron)* sp. | 1 | Crusoe Beach, WA | JN566262 | JN566313 | JN578902 | JN578947 |
| *Hylaeus (Prosopisteron)* sp. 3 | 1 | Geraldton, WA | JN566270 | JN566321 | JN578906 | JN578954 |
| *Hylaeus (Prosopisteron) aralis* | 1 | GenBank | DQ768590 | DQ884680 | - | - |
| *Hylaeus (Prosopisteron) bidentatus* | 1 | GenBank | DQ768614 | DQ884702 | - | - |
| *Hylaeus (Prosopisteron) burnsi* | 2 | GenBank | DQ768620 | DQ884708 | - | - |
| *Hylaeus (Prosopisteron) chlorosomus* | 1 | Whyalla, SA | JN566263 | JN566314 | JN578903 | JN578948 |
| *Hylaeus (Prosopisteron) cyaneomicans* | 1 | GenBank | DQ768616 | DQ884704 | - | - |
| *Hylaeus (Prosopisteron) eugeniellus* | 1 | GenBank | DQ768594 | DQ884684 | - | - |
| *Hylaeus (Prosopisteron) microphenax* | 1 | GenBank | DQ768595 | DQ884685 | - | - |
| *Hylaeus (Prosopisteron) perhumilis* | 1 | Port Lincoln, SA | JN566259 | JN566310 | JN578901 | JN578944 |
| *Hylaeus (Prosopisteron) primulipictus* | 1 | GenBank | DQ768593 | DQ884683 | - | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Hylaeus (Prosopisteron) quadratus* | 1 | Augusta, WA | JN566265 | JN566316 | - | - |
| *Hylaeus (Prosopisteron) simplus* | 1 | GenBank | DQ768573 | DQ884664 | - | - |
| *Hylaeus (Prosopisteron) vittatifrons* | 1 | Alawoona, SA | JN566266 | JN566317 | JN578904 | JN578950 |
| *Hylaeus (Rhodohylaeus)* sp. 4 | 2 | Perth, WA | JN566273 | JN566324 | - | JN578957 |
| *Hylaeus (Rhodohylaeus)* sp. 3 | 2 | Nelson, Vic | JN566275 | - | JN578909 | JN578959 |
| *Hylaeus (Rhodohylaeus)* sp. | 2 | Whyalla, SA | JN566274 | JN566326 | JN578908 | JN578958 |
| *Hylaeus (Rhodohylaeus)* sp. 1 | 2 | Halls Creek, WA | JN566276 | JN566327 | JN578910 | JN578960 |
| *Hylaeus (Rhodohylaeus)* sp. 2 | 2 | Kojunup, WA | JN566277 | JN566328 | JN578911 | JN578961 |
| *Hylaeus (Rhodohylaeus) constrictiformis* | 2 | GenBank | DQ768596 | DQ884686 | DQ872733 | - |
| *Hylaeus (Rhodohylaeus)* near *lateralis* | 2 | Kojunup, WA | - | JN566325 | - | - |
| *Hylaeus (Rhodohylaeus) proximus* | 2 | GenBank | AY654493 | AY585130 | DQ87273 | - |
| *Hylaeus (Sphaerhylaeus)* sp. 4 | 1 | Ngarakat, SA | JN566280 | JN566331 | JN578913 | JN578964 |
| *Hylaeus (Xenohylaeus) leptospermi* | 2 | GenBank | DQ768612 | DQ884701 | - | - |
| *Meroglossa striaticeps* | 1 | Qld | JN566283 | - | JN578917 | JN578966 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Meroglossa* near *eucalypti* | 1 | ? | JN566281 | JN566333 | JN578915 | JN578965 |
| *Meroglossa eucalypti* group | 1 | 90 km north of Adelaide River, NT | - | JN566332 | JN578914 | - |
| *Meroglossa torrida* group | 1 | 30 km north of Katherine, NT | - | JN566335 | JN578918 | - |
| *Meroglossa rubricata* | 1 | Gerladton, WA | JN566282 | JN566334 | JN578916 | - |

Table 3b. Collection localities and GenBank accession numbers of Euryglossinae species used in this study. 'Group' refers to which of the two major clades (Group A or Group B) in the phylogeny the species belongs to (see Fig. 3). Abbreviated collection localities refer to states in Australia: NT=Northern Territory; Qld=Queensland; SA=South Australia; Vic=Victoria, WA=Western Australia. GenBank accession numbers beginning with JN are from a previous study of the authors', those beginning with AY or DQ are from two other studies.

| Species | Group | Location | 28S | EF1 | COI | |
|---|---|---|---|---|---|---|
| *Brachyhesma* sp. | B | Adelaide, SA | JN603373 | JN603438 | - | JN603396 |
| *Callohesma* sp. | B | NT | JN603376 | JN603441 | JN603421 | JN603399 |
| *Callohesma* sp. 1 | B | Adelaide, SA | JN603377 | JN603442 | JN603422 | JN603400 |
| *Callohesma* sp. 2 | B | Ngaraket, SA | JN603378 | JN603443 | JN603423 | JN603401 |
| *Callohesma lacteipennis* | B | Scott Creek, SA | JN603375 | JN603440 | JN603420 | JN603398 |
| *Callohesma sinapipes* | B | Whyalla, SA | JN566238 | JN566284 | JN57886 | JN578919 |
| *Euhesma* sp. 1 | B | Cooee Bay, Qld | JN603379 | JN603444 | JN603424 | JN603402 |
| *Euhesma* sp. 2 | B | Cooee Bay, Qld | JN603380 | JN603445 | JN603425 | JN603403 |
| *Euhesma* sp. 3 | B | 94 km south of Williams, WA | JN603381 | JN603446 | JN603426 | JN603404 |
| *Euryglossina* sp. 1 | A | Broome, WA | JN603382 | JN603447 | - | JN603405 |
| *Euryglossina* sp. 2 | A | Qld | JN603383 | JN603448 | JN603427 | JN603406 |
| *Euryglossina* sp. 3 | A | Whyalla, SA | JN603384 | JN603449 | JN603428 | JN603407 |
| *Euryglossula* sp. 1 | A | Larrimah, NT | JN603385 | JN603450 | JN603429 | JN603408 |
| *Euryglossula* sp. 2 | A | Jurien Bay, WA | JN603386 | JN603451 | JN603430 | JN603409 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Euryglossula* sp. 3 | A | Charters Towers, Qld | JN603387 | JN603452 | JN603431 | JN603410 |
| *Hyphesma* sp. 1 | B | Kojonup, WA | JN603388 | JN603453 | JN603432 | JN603411 |
| *Hyphesma* sp. 2 | B | Scott Creek, SA | JN603389 | JN603454 | JN603433 | JN603412 |
| *Pachyprosopis* (*Pachyprosopula*) sp. | A | Warwick, Qld | JN603389 | JN603455 | JN603434 | JN603413 |
| *Xanthesma furcifera* | B | GenBank | DQ872770 | AY585140 | - | - |

Table 4a. Results of likelihood ratio tests (LRT) for hylaeines under Maximum Likelihood (ML) for the 500 randomly selected chronograms as well as the consensus chronogram (with polytomies excluded). Lh = log-likelihood as calculated by *Continuous*. LRT = (lLh [better fitting model] – (Lh [worse fitting model]). The positive LRT in each case suggest that the estimated values of each parameter fit the phylogeny better than 1.0 for all parameters, and better than 0.0 in the case of λ (lambda) and κ (kappa). δ (delta) was not tested against 0.0, as this value cannot be defined (see Table 1). Parameter values that were significantly different from both 0.0 and 1.0 (i.e. LRT > 2) are indicated by an asterisk. The ancestral inter-tegular widths of the root node and two internal nodes were estimated using the estimated parameter values and are shown in Figure 2.

| **500 random chronograms** | | | **Consensus chronogram** | | |
|---|---|---|---|---|---|
| **Parameter** | **Lh** | **LRT** | **Parameter** | **Lh** | **LRT** |
| λ=0 | -26.530086 | 19.49 | λ=0 | -25.235282 | 16.79 |
| **λ=1** | -27.444083 | 21.32 | **λ=1** | -21.805484 | 9.93 |
| λ = **0.892*** | -16.78484 | | λ = **0.890*** | -16.83944 | |
| δ=1 | -16.790428 | 1.50 | δ=1 | -16.83944 | 1.51 |
| δ = **0.495** | -16.039822 | | δ = **0.465** | -16.085946 | |
| κ=0 | -20.833624 | 9.87 | κ=0 | -20.299201 | 8.26 |
| κ=1 | -16.791819 | 1.78 | κ=1 | -16.870059 | 1.40 |
| κ = **1.822** | -15.900471 | | κ = **1.876** | -16.167817 | |

Table 4b. Results of likelihood ratio tests (LRT) for euryglossines under Maximum Likelihood (ML) for the 500 randomly selected chronograms as well as the consensus chronogram. Lh = log-likelihood as calculated by *Continuous*. LRT = (lLh [better fitting model] – (Lh [worse fitting model]). The positive LRT in each case suggest that the estimated values for each parameter fit the phylogeny better than 1.0 for all parameters, and better than 0.0 in the case of λ (lambda) and κ (kappa). δ (delta) was not tested against 0.0, as this value cannot be defined (see Table 1). Parameter values that were significantly different from both 0.0 and 1.0 (i.e LRT > 2) are indicated by an asterisk. The ancestral inter-tegular widths of the root node and two internal nodes were estimated using the estimated parameter values and are shown in Figure 3.

| **500 random chrongrams** | | | **Consensus chronogram** | | |
|---|---|---|---|---|---|
| **Parameter** | **Lh** | **LRT** | **Parameter** | **Lh** | **LRT** |
| λ=0 | -8.657375 | 10.72 | λ=0 | -8.657375 | 10.68 |
| **λ=1** | -8.5541904 | 10.52 | **λ=1** | -7.853778 | 9.07 |
| λ = **0.834*** | -3.2953136 | | λ = **0.835*** | -3.320073 | |
| δ=1 | -3.2975625 | 0.821 | δ=1 | -3.320073 | 0.78 |
| δ = **0.514** | -2.8869244 | | δ = **0.508** | -2.930436 | |
| κ=0 | -4.5976964 | 9.66 | κ=0 | -4.626271 | 10.13 |
| κ=1 | -3.2975625 | 7.06 | κ=1 | -3.380125 | 7.63 |
| κ = **2.999*** | 0.23198502 | | κ = **2.999*** | 0.436619 | |

