The development of nuclear protein coding genes as phylogenetic markers in bark and ambrosia beetles (Coleoptera: Curculionidae)

Dario Pistone

Thesis for the Degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2018

UNIVERSITY OF BERGEN

The development of nuclear protein coding genes as phylogenetic markers in bark and ambrosia beetles (Coleoptera: Curculionidae)

Dario Pistone

Thesis for the Degree of Philosophiae Doctor (PhD)
at the University of Bergen

2018

Date of defence: 20.03.2018

Year: 2018

Title: The development of nuclear protein coding genes as phylogenetic markers in bark and ambrosia beetles (Coleoptera: Curculionidae)

Name: Dario Pistone

Print: Skipnes Kommunikasjon / University of Bergen

The development of nuclear protein coding genes as phylogenetic markers in bark and ambrosia beetles (Coleoptera: Curculionidae)

Dissertation for the degree of philosophiae doctor (PhD)

University of Bergen, Norway - 2017

This thesis consists of a synthesis and three individual papers. The experimental PhD research activity was developed during three years (2012-2015).

Supervisor:

Associate Professor Bjarte Henry Jordal

Co-supervisor:

Professor Lawrence Kirkendall

"Coherence in insect systematics will ultimately depend on having a large database of homologous data. Currently, exploring a variety of markers is advantageous. However, direct comparisons among them should be requisite. It is fantasy to think that we will eventually fill in the gaps through random sequencing and that our studies will grow together and eventually fuse. It is necessary that we consciously work toward this goal."

Caterino et al. 2000

TABLE OF CONTENTS

Preface

In the city of Uppsala, in Sweden, there was a bar with a thick wooden counter. The upper part where beverages were placed was polished and smooth, but on large part of the lateral surface, towards the guests, the wood was covered by a fine system of dark-stained grooves. The drawings were the remnants of the galleries created by several broods of bark beetles with a polygynous mating system. The main pattern of the tunnels consisted of a series of single slightly curved central lines with a star-like configuration, from which departed perpendicularly other tightly packed small galleries created by larvae. Considering that in Sweden, the Norwegian spruce (*Picea abies*) is the main tree species used for wood construction and internal design, I could tentatively guess the beetle species. The tunneling system might have been created by *Pityogenes chalcographus*, the six toothed spruce bark beetle, one of the most common bark beetles in Europe, infesting mainly *P. abies* and other members of Pinaceae. At that time, a broader knowledge of these beetles would have probably helped me to support better the theory that the engraved drawings were made by beetles and disprove the antagonist less fascinating theory, suggested by other clients, who considered the drawings as the result of the work of a time-wasting human artist.

In those days, I was considering the possibility to move to Norway for working at the University Museum of Bergen on a PhD project in molecular systematics of bark and ambrosia beetle. At that time, my knowledge of this fascinating beetle group was limited to a few species causing coordinate tree-killing over large areas of North American forests or consistent economic loss in coffee production worldwide. However, reading the extensive literature on different biological and ecological aspects of Scolytinae and Platypodinae beetles, I was growing a mesmerizing interest for these insects.

As you can guess, I decided to accept this PhD project which largely consisted in a search for additional molecular markers in Scolytinae, but also in Platypodinae and other weevils; a real challenge, considering the scarce number of markers developed for beetle phylogenetics in the last decades, despite the large interest in the systematics of this extremely diversified insect order (Coleoptera). When I started working on the project in 2012, the rapid progresses of the Next Generation Sequencing technology required to be carefully considered as well. Indeed, the advantages, in terms of costs and benefits, of mining nuclear genes using a genome-scale Sanger sequencing approach were not so obvious. On the other hand, the majority of NGS options were still not tuned to deal efficiently with routine phylogeny matters. Nevertheless, genome assembly-free methods were emerging as a preferential choice in the systematics

field. It is worth mentioning that the first papers on genomic ultra-conserved elements and highly conserved anchor regions of genomes (also referred to as anchored hybrid enrichment) were published in the 2012. At the time of writing (2017), it seems clear that few protein coding genes could still represent a valuable alternative in phylogenetic studies, although NGS based data are more and more ready to claim their hegemony in insect molecular systematics.

This PhD research project was designed with the aim to remedy the lack of ready-to-use nuclear markers in 'classic' beetle phylogenetics. Here, I report on the multiple level optimization procedure to select nuclear protein coding genes and test their phylogenetic utility within the weevil superfamily Curculionoidea, with emphasis on the wood boring lineages grouped in the subfamilies Scolytinae and Platypodinae (family Curculionidae). According to the results obtained in this study and to the experience acquired during this research project, I can vouch for and encourage researchers in beetle systematics to test the 16 selected markers or some of the less characterized 18 markers.

Abstract

Bark and ambrosia beetles are grouped into two different subfamilies (Scolytinae and Platypodinae), within the superfamily Curculionoidea (more than 60,000 described species). These insects constitute a large part (circa 8,000 species) of the advanced weevils (family Curculionidae). The subfamilies Scolytinae and Platypodinae were traditionally considered closely related, due to anatomical affinities and similar ecological behavior of their members. Indeed, these beetles present morphological modifications which allow them to spend almost all the entire lifecycle in tunnels constructed mainly in dead wood, though showing extraordinary variation in ecological adaptations to thrive in different niches. Despite the large interest focused on Scolytinae and Platypodinae which include economically important pests, the evolutionary history of these two groups is largely unclear (especially for Scolytinae) as well as their precise placement in the weevil tree. Due to the high number of species and the lack of molecular markers, obtaining high phylogenetic resolution for framing the timing and ecological circumstances under which each of the largest radiations originated still represent a great challenge. Even though this is one of the beetle taxa where more efforts were concentrated in collecting molecular data, the low phylogenetic resolution at deeper nodes has not been markedly improved adding only a few protein coding genes. Morphological characters in larvae, pupae and adults together with few mitochondrial and nuclear molecular markers clarified only a limited number of important evolutionary issues in Scolytinae, while Platypodinae phylogeny is significantly more resolved.

This PhD research project focused on the development and standardization of nuclear protein coding genes as phylogenetic markers for weevils. One hundred genes were tentatively PCR amplified and sequenced with 'classic' Sanger technology for different species of Scolytinae, Platypodinae and other weevils. After this preliminary screening, unsuitable genes were discarded and the most promising ones were further tested in their capacity to recover monophyly for well-supported tribes. A total of sixteen protein coding genes emerged as first choice markers for reconstructing the phylogeny of Scolytinae, a subset of them were tested in other members of Curculionoidea and additional eighteen markers were shown to present different degree of utility for shallow level phylogenetics in weevils (e.g. tribes, genera and at population level).

In the first section of this study (**paper I**), the procedure of development and optimization of each selected marker was described. Information on the intron length and number were reported for all the sixteen nuclear genes. Problems of unspecific amplification or primer

failure in particular taxa were also emphasized. Finally, the novel genes were tested under different methods of phylogeny reconstruction (NJ, maximum parsimony and Bayesian inference), for their ability to recover well-established relationships among closely related species. The integrative knowledge provided by comparison among the different analyses allowed ranking the selected markers according to their utility for higher level phylogenetics in Scolytinae.

In the second section of this work, a total of 18 markers (five previously defined and 13 out of the 16 developed in this study) were used to reconstruct the phylogeny of the subfamily Scolytinae applying two different phylogenetic methods: maximum parsimony and Bayesian inference. Among the major findings, the tribe Scolytini and the genus *Microborus* were confirmed to be early divergent lineages. However, their placement at the base of the Scolytinae tree or close to other subfamilies in the weevil tree remains to be clarified. The tribe Hypoborini was recovered as the sister lineage to a group containing the species-rich Dryocoetini and Ipini. Better resolution was achieved within different tribes and the placement of a few enigmatic species was unambiguously solved, but the relationships among older tribes remained elusive (**paper II**).

Finally, ten genes (five developed in this study) were used to reconstruct the phylogeny of different weevil families and subfamilies (**paper III**). All the analyses placed the subfamily Platypodinae as the sister lineage to Dryophthorinae with high node support, therefore more distantly related to Scolytinae.

1 Introduction

1.1 The changing landscape of insect molecular systematics

Comparative analyses of homologous morphological structures represented for a long time the only strategy for resolving the insect tree of life (Wille, 1960, Crampton, 1938, Hennig, 1969). However, the utility of morphological characters can be limited in species-rich taxa where sometimes convergent evolution has masked true indicators of relationships (Haas and Kukalova-Peck, 2001). Therefore, in morphology-based phylogenies, obtaining a high degree of confidence for relationships among and especially within hyper-diverse insect families and subfamilies can be difficult. Molecular systematics emerged as a more promising tool to disentangle such relationships and achieve high resolution at different taxonomic levels (Mardulyn and Whitfield, 1999, Field et al., 1988, Cognato and Sperling, 2000, Russo et al., 1995). Whereas decades of morphological studies in insects have intensely explored a vast assortment of different character systems, we are only in an early phase of exploring genomic regions at large scale for insect phylogenetics. So far, the majority of PCR and sequencing based phylogenetic studies have largely relied on RNA sequences from both mitochondrial and nuclear genomes, and a few protein coding genes. Resolution of insect relationships has not been without problems, with results often highly influenced by the choice of markers, in addition to suboptimal use of search algorithms and evolutionary models. Since a large amount of molecular data can be required to resolve ancient divergences in highly variable groups, the selection of an adequate number of markers maintains a key importance, but this requisite was not always easy to fulfill. For the majority of the insect orders, only a limited number or protein coding genes were tested and developed as molecular markers for reconstructing phylogenetic relationships.

From a couple of decades ago, ribosomal, mitochondrial and to some extent nuclear protein coding genes started to be explored for insect systematics, with a slow but continuous development of phylogenetically informative gene fragments (Baker et al., 2001, Fang et al., 1997, Friedlander et al., 1998, Friedlander et al., 1992, Pelandakis et al., 1991). Certainly, a large amount of works on mitochondrial genes (Weirauch and Munro, 2009, Maekawa et al., 2001, Liu and Beckenbach, 1992, Howland and Hewitt, 1995, Scheffer and Wiegmann, 2000) and ribosomal structural RNAs (Weller et al., 1992, Shull et al., 2001, Whiting et al., 1997, Carmean et al., 1992) paved the way of insect phylogenetics, but more recent studies which included nuclear protein coding genes provided further advancements in the field (Gibson et

al., 2011, Winkler et al., 2015, Wild and Maddison, 2008, Sahoo et al., 2016, Wahlberg and Wheat, 2008, Regier et al., 2013, Wahlberg et al., 2016).

Ribosomal and mitochondrial genomic regions are still widely used in insect systematics, but only a limited number of nuclear protein coding genes (e.g. *EF-1α*, *CAD*, *ArgK*, *PEPCK* and *wingless*), can be considered as 'common' markers used across several insect orders (Maddison, 2012, Jordal and Cognato, 2012, Kim and Farrell, 2015, Riedel et al., 2016, Jordal et al., 2011). Beside this limited number of well-characterized phylogenetic markers, other nuclear genes were implemented in different insect taxa, often following independent routes for marker selection and optimization (Cruaud et al., 2013, Senatore et al., 2014). Noticeably, several groups in Lepidoptera and Hymenoptera received much more attention in this perspective and as a consequence, these orders currently have the highest number of standardized protein coding genes (more than 20) that can be selected for phylogeny reconstruction at various ranks (Regier et al., 2013, Mutanen et al., 2010, Danforth et al., 2004, Danforth et al., 2013, Hedtke et al., 2013, Wahlberg et al., 2016).

More than fifteen years ago, Caterino et al. (2000) highlighted a tendency for lack of coordinated efforts among different research groups to define a set of common nuclear genes for insect systematics. The few 'standard' genes (e.g. mitochondrial genes, structural RNAs and *EF-1α*) were often amplified and sequenced with different primer pairs according to the PCR amplification and sequencing success in different groups (Caterino et al., 2000). As a result, protein coding genes for insect phylogenetics are currently organized in a sort of 'tower of Babel' of markers which makes it difficult or impossible to compare or to predict the phylogenetic utility of such genes in various insect orders. Different degrees of complexity in gene structure, intron pattern, hypervariable regions and/or presence of paralogous copies are other factors that seriously complicate routine sequence production across different taxa (Yenerall et al., 2011, Hardy, 2007). Therefore, marker-specific intricacy, together with technical problems such as inconsistent or unspecific PCR amplification might be likely reasons behind an irregular development of protein coding genes in insect phylogeny (Wahlberg and Wheat, 2008).

Although the lack of nuclear markers remains a considerable limit for studying the evolution of several insect groups, the growing number of sequenced genomes provides a good source for selecting novel genes. With the huge amount of molecular data currently available in the public databases both in terms of genomes and transcriptomes, the scarceness of standardized nuclear markers could be potentially overcome. Nevertheless, gene exploration, testing, and

phylogenetic utility evaluation is a long and complex procedure, often with results that are difficult to predict 'a priori' as demonstrated during this PhD research project.

It is generally accepted that confidence in phylogenetic reconstruction can be obtained only through analyses of a large amount of molecular data. However, phylogenetic studies which combine information from five or more protein coding genes (excluding Next Generation Sequencing based study) are not common in insects (Wiegmann et al., 2009, Winkler et al., 2015, Maddison, 2012). Sanger sequencing applied to phylogenetic studies started more than two decades ago, and since the early application of this technology in phylogenetics, increasing the amount of data has always been one of the major concerns. On the other hand, molecular studies which include few mitochondrial and nuclear genes continue to provide more resolution in several insect taxa (Baca et al., 2016, Vuataz et al., 2016).

Consequently, there is still no consensus on the optimal number of nuclear genes required for resolving relationships, especially among old and species rich insect lineages. Thus, even if the primary goal of entomologists working in insect systematics is still to increase the number of characters to obtain more robust phylogenetic inference, the question is: how many nuclear genes are necessary to resolve such relationships?

Different studies seem to suggest that a 'PCR based' molecular strategy should be based on 15-20 genes to solve phylogenies at family and subfamily level; even if such datasets still represent a small fraction of the entire genome, they are more resistant to large fluctuations in tree topology and node support that are otherwise observed with fewer markers (Rokas et al., 2003, Ruane et al., 2015, Rokas and Carroll, 2005, Edwards et al., 2007). On the other hand, the fact that few genes with strong phylogenetic signal can be more useful than quantitative information is lately getting strong support (see Shen et al., 2017). However, the number of genes required for achieving good resolution and node support is largely dependent on the rank at which a particular phylogeny is investigated, the number of taxa included and the age of diversification of the group. Remarkably, a relatively small multiple-gene nucleotide dataset (6 genes) was capable of recovering deep divergences among Holometabola orders with high node support (Wiegmann et al., 2009). Another study based only on three nuclear genes showed that they were sufficient to support monophyly of major insect lineages with robust node support (Sasaki et al., 2013). Furthermore, it has been demonstrated that a small gene dataset of four nuclear markers resolved Halictidae subfamilies relationships in bees (Danforth et al., 2004). Finally, the combination of ribosomal genes, *COI* and *CAD*, resolved relationships in the Diptera infraorder Bibionomorpha (Sevcik et al., 2016). Hence, the debate on the essentiality of large genetic data volumes in insect phylogenetics is far from settled.

Phylogenomics is rapidly changing the need for an elaborate and time-consuming selection of protein coding genes (Trautwein et al., 2012, Misof et al., 2014, Kawahara and Breinholt, 2014, Crampton-Platt et al., 2015). Next Generation Sequencing technology can provide large amount of data that enable higher level of phylogenetic resolution compared to phylogenies based on few genes obtained with Sanger sequencing technology. As an example, NGS mitogenomics is rapidly gaining insights into weevil phylogeny (Gillett et al., 2014, Haran et al., 2013). Anchored hybrid enrichment (AHE) targeting ultra-conserved elements (UCEs) is an NGS technique that uses oligonucleotide probes to capture conserved regions of the genome flanked by less conserved areas in order to acquire useful data for phylogenetic inference from a broad range of taxa. Once a probe kit is developed, such approach is superior to traditional PCR-based Sanger sequencing in terms of both the amount of genomic data that can be recovered and effective cost (Young et al., 2016, Haddad et al., 2017). Therefore, such genome assembly-free methods are becoming preferential choices in the systematics field, especially in large-scale phylogenetics projects. In addition, high-throughput NGS sequencing of genomes and transcriptomes allowed a cost-effective way for the rapid development of phylogenetic markers for later Sanger sequencing (Rutschmann et al., 2017). In fact, Rutschmann et al. (2017) offers an interesting and uncommon example of how NGS and Sanger sequencing can be combined in an effective way.

For the large majority of researchers working in insect systematics, Sanger sequencing based phylogenetics can still be a relatively fast, economic and informative strategy. As a final consideration, it was recently demonstrated that a limited number of genes (15-20) can generate a phylogenetic tree highly congruent (with similar node support for several clades) with UCE based analyses (Ruane et al., 2015, Blaimer et al., 2015).

1.2 Molecular markers in beetles

Hitherto, exploring nuclear protein coding genes, using a PCR and Sanger sequencing approach, with the ultimate goal to evaluate and select such kind of markers for beetle molecular systematics has not been an easy task. David Maddison, co-author of the research article ‘Evaluating nuclear protein-coding genes for phylogenetic utility in beetles’ (Wild and Maddison, 2008) summarized the main findings of their study with a short sentence: ‘Hey guys! New genes! (https://myrmecos.wordpress.com/2008/08/13/new-genes-for-studying-beetle-evolution-or-blogging-my-own-research/). This short slogan was more than sufficient to communicate the importance of such article to researchers working on beetle phylogeny.

With the development of lab protocols for PCR amplifying and sequencing eight nuclear genes in Coleoptera (three genes previously unused in beetles – five already in use), this study represents the most important and successful attempt of 'developing' new genes for beetle phylogenetics (24 genes were considered and tested in the genus *Bembidion* and in other beetle groups).

The limited availability of nuclear markers is a relatively common situation in several insect taxa, especially in Coleoptera. Application of molecular markers in phylogenetic studies of beetles has not yet lead to a deep understanding of the evolutionary history of this order. Indeed, phylogenetic resolution in some part of the beetle tree such as for the weevils is still relatively low. More in general, the relationships among the four suborders are currently debated and the phylogeny of the extremely species rich suborder Polyphaga remains incompletely resolved, with medium-low resolution for several families and subfamilies (Yuan et al., 2016, Zhang et al., 2016, Lawrence et al., 2011). Understanding the phylogeny of the main beetle taxa represents a great challenge in phylogenetics.

Nevertheless, important advancements were recently achieved, especially in those groups on which more attention was focused, with studies including a large number of species and several markers (Maddison, 2012, Gunter et al., 2014, Jordal, 2015, Kim and Farrell, 2015).

Although the number of protein coding genes for beetle phylogenetics has increased over the last years, studies which include multiple nuclear genes are uncommon (Maddison, 2012, Sota and Vogler, 2001, McKenna et al., 2015). Large-scale studies, where combinations of morphological and molecular data were used, are also relatively rare (Bernhard et al., 2009, Whiting et al., 1997).

The early phase of molecular systematics of Coleoptera was largely based on mitochondrial and ribosomal markers (Sikes and Venables, 2013, Maddison, 2012, Maddison et al., 2013, Maus et al., 2001). Two cytochrome oxidase genes (*COI* and *COII*) are among the most frequently used genes and they are useful mainly to resolve recent divergences (Cognato and Sperling, 2000, Dobler and Muller, 2000, Martinez-Navarro et al., 2005). In several studies, both mitochondrial and ribosomal genes were combined, or they were often used in combination with protein coding genes (Cryan et al., 2001, Maddison, 2012, Bernhard et al., 2009, Ahrens et al., 2011, Ruiz et al., 2010, Sequeira et al., 2000). However, only five nuclear genes can be considered frequently implemented phylogenetic markers across this hyper diverse order, and they are the same genes with large utility in other insect groups. The gene *wingless* is quite popular in beetle phylogenetics and it was included in studies on different taxonomic groups often in combination with mitochondrial and ribosomal genes (Kim and

Farrell, 2015, Maddison, 2012, Zhang and Zhou, 2013, Tarasov and Dimitrov, 2016). Schubert et al. (2000) suggest extreme caution when analyzing *wingless* sequences which might occur in multiple copies (at least three) in insects. Similar paralog-related problems can potentially affect a number of low copy genes. Widely used markers in beetle phylogenetics such as *elongation factor 1 α* (*EF-1α*) and *enolase* are present in multiple copies in insect genomes. Nevertheless, they have been proven to be suitable for elucidating relationships between weevil genera, either alone or in combination with mitochondrial and nuclear genes (Farrell et al., 2001, Sequeira and Farrell, 2001, Normark et al., 1999, Jordal, 2002). More specifically, paralogy for *EF-1α* genes - two copies in beetles (Jordal, 2002), bees (Danforth and Ji, 1998) and flies (Hovemann et al., 1988) – is not particularly problematic because the paralogs can be distinguished for the presence of copy-specific introns (this is also true for the *enolase* gene). *CAD* and *ArgK* were used in bark and ambrosia beetle phylogenetic studies, in carabids and in staphylinid beetles (Jordal and Cognato, 2012, Maddison, 2012, Song and Ahn, 2017).

Other genes were sporadically included in beetle phylogenetics: *topoisomerase I*, *Histone III*, *DDC*, *white*, *opsin*, *period*, *hunchback* and others (Tarasov and Dimitrov, 2016, Cameron and Mardulyn, 2003, Polak et al., 2016, Fang et al., 1997, Caterino et al., 2000, Regier et al., 1998, Danforth et al., 2003, Baker et al., 2001, Tatarenkov et al., 1999, Tanzler et al., 2014).

1.3 The phylogeny of the superfamily Curculionoidea

The most easily recognized characteristic of weevils is the presence of a long rostrum, though it can be reduced or absent in some lineages (e.g. Entiminae, Cossoninae, Scolytinae and Platypodinae). The rostrum represents a key innovation that has been implicated in the evolutionary success of this group (Davis, 2014). Apart from a striking and unparalleled diversification, weevils have a tremendous economic impact on worldwide agriculture, wood trade and vegetal food transport and storage (Mariño et al., 2017a, Correa et al., 2013, Fettig et al., 2007).

The weevils constitute one of the largest superfamilies (Curculionoidea) in the animal kingdom, and increasing phylogenetic resolution at different ranks continues to be a great challenge. High species diversity, limited lineage extinction and the simultaneous origin of some families and subfamilies are all plausible reasons for such complexity (Gillett et al., 2014, McKenna et al., 2015, Marvaldi et al., 2002, Jordal et al., 2011). The low number of highly-informative phylogenetic markers constitutes another complicating factor, at least for

the classic PCR based phylogenetics. Despite two decades long effort in collecting molecular data, the phylogeny of weevils is still debated at higher ranks and resolution at shallow level (among tribes and genera) is highly variable. Key advancements were obtained within some groups such as Scolytinae, Cryptorhynchinae and Platypodinae. However, large dataset, comprehending multiple molecular markers and high number of species, were still not enough to solve the majority of late-Cretaceous nodes (Jordal and Cognato, 2012, Riedel et al., 2016, Jordal et al., 2011). Platypodinae represents an exception, and a comparatively more resolved phylogeny at deeper nodes was obtained using only five markers (see Jordal, 2015).

The uncertainty regarding the placement and rank of several subfamilies was emphasized by recent works (Alonso-Zarazaga and Lyal, 1999, Oberprieler et al., 2007). Lately, morphological character based analyses, supported by increasingly larger amount of molecular data allowed moving towards a gradually unified classification (Marvaldi et al., 2002, Gillett et al., 2014, McKenna et al., 2009, Gunter et al., 2015). Among the early diverging weevil lineages, the beetles possessing straight antennae in the families Antribidae, Attelabidae, Caride, Brentidae, Belidae and Nemonychidae constitute a grade. The phylogeny of the advanced weevils (with geniculate antennae) in the family Curculionidae remains unclear. The basal positions are currently occupied by some of the broad-nosed lineages (Brachycerinae sensu latu), monocot-associated taxa (Dryophthorinae) and the contended Platypodinae, potentially sharing pedotectal male genitalia (orthocerous-type). The remaining subfamilies are classified in the Curculionidae *sensu stricto*, with pedal genitalia (gonatocerous-type), and represent a derived lineage (Kuschel, 1995, Thompson, 1992).

The phylogenetic placement and current classification of the wood boring lineages, especially Platypodinae and Scolytinae, remains one of the more problematic issues (see Kuschel, 1995, Kuschel et al., 2000, Wood, 1986, Wood and Bright, 1992, Jordal et al., 2014). While several studies clearly indicate a nested position of Scolytinae within a narrowly defined Curculionidae (sensu Alonso-Zarazaga and Lyal, 1999), the placement of Platypodinae is more uncertain. Two contrasting hypotheses place these two families as sister-groups, or alternatively as more distantly related clades (**Figure 1**). Scolytinae as sister to Platypodinae, within Curculionidae, is one of the hypotheses suggested by a large number of morphological and by some molecular studies (Kuschel, 1995, Farrell et al., 2001, Marvaldi and Morrone, 2000, Jordal et al., 2011, Alonso-Zarazaga and Lyal, 1999, Lawrence et al., 2011, Marvaldi et al., 2002, Crowson, 1955, May, 1993, Zherikhin and Gratshev, 1995, Lawrence and Newton, 1995). In addition, some authors suggested a close relationship between Scolytinae and Platypodinae to the subfamily Cossoninae (Marvaldi, 1997, Kuschel et al., 2000).

Occasionally, these subfamilies were also elevated at family rank outside all other Curculionidae without solid evidence (Bright, 2014, Morimoto and Kojima, 2003, Wood, 1986, Wood, 1993, Wood and Bright, 1992). Finally, more recent molecular studies based on larger data volumes support the fact that Scolytinae and Platypodinae, even though they are adapted to similar life styles, they might be more distantly related, with the letter being the sister group to Dryophthorinae (Gillett et al., 2014, Gunter et al., 2015, McKenna et al., 2009, Haran et al., 2013).

Morphological and ecological similarities among Platypodinae and the ambrosia beetles in the subfamily Scolytinae suggest a relationship between these subfamilies. Platypodinae features such as the elongated body shape, the long tarsal segment 1 (relative to tarsae 2-5) different male genitalia and different larval morphology point towards a case of convergent evolution driven by similar niche utilization.

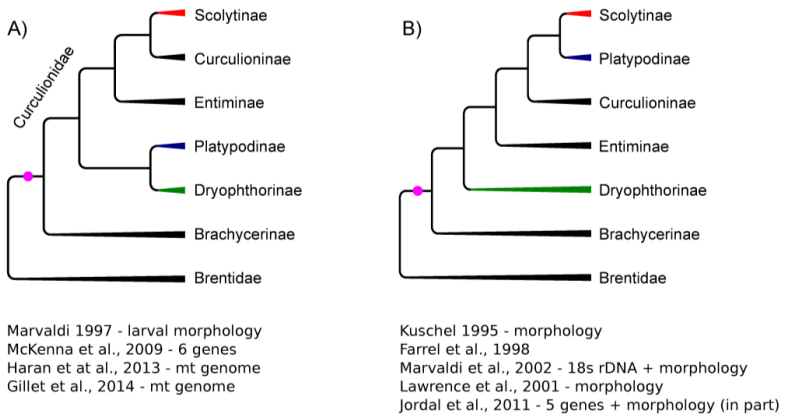


Figure 1 – The two main hypotheses regarding the possible relationships among Scolytinae and Platypodinae. A) Platypodinae as sister group to Dryophthorinae is mainly supported by larval morphology and molecular studies based on multiple genes or mithocondrial genomes; B) Scolytinae and Platypodinae as sister groups is supported by adult morphology and mainly ribosomal genes, sometimes associated with other markers.

1.4 Bark and ambrosia beetles

Bark and ambrosia beetles are an extraordinarily diverse group of insects which represent an interesting and ecologically variable model system for studying diversification processes (Jordal and Cognato, 2012, Gohli et al., 2017). These wood boring beetles are grouped in two of the most species-rich taxa within the advanced weevils. More than 6,000 described species currently belong to the subfamily Scolytinae and more than 1,500 species were described in the subfamily Platypodinae (McKenna et al., 2009, Kirkendall et al., 2015).

Outbreaks and damage to timber and to other forest products are typical for few species, but have nevertheless made these beetles known to a broader audience (Linnakoski et al., 2012). The evolution of various lifestyles, in general associated with decomposition of (mainly) dead plant material, originated from a phytophagous feeding behavior common in the large majority of weevils (Oberprieler et al., 2007).

The unstable classification of these two weevil subfamilies can still create some taxonomic confusion for non-experts due to the use of the same term 'ambrosia beetles' for taxa in two different subfamilies. The term denotes an ecological adaptation associated with the cultivation of fungal gardens for feeding, but does not define a taxonomic group. 'Ambrosia feeding' indicates a highly specialized and irreversible feeding mode which evolved independently in Platypodinae and, independently, in a minimum of ten lineages in Scolytinae. Obligate fungus feeding has also evolved in a single clade of ants (subfamily Myrmicinae - Attini tribe) and in termites (subfamily Macrotermitinae). In all three insect groups, the fungi are transported, actively cultivated and propagated as clones with some degrees of similarity (Farrell et al., 2001, Mueller and Gerardo, 2002).

Ambrosia beetles show strict mycophagy, derived from an obligate mutualistic symbiosis where fungi serve as the only food source for larvae and adults. Ambrosia fungi are mainly species in the orders Ophiostomatales, Microascales, and occasionally Hypocreales (Ascomycota) – but fungi in Basidiomycota can be involved in the symbiosis as well, e.g. *Flavodon ambrosius*, which was lately found to dominate the symbiotic community in certain *Ambrosiodmus* species (Kostovcik et al., 2015, Li et al., 2017). In different species, the relationship between the beetle and the community of fungi can range from stringent to promiscuous (Hulcr and Stelinski, 2017). In general, a diet based entirely on fungi allows ambrosia beetles to be ecological generalists in host plant selection compared to true bark beetles. Since these beetles do not feed directly on the host tissues, they can more easily attack and successfully colonize different plant species (Hulcr et al., 2007). Therefore, some ambrosia beetles are considered important pests and others may easily turn into invasive species when accidentally introduced to new areas (Carrillo et al., 2016, Rassati et al., 2016a, Jordal, 2002, Rassati et al., 2016b). These beetles are not able to survive and develop on a fungus-free diet composed only of plant tissue (Kok et al., 1970, Beaver, 1989).

The most striking morphological characteristic of ambrosia beetles is the presence of mycangia, structures which can be located in different part of the insect cuticula (e.g. mesonotum, mandibles, coxae) and form pockets to protect and transport fungal spores (Six,

2012, Mayers et al., 2015). Ambrosia beetles actively transport symbiotic fungi to new hosts, inoculate them into the colonized trees where these fungi are actively cultivated.

True bark beetle species are also somehow associated with fungi, but are not forming obligate symbioses, even though relatively advanced mycangia can be found (Six et al., 2003, Beaver, 1989). Sometimes wood boring beetles, irrespective of the level of symbiosis, can benefit from a transient association with phyto-pathogenic or other fungi which can weaken the host plant, overcome chemical and mechanical defenses and facilitate beetle colonization as well (Persson et al., 2009, Miller et al., 2016). Bark beetles are mainly phloem feeders that live and develop within the cambium layer of secondary phloem (a relatively richer food source) just under the outer bark of trees. In these beetles, fungi can provide nutrients (e.g. nitrogen and sterol) supplementing deficient compounds in an unbalanced diet. Fungi were also shown to have positive fitness effects in these beetles, with higher offspring survival rates when ophiostomatoid fungi were intermixed with phloem (Six and Paine, 1996, Ayres et al., 2000, Six et al., 2003). The two partners in the symbiotic relationships display a wide-range of associations, from entirely mutualistic to merely commensal and from facultative to obligate. The importance of fungal microbes for bark and ambrosia beetles is clear, nonetheless, we still have a poor understanding of the evolutionary processes that shape most of these interactions.

1.5 The subfamily Scolytinae

The subfamily Scolytinae is currently divided into 26 tribes. Only a minority of these tribes, as defined by Wood (1986), was monophyletic in recent molecular studies (Jordal and Cognato, 2012, Jordal and Kaidel, 2016, Jordal et al., 2011). Well-defined subgroups were recognized and to a certain extent correlated with hosts preference, feeding and mating strategies (Kirkendall et al., 2015). Several tribes are suspected to be paraphyletic (Jordal and Cognato, 2012).

Bark and ambrosia beetles are highly adapted to a life in tunnels (**Figure 2**). They have a quite small (0.1-12 mm) cylindrical and compact body with robust appendages. Legs often present robust spine for securing to the wood substrate, the rostrum is markedly reduced and the eyes are flat and elongated (Hulcr et al., 2015). All these characteristics confer advantages to thrive in concealed plant niches such as deep inside wood, under the bark, but also in other plant parts (e.g. seeds, petioles, root and fruits).

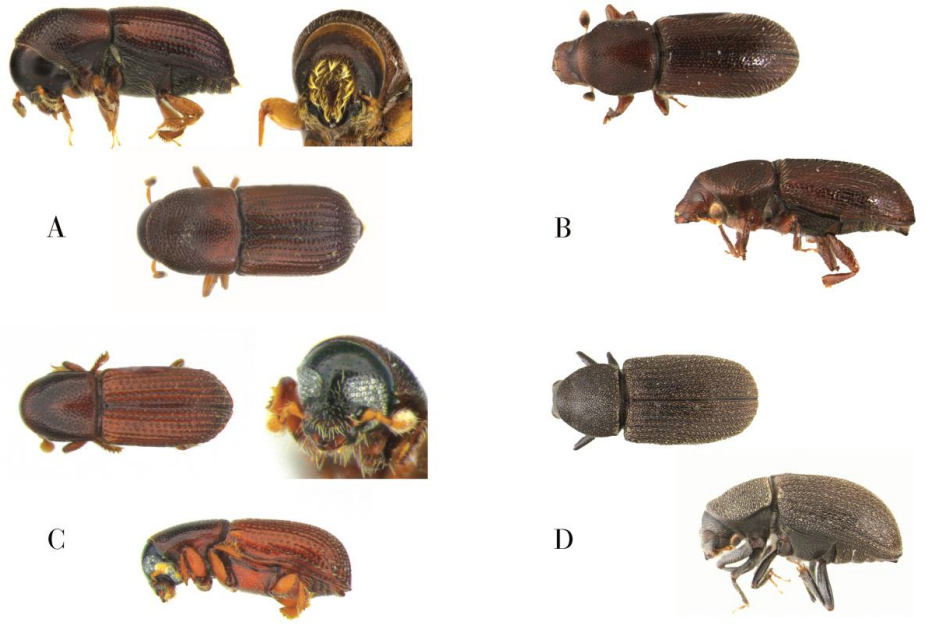


Figure 2 – Morphological diversity in Scolytinae: A) *Scolytodes pelicipennis*; B) *Dolurgocleptes malgassicus*; C) *Microborus brevisetosus*; D) *Dolurgocleptes punctifer*.

Only a limited number of species in the subfamily Scolytinae are capable of attacking and killing leaving trees. These species can be extremely invasive when introduced to non-native areas such as well-known pests with a vast economic impact (e.g. *Hypothenemus hampei*, *Euwallacea* spp.). Scolytinae have advanced abilities for detecting plant-produced (host) compounds and pheromones for mate location or aggregation which have important implications for coordinated tree killing as well.

Certain scolytine beetles are also vectors of plant-pathogenic fungi that can have considerable impacts on timber industry and agriculture (Hulcr and Dunn, 2011). Bark and especially ambrosia beetle associations with fungi extend far beyond a simple vector role (section 1.5).

The symbiosis with multicellular fungi delineates one of the two different feeding modes (ancestral and derived) which can be recognized within the subfamily. Fungal symbiosis not only opened up a whole range of new ecological opportunities for bark beetles, it also made the foundation for the development of more advanced interactions – from collective feeding to complex division of labour - among individuals. Parental care towards larvae might have been

promoted by aggregation close to the food source (fungi) that also keep siblings together for efficient mating (Kirkendall, 1997). Spending the entire lifecycles in such restricted and hidden niches, was probably another key factor in the evolution of such sub-social behavior. As one of most unusual mating systems found in Scolytinae, regular inbreeding by sibling mating (and haplodiploidy) is the most successful in terms of extant species. The evolutionary transition from ordinary outbreeding to regular inbreeding in these lineages is not known in detail. Bark and ambrosia beetles also exhibit a wide array of other complex mating strategies and genetic systems which probably flourished after the colonization of concealed niches (Kirkendall, 1983, Kirkendall, 1997).

1.6 The subfamily Platypodinae

Platypodinae are ambrosia beetles commonly referred to as 'pinhole borers' with more than 1,400 species grouped into 34 genera in two tribes, Platypodini and Tesserocerini (Wood, 1993, Wood and Bright, 1992, Jordal, 2015). With the exclusion of *Schedlarius* and *Mecopelmus* ,the core Platypodinae are characterized by a strict association with symbiotic fungi which allow several species to have a relatively large host range, including different plant families (Hulcr et al., 2007, Hulcr and Dunn, 2011). All core Platypodinae are monogamous. The male initiates the gallery excavation, mates with a single female and remains with her during brood development (Kirkendall, 1983). With an origin of the ambrosia symbiosis more than 80 Ma, Platypodinae is likely the oldest known group of fungus-cultivating insects. Pinhole borers are restricted to tropical or subtropical regions and only a minority of species has been able to colonize temperate areas. Furthermore, almost all genera in this subfamily have a distribution restricted to a single continent, demonstrating a high degree of endemism. Only a few species in the genera *Euplatypus*, *Megaplatypus*, and *Crossotarsus* are distributed more widely, possibly by recent introductions into new areas due to wood trade (Kirkendall and Faccoli, 2010). Although the group is about the same age as Scolytinae, is possible to reconstruct their evolutionary history and biogeography with limited molecular data (Jordal, 2015).

2 Aims of this study

This PhD research project describes the procedure of selection, optimization and standardization of novel protein coding genes for phylogenetics of Scolytinae and related beetles. Primers for amplification of selected fragments of nuclear genes were tested in different weevil species. Good quality sequences were aligned and each successful marker was tested for reconstructing relationships across species, weevil families and subfamilies.
The main goals of this project were:

a) to increase the number of nuclear protein coding genes available for beetle phylogenetics;
b) to increase resolution and node support in the phylogeny of the family Scolytinae using newly selected markers;
c) to confirm the utility of these markers in other weevil families and subfamilies, focusing on the placement of Platypodinae in the weevil tree and testing hypotheses on its sister group.

The main approach was mining of nuclear protein coding genes and characterizing these markers in terms of copy number, intron borders and phylogenetic signal. Therefore, one of the most innovative aspects of this study was a more restrictive approach to phylogenomics, focusing on a limited number of more properly characterized gene sequences. Scolytinae represent one of the few beetle groups were molecular markers were developed regularly, often adapting them from other insect taxa. Genes previously optimized, includes *COI* (mtDNA), *28S rRNA*, *EF-1α*, *ArgK*, *CAD* and *Enolase* (Jordal et al., 2011). Recent optimization efforts have focused on other genes previously used in insect phylogenetics such as *Histone H3* and *Polymerase II* which proved problematic in terms of paralogous copies (Jordal, 2007). Other genes were also considered by the same author: *NaK*, *TPI*, *PEPCK*, *wingless*, *gadh*, *RpS5*, *aats*, *IDH*, *LWR*, and *ddc*. With the exclusion of a few markers (e.g. *TPI*, *NaK*, *aats* and *IDH*), the nuclear genes developed in this study were not previously developed and used for Scolytinae (and beetles) phylogenetics. The final goal of this PhD research project was to define novel nuclear markers for Scolytinae and other weevils. We aimed at obtaining higher resolution for deeper nodes in Scolytinae and to clarify relationships at tribal levels and the placement of some particularly enigmatic species. Furthermore, our goal was to investigate the utility of these markers in non-Scolytinae beetles in an attempt of defining the sister group of Platypodinae.

3 Materials and Methods

3.1 Taxon selection

Most of the species selected for this study were previously collected by Bjarte Jordal, Lawrence Kirkendall and other members of the staff at the University of Bergen (Norway) and some specimens were donated by collaborators (J. Hulcr, A. Cognato, M. Knizek). All specimens used in this project derived from previous field collections in USA and Canada, Mexico, Costa Rica, Guyana, Argentina, Scandinavian countries, Russia Far East, Morocco, Sierra Leone, Ghana, Cameroon, Uganda, Tanzania, South Africa, China, Laos, Thailand, Malaysia including Borneo, Papua New Guinea and Australia. Colleagues donated some important specimens from field collections in New Zealand (M. Knizek), New Caledonia (R. Mecke), Western Russia, Ukraine (M. Mandelshtam), Malawi (S. Roth) and Bolivia (A. Petrov). Some species were collected during field work in Madagascar (Ranomafana National Park) at the beginning of this PhD project (September – October 2012).