Table 5a. Results of likelihood ratio tests (LRT) for hylaeines under MCMC for the 500 randomly selected chronograms. Lh = log-likelihood as calculated by *Continuous*. LRT = (lLh [better fitting model] – (Lh [worse fitting model]). The positive LRT in each case suggest that the estimated values for each parameter fit the phylogeny better than 1.0 for all parameters, and better than 0.0 in the case of λ (lambda) and κ (kappa). δ (delta) was not tested against 0.0, as this value cannot be defined (see Table 1). Parameter values that were significantly different from both 0.0 and 1.0 (i.e LRT > 2) are indicated by an asterisk. The ancestral inter-tegular widths of two internal nodes were estimated using the estimated parameter values and are shown in Figure 2.

| **MCMC random chrongrams** | | |
|---|---|---|
| **Parameter** | **Lh** | **LRT** |
| λ=0 | -27.047032 | 19.12 |
| λ=1 | -36.645981 | 0.08 |
| λ = 0.9994 | -36.608663 | |
| δ=1 | -21.03206759 | 4.35 |
| δ = 0.245* | -18.85536053 | |
| κ=0 | -19.7624329 | 1.79 |
| κ=1 | -34.994569 | 32.26 |
| κ = 0.47 | -18.86602248 | |

Table 5b. Results of likelihood ratio tests (LRT) for euryglossines under MCMC for the 500 randomly selected chronograms. Lh = log-likelihood as calculated by *Continuous*. LRT = (lLh [better fitting model] – (Lh [worse fitting model]). The positive LRT in each case suggest that the estimated values for each parameter fit the phylogeny better than 1.0 for all parameters, and better than 0.0 in the case of λ (lambda) and κ (kappa). δ (delta) was not tested against 0.0, as this value cannot be defined (see Table 1). Parameter values that were significantly different from both 0.0 and 1.0 (i.e LRT > 2) are indicated by an asterisk. The ancestral inter-tegular widths of two internal nodes were estimated using the estimated parameter values and are shown in Figure 3.

| **MCMC random chrongrams** | | |
|---|---|---|
| **Parameter** | **Lh** | **LRT** |
| λ=0 | -9.202168554 | 9.81 |
| λ=1 | -7.5426728 | 6.49 |
| λ =0.733* | -4.2997428 | |
| δ=1 | -3.999339882 | 0.33 |
| δ = 0.94 | -4.162306777 | |
| κ=0 | -5.179451754 | 8.69 |
| κ=1 | -4.0441401 | 6.42 |
| κ = 3.38* | -0.834467592 | |

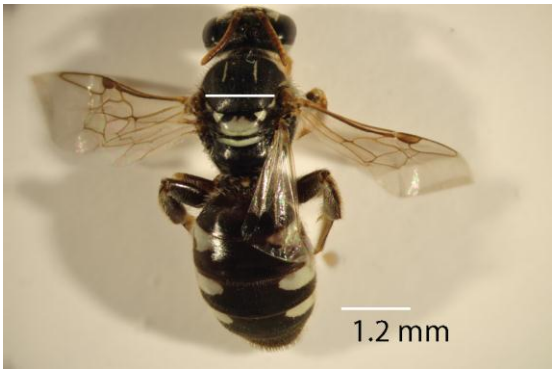



Figure 1. *Callohesma* sp. 2 (Euryglossinae) with inter-tegular width marked by white line. See Figure 2 (hylaeine) and Figure 3 (euryglossine) for list of inter-tegular widths of other species.

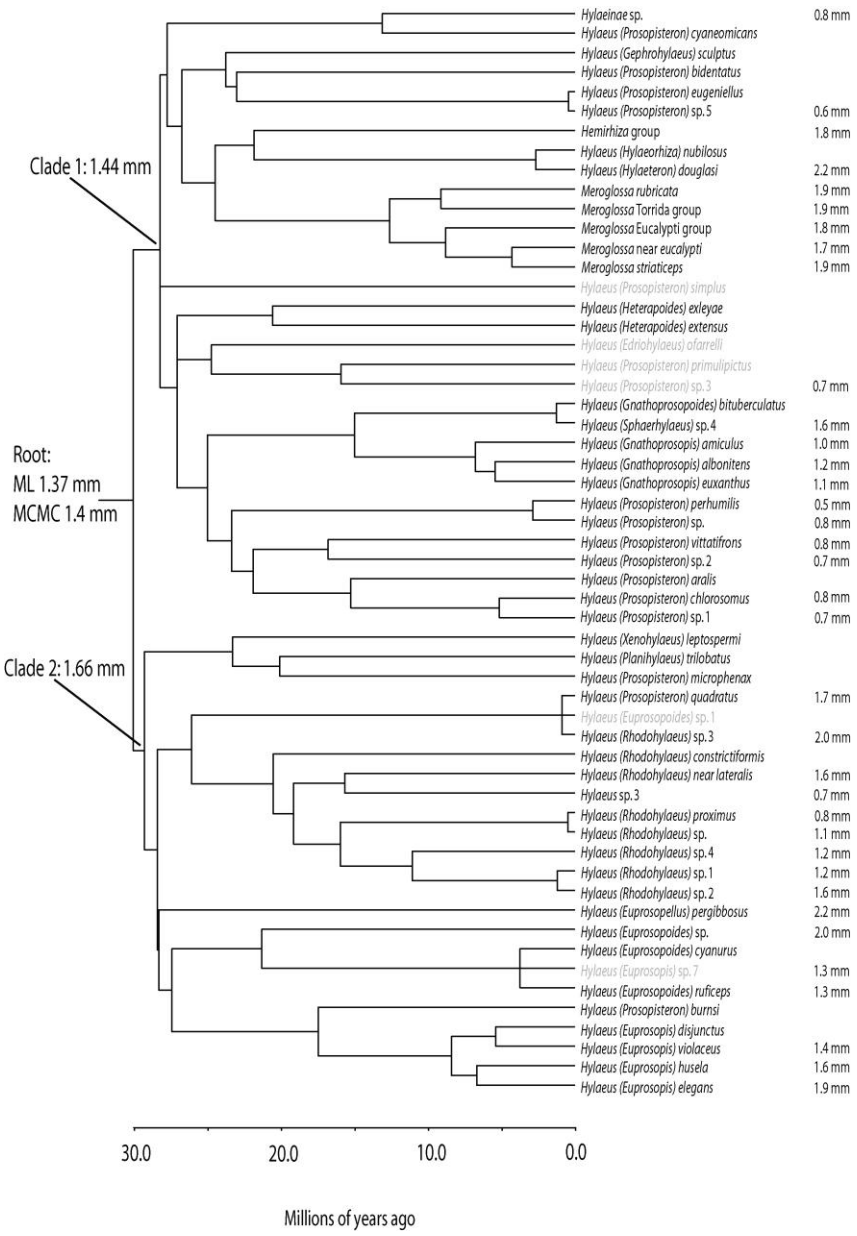



Figure 2. Consensus chronogram of hylaeine species used in character state reconstructions. Inter-tegular widths are listed next to those species for which data was available. Species highlighted in grey were pruned from the consensus tree for *Continuous* analyses, which does not accept polytomies. Estimated ancestral inter-tegular widths are given for three nodes. Ancestral inter-tegular widths for the root node were estimated using both a maximum likelihood (ML) and MCMC approach. Clade 1 and Clade 2 represent two major clades in the phylogeny for which ancestral states were estimated – internal nodes could only be estimated using MCMC.

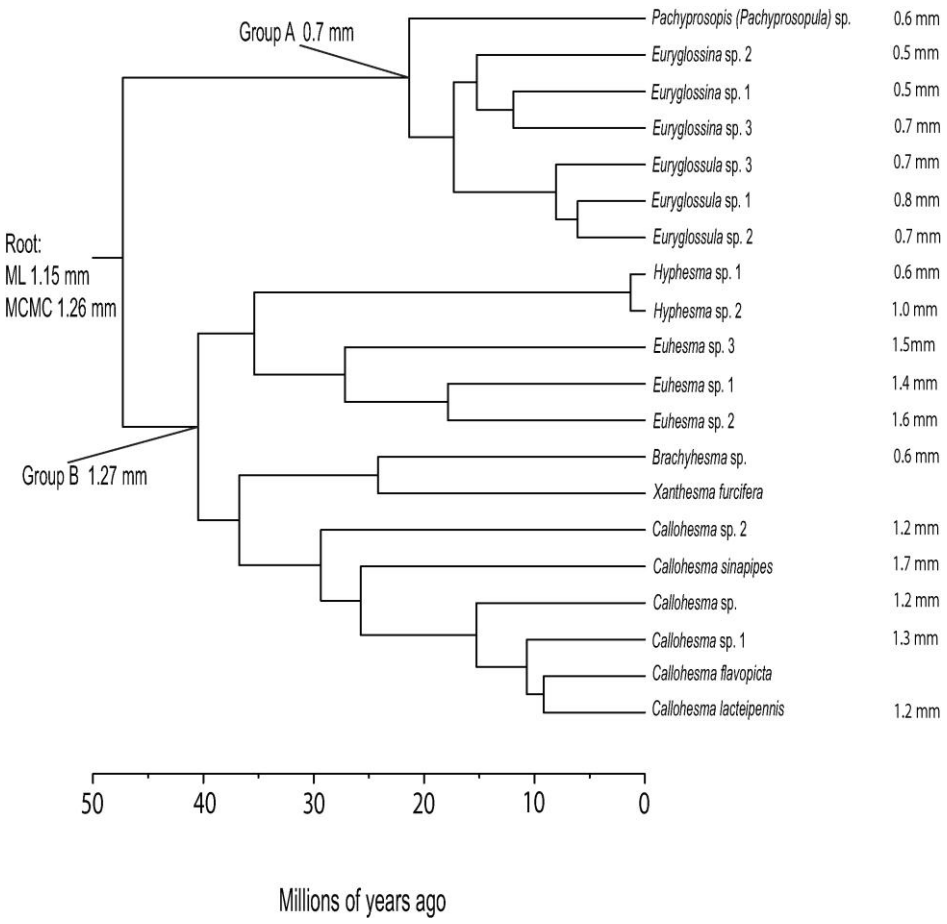
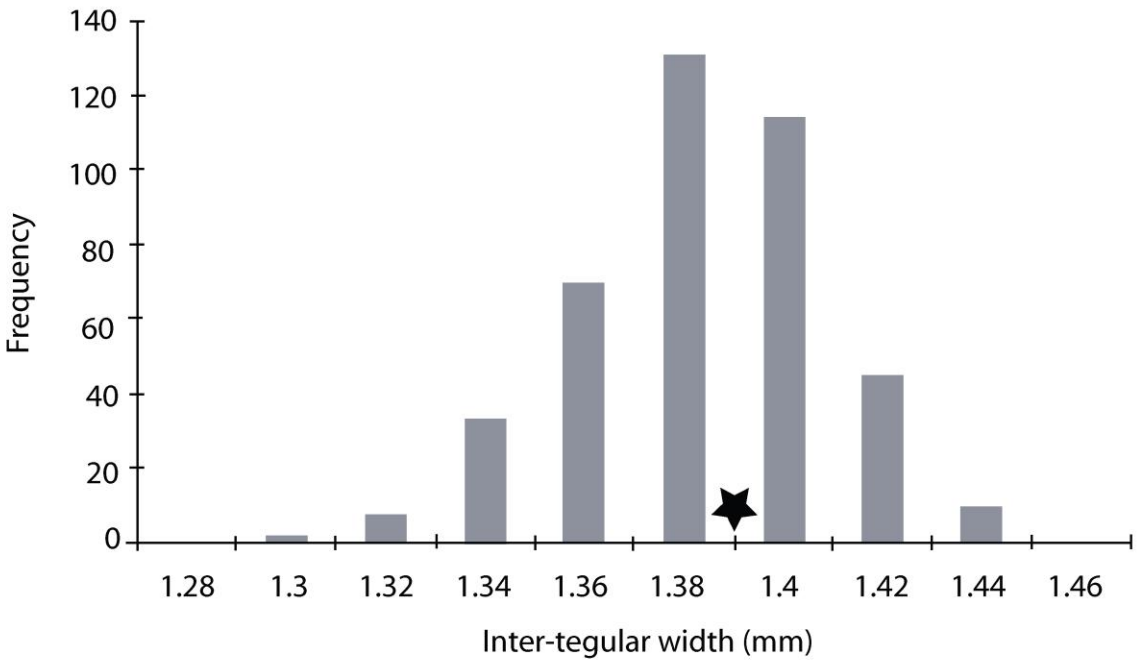