Eight beetle species belonging to seven tribes within the subfamily Scolytinae and one species of Platypodinae were initially selected for testing primers and sequencing genes. Additionally, 18 selected species (ten Scolytinae species plus seven species belonging to four different curculionid subfamilies: Platypodinae (3), Molytinae (2), Cossoninae (2) and Lixinae (1), and one individual representing the family Brentidae) were used to test the recovery of 'known' phylogenetic relationships among closely related species. These 26 species were selected for preliminary primer testing and evaluation of PCR and sequencing success, to test the capacity of these markers to reconstruct phylogenetic relationships within Scolytinae tribes and to test the correct amplification of the same targeted genes in other weevils (**paper I**). For investigating weevil phylogeny, 72 species were added to the original 26 (**paper II**), while a total of 186 species were used for reconstructing the large scale phylogeny of Scolytinae (**paper III**).

3.2 DNA extraction, PCR and Sanger sequencing

DNA was extracted from individual specimens using DNeasy Blood & Tissue kit (Qiagen) following the manufacturer's instructions. The PCR reaction mixture contained 2.5 µl 10x PCR buffer (Qiagen), in which the final concentration of $MgCl_2$ was 2.0 mM, 200 µM of each dNTP (Sigma Aldrich), 0.5 µM of each primer, 0.125 units Hot Start Taq1 DNA polymerase (Qiagen), 2 µl DNA, with water added to a final volume of 25 µl. A negative control (sterile water) was included in each test. The PCR was performed using a S1000TM Thermal Cycler

(BIO-RAD Laboratories, Inc.). Three standard cycle programs were used for the initial screening: denaturation step at 95°C for 5 minutes, 35 cycles of 30 seconds at 95°C, 30 seconds at 48, 52 and 58°C, 60 seconds at 72°C, and finally 5 minutes extension at 72°C. Further optimization included a gradient of annealing temperatures in the range of 44–62°C, modulating the extension time depending on the expected PCR product length, and $MgCl_2$ concentration. All PCR products were tentatively sequenced with the same primers as those used for amplification. DNA sequences of both strands were obtained using the BigDye Terminator cycle sequencing ready reaction kit (Applied Biosystems Inc.) using an automated DNA sequencer (Applied Biosystems Prism 3700). Sequencing was carried out at the sequencing facility of the Høyteknologisenteret i Bergen, at the University of Bergen – Norway (http://www.uib.no/en/seqlab).

3.3 Data analyses and phylogeny estimation

Amplified and sequenced gene fragments were blasted in the GenBank database for orthology verification, accepting a minimum threshold of E-value=1E-4. The genes were further investigated for possible theoretical indication of paralogy and/or multiple copies in the OrthoDB database. The majority of selected genes were confirmed to be single or low copy in other insect and in many arthropod genomes as well (**paper I**). Sequences were aligned using BioEdit v.7.2.5 with manual adjustments, locating intron borders based on GT-AG nucleotides or alternative splicing sites. MAFFT v.7 was used to align protein coding genes with indels-rich regions and the ribosomal gene *28S rRNA*. Gblocks v.0.91b was used to trim ambiguously aligned regions in *28S rRNA*. Introns were removed from protein coding genes before all phylogenetic analyses.

Three different methods were used to reconstruct phylogenetic trees: Neighbour-joining (NJ), Maximum Parsimony (MP) and Bayesian inference (BI). Beetle phylogenies were inferred in a Bayesian framework in the software MrBayes v3.2.5. Phylogenetic analyses using MP and NJ were performed using PAUP* v4b10 with heuristic searches using TBR branch swapping for parsimony based analyses. Two preferential partition schemes were used: a) by gene; b) by codon position. The best evolutionary model for each data partition was estimated using jModelTest and MrModeltest v2.3 (for larger dataset) according to AIC criterion (**paper I-III**). For the complete Scolytinae data set (18 genes for 186 species) the final alignment was also examined in PartitionFinder v1.1.1 which defined partitions (29) and best-fitting models for each of them (**paper II**). The divergence times for Scolytine beetles were estimated using

the software BEAST v1.8.2, with input files generated in BEAUti. Biogeographical inference was obtained by applying statistical DEC Lagrange (S-DEC) analysis and by Bayesian binary MCMC (BBM) analyses as implemented in RASP - Reconstruct Ancestral State in Phylogenies (details and references in **paper I-III)**.

NJ based phylogenies were used in the preliminary analyses of the sequences to identify gene duplication or paralogous copies. MP and BI were used for large-scale analyses and for reporting the main findings (**paper I–III**). In general, Bayesian Inference was more informative compared to Maximum Parsimony when applied to our data sets in terms of obtaining more resolved trees. Our main conclusions are largely based on Bayeasian Inference which also provided resolution for some deeper nodes; recent relationships and node support within monophyletic tribe are frequently supported by Maximum Parsimony analyses as well. Different genes can have variable degree of utility for specific time frames in accordance with their evolutionary rate; hidden phylogenetic signal can emerge when such genes are concatenated (Olmstead and Sweere, 1994).

In order to evaluate the ability of different analyses to recover phylogenetic information from different data partitions, we analyzed each gene separately, and concatenated, applying the selected models for each partition (BI). Contradictory tree topologies, for some taxa, were observed when different analytical approaches (MP and BI) were used on the same data set. However, the comparison of clades, branch lengths and support values obtained with the two different methods allowed testing for congruent results. Those clades resilient to changes, with maximum or high node support under the two methods applied (MP and BI) and partition scheme used, were considered more reliable.

A Bayesian MCMC approach present higher sensitivity to phylogenetic signal, as demonstrated by simulation studies which showed the capacity of BI to detect small amounts of signal in a data set (Alfaro et al., 2003). Indeed, phylogenetic information in data sets which may contain homoplasious characters can be more easily used by appropriate models of DNA evolution in Bayeasian Inference, compared with simpler models underlying Maximum Parsimony (Alfaro et al., 2003, Brandley et al., 2009, Ronquist and Huelsenbeck, 2003).

4 List of publications

Paper I: Pistone D, Mugu S and Jordal BH (2016) Genomic mining of phylogenetically informative nuclear markers in bark and ambrosia beetles. PLOS ONE 11: e0163529. doi:10.1371/journal.pone.0163529

Paper II: Pistone D, Gohli J and Jordal BH (2017) Molecular phylogeny of bark and ambrosia beetles (Coleoptera: Scolytinae) based on 18 molecular markers. Systematic Entomology. doi:10.1111/syen.12281

Paper III: Mugu D, Pistone D and Jordal BH (2017) New molecular markers resolve the phylogenetic position of the enigmatic wood-boring weevils Platypodinae (Coleoptera: Curculionidae). Accepted for publication in: Arthropod Systematics and Phylogeny

5 Abstract of scientific articles

Paper I

Deep level insect relationships are generally difficult to resolve, especially within taxa of the most diverse and species rich holometabolous orders. In beetles, the major diversity occurs in the Phytophaga, including charismatic groups such as leaf beetles, longhorn beetles and weevils. Bark and ambrosia beetles are wood boring weevils that contribute 12 percent of the diversity encountered in Curculionidae, one of the largest families of beetles with more than 50,000 described species. Phylogenetic resolution in groups of Cretaceous age has proven particularly difficult and requires large quantity of data. In this study, we investigated 100 nuclear genes in order to select a number of markers with low evolutionary rates and high phylogenetic signal. A PCR screening using degenerate primers was applied to 26 different weevil species. We obtained sequences from 57 of the 100 targeted genes. Sequences from each nuclear marker were aligned and examined for multiple copies, pseudogenes and introns. Phylogenetic informativeness (PI) and the capacity for reconstruction of previously established phylogenetic relationships were used as proxies for selecting a subset of the 57 amplified genes. Finally, we selected 16 markers suitable for large-scale phylogenetics of Scolytinae and related weevil taxa.

Paper II

The phylogeny of the large weevil subfamily Scolytinae has been difficult to resolve based on a limited number of genetic markers. With more than 6,000 nominal species in the subfamily, the general lack of resolution at deeper nodes indicates that large sequence volumes are needed to solve this problem. We have therefore assembled a large molecular dataset consisting of more than 10 kb of nucleotides from 18 gene fragments, for 182 species. Nucleotide and amino acid translated data were analyzed using Bayesian and parsimony based approaches, which gave largely congruent results. Compared to previous analyses, we obtained greater resolution for some of the deeper nodes, and detected many unexpected relationships that were strongly supported by our data. The tribe Scolytini was recovered as the earliest divergent lineage in Scolytinae, sometimes placed together with the hexacoline genus *Microborus*. Among the currently 26 recognized tribes, 15 were monophyletic, whereas the remaining tribes were largely paraphyletic. The majority of species in the tribe Hypoborini was recovered as the sister lineage to a large group containing the species-rich tribe

Dryocoetini, which includes the recently radiated ambrosia beetles in Xyleborini, and Ipini, which includes another recent group of ambrosia beetles in Premnobiina. Cryphalini, Hylesinini and Hylurgini were strikingly polyphyletic tribes each consisting of several independent lineages. Subgroups were to a large degree defined by geographical affinities, showing a clear distinction between the northern and southern hemispheres. The affiliation of the inbreeding genus *Hypothenemus* was revealed with strong support, as the sister group to the Malagasy and East African species of the genus *Cosmoderes*. *Cryptocarenus* was previously assumed to be the sister lineage of *Hypothenemus*, but was here found to be part of *Corthylini*, near *Araptus*. These and many other findings document the need for a thorough revision of the current classification of genera and tribes, including a systematic re-evaluation of morphological characters.

Paper III

The precise phylogenetic position of the weevil subfamily Platypodinae continues to be one of the more contentious issues in weevil systematics. Morphological features of adult beetles and similar ecological adaptations point towards a close relationship with the wood boring Scolytinae, while some recent molecular studies and larval morphology have indicated a closer relationship to Dryophthorinae. To test these opposing hypotheses, a molecular phylogeny was reconstructed using 5,966 nucleotides from ten gene fragments. Five of these genes are used for the first time to explore beetle phylogeny, i.e. the nuclear protein coding genes *PABP1*, *UBA5*, *Arr2*, *TPI*, and *Iap2*, while five markers have been used in earlier studies (*28S*, *COI*, *CAD*, *ArgK* and *EF-1α*). Bayesian, maximum likelihood and parsimony analyses of the combined data strongly support a monophyletic Curculionidae (the advanced weevils with geniculate antennae), where Brachycerinae, Platypodinae and Dryophthorinae formed the earliest diverging groups. Dryophthorinae and core Platypodinae were sister groups with high support, with the contentious genera *Mecopelmus* Blackman, 1944 and *Coptonotus* Chapuis, 1873 placed elsewhere. Other lineages of wood boring weevils such as Scolytinae, Cossoninae and Conoderinae were part of a derived, but less resolved, clade forming the sister group to Entiminae. Resolution among major curculionid subfamilies was ambiguous, emphasizing the need for large volumes of data to further improve resolution in this most diverse section of the weevil tree.

6 Discussion

6.1 Experimental design

Nuclear protein coding genes offer an ideal source of informative data for phylogenetic studies. Although difficult to develop and standardize, the wide variability in evolutionary rates render such molecular markers very useful to resolve phylogenies at different ranks. Therefore, new nuclear markers for molecular phylogenetics were developed to achieve the main goal of this research project which was to increase resolution in bark and ambrosia beetle phylogeny. Almost the entire set of novel protein coding genes (13 out of 16 developed), in combination with the previously defined markers (5), was used to explore phylogenetic relationships within Scolytinae. A subset of the selected protein coding genes (5) were used in association with previously developed markers to investigate the weevil phylogeny and to test the monophyly of Scolytinae and their placement in the weevil tree, particularly with respect to Platypodinae and other wood boring taxa.

The entire research project was conceived as a multistep unidirectional workflow. A small number of selected representative species (26) were used for primer testing. Although the initial taxon sample was quite broad, it did not represent the entire variability encountered in weevils. As a consequence, the number of sequences obtained and hence the missing data for each gene were highly variable. Therefore, the phylogenetic trees based on single protein coding genes were often rather difficult to compare with each other and to evaluate for predicting the utility of these markers. Nevertheless, sixteen protein coding genes were selected based on an evaluation of the advantageous and unfavorable properties. Thus, these markers were tentatively amplified for a large number of species (more than 250) and then used to investigate phylogenetic relationships in Scolytinae and in other weevils. This constitute a sort of paradox, since only after extensive primers testing, sequences analyses, paralogy evaluation and reconstruction of single-gene phylogeny including a large number of species, more reliable information can be obtained.

6.2 The development of novel nuclear genes for beetle phylogenetics

With the selection and optimization of 16 nuclear protein genes for phylogeny reconstruction in Scolytinae and weevils, the main goal of this PhD project was realized. Sixteen new markers with potential broad application in beetle phylogenetics contribute a modest improvement. Nonetheless, this study more than doubled both the number of nuclear genes and nucleotides ever used in reconstructing molecular phylogenies in Coleoptera using a PCR

and Sanger sequencing approach. Similar large-scale gene screening on beetles has only occurred in a few studies, but not with a similarly high number of genes and species investigated (Wild and Maddison, 2008, Tarasov and Dimitrov, 2016). During this PhD project, 57 markers (100 genes were screened) resulted in one or more sequences showing high homology with expected target genes in nucleotide BLASTn searches (E-value = 1E-4). On the contrary, the remaining 43 genes tested, were discarded in the early phase due to lack of PCR amplification, sequencing failure and/or non-specific amplification. Several other genes were previously screened and only a few of these were selected and optimized in bark and ambrosia beetles during the past years (Jordal and Cognato, 2012, Jordal et al., 2011, Jordal, 2007). Therefore, this large-scale study largely confirmed that the development of protein coding genes for phylogenetic analyses in these beetle taxa is an arduous task (**paper I**). In addition, using only one or two potentially suboptimal primer pairs to amplify the target genes may have excluded 'phylogenetically informative' markers from the screening procedure at an early stage. Gene fragment length, the number of introns and the variability of the intron pattern in different species were additional criteria for marker selection. However, as a result of this strategy, useful and informative genes might have been excluded.

In order to cover the genetic variability within the different Scolytinae tribes and other Curculionidae, degenerate primers were designed on conserved exons. The use of degenerate primers, offered the advantage to amplify and sequence a wider array of diverse beetle species, although causing frequent unspecific amplification problems. Some of these sequences were ascribed to non-targeted genomic regions of the beetles, but also to different organisms (e.g. fungi, nematode and bacteria) – both with and without gene specificity.

In addition, preliminary data showed the complexity of weevil genomes, with high levels of inter- or intra-specific genetic variability, especially in terms of non-conserved intron patterns, hypervariable regions and paralogous copies (**Figure 3**). This implies that considerable effort was required for PCR optimization and efficient Sanger sequencing. Obtaining good quality sequences is the first important requisite that must be fulfilled before any further phylogenetic evaluation of the gene can be attempted. Thus, this preliminary step represented the first bottleneck in the gene selection procedure.

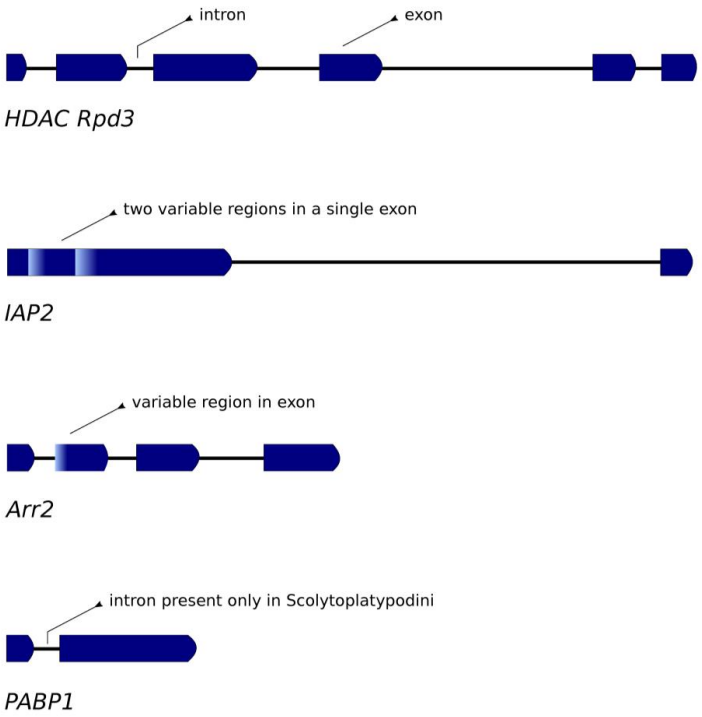


Figure 3 – Graphical representations of different degrees of complexity in four of the selected genes. The number and length of an intron is highly variable between closely related species. The shaded areas (light blue) indicate hyper variable regions or indels within the exon (dark blue) which can be translated into proteins with different amino acid length.

Concordance, defined as the capacity to recover relationships previously established by morphology or other molecular markers, remains a valuable method to evaluate the phylogenetic utility of new markers (Cho et al., 1995, Mardulyn and Cameron, 1999). Low phylogenetic signal from single gene phylogenetic analyses can generate polytomies for well-established clades and therefore such results must be carefully interpreted. Indeed, single-gene analyses are expected to provide just an indication of the gene performance; phylogenetic signal of a gene in a dataset can be perhaps more easily tested only under more exhaustive taxon sampling or in combined analyses with other genes.

Although the orthology assessment for the selected protein coding genes could be evaluated in several ways (e.g. OrthoDB database and taxa-monophyly recovery in test phylogenies - **paper I**) and it is strongly supported for the majority of the genes, the possibility that

undetected paralogs might have been included in the phylogenetic analyses cannot be entirely excluded. However, the presence of paralogous copies was unambiguously demonstrated only for one of the 16 selected nuclear genes. The heat shock protein 70 (*hsp70*) is reported to be present in multiple copies in the large majority (97.7%) of arthropod genomes in orthoDB database. This information made us suspect paralogy, and was later confirmed by our analyses of sequences that revealed presence of a single amino acid insertion in only some phylogenetically unrelated species. These taxa grouped together in a well-defined cluster in the test phylogeny for this particular gene (**paper I**). The fact that a gene is present in multiple copies is not necessarily problematic for phylogenetic inference. Understanding gene variability (e.g. paralogs and pseudogenes) between species can be time-consuming and the development and optimization of copy-specific primers require much effort (Danforth and Ji, 1998, Jordal, 2002). Nonetheless, such kind of nuclear markers were successfully developed for insect phylogenetics (i.e. *EF—1α*, *enolase* and *wingless*).

Even if a deep knowledge of insect genomes in a comparative perspective is currently rather limited (Krauss et al., 2008), a higher number of introns seem to be present in beetles than in other insect groups (Dolezelova et al., 2006, Wild and Maddison, 2008). In our study, the sequences of the amplified genes showed large variability in the number and length of introns (**Figure 3**). Nevertheless, a high degree of interspecific variability in gene structure is not always equivalent to problems in phylogenetic inference, especially if such variability is highly informative (e.g. clade defining indels - **paper I** - **III**).

What will be the next step in Scolytinae phylogenetics? Is it time to abandon multiple genes approach based on Sanger sequencing? At the time of writing, a ground-breaking paper which defined novel protein coding genes with proved utility in beetle phylogenetics was published (Che et al., 2017). The authors applied a whole-genome scan to investigate several insect genomes for selecting useful single copy nuclear genes. They were able to define 95 markers with an impressive amplification and sequencing success rate (90%). In order to select these markers, 1489 genes which showed the desired features were initially considered. In my opinion, the reasons behind the success of this study compared to my PhD project based on Sanger sequencing which achieve less outstanding accomplishments might be traced in the decision of targeting and design primers only on single exons, amplify these regions through a more specific and sensitive nested-PCR and use NGS technology for sequencing. In particular, NGS might have guaranteed the solution to one of the main problems in Sanger sequencing approach, namely, the reduction of missing data though a more efficient detection and removal of unspecific sequences.

6.3 Adding resolution to Scolytinae phylogeny with 18 molecular markers

A large-scale phylogenomic approach which includes several protein coding genes should secure sufficient data to enable reconstruction of phylogenies where inference of evolutionary transitions at different hierarchical levels is possible. This study included the largest number of molecular markers ever used in Sanger-based beetle phylogeny. Sixteen nuclear protein coding genes were used for reconstructing the phylogenetic branching pattern in the Scolytinae subfamily, including a high number of taxonomic representative species as well. However, a total of 18 markers was shown to be insufficient for obtaining maximum phylogenetic resolution of this species-rich, old and hypervariable group. Nevertheless, important events in the evolutionary history of Scolytinae were clarified and previous findings were supported by the new data. On the other hand, the early evolution and diversification of Scolytinae remained unclear and the relationships among the majority of the tribes were not resolved. This suggests that certain relationships may be inherently difficult to resolve or that more data will be required, or both (**paper II**).

All the analyses pointed out that the widely used current classification of Scolytinae (Wood, 1982, Wood, 1986) is wrought with para- and polyphyletic genera and tribes. While the monophyly of fifteen tribes (*sensu* Wood) was confirmed, paraphyly and polyphyly were clearly demonstrated for several of the remaining tribes, in particular Hylurgini, Hylesinini and Cryphalini (**paper III** - **Figure 4**). Two early diverging lineages which were monophyletic and stable at the base of the tree (e.g. tribe Scolytini and the genus *Microborus*– currently in Hexacolini) were recognized. The tribe Scolytini is characterized by morphological characters not shared by other bark and ambrosia beetles which might justify the placement of this tribe as an early diverging or even a separate lineage. For the first time, the tribes Hypoborini, Dryocoetini (with nested Xyleborini), and Ipini (with nested Premnobiina) were shown to share a common origin, with divergences among the main groups which were dated to more than 80 million years ago (**paper II**). More precisely, Hypoborini were recovered as the sister lineage to (Ipini + Premnobiina) + (Dryocoetini + Xyleborini) in all the analyses and in general with high or maximum support values. Therefore, a sister relationship between Hypoborini and Micracidini, as previously suggested both based on molecular data and morphological similarities in protibial, antennal and proventricular characters, was not confirmed in this study (Jordal and Kaidel, 2016). However, the findings of this recent phylogenetic analysis, namely a single trans-Atlantic

disjunction for Micracidini and the inclusion of *Cactopinus* and their closely related genus *Phloeocleptus* in the American clade, were here confirmed and strongly supported (**paper II**). The general resolution within tribes increased as well. Some examples of advancements in Scolytinae classification can be found in the recognition of the sister relationship of the inbreeding genus *Hypothenemus* and a Madagascar genus near *Cosmoderes*, and in the new placement of the genus *Cryptocarenus* in Corthylini, near the genus *Araptus*. These new findings have important implication on our understanding of Scolytinae evolution, for example in the origin of inbreeding. Further analyses at tribal level will probably require the inclusion of more taxa rather than more characters for resolving other contentious relationships. On the other hand, the resolution remained rather low at deeper nodes which could be interpreted as an indication of lack of sufficient molecular data (i.e. phylogenetic signal) in rapidly evolving genes (see **paper I**). In general, tribes of Cenozoic age were easier to resolve – also with fewer genes – while older splits will necessarily require much more molecular markers. Young groups such as Ipini and their sister group Dryocoetini were well resolved also in previous studies based on five molecular markers. In our study, the inadequacy of the novel genes for resolving ancient divergences could be considered the most logical explanation for the low resolution at deeper nodes in the Scolytinae tree. However, considering that the phylogeny of Platypodinae - a group of comparable age to Scolytinae - was considerably more resolved based only on data from the same five markers previously used for Scolytinae, the new markers could only be part of the problem. In addition to the high evolutionary rate of the majority of the novel genes (**paper I**), deep coalescence could be a major challenge in resolving relationships between Scolytinae tribes with similar stem ages (**paper II.**). Therefore, it is quite possible that the polytomy observed in Scolytinae deep nodes might derive from rapid radiation events (hard polytomies). Further studies should focus on testing the hard polytomy hypotheses in Scolytinae which imply that deeper nodes might be intrinsically difficult to resolve even with considerable data accumulation. On the contrary, phylogenetic studies in other insect taxa have confirmed that the addition of a handful of new nuclear genes increased both resolution and node support in phylogenetic trees, often solving polytomies that were previously present due to low amount of data (Kodandaramaiah et al., 2010, Ruiz et al., 2009). With all these considerations, the question is: why are Scolytinae so difficult to resolve phylogenetically using 16 markers? Is the evolutionary history of this group particularly complex? Or do the novel protein genes have high evolutionary rates which cause reduced phylogenetic signal for ancient divergences? Or

in other terms, as presented in a recent molecular phylogeny of Diptera, are we dealing with an explosive radiation or uninformative genes (Winkler et al., 2015)?

In general, a more careful evaluation of paraphyletic groups (e. g. Hylurgini, Hylesinini and Cryphalini) will be required. Discrepancies among current classification and phylogeny were quite evident, and emerged even more strikingly when the results of biogeographic analyses were considered. In the pre-molecular era of insect systematics, researchers were occasionally misled by convergent morphological characters which lead to an incorrect classification. The previous placement of *Premnobius* in Xyleborini and the definition of Hylurgini and Hylastini as independent lineages represent two of the issues that were recently changed (Cognato, 2013, Jordal et al., 2014). The presumed sister relationships among Hypoborini and Micracidini suggested by recent molecular analyses was apparently supported by morphological characters (Jordal and Kaidel, 2016). The phylogeny based on 18 markers showed that the previous finding was probably based on limited data and therefore convergent evolution could more easily explain morphological similarities (**paper II**). A long history of incorrect classification due to the intrinsic limits of morphology based taxonomy should also be considered to get a better explanation of part of the highly supported results in the latest analyses. The genus *Cryptocarenus* represents another example of such obvious taxonomic mistakes, but similar cases can also be found in the tribe Hylurgini and Hylesinini where the relationships among few species are largely in agreement with our biogeographic analyses resulting in northern or southern hemisphere groups (**paper II**). Additionally, a high level of complexity describes the biogeographic history of this beetle group, which includes frequent occurrences of long distance dispersal and a general lack of endemism.

Stability in tree topology in different analyses can be considered a strong confidence measure for a phylogenetic hypothesis - a long debated philosophical approach to phylogenetic consensus (see Miyamoto and Fitch, 1995). Changing the analysis assumptions can provide information about the strength of phylogenetic signal in a data set. Data sets with weak phylogenetic signals can be strongly influenced by such changes. Congruence in tree topology and similar node support values when using different methods of analysis (e.g. Bayesian analyses and parsimony) or different partition schemes can be considered as a sign of a predictable dataset for phylogenetic inference. In the same way, contentious or weakly supported relationships can be easily highlighted by changes in topology.

Missing data, which almost always was less than 50% (close to 70% only for two genes), are expected to have a negligible negative effect on tree resolution (Wiens and Morrill, 2011). The unstable placement of some species (e.g. *Acacicis* and *Halystus*), might have been caused

by low amplification rate for several genes (missing data), but in general problematic taxa could not be explained by low quality or missing data.

Many deep level branching patterns are still difficult to reconstruct, the resolution of which will require further data. Therefore, taxonomic changes can only be tentatively suggested and additional studies, including morphological analyses and more molecular markers will be required before a proper re-classification of taxa can be made.

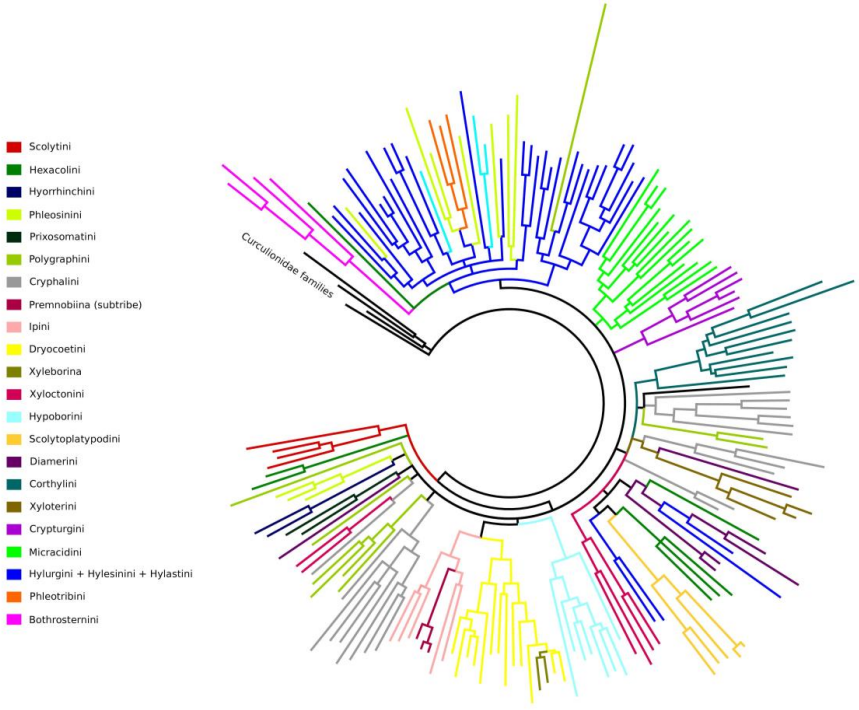


Figure 4 - Phylogenetic tree based on Bayesian analyses of the concatenated dataset (10156bp) partitioned by gene. Several tribes are monophyletic and highly supported under different partition schemes. Paraphyly and polyphyly is also evident, especially for Hylurgini (with subclades composed by some of the genera currently in Hylesinini and Hylastini) and Cryphalini (see **paper II)**.

Scolytinae beetles represent a unique system for testing hypotheses on the origin and evolution of unusual ecological and behavioral adaptations. Among the most fascinating aspects, permanent inbreeding, host plant association and fungus farming have been previously investigated from a phylogenetic perspective (Farrell et al., 2001, Jordal and Cognato, 2012, Gohli et al., 2017). An intricate pattern can now be more clearly derived for the origin and directionality of such key evolutionary events in Scolytinae (**Figure 5**). The host plant preference is still one of the most difficult traits to track and follow in the different

lineages. With the exception of a few groups such as the conifer feeding Ipini and the *Dendroctonus* clade, which represent species-rich conifer associations, the evolution of this trait in other lineages is more enigmatic. Same species within a genus can be polyphagous and feed also on angiosperm hosts or this feeding preference can be present only in one or few genera within a tribe - e.g. *Hylocurus*, *Pityophthorus* and *Scolytus* (Gohli et al., 2017, Avtzis et al., 2012). Obtaining a well resolved phylogeny of Scolytinae, in particular a better understanding of relationships within Hylurgini, Hylesinini and Hylastini, will be an important step in clarifying the evolution of host preference (**Figure 5**).

In general, the multiple origins of fungus farming (10) and inbreeding (6) are somewhat easier to date since many of these evolved in more recent times and therefore in more resolved clades. In some lineages, these two innovative traits originated together, such as in Xyleborini (circa 16 Ma) and in the subtribe Premnobiina (uncertain, but recent). There are also exceptions where the origin of the irreversible adaptation to 'ambrosia feeding' characterizing entire tribes (e.g. Scolytoplatypodini and Xyloterini) could be dated back to more than 50 Ma (**Figure 5 - paper II**).

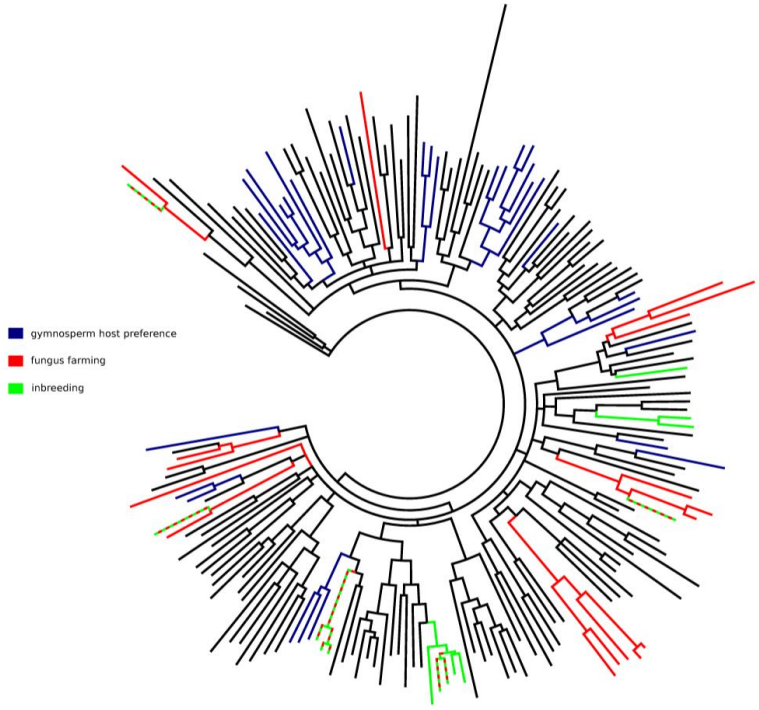


Figure 5 – Phylogenetic tree showing the distribution of three ecological adaptations (conifer host plant use, fungus farming and permanent inbreeding) which evolved multiple times in different lineages of Scolytinae.

6.4 Nuclear genes confirmed the phylogenetic placement of Platypodinae

Resolution of the weevil phylogeny at family and subfamily level represents a great challenge, also for mitochondrial genome studies (Gillett et al., 2014, Haran et al., 2013). At present, the phylogenetic relationships within Curculionidae remain the most controversial (see Chapter 1 – Section 1.5). How many nuclear genes are required to obtain a reliable phylogeny that can be used to trace patterns of evolution in the advanced weevils? Relationships within Curculionidae was difficult to estimate based on *18s rRNA* alone (Farrell, 1998, Marvaldi et al., 2002). The lack of resolution was inferred as a result of explosive radiation in weevils, predicting that further phylogenetic resolution would have required a much more extensive sampling of characters (more genes and/or additional morphological data) and taxa (Marvaldi et al., 2002). This prediction was largely demonstrated to be correct by a recent study that included more nuclear genes, which obtained better resolution, albeit with modest node support (McKenna et al., 2009).

Despite a continuous improvement in weevil classification based partly on molecular systematics, one of the long-lasting problems remains the relationship among Scolytinae and Platypodinae. In order to investigate this and other related phylogenetic issues, the weevil phylogeny was reconstructed using five newly developed markers (*Iap2*, *Arr2*, *TPI*, *UBA5* and *PABP1*) in combination with the five 'standard' markers (*28s rRNA*, *COI*, *CAD*, *EF-1α* and *ArgK*) previously used in other weevil groups. Results, based on these 10 markers, showed that the general placement of two major clades remained consistent inside the family Curculionidae under different analyses and partition schemes. The first clade grouped the core Platypodinae (excluding *Mecopelmus zeteki* and *Coptonotus cyclopus*) and Dryophthorinae, as highly supported sister lineages, and a second clade included Entiminae as the sister group to all other Curculionidae sensu stricto (**Figure 6**). The placement of Platypodinae, was consistent and highly supported, while the remaining phylogeny of advanced weevils was less resolved. However, all subfamilies sensu Alonso-Zarazaga and Lyal (1999) were largely monophyletic, with lack of structure between them, with the exception of Entiminae separated from the other Curculionidae sensu stricto (**paper III**).

The relationship among Platypodinae and Dryophthorinae was not so unexpected and it was already hypothesized based on morphology and more recently on large-scale molecular data (Gillett et al., 2014, Haran et al., 2013, Marvaldi et al., 2002). To a certain extent, it was surprising that this relationship was solved and maximally supported by using only 10 genes,

while a previous study using larger amount of molecular data (complete mitochondrial genomes), recovered this relationship with low node support (Gillett et al., 2014).

Scolytinae are not closely related to Platypodinae, closing a long debate on their hypothetical sister relationship, repeatedly proposed in weevil classification (Schedl, 1962, Bright, 2014, Kuschel, 1995, Kuschel et al., 2000, Wood, 1986). The recognition of Scolytinae and Platypodinae as unrelated subfamilies emphasize that probably wood boring evolved independently in other weevil groups (e.g. Cossoninae and Conoderinae). The monophyly of Scolytinae is well supported and generally not questioned, but according to results of the phylogenetic analyses based on the 10 markers, this subfamily might consist of at least two paraphyletic groups (paper III; see also Gillett et al., 2014). The paraphyly of Scolytinae might be due to low phylogenetic signal in the molecular markers, or it could indicate an independent origin of the two lineages (see **paper II**).

Primers were designed for PCR amplification and sequencing in species belonging to the subfamily Scolytinae, but were also demonstrated to be useful in closely related weevil families and subfamilies: this finding indicates a high or moderate level of conservation for the selected protein coding genes across the weevils (**paper I** and **paper III**). The initial predicted utility of the primer was empirically demonstrated as well as the utility of the genes in investigating phylogenies of beetle taxa at various ranks. Therefore, the degenerate primers might potentially be useful in solving phylogenetic relationships in other Curculionoidea families and subfamilies. However, taxon-specific primers can certainly increase PCR and sequencing success and therefore decrease the missing data in the final alignments.

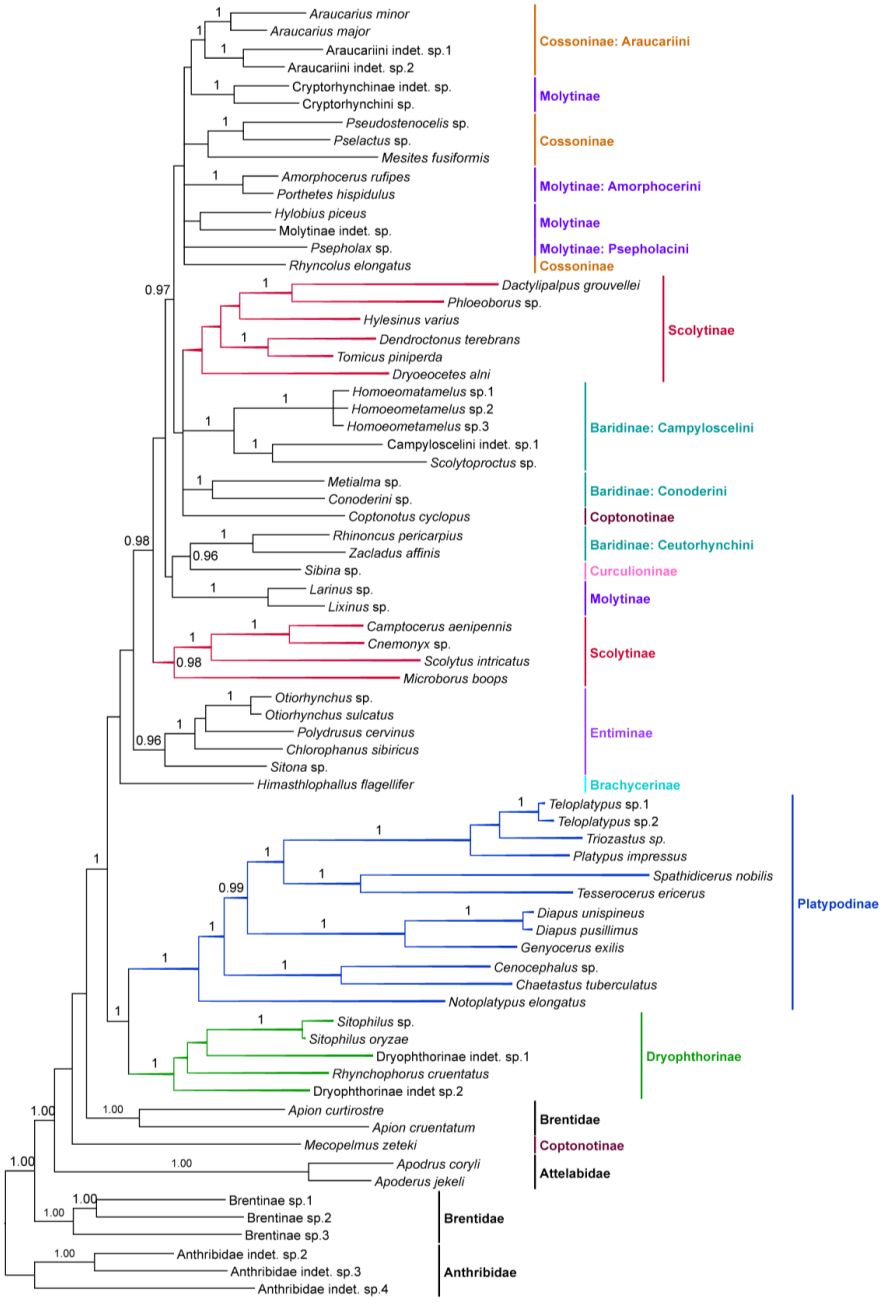


Figure 6 – Phylogenetic tree based on the Bayesian analysis of 10 molecular markers using seven partitions (**paper III**). Family and subfamily names follow Oberprieler et al. (2007).