Figure 3. Consensus chronogram of euryglossines used in character state reconstruction. Inter-tegular widths are listed next to those species for which data were available. Estimated ancestral inter-tegular widths are given for three nodes. Ancestral inter-tegular widths for the root node were estimated using both a maximum likelihood (ML) and MCMC approach. Group A and Group B represent two major clades within the phylogeny for which ancestral states were estimated – internal nodes could only be estimated using MCMC.

4a




4b

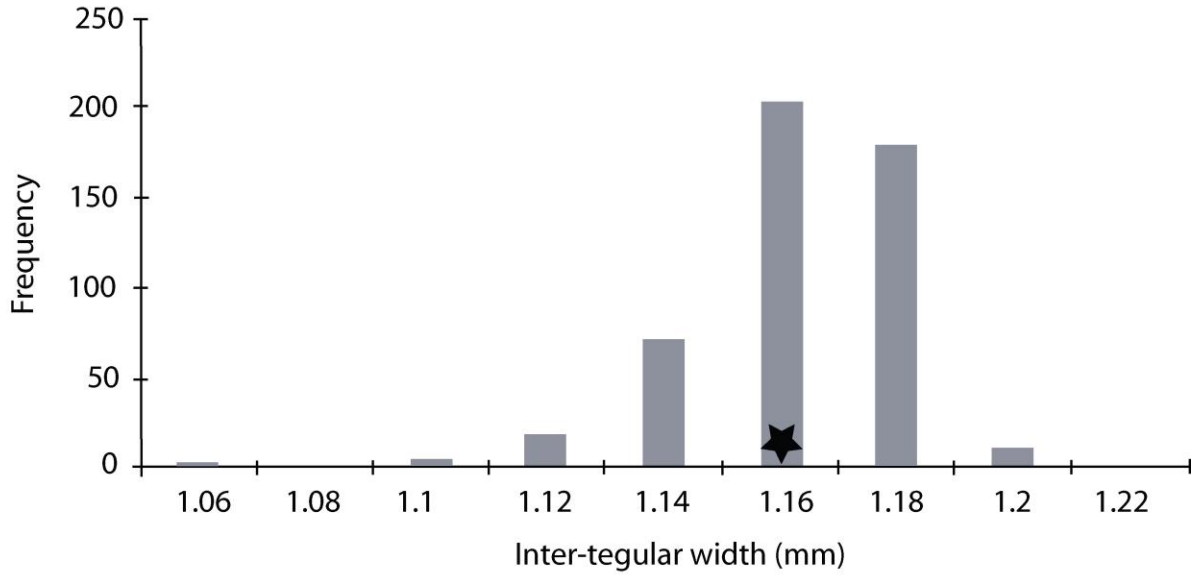



Figure 4. Histograms of estimated maximum likelihood (ML) ancestral values across 500 random chronograms for (a) hylaeines, and (b) euryglossines. Stars represent the ancestral values estimated from consesnsus chronograms for (a) hylaeines, and (b) euryglossines.

# GENERAL DISCUSSION

## SUMMARY OF FINDINGS

The analyses presented here on hylaeine and euryglossine bees add to recent estimates of diversification and dispersal events in several other major bee groups, covering the period 50-30 Mya. These other bee groups include the Allodapini (Schwarz et al. 2006; Chenoweth and Schwarz 2010), Bombini (Hines 2008), Ceratinini (Rehan et al. 2010), Ctenoplecterini (Schaefer and Renner 2008), Halictini (Danforth et al. 2004) and Meliponini (Rasmussen and Cameron 2010), and importantly, the estimated dates for euryglossines and hylaeines presented here are similar to those estimated by Almeida et al. (in press).

An origin of *Hylaeus* approximately 30 Mya (Chapter 1) is too recent for the non-Australasian clade to have diverged as a result of vicariance, and suggests that the global distribution of *Hylaeus* occurred via dispersal. There is no indication from the molecular data presented here, or from available taxonomic data, of subsequent dispersal events (other than to New Zealand) despite the northward drift of Australia towards south-east Asia providing a weakening physical barrier for dispersal into the northern hemisphere. This possibly reflects the significance that available niche space has on successful colonisation as opposed to weak physical barriers. *Hylaeus* diversity outside of Australia is comprised of over 500 species, however, other Hylaeinae genera have not been successful in dispersing and/or establishing outside of Australia and New Guinea. This might be related to competitive restraints and ecological limitations outside of Australia and New Guinea, similar to factors that have restricted Euryglossinae success to Australia, with a very limited presence in New Zealand and New Caledonia. Log-lineages through time (LTT) plots suggest an early burst in the diversification of the Australian *Hylaeus* species (including *Meroglossa* and *Hemirhiza*), with a decrease in diversification around 25 Mya. A sudden decrease in diversification could be attributable to a large extinction event within a clade, but those events are more often represented by an anti-sigmoidal curve. The Moran-type plot found here for *Hylaeus* is a pattern generally associated with an adaptive radiation (Crisp and Cook 2009), whereby available niche space has enabled rapid diversification, with the slowing of

diversification reflecting the decreasing availability of niche space. Two of the eight species known from New Zealand fall within the Australian *Hylaeus* and probably dispersed from Australia approximately 25 Mya. Representatives of *Hylaeus* that are found outside of Australia and New Zealand are strongly supported to be a monophyletic group sister to *Hylaeus (Euprosopis) + H. (Prosopisteron) burnsi*. Strong support for the monophyly of all *Hylaeus* species from outside of Australia and New Zealand supports the taxonomic division of those species into subgenera not found in Australasia (Michener 1965, 2007). However, the taxonomy of the Australian Hylaeinae is in need of revision, with species from the genera *Meroglossa* and *Hemirhiza* falling within *Hylaeus*.

Although the estimated 380 species missing from the Euryglossinae phylogeny (Chapter 2) might limit the extent of estimates of diversification and divergence times, the topology inferred from the molecular phylogeny does agree largely with the separation of Euryglossinae into two clades (Groups A and B) based on wing morphology (Michener 2007). Along with certain discriminating wing characteristics, Group A also includes the only known stem-nesting Euryglossinae species and parsimony-based anlyses inferred that that there has been a reversion to a stem-nesting habit within Group A as well as retaining the soil nesting habit found in Group B species and the sister subfamily to the Euryglossinae, the sub-Saharan Scrapterinae. However, the Bayesian approach inferred a wood nesting ancestor at major nodes, indicating that soil-nesting is a derived state within Scrapterinae and Euryglossinae. Because descriptions of nesting habit in Euryglossinae species are few, evidence to conclusively support the assumption that species in Group B do not also have a stem-nesting habit is not sufficient, and both analyses into ancestral nesting state should be treated with caution. The sister-group relationship between Scrapterinae and Euryglossinae inferred by a previous study (Almeida and Danforth 2009) is supported. A divergence between the two subfamilies of approximately 60 Mya (Almeida et al. in press) and the probable ancestral range of this lineage in Australia, is an indication of a trans-oceanic dispersal event from Australia/Antarctica/South America to Africa. Subsequent long-range dispersal events for the Euryglossinae can only be inferred by the presence of a few species in New Zealand and New Caledonia, and long-range dispersal has not occurred for the Scrapterinae, which is restricted to southern Africa. Like Hylaeinae, it seems that the

high diversity and abundance of Euryglossinae in Australia is not matched by high dispersal rates from Australia to other regions.

Phylogenetic signal was detected in the evolution of body size in both the euryglossines and hylaeines (Chapter 3). In both groups a gradual evolutionary change was inferred, but within the Euryglossinae more change is associated in lineages with longer branches. The ancestral state in both groups was inferred to be of an intermediate state, suggesting that decreases in body size are just as likely as increases. One of the striking morphological traits shared by all species of Hylaeinae and Euryglossinae, but only by a few species within other bee groups, is the absence of pollen-collecting scopae. Strong evidence that hylaeines and euryglossines are not sister-clades suggests that the loss of scopae has occurred twice in the Colletidae, and it has been suggested that this feature has enabled species of these groups to have a smaller body size (Pasteels and Pasteels 1976 in Davies and Brothers 2006, p. 160), relative to other bees. I argue that together, the prevelance of nectar-rich Myrtaceae in Australia and the nectar-rich provisions provided by colletid bees to their young has led to the adaptation of internal pollen transport in hylaeines and euryglossines, and to the reduction of scopae.

**HYLAEINE TAXONOMY**

Morphological similarities that could potentially be adapted into taxonomic revisions of Hylaeinae to reassign some genera as a *Hylaeus* subgenus, or vice versa are: the presence of epidermal glands in the scapes of males in some *Meroglossa* and some *Hylaeus* (Houston 1975), and the form of the spermathecal duct in females of some *Hylaeus*, *Meroglossa*, and *Hemirhiza* species (Houston 1975).

There are examples from several other bee groups where current taxonomies do not always reflect phylogeny. For example, on cataloguing the Afrotropical bees, Eardley and Urban (2010) highlighted that several species were recorded in genera that are not Afrotropical, e.g., *Halictus* and *Osmia*, indicating the need for revision of those groups. Taxonomic revisions within Colletidae have particularly been proposed for cases within the subfamily Colletinae. *Scrapter* (originally placed within Colletinae) is now classed as monogeneric family, Scrapterinae, (Ascher and Engel 2006; Almeida and Danforth 2009), with strong support for that reassignment from molecular evidence (Almeida and Danforth 2009; Chapter 2). This is also the case for *Callomelitta*, which is now placed in

a monogeneric subfamily Callomelittinae (Almeida 2008a). As mentioned by Almeida (2008b), *Callomelitta* larvae were described as being very different from Paracolletinae larvae by McGinley in 1981. That *Callomelitta* has only been re-assigned nearly 20 years later indicates the need to bring together information from a variety of sources. In this way, molecular phylogenies can be effectively used to, amongst other things, corroborate or question current taxonomic standings.