7 Concluding remarks and future perspectives

7.1 Next frontiers in bark and ambrosia beetles phylogenetics and evolution

An updated and more resolved phylogeny of bark and ambrosia beetles will provide a reliable framework to better test hypotheses on the origin of mating systems, host plant preference and fungus farming. Improving our understanding of Scolytinae evolution allows introduction to the next level of complexity, attempting to elucidate the possible role that different microbes could have played in the diversification process. Bark and ambrosia beetles show variable degree of association with diverse organisms, including bacteria, yeast, and mycelial fungi. While the symbiotic multicellular fungi contribute an important but variable component of the diet of these wood boring insects, the role of other microbes remains more unclear. As an example of the complexity of possible interactions, the mycangia of ambrosia beetles host diverse communities of vertically transmitted bacteria (Hulcr et al., 2012), but the intimate relationships of these bacteria with the host and the symbiotic fungi are still far from being completely understood. This research line certainly deserves more attention, since the beetle-associated microbiome might have had a prominent role in the evolutionary success of Scolytinae and other wood boring lineages. Bacteria were implicated in the modification of mating systems, degradation of plant compounds, synthesis of vitamins and amino acids in different insect species (Sudakaran et al., 2017, Werren et al., 2008, Nikoh et al., 2014). On the contrary, a limited number of studies investigated the role of bacterial microorganisms in promoting ecological specialization and refining metabolic capacity in Scolytinae and Platypodinae (Fabryova et al., 2017, Hernandez-Garcia et al., 2017, Popa et al., 2012).

Bacterial endosymbiosis (intracellular symbiosis) is extremely common in insects, where it has been shown to be implicated in the evolutionary success of groups with specialized diets (e.g. sap and phloem feeders and hematophagous insects) and some of these host-bacteria interactions have been widely studied (Kikuchi, 2009, Clark et al., 2010, Skidmore and Hansen, 2017). In beetles, microscopy and molecular analyses confirmed the presence of symbiotic bacteria in many species from different families and in particular in the weevils (Sontowski et al., 2015, Kellner, 2002, Masson et al., 2015a, Kuriwada et al., 2010, Lefevre et al., 2004). Specialized symbiotic organs (bacteriome or mycetome) harboring bacterial endosymbionts have also been described (Toju et al., 2013, Masson et al., 2015b). The endosymbiotic bacteria reported in weevils generally grouped in the alpha-proteobacteria (genera *Wolbachia* and *Rickettsia*) and gamma-proteobacteria (genera *Nardonella*, *Sodalis*,

Curculioniphilus and *Arsenophorus*) (Lefevre et al., 2004, Conord et al., 2008, Kuriwada et al., 2010).

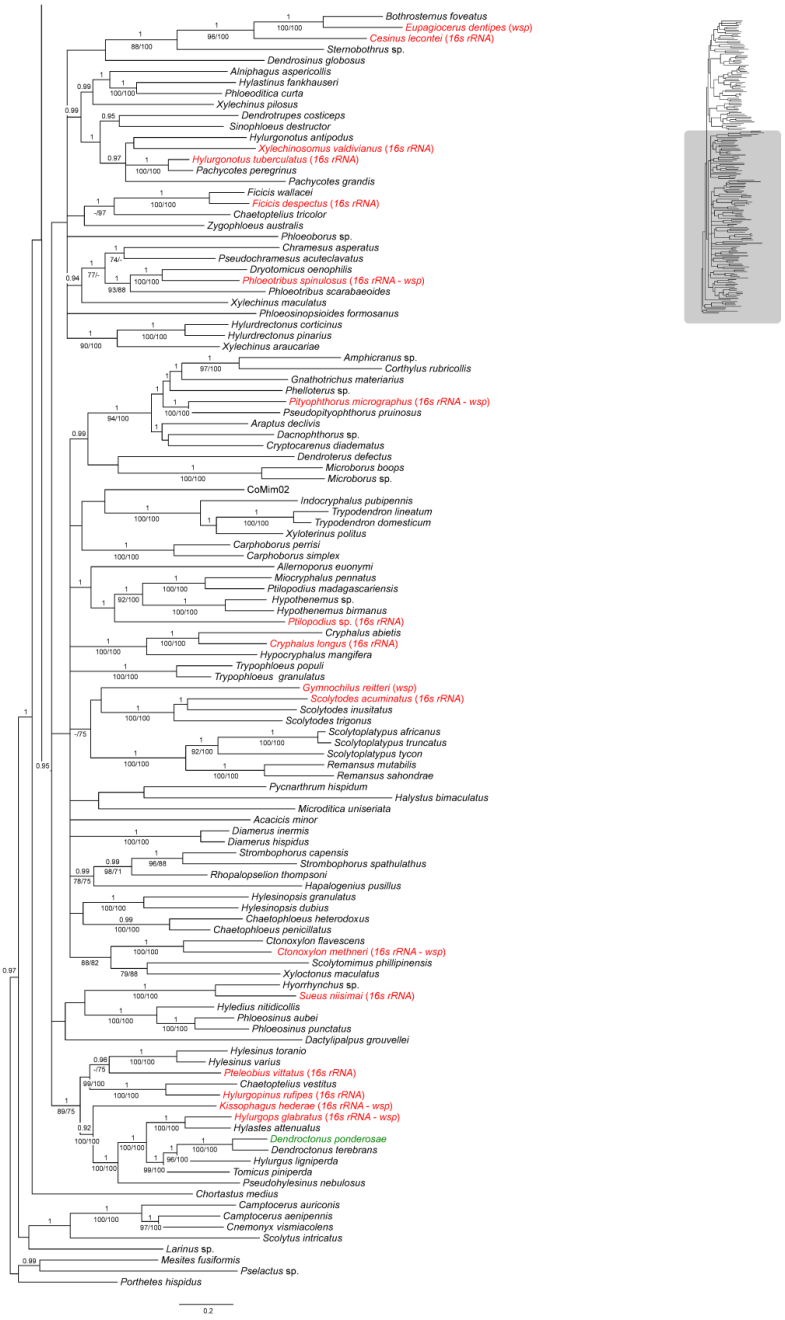
The ubiquitous microorganisms of the genus *Wolbachia* are one of the bacteria more widely investigated in bark and ambrosia beetles (Kawasaki et al., 2016, Arthofer et al., 2009, Lachowsky et al., 2015). *Wolbachia* are the endosymbionts most frequently reported in insects, but also present in other arthropod taxa such as spiders, mites and crustaceans and in filarial nematodes (Ilinsky and Kosterin, 2017, Comandatore et al., 2015, Werren et al., 2008). In general, *Wolbachia* preferentially reside inside the cells of the gonads where they are capable of manipulating the reproduction of their hosts (e.g. cytoplasmic incompatibility, parthenogenesis), ensuring their vertical transmission as well. A phylogenetic study on *Wolbachia* in Scolytinae beetles investigated the presence of these bacteria in 23 species detecting PCR positivity for eight (mainly haplodiploid) species and a role in the development of such reproductive strategy was hypothesized (Kawasaki et al., 2016). The coffee berry borer (*Hypothenemus hampei*) is another haplodiploid and inbreeding species where *Wolbachia* was shown to play an essential role in beetle reproduction and fitness (Mariño et al., 2017b).

Occasionally, during the PCR and Sanger sequencing of the protein coding genes in this project, amplification by our primers produced unspecific sequences which showed high identity (>90%) with alpha- (*Wolbachia* and *Rickettsia*) and gamma-proteobacteria (*Sodalis* spp., *Arsenophonus* spp.) as identified by BLASTn database search. Frequently, the low identity with uncharacterized regions of bacterial genomes in the databases did not allow a proper identification of the bacteria.

Based on preliminary evidence for the presence of *Wolbachia* and other bacteria in Scolytinae, a molecular screening for *Wolbachia* was carried out using specific primes targeting the two genes *16 rRNA* and *Wolbachia surface protein – wsp* (Werren and Windsor, 2000, Baldo et al., 2005). The presence of *Wolbachia* in different lineages was mapped on the phylogeny of Scolytinae (**Figure 7**). This preliminary screening constitutes the first large scale analysis on the distribution of *Wolbachia* in bark and ambrosia beetles, with the possibility for testing of co-evolution or horizontal transmission patterns. The prevalence of *Wolbachia* was 16.4% (30 out of 182) in the Scolytinae species screened in this study (17.5% when two other species previously reported as infected with *Wolbachia* are included). Thirty species were PCR positive for 16s *rRNA* gene, fourteen of them were confirmed positive for *wsp* and three additional species (*Eupagiocerus dentipes*, *Gymnochilus reitteri* and *Xyleborus affinis*) were positive only for this second marker (unpublished data – **Figure 7**).

A PCR screening cannot be considered a solid proof for the presence of these bacteria in beetles since positivity might derive from other sources such as parasitoids or other invertebrates (Plantard et al., 2012, Brown et al., 2016). The distribution of *Wolbachia* in bark beetles seems to follow a rather casual pattern in most lineages, but a non-random distribution is hypothesized for Xyleborina (two species – both positive) and Dryocoetini (7/13 positive species to *16s rRNA* and/or *wsp*). However, since this molecular screening was not carried out at the population level, it is not possible to determine whether these bacteria were strictly associated with the beetles or occasional tenants (non-primary endosymbionts normally show less than 100% prevalence) and any speculation on the role of these bacteria in driving to haplodiploidy would be premature. On the other hand, considering that this molecular survey most likely represents an underestimation of the real prevalence – assuming an infection rate lower than 100% for the large majority of the investigated species – *Wolbachia* seems to be widespread in Scolytinae. These results are particularly striking when compared to a previous study which found *Wolbachia* to be rare in the bark beetle *Pityogenes chalcographus* where it is present at the limit of PCR detection level and a nested PCR was required (Arthofer et al., 2009). Further development of the current screening program must involve different populations of the same hosts to achieve a better understanding of *Wolbachia* prevalence and patterns of co-evolution.

The presence of *Wolbachia* in arthropods has attracted the interests of entomologist for decades, but nevertheless is very little known about the role of this symbiont in this group of wood boring insects. The bacterial community associated with Scolytinae represents an extremely interesting research theme that can be explored more efficiently applying NGS technology. Such studies will gather preliminary information on the distribution and abundance of endosymbionts and the interaction with other members of the microbial community such as bacteria and fungi.


Bothrostemus foveatus
Eupagiocerus dentipes (wsp)
Cesinus lecontei (16s rRNA)
Sternobothrus sp.
Dendrosinus globosus
Alniphagus aspericollis
Hylastinus fankhauseri
Phloeoditica curta
Xylechinus pilosus
Dendrotrupes costiceps
Singhioeus destructor
Hylurgonotus antipodus
Xylechinosomus valdivianus (16s rRNA)
Hylurgonotus tuberculatus (16s rRNA)
Pachycotes peregrinus
Pachycotes grandis
Ficicis wallacei
Ficicis despectus (16s rRNA)
Chaetoptelius tricolor
Zygophloeus australis
Phloeoborus sp.
Chramesus asperatus
Pseudochramesus acuteclavatus
Dryotomicus oenophilis
Phloeotribus spinulosus (16s rRNA - wsp)
Phloeotribus scarabaeoides
Xylechinus maculatus
Phloeosinopsioides formosanus
Hylurdrectonus corticinus
Hylurdrectonus pinarius
Xylechinus araucariae
Amphicranus sp.
Corthylus rubricollis
Gnathotrichus materiarius
Phelloterus sp.
Pityophthorus micrographus (16s rRNA - wsp)
Pseudopityophthorus pruinosus
Araptus declivis
Dacnophthorus sp.
Cryptocarenus diadematus
Dendroterus defectus
Microborus boops
Microborus sp.
CoMim02
Indocryphalus pubipennis
Trypodendron lineatum
Trypodendron domesticum
Xyloterinus politus
Carphoborus perrisi
Carphoborus simplex
Allernoporus euonymi
Miocryphalus pennatus
Ptilopodius madagascariensis
Hypothenemus sp.
Hypothenemus birmanus
Ptilopodius sp. (16s rRNA)
Cryphalus abietis
Cryphalus longus (16s rRNA)
Hypocryphalus mangifera
Trypophloeus populi
Trypophloeus granulatus
Gymnochilus reitteri (wsp)
Scolytodes acuminatus (16s rRNA)
Scolytodes inusitatus
Scolytodes trigonus
Scolytoplatypus africanus
Scolytoplatypus truncatus
Scolytoplatypus tycon
Remansus mutabilis
Remansus sahondrae
Pycnarthrum hispidum
Halystus bimaculatus
Microdilica uniseriata
Acacicis minor
Diamerus inermis
Diamerus hispidus
Strombophorus capensis
Strombophorus spathulathus
Rhopalopselion thompsoni
Hapalogenius pusillus
Hylesinopsis granulatus
Hylesinopsis dubius
Chaetophloeus heterodoxus
Chaetophloeus penicillatus
Clonoxylon flavescens
Ctonoxylon methneri (16s rRNA - wsp)
Scolytomimus phillipinensis
Xyloctonus maculatus
Hyorrhynchus sp.
Sueus nilsimai (16s rRNA)
Hyledius nitidicollis
Phloeosinus aubei
Phloeosinus punctatus
Dactylipalpus grouvellei
Hylesinus toranio
Hylesinus varius
Pteleobius vittatus (16s rRNA)
Chaetoptelius vestitus
Hylurgopinus rufipes (16s rRNA)
Kissophagus hederae (16s rRNA - wsp)
Hylurgops glabratus (16s rRNA - wsp)
Hylastes attenuatus
Dendroctonus ponderosae
Dendroctonus terebrans
Hylurgus ligniperda
Tomicus piniperda
Pseudohylesinus nebulosus
Chortastus mediius
Camptocerus auriconis
Camptocerus aenipennis
Cnemonyx vismiacolens
Scolytus intricatus
Larinus sp.
Mesites fusiformis
Pselactus sp.
Porthetes hispidus
0.2

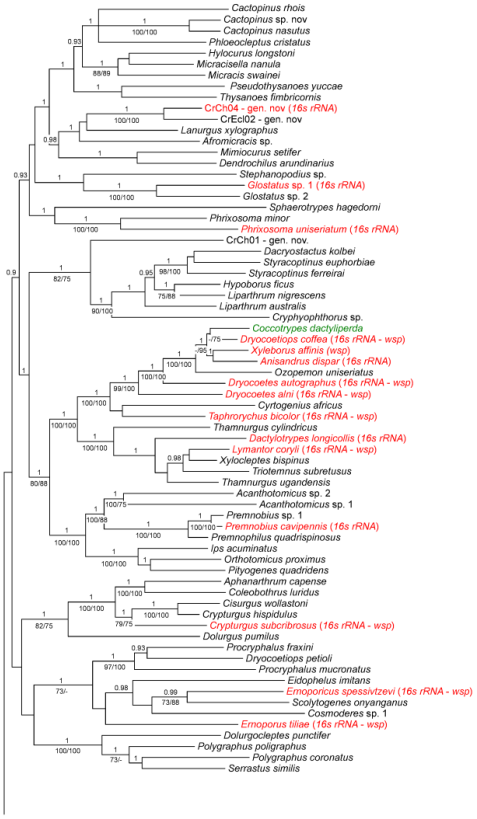


Figure 7 - Phylogenetic tree (BI - **paper II**) showing the presence of *Wolbachia* in Scolytinae species belonging to different tribes. The molecular screening targeting the 16s rDNA gene and *wsp* protein of *Wolbachia* resulted positive for 32 species (red); *Dendroctonus ponderosae* and *Coccotrypes dactyliperda* (green) are two of the several species – not used in our study - which were previously reported to be infected with *Wolbachia* bacteria.

7.2 Final considerations

This PhD project contributed to the field of weevil molecular systematics by characterizing novel protein coding genes for phylogenetic analyses. These genes provide different degrees of phylogenetic signal for resolving phylogenies at various ranks. Indeed, with the selection and optimization of 16 novel protein coding genes for Scolytinae, and at least five of them more broadly tested in weevils, this work significantly increases the number of available markers for PCR based weevil phylogenetics (**paper I – paper III**).

Thirteen of these nuclear markers were used to test phylogenetic relationships between tribes and genera in the subfamily Scolytinae. From a strictly systematics perspective, this PhD project achieved advances in understanding Scolytinae evolution and pointed out a series of mistakes in the current classification. More specifically, the majority of relationships within tribes were well resolved and highly supported, confirming fifteen of them as monophyletic, but also highlighting paraphyletic assemblages. However, a general lack of resolution persisted at deeper nodes and inter-tribal relationships remained largely unclear. Nevertheless, this study provides a new phylogenetic framework, highlighting taxonomic groups in need of revision.

Furthermore, the newly selected nuclear protein coding genes were also used to investigate family and subfamily level relationships within the superfamily Curculionoidea. A subset of these new markers (5 new genes + 5 previously defined), were useful for phylogenetic inference among different taxa. Indeed, Platypodinae were unambiguously placed as the sister clade to Dryopthorinae, although the selected markers were not sufficient to solve other problematic divergences in the weevil tree.

This study examined and characterized the highest number of protein coding genes in weevils using a PCR and Sanger sequencing approach. A priori, we expected to optimize at least 30 markers with sufficient phylogenetic signal, after screening 100 genes. We also predicted that 15-20 genes would be sufficient to solve relationships between tribes in Scolytinae. Although the results did not meet the expectations, this PhD project has made a substantial step forward in PCR and Sanger-based sequencing in weevil phylogenetics. However, the possibility that the addition of few more protein coding genes will provide enough phylogenetic signal to guarantee the resolution of ancient divergences in Scolytinae remains doubtful. If, as we now anticipate, some of the polytomies in deep nodes are due to rapidly diverging lineages, it might also provide a challenge for NGS technology.

In conclusion, PCR and Sanger sequencing based approach still claim a relevant place in insect molecular systematics given a proper availability of effective PCR primers. Other important advantages such as low costs and the possibility to amplify DNA from deteriorated and old samples, could guarantee even a longer survival of this popular technique. When the first NGS solutions appeared on the market, it could be tempting to predict that the Sanger sequencing period had its days, however many years later, it seems that this long-standing technology is not yet ready to step aside.

8 References

AHRENS, D., SCOTT, M. & VOGLER, A. P. 2011. The phylogeny of monkey beetles based on mitochondrial and ribosomal RNA genes (Coleoptera: Scarabaeidae: Hopliini). *Mol Phylogenet Evol,* 60**,** 408-415.

ALFARO, M. E., ZOLLER, S. & LUTZONI, F. 2003. Bayes or Bootstrap? A Simulation Study Comparing the Performance of Bayesian Markov Chain Monte Carlo Sampling and Bootstrapping in Assessing Phylogenetic Confidence. *Mol Biol Evol,* 20**,** 255-266.

ALONSO-ZARAZAGA, M. A. & LYAL, C. H. C. 1999. A world catalogue of families and genera of Curculionoidea (Insecta: Coleoptera) (excepting Scolytidae and Platypodidae). Entomopraxis, S. C. P. Edition 1999.

ARTHOFER, W., M., R., AVTZIS, D. & STAUFFER, C. 2009. Evidence for low-titre infections in insect symbiosis: Wolbachia in the bark beetle Pityogenes chalcographus (Coleoptera, Scolytinae). *Environ Microbiol,* 11**,** 1923-1933.

AVTZIS, D. N., BERTHEAU, C. & STAUFFER, C. 2012. What is Next in Bark Beetle Phylogeography? *Insects,* 3**,** 453-472.

AYRES, M. P., WILKENS, R. T., RUEL, J. J., LOMBARDERO, M. J. & VALLERY, E. 2000. Nitrogen budgets of phloem-feeding bark beetles with and without symbiotic fungi (Coleoptera: Scolytidae). *Ecology,* 81**,** 2198–2210.

BACA, S. M., TOUSSAINT, E. F., MILLER, K. B. & SHORT, A. E. 2016. Molecular phylogeny of the aquatic beetle family Noteridae (Coleoptera: Adephaga) with an emphasis on data partitioning strategies. *Mol Phylogenet Evol,* 107**,** 282-292.

BAKER, R. H., WILKINSON, G. S. & DESALLE, R. 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). *Syst Biol,* 50**,** 87-105.

BALDO, L., LO, N. & WERREN, J. H. 2005. Mosaic nature of the wolbachia surface protein. *J Bacteriol,* 187**,** 5406-5418.

BEAVER, R. A. 1989. Insect-fungus relationships in the bark and ambrosia beetles. *In:* WILDING, N., COLLINS, N. M., HAMMOND, P. M. & WEBBER, J. F. (ed.) *Insect-Fungus Interactions.* London: Academic Press.

BERNHARD, D., KOMAREK, A., BEUTEL, R. & RIBERA, I. 2009. Phylogenetic analysis of Hydrophiloidea (Coleoptera: Polyphaga) based on molecular data and morphological characters of adults and immature stages. *Insect Syst Evol,* 40**,** 3-41.

BLAIMER, B. B., BRADY, S. G., SCHULTZ, T. R., LLOYD, M. W., FISHER, B. L. & WARD, P. S. 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evol Biol,* 15**,** 271.

BRANDLEY, M. C., WARREN, D. L., LEACHE, A. D. & MCGUIRE, J. A. 2009. Homoplasy and clade support. *Syst Biol,* 58**,** 184-198.

BRIGHT, D. E. 2014. A Catalog of Scolytidae and Platypodidae (Coleoptera), Supplement 3 (2000-2010), with notes on subfamily and tribal reclassifications. *Insecta Mundi,* 356**,** 1-336.

BROWN, A. M., WASALA, S. K., HOWE, D. K., PEETZ, A. B., ZASADA, I. A. & DENVER, D. R. 2016. Genomic evidence for plant-parasitic nematodes as the earliest Wolbachia hosts. *Sci Rep,* 6**,** 34955.

CAMERON, S. A. & MARDULYN, P. 2003. The major opsin gene is useful for inferring higher level phylogenetic relationships of the corbiculate bees. *Mol Phylogenet Evol,* 28**,** 610-630.

CARMEAN, D., KIMSEY, L. S. & BERBEE, M. L. 1992. 18S rDNA sequences and the holometabolous insects. *Mol Phylogenet Evol,* 1**,** 270-280.

CARRILLO, D., CRUZ, L. F., KENDRA, P. E., NARVAEZ, T. I., MONTGOMERY, W. S., MONTERROSO, A., DE GRAVE, C. & COOPERBAND, M. F. 2016. Distribution, Pest Status and Fungal Associates of Euwallacea nr. fornicatus in Florida Avocado Groves. *Insects,* 7.

CATERINO, M. S., CHO, S. & SPERLING, F. A. 2000. The current state of insect molecular systematics: a thriving Tower of Babel. *Annu Rev Entomol,* 45**,** 1-54.
CHE, L. H., ZHANG, S. Q., LI, Y., LIANG, D., PANG, H., SLIPINSKI, A. & ZHANG, P. 2017. Genome-Wide Survey of Nuclear Protein-Coding Markers for Beetle Phylogenetics and Their Application in Resolving both Deep and Shallow-Level Divergences. *Mol Ecol Resour***,** 1342-1358.
CHO, S. W., MITCHELL, A., REGIER, J. C., MITTER, C., POOLE, R. W., FRIEDLANDER, T. P. & ZHAO, S. W. 1995. A highly conserved nuclear gene for low-level phylogenetics - Elongation Factor-1-alpha recovers morphology-based tree for Heliothine moths. *Mol Biol Evol,* 12**,** 650-656.
CLARK, E. L., KARLEY, A. J. & HUBBARD, S. F. 2010. Insect endosymbionts: manipulators of insect herbivore trophic interactions? *Protoplasma,* 244**,** 25-51.
COGNATO, A. I. 2013. Molecular phylogeny and taxonomic review of Premnobiini Browne, 1962 (Coleoptera: Curculionidae: Scolytinae). *Front Ecol Evol*.
COGNATO, A. I. & SPERLING, F. A. 2000. Phylogeny of Ips DeGeer species (Coleoptera: scolytidae) inferred from mitochondrial cytochrome oxidase I DNA sequence. *Mol Phylogenet Evol,* 14**,** 445-60.
COMANDATORE, F., CORDAUX, R., BANDI, C., BLAXTER, M., DARBY, A., MAKEPEACE, B. L., MONTAGNA, M. & SASSERA, D. 2015. Supergroup C Wolbachia, mutualist symbionts of filarial nematodes, have a distinct genome structure. *Open Biol,* 5**,** 150099.
CONORD, C., DESPRES, L., VALLIER, A., BALMAND, S., MIQUEL, C., ZUNDEL, S., LEMPERIERE, G. & HEDDI, A. 2008. Long-term evolutionary stability of bacterial endosymbiosis in curculionoidea: additional evidence of symbiont replacement in the dryophthoridae family. *Mol Biol Evol,* 25**,** 859-68.
CORREA, A. S., ORLANDO DE OLIVEIRA, L., BRAGA, L. S. & GUEDES, R. N. 2013. Distribution of the related weevil species Sitophilus oryzae and S. zeamais in Brazil. *Insect Sci,* 20**,** 763-770.
CRAMPTON-PLATT, A., TIMMERMANS, M. J., GIMMEL, M. L., KUTTY, S. N., COCKERILL, T. D., VUN KHEN, C. & VOGLER, A. P. 2015. Soup to Tree: The Phylogeny of Beetles Inferred by Mitochondrial Metagenomics of a Bornean Rainforest Sample. *Mol Biol Evol,* 32**,** 2302-2316.
CRAMPTON, G. C. 1938. The Interrelationships and Lines of Descent of Living Insects. *Psyche,* 45**,** 165-181.
CROWSON, R. A. 1955. The natural classification of the families of Coleoptera. *Nathaniel Lloyd & Co., London***,** 187.
CRUAUD, A., UNDERHILL, J. G., HUGUIN, M., GENSON, G., JABBOUR-ZAHAB, R., TOLLEY, K. A., RASPLUS, J. Y. & VAN NOORT, S. 2013. A multilocus phylogeny of the world Sycoecinae fig wasps (Chalcidoidea: Pteromalidae). *PLoS One,* 8**,** e79291.
CRYAN, J. R., LIEBHERR, J. K., FETZNER, J. W., JR. & WHITING, M. F. 2001. Evaluation of relationships within the endemic Hawaiian Platynini (Coleoptera: Carabidae) based on molecular and morphological evidence. *Mol Phylogenet Evol,* 21**,** 72-85.
DANFORTH, B. N., BRADY, S. G., SIPES, S. D. & PEARSON, A. 2004. Single-copy nuclear genes recover cretaceous-age divergences in bees. *Syst Biol,* 53**,** 309-26.
DANFORTH, B. N., CARDINAL, S., PRAZ, C., ALMEIDA, E. A. & MICHEZ, D. 2013. The impact of molecular data on our understanding of bee phylogeny and evolution. *Annu Rev Entomol,* 58**,** 57-78.
DANFORTH, B. N., CONWAY, L. & JI, S. 2003. Phylogeny of eusocial Lasioglossum reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). *Syst Biol,* 52**,** 23-36.
DANFORTH, B. N. & JI, S. 1998. Elongation factor-1 alpha occurs as two copies in bees: implications for phylogenetic analysis of EF-1 alpha sequences in insects. *Mol Biol Evol,* 15**,** 225-35.
DAVIS, S. R. 2014. Morphology, phylogeny, and evolutionary development in the weevils (Insecta: Coleoptera: Curculionoidea). *PhD Thesis,* University of Kansas**,** 1-482.
DOBLER, S. & MULLER, J. K. 2000. Resolving phylogeny at the family level by mitochondrial cytochrome oxidase sequences: phylogeny of carrion beetles (Coleoptera, Silphidae). *Mol Phylogenet Evol,* 15**,** 390-402.

DOLEZELOVA, E., ZUROVEC, M., BOHMOVA, M. & SEHNAL, F. 2006. Use of two transcription starts in the G6PD gene of the bark beetle Ips typographus. *Insect Mol Biol,* 15**,** 25-32.

EDWARDS, S. V., LIU, L. & PEARL, D. K. 2007. High-resolution species trees without concatenation. *Proc Natl Acad Sci U S A,* 104**,** 5936-5941.

FABRYOVA, A., KOSTOVCIK, M., DIEZ-MENDEZ, A., JIMENEZ-GOMEZ, A., CELADOR-LERA, L., SAATI-SANTAMARIA, Z., SECHOVCOVA, H., MENENDEZ, E., KOLARIK, M. & GARCIA-FRAILE, P. 2017. On the bright side of a forest pest-the metabolic potential of bark beetles' bacterial associates. *Sci Total Environ,* 619-620**,** 9-17.

FANG, Q. Q., CHO, S., REGIER, J. C., MITTER, C., MATTHEWS, M., POOLE, R. W., FRIEDLANDER, T. P. & ZHAO, S. 1997. A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). *Syst Biol,* 46**,** 269-283.

FARRELL, B. D. 1998. "Inordinate Fondness" explained: why are there so many beetles? *Science,* 281**,** 555-559.

FARRELL, B. D., SEQUEIRA, A. S., O'MEARA, B. C., NORMARK, B. B., CHUNG, J. H. & JORDAL, B. H. 2001. The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution,* 55**,** 2011-2027.

FETTIG, C. J., KLEPZIG, K. D., BILLINGS, R. F., MUNSON, A. S., NEBEKER, T. E., NEGRON, J. F. & NOWAK, J. T. 2007. The effectiveness of vegetation management practices for prevention and control of bark beetle infestations in coniferous forests of the western and southern United States. *For Ecol Manage,* 238**,** 24 -53.

FIELD, K. G., OLSEN, G. J., LANE, D. J., GIOVANNONI, S. J., GHISELIN, M. T., RAFF, E. C., PACE, N. R. & RAFF, R. A. 1988. Molecular phylogeny of the animal kingdom. *Science,* 239**,** 748-753.

FRIEDLANDER, T. P., HORST, K. R., REGIER, J. C., MITTER, C., PEIGLER, R. S. & FANG, Q. Q. 1998. Two nuclear genes yield concordant relationships within Attacini (Lepidoptera: Saturniidae). *Mol Phylogenet Evol,* 9**,** 131-140.

FRIEDLANDER, T. P., REGIER, J. C. & MITTER, C. 1992. Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. *Syst Biol***,** 483–490.

GIBSON, J. F., KELSO, S., JACKSON, M. D., KITS, J. H., MIRANDA, G. F. G. & SKEVINGTON, J. H. 2011. Diptera-Specific Polymerase Chain Reaction Amplification Primers of Use in Molecular Phylogenetic Research. *Ann Entomol Soc Am,* 104**,** 976-997.

GILLETT, C. P., CRAMPTON-PLATT, A., TIMMERMANS, M. J., JORDAL, B. H., EMERSON, B. C. & VOGLER, A. P. 2014. Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). *Mol Biol Evol,* 31**,** 2223-2237.

GOHLI, J., KIRKENDALL, L. R., SMITH, S. M., COGNATO, A. I., HULCR, J. & JORDAL, B. H. 2017. Biological factors contributing to bark and ambrosia beetle species diversification. *Evolution*.

GUNTER, N., OBERPRIELER, R. & CAMERON, S. 2015. Molecular phylogenetics of Australian weevils (Coleoptera: Curculionoidea): exploring relationships in a hyperdiverse lineage through comparison of independent analyses. *Aust Entomol,* 55**,** 217–233.

GUNTER, N. L., LEVKANICOVA, Z., WEIR, T. H., SLIPINSKI, A., CAMERON, S. L. & BOCAK, L. 2014. Towards a phylogeny of the Tenebrionoidea (Coleoptera). *Mol Phylogenet Evol,* 79**,** 305-312.

HAAS, F. & KUKALOVA-PECK, J. 2001. Dermaptera hindwing structure and folding: New evidence for familial, ordinal and superordinal relationships within Neoptera (Insecta). *European Journal of Entomology,* 98**,** 445-509.

HADDAD, S., SHIN, S., LEMMON, A. R., LEMMON, E. M., SVACHA, P., FARRELL, B., ŚLIPIŃSKI, A., WINDSOR, D. & MCKENNA, D. D. 2017. Anchored hybrid enrichment provides new insights into the phylogeny and evolution of longhorned beetles (Cerambycidae). *Syst Entomol*.

HARAN, J., TIMMERMANS, M. J. T. N. & VOGLER, A. P. 2013. Mitogenome sequences stabilize the phylogenetics of weevils (Curculionoidea) and establish the monophyly of larval ectophagy. *Mol Phylogenet Evol,* 67**,** 156-166.

HARDY, N. B. 2007. Phylogenetic utility of dynamin and triose phosphate isomerase. *Systematic Entomology,* 32**,** 396-403.

HEDTKE, S. M., PATINY, S. & DANFORTH, B. N. 2013. The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography. *BMC Evol Biol,* 13**,** 138.
HENNIG, W. 1969. Die Stammesgeschichte der Insekten. *Verlag von Waldemar Kramer,* Frankfurt am Main.
HERNANDEZ-GARCIA, J. A., BRIONES-ROBLERO, C. I., RIVERA-ORDUNA, F. N. & ZUNIGA, G. 2017. Revealing the gut bacteriome of Dendroctonus bark beetles (Curculionidae: Scolytinae): diversity, core members and co-evolutionary patterns. *Sci Rep,* 7**,** 13864.
HOVEMANN, B., RICHTER, S., WALLDORF, U. & CZIEPLUCH, C. 1988. Two genes encode related cytoplasmic elongation factors 1 alpha (EF-1 alpha) in Drosophila melanogaster with continuous and stage specific expression. *Nucleic Acids Res,* 16**,** 3175-3194.
HOWLAND, D. E. & HEWITT, G. M. 1995. Phylogeny of the Coleoptera based on mitochondrial cytochrome oxidase I sequence data. *Insect Mol Biol,* 4**,** 203-215.
HULCR, J., ATKINSON, T., COGNATO, A., JORDAL, B. & MCKENNA, D. 2015. Morphology, Taxonomy and Phylogenetics of Bark Beetles. *In:* PRESS, A. (ed.) *Bark Beetles: Biology and Ecology of Native and Invasive Species.* Fernando E. Vega and Richard W. Hofstetter.
HULCR, J. & DUNN, R. R. 2011. The sudden emergence of pathogenicity in insect-fungus symbioses threatens naive forest ecosystems. *Proc Biol Sci,* 278**,** 2866-2873.
HULCR, J., MOGIA, M., ISUA, B. & NOVOTNY, V. 2007. Host specificity of ambrosia and bark beetles (Col., Curculionidae: Scolytinae and Platypodinae) in a New Guinea rainforest. *Ecological Entomology,* 32**,** 762-772.
HULCR, J., ROUNTREE, N. R., DIAMOND, S. E., STELINSKI, L. L., FIERER, N. & DUNN, R. R. 2012. Mycangia of ambrosia beetles host communities of bacteria. *Microb Ecol,* 64**,** 784-793.
HULCR, J. & STELINSKI, L. L. 2017. The Ambrosia Symbiosis: From Evolutionary Ecology to Practical Management. *Annu Rev Entomol,* 62**,** 285-303.
ILINSKY, Y. & KOSTERIN, O. E. 2017. Molecular diversity of Wolbachia in Lepidoptera: Prevalent allelic content and high recombination of MLST genes. *Mol Phylogenet Evol,* 109**,** 164-179.
JORDAL, B. H. 2002. Elongation Factor 1 α resolves the monophyly of the haplodiploid ambrosia beetles Xyleborini (Coleoptera: Curculionidae). *Insect Mol Biol,* 11.
JORDAL, B. H. 2007. Reconstructing the phylogeny of Scolytinae and close allies: major obstacles and prospects for a solution. *U S Forest Service RMRS-P,* 45**,** 3-9.
JORDAL, B. H. 2015. Molecular phylogeny and biogeography of the weevil subfamily Platypodinae reveals evolutionarily conserved range patterns. *Mol Phylogenet Evol,* 92**,** 294-307.
JORDAL, B. H. & COGNATO, A. I. 2012. Molecular phylogeny of bark and ambrosia beetles reveals multiple origins of fungus farming during periods of global warming. *BMC Evol Biol,* 12**,** 133.
JORDAL, B. H. & KAIDEL, J. 2016. Phylogenetic analysis of Micracidini bark beetles (Coleoptera: Curculionidae) demonstrates a single trans-Atlantic disjunction and inclusion of *Cactopinus* in the New World clade. *Can Entomol***,** 1-18.
JORDAL, B. H., SEQUEIRA, A. S. & COGNATO, A. I. 2011. The age and phylogeny of wood boring weevils and the origin of subsociality. *Mol Phylogenet Evol,* 59**,** 708-724.
JORDAL, B. H., SMITH, S. M. & COGNATO, A. I. 2014. Classification of weevils as a data-driven science: leaving opinion behind. *ZooKeys,* 439**,** 1-18.
KAWAHARA, A. Y. & BREINHOLT, J. W. 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proc Biol Sci,* 281**,** 20140970.
KAWASAKI, Y., SCHULER, H., STAUFFER, C., LAKATOS, F. & KAJIMURA, H. 2016. Wolbachia endosymbionts in haplodiploid and diploid scolytine beetles (Coleoptera: Curculionidae: Scolytinae). *Environ Microbiol Rep,* 6**,** 680–688.
KELLNER, R. L. 2002. Molecular identification of an endosymbiotic bacterium associated with pederin biosynthesis in Paederus sabaeus (Coleoptera: Staphylinidae). *Insect Biochem Mol Biol,* 32**,** 389-395.
KIKUCHI, Y. 2009. Endosymbiotic bacteria in insects: their diversity and culturability. *Microbes Environ,* 24**,** 195-204.

KIM, S. I. & FARRELL, B. D. 2015. Phylogeny of world stag beetles (Coleoptera: Lucanidae) reveals a Gondwanan origin of Darwin's stag beetle. *Mol Phylogenet Evol,* 86**,** 35-48.
KIRKENDALL, L. 1983. The evolution of mating systems in bark and ambrosia beetles (Coleoptera: Scolytidae and Platypodidae). *Zool J Linnean Soc,* 77**,** 293-352.
KIRKENDALL, L., BIEDERMANN, P. & JORDAL, B. 2015. Evolution and diversity of bark and ambrosia beetles. *In:* PRESS, A. (ed.) *Bark Beetles: Biology and Ecology of Native and Invasive Species.* Fernando E. Vega and Richard W. Hofstetter.
KIRKENDALL, L. & FACCOLI, M. 2010. Bark beetles and pinhole borers (Curculionidae, Scolytinae, Platypodinae) alien to Europe. *Zookeys***,** 227-51.
KIRKENDALL, L. R. K., D S, RAFFA, K F 1997. Interactions among males, females and offspring in bark and ambrosia beetles: The significance of living in tunnels for the evolution of social behavior. *In:* CRESPI, J. C. B. (ed.) *The Evolution of Social Behaviour in Insects and Arachnids.* Cambridge University Press.
KODANDARAMAIAH, U., PENA, C., BRABY, M. F., GRUND, R., MULLER, C. J., NYLIN, S. & WAHLBERG, N. 2010. Phylogenetics of Coenonymphina (Nymphalidae: Satyrinae) and the problem of rooting rapid radiations. *Mol Phylogenet Evol,* 54**,** 386-394.
KOK, L. T., NORRIS, D. M. & CHU, H. M. 1970. Sterol metabolism as a basis for a mutualistic symbiosis. *Nat Rev Microbiol,* 225**,** 661-662.
KOSTOVCIK, M., BATEMAN, C. C., KOLARIK, M., STELINSKI, L. L., JORDAL, B. H. & HULCR, J. 2015. The ambrosia symbiosis is specific in some species and promiscuous in others: evidence from community pyrosequencing. *ISME J,* 9**,** 126-138.
KRAUSS, V., THUMMLER, C., GEORGI, F., LEHMANN, J., STADLER, P. F. & EISENHARDT, C. 2008. Near intron positions are reliable phylogenetic markers: an application to holometabolous insects. *Mol Biol Evol,* 25**,** 821-830.
KURIWADA, T., HOSOKAWA, T., KUMANO, N., SHIROMOTO, K., HARAGUCHI, D. & FUKATSU, T. 2010. Biological role of Nardonella endosymbiont in its weevil host. *PLoS One,* 5.
KUSCHEL, G. 1995. A phylogenetic classification of Curculionoidea to families and subfamilies. *Mem Ent Soc Wash,* 14**,** 5-33.
KUSCHEL, G., LESCHEN, R. A. B. & ZIMMERMAN, E. C. 2000. Platypodidae under scrutiny. *Invertebrate Taxonomy,* 14**,** 771-805.
LACHOWSKY, L., LALONDE, R. & REID, M. 2015. Can Wolbachia (Rickettsiaceae) explain female-biased sex ratios in mountain pine beetles (Coleoptera: Curculionidae)? *Can Entomol,* 47**,** 732-736.
LAWRENCE, J. F. & NEWTON, A. F. J. 1995. Families and subfamilies of Coleoptera (with selected genera, notes, references and data on family-group names). *In:* (EDS.), J. P. S. A. S. (ed.) *Biology, Phylogeny, and Classification of Coleoptera: Papers Celebrating the 80th Birthday of Roy A. Crowson.* Museum i Instytut Zoologii PAN.
LAWRENCE, J. F., SLIPINSKI, A., SEAGO, A. E., THAYER, M. K., NEWTON, A. F. & MARVALDI, A. E. 2011. Phylogeny of the Coleoptera based on morphological characters of adults and larvae. *Annales Zoologici,* 61**,** 1-217.
LEFEVRE, C., CHARLES, H., VALLIER, A., DELOBEL, B., FARRELL, B. & HEDDI, A. 2004. Endosymbiont phylogenesis in the dryophthoridae weevils: evidence for bacterial replacement. *Mol Biol Evol,* 21**,** 965-973.
LI, Y., BATEMAN, C. C., SKELTON, J., JUSINO, M. A., NOLEN, Z. J., SIMMONS, D. R. & HULCR, J. 2017. Wood decay fungus Flavodon ambrosius (Basidiomycota: Polyporales) is widely farmed by two genera of ambrosia beetles. *Fungal Biol,* 121**,** 984-989.
LINNAKOSKI, R., DE BEER, Z. W., NIEMELA, P. & WINGFIELD, M. J. 2012. Associations of Conifer-Infesting Bark Beetles and Fungi in Fennoscandia. *Insects,* 3**,** 200-227.
LIU, H. & BECKENBACH, A. T. 1992. Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of insects. *Mol Phylogenet Evol,* 1**,** 41-52.
MADDISON, D. R. 2012. Phylogeny of Bembidion and related ground beetles (Coleoptera: Carabidae: Trechinae: Bembidiini: Bembidiina). *Mol Phylogenet Evol,* 63**,** 533-576.

MADDISON, D. R., TOLEDANO, L., SALLENAVE, S. & ROIG-JUNENT, S. 2013. Phylogenetic relationships of the South American ground beetle subgenus Chilioperyphus Jeannel (Coleoptera: Carabidae: Trechinae: Bembidiini: Bembidion Latreille). *Zootaxa,* 3636**,** 547-560.

MAEKAWA, K., KON, M., ARAYA, K. & MATSUMOTO, T. 2001. Phylogeny and biogeography of wood-feeding cockroaches, genus Salganea Stal (Blaberidae: Panesthiinae), in Southeast Asia based on mitochondrial DNA sequences. *J Mol Evol,* 53**,** 651-659.

MARDULYN, P. & CAMERON, S. A. 1999. The major opsin in bees (Insecta: Hymenoptera): A promising nuclear gene for higher level phylogenetics. *Mol Phylogenet Evol,* 12**,** 168-176.

MARDULYN, P. & WHITFIELD, J. B. 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.

MARIÑO, Y. A., VEGA, V. J., GARCIA, J. M., VERLE RODRIGUES, J. C., GARCIA, N. M. & BAYMAN, P. 2017a. The Coffee Berry Borer (Coleoptera: Curculionidae) in Puerto Rico: Distribution, Infestation, and Population per Fruit. *J Insect Sci,* 17.

MARIÑO, Y. A., VERLE RODRIGUES, J. C. & BAYMAN, P. 2017b. Wolbachia Affects Reproduction and Population Dynamics of the Coffee Berry Borer (Hypothenemus hampei): Implications for Biological Control. *Insects,* 8.

MARTINEZ-NAVARRO, E. M., GALIAN, J. & SERRANO, J. 2005. Phylogeny and molecular evolution of the tribe Harpalini (Coleoptera, Carabidae) inferred from mitochondrial cytochrome-oxidase I. *Mol Phylogenet Evol,* 35**,** 127-146.

MARVALDI, A. E. 1997. Higher level phylogeny of Curculionidae (Coleoptera: Curculionoidea) based mainly on larval characters, with special reference to broad-nosed weevils. *Cladistics,* 13**,** 285-312.

MARVALDI, A. E. & MORRONE, J. J. 2000. Phylogenetic systematics of weevils (Coleoptera: Curculionoidea): a reappraisal based on larval and adult morphology. *Insect Syst Evol,* 31**,** 43-58.

MARVALDI, A. E., SEQUEIRA, A. S., O'BRIEN, C. W. & FARRELL, B. D. 2002. Molecular and morphological phylogenetics of weevils (Coleoptera, Curculionoidea): do niche shifts accompany diversification? *Syst Biol,* 51**,** 761-785.

MASSON, F., MONE, Y., VIGNERON, A., VALLIER, A., PARISOT, N., VINCENT-MONEGAT, C., BALMAND, S., CARPENTIER, M. C., ZAIDMAN-REMY, A. & HEDDI, A. 2015a. Weevil endosymbiont dynamics is associated with a clamping of immunity. *BMC Genomics,* 16**,** 819.

MASSON, F., VALLIER, A., VIGNERON, A., BALMAND, S., VINCENT-MONEGAT, C., ZAIDMAN-REMY, A. & HEDDI, A. 2015b. Systemic infection generates a local-like immune response of the bacteriome organ in insect symbiosis. *J Innate Immun,* 7**,** 290-301.

MAUS, C., PESCHKE, K. & DOBLER, S. 2001. Phylogeny of the genus Aleochara inferred from mitochondrial cytochrome oxidase sequences (Coleoptera: Staphylinidae). *Mol Phylogenet Evol,* 18**,** 202-216.

MAY, B. M. 1993. Larvae of Curculionoidea (Insecta: Coleoptera): a systematic overview. 28.

MAYERS, C. G., MCNEW, D. L., HARRINGTON, T. C., ROEPER, R. A., FRAEDRICH, S. W., BIEDERMANN, P. H., CASTRILLO, L. A. & REED, S. E. 2015. Three genera in the Ceratocystidaceae are the respective symbionts of three independent lineages of ambrosia beetles with large, complex mycangia. *Fungal Biol,* 119**,** 1075-1092.

MCKENNA, D. D., SEQUEIRA, A. S., MARVALDI, A. E. & FARRELL, B. D. 2009. Temporal lags and overlap in the diversification of weevils and flowering plants. *Proc Natl Acad Sci U S A,* 106**,** 7083-7088.

MCKENNA, D. D., WILD, A. L., KANDA, K., BELLAMY, C. L., BEUTEL, R. G., CATERINO, M. S., FARNUM, C. W., HAWKS, D. C., IVIE, M. A., JAMESON, M. L., LESCHEN, R. A. B., MARVALDI, A. E., MCHUGH, J. V., NEWTON, A. F., ROBERTSON, J. A., THAYER, M. K., WHITING, M. F., LAWRENCE, J. F., ŚLIPIŃSKI, A., MADDISON, D. R. & FARRELL, B. D. 2015. The beetle tree of life reveals that

Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. *Syst Ent,* 40**,** 835-880.

MILLER, K. E., HOPKINS, K., INWARD, D. J. & VOGLER, A. P. 2016. Metabarcoding of fungal communities associated with bark beetles. *Ecol Evol,* 6**,** 1590-1600.

MISOF, B., LIU, S., MEUSEMANN, K., PETERS, R. S., DONATH, A., MAYER, C., FRANDSEN, P. B., WARE, J., FLOURI, T., BEUTEL, R. G., NIEHUIS, O., PETERSEN, M., IZQUIERDO-CARRASCO, F., WAPPLER, T., RUST, J., ABERER, A. J., ASPOCK, U., ASPOCK, H., BARTEL, D., BLANKE, A., BERGER, S., BOHM, A., BUCKLEY, T. R., CALCOTT, B., CHEN, J., FRIEDRICH, F., FUKUI, M., FUJITA, M., GREVE, C., GROBE, P., GU, S., HUANG, Y., JERMIIN, L. S., KAWAHARA, A. Y., KROGMANN, L., KUBIAK, M., LANFEAR, R., LETSCH, H., LI, Y., LI, Z., LI, J., LU, H., MACHIDA, R., MASHIMO, Y., KAPLI, P., MCKENNA, D. D., MENG, G., NAKAGAKI, Y., NAVARRETE-HEREDIA, J. L., OTT, M., OU, Y., PASS, G., PODSIADLOWSKI, L., POHL, H., VON REUMONT, B. M., SCHUTTE, K., SEKIYA, K., SHIMIZU, S., SLIPINSKI, A., STAMATAKIS, A., SONG, W., SU, X., SZUCSICH, N. U., TAN, M., TAN, X., TANG, M., TANG, J., TIMELTHALER, G., TOMIZUKA, S., TRAUTWEIN, M., TONG, X., UCHIFUNE, T., WALZL, M. G., WIEGMANN, B. M., WILBRANDT, J., WIPFLER, B., WONG, T. K., WU, Q., WU, G., XIE, Y., YANG, S., YANG, Q., YEATES, D. K., YOSHIZAWA, K., ZHANG, Q., ZHANG, R., ZHANG, W., ZHANG, Y., ZHAO, J., ZHOU, C., ZHOU, L., ZIESMANN, T., ZOU, S., LI, Y., XU, X., ZHANG, Y., YANG, H., WANG, J., WANG, J., KJER, K. M., et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science,* 346**,** 763-777.

MIYAMOTO, M. M. & FITCH, W. M. 1995. Testing Species Phylogenies and Phylogenetic Methods with Congruence. *Syst Biol,* 44**,** 64-76.

MORIMOTO, K. & KOJIMA, H. 2003. Morphologic Characters of the Weevil Head and Phylogenetic Implications (Coleoptera, Curculionoidea). *Esakia: occasional papers of the Hikosan Biological Laboratory in Entomology,* 43**,** 133-169.

MUELLER, U. G. & GERARDO, N. 2002. Fungus-farming insects: multiple origins and diverse evolutionary histories. *Proc Natl Acad Sci U S A,* 99**,** 15247-9.

MUTANEN, M., WAHLBERG, N. & KAILA, L. 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. *Proc Biol Sci,* 277**,** 2839-2848.

NIKOH, N., HOSOKAWA, T., MORIYAMA, M., OSHIMA, K., HATTORI, M. & FUKATSU, T. 2014. Evolutionary origin of insect-Wolbachia nutritional mutualism. *Proc Natl Acad Sci U S A,* 111**,** 10257-10262.

NORMARK, B. B., JORDAL, B. H. & FARRELL, B. 1999. Origin of a haplodiploid beetle lineage. *Proc R Soc Lond B Biol Sci***,** 2253-2259.

OBERPRIELER, R. G., MARVALDI, A. E. & ANDERSON, R. S. 2007. Weevils, weevils, weevils everywhere. *Zootaxa***,** 491-520.

OLMSTEAD, R. G. & SWEERE, J. A. 1994. Combining Data in Phylogenetic Systematics: An Empirical Approach Using Three Molecular Data Sets in the Solanaceae. *Syst Biol,* 43**,** 467-481.

PELANDAKIS, M., HIGGINS, D. G. & SOLIGNAC, M. 1991. Molecular phylogeny of the subgenus Sophophora of Drosophila derived from large subunit of ribosomal RNA sequences. *Genetica,* 84**,** 87-94.

PERSSON, Y., VASAITIS, R., LANGSTROM, B., OHRN, P., IHRMARK, K. & STENLID, J. 2009. Fungi vectored by the bark beetle Ips typographus following hibernation under the bark of standing trees and in the forest litter. *Microb Ecol,* 58**,** 651-659.

PLANTARD, O., BOUJU-ALBERT, A., MALARD, M. A., HERMOUET, A., CAPRON, G. & VERHEYDEN, H. 2012. Detection of Wolbachia in the tick Ixodes ricinus is due to the presence of the hymenoptera endoparasitoid Ixodiphagus hookeri. *PLoS One,* 7**,** e30692.

POLAK, S., DELIĆ, T., KOSTANJSEK, R. & TRONTELJ, P. 2016. Molecular phylogeny of the cave beetle genus Hadesia (Coleoptera: Leiodidae: Cholevinae: Leptodirini), with a description of a new species from Montenegro. *Arthropod Syst Phylogeny,* 7412**,** 241-254.

POPA, V., DEZIEL, E., LAVALLEE, R., BAUCE, E. & GUERTIN, C. 2012. The complex symbiotic relationships of bark beetles with microorganisms: a potential practical approach for biological control in forestry. *Pest Manag Sci,* 68**,** 963-975.

RASSATI, D., FACCOLI, M., BATTISTI, A. & MARINI, L. 2016a. Habitat and climatic preferences drive invasions of non-native ambrosia beetles in deciduous temperate forests. *Biol Invasion,* 18**,** 764-779.

RASSATI, D., LIEUTIER, F. & FACCOLI, M. 2016b. Alien Wood-Boring Beetles in Mediterranean Regions. *In:* TIMOTHY D. PAINE, F. L. (ed.) *Insects and Diseases of Mediterranean Forest Systems.* Springer International Publishing.

REGIER, J. C., FANG, Q. Q., MITTER, C., PEIGLER, R. S., FRIEDLANDER, T. P. & SOLIS, M. A. 1998. Evolution and phylogenetic utility of the period gene in Lepidoptera. *Mol Biol Evol,* 15**,** 1172-82.

REGIER, J. C., MITTER, C., ZWICK, A., BAZINET, A. L., CUMMINGS, M. P., KAWAHARA, A. Y., SOHN, J. C., ZWICKL, D. J., CHO, S., DAVIS, D. R., BAIXERAS, J., BROWN, J., PARR, C., WELLER, S., LEES, D. C. & MITTER, K. T. 2013. A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). *PLoS One,* 8**,** e58568.

RIEDEL, A., TÄNZLER, R., PONS, J., SUHARDJONO, Y. R. & BALKE, M. 2016. Large-scale molecular phylogenyof Cryptorhynchinae (Coleoptera, Curculionidae) from multiple genes suggests American origin and later Australian radiation. *Syst Entomol,* 41**,** 492-503.

ROKAS, A. & CARROLL, S. B. 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol Biol Evol,* 22**,** 1337-1344.

ROKAS, A., WILLIAMS, B. L., KING, N. & CARROLL, S. B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature,* 425**,** 798-804.

RONQUIST, F. & HUELSENBECK, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* 19**,** 1572-1574.

RUANE, S., RAXWORTHY, C. J., LEMMON, A. R., LEMMON, E. M. & BURBRINK, F. T. 2015. Comparing species tree estimation with large anchored phylogenomic and small Sanger-sequenced molecular datasets: an empirical study on Malagasy pseudoxyrhophiine snakes. *BMC Evol Biol,* 15**,** 221.

RUIZ, C., JORDAL, B. & SERRANO, J. 2009. Molecular phylogeny of the tribe Sphodrini (Coleoptera: Carabidae) based on mitochondrial and nuclear markers. *Mol Phylogenet Evol,* 50**,** 44-58.

RUIZ, C., JORDAL, B. H., EMERSON, B. C., WILL, K. W. & SERRANO, J. 2010. Molecular phylogeny and Holarctic diversification of the subtribe Calathina (Coleoptera: Carabidae: Sphodrini). *Mol Phylogenet Evol,* 55**,** 358-371.

RUSSO, C. A., TAKEZAKI, N. & NEI, M. 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol,* 12**,** 391-404.

RUTSCHMANN, S., DETERING, H., SIMON, S., FUNK, D. H., GATTOLLIAT, J. L., HUGHES, S. J., RAPOSEIRO, P. M., DESALLE, R., SARTORI, M. & MONAGHAN, M. T. 2017. Colonization and diversification of aquatic insects on three Macaronesian archipelagos using 59 nuclear loci derived from a draft genome. *Mol Phylogenet Evol,* 107**,** 27-38.

SAHOO, R. K., WARREN, A. D., WAHLBERG, N., BROWER, A. V., LUKHTANOV, V. A. & KODANDARAMAIAH, U. 2016. Ten genes and two topologies: an exploration of higher relationships in skipper butterflies (Hesperiidae). *PeerJ,* 4**,** e2653.

SASAKI, G., ISHIWATA, K., MACHIDA, R., MIYATA, T. & SU, Z. H. 2013. Molecular phylogenetic analyses support the monophyly of Hexapoda and suggest the paraphyly of Entognatha. *BMC Evol Biol,* 13**,** 236.

SCHEDL, W. 1962. Ein beitrag zur kenntnis der Pilzubertragungsweisebei xylomycetophagen Scolytiden. *Osterreichische der Akademie der Wissenschaften mathematisch-naturwissenschaftliche Klasse,* 171**,** 363-387.

SCHEFFER, S. J. & WIEGMANN, B. M. 2000. Molecular phylogenetics of the holly leafminers (Diptera: Agromyzidae: Phytomyza): species limits, speciation, and dietary specialization. *Mol Phylogenet Evol,* 17**,** 244-255.

SCHUBERT, M., HOLLAND, L. Z., HOLLAND, N. D. & JACOBS, D. K. 2000. A phylogenetic tree of the Wnt genes based on all available full-length sequences, including five from the cephalochordate amphioxus. *Mol Biol Evol,* 17**,** 1896-1903.

SENATORE, G. L., ALEXANDER, E. A., ADLER, P. H. & MOULTON, J. K. 2014. Molecular systematics of the Simulium jenningsi species group (Diptera: Simuliidae), with three new fast-evolving nuclear genes for phylogenetic inference. *Mol Phylogenet Evol,* 75**,** 138-148.

SEQUEIRA, A. & FARRELL, B. 2001. Evolutionary origins of Gondwanan interactions: How old are Araucaria beetle herbivores? *Biol J Linnean Soc,* 74**,** 459-474.

SEQUEIRA, A. S., NORMARK, B. B. & FARRELL, B. D. 2000. Evolutionary assembly of the conifer fauna: distinguishing ancient from recent associations in bark beetles. *Proc Biol Sci,* 267**,** 2359-2366.

SEVCIK, J., KASPRAK, D., MANTIC, M., FITZGERALD, S., SEVCIKOVA, T., TOTHOVA, A. & JASCHHOF, M. 2016. Molecular phylogeny of the megadiverse insect infraorder Bibionomorpha sensu lato (Diptera). *PeerJ,* 4**,** e2563.

SHEN, X.-X., HITTINGER, C. T. & ROKAS, A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology & Evolution,* 1**,** 0126.

SHULL, V. L., VOGLER, A. P., BAKER, M. D., MADDISON, D. R. & HAMMOND, P. M. 2001. Sequence alignment of 18S ribosomal RNA and the basal relationships of Adephagan beetles: evidence for monophyly of aquatic families and the placement of Trachypachidae. *Syst Biol,* 50**,** 945-969.

SIKES, D. S. & VENABLES, C. 2013. Molecular phylogeny of the burying beetles (Coleoptera: Silphidae: Nicrophorinae). *Mol Phylogenet Evol,* 69**,** 552-565.

SIX, D. L. 2012. Ecological and Evolutionary Determinants of Bark Beetle -Fungus Symbioses. *Insects,* 3**,** 339-366.

SIX, D. L., HARRINGTON, T. C., STEIMEL, J., MCNEW, D. & PAINE, T. D. 2003. Genetic relationships among *Leptographium terebrantis* and the mycangial fungi of three western *Dendroctonus* bark beetles. *Mycologia,* 95**,** 781-792.

SIX, D. L. & PAINE, T. D. 1996. *Leptographium pyrinum* is a mycangial fungus of *Dendroctonus adjunctus*. *Mycologia***,** 739–744.

SKIDMORE, I. H. & HANSEN, A. K. 2017. The evolutionary development of plant-feeding insects and their nutritional endosymbionts. *Insect Sci,* 24**,** 910–928.

SONG, J. & AHN, K. 2017. Species trees, temporal divergence and historical biogeography of coastal rove beetles (Coleoptera: Staphylinidae) reveal their early Miocene origin and show that most divergence events occurred in the early Pliocene along the Pacific coasts. *Cladistics*.

SONTOWSKI, R., BERNHARD, D., BLEIDORN, C., SCHLEGEL, M. & GERTH, M. 2015. Wolbachia distribution in selected beetle taxa characterized by PCR screens and MLST data. *Ecol Evol,* 5**,** 4345-4353.

SOTA, T. & VOGLER, A. P. 2001. Incongruence of mitochondrial and nuclear gene trees in the Carabid beetles Ohomopterus. *Syst Biol,* 50**,** 39-59.

SUDAKARAN, S., KOST, C. & KALTENPOTH, M. 2017. Symbiont Acquisition and Replacement as a Source of Ecological Innovation. *Trends Microbiol,* 25**,** 375-390.

TANZLER, R., TOUSSAINT, E. F., SUHARDJONO, Y. R., BALKE, M. & RIEDEL, A. 2014. Multiple transgressions of Wallace's Line explain diversity of flightless Trigonopterus weevils on Bali. *Proc Biol Sci,* 281**,** 20132528.

TARASOV, S. & DIMITROV, D. 2016. Multigene phylogenetic analysis redefines dung beetles relationships and classification (Coleoptera: Scarabaeidae: Scarabaeinae). *BMC Evol Biol,* 16**,** 257.

TATARENKOV, A., KWIATOWSKI, J., SKARECKY, D., BARRIO, E. & AYALA, F. J. 1999. On the evolution of Dopa decarboxylase (Ddc) and Drosophila systematics. *J Mol Evol,* 48**,** 445-462.

THOMPSON, R. T. 1992. Observations on the morphology and classification of weevils (Coleoptera, Curculionoidea) with a key to major groups. *J Nat History,* 26**,** 835-891.

TOJU, H., TANABE, A. S., NOTSU, Y., SOTA, T. & FUKATSU, T. 2013. Diversification of endosymbiosis: replacements, co-speciation and promiscuity of bacteriocyte symbionts in weevils. *Isme j,* 7**,** 1378-1390.

TRAUTWEIN, M. D., WIEGMANN, B. M., BEUTEL, R., KJER, K. M. & YEATES, D. K. 2012. Advances in insect phylogeny at the dawn of the postgenomic era. *Annu Rev Entomol,* 57**,** 449-468.

VUATAZ, L., RUTSCHMANN, S., MONAGHAN, M. T. & SARTORI, M. 2016. Molecular phylogeny and timing of diversification in Alpine Rhithrogena (Ephemeroptera: Heptageniidae). *BMC Evol Biol,* 16**,** 194.

WAHLBERG, N., PENA, C., AHOLA, M., WHEAT, C. W. & ROTA, J. 2016. PCR primers for 30 novel gene regions in the nuclear genomes of Lepidoptera. *Zookeys***,** 129-141.

WAHLBERG, N. & WHEAT, C. W. 2008. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of lepidoptera. *Syst Biol,* 57**,** 231-242.

WEIRAUCH, C. & MUNRO, J. B. 2009. Molecular phylogeny of the assassin bugs (Hemiptera: Reduviidae), based on mitochondrial and nuclear ribosomal genes. *Mol Phylogenet Evol,* 53**,** 287-299.

WELLER, S. J., FRIEDLANDER, T. P., MARTIN, J. A. & PASHLEY, D. P. 1992. Phylogenetic studies of ribosomal RNA variation in higher moths and butterflies (Lepidoptera: Ditrysia). *Mol Phylogenet Evol,* 1**,** 312-337.

WERREN, J. H., BALDO, L. & CLARK, M. E. 2008. Wolbachia: Master manipulators of invertebrate biology. *Nat Rev Microbiol* 6**,** 741-751.

WERREN, J. H. & WINDSOR, D. M. 2000. Wolbachia infection frequencies in insects: evidence of a global equilibrium? *Proc Biol Sci,* 267**,** 1277-1285.

WHITING, M. F., CARPENTER, J. C., WHEELER, Q. D. & WHEELER, W. C. 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst Biol,* 46**,** 1-68.

WIEGMANN, B. M., TRAUTWEIN, M. D., KIM, J. W., CASSEL, B. K., BERTONE, M. A., WINTERTON, S. L. & YEATES, D. K. 2009. Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol,* 7**,** 34.

WIENS, J. J. & MORRILL, M. C. 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Syst Biol,* 60**,** 719-731.

WILD, A. L. & MADDISON, D. R. 2008. Evaluating nuclear protein-coding genes for phylogenetic utility in beetles. *Mol Phylogenet Evol,* 48**,** 877-891.

WILLE, A. 1960. The phylogeny and relationships between the insect orders. *Revista de Biología Tropical,* 8**,** 93-123.

WINKLER, I. S., BLASCHKE, J. D., DAVIS, D. J., STIREMAN, J. O., 3RD, O'HARA, J. E., CERRETTI, P. & MOULTON, J. K. 2015. Explosive radiation or uninformative genes? Origin and early diversification of tachinid flies (Diptera: Tachinidae). *Mol Phylogenet Evol,* 88**,** 38-54.

WOOD, S. L. 1982. The bark and ambrosia beetles of North and Central America (Coleoptera: Scolytidae), a taxonomic monograph. *Gr Basin Nat Mem***,** 1-1359.

WOOD, S. L. 1986. A Reclassification of the Genera of Scolytidae (Coleoptera). *Gr Basin Nat Mem,* 10**,** 1-126.

WOOD, S. L. 1993. Revision of the genera of Platypodidae (Coleoptera). *Gr Basin Nat Mem,* 53**,** 259-281.

WOOD, S. L. & BRIGHT, D. E. 1992. A catalog of Scolytidae and Platypodidae (Coleoptera) Part 2: Taxonomic index. *Gr Basin Nat Mem,* 13**,** 1 - 1553.

YENERALL, P., KRUPA, B. & ZHOU, L. 2011. Mechanisms of intron gain and loss in Drosophila. *BMC Evol Biol,* 11**,** 364.

YOUNG, A. D., LEMMON, A. R., SKEVINGTON, J. H., MENGUAL, X., STAHLS, G., REEMER, M., JORDAENS, K., KELSO, S., LEMMON, E. M., HAUSER, M., DE MEYER, M., MISOF, B. & WIEGMANN, B. M. 2016. Anchored enrichment dataset for true flies (order Diptera) reveals insights into the phylogeny of flower flies (family Syrphidae). *BMC Evol Biol,* 16**,** 143.

YUAN, M. L., ZHANG, Q. L., ZHANG, L., GUO, Z. L., LIU, Y. J., SHEN, Y. Y. & SHAO, R. 2016. High-level phylogeny of the Coleoptera inferred with mitochondrial genome sequences. *Mol Phylogenet Evol,* 104**,** 99-111.

ZHANG, H., LIU, N., HANA, Z. & LIU, J. 2016. Phylogenetic analyses and evolutionary timescale of Coleoptera based on mitochondrial sequence. *Biochemi Syst Ecol,* 66**,** 229-238.

ZHANG, X. & ZHOU, H. Z. 2013. How old are the rove beetles (Insecta: Coleoptera: Staphylinidae) and their lineages? Seeking an answer with DNA. *Zoolog Sci,* 30**,** 490-501.

ZHERIKHIN, V. V. & GRATSHEV, V. G. 1995. A comparative study of the hind wing venation of the superfamily Curculionoidea, with phylogenetic implications. *In:* (EDS.), J. P. S. A. S. (ed.) *Biology, phylogeny, and classification of Coleoptera: papers celebrating the 80th birthday of Roy A. Crowson.* Muzeum i Instytut Zoologii PAN.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Bjarte Jordal for the continuous support during the PhD project, for his patience, motivation and his enthusiasm in sharing the vast knowledge on bark and ambrosia beetles. His guidance helped me during all the research activity and the writing of this thesis. I could not have imagined having a better advisor and mentor for this scientific adventure. A deep thankfulness goes also to my co-supervisor Lawrence Kirkendall for his insightful comments and encouragement during the thesis writing process, but also for his broaden interest in several aspect of bark and ambrosia beetles biology and ecology which incented me to widen my research from various perspectives.

My sincere thanks also go to Louse Lindblom and Kenneth Meland who gave me access to the laboratory and research facilities. Without they precious support it would not be possible to conduct this research; their high level of organization and the strict rules applied in the laboratory allowed me to carry out my research without ever experiencing serious PCR product contamination problems.

I would like to thank my colleagues at the University of Bergen (Richard, Steffen, Jostein, Manuel, Per and all the others) for the stimulating discussions, for the advices and for all the fun we have had during these years. Also I thank all my friends in the other institutions. In particular, I am grateful to Massimo Pajoro for helping me in developing other side projects and to Matteo Montagna for his fondness for beetles.

I want to express my deep gratitude and love to my family. I cannot express in words my profound admiration for unconditioned love and great support I received in these years, also in the craziest decisions we took together. Last but not the least; I would like to thank my family sensu lato: my parents, my brother and sister uncles, aunts and my grandma for supporting me throughout writing this thesis and my life in general.

I

PLOS ONE

RESEARCH ARTICLE

Genomic Mining of Phylogenetically Informative Nuclear Markers in Bark and Ambrosia Beetles

Dario Pistone, Sigrid Mugu, Bjarte Henry Jordal*

Department of Natural History, The University Museum, University of Bergen, PB7800, NO-5020 Bergen, Norway

* Bjarte.Jordal@uib.no

OPEN ACCESS

Citation: Pistone D, Mugu S, Jordal BH (2016) Genomic Mining of Phylogenetically Informative Nuclear Markers in Bark and Ambrosia Beetles. PLoS ONE 11(9): e0163529. doi:10.1371/journal.pone.0163529

Editor: Ben J Mans, Onderstepoort Veterinary Institute, SOUTH AFRICA

Received: April 13, 2016

Accepted: September 10, 2016

Published: September 26, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This project was funded by grant 214232/F20 from the Norwegian Research Council. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Deep level insect relationships are generally difficult to resolve, especially within taxa of the most diverse and species rich holometabolous orders. In beetles, the major diversity occurs in the Phytophaga, including charismatic groups such as leaf beetles, longhorn beetles and weevils. Bark and ambrosia beetles are wood boring weevils that contribute 12 percent of the diversity encountered in Curculionidae, one of the largest families of beetles with more than 50000 described species. Phylogenetic resolution in groups of Cretaceous age has proven particularly difficult and requires large quantity of data. In this study, we investigated 100 nuclear genes in order to select a number of markers with low evolutionary rates and high phylogenetic signal. A PCR screening using degenerate primers was applied to 26 different weevil species. We obtained sequences from 57 of the 100 targeted genes. Sequences from each nuclear marker were aligned and examined for detecting multiple copies, pseudogenes and introns. Phylogenetic informativeness (PI) and the capacity for reconstruction of previously established phylogenetic relationships were used as proxies for selecting a subset of the 57 amplified genes. Finally, we selected 16 markers suitable for large-scale phylogenetics of Scolytinae and related weevil taxa.

Introduction

In the postgenomic era, obtaining well resolved and highly supported molecular phylogenies of hyper-diverse eukaryotic lineages continues to represent a major challenge. Previous attempts on investigating phylogenetic relationships in beetles have demonstrated recurrent problems in resolving deeper relationships such as those between the four beetle suborders, but also much younger divergences [1–4]. One of the most problematic groups includes the weevils, where the majority of tribes and subfamilies remain unresolved despite considerable efforts in assembling molecular data [5–8]. Bark and ambrosia beetles in the subfamily Scolytinae represent a weevil lineage where much effort has been invested in developing molecular markers for phylogenetic analysis [9, 10]. Nevertheless, resolution between many Cretaceous relationships

remains rather low [11], emphasizing the scarceness of molecular markers to resolve this particular phylogeny.

So far, the vast majority of phylogenetic studies on beetles were based on markers such as ribosomal RNAs and mitochondrial cytochrome oxidase I and II genes [8, 12–15]. With the exception of nuclear ribosomal genes (*18s* and *28s rRNAs*) are most markers useful for the resolution of Cenozoic divergences, showing lack of phylogenetic signal for Cretaceous time frames [10]. In the last years, a growing number of phylogenetic studies on beetles have started to include nuclear protein coding genes, especially *EF-1α*, *CAD*, *ArgK*, and *wingless* [11, 16, 17], which are also widely used in other insect taxa [18–21]. However, a relatively limited amount of work has been done to discover and select additional nuclear genes for beetle systematics [22, 23–25], and all studies to date were based on less than 10 molecular markers [26, 27]. Therefore, obtaining a high degree of phylogenetic resolution in beetles is difficult; a direct consequence of high species diversity and a limited number of informative markers.

The first studies on the utility of protein coding genes in insect systematics date back to more than 20 years ago [28–30]. The advancement of insect phylogenies has largely been driven by the development of new markers in Lepidoptera [31]. At present, dozens of nuclear markers can be chosen to investigate Lepidoptera phylogeny at various ranks [18, 32–36]. Hymenoptera is another group where a consistent number of nuclear markers have been developed [37–39]. Although similar studies have been carried out in other insect groups such as Diptera [40–42], the majority of the remaining insect orders present a situation more similar to Coleoptera with few published markers conserved across different families [43, 44]. Thus, increasing the number of phylogenetic characters from protein coding nuclear genes is of mandatory importance for achieving robust phylogenetic hypotheses in beetle systematics.

Recently, the advent of next generation sequencing (NGS) technologies has contributed to additional ground-breaking advancements in the systematics field, profoundly increasing the level of resolution compared to previous phylogenies based on single or few genes [45]. Genomic and transcriptomic data obtained from NGS based research has led to predictive insect phylogenies, which now more clearly reveal key events in insect evolutionary history [46–50]. New developments based on ultra-conserved elements (UCEs) or RAD-sequencing will increase resolution also at lower taxonomic ranks in insects [51, 52]. However, the benefits of NGS are generally counterbalanced by the high cost and computationally demanding analyses of such high throughput data. The utility of few well-characterized markers should not be underestimated as they represent a rapid and cost effective approach for resolving small scale phylogenies.

Bark and ambrosia beetles in the subfamily Scolytinae constitute a group of highly derived, small wood boring weevils capable of excavating galleries into different parts of dead trees, shrubs and bushes, as well as in lianas and other plant tissues in different forest habitats throughout the world [53]. Scolytinae is generally regarded as a well-supported clade of more than 6000 described species representing approximately 12 percent of the entire diversity in the family Curculionidae [5, 54, 55]. A tremendous variability in life cycles, reproductive strategies, mating systems, host plants interactions, feeding behavior and ecology has been documented [56, 57], which makes this group of beetles particularly interesting to study in a phylogenetically comparative context. Phylogenies of Scolytinae have so far relied on a combination of five molecular markers (one mitochondrial and four nuclear genes) and eventually morphological characters. Given the high diversity of Scolytine species, additional data are needed to obtain sufficient resolution at deeper nodes.

In order to select new phylogenetic markers, 100 different nuclear genes were screened by PCR using degenerate primers and tested in a restricted but representative group of Scolytinae and other weevils. With the aim of developing slowly evolving genes, the properties of each gene

fragment were evaluated based on PCR amplification and sequencing success and their phylogenetic performance. This study reports on the development and utility of 16 novel markers for weevils, with a particular focus on bark and ambrosia beetles in the subfamily Scolytinae.

Materials and Methods

We included 18 species of bark and ambrosia beetles and 8 additional weevils from other subfamilies for primer screening (Table 1 and S1 Table). These beetles were collected by one of the authors (BHJ) during fieldwork in tropical forests (1998–2012). Collection permits were requested from authorities in Uganda, Tanzania, Cameroun, South Africa and Madagascar. Ethical guidelines were followed. Voucher specimens are deposited in the Coleoptera collection of the University Museum of Bergen, University of Bergen, Norway. All weevils, Platypodinae and Scolytinae species used in this study were previously described in other phylogenetic studies [7, 11, 58].

The procedure for primer selection can be summarized as follows: 1) putatively single copy expressed sequence tags (ESTs) longer than 800 base pairs were selected in GenBank for two different beetle species, *Tribolium castaneum* and *Dendroctonus ponderosae*; 2) preliminary

Table 1. Weevil species included in this study.

Species	Code	Subfamily	Tribe	Country
Brentidae sp.	BrBre05	Brentidae (familiy)	Brentinae	Cameroon
Mesites fusiformis	CsMes01	Cossoninae	Cossonini	Spain
Pselactus sp.	CsPse01	Cossoninae	Onycholipini	Portugal (Madeira)
Larinus sp.	ClLar01	Lixinae	Cleonini	Russia
Porthetes hispidus	MoPor01	Molytinae	Amorphocerini	South-Africa
Platypus impressus	PlPla07	Platypodinae	Platypodini	Tanzania
Triozastus marshalli	PlTri02	Platypodinae	Platypodini	Cameroon
Chaetastus tuberculatus	TsCha02	Platypodinae	Tesserocerini	Cameroon
Pityophthorus micrographus	CoPit01	Scolytinae	Corthylini	Norway
Diamerus inermis / D. hispidus	DiDia03 / DiDia04	Scolytinae	Diamerini	Tanzania / Madagascar
Dryocoetes autographus	DrDry01	Scolytinae	Dryocoetini	Russia
Ozopemon uniseriatus	DrOzo02	Scolytinae	Dryocoetini	Papua New Guinea
Hylastes attenuatus	HtHyt06	Scolytinae	Hylastini	Sweden
Hylesinus varius	HlHyl02	Scolytinae	Hylesinini	Sweden
Kissophagus hederae	HlKis01	Scolytinae	Hylesinini	Austria
Chaetoptelius vestitus	ToCha01	Scolytinae	Hylurgini	Morocco
Dendroctonus terebrans / D. micans	ToDen02 / ToDen01	Scolytinae	Hylurgini	USA
Tomicus piniperda	ToTom01	Scolytinae	Hylurgini	Norway
Acanthotomicus sp.	IpAca01	Scolytinae	Ipini	Cameroon
Pityogenes quadridens	IpPit03	Scolytinae	Ipini	Sweden
Premnobius cavipennis	PrPre01	Scolytinae	Premnobiini	Sierra Leone
Camptocerus aenipennis	ScCam02	Scolytinae	Scolytini	Guyana
Cnemonyx vismiacolens	ScCne01	Scolytinae	Scolytini	Guyana
Scolytus intricatus	ScScl02	Scolytinae	Scolytini	Czech Republic
Xyleborus affinis	XyXyl00	Scolytinae	Xyleborini	Cameroon
Xyleborus monographus	XyXyl03	Scolytinae	Xyleborini	Czech Republic

Degenerate primers were designed on conserved regions in the alignment of insect nucleotide sequences that were available from genomic and transcriptomic sources. Two or more consecutive degenerate sites were preferentially avoided as well as the use of completely degenerate sites (N). A total of 274 primers were designed (Table 2 - only successful primers reported).

doi:10.1371/journal.pone.0163529.t001

BLAST searches were performed to discard unsuitable markers, based on the evidence for multiple paralogous copies (e.g. large gene families) or ambiguous genomic characterization (e.g. similar matching values for different proteins); 3) available sequences for each selected gene were aligned, including annotated genomic and transcriptomic sequences from model organisms (e.g. *Drosophila melanogaster*, *Apis mellifera* and *Bombyx mori*) to determine intron-exon structure; 4) degenerate primers were designed; 5) a PCR screening was run and products with the expected correct size (albeit highly variable due to presence of introns) were sequenced; 6) markers reaching a minimum PCR and sequencing success of 20% were used to reconstruct single gene phylogenies (Bayesian) and trees were compared to previously established and well-supported clades [5, 7, 10, 11].

DNA was extracted from individual specimens using DNeasy Blood & Tissue kit (Qiagen) following the manufacturer's instructions. The PCR reaction mixture contained 2.5 μl 10x PCR buffer (Qiagen), in which the final concentration of $MgCl_2$ was 2.0 mM, 200 μM of each dNTP (Sigma Aldrich), 0.5 μM of each primer, 0.125 units Hot Start Taq ® DNA polymerase (Qiagen), 2 μl DNA, with water added to a final volume of 25 μl. A negative control (sterile water) was included in each test. The PCR was performed using a S1000™ Thermal Cycler (BIO-RAD Laboratories, Inc.). Three standard cycle programs were used for the initial screening: denaturation step at 95°C for 5 minutes, 35 cycles of 30 seconds at 95°C, 30 seconds at 48, 52 and 58°C, 60 seconds at 72°C, and finally 5 minutes extension at 72°C. Further optimization included a gradient of annealing temperatures in the range of 44–62°C, modulating the extension time depending on the expected PCR product length, and $MgCl_2$ concentration. We also considered two different touch-down PCR protocols for two of these genes (see Table 2 for details).

PCR products were sequenced with the same primers as those used for amplification. DNA sequences of both strands were obtained using the BigDye Terminator cycle sequencing ready reaction kit (Applied Biosystems Inc.) using an automated DNA sequencer (Applied Biosystems Prism 3700) following the manufacturer's instructions.

All obtained sequences were submitted to BLAST analyses, accepting a correct gene target if the cutoff value was below 1E-4. All sequences for each gene were aligned with other insect sequences for a preliminary NJ analysis in PAUP* 4.0 [59] to detect deviant sequences. The sequences were checked by eye and using Bioedit 7.2.5 [60] and MAFFT [61] to align gene fragments with complex structure, caused either by to the presence of indels of coding triplets, or less frequently by long introns marked by unusual exon-intron borders such as the most common alternative splice site GC—AG [62].