## THE INFLUENCE OF CLIMATE ON DIVERSIFICATION, LONG-RANGE DISPERSAL, AND BODY SIZE

The detection of a period of rapid diversification, as seen in *Hylaeus* (Chapter 1), can be a reflection of a clade that has a greater number of species (e.g. Mullen et al. 2011), or is older (McPeek and Brown 2007). A log-LTT plot for the whole Colletidae family (Almeida et al. in press) suggested that diversification within the family was initially constant but with an increase in diversification around 30 Mya. Almeida et al. (in press) suggest that this increase in diversification rate may be associated with the inferred origins of *Hylaeus* and *Colletes* (Colletinae) 30 Mya. *Colletes* is the second most speciose colletid genus after *Hylaeus* and is also widely distributed (Michener 2007), albeit absent from Australia, as well as Madagascar. The subfamily Hylaeinae has an estimated crown age of 45 Mya, which is similar to the estimated Euryglossinae crown age of approximately 48 Mya (Almeida et al. in press), and yet the majority of the 900 Hylaeinae species arose from a rapid diversification 30 Mya. Combined, these inferences suggest that diversification rate rather than clade age is the driving factor behind the level of species richness within clades and across the colletids as a whole.

Diversification and range expansion are often associated with climatic features, whereby environmental change can lead to habitat modification and consequently open up previously unavailable niches. Of course climatic patterns are complex and it is almost impossible to extract detailed climatic data from millions of years ago (e.g. Zachos et al. 2001). There is, however, evidence for the expansion of bee-friendly temperate environments from the Eocene onwards that slowly began to replace the pre-existing tropical biota (Zachos et al. 2001; Byrne et al. 2011). Although generic diversity in bees is greatest in the Neotropics (Wcislo and Cane 1996), bees are more speciose throughout the world in xeric and temperate regions where a ‘Mediterranean’ climate dominates

(Michener 1979; Wcislo and Cane 1996). This is also true of the Australian colletids, which are more generally distributed in xeric regions of Australia (Michener 1965, 2007), and for which diversification has occurred since the Eocene – the same period in which Australia began its transition into a dry continent (Byrne et al. 2011). Significantly for the Australian colletids, Myrtaceae became one of the dominant angiosperms since the early Miocene (Byrne et al. 2011).

Although it has not been formally addressed, the assumption that bee diversification is influenced by the historical diversification in angiosperms has been a common proposal (Michener 1979; Grimaldi 1999; Danforth et al. 2006). This assumption provides a good basis for comparing divergence times for bees given that pollens, nectars, and oils from angiosperms are major resources for bees. Linked to this is the specialisation of some bees on certain angiosperms, and this in turn has been associated with climatic features. Xeric and Mediterranean climates apparently support higher numbers of oligolectic species, whereas generalist species predominate in the tropics (Wcislo and Cane 1996). Furthermore, it has been noted that higher rates of speciation have been detected in oligolectic bee species (Danforth et al. 2003), which suggests there is a correlation between oligolecty and species richness in drier climates (Danforth et al. 2003). Using the definitions of provisioning behaviour in bees summarised by Wcislo and Cane (1996), most colletid genera contain both oligolectic and polylectic species, but with polylecty more frequent in Hylaeinae than in Euryglossinae (Michener 1965; Exley 1967, 1968, 1972, 1974, 1975, 1976, 1978, 2001; Houston 1975, 1981).

Aridity has been increasing in Australia since around 20 Mya (Crisp et al. 2004) and we therefore might also expect increased speciation in bees around the same time (Austin et al. 2004; Byrne et al. 2008). Consequently (possibly as a result of aridification), the same period saw the rapid radiation of the Australian schlerophyll biome (consisting of *Banksia*, eucalypts, and legumes) 25-10 Mya (Crisp et al. 2004; Austin et al. 2004; Byrne et al. 2008). In contrast, taxa restricted to the aseasonal-wet biome of Australia did not show signs of such radiation, or alternatively were depleted by extinction (Crisp et al. 2004; Byrne et al. 2011). From the known distributions of euryglossine genera, the same patterns can be detected – whereby species richness is far less in those genera restricted to the eastern seaboard of Australia (Michener 2007). The high level of

diversification and species richness in Hylaeinae is mainly restricted to a single genus, *Hylaeus* (Michener 2007; Almeida et al. in press; Chapter 1). In comparing Australian Hylaeinae and Euryglossinae only, the latter are more speciose (Michener 2007), and this may be the influence of differences in life-history traits.

The high extinction ratio in Euryglossinae (Chapter 2) might be identified from lineages that now have few species and are restricted to the mesic biome of Australia, which has undergone contraction and fragmentation over the last 30 Mya (Byrne et al. 2011). Michener's (2007) taxonomic treatment of Euryglossinae (Group B) includes six described genera that all contain fewer than five species but are not included in the Euryglossinae phylogeny (Chapter 2). These species are relatively large in size and are largely restricted to the cool temperate to semi-tropical areas of eastern Australia with limited distribution in the xeric regions of the country where the more species-rich genera dominate (Michener 2007). The two monospecific genera, *Sericogaster* and *Stenohesma* are large for Euryglossinae (7-11mm, and 7.5-10mm long) and are restricted to the less arid eastern regions of Australia (New South Wales to northern Qld) (Michener 2007). Similarly, there are only two described *Heterohesma* species, both of which are also relatively large (~10mm) and restricted to mountainous regions, ranging from northern NSW to Tasmania (Michener 2007).

These larger species may exhibit size-related distribution patterns associated with latitude and climate, consistent with other Hymenoptera that are generally larger in tropical than in temperate regions (Schoener and Janzen 1968). Such patterns would suggest that species from arid regions of Australia are generally smaller than those found in mesic environmetns. The mesic areas of Australia constitute the eastern coast and south-western corner of Australia, and represent biomes that have contracted and fragmented into the marginal fringes of Australia since drying of the continent began in the Eocene (Byrne et al. 2011). The arid regions have expanded to make up around 70 per cent of the continent in the past 30-20 Mya (Crisp et al. 2004). Aridification might have been an important process leading to the minimal representation of larger sized Euryglossinae that are restricted to the eastern seaboard (Michener 2007), either in the form of an extended extinction event as the wetter regions contracted, or by extremely low levels of diversification. With extensive aridification in Australia occurring in the

last 5 My (Crisp et al. 2004), extinction and/or retreat may have occurred relatively recently.

Evidence that climatic changes (e.g. warming and cooling, aridification) have not been even across all regions of the World (Zachos et al. 2001) indicates that other factors are also relevant to the diversification of bees 50-30 Mya. This work highlights that opportunities for long distance dispersal might be less of a predictor for successful colonisation than the availability of novel environments (Chapter 1 and Chapter 2). When assessing biogeographical relationships, it is not only important to consider past tectonic movement, geological disjunction, and the extent of physical barriers, but also to consider any 'ecological' barriers (Wiens and Donaghue 2004; Goldberg and Lande 2007). Where we don't have closely related fossilised lineages, we can only use the distribution of the extant focal taxa as a conservative estimate of the distribution of extinct lineages. In other words, the resources used and the biomes inhabited by extant taxa can be an indication of the 'preferences' of related ancestral lineages. It is therefore important to assess whether the proposed dispersal route had adequate resources, and climatic features for the adaptation of the colonising lineage (Wiens and Donoghue 2004). Successful diversification and colonisation relies partly on life-history traits that are suitable for adaptation. The potential dispersal of stem-nesting bee species within their nests and their ability to more readily maintain nesting opportunities in high humidity and rainfall might be a fundamental reason as to the relative scarcity of soil nesting species in the tropics (Michener 2007). In contrast, a dry climate where vegetation is sparser would enable more nesting opportunities for soil-nesters, and the availability of suitable stem or wood would be more limiting.

## LIMITATIONS OF THIS WORK AND FUTURE DIRECTIONS

This study provides the largest molecular phylogenies of Australian representatives of the euryglossines and hylaeines, and is the first to examine diversification in these groups from a detailed molecular framework. Future work would benefit greatly from increased sampling, as well as more sequence data. The inclusion of monospecific hylaeine and euryglossine genera and subgenera in future dating work might provide further clues to extrinsic and intrinsic influences on diversification. These monospecies have presumably diverged from a most recent common ancestor tens of millions of years

ago, and it is likely that they are remnant species of what were once more diverse lineages that have gone extinct. Establishing minimum stem ages for those monospecies might provide some clues as to whether extinction events are linked to major climatic shifts, and/or periods of extensive change to the Australian biome. This would provide further understanding of bee and invertebrate diversity and diversification in Australia, as well as potentially enhance our understanding of biogeographical processes.

This research only touches the surface on the possible factors that may have influenced such diversity of some groups in Australia. It is important to continue gaining insight. Importantly, studies such as Paini and Roberts' (2005) that found fecundity was reduced in a *Hylaeus* species where honey bees were present, are needed to directly assess the impact of human actions on bee diversity. Understanding the patterns of the natural world has become more relevant as biodiversity is becoming more valued. It is known that cyclical events have affected the Earth's climate and the patterns in the biota are sometimes linked to these events. No doubt climatic events and geological processes have played a significant role in shaping the bee fauna.

**LITERATURE CITED**

Almeida. E.A. 2008. Revised species checklist of the Paracolletinae (Hymenoptera, Colletidae) of the Australian Region, with the description of new taxa. Zootaxa. 1891:1–24.

Almeida, E.A.B., and B.N. Danforth. 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50:290-309.

Almeida, E.A.B., M.R. Pie, S.G. Brady, and B.N. Danforth. In press. Biogeography and diversification of Colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. J. Biogeogr.

Ascher, J.S., and M.S. Engel. 2006. On the availability of family-group names based on *Scrapter* (Hymenoptera: Colletidae). Entomol. News. 117:117-119.

Austin, A.D., D.K. Yeates, G.Cassis, M.J. Fletcher, J. La Salle, J.F. Lawrence, P.B. McQuillan, L.A. Mound, D.J Bickel, P.J. Gullan, D.F Hales, and G.S. Taylor. 2004. Insects 'Down Under' – diversity, endemism and evolution of the Australian insect fauna: examples from select orders.
Aust. J. Entomol. 43:216–234.