Introns were trimmed and the coding fragments were translated into amino acid sequences using Bioedit 7.2.5 to check for translational errors (stop codons). All these preliminary analyses had the purpose of detecting pseudogenes or early signs of possible paralogs (e.g. high degree of amino acid substitutions). In addition, the amino acid sequences of the selected markers were examined in OrthoDB v9 to assess gene orthology [63, 64]. The orthology for each gene was confirmed by cluster of orthologous groups (COGs) comparison among arthropod sequences in the database. Ambiguous nucleotide positions in the coding region that were difficult to align were tentatively excluded (in *Arr2* and *Iap2*) to create an alternative alignment for comparisons (see results and discussion).

Phylogenetic analyses were performed on unambiguously aligned sequences obtained from a minimum of 5 species. Phylogenetic inference was based on Bayesian and maximum parsimony analyses, the latter as implemented in PAUP* 4.0. Node support in the parsimony analyses was estimated by bootstrap analyses using 20 random additions of heuristic searches for each of 200 bootstrap replicates. Bayesian phylogenetic analyses were performed in MrBayes 3.2 [65]. The most appropriate model for base frequencies and substitution rates was determined by jModelTest [66], using the Akaike information criterion (AIC). MrBayes searches were run for each gene

Table 2. Primer sequences and annealing temperature for the nuclear markers selected in this study. Furthermore, primers for additional genes for lower level phylogenetics are reported.

Gene acronym	Primer forward (5'-3')	Primer reverse (5'-3')	Annealing T°C
EF2	CGTTTCTAYGCBTTYGGHCGTG	CCYTCYTTRGTGGCCCAYTGG	TD 58 (10 cy) 44 (25cy)
	ATGATGGGYCGTTAYGTWGARGC		TD 58 (10 cy) 44 (25cy)
Hsp70	CAAGCYGACATGAAGCAYTGGCC	CGGGTGATGGAGGTGTAGAARTC	58
	GAYGGTATCTTYGARGTMAAGTC	CGRCCYTTGTCRTTRGTGATGG	55
CCNC	ATGGCTGGMAAYTTTTGGCARAG	TCGAGCAGATARAAYTCRCAYTC	52
HDAC Rpd3	ATGAARCCSCACMGSATAMGSATGAC	GTAGTCGTTRTARGGSAGYTCRTTGGC	53
		GCCACSGAAGTYTCRTASGTCCA	53/50
Arr2	CGYGARGAGGAYGARGTYATGGG	ACCATSGTRACYTCGCAATGYTGCAC	52
		CTCAAARACKATRTTGTCGTCRTCGTC	52
Iap2	TGGAAYTAYGGRGACCAAGTRATGGC	CCATCKGGCRTGYTCYGTCCAWGGATC	52
PABP1	CCRATTCGYATYATGTGGTC	GAARGCRACAAAWCCRAAWCC	50
Prp1	ATGTCSGCKACTYTRGAYGCWGG	GGRTASGTGTTRTCYTGCATYTC	44
CTR9	GAAGGYGATAARATGGAWCARGC	TCGAAACAYTGKGCKGCATTTTC	52
RCC1	GGKTGYAATGACGARGGSGC	CGGCCCAATTGTCCYTGYTC	52
SOD1	TCCACATYCAYGARTTYGGGG	CCTTKKCCCAAATCATCMGG	TD 52 (10 cy) 46 (25cy)
TPI	CGHAAATTCGTWGTYGGWGGHAACTGG	CKGARCCYCCRTATTGRATTC	50
	GGTGGHAACTGGAARATGAACGG		52
ADA2	GAYATGYTDGAYGTVCATGC	ACAGGRCCRGCTTCRCCRCAATG	52
	AARTTYAATGCCAAATAYAAYCC	GGWCCRGCTTCACCRCARTGWGG	48/52
UBA5	TTGGKAGYGTAACWGCRGAAATG	ATATGGCCWGARACSGCRTTTTC	52
Cda4	TACGARGARTGGGTKGGRGARATG	AACCAATTMGTRTGRAASGGCATC	48
FEN1	GARGCCCCYTGYGARGCKGARGC	TCACCATGCCYTCYTCRTCMGG	48
ACTB	CTGAAGCCCCMTTGAACCCMAAGGC	GAGATCCACATCTGYTGGAARGTGG	
CXorf56	GAAGYATTGCRTGTTCSGAYAC	GTCACMGAACTGAAYTTKCCC	
eRF1	GTTGGCAGATGAATTTGGAACRGC	CCRAABAGAGCTCCRTTACCATCC	
U2AF	ATYGCTGGATTWAAYGGRATGC	TCTCKTCTRTGRTACTTRTCSGGWTC	
MAD	YAAYTTYCCWGCYATGRTWCC	ACACCRTGRTTYTTWGCWCC	
mp20	GACAAGGARGCCCARGARTGGATCG	TCCCACAGRTCAACTGTYTGGAARAC	
		GGTCCGGGCCCAYTCRGRGTGCYTGTTAGG	
5MP	CATGACKTTTATGMGKGCKTTC	CTTCYTCRGCGTTTTGWAGCC	
Pi4k	TGYTGYCCKTGYTGYTTYGG	TGGTAYGGRTASGCYCGCC	
Gel	GAYGAGGGCSGGWTCSGCWGC	AGGATRAAGCARTCRCCTTTGTTC	
C1-THF	CATYTRACYGGYGAYATYCATGC	ACAGCYCCYGTKGCYCCCAAATC	
alpha-Spec	CAYGCHAATGCWTTCCATCARTGG	GGYTGKCCYTCYTCWACCATYGG	
AATS	CATCAYACGTTTTTTGAGATG	GCATGRTCNGCTAARACNCGRTARGCC	
Hsp90	GATCATCAATATSTTCTACTC	TCTCCGGTGATGWARTAGATG	
dldE3	GGRGAYTGTATWCATGGRCC	GCYTCRTTRATBARTTCRCC	
	CATCCWGAAGTKGGMTGGGTKGG		
Mpgt	AAACCSCTGTTYCCMGTTGCKGG	GCMGTTTTYAACTGSGACCACC	
NaK	GGYGGTTTCGCSWTGYTGYGTGGATCGG	GCGACGATGATACCGATCARGAAGATGACAGC	
Fbox11	AATGCWTTRGCTGGWATYTGGG	CCRCCRTGYTGACCRTGRTG	
UDE	AAGCCRGACACCGTWCCCGG	CTGGCWTCRGGRCTGTACGCCC	
GTPbp	ATTARAAYGTAKCCATCGTTRCCCC	GTGTTGATAATWGASGACTTGCC	
CatL	CACATTTACACTTTYAACCCRATG	ACCARCTGTTYTTMACCARCCAGTA	
TpC	CTTCCCSCMGARCARATYGCCG	CCTCSCCRGTCATCATCTCCATG	
PGI	GGCCCSCTKATGGTRACCGAAGC	CCCAGCTCCACKCCCCATTGGTC	

(*Continued*)

Table 2. (*Continued*)

Gene acronym	Primer forward (5'-3')	Primer reverse (5'-3')	Annealing T°C
AcCoA	GGTGTACTGCKGAYATTGGYTGGATCAC	GGAAACSCAGCMGCKCCWGGYTTCAT	
		CATCAGRTGYCCKGASACGTTYARCAT	
Ucdk	GAGCACKGTWTGCAARCGYATWATGG	CCYCTWGGAATRATRACATCAGC	
PPO1	AAYCTSCACCAYTGGCAYTGGC	CGGAASGTSCKCTCRAASGG	
Prp6	AATCCSAATCATCCWCCGGCKTGG	TTCTTCCAGYTTRGCSGCRGTWGTCC	
Mxp	TAMGSACRGCSTAYACSAACAC	CGCTTGTGYTTCATSCKCCG	
Npl4	CTCGYTGYGTSCAYTGCTC	TCGCGCACYAGCGCCATRCAYTG	
Cam1	GAYGGMGATGGCACRATYACTACC	TCRTAATTGACCTGACCGTCRCC	
STX1A	ATGACYAARGAYAGATTRGCRGC	GCCATRTCCATRAACATRTCRTG	
TP120b	TWGGRAATGTCAAYGTYTC	AAGCTCAACCCKCKCCACATCC	
CHS1	CATATMTTYTTCGAYGAYGC	CAACGATCYTCKCCYTGATC	
DDX49	AARGCTATACGARGAYCCWTATGG	TGCCTGCYCTAGCWGTYCTYCC	
GTF2H3	CTCGCATTTGATGCAGAAGGC	CARATYGGRCTAAACTTGCA	
IF3	ACTCGCTYTACAAAATGTTGGG	CTTTSGTRTCGGCRATATGRATC	
TIF6	GACACRATWCCSGTGGTSCATGC	CTACCWCARTTWACYGTTCC	
IDH	TACAAYGTWGGAATWAARTGTGC	CAMACAAARCCYCCYTCMGATTTC	
Ecr	GAAGTKATGATGTTCMGRATGGC	GAWGCACATYTCDGARTTYTG	

doi:10.1371/journal.pone.0163529.t002

separately and for concatenated datasets (8109 bp– 2702 aa) using the suggested models for each gene partition and a mixed model for amino acid substitution. In both cases, the search consisted of 2000000 generations with two independent runs, each with four simultaneous chains, and trees sampled every 1000 generations. The convergence diagnostics (SDSF, PSRF) and parameter sample plots were evaluated using the software Tracer 1.6 [67].

An indirect measure of the phylogenetic signal in each marker was assessed through topological congruence with previously well documented clades [5–7, 10, 11, 68] which were used to derive a scheme of the current classification of Curculionoidea (Fig 1). These clades belong to six tribes of Scolytinae (A = Dryocoetini including Xyleborini, B = Ipini, C = Hylurgini + Hylesinini, D = Scolytini) and the subfamily Platypodinae (E). Rooting of the trees was

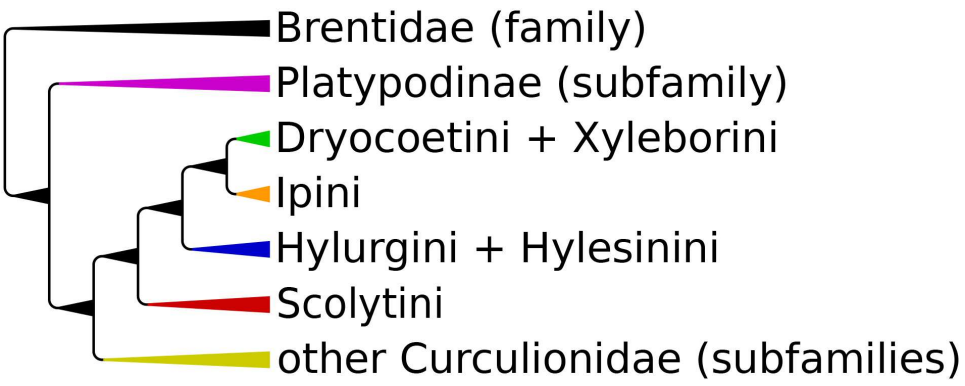


Fig 1. Schematic tree showing well supported relationships between tribes within the subfamily Scolytinae and other weevil families and subfamilies considered in this study.

doi:10.1371/journal.pone.0163529.g001

dependent on the sequences available, and used in the following order: 1) Brentidae, 2) Platypodinae, 3) Cossoninae, Molytinae and Lixinae, 4) Scolytini [5, 6].

Basic properties of each gene, including the overall mean divergence of sequences (*p*-distance) and the variation in first, second and third positions, were calculated for each gene fragment using MEGA 6.0 [69]. Parsimony informative sites were calculated together with the homoplasy and retention indices (respectively HI and RI–S2 Table) using PAUP * 4.0. A phylogenetic informativeness profile (PI) was obtained for each gene using PhyDesign [70], an online program developed from a previous study [71]. Substitution rates for each position were calculated using HyPhy implemented in PhyDesign, selecting a K2P model (base frequencies = 0.25, transitions = 2, transversions = 1). The input time tree was obtained using Beast v1.8.2 [72], with topology constraints following previously published phylogenies of weevils and Scolytinae [5, 6, 11]. The tree was reconstructed using a concatenated dataset of 16 genes, using a GTR+I+ Γ model for each gene partition, and a Yule speciation process. We selected an uncorrelated lognormal relaxed molecular clock and used default priors as suggested by the authors (see XML S1 file in Supplementary information). Two calibration points were used: 116 Ma for the node subtending Scolytinae and other weevil subfamilies, and 30 Ma for clade A (Dryocoetini+Xyleborini).

Results

Sequences were obtained for 57 different genes, whereas 43 primer sets never amplified the correct gene. A total of 798 sequences were obtained, but only 510 of these (64%) were unambiguously characterized as beetle orthologs in BLASTN search. Among the remaining 288 sequences, 53 were identified as non-beetle sequences (mainly from bacteria, fungi or nematodes associated with beetles) with different degree of confidence in gene identity. The remaining 235 sequences resulted in unreadable or poor quality sequences without a clear match in GenBank (E value $>$ 1E-4, query coverage $<$ 30% and/or less than 30% identity).

The evaluation of the 57 markers with readable sequences was based on the number of sequences obtained and their phylogenetic performance. When only one or two sequences were obtained for a gene (e.g. *cathepsin L*, *troponin C*, *acetyl coenzima A synthetase*, *maxillopedia*, *calmodulin 1*), the phylogenetic utility was not possible to assess. Other excluded markers produced a higher number of sequences, such as *odorant binding protein* (8 sequences) and *glycoside hydrolase family 31* (11), but these were largely unalignable. Another group of failed markers produced sequences from non-target organisms, such as *6-phosphogluconate dehydrogenase* of fungi, or *phosphoglucose isomerase* of bacteria. A total of 23 genes were discarded due to low amplification rates, high levels of non-beetle amplification, or generally low degree of gene orthology.

The remaining 34 genes showed differing degree of PCR and sequencing success (from 5 to 26 sequences obtained), and were further evaluated based on their capacity to recover known relationships at various taxonomic levels. Eighteen of these markers were found insufficiently informative for higher level phylogenetics, because no more than two of the predefined clades were reconstructed correctly. However, most discarded markers nevertheless revealed some phylogenetic utility at lower taxonomic level; including populations (see S3 Table for further details).

We selected 16 genes that revealed a relatively high and stable PCR and sequencing success (from 50 to 100%) as the best candidates for Scolytinae phylogenetics (Table 3). All the verified sequences obtained in this study were deposited in GenBank database under the accession numbers KX160539—KX160803 (S1 Table). The species *Xyleborus affinis* was the most successful in PCR and sequencing (15 out of 16 possible sequences obtained); the other samples

Table 3. PCR and sequencing success for 16 selected genes.

GENE ACRONYM	A	B	C	D	E	F	G	H	Total (%)
PABP1	4	3	6	3	3	4	2	1	26 (100%)
TPI	4	2	6	-	2	2	2	-	18 (69%)
UBA5	3	3	5	3	2	2	1	1	20 (77%)
Iap2	3	3	1	2	1	4	2	-	16 (62%)
SOD1	2	1	4	3	2	3	1	-	16 (62%)
Prp1	3	3	5	1	3	1	2	-	18 (69%)
ADA2	3	2	2	2	3	-	2	-	14 (54%)
CTR9	2	2	4	2	-	1	2	-	13 (50%)
CCNC	4	2	5	2	2	2	2	1	20 (77%)
Cda4	2	1	4	-	3	1	1	1	13 (50%)
HDAC Rpd3	3	1	4	-	2	2	1	-	13 (50%)
Arr2	4	2	4	3	3	2	2	-	20 (77%)
FEN1	3	2	4	2	1	-	2	1	15 (58%)
EF2	2	2	3	2	3	-	2	-	14 (54%)
Hsp70	1	1	5	2	1	1	2	1	14 (54%)
RCC1	2	2	4	-	2	2	1	-	13 (50%)

The number of sequences obtained was reported for the following groups: A = Xyleborini + Dryocoetini, B = Ipini, C = Hylurgini + Hylesinini, D = Scolytini, E = Platypodinae, F = other Curculionidae subfamilies, G = other Scolytinae, H = Brentidae.

doi:10.1371/journal.pone.0163529.t003

varied considerably in this respect with only 4 sequences obtained for *Larinus* sp. (S1 Table). The total fragment length, the presence of length-variable regions, and the number and position of introns, were mapped on the annotated genomes of *T. castaneum* and *D. ponderosae* (eventually transcriptomic and genomic data of other insect species) to create a map of the gene structure (Fig 2; see also Table 4).

OrthoDB analyses showed that 12 out of 16 genes selected in this study are present in single copy in more than 70% of the arthropod species currently in the database (133). *PABP1* and *UBA5* are in single copy in 96% of these species, followed by *HDAC Rpd3* (95%), *CCNC* (94%), *Prp1* (92%), *TPI*, *CTR9* and *FEN1* (90%), *Cda4* (89%), *EF2* (84%), *RCC1* (81%) and *ADA2* (74%). Only five genes are frequently in multi-copy status in arthropod genomes: *Hsp70* (single copy only in 2% of the species in the database), *Arr2* (4.5%), *Iap2* (8.3%) and *SOD1* (22%).

The best evolutionary model for the majority of the genes was GTR+I+Γ, except for *SOD1* and *Iap2* in which SYM+I+Γ and GTR+Γ were selected. Bayesian analysis of the concatenated nucleotide and amino acid data from 16 genes showed a well resolved tree topology (S1 Fig) with all expected clades recovered with maximum support, except Scolytini (pp = 0.75). The overall tree topology was correct with the exception of four weevil species that were nested inside Scolytinae as the sister lineage to Hylurgini (weakly supported in the amino acid analysis). Parsimony analyses of the concatenated dataset revealed similar results both for the nucleotide and amino acid datasets, with all major clades recovered with medium to high bootstrap support. However, the sub-family Scolytinae was not monophyletic in respect to the other advanced weevil species (S2 Fig).

Single gene analyses resulted in partially resolved phylogenies, mainly recovering a monophyletic Scolytinae, the majority of the predefined subgroups of Scolytinae (A-B-C-D), and the subfamily Platypodinae (Fig 3). All selected genes enabled the correct reconstruction of the most recent clade (A), with 3 genes obtaining the correct sister group (B). None of the selected genes showed high degree of incongruence that received high node support. Overall mean

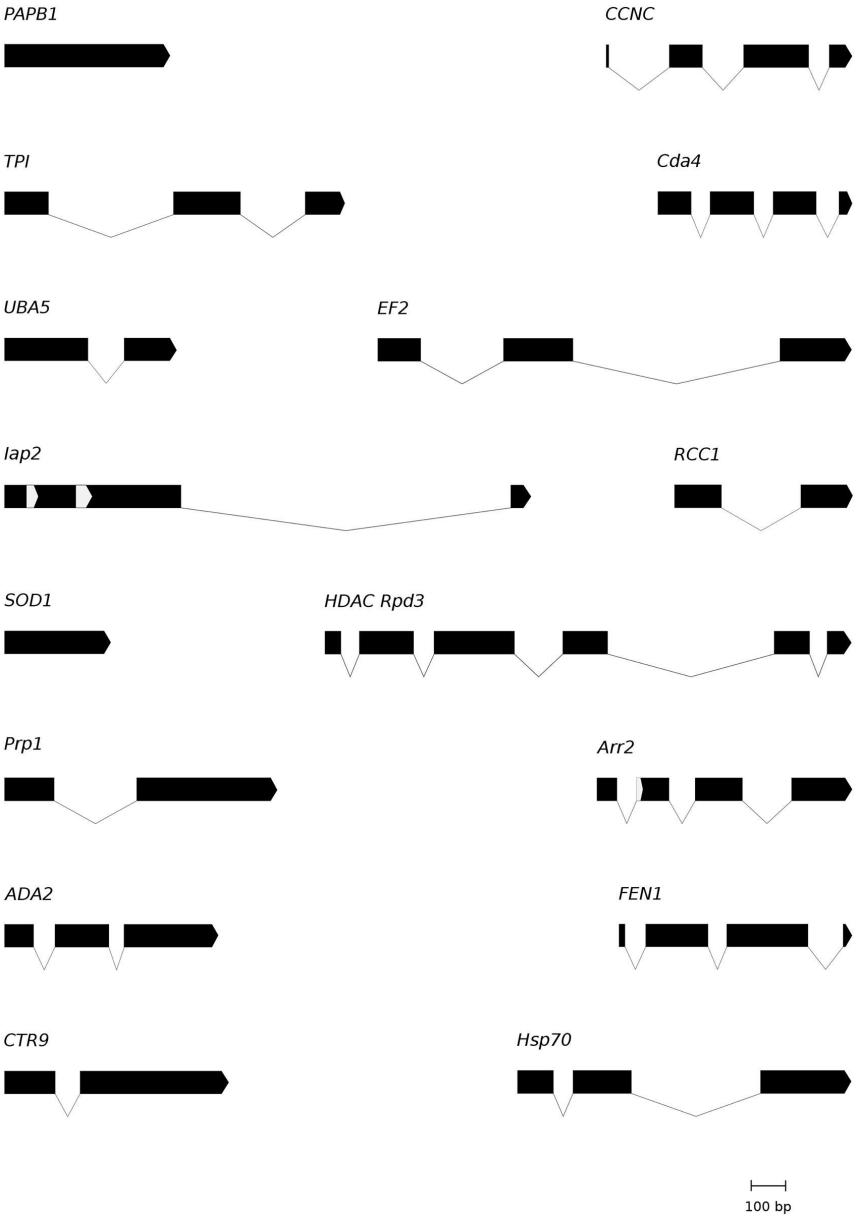


Fig 2. Structure of the PCR amplified gene fragments. The graphics illustrate intron-exon patterns in 16 markers with coding regions shown as black bars and introns as thin black lines. Length variable coding regions (indels) were colored in light grey (*Iap2* and *Arr2*).

doi:10.1371/journal.pone.0163529.g002

Table 4. Gene information.

Acronym	nucs	aa	Intron	Intron range (per intron)
PABP1	435	145	0	-
TPI	547	182	0–2	(457–51)(237–48)
UBA5	348	116	1	(94–48)
Iap2	672*	224*	1	(1131–50)
SOD1	213	71	0	-
Prp1	582	194	0–1	(258–55)
ADA2	624	208	2	(70–39) (105–53)
CTR9	627	209	0–1	(81–59)
CCNC	384	128	3	(200–69)(134–49)(71–58)
Cda4	410	136	0–3	(68–51)(63–56)(53)
HDAC Rpd3	858	286	3–5	(69–53)(70–54)(165–48)(564–54)(66–55)
Arr2	501*	167*	0–3	(110–51)(84–53)(158–55)
FEN1	417	139	1–3	(63–46)(55–42)(93–44)
EF2	621	207	1–2	(398–183)(702–84)
Hsp70	567	189	0–2	(61-?)(317–187)
RCC1	303	101	0–1	(250–51)

For each marker, the length of the sequenced coding region is given as the number of nucleotides and amino acids, together with the number and length of intron(s). The symbol * indicates genes with sequence length variability due to exonic indels.

doi:10.1371/journal.pone.0163529.t004

divergence in nucleotide sequences was reported for each codon position for each gene (S3 Fig).

Selected genes for Scolytinae phylogeny

Polyadenylate binding protein 1 (PABP1). *PABP1* was the most successful marker, with sequences obtained from all 26 species. The amplified fragment was 435 bp long, contained no introns, and translated into 145 amino acids. The phylogenetic analyses recovered almost all pre-defined clades (Fig 3a), but only two of them were highly supported (B, pp = 0.98; E, pp = 1). The tribe Scolytini was placed outside a polytomy including the remaining species of Scolytinae, the subfamily Platypodinae and the various other weevil subfamilies. No clear evidence of paralogs emerged from the analyses. Preliminary studies indicated increased phylogenetic performance with broader taxon coverage.

Triose-phosphate isomerase (TPI). A combination of two primer pairs (two forward, one reverse) resulted in 67% PCR amplification and sequencing success. The aligned fragments consisted of 547 bp after removal of introns, which translated into 182 amino acids. Two introns were located in this gene fragment (Fig 2, Table 4). The phylogeny based on this marker confirmed the monophyly of Platypodinae (pp = 1), while Scolytinae formed a large polytomy including two advanced weevil species. Furthermore was Cossoninae monophyletic (pp = 1), in addition to one scolytine subgroup (A, pp = 1), and subgroup C almost so (Fig 3b).

Ubiquitin-like modifier activating enzyme 5 (UBA5). The *UBA5* gene fragment is 348 bp long and translated into 116 amino acids. It was amplified from 20 different species (77%) in all main clades and contained one short intron in all species. The phylogeny recovered the monophyly of clades A and E with high node support (pp = 0.99 and 1, respectively) while clade D (pp = 1) had *Scolytus intricatus* excluded. Clade B and C were weakly supported (pp<0.95) and *Kissophagus hederae* was not included in Hylurgini (Fig 3c).

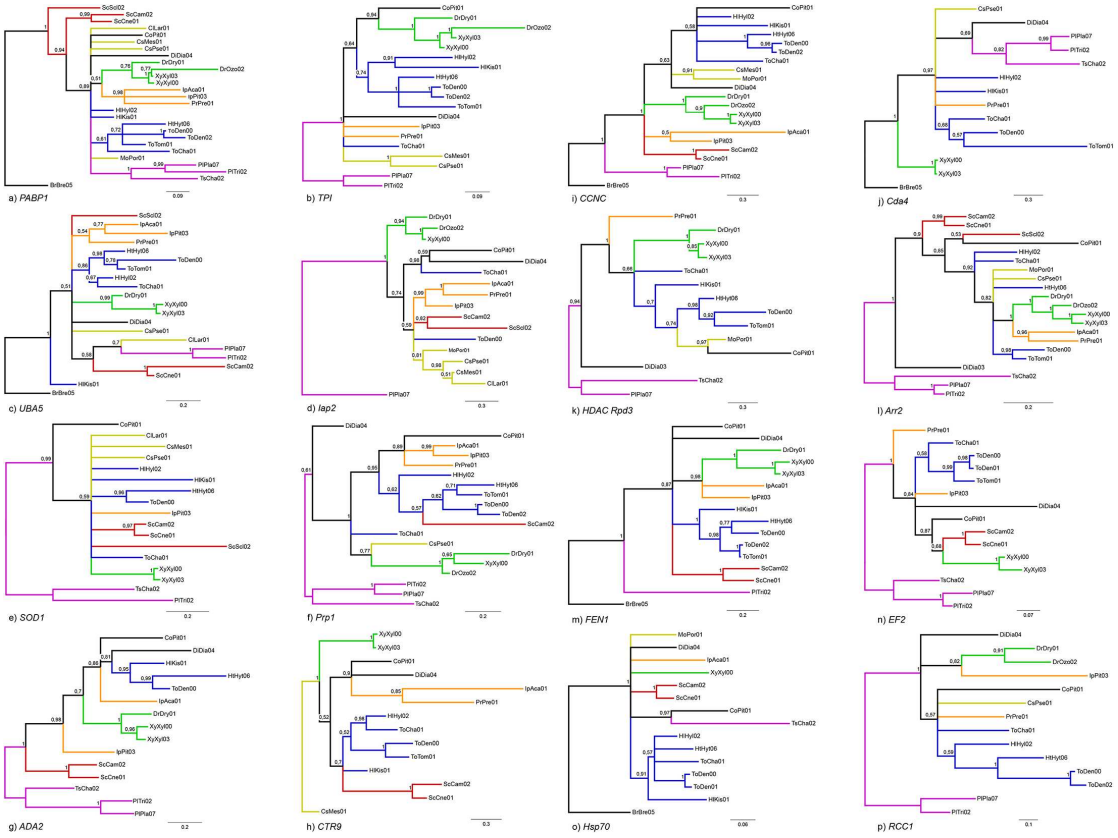


Fig 3. Phylogenetic trees based on Bayesian analyses of 16 selected genes. Trees were rooted with the most distant outgroup available for each marker. Posterior probabilities are given to the left of the nodes. Sequences of *D. ponderosae* (ToDen00) were obtained from GenBank.

doi:10.1371/journal.pone.0163529.g003

Inhibitor of apoptosis 2 (Iap2). A total of 16 sequences (62%) were obtained from partial *Iap2*. This gene was amplified for only one species in Hylurgini (*Chaetoptelius vestitus*). The amplified fragments contained one long intron and a coding region of variable length up to 672 bp. Two hypervariable regions in the first exon were characterized by a series of indels of up to a maximum of six and ten triplets, respectively, consisting of serine-rich strings of amino acids. The intron range is within 50–80 bp in the majority of the species, but *D. ponderosae* (obtained from GenBank) contained a very long intron (1131 bp). BLASTN search indicated that a baculoviral Iap repeat is located between the two hypervariable regions. The phylogenetic analyses resulted in four monophyletic groups (clade A, pp = 0.94; clade B, pp = 0.99; D and F, pp<0.95), with no phylogenetic evidence of paralogs (Fig 3d).

Cu-Zn superoxide dismutase 1 (SOD1). We amplified a short fragment (213 bp) of the cytoplasmic copper/zinc superoxide dismutase (*SOD1*), which contained no intron. We obtained 14 orthologous beetles sequences (54%) and five non-beetle sequences, but also amplified other genomic regions, suggesting non-specificity for this primer pair. The

phylogeny contained several polytomies, with only one clade (A) receiving maximum support. Two internal nodes in the C and D clades were also recovered (pp>0.95). The tree was rooted with a monophyletic Platypodinae (Fig 3e).

Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP1 (Prp1). A fragment of the *Prp1* gene with the length of 582 bp (intron excised) corresponding to 194 amino acids, was amplified from 18 different species (70%). The presence of a single intron was observed in the majority of the species except three unrelated Scolytinae species and one Platypodinae. The phylogeny revealed two monophyletic groups (A, pp = 1; E, pp<0.95) and three groups which contained highly supported internal nodes (B, C and E), and a series of weakly supported incongruent relationships (Fig 3f). The tree was rooted on a monophyletic Platypodinae.

Adenosine deaminase 2 (ADA2). We amplified and sequenced the *ADA2* gene from 14 species (54%). Failures were most frequent in weevils other than Scolytinae and Platypodinae. The tree topology (Fig 3g) was largely congruent with our predefined clades (A, C, D, E; all pp≥0.95), except Ipini (clade B). The tree was rooted on a monophyletic Platypodinae.

RNA-associated protein CTR9 (CTR9). A single primer pair resulted in the amplification and sequencing of 13 sequences (50%), mainly in Scolytinae, with much lower amplification rates in other weevil subfamilies (1 sequence). The amplified gene fragment revealed a simple structure with a single intron in many species, but was absent in the entire tribe Scolytini and a few other Scolytinae species. The two exons presented a total sequence length of 627 bp (209 amino acids). The phylogeny recovered three pre-defined clades (A, B and D), two of them highly supported (A and D) while resolution at deeper nodes was generally low (Fig 3h).

Cyclin-C (CCNC). A 384 bp fragment (introns excised) was amplified for 20 species (77%), with relatively good taxon coverage among the different groups. The alignment included three long introns which may cause amplification and sequencing problems. The phylogeny based on this marker revealed a monophyletic Platypodinae (pp = 1) that formed the sister group to the advanced weevils (Curculionidae sensu Alonso-Zarazaga and Lyal 1999, pp = 1). All smaller clades were congruent with previous phylogenies, albeit only three clades were strongly supported (A, D and E, pp = 1), whereas the larger group of Scolytinae was paraphyletic with respect to two other weevil species (Fig 3i).

Chitin deacetylase 4 (Cda4). *Cda4* sequences were obtained from a total of 13 beetle species (50%). This marker amplified few weevils other than Scolytinae (2 sequences) and failed to amplify species in the tribe Scolytini. The gene structure was relatively simple with 3 short introns (<100bp), with the first and the third intron present in the majority of the species, while the second one was absent in all Platypodinae and Hylurgini species. The phylogeny based on a 410 bp long coding fragment (136 amino acids) showed monophyly for group A (pp = 1) and E, while Hylurgini (group C) was paraphyletic (Fig 3j).

Histone deacetylase Rpd3 (HDAC Rpd3). *HDAC Rpd3* represents the longest gene fragment selected in this study. This gene was amplified and sequenced for 13 species (50%), with the longest fragments reaching more than 1700 bp due to the presence of introns. A total of 5 introns were present in one species (*Platypus impressus*), while the other species showed a high variability in intron numbers (1–4) with intron 4 particularly long in *Kissophagus hederae* (571 bp). The final alignment, with introns removed, resulted in 858 nucleotide positions coding for 286 amino acids. We did not amplify any species in the tribe Scolytini (clade D) and we had limited success with Ipini (B) and in weevils other than Scolytinae and Platypodinae (Fig 3k). The phylogeny based on these sequences showed a largely unstructured tree, with only clades A and F recovered (pp = 1 and pp = 0.94 respectively), and partially so in Hylurgini (clade C: *Hylastes attenuatus*, *Tomicus piniperda* and *D. ponderosae*, pp = 0.98).

Arrestin 2 (Arr2). *Arr2* showed high degree of PCR and sequencing success in Scolytinae and in some other weevils, obtaining a total of 20 sequences (77%). The alignment of our new

Arr2 sequences contained three introns. At the beginning of the second exon, the coding region varied in length due to triplet indels. One example of atypical intron borders was encountered in the first intron (GC-AG), in *Premnobius caevipennis*. Three predefined clades were recovered (A, pp = 1; B, pp = 0.96; E, pp = 1), with two other groups only partly resolved (clade C, pp = 0.98; D, pp = 0.99). The overall tree topology was largely congruent with established phylogenies, where clades A and B were recognized as sister lineages with maximum node support (Fig 3l). The tree was rooted on a monophyletic Platypodinae.

Flap endonuclease 1 (FEN1). *FEN1* sequences were obtained from 15 different species (58%). The alignment of nucleotide sequences revealed three introns that were present in the majority of the species. The coding region was 417 bp long and translated into 139 amino acids. The phylogeny was well resolved and recovered highly supported monophyletic groups corresponding to the clades A, B, C, and D (Fig 3m). In addition, the sister clades A and B were correctly reconstructed (pp = 0.98), and Platypodinae (one species) was, in the absence of other advanced weevils, placed as sister to Scolytinae.

Elongation factor 2 (EF2). We obtained *EF2* sequences from 14 species (54%), but only from species in Scolytinae and Platypodinae. Additional unspecific amplifications of *EF2* were also obtained (7 sequences), mainly from fungi and nematodes. The amplified fragment contained two long introns up to 300 bp, but occasionally longer in a few species (Table 2). Bayesian analysis of 621 aligned nucleotides (207 amino acids) showed a partially correct phylogeny that included several highly supported clades (A, D and E, all with pp = 1). The monophyly of Hylurgini (clade C) was only weakly supported (Fig 3n). The tree was rooted on a monophyletic Platypodinae.

Heat shock protein 70 (Hsp70). Partial *Hsp70* gene was amplified in 14 species (54%) and contained one or two introns. Only the second intron was present in the majority of amplified species. With introns excised, the alignment consisted of 567 nucleotides coding for 189 amino acids. This marker performed particularly well in Hylurgini and Hylesinini (clade C) with 5 out of 6 samples amplified. The phylogeny contained a well resolved clade C (pp = 0.91) and D (Scolytini, pp = 1), while the remaining parts of the tree topology formed largely a polytomy (Fig 3o). Unspecific PCR amplification and sequencing of fungi and nematodes occurred in four samples. Furthermore, paralogous copies, characterized by a triplet insertion in weevils, were identified based on phylogenetic analysis of all available sequences (S3 Fig).

Regulator of chromosome condensation 1 (RCC1). A short fragment consisting of 303 bp (intron excised) was amplified for 13 species (50%). The sequenced gene fragment contained one intron in all species, except *Hylesinus varius*, and the exons could be translated into 101 amino acids. The primers showed very low success in weevils other than Scolytinae, amplifying only two species in group E (Platypodinae) and one species of Cossoninae. The primers did not amplify this gene in the tribe Scolytini (D). Occasional unspecific amplifications were observed (4 sequences, from fungi and nematodes). The phylogeny based on this marker was mainly congruent with established relationships and showed no evidence of multiple copies (Fig 3p). Platypodinae (E, pp = 1), Dryocoetini (A, pp = 0.91) and a subclade of Hylurgini (C, pp = 1) were recovered.

Phylogenetic signal

Phylogenetic informativeness (PI) profiles varied considerably between the selected markers, showing different degrees of signal across the more than 100 Ma of weevil evolutionary history (Fig 4). The net PI values showed a marked decline for all markers towards the Cretaceous era. *Iap2* displayed the highest PI peak in recent times, followed by four other markers with lower PI profiles (*TPI*, *Prp1* and *Arr2*, *FEN1*). The gene *EF2* showed a diverse profile, having lower PI

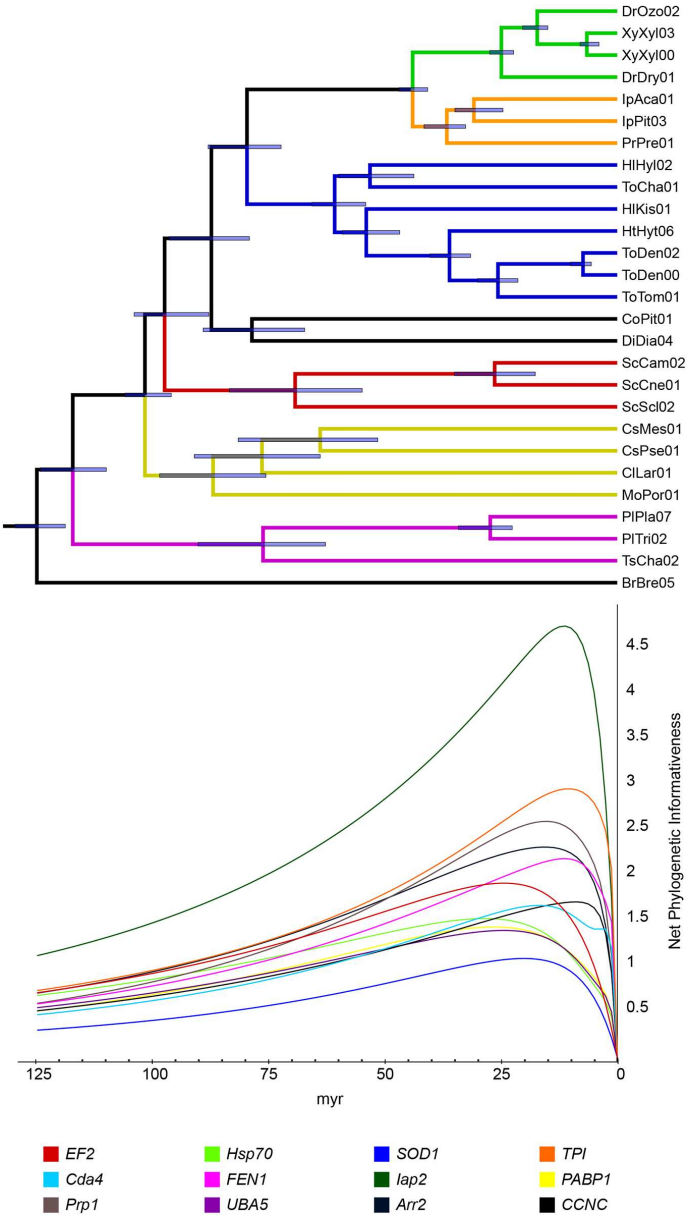


Fig 4. Phylogentic informativeness profiles. The K2P model was used to estimate substitution rates in HyPhy as implemented in the software PhyDesign. Different evolutionary models produced similar results (data not shown). The dated phylogenetic tree was obtained using BEAST v1.8.2.

doi:10.1371/journal.pone.0163529.g004

for recent times but relatively more PI than *FEN1* and *Arr2* at more ancient times. *PABP1*, which presented the highest homoplasy level among the selected genes (S3 Table), showed an intermediate PI profile, following the same trend of *Hsp70*, *Cda4*, *CCNC* and almost identical to *UBA5*. *Cda4* and *CCNC* showed higher PI in recent times while *Hsp70* maintained marginally higher PI for ancient times. The gene with the lowest PI value was *SOD1*. Four markers (*HDAC Rpd3*, *ADA2*, *RCC1* and *CTR9*) were not included in the analysis due to missing data.

Additional genes for lower level phylogenetics

One of the main characteristics shared by several of the 18 genes that were not selected was the generally low, and sometimes clade-specific, PCR and sequencing success (S3 Table). These genes also exhibited many problems in phylogeny reconstruction when sufficient data were obtained, including failure to recover well-established clades (Fig 5). For example, very few sequences were acquired for *α-spectrin*, with no sequences obtained for three of the groups (B, C and D), producing a tree topology with only one correct clade recovered (A, pp = 1) and therefore difficult to evaluate (Fig 5a). A similar situation was reported for *phosphatidylinositol 4-kinase type 2-alpha* (*Pi4k*) where no sequences were obtained for the clades D and E, but two clades (A, pp = 1 and C, pp = 0.97) were recovered correctly (Fig 5b), and a third group was nearly monophyletic (B, excluding *Pityogenes quadridens*, pp = 0.96). For *muscular protein 20* (*mp20*) we obtained a higher number of sequences (12), with two monophyletic groups recovered (clade A, pp = 1 and B, pp<0.95), but with group D (Scolytini) not monophyletic (Fig 5c). In the case of the *beta-actin* gene (*ACTB*), sequences were obtained from 18 different species, including 5 species of Hylurgini. However, the phylogeny recovered only one of the youngest clades (B, pp = 1), while all other groups were largely paraphyletic (Fig 5d). In the *chromosome X open reading frame 56* gene (*CXorf56*), only the youngest group (clade A, pp = 0.99) was correctly recovered (Fig 5e) whereas closely related species did not group together. Another poorly performing gene was *MAD*, with a phylogenetic tree showing a large polytomy that included a highly paraphyletic Hylurgini (clade C). This gene nevertheless distinguished Platypodinae (pp = 0.96) from all other advanced weevils at the root of the tree (Fig 5f). A similar situation was also observed for the *eukaryotic peptide chain release factor subunit 1* (*eRF1*) gene. The phylogeny largely formed a polytomy (Fig 5g), and included many paraphyletic groups, including Platypodinae (clade E). The phylogeny for *splicing factor U2F* showed a largely unstructured tree with generally low support (Fig 5h), with only Scolytini monophyletic (clade D, pp<0.95).

The remaining 10 of the 18 genes with shallow level phylogenetic utility generally exhibited low PCR and sequencing success (5–9 sequences), and showed clade-specific amplification (see S3 Table). A correct tree topology was recovered for *dihydrolipoamide dehydrogenase E3* (*dldE3*) which showed a congruent and well supported phylogeny for three clades (A, B and C, all with pp>0.95) and also recovered a node including A+B (pp = 0.99). The low number of sequences obtained (7) was the main reason to exclude this gene. *Alanyl-tRNA synthetase* (*AATS*), *F-box only protein 11* and *Na+/K+ ATPase alpha subunit* (*NaK*) displayed very low PCR and sequencing success. The first of these recovered clade A (pp = 1), the second clade A and B (pp = 1 and pp = 0.99 respectively) while the third one did not produce enough sequences to enable hypotheses testing. *Hsp90* revealed amplification of eight species in Scolytinae, but not other weevils. The phylogeny was consistent with clade A (pp = 0.97) and partially so for clade C (3 species pp<0.95). The alignment of *Hsp90* revealed no intron but the coding region presented variable length due to the presence of indels. Primers for the two genes *mannose-1-phosphate guanyltransferase α C1-tetrahydrofolate synthase* (*C1-THF*) and *uracil-DNA degrading factor* amplified well in Hylurgini. Finally, *gelsolin* and *elongation initiation factor*

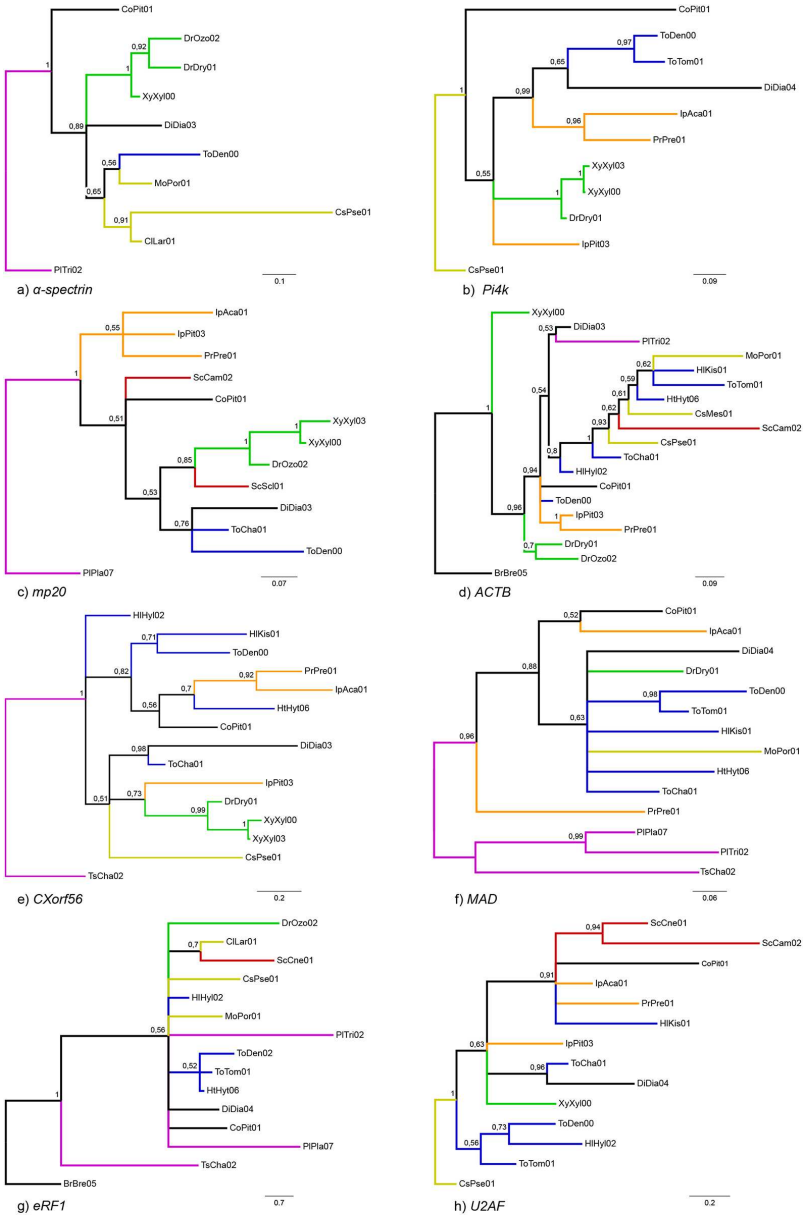


Fig 5. Phylogenetic trees resulting from Bayesian analyses of 8 excluded gene fragments.

doi:10.1371/journal.pone.0163529.g005

5C (also known as *krasavietz - 5MP*) revealed unstructured tree topologies. The first gene recovered only clade D (pp = 1) while the second supported clade E (pp = 1) and in part clade C. Additional information on suggested subfamily/tribe/genus specific markers were reported in supplementary material (S3 Table).

Discussion

Phylogenetic studies on insects have generally suffered from a lack of coordination in establishing a common set of nuclear markers [73]. Most efforts were invested in butterflies and bees [31, 37], with other related groups occasionally taking advantage of such developments [74]. Beetles are one of the many groups lagging behind in terms of phylogenetic marker availability. With the presentation of 16 protein coding genes, which are here shown to be informative in weevil phylogenetics, and the suggestion of 18 additional, but less developed, genes as potential phylogenetic markers at various taxonomic levels, we have at least partly remedied this situation. Indeed, many of the 16 best markers were relatively easy to amplify with one or two primer pairs, with a PCR success ratio between 50 and 100%. Direct sequencing was facilitated by the high proportion of single bands produced in the PCR of these genes. Only occasional events of unspecific amplification occurred and most sequences could be aligned unambiguously and translated into amino acids.

Further optimization of primers is required to enable amplification across a broader range of weevils and other beetle groups. This is particularly relevant to the many unsuccessful genes that we screened, which may amplify with a better design of primers. In such a brief screening of candidate genes it is likely that promising markers were overlooked. The gene *α-spectrin*, as one example, may deserve further attention as one of very few genes previously screened for beetles [23]. Unfortunately, the primers designed in this study amplified mainly Xyleborini and Dryocoetini, but not the majority of other tribes. We also continued our previous screening of the *NaK* gene [9], which again was particularly positive for Ipini, Dryocoetini and Xyleborini, with potential application at lower level phylogeny.

Only one marker amplified in all samples (*PABP1*). This gene, and three additional ones (*TPI*, *UBA5* and *Prp1*) with comparable high amplification rates, shared a pattern of simple intron structure, which may facilitate the amplification process. Other genes could be almost as easily amplified (*Arr2*, *Iap2*, *CCNC*), but required more efforts in the alignment procedure due to the presence of highly variable regions and/or introns. For all the other genes, improved primer design seems required to obtain PCR and sequencing regularity at appreciable levels such as in nymphalid butterflies [32] or dolichoderine ants [39]. Suboptimal primer design was most evident in cases where failures in amplification were taxon-specific, for instance *TPI*, *HDAC Rpd3*, *Cda4* and *RCC1* in species of the tribe Scolytini. Other genes such as *ADA2*, *Hsp70*, *FEN1* and *CTR9* were amplifying Scolytinae, which was our main target group, but failed in most other weevils.

Degenerate primers tend to amplify non-targeted regions for several of the screened genes. However, only two genes with short amplified fragments (*SOD1* and *RCC1*) were regularly affected by this kind of problem, and occurred less frequently in *CXorf56*, *Hsp90* and *eRF1*. The amplification of other gene copies is a relatively common problem in PCR based methods and at least three routinely used markers (*COI*, *EF-1α*, *enolase*) in bark and ambrosia beetles are occasionally burdened with such complexity [10, 68, 75]. In other cases, such as *EF2* and *Hsp70*, the same gene copy was unintentionally amplified from other beetle-associated organisms (fungi and nematodes), probably due to the conserved nature of these genes [68]. When we tested nuclear markers for orthology assessment in arthropods (OrthoDB v9), *Hsp70* was one of the few genes which resulted present in multiple copies in the large majority of the

species in the database (98%). In our study, the presence of *Hsp70* paralogs was clearly demonstrated based on BLAST search, strongly deviating amino acid substitution patterns and long phylogenetic branches of paraphyletic groups (S4 Fig). Although three other genes (*Iap2*, *Arr2* and *SOD1*) are rarely in single copy in the arthropod genomes, our study did not provide any clear evidence of paralogy in beetles.

Two markers (*HDAC Rpd3* and *CCNC*) were particularly problematic due to the many long introns they contained (up to 5 in *HDAC Rpd3*) and they require internal primers for more effective amplification and sequencing. The presence of long and/or numerous introns seems widespread in beetles. This insect order has generally a higher number of introns compared to other insects [76], particularly so in the phytophagan beetles [23]. For example, a 300 bp short fragment of the gene *Wingless*, which is widely used in insect phylogeny, contains three complicated introns in weevils, but it is intron free in adephagan beetles and most other insect orders. On the other side are weevil sequences of *TPI* simpler than those of coccoidean Hemiptera which have two extra introns and one hypervariable indels region [77]. Only two introns were present in the majority of weevils, although highly variable in Hylurgini and four additional species. Similar situations, with lack of conserved intron patterns within clades, were observed for genes such as *CTR9*, *HDAC Rpd3* and *Cda4*, contrasting the long held argument that intron structure is a conserved and therefore useful phylogenetic marker [78, 79].

A further complicating feature in the alignments of *Arr2* and *Iap2* involved variable coding regions that contained different numbers of triplet nucleotide indels. Because indel-rich regions are difficult to align, they could potentially introduce unwarranted noise in the phylogenetic signal. However, the removal of these ambiguous regions did not affect tree topologies resulting from independent analyses of each of these genes. Indel-rich regions of *Arr2* occur in species from other insect orders (BLAST analyses), which further document natural and widespread variation in this trait. *Iap2* is much less known in terms of indels variation and our data were only comparable to other GenBank sequences in the second more conserved exon.

The process of evaluating and ranking different markers in terms of phylogenetic utility is a complex task. Rates are not always inversely correlated with phylogenetic resolution and clade support [80] and only the implementation in large taxonomic samples represents the ultimate test of a phylogenetic marker performance. Our gene classification based on phylogenetic utility that was assessed according to clade congruence and phylogenetic informativeness (PI) must therefore be taken as a preliminary proxy for a marker's phylogenetic signal [81, 82, 83]. It will be particularly interesting to observe the contribution of *Iap2* in a larger data set given its much higher PI compared with other markers. *Iap2* is a fast evolving gene which, likewise *TPI*, *Prp1*, *FEN1* and *Arr2*, showed a high peak for the Miocene epoch, but it differs from the other genes by maintaining a stronger phylogenetic signal over time. Even though this marker has two variable regions that could have biased the PI profile estimate, the average level of homoplasy was also the lowest for this gene. On the other hand, the tree topology resulting from the phylogenetic analyses was not particularly congruent with previously established relationships.

Only one gene (*FEN1*) produced a tree topology that was largely congruent with all predefined clades, and only three genes (*PABP1*, *FEN1*, *Arr2*) were congruent with the most recent split—between Ipini and Dryocoetini/Xyleborini (Paleocene age)—indicating high substitution rates for most genes in our screening. However, a perfect match between a gene tree and the species tree is rarely observed [84]. Dense taxon sampling and simultaneous analyses of many genes will usually overcome such limitations, building on the hidden support from many genes not visible in single gene analyses [85, 86].

Large amounts of data are usually required to obtain resolution between more ancient groups such as insect orders and families. It is therefore a possibility that 15–20 markers are not sufficient to resolve the weevil phylogeny, including relationships among bark and

ambrosia beetles. Data volume is by itself useful as demonstrated by studies on the complete mitochondrial genome of weevils that resolve certain parts of the tree topology [6, 87]. Limiting mitochondrial data to a handful of genes illustrates this point well as resolution fades rapidly [8]. Larger data volumes are now available from nuclear genome sequencing, either in terms of entire genomes [88–90], or transcribed genomes [91, 92]. Each of these approaches has their own disadvantages with respect to high cost and labor intensity. Transcriptome data are furthermore burdened with highly biased gene expressions, for instance the overexpression of ribosomal proteins in ESTs of beetles [93]. A targeted PCR-based approach to sequencing has on these grounds been recommended in phylogenetic analyses [94].

New NGS technologies have lately enabled more specific amplification of conserved sequence regions, bypassing complete genomic or transcriptomic assembly, and thereby reducing the dataset to a core of comparable informative sequences which are more suitable for phylogenetics [95, 96]. Sequence capture of ultra-conserved elements (UCEs) has enabled high sequence homology [51, 97, 98] and hence, these results are more directly comparable to PCR based sequences. UCEs have a phylogenetic information potential comparable to protein coding genes at the per nucleotide level; however, the large volume of data involving hundreds of loci and more than 100,000 nucleotides provide better resolution and higher support at deep phylogenetic level [99, 100].

It is increasingly being argued that PCR-based methods are becoming redundant in the age of NGS, but this is largely an overstatement. Most sequencing, in fact, occurs at a routine basis, as a tool in integrative taxonomy where a handful of sequences from established markers are sufficient to place a new species in the tree of life. Most laboratories in the world are not yet rigged for the latest NGS in terms of equipment, labor and budgetary concerns. As long as the monthly turnaround rate involves less than 10 genes and 100 taxa, the time and cost doing traditional PCR and sequencing is much lower [100]. Recognizing that small data sets are not only less expensive, but also can be sufficiently informative, the reliance on PCR and Sanger sequencing will continue as the best option for many small scale studies also in the future. In fact, modest data sets of a few thousands of nucleotides (5–10 genes) can be almost as informative as large collections of UCEs [100, 101]. With approximately 80–90% congruence in topology, one may reconsider if sequencing of UCEs is always the best option despite the generally higher node support obtained for this type of data.

Conclusion

This study has revealed the many difficulties in selecting and optimizing new markers for weevil phylogenetics. Other beetle groups may be less problematic than weevils [23], but beetles in general are much more challenging in this respect as compared to Hymenoptera and Lepidoptera [32, 36, 86, 102]. Nevertheless, this study provides a step forward in PCR-based sequencing of beetles and we hope that these new markers will provide a useful toolbox for beetle phylogenetics, particularly in studies on more recent divergences where a limited amount of genetic data can enable accurate inference of past evolutionary events.

Supporting Information

S1 Fig. Phylogenetic tree based on Bayesian analyses of 16 concatenated genes both for nucleotides and amino acids. Posterior probability values are reported below the node for the nucleotides analysis (8109 bp), while the pp values above the node refer to the amino acids analysis (2702 aa).
(TIF)

S2 Fig. Phylogentic tree based on parsimony analyses of 16 concatenated genes both for nucleotides and amino acids. Bootstrap support values are reported below the node for the nucleotides analysis (8109 bp), while the values above the node indicate the bootstrap support for amino acids analysis (2702 aa).
(TIF)

S3 Fig. Average genetic variation for each marker. *p*-distance values for each position and for each gene were calculated across the entire sample, excluding Brentidae to avoid missing data.
(TIF)

S4 Fig. Phylogenetic tree based on a fragment of the gene *Hsp70*. Results of Bayesian analysis based on *Hsp70* sequences of weevils and Scolytinae; three different copies of *D. ponderosae Hsp70* were included in order to test for paralogs. Six more species were also included in the analysis (CuSib01 = *Sibinia* sp. CgAph02 = *Aphanarthrum capense*, MiLan01 = *Lanurgus xylographus*, MoAmo01 = *Amorphocerus rufipes*, DrCyr02 = *Acanthotomicus* sp. and TsCen01 = *Cenocephalus* sp.). Three different *Hsp70* groups were identified. One group consisted of paralogous copies of *Hsp70* (A), plus two clusters of sequences from fungi (B) and nematodes (C).
(TIF)

S1 File. XML file used for analyses in BEAST v1.8.2. The file was generated using BEAUTI v 1.8.2.
(XML)

S2 File. Additional information on 16 PCR amplified and sequenced genes.
(DOCX)

S1 Table. GenBank accession numbers for each of the 16 selected genes sequenced in this study.
(DOCX)

S2 Table. Estimates of evolutionary divergence (*p*-distance) between sequences. For each of the 16 genes, the proportion of different nucleotide sites between sequences was calculated. The most frequently PCR amplified species (*Xyleborus affinis*) was compared with members of the other tribes and subfamilies and the lower value was reported. PIC = Parsimony informative characters, HI = Homoplasy index and RI = Retention index.
(DOCX)

S3 Table. Information on markers not developed for higher level phylogenetics. The main problems for further development are reported, together with data on fragment length, and number and length of introns for 18 of these markers. The same information could not be derived for markers with low number of sequences.
(DOCX)

Acknowledgments

This project was funded by grant 214232/F20 from the Norwegian Research Council.

Author Contributions

Conceptualization: BHJ DP.

Data curation: BHJ DP.

Formal analysis: DP.

Funding acquisition: BHJ.

Investigation: DP SM.

Methodology: BHJ DP.

Project administration: BHJ.

Resources: BHJ.

Supervision: BHJ.

Validation: BHJ.

Visualization: DP SM.

Writing – original draft: BHJ DP.

Writing – review & editing: BHJ DP.

References

1. Maddison DR, Moore W, Baker MD, Ellis TM, Ober KA, Cannone JJ, et al. Monophyly of terrestrial adephagan beetles as indicated by three nuclear genes (Coleoptera: Carabidae and Trachypachidae). Zool Scr. 2009; 38(1): 43–62. Epub 2009/10/01. doi: 10.1111/j.1463-6409.2008.00359.x PMID: 19789725; PubMed Central PMCID: PMCPmc2752903.
2. Pons J, Ribera I, Bertranpetit J, Balke M. Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. Molecular phylogenetics and evolution. 2010; 56(2): 796–807. Epub 2010/02/16. doi: 10.1016/j.ympev.2010.02.007 PMID: 20152911.
3. Hunt T, Bergsten J, Levkanicova Z, Papadopoulou A, John OS, Wild R, et al. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. Science. 2007; 318(5858): 1913–1916. Epub 2007/12/22. doi: 10.1126/science.1146954 PMID: 18096805.
4. Shull VL, Vogler AP, Baker MD, Maddison DR, Hammond PM. Sequence alignment of 18S ribosomal RNA and the basal relationships of Adephagan beetles: evidence for monophyly of aquatic families and the placement of Trachypachidae. Syst Biol. 2001; 50(6): 945–969. Epub 2002/07/16. PMID: 12116642.
5. McKenna DD, Sequeira AS, Marvaldi AE, Farrell BD. Temporal lags and overlap in the diversification of weevils and flowering plants. Proc Natl Acad Sci U S A. 2009; 106(17): 7083–7088. Epub 2009/04/15. doi: 10.1073/pnas.0810618106 PMID: 19365072; PubMed Central PMCID: PMCPmc2678426.
6. Gillett CP, Crampton-Platt A, Timmermans MJ, Jordal BH, Emerson BC, Vogler AP. Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). Molecular biology and evolution. 2014; 31(8): 2223–2237. Epub 2014/05/08. doi: 10.1093/molbev/msu154 PMID: 24803639; PubMed Central PMCID: PMCPmc4104315.
7. Jordal BH, Sequeira AS, Cognato AI. The age and phylogeny of wood boring weevils and the origin of subsociality. Molecular phylogenetics and evolution. 2011; 59(3): 708–724. Epub 2011/03/26. doi: 10.1016/j.ympev.2011.03.016 PMID: 21435394.
8. Gunter N, Oberprieler R, and Cameron S. Molecular phylogenetics of Australian weevils (Coleoptera: Curculionoidea): exploring relationships in a hyperdiverse lineage through comparison of independent analyses. Austral Entomology. 2015.
9. Jordal BH. Reconstructing the Phylogeny of Scolytinae and Close Allies: Major Obstacles and Prospects for a Solution. Proceedings RMRS. 2007.
10. Farrell BD, Sequeira AS, O'Meara BC, Normark BB, Chung JH, Jordal BH. The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). Evolution. 2001; 55(10): 2011–2027. Epub 2002/01/05. PMID: 11761062.
11. Jordal BH, Cognato AI. Molecular phylogeny of bark and ambrosia beetles reveals multiple origins of fungus farming during periods of global warming. BMC Evol Biol. 2012; 12: 133–140. Epub 2012/08/03. doi: 10.1186/1471-2148-12-133 PMID: 22852794; PubMed Central PMCID: PMCPmc3514184.
12. Sikes DS, Venables C. Molecular phylogeny of the burying beetles (Coleoptera: Silphidae: Nicrophorinae). Molecular phylogenetics and evolution. 2013; 69(3): 552–565. Epub 2013/08/06. doi: 10.1016/j.ympev.2013.07.022 PMID: 23911726.

13. Becerra JX. Molecular systematics of Blepharida beetles (Chrysomelidae: Alticinae) and relatives. Molecular phylogenetics and evolution. 2004; 30(1): 107–117. Epub 2004/03/17. PMID: 15022762.

14. Duan Y, Kerdelhue C, Ye H, Lieutier F. Genetic study of the forest pest *Tomicus piniperda* (Col., Scolytinae) in Yunnan province (China) compared to Europe: new insights for the systematics and evolution of the genus *Tomicus*. Heredity (Edinb). 2004; 93(5): 416–422. Epub 2004/07/29. doi: 10.1038/sj.hdy.6800518 PMID: 15280894.

15. Shaw-Lee RL, Lissemore JL, Sullivan DT. Structure and expression of the triose phosphate isomerase (Tpi) gene of *Drosophila melanogaster*. Mol Gen Genet. 1991; 230(1–2): 225–229. Epub 1991/11/01. PMID: 1720860.

16. Dole SA, Jordal BH, Cognato AI. Polyphyly of Xylosandrus Reitter inferred from nuclear and mitochondrial genes (Coleoptera: Curculionidae: Scolytinae). Molecular phylogenetics and evolution. 2010; 54(3): 773–782. Epub 2009/11/21. doi: 10.1016/j.ympev.2009.11.011 PMID: 19925873.

17. Maddison DR. An unexpected clade of South American ground beetles (Coleoptera, Carabidae, Bembidion). Zookeys. 2014; (416:): 113–155. Epub 2014/07/26. doi: 10.3897/zookeys.416.7706 PMID: 25061349; PubMed Central PMCID: PMCPmc4109512.

18. Zaspel JM, Weller SJ, Wardwell CT, Zahiri R, Wahlberg N. Phylogeny and evolution of pharmacophagy in tiger moths (Lepidoptera: Erebidae: Arctiinae). PLoS One. 2014; 9(7): e101975. Epub 2014/07/19. doi: 10.1371/journal.pone.0101975 PMID: 25036028; PubMed Central PMCID: PMCPmc4103773.

19. Schmidt C. Molecular phylogenetics of ponerine ants (Hymenoptera: Formicidae: Ponerinae). Zootaxa. 2013; 3647: 201–250. Epub 2013/01/01. PMID: 26295106.

20. Lucky A. Molecular phylogeny and biogeography of the spider ants, genus *Leptomyrmex* Mayr (Hymenoptera: Formicidae). Molecular phylogenetics and evolution. 2011; 59(2): 281–292. Epub 2011/03/17. doi: 10.1016/j.ympev.2011.03.004 PMID: 21406240.

21. Urban JM, Cryan JR. Entomologically famous, evolutionarily unexplored: the first phylogeny of the lanternfly family Fulgoridae (Insecta: Hemiptera: Fulgoroidea). Molecular phylogenetics and evolution. 2009; 50(3): 471–484. Epub 2009/01/03. doi: 10.1016/j.ympev.2008.12.004 PMID: 19118634.

22. Angelini DR, Jockusch EL. Relationships among pest flour beetles of the genus Tribolium (Tenebrionidae) inferred from multiple molecular markers. Molecular phylogenetics and evolution. 2008; 46(1): 127–141. Epub 2007/11/21. doi: 10.1016/j.ympev.2007.08.017 PMID: 18024090; PubMed Central PMCID: PMCPmc2292397.

23. Wild AL, Maddison DR. Evaluating nuclear protein-coding genes for phylogenetic utility in beetles. Molecular phylogenetics and evolution. 2008; 48(3): 877–891. Epub 2008/07/23. doi: 10.1016/j.ympev.2008.05.023 PMID: 18644735.

24. Pons J, Barraclough T, Theodorides K, Cardoso A, Vogler A. Using exon and intron sequences of the gene *Mp20* to resolve basal relationships in *Cicindela* (Coleoptera:Cicindelidae). Syst Biol. 2004; 53(4): 554–570. Epub 2004/09/17. doi: 10.1080/10635150490472940 PMID: 15371246.

25. Maddison DR. Phylogeny of Bembidion and related ground beetles (Coleoptera: Carabidae: Trechinae: Bembidiini: Bembidiina). Molecular phylogenetics and evolution. 2012; 63(3): 533–576. Epub 2012/03/17. doi: 10.1016/j.ympev.2012.01.015 PMID: 22421212.

26. McKenna DD, Wild AL, Kanda K, Bellamy CL, Beutel RG, Caterino MS, et al. The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. Systematic Entomology. 2015; 40(4): 835–880. doi: 10.1111/syen.12132

27. Wiegmann BM, Trautwein MD, Kim JW, Cassel BK, Bertone MA, Winterton SL, et al. Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. BMC Biol. 2009; 7: 34. Epub 2009/06/26. doi: 10.1186/1741-7007-7-34 PMID: 19552814; PubMed Central PMCID: PMCPmc2709105.

28. Russo CA, Takezaki N, Nei M. Molecular phylogeny and divergence times of drosophilid species. Molecular biology and evolution. 1995; 12(3): 391–404. Epub 1995/05/01. PMID: 7739381.

29. Cho S, Mitchell A, Regier JC, Mitter C, Poole RW, Friedlander TP, et al. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 alpha recovers morphology-based tree for heliothine moths. Molecular biology and evolution. 1995; 12(4): 650–666. Epub 1995/07/01. PMID: 7659020.

30. Friedlander TP, Regier JC, Mitter C. Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. Syst Biol. 1992; 41: 483–490.

31. Regier JC, Mitter C, Zwick A, Bazinet AL, Cummings MP, Kawahara AY, et al. A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). PLoS One. 2013; 8(3): e58568. Epub 2013/04/05. doi: 10.1371/journal.pone.0058568 PMID: 23554903; PubMed Central PMCID: PMCPmc3595289.

32. Wahlberg N, Wheat CW. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of lepidoptera. Syst Biol. 2008; 57(2): 231–242. Epub 2008/04/10. doi: 10.1080/10635150802033006 PMID: 18398768.

33. Ohshima I, Tanikawa-Dodo Y, Saigusa T, Nishiyama T, Kitani M, Hasebe M, et al. Phylogeny, biogeography, and host-plant association in the subfamily Apaturinae (Insecta: Lepidoptera: Nymphalidae) inferred from eight nuclear and seven mitochondrial genes. Molecular phylogenetics and evolution. 2010; 57(3): 1026–1036. Epub 2010/10/06. doi: 10.1016/j.ympev.2010.09.018 PMID: 20920592.

34. Aduse-Poku K, Brattstrom O, Kodandaramaiah U, Lees DC, Brakefield PM, Wahlberg N. Systematics and historical biogeography of the old world butterfly subtribe Mycalesina (Lepidoptera: Nymphalidae: Satyrinae). BMC Evol Biol. 2015; 15: 167. Epub 2015/08/21. doi: 10.1186/s12862-015-0449-3 PMID: 26289424; PubMed Central PMCID: PMCPmc4545879.

35. Campbell DL, Brower AV, Pierce NE. Molecular evolution of the wingless gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera: papilionoidea). Molecular biology and evolution. 2000; 17(5): 684–696. Epub 2000/04/26. PMID: 10779529.

36. Sohn J-C, Regier JC, Mitter C, Adamski D, Landry J-F, HeikkilÄ M, et al. Phylogeny and feeding trait evolution of the mega-diverse Gelechioidea (Lepidoptera: Obtectomera): new insight from 19 nuclear genes. Systematic Entomology. 2016; 41(1): 112–132. doi: 10.1111/syen.12143

37. Hedtke SM, Patiny S, Danforth BN. The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography. BMC Evol Biol. 2013; 13: 138. Epub 2013/07/05. doi: 10.1186/1471-2148-13-138 PMID: 23822725; PubMed Central PMCID: PMCPmc3706286.

38. Danforth BN, Cardinal S, Praz C, Almeida EA, Michez D. The impact of molecular data on our understanding of bee phylogeny and evolution. Annu Rev Entomol. 2013; 58: 57–78. Epub 2012/09/01. doi: 10.1146/annurev-ento-120811-153633 PMID: 22934982.

39. Ward PS, Brady SG, Fisher BL, Schultz TR. Phylogeny and biogeography of dolichoderine ants: effects of data partitioning and relict taxa on historical inference. Syst Biol. 2010; 59(3): 342–362. Epub 2010/06/09. doi: 10.1093/sysbio/syq012 PMID: 20525640.

40. Zhao L, Annie AS, Amrita S, Yi SK, Rudolf M. Does better taxon sampling help? A new phylogenetic hypothesis for Sepsidae (Diptera: Cyclorrhapha) based on 50 new taxa and the same old mitochondrial and nuclear markers. Molecular phylogenetics and evolution. 2013; 69(1): 153–164. Epub 2013/05/28. doi: 10.1016/j.ympev.2013.05.011 PMID: 23707858.

41. Winkler IS, Blaschke JD, Davis DJ, Stireman JO 3rd, O'Hara JE, Cerretti P, et al. Explosive radiation or uninformative genes? Origin and early diversification of tachinid flies (Diptera: Tachinidae). Molecular phylogenetics and evolution. 2015; 88: 38–54. Epub 2015/04/07. doi: 10.1016/j.ympev.2015.03.021 PMID: 25841383.

42. Gibson JF, Kelso S, Jackson MD, Kits JH, Miranda GFG, Skevington JH. Diptera-Specific Polymerase Chain Reaction Amplification Primers of Use in Molecular Phylogenetic Research. Annals of the Entomological Society of America. 2011; 104(5): 976–997. http://dx.doi.org/10.1603/AN10153.

43. Ogden TH, Whiting MF. Phylogeny of Ephemeroptera (mayflies) based on molecular evidence. Molecular phylogenetics and evolution. 2005; 37(3): 625–643. Epub 2005/10/11. doi: 10.1016/j.ympev.2005.08.008 PMID: 16214375.

44. Ishiwata K, Sasaki G, Ogawa J, Miyata T, Su ZH. Phylogenetic relationships among insect orders based on three nuclear protein-coding gene sequences. Molecular phylogenetics and evolution. 2011; 58(2): 169–80. Epub 2010/11/16. doi: 10.1016/j.ympev.2010.11.001 PMID: 21075208.

45. Chang ES, Neuhof M, Rubinstein ND, Diamant A, Philippe H, Huchon D, et al. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc Natl Acad Sci U S A. 2015; 112(48): 14912–14917. Epub 2015/12/03. doi: 10.1073/pnas.1511468112 PMID: 26627241.

46. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014; 346(6210): 763–777. Epub 2014/11/08. doi: 10.1126/science.1257570 PMID: 25378627.

47. Kawahara AY, Breinholt JW. Phylogenomics provides strong evidence for relationships of butterflies and moths. Proceedings Biological sciences / The Royal Society. 2014; 281(1788): 20140970. Epub 2014/06/27. doi: 10.1098/rspb.2014.0970 PMID: 24966318; PubMed Central PMCID: PMCPmc4083801.

48. Johnson BR, Borowiec ML, Chiu JC, Lee EK, Atallah J, Ward PS. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. Current biology: CB. 2013; 23(20): 2058–2062. Epub 2013/10/08. doi: 10.1016/j.cub.2013.08.050 PMID: 24094856.

49. Peters RS, Meusemann K, Petersen M, Mayer C, Wilbrandt J, Ziesmann T, et al. The evolutionary history of holometabolous insects inferred from transcriptome-based phylogeny and comprehensive morphological data. BMC Evol Biol. 2014; 14(1): 52. Epub 2014/03/22. doi: 10.1186/1471-2148-14-52 PMID: 24646345; PubMed Central PMCID: PMCPmc4000048.

50. Savard J, Tautz D, Richards S, Weinstock GM, Gibbs RA, Werren JH, et al. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. Genome Res. 2006; 16(11): 1334–1338. Epub 2006/10/27. doi: 10.1101/gr.5204306 PMID: 17065606; PubMed Central PMCID: PMCPmc1626634.

51. McCormack JE, Faircloth BC, Crawford NG, Gowaty PA, Brumfield RT, Glenn TC. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 2012; 22(4): 746–754. doi: 10.1101/gr.125864.111 PMID: 22207614; PubMed Central PMCID: PMCPMC3317156.

52. Cruaud A, Gautier M, Galan M, Foucaud J, Saune L, Genson G, et al. Empirical assessment of RAD sequencing for interspecific phylogeny. Molecular biology and evolution. 2014; 31(5): 1272–1274. Epub 2014/02/06. doi: 10.1093/molbev/msu063 PMID: 24497030.

53. Hulcr J, Atkinson T, Cognato A, Jordal B, McKenna D. Morphology, Taxonomy and Phylogenetics of Bark Beetles. Edited by: Vega FE and Hofstetter RW. Bark Beetles: Biology and Ecology of Native and Invasive Species. 2014. Pp: 41–84.

54. Marvaldi AE, Sequeira AS, O'Brien CW, Farrell BD. Molecular and morphological phylogenetics of weevils (Coleoptera, Curculionoidea): do niche shifts accompany diversification? Syst Biol. 2002; 51 (5): 761–785. Epub 2002/10/25. doi: 10.1080/10635150290102465 PMID: 12396590.

55. Oberprieler RG, Marvaldi AE, Anderson RS. Weevils, weevils, weevils everywhere. Zootaxa. 2007; (1668:): 491–520. WOS:000251791200023.

56. Kirkendall L. The evolution of mating systems in bark and ambrosia beetles (Coleoptera: Scolytidae and Platypodidae). Zoological journal of the Linnenn Society. 1983; 77: 293–352.

57. Kirkendall L, Biedermann P, Jordal B. Diversity and evolution of Bark Beetles. Edited by: Vega FE and Hofstetter RW. Bark Beetles: Biology and Ecology of Native and Invasive Species 2014. Pp: 85–156.

58. Jordal BH. Molecular phylogeny and biogeography of the weevil subfamily Platypodinae reveals evolutionarily conserved range patterns. Molecular phylogenetics and evolution. 2015; 92: 294–307. Epub 2015/07/21. doi: 10.1016/j.ympev.2015.05.028 PMID: 26190520.

59. Swofford DL. PAUP*: phylogenetic analysis using parsimony, version 4.0b10. 2011. citeulike-article-id:2347167.

60. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 1999; 41: 95–98.

61. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular biology and evolution. 2013; 30(4): 772–780. doi: 10.1093/molbev/mst010 PMID: 23329690

62. Kitamura-Abe S, Itoh H, Washio T, Tsutsumi A, Tomita M. Characterization of the splice sites in GT-AG and GC-AG introns in higher eukaryotes using full-length cDNAs. J Bioinform Comput Biol. 2004; 2(2): 309–331. Epub 2004/08/07. PMID: 15297984.

63. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic acids research. 2013; 41(Database issue): D358–365. Epub 2012/11/28. doi: 10.1093/nar/gks1116 PMID: 23180791; PubMed Central PMCID: PMCPmc3531149.

64. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, et al. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic acids research. 2015; 43(Database issue): D250–256. Epub 2014/11/28. doi: 10.1093/nar/gku1220 PMID: 25428351; PubMed Central PMCID: PMCPmc4383991.

65. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19(12): 1572–1574. doi: 10.1093/bioinformatics/btg180 PMID: 12912839

66. Posada D. jModelTest: phylogenetic model averaging. Molecular biology and evolution. 2008; 25(7): 1253–1256. Epub 2008/04/10. doi: 10.1093/molbev/msn083 PMID: 18397919.

67. Rambaut A, Drummond A. Tracer version 1.4. Computer program and documentation distributed by the author, website http://beastbioedacuk/Tracer. 2007.

68. Jordal BH. Elongation Factor 1 alpha resolves the monophyly of the haplodiploid ambrosia beetles Xyleborini (Coleoptera: Curculionidae). Insect Mol Biol. 2002; 11(5): 453–465. Epub 2002/09/17. PMID: 12230544.

69. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution. 2013; 30(12): 2725–2729. Epub 2013/10/18. doi: 10.1093/molbev/mst197 PMID: 24132122; PubMed Central PMCID: PMCPmc3840312.

70. Lopez-Giraldez F, Townsend JP. PhyDesign: an online application for profiling phylogenetic informativeness. BMC Evol Biol. 2011; 11: 152. Epub 2011/06/02. doi: 10.1186/1471-2148-11-152 PMID: 21627831; PubMed Central PMCID: PMCPmc3124428.

71. Townsend JP. Profiling phylogenetic informativeness. Syst Biol. 2007; 56(2): 222–231. Epub 2007/04/28. doi: 10.1080/10635150701311362 PMID: 17464879.

72. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution. 2012; 29(8): 1969–1973. Epub 2012/03/01. doi: 10.1093/molbev/mss075 PMID: 22367748; PubMed Central PMCID: PMCPmc3408070.

73. Caterino MS, Cho S, Sperling FA. The current state of insect molecular systematics: a thriving Tower of Babel. Annu Rev Entomol. 2000; 45: 1–54. Epub 2000/04/13. doi: 10.1146/annurev.ento.45.1.1 PMID: 10761569.

74. Ward PS, Brady SG, Fisher BL, Schultz TR. The evolution of myrmicine ants: phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). Systematic Entomology. 2015; 40 (1): 61–81. doi: 10.1111/syen.12090

75. Jordal BH, Kambestad M. DNA barcoding of bark and ambrosia beetles reveals excessive NUMTs and consistent east-west divergence across Palearctic forests. Mol Ecol Resour. 2014; 14(1): 7–17. Epub 2013/08/08. doi: 10.1111/1755-0998.12150 PMID: 23919425.

76. Dolezelova E, Zurovec M, Bohmova M, Sehnal F. Use of two transcription starts in the G6PD gene of the bark beetle Ips typographus. Insect Mol Biol. 2006; 15(1): 25–32. Epub 2006/02/14. doi: 10.1111/j.1365-2583.2006.00604.x PMID: 16469065.

77. Hardy NB. Phylogenetic utility of dynamin and triose phosphate isomerase. Systematic Entomology. 2007; 32(2): 396–403. doi: 10.1111/j.1365-3113.2007.00377.x

78. Danforth BN, Ji S. Elongation factor-1 alpha occurs as two copies in bees: implications for phylogenetic analysis of EF-1 alpha sequences in insects. Molecular biology and evolution. 1998; 15(3): 225–235. Epub 1998/03/21. PMID: 9501490.

79. Rokas A, Holland PW. Rare genomic changes as a tool for phylogenetics. Trends Ecol Evol. 2000; 15(11): 454–459. Epub 2000/10/26. PMID: 11050348.

80. Källersjö M, Albert VA, Farris JS. Homoplasy increases phylogenetic structur. Cladistics. 1999; (15:): 91–93.

81. Dornburg A, Townsend JP, Friedman M, Near TJ. Phylogenetic informativeness reconciles ray-finned fish molecular divergence times. BMC Evol Biol. 2014; 14: 169. Epub 2014/08/12. doi: 10.1186/s12862-014-0169-0 PMID: 25103329; PubMed Central PMCID: PMCPmc4236503.