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

Byrne, M., D.A. Steane, L. Joseph, D.K. Yeates, G.J. Jordan, D. Crayn, K. Aplin, D.J. Cantrill, L.G. Cook, M.D. Crisp, J.S. Keogh, J. Melville, C. Moritz, N. Porch, J.M.K. Sniderman, P. Sunnucks, and P.H. Weston. 2011. Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. J. Biogeogr. doi:10.1111/j.1365-2699.2011.02535.x.

Chenoweth, L.B., and M.P. Schwarz. 2011. Biogeographical origins and diversification of the exoneurine allodapine bees of Australia (Hymenoptera, Apidae). J. Biogeogr. 38:no. doi: 10.1111/j.1365-2699.2011.02488.x.

Crisp, M., L. Cook, and D. Steane. 2004. Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? Philos. T. Roy. Soc. B. 359:1551–1571.

Crisp MD, and L.G. Cook. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. Evolution. 63(9):2257-2265.

Danforth, B.N., S. Ji, and L.J. Ballard. 2003. Gene flow and population structure in an oligolectic desert bee, *Macrotera* (*Macroteropsis*) *portalis* (Hymenoptera: Andrenidae). J. Kan. Entomolog. Soc. 76(2):221-235.

Danforth, B.N., S.G. Brady, S.D. Sipes, and A. Pearson. 2004. Single-copy nuclear genes recover cretaceous-age divergences in bees. Syst. Biol. 53(2):309-326.

Danforth B.N, S. Sipes, J. Fang, and S.G. Brady. 2006. The history of early bee diversification based on five genes plus morphology. P. Natl. Acad. Sci. USA. 103:15118-15123.

Davies, G.B.P., and D.J. Brothers. 2006. Morphology of *Scrapter* (Hymenoptera: Anthophila: Colletidae), with description of three new species and taxonomic status of five Cockerell taxa. Afr. Invertebr. 47:135-183.

Eardley, C., and R. Urban. 2010. Catalogue of Afrotropical bees (Hymenoptera: Apoidea: Apiformes). Zootaxa. 2455:1-548.

Exley, E.M. 1967. Revision of the genus *Euryglossula* Michener (Apoidea: Colletidae). Aust. J. Zool. 16:203-217.

Exley, E.M. 1968. Revision of the genus *Brachyhesma* Michener. (Apoidea: Colletidae). Aust. J. Zool. 16:167-201.

Exley, E.M. 1972. Revision of the genus *Pachyprosopis* Perkins (Apoidea: Colletidae). Aust. J. Zool. Suppl. Ser. No. 10:1-43.

Exley, E.M. 1974. Revision of the subgenus *Callohesma* Michener (Apoidea: Colletidae). Aust. J. Zool. Suppl. Ser. No. 26:1-58.

Exley, E.M. 1975. Revision of the genus *Hyphesma* Michener (Apoidea: Colletidae). Aust. J. Zool. 23: 277-291.

Exley, E.M. 1976. Revision of the subgenus *Euryglossa* Smith (Apoidea: Colletidae: Euryglossinae). Aust. J. Zool. Suppl. Ser. No. 41: 1-72.

Exley, E.M. 1978. *Chaetohesma* - a new genus of Australian bees (Apoidea: Colletidae: Euryglossinae) Aust. J. Zool. 26:373-97.

Exley, E.M. 2001. The *walkeriana* species-group of *Euhesma* Michener (Hymenoptera: Colletidae: Euryglossinae). Aust. J. Entomol. 40(2):102–112.

Goldberg, E.E., and R. Lande. 2007. Species' borders and dispersal barriers. Am. Nat. 170(2):297-304.

Grimaldi, D., 1999. The co-radiations of pollinating insects and angiosperms in the Cretaceous. Ann. Miss. Bot. Gdns. 86(2):373-406.

Hines, H.M. 2008. Historical biogeography, divergence times, and diversification patterns of Bumble bees. Syst. Biol. 57(1):58–75.

Houston, T.F. 1975. A revision of the Australian Hylaeine bees (Hymenoptera : Colletidae). Aust. J. Zool. Suppl. Ser. No. 36:1-135.

Houston, T.F. 1981. A revision of the Australian Hylaeine bees (Hymenoptera : Colletidae) II. Aust. J. Zool. Suppl. Ser. No. 80:1-128.

McGinley, R.J. 1981 Systematics of the Colletidae based on mature larvae with phenetic analysis of apoid larvae (Hymenoptera: Apoidea). University of California Publications in Entomology.

McPeek, M.A., and J.M. Brown. 2007. Clade age and not diversification rate explain species richness among animal taxa. Am. Nat. 169:E97-106.

Michener, C.D. 1965. A classification of the bees of the Australian and South Pacific regions. Bull. Am. Mus. Nat. Hist. 130:1-362.

Michener, C.D. 1979. The biogeography of the bees. Ann. Mo. Bot. Gard.66(3):277-347.

Michener, C.D. 2007. The Bees of the World, 2nd edn. Johns Hopkins University Press, Baltimore, MD.

Mullen, S.P., W.K. Savage, N. Wahlberg, and K.R. Willmott. 2011. Rapid diversification and not clade age explains high diversity in neotropical *Adelpha* butterflies. Proc. R. Soc. B. 278:1777-1785.

Paini, D.R., and J.D. Roberts. 2005. Commercial honey bees (*Apis mellifera*) reduce the fecundity of an Australian native bee (*Hylaeus alcyoneus*). Biol. Conserv. 123:103-112.

Rasmussen, C., and S.A. Cameron. 2010. Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. Biol. J. Linn. Soc. 99:206–232.

Rehan, S.M., T.W. Chapman, A.I. Craigie, M.H. Richards, S.J.B. Cooper, and M.P. Schwarz. 2010. Molecular phylogeny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. Mol. Phylogen. Evol. 55(3):1042-1054.

Schaefer, H., and S. Renner. 2008. Phylogeny of the oil bee tribe Ctenoplectrini (Hymenoptera: Anthophila) based on mitochondrial and nuclear data: evidence for early Eocene divergence and repeated out-of-Africa dispersal. Mol. Phylogenet. Evol. 47(2):799-811.

Schoener, T.W., and D.H. Janzen. 1968. Notes on environmental determinants of tropical versus temperate insect size patterns. Am. Nat. 102:207-224.

Schwarz, M.P., S. Fuller, S.M. Tierney, and S.J.B Cooper. 2006. Molecular phylogenetics of the exoneurine allodapine bees reveal an ancient and puzzling dispersal from Africa to Australia. Syst. Biol. 55(1):31–45.

Wcislo, W.T., and J.H. Cane. 1996. Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies. Annu. Rev. Entomol. 41:257-286.

Wiens, J.J., and M.J. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol.19(12): 639-644.

Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms and aberrations in global climate 65 Ma to present. Science. 292:686-693.

# APPENDIX

While in the field in South Africa I collected a stem nest that contained a colony of an undescribed allodapine bee species. The following is a version of a manuscript describing these bees, which is currently in press:

**Kayaalp, P., Craigie, A.I., and Schwarz, M.P. Description of a new subgenus of *Allodapula* (Apidae: Allodapini) based on larval morphology and DNA sequence data. African Entomology. Volume 19.3.**

## *Abstract*

Allodapine bees provide unique insights into social evolution because their method of provisioning brood in communal chambers differs from nearly all other social insects except ants. This method of brood rearing in allodapines is associated with the evolution of remarkable larval morphologies. Interestingly, allodapine larval morphology seems to closely correspond to phylogeny, more so than adult morphology. Here we describe a very unusual form of larval morphology in an African allodapine bee, *Compsomelissa zaxantha*, and show that this corresponds with phylogenetic analyses of DNA sequence data indicating that this species is much more closely allied with the genus *Allodapula*, but represents an evolutionary clade that differs strongly from currently described subgenera of *Allodapula*. We erect a new subgenus, *Mhkuze*, to contain this species and discuss the implications of this species for understanding the radiation in allodapine larval morphology.

**Keywords:** Allodapine, larval morphology, *Allodapula*, *Compsomelissa*

## *Introduction*

The allodapine bees (Family Apidae, tribe Allodapini) have received attention because of their utility for understanding social evolution (review in Schwarz *et al*. 2007) and because of patterns in their historical biogeography (Fuller *et al*. 2005; Schwarz 2006). In both of these research areas it is important to have a sound understanding of phylogeny and systematics.

Very recently our understanding of allodapine systematics has been changing due to information from both molecular phylogenetics (Schwarz *et al*. 2007; Chenoweth *et al*. 2007) and the discovery of larvae in taxa where larval morphology was previously unknown (Chenoweth *et al*. 2008). Larval morphology in allodapines is more variable than in all other bee groups combined (Michener 1977). Larvae are reared in a communal nest of a hollowed pithy stem or twig, not much wider than the adult form (Schwarz 1988). The narrow and shared nest means that larvae are in constant contact with other larvae and with adults (Michener 2007) and this highly unusual nesting strategy is possibly the reason for such variation in larval morphology. Larval morphology shows very strong covariation with phylogeny and is often a clearer indicator of generic boundaries than is adult morphology, which tends to be conserved (Michener 1976, 1977, 2007).

Michener (2007) treated the genus *Compsomelissa* as containing two subgenera, *Compsomelissa s. s.* and *Halterapis*, whereas Michener (1975a) and Reyes (1998) accorded generic rank to the two subgenera, and generic rank has also been used in molecular phylogenetic studies by Schwarz *et al*. (2003, 2006) and Chenoweth *et al*. (2007, 2008). Here we accord separate generic status for *Compsomelissa* and *Halterapis* because of phylogenetic patterns that will be presented below. Studies of larval morphology in one species of *Compsomelissa*, *C. stigmoides*, and two species of *Halterapis*, *H. nigrinervis* and *H. angustula*, indicated that those species share very similar larval morphology, characterised by two small lateral projections on the second and third body segments, but without projections or tubercles on other body segments, with either a few short hairs or no hairs on the body and head, and with the body distinctly curled ventrally (Michener 1976). In both genera the male genitalia have a

long apically directed mesal projection arising from the ventroapical plate of the gonocoxite, and this is regarded as one of the diagnostic features of *Compsomelissa s. l.* (Michener 2007). *Compsomelissa zaxantha* is a largely yellow species with the second recurrent vein present – a vein that is absent in other species of *Compsomelissa* but present in *Halterapis*, *Allodapula*, *Allodape* and *Braunsapis* (Michener 2000).

Molecular phylogenetic studies of allodapine bees have consistently grouped *Halterapis* and *Compsomelissa* in a clade that also contains the African genus *Allodapula* and the rare Middle Eastern genus *Exoneuridia*, and we will subsequently refer to this combined clade as the '*Allodapula* group'. This clade is sister to the genera *Braunsapis* and *Allodape* which form the most speciose monophyletic group of allodapines, dominating the allodapine fauna of sub-Saharan Africa (Michener 1976) with *Braunsapis* the only allodapine element in Asia minor.