82. Lopez-Giraldez F, Moeller AH, Townsend JP. Evaluating phylogenetic informativeness as a predictor of phylogenetic signal for metazoan, fungal, and mammalian phylogenomic data sets. Biomed Res Int. 2013; 2013: 621604. Epub 2013/07/24. doi: 10.1155/2013/621604 PMID: 23878813; PubMed Central PMCID: PMCPmc3708382.

83. Hilu KW, Black CM, Oza D. Impact of gene molecular evolution on phylogenetic reconstruction: a case study in the rosids (Superorder Rosanae, Angiosperms). PLoS One. 2014; 9(6): e99725. Epub 2014/06/17. doi: 10.1371/journal.pone.0099725 PMID: 24932884; PubMed Central PMCID: PMCPmc4059714.

84. Degnan JH, Rosenberg NA. Discordance of species trees with their most likely gene trees. PloS Genet. 2006; 2(5): e68. Epub 2006/05/31. doi: 10.1371/journal.pgen.0020068 PMID: 16733550; PubMed Central PMCID: PMCPmc1464820.

85. Gatesy J, Baker RH. Hidden likelihood support in genomic data: can forty-five wrongs make a right? Syst Biol. 2005; 54(3): 483–492. Epub 2005/07/14. doi: 10.1080/10635150590945368 PMID: 16012113.

86. Kawahara AY, Ohshima I, Kawakita A, Regier JC, Mitter C, Cummings MP, et al. Increased gene sampling strengthens support for higher-level groups within leaf-mining moths and relatives (Lepidoptera: Gracillariidae). BMC Evol Biol. 2011; 11: 182. Epub 2011/06/28. doi: 10.1186/1471-2148-11-182 PMID: 21702958; PubMed Central PMCID: PMCPmc3145599.

87. Crampton-Platt A, Timmermans MJ, Gimmel ML, Kutty SN, Cockerill TD, Vun Khen C, et al. Soup to Tree: The Phylogeny of Beetles Inferred by Mitochondrial Metagenomics of a Bornean Rainforest Sample. Molecular biology and evolution. 2015; 32(9): 2302–2316. Epub 2015/05/10. doi: 10.1093/molbev/msv111 PMID: 25957318; PubMed Central PMCID: PMCPmc4540967.

88. Keeling CI, Yuen MM, Liao NY, Docking TR, Chan SK, Taylor GA, et al. Draft genome of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major forest pest. Genome Biol. 2013; 14(3): R27. Epub 2013/03/30. doi: 10.1186/gb-2013-14-3-r27 PMID: 23537049; PubMed Central PMCID: PMCPmc4053930.

89. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, et al. The genome of the model beetle and pest *Tribolium castaneum*. Nature. 2008; 452(7190): 949–955. Epub 2008/03/26. doi: 10.1038/nature06784 PMID: 18362917.

90. Vega FE, Brown SM, Chen H, Shen E, Nair MB, Ceja-Navarro JA, et al. Draft genome of the most devastating insect pest of coffee worldwide: the coffee berry borer, *Hypothenemus hampei*. Sci Rep. 2015; 5: 12525. Epub 2015/08/01. doi: 10.1038/srep12525 PMID: 26228545; PubMed Central PMCID: PMCPmc4521149.

91. Yin A, Pan L, Zhang X, Wang L, Yin Y, Jia S, et al. Transcriptomic study of the red palm weevil *Rhynchophorus ferrugineus* embryogenesis. Insect Sci. 2015; 22(1): 65–82. Epub 2013/12/19. doi: 10.1111/1744-7917.12092 PMID: 24347559.

92. Firmino AA, Fonseca FC, de Macedo LL, Coelho RR, Antonino de Souza JD Jr., Togawa RC, et al. Transcriptome analysis in cotton boll weevil (*Anthonomus grandis*) and RNA interference in insect pests. PLoS One. 2013; 8(12): e85079. Epub 2014/01/05. doi: 10.1371/journal.pone.0085079 PMID: 24386449; PubMed Central PMCID: PMCPmc3874031.

93. Hughes J, Longhorn SJ, Papadopoulou A, Theodorides K, de Riva A, Mejia-Chang M, et al. Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (beetles). Molecular biology and evolution. 2006; 23(2): 268–278. Epub 2005/10/21. doi: 10.1093/molbev/msj041 PMID: 16237206.

94. Philippe H, Brinkmann H, Lavrov DV, Littlewood DT, Manuel M, Worheide G, et al. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 2011; 9(3): e1000602. Epub 2011/03/23. doi: 10.1371/journal.pbio.1000602 PMID: 21423652; PubMed Central PMCID: PMCPmc3057953.

95. Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics. 2012; 13: 403. Epub 2012/08/21. doi: 10.1186/1471-2164-13-403 PMID: 22900609; PubMed Central PMCID: PMCPmc3472323.

96. Lemmon AR, Emme SA, Lemmon EM. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst Biol. 2012; 61(5): 727–744. Epub 2012/05/19. doi: 10.1093/sysbio/sys049 PMID: 22605266.

97. Faircloth BC, Branstetter MG, White ND, Brady SG. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol Ecol Resour. 2015; 15(3): 489–501. Epub 2014/09/11. doi: 10.1111/1755-0998.12328 PMID: 25207863; PubMed Central PMCID: PMCPmc4407909.

98. Faircloth BC, Sorenson L, Santini F, Alfaro ME. A Phylogenomic Perspective on the Radiation of Ray-Finned Fishes Based upon Targeted Sequencing of Ultraconserved Elements (UCEs). PLoS One. 2013; 8(6): e65923. Epub 2013/07/05. doi: 10.1371/journal.pone.0065923 PMID: 23824177; PubMed Central PMCID: PMCPmc3688804.

99. Gilbert PS, Chang J, Pan C, Sobel EM, Sinsheimer JS, Faircloth BC, et al. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Molecular phylogenetics and evolution. 2015; 92: 140–146. doi: 10.1016/j.ympev.2015.05.027 PMID: 26079130; PubMed Central PMCID: PMCPMC4583375.

100. Blaimer BB, Brady SG, Schultz TR, Lloyd MW, Fisher BL, Ward PS. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol Biol. 2015; 15(1): 271. Epub 2015/12/08. doi: 10.1186/s12862-015-0552-5 PMID: 26637372; PubMed Central PMCID: PMCPmc4670518.

101. Ruane S, Raxworthy CJ, Lemmon AR, Lemmon EM, Burbrink FT. Comparing species tree estimation with large anchored phylogenomic and small Sanger-sequenced molecular datasets: an empirical study on Malagasy pseudoxyrhophiine snakes. BMC Evol Biol. 2015; 15: 221. Epub 2015/10/16. doi: 10.1186/s12862-015-0503-1 PMID: 26459325; PubMed Central PMCID: PMCPmc4603904.

102. Brady SG, Fisher BL, Schultz TR, Ward PS. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol Biol. 2014; 14: 93. Epub 2014/06/03. doi: 10.1186/1471-2148-14-93 PMID: 24886136; PubMed Central PMCID: PMCPmc4021219.

II

Systematic Entomology

Systematic Entomology (2017), DOI: 10.1111/syen.12281

Molecular phylogeny of bark and ambrosia beetles (Curculionidae: Scolytinae) based on 18 molecular markers

DARIO PISTONE, JOSTEIN GOHLI and BJARTE H. JORDAL

University Museum of Bergen, University of Bergen, Bergen, Norway

Abstract. The phylogeny of the large weevil subfamily Scolytinae has been difficult to resolve based on a limited number of genetic markers. With more than 6000 nominal species in the subfamily, the general lack of resolution at deeper nodes indicates that large sequence volumes are needed to solve this problem. We have therefore assembled a large molecular dataset consisting of more than 10 kb of nucleotides from 18 gene fragments, for 182 species. Nucleotide and amino acid translated data were analysed using Bayesian and parsimony-based approaches, which gave largely congruent results. Compared with previous analyses, we obtained greater resolution for some of the deeper nodes, and detected many unexpected relationships that were strongly supported by our data. The tribe Scolytini was recovered as the earliest divergent lineage in Scolytinae, sometimes placed together with the hexacoline genus *Microborus*. Among the 26 currently recognized tribes, 15 were monophyletic, whereas the remaining tribes were largely paraphyletic. The majority of species in the tribe Hypoborini were recovered as the sister lineage to a large group containing the species-rich tribe Dryocoetini, which includes the recently radiated ambrosia beetles in Xyleborini, and Ipini, which includes another recent group of ambrosia beetles in Premnobiina. Cryphalini, Hylesinini and Hylurgini were strikingly polyphyletic tribes, each consisting of several independent lineages. Subgroups were to a large degree defined by geographical affinities, showing a clear distinction between the northern and southern hemispheres. The affiliation of the inbreeding genus *Hypothenemus* was revealed with strong support as the sister group to the Malagasy and East African species of the genus *Cosmoderes*. *Cryptocarenus* was previously assumed to be the sister lineage of *Hypothenemus*, but was here found to be part of Corthylini, near *Araptus*. These and many other findings document the need for a thorough revision of the current classification of genera and tribes, including a systematic re-evaluation of morphological characters.

Introduction

The weevil subfamily Scolytinae Latreille constitutes a highly diversified group of beetles with more than 6000 described species, currently grouped into 26 tribes and 246 genera (Hulcr *et al.*, 2015). These insects are found in fairly equal proportions on all forested continents, with the highest diversity in the tropics. Commonly referred to as bark and ambrosia beetles, they are among the most important wood-decomposing insects in any forests, making characteristic wood burrows deep into the wood or fine engravings just under the bark of dead trees. Only a few species are capable of attacking living trees (Hulcr & Dunn, 2011; Ranger *et al.*, 2015), whereas most species colonize woody substrates of dead plants. Some species of bark beetles are not found in logs and branches, but instead feed and reproduce in seeds, petioles, or seedlings of trees, occasionally in woody herbs and ferns, adding to the broad range of host plant relationships in this group of beetles (Jordal & Kirkendall, 1998; Kirkendall *et al.*, 2015).

A complete life cycle inside dead plant tissue makes life relatively protected in concealed niches, and has resulted in

Correspondence: Bjarte Jordal, University Museum of Bergen, University of Bergen, PO7800, NO-5020 Bergen, Norway. E-mail: bjarte.jordal@uib.no

the evolution of many different kinds of reproductive systems (Kirkendall, 1983). Guided by advanced pheromone attractants or sound production in the establishment of new nests, these insects have among the most advanced mating systems in animals. Subsocial life in tunnels and caves is characterized by interactions between parents and their offspring (Kirkendall *et al.*, 1997), where diverse microbial communities potentially play an important role (Six, 2012; Dohet *et al.*, 2016; Kawasaki *et al.*, 2016; Mariño *et al.*, 2017). A range of symbiotic relationships between microbes and bark beetles has evolved since Cretaceous times, particularly the cultivation of fungi seen in at least 10 independent lineages of Scolytinae (Farrell *et al.*, 2001; Massoumi Alamouti *et al.*, 2009; Jordal & Cognato, 2012; Kirkendall *et al.*, 2015; Li *et al.*, 2015). The adaptation to symbiosis by carrying fungal spores in special cavities in the beetle body (mycangia), and the obligate dependence on the fungal constituents for larval development, makes a strong case for coevolution. Repeated evolution of these fascinating traits therefore makes bark and ambrosia beetles ideal for testing evolutionary hypotheses in a phylogenetic context (Gohli & Jordal, 2017; Gohli *et al.*, 2017).

However, the phylogenetic framework for testing evolutionary hypotheses in Scolytinae has remained poorly developed, despite substantial attention from forest entomologists, and several attempts on resolving the molecular phylogeny of the group (Farrell *et al.*, 2001; Jordal & Cognato, 2012). The most recent and well sampled study was based on one mitochondrial and four nuclear genes for nearly 200 taxa, but nevertheless showed a widespread lack of resolution between tribes and deeper relationships (Jordal & Cognato, 2012). Although monophyly in eight of the tribes is currently supported by both morphological and molecular data, the majority of tribes and many genera seem paraphyletic, although this is not yet sufficiently supported to justify changes in the classification. Much more data are therefore needed to obtain better resolution and increased node support.

As a means of improving phylogenetic resolution in bark and ambrosia beetles, we add 13 new markers (Pistone *et al.*, 2016) to the previous five standard markers used for weevil systematics (McKenna *et al.*, 2009; Jordal *et al.*, 2011). Our main hypothesis implies that morphologically defined tribes as proposed by Wood (1986) are monophyletic, except for recently resolved groups including Micracidini LeConte, Ipini Bedel and Dryocoetini Lindemann (Jordal *et al.*, 2002; Jordal & Cognato, 2012; Cognato, 2013; Jordal & Kaidel, 2016). Our alternative hypothesis states that paraphyletic tribes are structured geographically between continents, such as Micracidini (Jordal & Kaidel, 2016), as opposed to the traditional classification (Wood, 1986), which indicates very little geographical structure within or between tribes.

Material and methods

We included 182 species, from 24 of the 26 currently recognized tribes (Table S1, Table 1, Fig. 1). As many genera as possible were represented, except for the previously well sampled Xyleborini. Some genera were occasionally represented by two or three species when paraphyly was suspected (Jordal & Cognato, 2012; Jordal & Kambestad, 2014). Four species in three different subfamilies of Curculionidae were used as outgroup (Lixinae, Cossoninae and Molytinae). A fragment of the mitochondrial gene cytochrome oxidase I (*CO1*), the D2–D3 segment of the nuclear large ribosomal subunit (*28S rRNA*), elongation factor 1 alpha (*EF-1α*), arginine kinase (*ArgK*) and carbamoyl-phosphate synthase 2-aspartate transcarbamylase – dihydroorotase (*CAD*) were amplified and sequenced using previously published protocols (Jordal *et al.*, 2011). Furthermore, we obtained sequences from 13 additional gene fragments (Table S2): poly-A binding protein (*PABP1*), triosephosphate isomerase (*TPI*), inhibitor of apoptosis 2 (*Iap2*), adenosine deaminase 2 (*ADA2*), ubiquitin-like modifier activating enzyme 5 (*UBA5*), RNA polymerase-associated protein CTR9 (*CTR9*), cyclin C (*CCNC*), chitin deacetylase 4 (*Cda4*), histone deacetylase Rpd3 (*HDAC Rpd3*), arrestin 2 (*Arr2*), Cu-Zn superoxide dismutase 1 (*SOD1*), pre-mRNA-splicing factor ATP-dependent RNA helicase PRP1 (*Prp1*) and Flap endonuclease 1 (*FEN1*). DNA extraction, PCR reaction and sequencing followed recently developed protocols (Pistone *et al.*, 2016). We designed additional primers for *FEN1* and *HDAC Rpd3* to increase PCR and sequencing success – FEN1 forward GCCACHGCHACYGARGAYATGG, and reverse TCACCATGCCYTCTTCGTCCGG, and

Table 1. The number of genera and species included in this study, compared to the total diversity in the subfamily.

Tribe	Genera	Species	Included genera	Included species
Amphiscolytini	1	1	–	–
Bothrosternini	6	131	4	4
Cactopinini	1	21	1	3
Carphodicticini	3	5	–	–
Corthylini	30	1211	11	11
Cryphalini	25	702	18	20
Crypturgini	5	55	5	6
Diamerini	7	132	4	6
Dryocoetini	18	474	12	14
Hexacolini	4	242	4	7
Hylastini	3	55	2	2
Hylesinini	14	164	11	14
Hylurgini	14	130	14	21
Hyorrhynchini	2	19	2	2
Hypoborini	9	74	6	11
Ipini	9	230	4	5
Micracidini	14	298	9	10
Phloeosinini	15	227	8	9
Phloeotribini	3	110	2	3
Phrixosomatini	1	25	1	2
Polygraphini	9	154	6	8
Scolytini	6	209	3	4
Scolytoplatypodini	2	53	2	5
Xyleborini	34	1168	2	2
Xyloctonini	5	78	4	6
Xyloterini	3	22	3	4

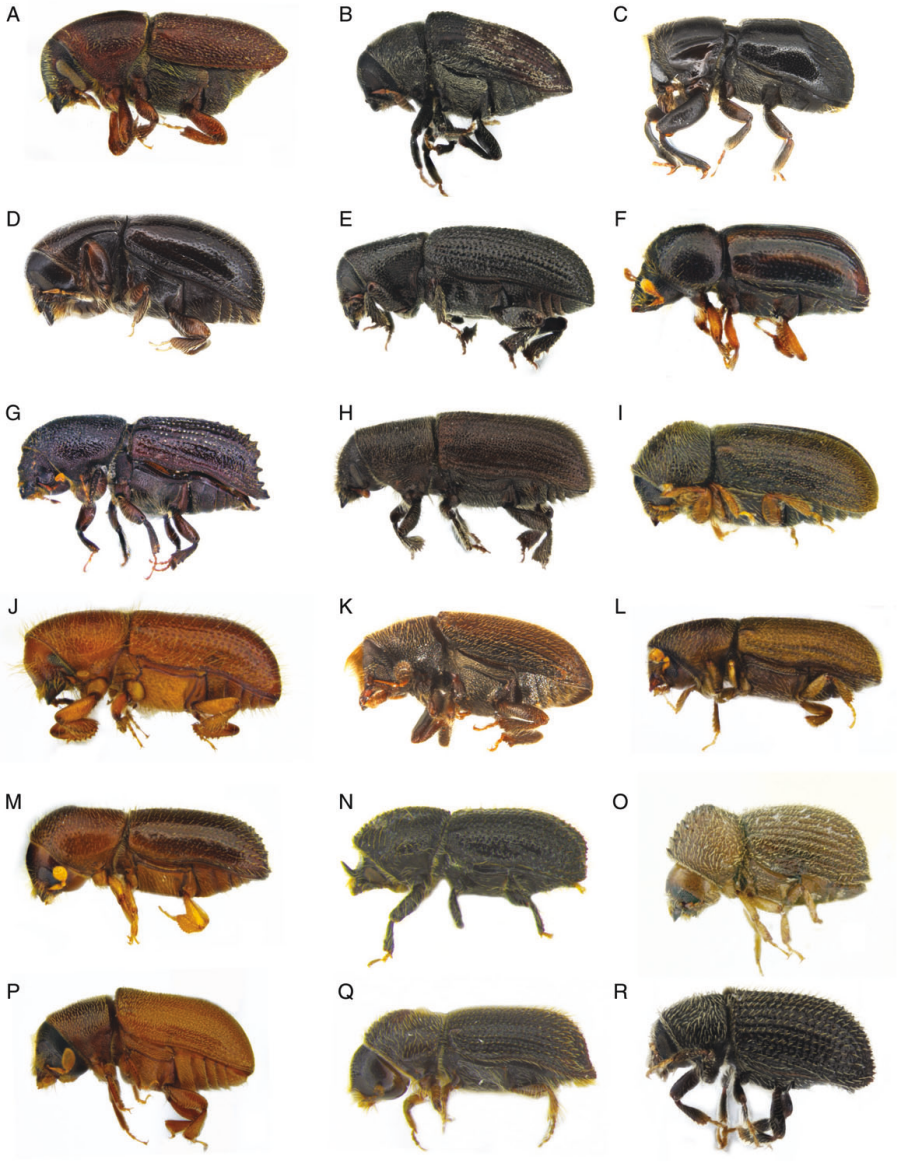


Fig. 1. Morphological diversity in Scolytinae represented by 18 genera in 18 tribes: (A) Scolytini – *Scolytus intricatus* (Ratzeburg); (B) Hylesinini – *Hylesinus eos* Spessivtsev; (C) Scolytoplatypodini – *Scolytoplatypus permirus* Schaufuss; (D) Hexacolini – *Scolytodes fraterniatratus* Jordal; (E) Hylastini – *Hylastes salebrosus* Eichhoff; (F) Xyloterini – *Trypodendron lineatum* (Olivier); (G) Ipini – *Ips sexdentatus* (Börner); (H) Hylurgini – *Hylurgus ligniperda* (Fabricius); (I) Cryphalini – *Cryphalus asperatus* (Gyllenhal); (J) Dryocoetini – *Dryocoetes autographus* (Ratzeburg); (K) Phrixosomatini – *Phrixosoma concavifrons* Jordal; (L) Crypturgini – *Crypturgus subcribrosus* Eggers; (M) Xyleborini (–a) – *Xyleborinus saxeseni* (Ratzeburg); (N) Cactopinini (–a) – *Cactopinus nasutus* Wood; (O) Xyloctonini – *Stephanopodius dispar* (Eggers); (P) Polygraphini – *Polygraphus proximus* Blandford; (Q) Micracidini – *Phloeocleptus cristatus* Wood; (R) Phloeotribini – *Phloeotribus lecontei* Schedl.

HDAC Rpd3_for1 GARTAYAAYAARCARATGC and HDAC Rpd3_for2 CRGARATYTGYATHAAYTGGGG – to use with previously published reverse primers (Pistone *et al.*, 2016). The optimal annealing temperature for these primers was 52°C for FEN1 and 62°C for *HDAC Rpd3*. Nucleotide sequences were blasted in GenBank for gene target verification, accepting a minimum E value threshold of 1E−4. All genes considered in this study were previously examined for paralogy and/or multiple copies in the OrthoDB database (Waterhouse *et al.*, 2013) and the majority of them (with the exclusion of *Arr2*, *Iap2* and *SOD1*) were frequently present in single copy in arthropod genomes (Pistone *et al.*, 2016). Careful phylogenetic analyses of each of the three genes with potential multiple copies did not indicate signs of anomalous relationships between sequences (Pistone *et al.*, 2016). Sequences with simple intron-exon patterns were aligned using BIOEDIT v. 7.2.5 (Hall, 1999) with manual adjustments, locating intron borders based on previously published alignments (Pistone *et al.*, 2016). MAFFT v.7 (Katoh *et al.*, 2002) was used to align protein coding genes with complicated structures, such as indel-rich regions, and the ribosomal gene 28S. GBLOCKS v. 0.91b (Castresana, 2000) was used to trim ambiguously aligned regions in 28S, applying the following settings: less strict flanking positions, gap positions allowed within blocks, allow smaller final blocks. The resulting 28S alignment contained 670 positions. Introns were removed from protein coding genes before phylogenetic analyses. The final alignment of all 18 gene fragments consisted of 10 156 nucleotides. We also analysed a dataset consisting of 3162 amino acids from the 17 protein coding genes, excluding 28S rRNA. All the new sequences of Scolytinae species used in this study were deposited in GenBank under the accession number MF771267–MF772316 (Table S2).

In order to assess the best partition scheme, the alignment was examined in PARTITIONFINDER v1.1.1 (Lanfear *et al.*, 2012), which defined 29 partitions and best-fitting models for each of them: 28S rRNA (SYM + I + G); COIpos1 (SYM + I + G); COIpos2 (GTR + I + G); COIpos3 (HKY + I + G); EF-1αpos1(TrN + I + G); ArgKpos2, EF-1αpos2 (SYM + I + G); EF-1αpos3 (TIM + I + G); ArgKpos1, CADpos1 (SYM + I + G); Arr2pos2, CADpos2, Endo1pos2, PyApos2, SODpos2 (GTR + I + G); CADpos3, Endo1pos3 (TrN + I + G); ArgKpos3 (SYM + I + G); Dea2pos1, PABP1pos1, UBA5pos1, CCNCpos1 (SYM + G); PABP1pos3 (TIM + I + G); Prp1pos1, TPIpos1 (GTR + G); Dea2pos2, TPIpos2, CCNCpos2 (GTR + G); Prp1pos3, Tpipos3 (TVM + I + G); Arr2pos1, SODpos1, UBA5pos2 (TVMef + I + G); IAP2pos3, UBA5pos3 (GTR + I + G); IAP2pos1 (TVMef + I + G); IAP2pos2 (GTR + I + G); SODpos3 (TrN + G); RNAelpos2 (TrNef + G); Dea2pos3, HDACpos3 (TIMef + I + G); CTR9pos1, HDACpos1 (SYM + G); CTR9pos2, Cda4pos2 (TrN + I + G); CTR9pos3, Cda4pos3, CCNCpos3 (GTR + I + G); Cda4pos1, FEN1pos1 (SYM + G); HDACpos2 (JC + G); Arr2pos3 (TIM + I + G). GTR + I + G was used for those models that could not be implemented in MRBAYES.

Two alternative partition schemes were also used: (i) based on genes (18 partitions), and (ii) based on 28s rRNA, plus codon positions for each genome (seven partitions: 28 s rRNA, first, second and third position for mitochondrial COI and nuclear genes). Furthermore, we assessed topological stability by excluding putatively problematic genes, and by excluding third codon positions for COI. The best nucleotide substitution model for these partitions was selected using the Akaike information criterion (AIC) in MRMODELTEST 2.3 (Nylander, 2004). The GTR + I + G model was selected for all partitions.

We ran the 29 partition analysis in MRBAYES v3.2 (Ronquist & Huelsenbeck, 2003) using two parallel runs of four chains running for 100 million generations, sampling every 10 000 generation. The first 50% of the tree sample were discarded as burn-in. For all other analyses (partition by gene and codon position) we used 50 million generations as sufficient based on parameter analyses in TRACER v1.6 (Rambaut & Drummond, 2007).

Parsimony analyses were conducted in PAUP* 4.0 (Swofford, 2011) and consisted of 1000 heuristic searches with 20 random additions and tree bisection and reconnection swapping for each search. Node support was estimated by 100 bootstrap replicates of 20 random addition replicates each.

Analyses of the amino acid dataset were made in MRBAYES v3.2 and PAUP* 4.0 using parsimony settings specified for the nucleotide dataset, and a mixed model for the Bayesian analysis. The amino acid dataset was also run in PHYLOBAYES v4.1 (Lartillot *et al.*, 2009) using a CAT + GTR model. This model is usually the model with the highest fit for the data among all models implemented in PHYLOBAYES (except for small datasets) and is more robust against long-branch attraction artifacts compared with all other models. Two independent chains were run until the maxdiff parameter was less than 0.3 and the minimum effective size was higher than 50. Convergence of the chains was checked using the *bpcomp* command and a consensus tree was built, discarding the first 1000 generations as burn-in. We analysed the *Iap2* and *Arr2* nucleotide sequences (100 and 89 species, respectively) separately to assess the influence of including hypervariable regions, and to examine the phylogenetic utility of clade defining indels (CDIs). Analyses were run in MRBAYES for 50 million generations, sampling every 10 000 generations, discarding the first 50% of the tree sample as burn-in after assessment of chain convergence in TRACER v1.6.

We estimated divergence times using the software BEAST v1.8.2 (Drummond *et al.*, 2012), with input files generated in BEAUTI. Three different partition schemes were used: per gene (18 partitions), codon positions per genome plus 28 s (7 partitions), or 29 partitions as defined by PARTITIONFINDER. The tree was calibrated with four fossils, using a normal distribution for fossil age. The following calibration points were used: 116 ± 20 Ma (Lebanese amber), stem Scolytinae (Kirejtshuk *et al.*, 2009); 100 ± 20 Ma (Burmese amber), stem *Microborus* Blandford (see Cognato & Grimaldi, 2009); 40 ± 10 Ma (Baltic amber), stem Hylastini LeConte (*Hylurgops* LeConte + *Hylastes* Erichson); and 20 ± 5 Ma (Dominican amber), stem *Lymantor* Perris + *Xylocleptes* Olivier (see Jordal *et al.*, 2011). The analysis was run for 200 million generations, using recommended priors, a GTR + I + G model, an uncorrelated lognormal relaxed clock with estimated rates, and a Yule

speciation prior, with a total of 50 000 trees sampled, deleting the first 25 000 trees as burn-in.

Biogeographical inference was obtained by applying statistical Dispersal-Extinction-Cladogenesis (S-DEC, Lagrange) and Bayesian binary Markov chain Monte Carlo (BBM) analyses as implemented in RASP – 'Reconstruct Ancestral State in Phylogenies' (Yu *et al.*, 2015). Because very little is known about biogeographical relations in Scolytinae, we used two relatively simple analyses to make a first basic interpretation of the main patterns for the group as a whole. For the BBM analysis we used 50 000 cycles of 10 Markov chain Monte Carlo chains sampling every 100 generations, implementing an estimated F81 model and allowing for a maximum of two ancestral areas. Broadly defined biogeographical hypotheses were tested, based on six areas defined by major geographical affinities: A, Afrotropical; B, Palearctic; C, Nearctic; D, Neotropical; E, Indomalayan and F, Australasian.

Results

Bayesian analysis of the nucleotide data divided into 29 partitions resulted in a tree topology which was well resolved within genera and tribes, but considerably less so between tribes and older lineages (Fig. 2). The monophyletic tribe Scolytini Latreille [posterior probability (PP) = 1] was the first diverging lineage in the subfamily Scolytinae and represented, together with the molytine genus *Larinus* Schaller, the sister clade to all remaining Scolytinae (PP = 1). One of the most well resolved tribal-level relationships was Dryocoetini and its sister lineage Ipini (PP = 1), with the inbreeding Xyleborini LeConte and Premnobiina Browne as nested subclades in each of these tribes. These well-supported groups were the closest relatives to the Hypoborini LeConte (PP = 1), which all together made the sister group to Micracidini sensu lato (Jordal & Kaidel, 2016) and the genera *Phrixosoma* Blandford, *Sphaerotrypes* Blandford and *Glostatus* Schedl (including *Stephanopodius* Schedl).

Monophyly was recovered completely, or nearly so, for 15 tribes: Scolytini, Crypturgini LeConte, Phloeotribini Chapuis, Bothrosternini Blandford, Hyorrhynchini Hopkins, Scolytoplatypodini Blandford, Hylastini, Phrixosomatini Wood, Micracidini (sensu lato, including *Cactopinus* Schwarz – Cactopinini Chamberlin), Hypoborini Nüsslin (excluding *Chaetophloeus* LeConte and *Zygophloeus* Schedl), Dryocoetini (sensu lato, including Xyleborini), Ipini (sensu lato, including Premnobiina), and Xyloterini LeConte. The majority of these clades were recovered with maximum or high node support (PP ≥ 0.95). The core Hypoborini was furthermore joined by two undescribed species which grouped together with maximum node support and which represent new genera in this tribe.

Nine tribes (Phleosinini Nüsslin, Xyloctonini Eichhoff, Corthylini LeConte, Diamerini Hagedorn, Cryphalini Lindemann, Hexacolini Eichhoff, Polygraphini Chapuis, Hylesinini Erichson and Hylurgini Gistel) were not recovered as monophyletic. Genera in Hylurgini were distributed mainly in one northern hemisphere and two southern hemisphere clades and each was highly supported. The first group included 13 species in Hylurgini, Hylastini and Hylesinini, all with boreal distribution (PP = 1), including a subclade of *Dendroctonus* Erichson and five other conifer associated genera (PP = 1), with Hylastini (*Hylurgops* and *Hylastes*) as part of the *Dendroctonus* clade. A large group of *Araucaria*-associated species related to *Hylurgonotus* Eggers was recovered with maximum support, separate from a second *Araucaria*-associated linage consisting of *Hylurdrectonus* Schedl species and *Xylechinus araucariae* Schedl.

Species in two paraphyletic genera – *Xylechinus* Chapuis and *Chaetoptelius* Fuchs – were associated with species of other genera from the same geographical area. The New World *Xylechinus maculatus* Schedl grouped with species of the New World genera *Chramesus* LeConte, *Pseudochramesus* Schedl and *Phloeotribus* Bernard, including the recently erected *Dryotomicus* Wood. A clade of southern hemisphere broadleaf-associated beetles (*Ficicis* Blandford, *Chaetoptelius tricolor* Schedl and *Zygophloeus australis* Schedl) was recovered (PP = 0.84), and the single Palearctic species of *Chaetoptelius* grouped with northern boreal species of Hylesinini. Several other genera in the southern hemisphere, such as *Hylurgonotus* and *Pachycotes* Chapuis, were also paraphyletic.

Xyloctonini was paraphyletic with respect to *Glostatus* Schedl (and *Stephanopodius* Schedl). The two *Phrixosoma* Blandford species represent a monotypic tribe that grouped with *Sphaerotrypes hagedorni* Eggers (Diamerini). Three other genera in Diamerini were separated into three independent lineages. *Strombophorus* Eggers was strongly supported as a part of a Hylesinini subclade, as sister group to *Rhopalopselion* Hagedorn (PP = 0.99), both clustering with *Hapalogenius* Hagedorn. Two genera in Hexacolini (*Scolytodes* Ferrari and *Gymnochilus* Eichhoff) were recovered as the sister lineage to Scolytoplatypodini (*Remansus* Jordal and *Scolytoplatypus* Schaufuss), but with low node support.

The tribe Corthylini was paraphyletic with respect to the genera *Microborus* and the cryphaline genus *Cryptocarenus* Eggers, with *Dendroterus defectus* Wood grouping with *Microborus* in a basal position (PP = 0.99). *Cryptocarenus diadematus* Eggers was closely related to the genera *Araptus* Eichhoff and *Dacnophthorus* Wood (PP = 1). Other genera in Cryphalini were distributed on three main clades, the largest clade included species of *Ernoporus* Thomson, *Scolytogenes* Letzner and allies; a second clade included only *Cryphalus* Hopkins and *Hypocryphalus* Stebbing; and a third clade included *Hypothenemus* Westwood and part of the otherwise strongly polyphyletic *Ptilopodius* Hopkins and *Cosmoderes* Eichhoff.

Comparison among different analyses and partition schemes

Bayesian analyses using different partition schemes did not result in large differences in tree topology, except a few clades that were generally poorly supported in each analysis (Table 2). Analysis of 18 partitions (genes) placed the tribe Scolytini and the genus *Microborus* at the root of Scolytinae (Figure S1). Furthermore, the *Cryphalus-Hypocryphalus* clade grouped with Xyloterini, albeit weakly supported (PP = 0.8). Node support was overall lower in the gene partitioned analysis compared with

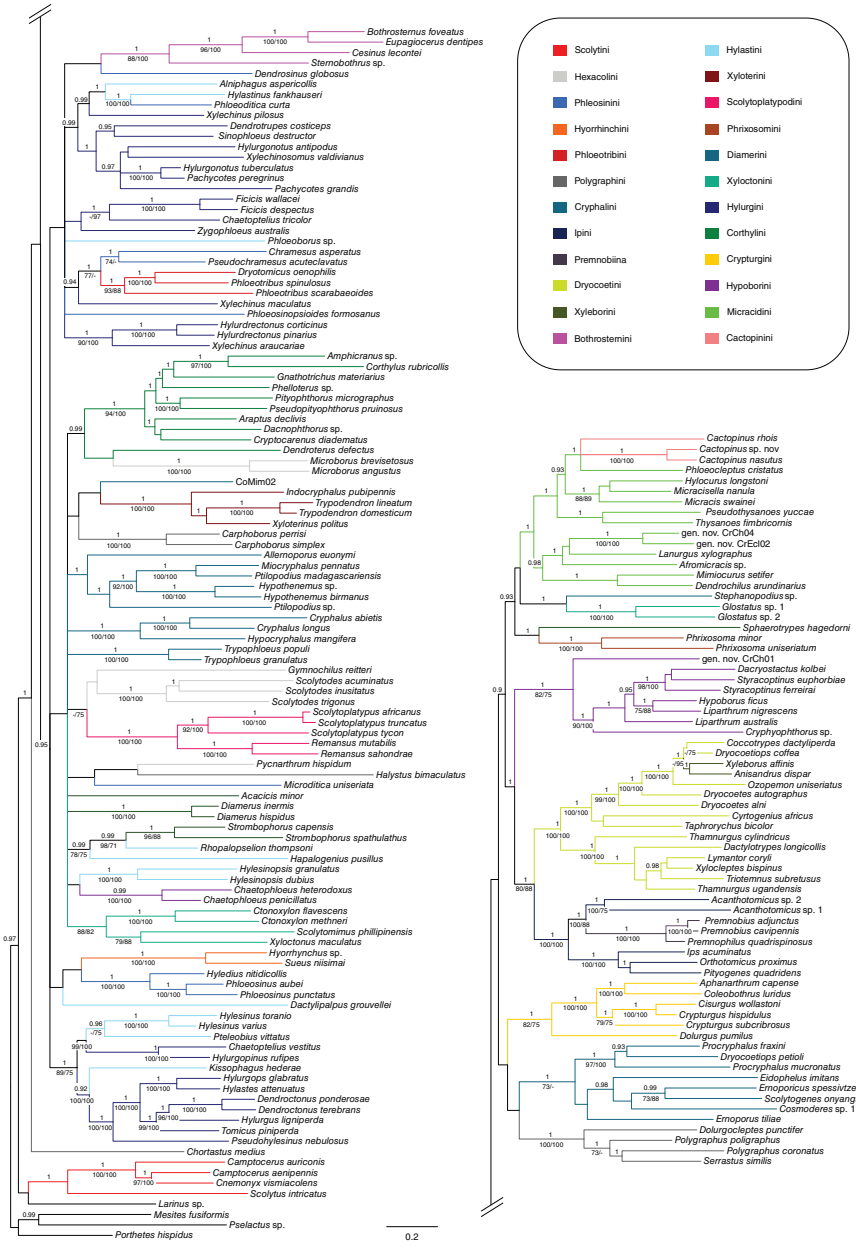


Fig. 2. Tree topology resulting from the Bayesian analysis of 29 partitions as defined by PARTITIONFINDER (standard deviation of split frequencies = 0.07, potential scale reduction factor = 1). Posterior probabilities > 0.9 are shown above branches, parsimony bootstrap (BS) values > 70 below (the first BS value refers to the analysis of all nucleotides, and the second BS value to the analysis with third codon position excluded).

Table 2. Node support for selected taxon groups based on Bayesian and parsimony analyses of the different data sets and partitioning schemes.

	Bayesian posterior probability				Parsimony bootstrap values		
Clades	29 partitions	18 partitions	7 partitions	aa	all data	3rds excl.	aa
Scolytinae, ex Scolytini	1	0.87	0.89	0.55	–	–	–
Hylurgini – *Dendroctonus* clade	1	1	1	0.99	89	75	–
Phloeotribus clade	1	1	1	0.54	–	–	–
Ipini + Dryocoetini/Xyleborini	1	1	1	0.99	80	88	–
Hypoborini, ex *Chaetophloeus*	1	1	1	1	82	75	83
Ipini + Dryocoetini/Xyleborini + Hypoborini	1	1	1	–	–	–	–
Northern hemisphere Phleosinini + Hyorrhynchini	0.8	0.67	–	–	–	–	–
Hypothenemus clade	1	1	1	0.99	92	–	75
Strombophorus clade	1	1	1	0.89	–	–	–
Hexacolini + Scolytoplatypodini	0.88	0.5	0.68	–	–	75	–
Corthylini (including *Cryptocarenus*)	-	1	1	0.99	–	–	–
Micracidini	1	0.99	1	0.79	64	63	–
Bothrosternini	1	0.87	1	1	88	100	58

3rds, third positions; aa, amino acids.

the analysis of 29 partitions (Table 2). The analysis based on seven partitions (28 s rDNA, codon positions for mitochondrial and nuclear genes) recovered much of the same relationships, including *Microborus* as the sister lineage to the Scolytini tribe (PP = 1). In both cases *Cryptocarenus* was deeply nested in Corthylini, with *Dendroterus* forming the first diverging lineage in that tribe.

The parsimony analyses resulted in lower resolution, and fewer tribes were monophyletic. Exclusion of nucleotides in the third position of protein coding genes did not provide better resolution, and node support was generally similar (Table 2). The only well-established relationship at tribe level that was recovered in this analysis was the Dryocoetini–Ipini affiliation (BS = 80; with third positions excluded, BS = 88); in the remaining parts of the tree, most tribes formed a polytomy. Nevertheless, tribes such as Bothrosternini, Xyloterini, Crypturgini, and Hypoborini (excluding *Chaetophloeus*) were recovered with maximum or medium bootstrap values (> 70%; see Table 3). In both analyses, Scolytini and Micracidini were not monophyletic and *Dendroterus defectus* grouped with the genus *Microborus*, although weakly supported (BS < 70%).

The tree based on the analysis of the amino acid dataset obtained in PHYLOBAYES confirmed most nodes found in the nucleotide analyses (Fig. 3). Some noteworthy changes from

Table 3. Average estimates and their 95% highest posterior density intervals for the crown and stem age based on three different partition schemes. Estimates from Jordal and Cognato (2012) are included for comparison.