*Allodapula* comprises three subgenera, *Allodapula s.s.*, *Dalloapula*, and *Allodapulodes* (Michener 2007). Because it can be difficult to reliably assign species to some genera on the basis of female adult morphology alone (Michener 1975a), we will focus on descriptions of larvae as well as male genitalia in the following outline. The most basic division at the subgeneric level in *Allodapula* is based on the size of larval tubercles (Michener 1975b). Where the larval body form of *C. stigmoides* and *C. borneri* are curved like other allodapine larvae, and have only a few short setae, if any (Michener 1976; MP Schwarz, pers. obs.), the larvae of *Allodapula* are more or less straight and hairy or multituberculate (Michener 1975b). The mature larvae of *A. (Allodapulodes)* have no larval tubercles but many short setae. In *A. (Dalloapula)* there are three rows of lateral and ventro-lateral tubercles, with the lateral tubercles moderate in size and the others small, and the body is bare. *Allodapula (Allodapula)* larvae also have three rows of tubercles similar in form to those of *A. (Dalloapula)* but they are larger, and setae are absent on the maxillae and labia of *A. (Allodapula) ornaticeps*, whereas they are present in the other three *A. (Allodapula)* species for which data are available. All subgenera of *Allodapula* have larvae that are not ventrally curled, and this is associated with a peculiar positioning of larvae in the nest, where they are assembled in a venter-to-venter arrangement (Michener 1977).

Male genitalia in *Allodapula* are highly variable and common genitalic features are not as pronounced in this genus as in others (Michener 1977). Importantly, though, there is a small mesal projection arising from the ventroapical plate of the gonocoxite in *A. (Allodapulodes) xerica* and a spine like mesal projection in *A. (Allodapula) variegata*, and it seems possible that these are homologous to the long projections in *Compsomelissa* and *Halterapis*.

Recent fieldwork by us recovered several colonies of *C. zaxantha*, one of which contained several mature larvae, and we were able to recover DNA sequences from this material. Larval descriptions below are based on larvae from one nest but DNA sequences indicate that these larvae are of the same species as the adults from their nest. The larvae of *C. zaxantha* were, to the best of our knowledge, collected for the first time. The form of these larvae is very different to other known *Compsomelissa* larvae and more similar to *Allodapula* larvae. Here we argue that DNA sequence and larval morphological data indicate that this species is more closely allied to *A. (Allodapula)* and *A. (Dalloapula)*, and raise issues regarding the early evolution of this major group of African allodapines. We present these data here and transfer *C. zaxantha* to a new subgenus of *Allodapula, Mkhuze*. Our findings further add to the diversity of larval forms found within the allodapines and illustrate the remarkable radiation of larval morphology in early allodapine evolution.

## *Methods*

### FIELD SITES

Three nests with multiple individuals were collected from uMkhuze National Park, northern Kwazulu Natal, South Africa (27°37.720’S, 32°14.104’E, 99m) on 20 January 2008 and two nests with one individual each were collected from Sodwana Bay, South Africa in July 2008. One nest from the January collection contained three fourth-instar larvae.

### COLLECTION AND PRESERVATION TECHNIQUES.

Colonies were collected in dead pithy stems of an unidentified shrub and preserved by dropping adults and immatures directly into absolute ethanol. Extensive experience by

us in the study of allodapine larvae indicates that preservation in ethanol does not lead to changes in gross larval morphology compared to more traditional fixing in Kahle's solution, nor does ethanol lead to changes in form compared to living specimens, though we have not looked at differences at a small-scale tissue level.

**DNA SEQUENCING.**

*DNA extraction.* All DNA extractions were from mesosomal or metasomal sections of one ethanol-preserved adult and an entire larval specimen using the Puregene DNA extraction kit (Gentra Systems Inc., Minneapolis, MN). 10 µL of Proteinase-K (20 mg/ml) was added to the extraction prior to incubation at 55°C overnight. DNA pellets were resuspended in 50 µL of TE buffer and stored frozen.

*PCR primers.* Genes and relevant primers were chosen in order to provide overlap with previous allodapine studies. One nuclear and two mitochondrial gene regions were amplified and sequenced bi-directionally. The nuclear exon region from the F2 copy of elongation factor 1α (EF-1α F2; 585bp excluding the intron) was used. The mitochondrial regions were from the protein coding genes cytochrome *b* (Cyt *b*; 374 bp) and cytochrome oxidase I (CO1; 640 bp). The primers used for PCR amplification of the EF-1α F2 region were the F2 specific forward primer HaF2For1: 5'G GGY AAA GGW TCC TTC AAR TAT GC-3', (unpublished, B. Danforth, Cornell University), and the F2-specific reverse primer, F2-rev-1: 5'-AAT CAG CAG CAC CTT TAG GTG G-3', designed by Danforth *et al*. (1999) for halictid bees. The primers used for the amplification of the Cyt *b* region were, cb1: 5'-TAT GTA CTA CCA TGA GGA CAA ATA TC-3' designed by Y.C. Crozier (Latrobe University, Melbourne, Australia), and cb2: 5'-ATT ACA CCT CCT AAT TTA TTA GGA AT-3', designed by Schwarz *et al*. (2004). The primers used for the amplification of the COI region were LC01490: 5'-CCT TTT ATA ATT GGA GGA TTT GG-3' (Folmer *et al*., 1994), and M399: 5'-TCA TCT AAA AAC TTT AAT TCC TG-3' designed by S. Cooper (Schwarz *et al*. 2004).

*PCR conditions.* PCR reactions were carried out in 25 µl volumes, each containing 2.5 µl of DNA, 2.5 µl 10x HotMaster$^{TM}$ Taq Buffer (Eppendorf), 2 µl dNTP mix (Bioline) (2.5 mM of each dNTP), 1 µl of each primer (5 mM), 15.9 µl $H_2O$, and 0.1 µl

HotMaster™ Taq DNA polymerase (5 U/µl). Amplification involved a single enzyme activation step of 94ºC for 2 min, followed by 35 cycles of denaturation (94ºC for 1 min); annealing (48-50ºC for 1 min); and elongation (72ºC for 1:30 min); and a single final extension step at 72ºC for 6 min. PCR products were purified using a MultiScreen $PCR_{384}$ Filter Plate (Millipore). Sequencing reactions were carried out in 10 µl volumes using the amplification primers with BigDye® Terminator chemistry v3.1, and sequenced using an ABI 3730 capillary sequencer.

*Specimens included.* One female from the nest containing larvae was used for DNA sequencing to include in the analyses, GenBank accession numbers for each gene are as follows: COI [HM013827], Cyt *b* [HM013828], EF1α(F2) [HM013829].

Our analyses included 28 allodapine species from 13 genera which have been used in previous phylogenetic studies (Bull *et al*. 2003; Schwarz *et al*. 2003; Schwarz *et al*. 2004; Tierney 2004; Fuller *et al*. 2005; Schwarz *et al*. 2005; Chenoweth *et al*. 2007) and NCBI accession numbers can be found in the relevant papers. Two or three species from each genus were used, except in the case of *Exoneuridia* and *Compsomelissa s. s.*, for which there is sequence data for only one species of each. We included three species of Ceratinini (the sister tribe to the allodapines) as the outgroup. Forward and reverse sequences were compared for each gene fragment, and sequences were edited manually using SeqEd version 1.03 (Applied Biosystems). The intron region of EF-1α was excluded from the phylogenetic analyses because large sections of the sequence were not alignable.

**MOLECULAR PHYLOGENETIC ANALYSES.**

To explore phylogenetic relationships we used a Bayesian MCMC method implemented in MrBayes version 3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Previous studies on allodapine molecular phylogenetics indicate that an extreme AT bias in the third codon position of mitochondrial genes is best accommodated by partitioning the first two codon positions together as a single unit and the third separately (Schwarz et al. 2004). We partitioned sequence data into four groups: $1^{st}$ + $2^{nd}$ codon positions and $3^{rd}$ codon positions for COI+cytb and, the nuclear gene EF-1α. We

used default MrBayes priors, which are generally uninformative, because we did not hold any prior beliefs about likely topology, branch lengths, or the dynamics of evolutionary change among the gene partitions. In this sense, we used an 'objective' Bayesian approach (Berger 2006). We used a GTR + I + Γ model (general time-reversible model with a proportion of invariant sites and where site-specific rates varied according to a gamma distribution) for each partition, which is the least restrictive model for substitutional change in MrBayes (Huelsenbeck & Ronquist 2001). Model parameters for each partition were unlinked between partitions, so that each partition was free to evolve with different substitution rates, different base-pair composition, and with different proportions of invariant sites. The model was therefore highly parameterised.

The Bayesian analysis was run for six million generations, sampling every 500$^{th}$ generation. Each run comprised two parallel MCMC processes, and each of these comprised three hot chains, where proposed changes to model parameters in each iteration could be large, and one cold chain, where proposed changes were small. The burnin was set at 1.5 million generations, well after convergence of parallel runs as assessed by the average standard deviation of split frequencies using the MrBayes run-time diagnostic. Analyses were run three times to check for convergence in topology and branch lengths, and no discernable differences were found between replicate runs.

## *Results*

**PHYLOGENETIC ANALYSES.**

The results of the Bayesian analysis are summarised in Figure 1 where posterior probability (PP) values for nodes less than 100 are given. The Bayesian analysis suggests moderate support (91 PP) for the *Allodapula* group being sister clade to the clade comprising *Allodape* and *Braunsapis*. Support for the divergence between *Compsomelissa borneri* and *Exoneuridia hakkariensis* is poor (87 PP), but support for their position as a sister clade to the other species in the *Allodapula*-group is high (100 PP). Support for the divergence of *Halterapis* from *Allodapula+Dalloapula+Compsomelissa* is low (79 PP). The two *Halterapis* species, both key out as *H. nigrinervis* in Michener's (1975a) revision, but differ in having either

red or black metasomas. As pointed out by Chenoweth *et al*. (2007), sequence divergence between the two colour morphs is indicative of species differentiation and this is consistent with Michener's (1975a) observation that *Halterapis* likely contains a number of cryptic species. The two *A. (Dalloapula)* species, *C. zaxantha* and the three *A. (Allodapula)* species form a sister clade to the rest of the *Allodapula*-group (95 PP).

Importantly, support for *C. zaxantha* being included in the *Allodapula* clade is very high (100 PP), and this concords with larval morphology. This means that neither larval morphology or molecular phylogenetics indicate that *C. zaxantha* belongs in the clade containing *C. borneri* or in the clade containing the two *Halterapis* species. Furthermore, any taxonomic arrangement that grouped *H. nigrinervis* and *C. borneri* within a single genus would be paraphyletic unless *A. (Allodapula)* and *A. (Dalloapula)* were also included in this genus.

**SUBGENERIC DIAGNOSIS.**

Sub-genus *Mkhuze* Kayaalp & Schwarz, subgen. n.

Type species: *Allodape zaxantha* Cockerell, 1934, p. 241, Type Female, BM, from Weenen, Natal, South Africa.

*Allodapula zaxantha* (Michener, 1966, p. 573)

*Compsomelissa zaxantha* (Michener, 1975a, p. 225).

*Etymology: Mkhuze* refers to the collection locality of the nest containing larvae, uMkhuze Reserve in South Africa.

Larval diagnosis:
*Head.* Antennae long and attenuated, blunt seta arising above each antennae, another pair of blunt setae arising at top of head. Another series of blunt setae arising from side of head and directed laterally. Clypeus laterally extended from base, forming conspicuous flaps. Labrum swollen and strongly projected dorso-anteriorly, projecting

anteriorly beyond closed mandible. Labium also swollen, trilobed with a large papilla arising from each lateral lobe and two from the medial lobe. Each labial lobe with irregularly shaped and blunt apically-directed setae.

*Body.* Body linear as in other *Allodapula*. First body segment with two appendages or projections. A series of enormous, tapering but irregularly shaped non-bifurcate appendages, beginning at second body segment and becoming smaller posteriorly. Two further latero-dorsal series of appendages, similarly shaped to the large series. Blunt setae prominent but sparse on body, largely restricted to bases of lateral appendages and also forming weak transverse lines at the middle part of anterior-most segments.

Adult diagnosis:
Both sexes: Body, head and mouthparts yellow with brown markings on vertex, extending to orbits, and on dorsal surfaces of metasoma and thorax and, in female, ventral parts of 6$^{th}$ tergum. Legs entirely yellow. Second recurrent vein present, stigma clear with brown margins, less than costal margin of marginal cell. Second submarginal cell half as long as first.
Female: Length greater than 5mm. Malar space greater than half width of scape. 6$^{th}$ tergum flattened with ventral parts separated from dorsal surface by a sharp angularity and almost emarginate.
Male: Penis valve lacking pegs. Ventroapical mesal projection of gonocoxite less than length of penis valve.

Distribution: Northeastern South Africa, adjacent to Mozambique. SOUTH AFRICA: Mafeking; Mfongosi; Ndumu Reserve; Tugela Crossing; 16 km. north of Ubombo; Weenen; uMkhuze Reserve; Sodwana Bay. MOZAMBIQUE: Lourenco Marques.

*Allodapula (Mkhuze) zaxantha*
Description of mature larva.

*Head* (Fig 2A). Antennae long and attenuated, about 2/3rds length of mandible; one blunt forwardly-directed seta arising above each antennae and as long as about half antennal length, another pair of blunt setae arising at top of head and about same length

as antennae. Another six, irregularly shaped blunt setae arising from side of head and directed laterally. Clypeus laterally extended from base, forming conspicuous flaps. Labrum swollen and strongly projected dorso-anteriorly, projecting anteriorly beyond closed mandible, with two large antero-lateral papillae. Labium also swollen, trilobed with a large papilla arising from each lateral lobe and another two papillae arising from the middle lobe. Each labial lobe with two apically directed setae on the ventral surfaces.

*Body* (Figs 2B-2C). Body not curved as in *Halterapis*, but straight as in *Allodapula*. First body segment with two appendages or projections, the first arising at base of head, bluntly rounded and laterally directed, the second appendages tapering irregularly and situated ventrally. A series of enormous, tapering but irregularly shaped appendages, beginning at second body segment and becoming smaller posteriorly. First two appendages as long as about 2/3 width of body. Two further latero-dorsal series of appendages, similarly shaped to the large series but 2/3 to 1/3 length of corresponding lateral appendages. No appendages are bifurcate as in *Allodapula s. s.* Mid-dorsal part of each segment with a sharply raised protuberance, becoming smaller posteriorly. Blunt setae sparse on body, largely restricted to bases of lateral appendages and also forming weak transverse lines at the mid-dorsal part of anterior-most segments.

*Adults.* Female type described by Cockerell (1934, p. 241) and male and female specimens described by Michener (1975a). We compared our specimens with the published descriptions, paying particular attention to the male genitalic characters, and found no differences between our specimens and the published descriptions.

Michener (2007) provides a key to the genera of the African allodapines based on adult morphology and mature larval morphology. *Allodapula zaxantha* will correctly key out as *Allodapula s. l.* using the larval key. On the basis of adult morphology, *A. zaxantha* keys out to *Compsomelissa*, but can be distinguished from this genus by the presence of the 2$^{nd}$ recurrent vein, having the length of the stigma less than the costal margin of the marginal cell, and the length of the 2$^{nd}$ submarginal cell more than 1/3 length of the 1$^{st}$. Males are described for only two species of *Compsomelissa*, and one of these species, *C. stigma*, is only known from a single specimen which is not associated with female

specimens. Males of *A. zaxantha* are readily distinguished from males of *C. stigmoides* by lacking pegs on the penis valve and on the basis of adult morphology described above.

## *Discussion*

Revision of taxonomic structure is an ongoing process, and will be influenced by new sources of data and improved forms of analysis. Our molecular phylogenetic and larval morphology data presented here indicate a very different systematic position for *Allodapula zaxantha* than was apparent from adult morphology alone. This is similar to a recent study (Chenoweth *et al*. 2008) of Madagascan bees previously placed in *Halterapis*, where larval morphology and molecular data indicated very different systematic relationships to other genera. Interestingly, for both our study and that of Chenoweth *et al*. (2008), both molecular and larval morphology conflicted with relationships inferred from adult morphology. Schwarz *et al*. (2003) found a similar problem when exploring basal relationships in Allodapini more generally, and suggested that some of these problems were due to coding of male genitalic characters and ability to recognize homoplasies.

Placement of *A. zaxantha* within *Compsomelissa* was largely based on three characters: (i) extensive yellow coloration in adults; (ii) a long mesal apical projection of the gonocoxite; and (iii) gonostylus fused to a ventral apical projection of the gonocoxite. We argue here that it is possible that these characters may not be reliable indicators of phylogeny. Extensive yellow coloration is common in many diverse allodapine lineages, including *Exoneurella eremophila* in Australia (Houston 1977), numerous *Allodape* species in central-west Africa (Michener 1975a), and *Exoneuridia oriola* in the southern Arabian Peninsula (Terzo 1999). A long ventroapical mesal projection of the gonocoxite is common to *Halterapis*, *Compsomelissa* and *A. zaxantha*, but the projection in *A. zaxantha* is much shorter than in *Halterapis* and *Compsomelissa* species, and a very short projection is present in at least *Allodapula variegata* and *A. xerica* (Michener 1975a). Lastly, the ventral apical projection of the gonocoxite in *A. zaxantha* is basally fused to the gonostylus, and gives rise to a bifurcate apically acute claw-like

structure, very similar to *Allodapula* species, and less similar to the gently-rounded bidentate structure in *C. stigmoides*, and very different from the bi-lamellate structures in *Halterapis* species. Taken together, these considerations of male genetical structure do not seem to provide strong evidence for placement of *A. zaxantha* in *Compsomelissa* as opposed to *Allodapula*.

Our phylogenetic analyses based on DNA sequence data show that the species previously classified as *C. zaxantha* is nested within the subgenera assigned to *Allodapula*. Furthermore, the extensive appendages and straight, rather than curved, body form of the larvae are more similar to other known *Allodapula* species and very unlike larvae of *Compsomelissa* or *Halterapis*. Because *A. zaxantha* is phylogenetically placed between the subgenera *Dalloapula* and *Allodapula s. s*., it could be placed within the latter or accorded separate subgeneric status. We choose the latter because the adult and larval morphology is so strikingly different from *Allodapula s. s.*, indeed to the point where it had previously been regarded as belonging to *Compsomelissa*.

Lastly, our study reveals a form of larval morphology in *A. zaxantha* that is extreme in its elaboration of appendages, even when considering the enormous range in allodapine larval morphology. It seems that Michener's (1977) observation that variation in allodapine larval morphology was greater than for all other bees combined was certainly not exaggerated. This raises the question of why larval morphology is so highly variable among allodapine clades when adult morphology is mostly conservative. Explaining this contrast in adult and larval morphological variation is likely to provide important insights into the factors that have driven diversification of bees.

**ACKNOWLEDGEMENTS**

This work was funded by an Australian Research Council grant awarded to M.P. Schwarz, and a South African National Research Foundation grant (GUN 61497) to S. van Noort of the Iziko South African Museum, Cape Town, whose assistance in the field is also greatly appreciated.

**LITERATURE CITED**

Berger, J. 2006. The case for objective Bayesian analysis. Bayes. Anal. 1(3): 385-402.

Bull, N.J., M.P. Schwarz, and S.J.B. Cooper. 2003. Phylogenetic divergence of the Australian allodapine bees (Hymenoptera: Apidae). Mol. Phylogen. Evol. 27(2): 212-222.

Chenoweth, L.B., S.M. Tierney, J.A. Smith, S.J.B. Cooper, and M.P. Schwarz. 2007. Social complexity in bees is not sufficient to explain lack of reversions to solitary living over long time scales. BMC Evol. Biol. 7:246.

Chenoweth, L.B., S. Fuller, S.M. Tierney, Y.C. Park, and M.P. Schwarz. 2008. *Hasinamelissa*: a new genus of allodapine bee from Madagascar revealed by larval morphology and DNA sequence data. Syst. Entomol. 33: 700-710.

Cockerell, T.D.A. 1934. Some new or little-known South African bees of the genus *Allodape* in the British Museum. Ann. Mag. Nat. Hist. (ser. 10) 14: 220-242.

Danforth, B.N., H. Saquet, and L. Packer. 1999. Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae) based on parsimony and likelihood analyses of nuclear EF-1$\alpha$ sequence data. Mol. Phylogen. Evol. 13(3):605-618.

Folmer, O., M. Black, W. Hoeth, R. Lutz, and R. Vrijenhoek. R. 1994. DNA primers for amplification of cytochrome oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3:294-299.

Fuller, S., S.M. Tierney, and M.P. Schwarz. 2005. Phylogenetics of the allodapine bee genus *Braunsapis*: implications for across sea dispersal of bees. J. Biogeogr. 32(12): 2135-2144.

Houston, T.F. 1977. Nesting biology of three allodapine bees in the subgenus

*Exoneurella* Michener (Hymenoptera: Anthophoridae). T. Roy. Soc. South Aust. 101:99-113.

Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

Michener, C.D. 1966. A new genus of creatinine bee from the Red Sea area. J. Kansas Entomol. Soc.. 39:572-576.

Michener, C.D. 1975a. A taxonomic study of African allodapine bees (Hymenoptera, Anthophoridae, Ceratinini). B. Am. Mus. Nat. Hist. 155:67-240.

Michener, C.D. 1975b. Larvae of African allodapine bees. 3. The genera *Allodapula* and *Eucondylops*. J. Entomol. Soc. S. Afr. 38(2): 243-250.

Michener, C.D. 1976. Larvae of African allodapine bees. 4. *Halterapis, Compsomelissa, Macrogalea,* and a key to African genera. J. Entomol. Soc. S. Afr. 39(1): 33-37.

Michener, C.D. 1977. Discordant evolution and the classification of allodapine bees. Syst. Zool. 26(1): 32-56.

Michener, C.D. 2000. The Bees of the World. 1st Edition. The Johns Hopkins University Press, Baltimore.

Michener, C.D. 2007. The Bees of the World. 2nd Edition. The Johns Hopkins University Press, Baltimore.

Reyes, S.G. 1998. A cladistic analysis of the bee tribe Allodapini. Philipp. Entomol. 12: 55-84.

Ronquist, F., and J.P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetics inference under mixed models. Bioinformatics 19:1572-1574.

Schwarz, M.P. 1988. Intra-specific mutualism and kin-association of cofoundresses in allodapine bees (Hymenoptera Anthrophoridae). Monit. Zool. Ital. 22(3): 245-254.

Schwarz, M.P., N.J. Bull, and S.J.B. Cooper. 2003. The molecular phylogenetics of allodapine bees, with implications for the evolution of sociality and progressive rearing. Syst. Biol. 52: 1-14.

Schwarz, M.P., S.M. Tierney, N.J. Bull, and S.J.B Cooper. 2004.Molecular phylogenetics of the allodapine bee genus *Braunsapis*: A-T bias and heterogeneous substitution parameters. Mol. Phylogen. Evol. 32:110-122.

Schwarz, M.P., S.M. Tierney, J. Zammit, and S. Fuller. 2005. Social and nesting biology of a Malagasy species of *Halterapis*: implications for understanding social evolution in the allodapine bees. Ann. Entomol. Soc. Am. 98:126-133.

Schwarz, M.P., S. Fuller, S.M. Tierney, and S.J.B Cooper. 2006. Molecular phylogenetics of the exoneurine allodapine bees reveal an ancient and puzzling dispersal from Africa to Australia. Syst. Biol. 55(1):31-45.

Schwarz, M.P., M.H. Richards, and B.N. Danforth. 2007. Changing paradigms in insect social evolution: new insights from halictine and allodapine bees. Ann. Rev. Entomol. 52:127-150.

Terzo, M. 1999. Révision du genre Exoneuridia Cockerell, 1911. Belg. J. Entomol. 1:137-152.

Tierney, S.M. 2004. Evolution of African allodapine bees. PhD Thesis, Flinders University of South Australia, Adelaide.

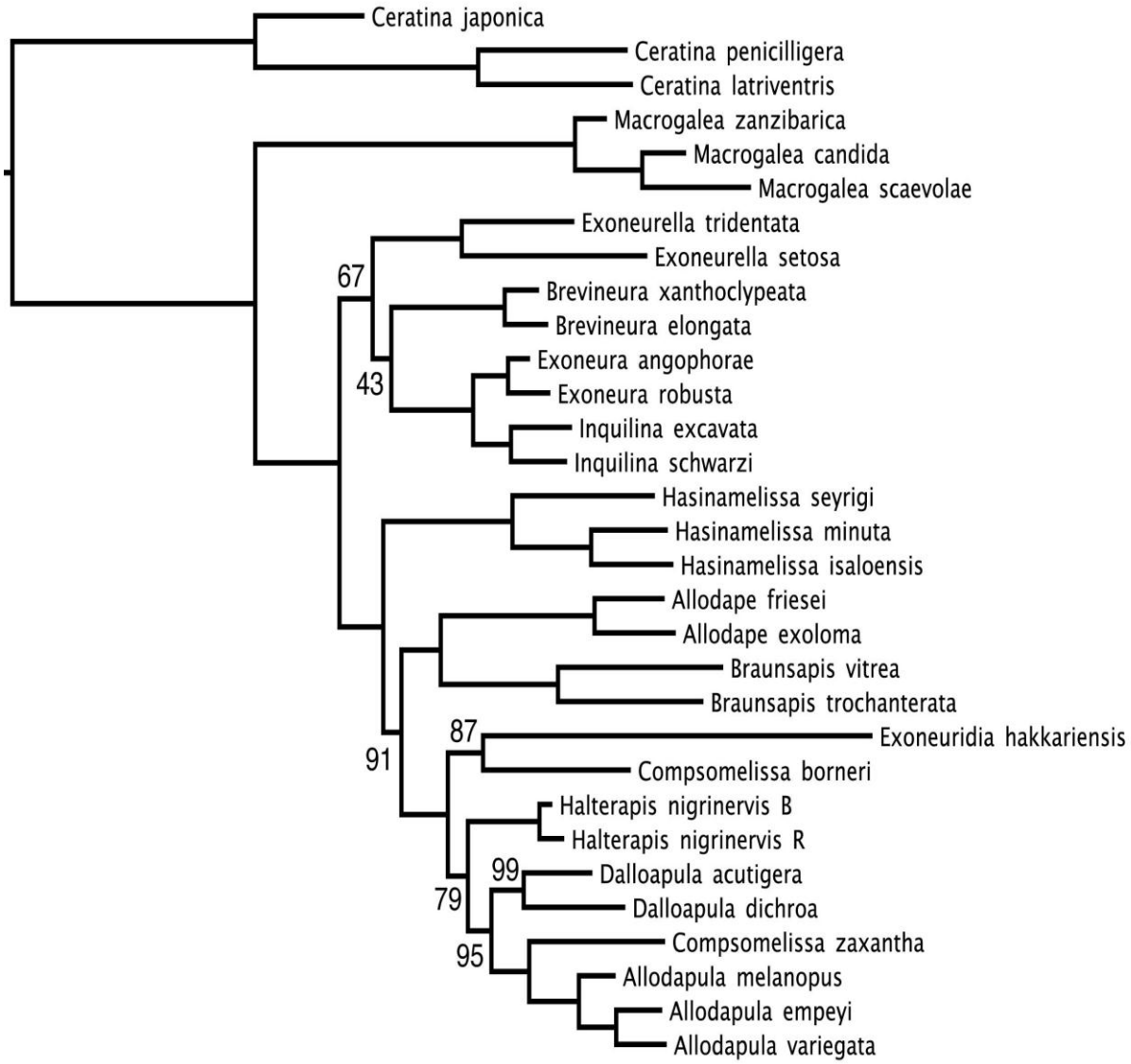
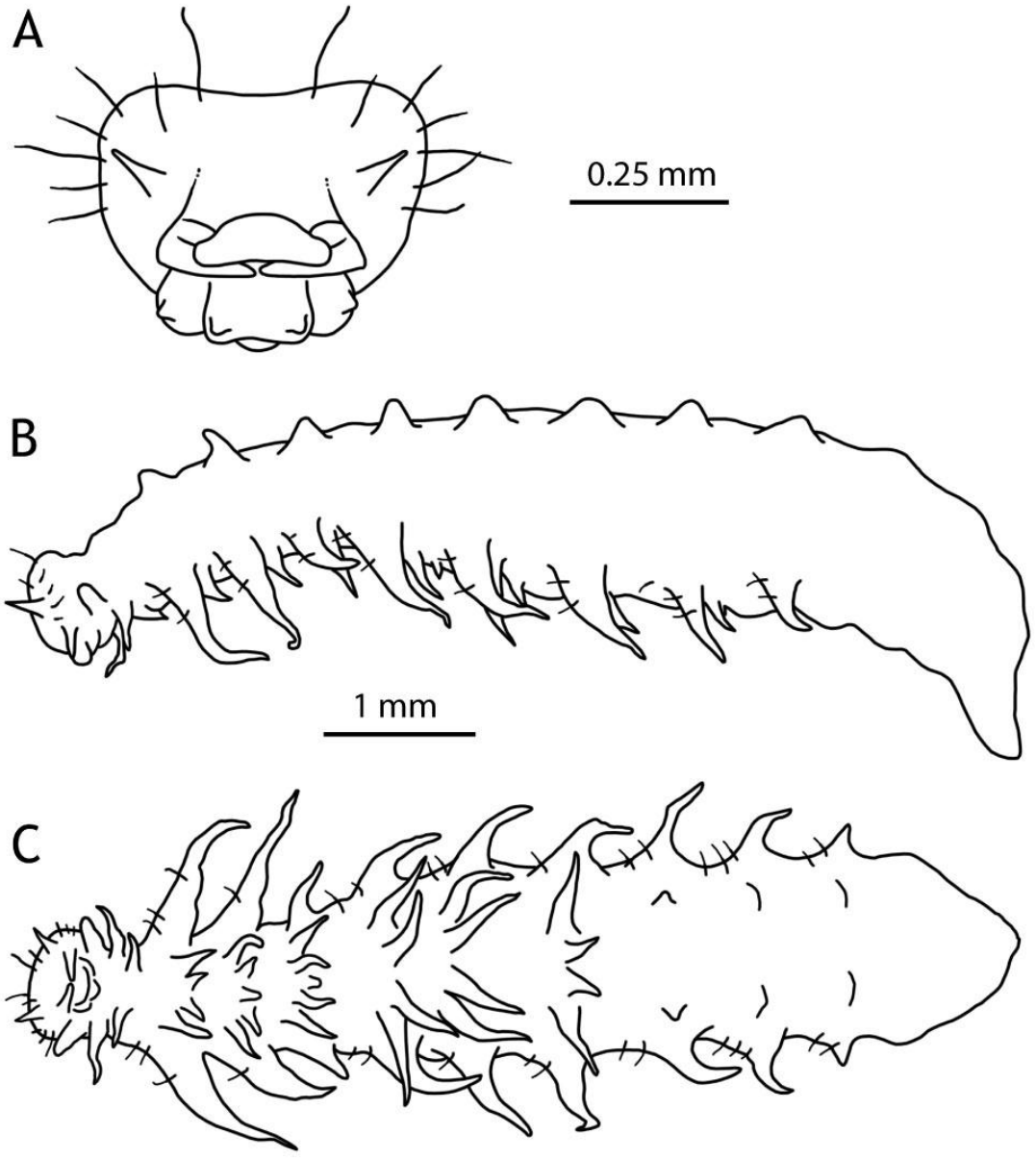



Fig. 1. Bayesian consensus phylogram with posterior probability (PP) values given for each node with less than 100 per cent support. The analysis suggests that the species previously classified as *Compsomelissa zaxantha* is nested within the subgenera assigned to *Allodapula*.




Figs. 2A-C. *Allodapula (Mkhuze) zaxantha* 4th instar larva. 1A, Head, frontal view; 1B, Body, lateral view; 1C, Body, ventral view.