	PARTITIONFINDER (29 partitions)		Genes (18 partitions)		Codon position (7 partitions)		Jordal & Cognato (2012) (7 partitions)	
Clade	Stem age	Crown age	Stem age	Crown age	Stem age	Crown age	Stem age	Crown age
A	96 + 11	78 ± 9 (37 ± 6)	108 ± 4	89 ± 10 (39 ± 10)	102 ± 3	88 ± 10 (38 ± 9)	110 ± 8 (38 ± 12)	81 ± 18 (35 ± 12)
B	97 ± 5	85 ± 6	96 ± 4	92 ± 4	94 ± 3	82 ± 4	83 ± 10	79 ± 10
C	104 ± 9	78 ± 11	97 ± 8	74 ± 10	101 ± 4	68 ± 11	79 ± 9	62 ± 13
D	94 ± 4	71 ± 7	96 ± 4	73 ± 4	88 ± 3	72 ± 5	82 ± 10	71 ± 10
E	71 ± 6	55 ± 9	73 ± 6	61 ± 8	75 ± 5	–	68 ± 10	56 ± 12
F	65 ± 5	38 ± 3	64 ± 5	37 ± 6	62 ± 5	37 ± 4	58 ± 8	46 ± 10
G	57 ± 5	–	68 ± 4	–	41 ± 12	–	–	–
H	59 ± 4	20 ± 9	56 ± 5	19 ± 8	57 ± 6	19 ± 6	54 ± 14	23 ± 8
I	16 ± 1	15 ± 1	18 ± 1	15 ± 1	16 ± 2	14 ± 1	23 ± 5	21 ± 4
J	57 ± 6	12 ± 2	57 ± 7	12 ± 2	55 ± 3	10 ± 2	37 ± 9	8 ± 4
K	81 ± 6	50 ± 5	70 ± 16	48 ± 9	79 ± 8	45 ± 4	70 ± 9	42 ± 9
M	101 ± 4	50 ± 12	89 ± 10	44 ± 17	90 ± 6	33 ± 19	72 ± 14	–
N	81 ± 3	52 ± 5	84 ± 5	52 ± 4	80 ± 4	50 ± 4	56 ± 10	37 ± 9
O	66 ± 6	46 ± 5	74 ± 7	53 ± 4	70 ± 7	52 ± 5	53 ± 8	50 ± 8
P	54 ± 4	36 ± 5	52 ± 4	31 ± 7	48 ± 4	32 ± 4	37 ± 8	24 ± 10

Clades: A, Scolytini (*Camptocerus*); B, Micracidini; C, Crypturgini; D, *Dendroctonus* clade (boreal distribution); E, *Pachycotes*, *Hylurgonotus*, *Sinophloeus*; F, conifer feeding Ipini (*Ips*, *Pityogenes*, *Orthotomicus*); G, *Cryptocarenus*; H, *Hypothenemus*; I, Xyleborini; J, Premnobiina (*Premnobius*, *Premnophilus*); K, Xyloterini; M, Hyorrynchini, N, Scolytoplatypodini; O, Corthylini subclade – *Corthylus*, *Amphicranus*; P, *Bothrosternus*, *Eupagiocerus*. D–F, conifer feeding; G–J, inbreeding; K–P, fungus farming.

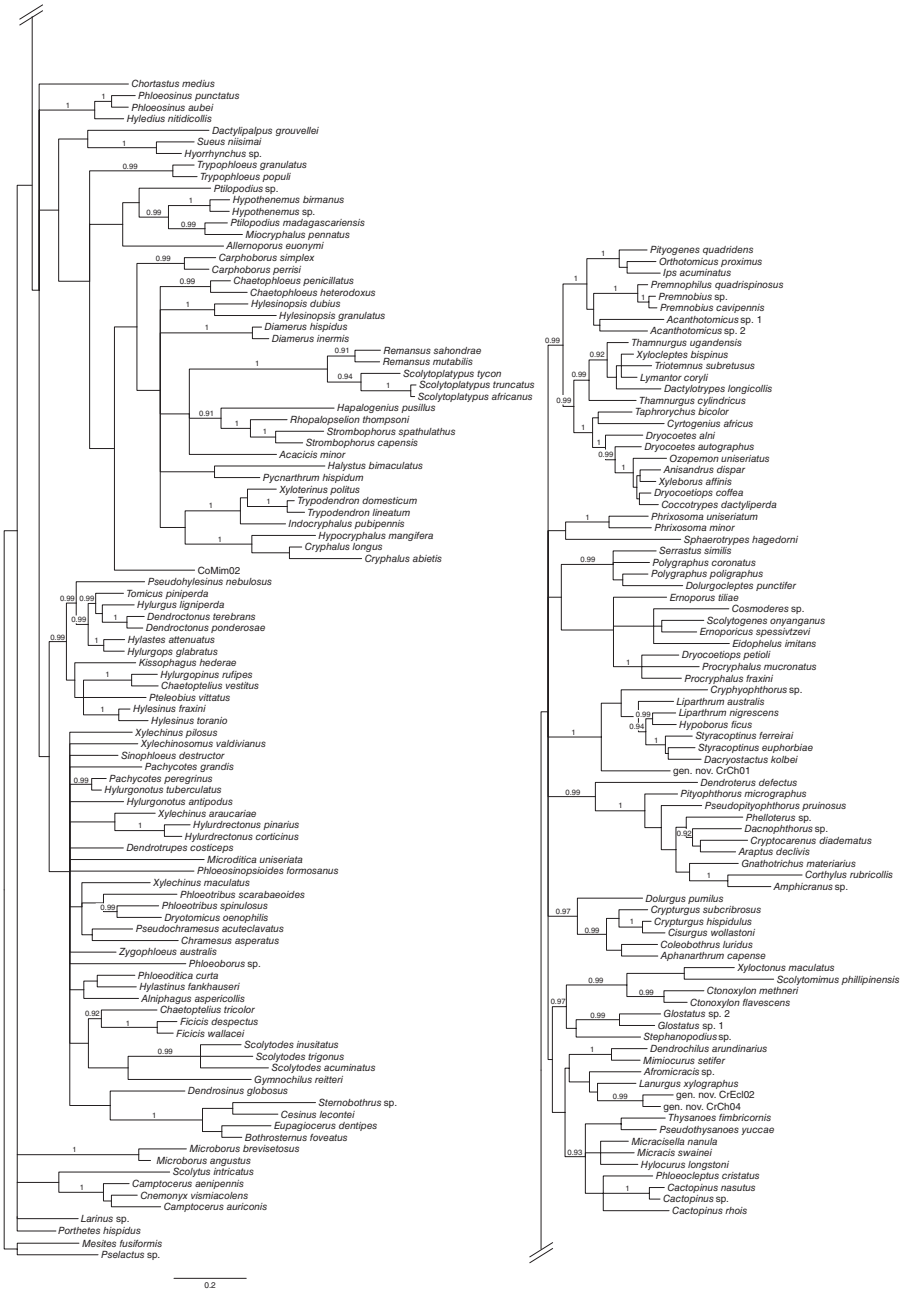


Fig. 3. Tree topology resulting from the PHYLOBAYES analysis based on the amino acid dataset using a CAT + GTR model. Posterior probability values are shown on branches.

the 29-partition analysis included *Microborus* at the base of tree close to the tribe Scolytini, the position of *Glostatus* and *Stephanopodius* together with the remaining genera of Xyloctonini (PP = 0.97), *Gnathotrichus materiarius* Fitch together with the two other fungus-farming genera in Corthylini, and *Kissophagus* Schmitt together with the *Hylesinus* Fabricius clade rather than the *Hylurgus* Fabricius clade. A large portion of the Hylurgini and Hylastini were, on the other hand, poorly resolved and generally formed a polytomy.

When analysing the amino acid dataset by maximum parsimony or Bayesian analysis in MRBAYES, a lower number of tribes was recovered as monophyletic (Figure S2; Table 2). The tribes Scolytoplatypodini, Bothrosternini, Hypoborini (excluding *Chaetophloeus*), Crypturgini, Hyorrhynchini, Ipini and Xyloterini were monophyletic with maximum or high node support and/or Bayesian PP (BS > 70 and PP ≥ 0.95). Other groups were also monophyletic but weakly supported in either the parsimony analysis (e.g. Corthylini with *Cryptocarenus*, BS = 65; Xyloctonini, BS = 56) or the Bayesian analysis (e.g. Dryocoetini, PP = 0.67, and Micracidini, PP = 0.87).

Excluding gene markers

The Bayesian tree based on a reduced gene sample of 13 new markers resulted in lower resolution and generally lower node support for the majority of clades (Figure S3). At the tribe level, only Hyorrhynchini, Xyloterini and Hypoborini (excluding *Chaetophloeus*) received high node support (PP = 0.97–0.99). Other tribes, e.g. Scolytini, Scolytoplatypodini, Dryocoetini and Ipini, received much lower node support in these analyses. By comparison, the analysis of the five established gene markers resulted in high support (> 0.95) in 10 tribes (Figure S4). We also explored the effect of removing the five least informative genes according to Pistone *et al.* (2016; low amplification and sequencing rate, short length of the gene fragment, low phylogenetic signal: *HDAC Rpd3*, *DEA2*, *SOD1*, *CDA4*, *CCNC*), together with the third codon position of COI. The tree topology experienced minimal changes compared with the analyses of the full dataset, with maximally supported clades obtained for 16 tribes (tree not shown), and *Microborus* as the sister lineage to Scolytini at the root of the tree (PP < 0.95).

Clade defining indels

The genes *Iap2* and *Arr2* contained length variable regions (Pistone *et al.*, 2016). These were aligned and examined for CDIs. Two variable regions were encountered in the first exon of *Iap2* (Fig. 4). The first variable region had a maximum length difference of 72 nucleotides (24 amino acids). The second region varied by 75 nucleotides (25 amino acids). Both variable regions in *Iap2* contained serine-rich repeats, with three to 11 consecutive serines in the first and most serine-rich region.

Arr2 contained only a single variable region located at the beginning of the second exon, just after the intron border. This region contained a maximum of 21 nucleotides, but insertion of only three or six nucleotides was the most prevalent pattern. Due to limited length variation, few consistent CDIs were found in this gene.

Several CDIs were identified in the *Iap2* gene, with the second variable region being more informative (Figs 4, 5). Clade defining indels were unambiguously identified in Corthylini (including *Cryptocarenus*), Ipini, subtribe Premnobiina, Xyleborini (with the haplodiploid species in Dryocoetini) and Hyorrhynchini. One subclade in Dryocoetini (*Lymantor coryli* Perris, *Taphrorychus bicolor* Herbst, *Triotemnus subretusus* Wollaston, *Xylocleptes bispinus* Duftschmid) was further characterized by a shared CDI sequence. A conserved motif was shared by a Cryphalini clade consisting of *Hypothenemus* and its sister clade of *Cosmoderes*-like genera. Furthermore, a specific insertion of the four amino acids LGAR was shared between two genera of Hypoborini only, congruent with the paraphyly of *Liparthrum* Wollaston in the combined analyses (Fig. 1).

Phylogenetic analysis of *Iap2* recovered several clades with high node support, congruent with the concatenated analyses (Fig. 5). By contrast, the phylogenetic tree based on the gene fragment of *Arr2* was largely unresolved, with fewer resolved tribes (Xyloterini, Crypturgini and Hypoborini), but some clades of closely related genera were recovered with high node support (Figure S5). Individual phylogenies based on *Arr2* and *Iap2* received higher node support for expected clades when the hypervariable regions were included rather than excluded (trees not shown), and were therefore included in all combined analyses.

Timing of evolutionary novelties

BEAST analyses under the three different partition schemes were largely congruent and similar to the Bayesian analysis of 29 partitions (Fig. 6). The three different analyses produced similar time estimates, with mean stem and crown ages in one analysis generally ranging well within the 95% highest posterior density intervals (HPD) of the other two analyses. We used the analyses based on seven partitions to present our findings for a direct comparison with Jordal & Cognato (2012) where the same partition scheme was used (Table 3). The oldest split in Scolytinae, equivalent to the stem age of the tribe Scolytini and *Microborus*, was 112 Ma, with an HPD of 108–116 Ma. Lineages classified as tribes differed four-fold in age, ranging from > 80 Ma in Scolytini, Corthylini and Micracidini to < 20 Ma in Xyleborini (Table 3).

The many origins of conifer feeding were derived within angiosperm-associated clades and occurred at the earliest possible age around 101 ± 4 Ma (stem age) in Crypturgini, 88 ± 3 Ma in boreal Hylurgini (*Dendroctonus* clade), and a stem age of 62 ± 5 Ma in boreal Ipini. The oldest crown age for any of these clades was 68 ± 11 Ma for Crypturgini (Table 3).

The first origin of permanent inbreeding by sibling mating was *Hypothenemus* with a stem age of 57 ± 6 Ma, with a minimum age indicated by its crown age of 19 ± 6 Ma. All other inbreeders evolved much later, including the haplodiploid clade

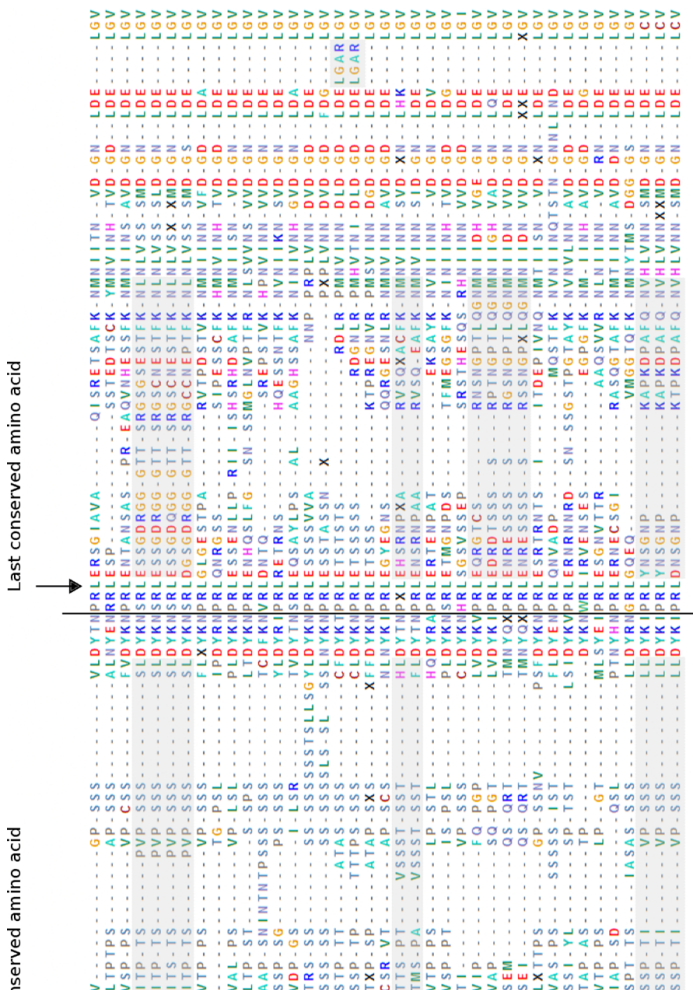


Fig. 4. The graphic illustrates the complexity of the two hypervariable regions of *IAP2* and the grey box highlights motives conserved among tribes and/or genera. A total of 73 conserved amino acids in between the two hypervariable regions were removed and only the variability between a subset of sequences is shown here.

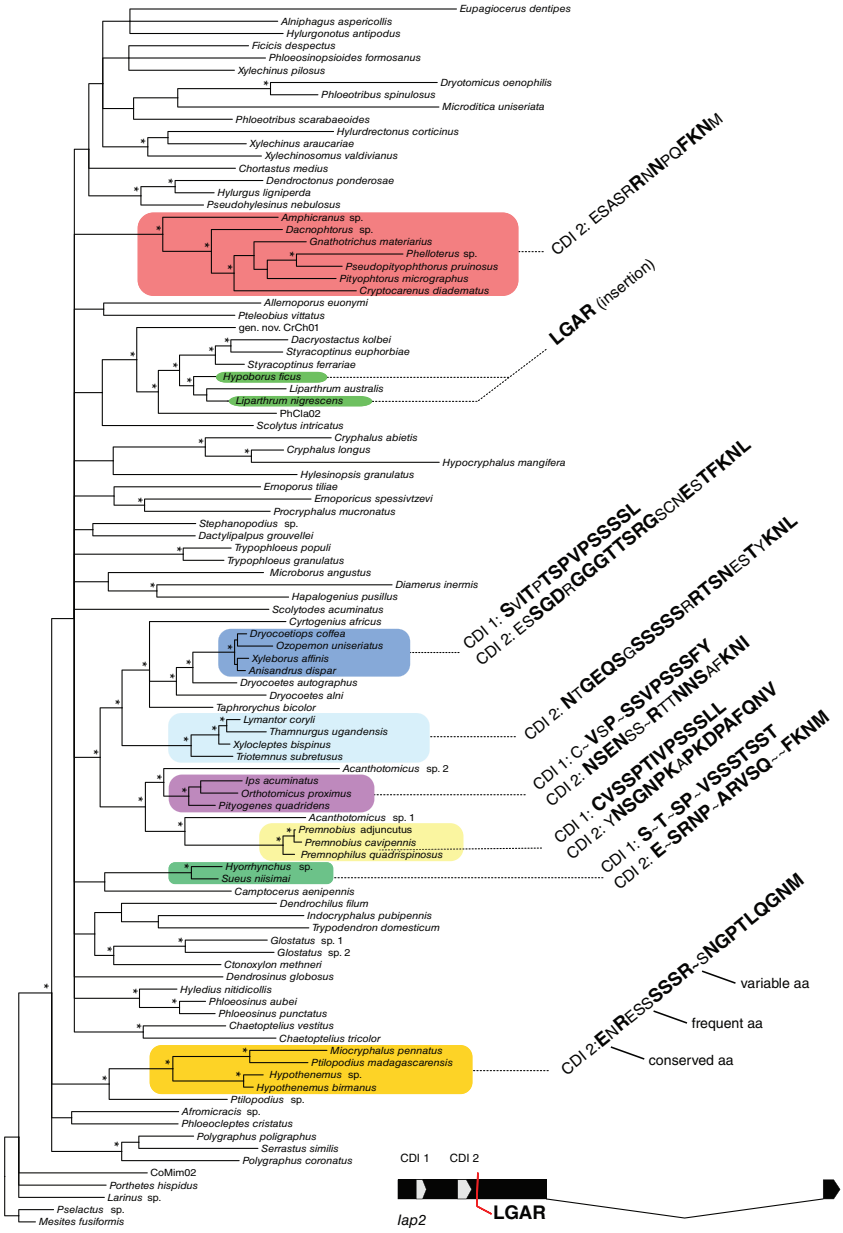


Fig. 5. Phylogenetic tree resulting from Bayesian analysis of the IAP2 gene. Clade defining indels (CDIs) of *Iap2* are reported for some tribes. Node support higher than 0.95 is marked by '*'.

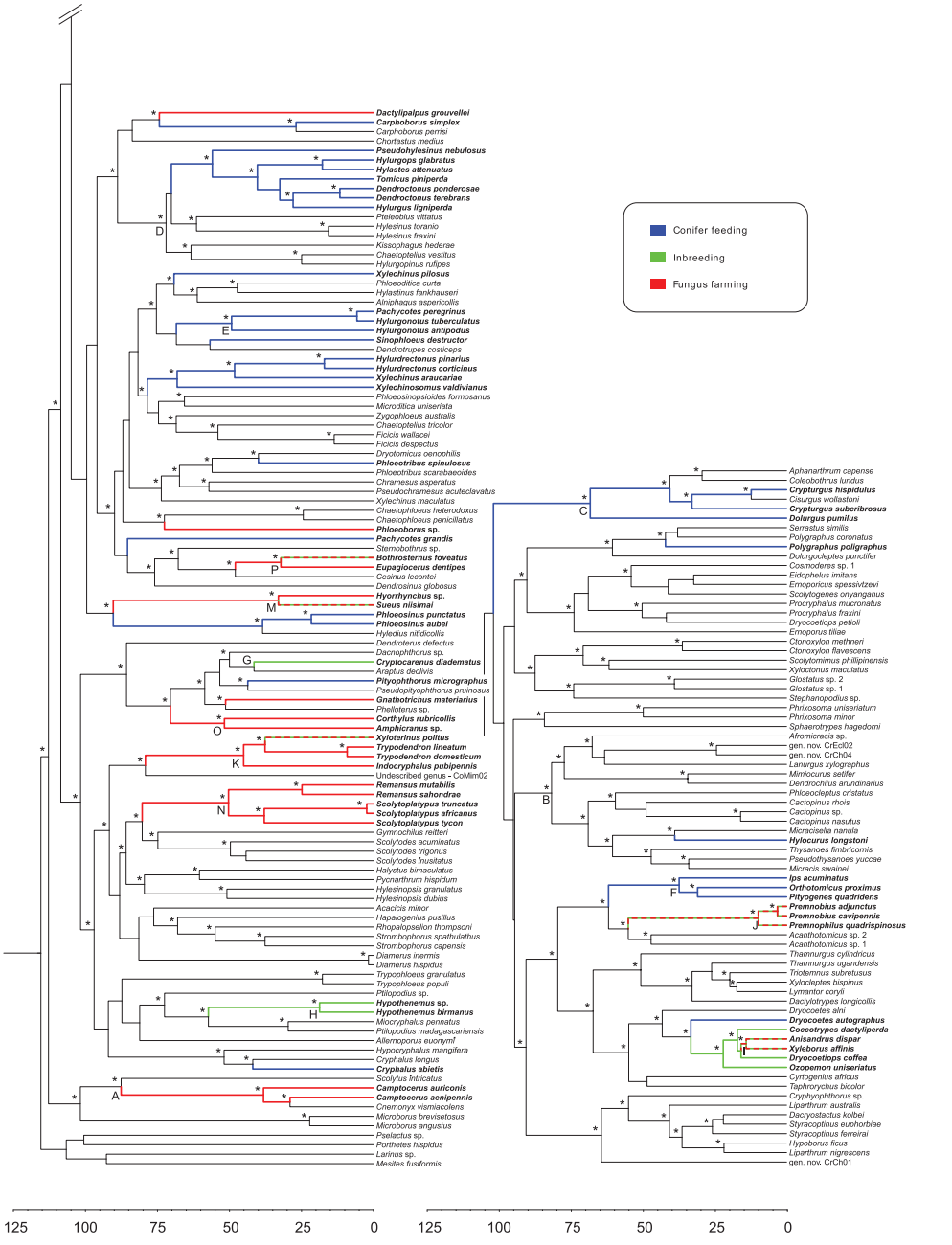


Fig. 6. BEAST dated phylogeny of Scolytinae with posterior probabilities > 0.95 indicated by '*' above the branch. Conifer-associated lineages are marked in blue, fungus farming in red, and permanent inbreeding in green. Capital letters (A–N) refer to the origin of traits in some important clades (see Table 4).

in Dryocoetini (*Ozopemon* Strohmeyer, *Coccotrypes* Eichhoff, *Dryocoetiops* Schedl and Xyleborini) at 22 ± 3 Ma.

The age of eight fungus-farming lineages could be estimated and five of them defined well-supported clades. Scolytoplatypodini, Xyloterini and Corthylina revealed crown ages of 50–45 Ma, representing minimum ages for the oldest origins of this peculiar feeding mode in Scolytinae. The tribe Corthylini showed two different origins for xylemycetophagy, with *Gnothotrichus* occurring slightly later than other fungus-farming Corthylina (*Corthylus* Erichson and *Amphicranus* Blandford). The youngest fungus-farming clade was Xyleborini, radiating no earlier than 16 ± 2 Ma, while all other potential young clades such as the Ipini subtribe Premnobiina (crown age 10 Ma) and the *Bothrosternus* Eichhoff–*Eupagiocerus* Blandford clade (crown age 26 Ma) had unreliably long stem ages.

Biogeographic analyses

Analyses by two different methods of ancestral area reconstruction (S-DEC Lagrange and BBM) were largely congruent; hence we report the results from the BBM analyses only (Fig. 7). Estimates for many of the deepest nodes were highly ambiguous, while many of the more recent nodes corresponding to tribes or younger categories were resolved with high probability (> 85%). The Afrotropical region was the most common ancestral region with two major Cretaceous origins: (i) Scolytoplatypodini, Hexacolini, Xyloterini and parts of Cryphalini; and (ii) a large clade consisting mainly of Xyloctonini, parts of Cryphalini and Polygraphini in one subclade, and Micracidini, Hypoborini, Dryocoetini and Ipini in another. Repeated colonization of the Palearctic occurred in each of these primarily Afrotropical clades, with characteristic elements evolving in the boreal region such as the conifer-associated Ipini, Cryphalini (part) and Polygraphini (part).

Two colonizations of the New World from the Afrotropical region were particularly well supported, including Hexacolini (minimum age of origin 74 Ma) and the New World clade of Micracidini (68 Ma). A younger Neotropical origin occurred in *Phrixosoma* (50 Ma). Several other Neotropical clades, such as Corthylini, Bothrosternini and Scolytini, were much older, and their ancestral history could not be reconstructed with certainty.

Species in the southern and northern hemispheres formed separate clades, with very few faunal connections between the two parts of the world. However, reconstructions of the geographical origin within each of these restricted clades were uncertain, except for the southern, mainly conifer-feeding, Hylurgini lineage that includes *Pachycotes*, *Hylurgonotus*, *Sinophloeus* Brethes and *Dendrotrupes*, which revealed a definite Neotropical origin. A second large southern, mainly Australasian, group contained a paraphyletic assemblage of *Hylurdrectonus* (Hylurgini), *Ficicis* and *Chaetoptelius* (Hylesinini), *Zygophloeus* (Hypoborini), *Microditica* Jordal and *Phloeosinopsioides* Schedl (Phloeosinini), which shared a mainly Australian distribution indicative of their ancestral origin.

Discussion

This study constitutes the most comprehensive phylogenetic assessment of Scolytinae to date, in terms of both taxon and gene sampling. Our data included 18 molecular markers sequenced for 182 species, which represent 24 of the currently recognized 26 Scolytinae tribes. With the addition of 20 new genera (*Dendrotrupes*, *Zygophloeus*, *Dryotomicus* Wood, *Cryptocarenus*, *Phelloterus* Wood, *Dacnophthorus*, new genus near *Cosmoderes*, *Remansus*, *Halystus* Schedl, *Hyorrhynchus* Blandford, *Kissophagus*, *Phloeocleptus* Wood, *Micracisella* Wood, *Dendrochilus* Schedl, *Mimiocurus* Schedl, *Stephanopodius* Schedl, *Dacryostactus* Schaufuss, *Cryphyophthorus* Schedl, *Premnophilus* Browne, *Eidophelus* Eichhoff) across 10 different tribes, we have tested the phylogenetic position of nearly 80% of the Scolytinae diversity at the tribal and genus levels (Hulcr *et al.*, 2015). However, despite the addition of 13 new molecular markers, the resolution was still ambiguous for some of the deepest nodes. The new data confirmed several previous results that suggest a basal or possibly separate position of the tribe Scolytini, clearly isolated from the remaining genera of Scolytinae (Jordal & Cognato, 2012; Gillett *et al.*, 2014; Mugu *et al.*, 2018). This is perhaps not surprising given that several deviant morphological features characterize this tribe (Smith & Cognato, 2014). A broader sample of weevils is nevertheless needed to test whether they form the sister group to other bark and ambrosia beetles, or if they form a separate and potentially unrelated lineage (Mugu *et al.*, 2018). A separate standing of Scolytini would have nomenclatural consequences for the usage of the name 'Scolytinae' (Jordal *et al.*, 2014).

Our new data supported a great number of relationships already suggested in previous molecular studies that were based on more limited sampling (Farrell *et al.*, 2001; Jordal *et al.*, 2002, 2008; Jordal & Cognato, 2012). This means that these relationships are predictable with increased data and therefore not supposed to change with additional genome-wide sequencing. Most notable in this respect was the solid relationship between Ipini and Dryocoetini, where the fungus-farming and permanently inbreeding species of Premnobiina and Xyleborini were nested in each of these clades. Other characteristic and well-supported relationships involved the nested position of *Cactopinus* in Micracidini (Jordal & Kaidel, 2016), the close relationship between Hylastini and genera affiliated with *Dendroctonus*, the sister relationship between Phloeotribini and *Chramesus*/ *Pseudochamesus*, between Scolytoplatypodini and part of Hexacolini, and the nested position of *Strombophorus* in African Hylesinini. The similarity between the original five-gene dataset and the current one suggests that a rather limited dataset can be equally adept at solving certain parts of the bark beetle tree as more substantial datasets.

A series of exciting new relationships was discovered with the addition of many new taxa, with significant implications for our understanding of bark beetle evolution (Gohli & Jordal, 2017; Gohli *et al.*, 2017). For the first time, we were able to reliably assess the sister group to the pygmy borers in the genus *Hypothenemus*. Famous for the world coffee pest

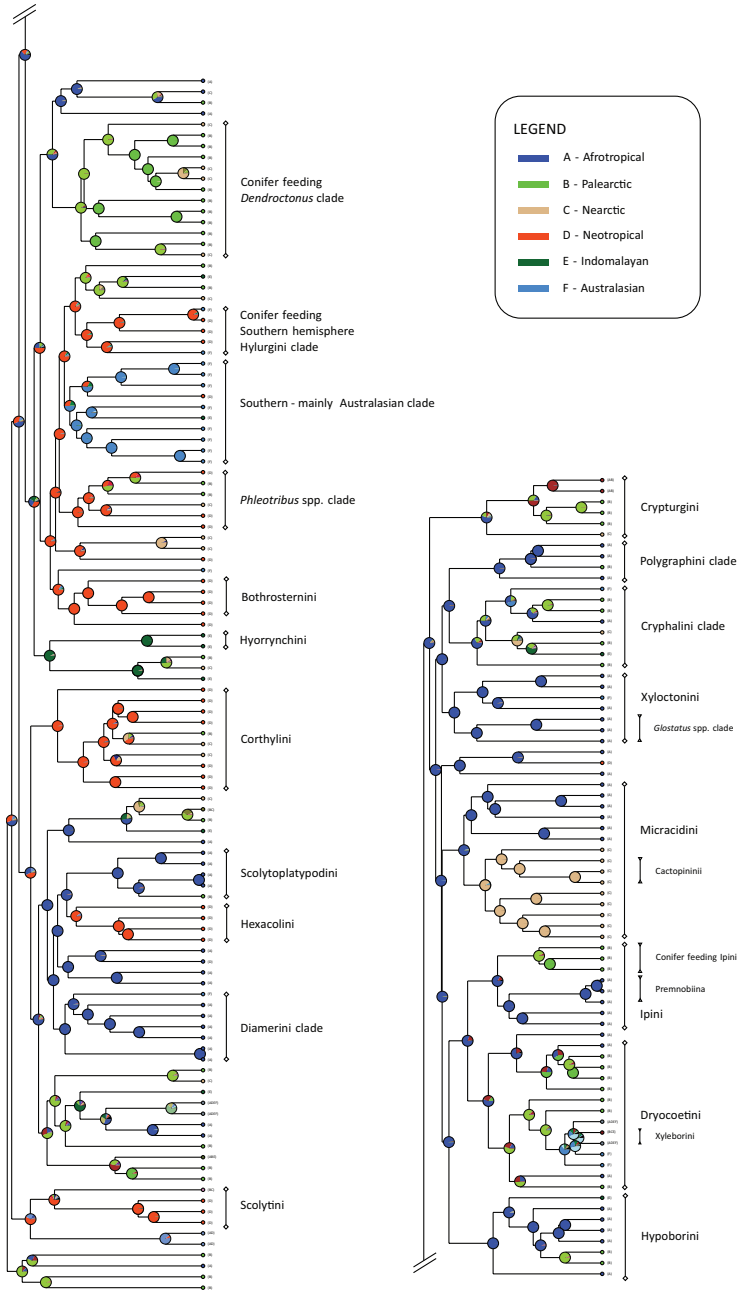


Fig. 7. Ancestral range reconstruction based on the Bayesian binary Markov chain Monte Carlo method in RASP using the tree sample from the BEAST analysis. Pie charts at nodes represent the probabilities for ancestral area.

H. hampei (Brun *et al.*, 1995), the genus is also highly diverse in terms of species, all of which inbreed by sibling mating. This is a very old lineage of permanent inbreeders which, since the split from an outbreeding Afrotropical genus about 57 Ma, has continued to produce many lineages containing some of the most abundant and omnipresent bark beetle complexes on earth (Kambestad *et al.*, 2017). The Malagasy sister group of *Hypothenemus* typically breeds in small-diameter lianas and twigs, but differs from *Hypothenemus* by having a normal monogamous mating system (B. H. Jordal, unpublished data; Schedl, 1977). Previous taxonomic treatments have placed the sib-mating genus *Cryptocarenus* as the closest relative of *Hypothenemus* (Wood, 1957, 1986); however, this genus shares the same ancestral lineage as *Araptus* and related genera in Corthylini. *Araptus* also contain a small clade of permanent inbreeders, typically breeding in petioles and twigs (Kirkendall *et al.*, 2015). It is noteworthy that many origins of inbreeding evolved lineages where species breed in twigs and lianas, including the inbreeding genera *Bothrosternus* and *Sueus* Eggers (Kirkendall, 1983).

Bad taxonomy, bad data, or just very old radiations?

One of the more peculiar incongruences with the current classification was the reciprocal paraphyly of Hylurgini and Hylesinini. Morphological differences between the two tribes are not very precisely defined (Wood, 1986) and many morphological features are even erroneously described for several genera (e.g. Mecke, 2004; Jordal *et al.*, 2011; Jordal & Kaidel, 2016). It is therefore not entirely unexpected to observe incongruence between molecular data and the current classification of these bark beetles.

Future revisions will result in several new and thereby less inclusive tribes, largely restricted to specific biogeographical regions, with a clear distinction between the northern and southern hemisphere faunas. Even though biogeographical patterns are overall not as stringent in Scolytinae as in the extremely endemic Platypodinae (Jordal, 2015), there are several other indications that biogeography is a stronger factor in bark beetle classification than was previously anticipated. Recent studies have demonstrated a strict separation of Afrotropical and Neotropical genera in Micracidini (Jordal & Kaidel, 2016) and Phrixosomatini (Jordal, 2012), and a similarly conservative pattern was revealed for Asian versus Afrotropical species of Scolytoplatypodini (Jordal, 2013). These examples fall in line with other groups that are restricted in distribution, such as the monophyly of tribes that are largely endemic to a single area or closely situated continents. Hence, the distinction between northern and southern clades of mixed Hylurgini and Hylesinini genera is probably illustrating insufficient taxonomic practice in the past rather than misleading molecular data in the current study. Detailed morphological studies must nevertheless be made to distinguish between insufficient taxonomic work and possible parallel evolution of similar morphotypes.

Some of the groups that are geographically restricted to one of the hemispheres are associated with conifers in the boreal zone. Most, if not all, of such clades are of Tertiary age and nested within much older angiosperm-associated lineages. Hence, the origin of conifer feeding did not generally stem from primitive associations as previously suggested (Sequeira *et al.*, 2000), but instead documented multiple recent radiations on these host plants (Gohli & Jordal, 2017; Gohli *et al.*, 2017). Another specialized feeding trait involved fungus farming – a factor associated with the largest and most recent large-scale species radiations in Scolytinae – in Xyleborini. Other fungus-farming lineages are older, albeit no more than 40–60 Ma, appearing with dramatically extended tropical forests at the onset of thermal maximum (Zachos *et al.*, 2003; Westerhold *et al.*, 2009). Estimates for some of these lineages (e.g. Hyorrhynchini, Scolytoplatypodini and *Camptocerus* Dejean) were older than in previous analyses (Jordal & Cognato, 2012), as they were moved back in time more than 20 Ma. Still, there are no fungus-farming lineages in Scolytinae comparable to the much older Platypodinae Shuckard (see Jordal, 2015).

Because deep divergence in Scolytinae was not associated with particular changes in ecology or biogeography of the group, it seems likely that deep coalescence, and thereby age, is the major explanation for the low resolution obtained for the deepest nodes. It is a well-known problem that phylogenetic resolution is particularly difficult for Cretaceous relationships (Cameron & Mardulyn, 2001; Moulton, 2003). Large data volumes may not necessarily overcome these problems, particularly in the face of high evolutionary rates typical for many of the new genes included (Pistone *et al.*, 2016). Lack of deep resolution could also be caused in part by irregularities in the amplification of one or several of the new genes included here. However, the existence of paralogous gene copies is rather unlikely for these genes based on OrthoDB analyses, and individual gene analyses did not indicate particular anomalies in branch lengths or relationships (Pistone *et al.*, 2016; Mugu *et al.*, 2018). Missing data could therefore have a more negative influence on tree resolution, even though *c*. 60% of the taxa were sequenced for more than half of the genes. Nevertheless, missing data did not appear to be particularly troublesome, because less inclusive gene sampling with a gradually increased focus on the best sampled genes did not increase resolution or node support. The highest average node support was obtained with all data included, although exclusion of some of the less well-performing gene fragments produced similar results. Many studies support the inclusion of all data, showing that high levels of missing data usually have negligible effect on tree resolution (Fulton & Strobeck, 2006; Wiens & Morrill, 2011).

Implications for revised classification

The updated phylogeny presented here, together with several previous studies, provides a new foundation to revise the classification of Scolytinae. However, several critical nodes are still insufficiently supported by molecular data and more data are needed to resolve some of the older groups in this subfamily. A formal reclassification is therefore not yet advisable and we will continue to use Wood's (1986) classification until such data

become available, hopefully in the near future. In the following section, we are pointing towards the most likely changes that will emerge in the next reclassification.

Tribe Scolytini. This tribe is characterized by morphological characters not shared by other bark and ambrosia beetles (Smith & Cognato, 2014), which agree with the placement as an early diverging lineage. These beetles represent a highly supported and morphologically well-defined tribe (Jordal & Cognato, 2012; Smith & Cognato, 2014). Most previous studies, including this one, place Scolytini at the root of Scolytinae. However, some studies also indicate the possibility that this group is not even member of the subfamily (Gillett *et al.*, 2014; Mugu *et al.*, 2018). If additional molecular data continue to support the separate standing of Scolytini, it will require a name shift for the bark and ambrosia beetle subfamily to Hylesininae.

Tribes Hylurgini, Hylastini and Hylesinini. The classification of each of these tribes needs substantial changes to reflect their true evolutionary history. *Hylastes* and *Hylurgops* are clearly nested in a boreal subclade of Hylurgini (Fig. 2), reflecting a much more recent origin than for other tribes (Fig. 6). Hylastini can therefore only be regarded as a subtribe, or dissolved altogether. Hylurgini and Hylesinini were paraphyletic with respect to each other and to genera in other tribes. The name-bearing genus *Hylurgus* defines the tribe, and in this study included in the same clade *Hylastes*, *Hylurgops*, *Dendroctonus*, *Tomicus* Wollaston and *Pseudohylesinus* LeConte. Hylesinini included in the same clade the name-bearing genus *Hylesinus*, and the closely related *Pteleobius* Fabricius*, Hylurgopinus* Eichhoff and the Palearctic species of *Chaetoptelius*. Southeast Asian and Australian species of *Chaetoptelius* and other genera in these regions grouped separately from the northern boreal taxa, warranting description of new tribes. A similar situation occurs in Hylurgini in which two different subclades associated with *Araucaria* hosts emerged, although some analyses placed these two lineages together. The genus *Xylechinus* is clearly a non-sense genus where a revision must take geographic distribution into consideration (Fig. 7). Some genera which are currently in Hylurgini or Hylesinini, such as *Rhopalopselion*, *Hapalogenius*, *Hylesinopsis* Eggers and *Dactylipalpus* Chapuis, grouped with members of different tribes. It is clear from several molecular studies (Jordal & Cognato, 2012) and morphology (Jordal & Kaidel, 2016) that *Strombophorus* is closely related to *Rhopalopselion*, *Hapalogenius* and will be the foundation of a new tribe. The observation of the primarily Neotropical *Phloeotribus* as nested in a southern hemisphere clade indicates a strong biogeographical influence on future classifications.

Tribe Phrixosomatini. The monotypic genus formed a well-defined clade divided in two main Eocene-aged clades – one in the Neotropics and one in the Afrotropics (Jordal & Cognato, 2012). The characteristic morphology of tibiae in this genus (Jordal, 2012) shows similarity with *Sphaerotrypes*, an otherwise unexpected sister group in our analyses. Although not conclusive, there may be reasons to place the two genera closer in a revised classification.

Tribe Hyorrhynchini. A close relationship between *Hyorrhynchus* and *Sueus* was confirmed by genetic data for the first time. The third genus *Pseudohyorrhynchus* Murayama is supposedly very closely related (Beaver & Gebhardt, 2004), which makes the tribe a coherent group. Affinities with other tribes remain unresolved.

Tribe Diamerini. This tribe is highly polyphyletic, as currently defined. The name-bearing genus *Diamerus* Erichson grouped occasionally with *Acacicis* Lea, but never with other genera in that tribe, and relationships with other genera were weakly supported. *Sphaerotrypes* may be the sister group to *Prixosoma*, whereas *Strombophorus* (and the morphologically nearly identical *Pernophorus* Strohmeyer) is a very close relative of the hylesinine genera *Rhopalopselion* and *Hapalogenius*, as firmly documented in all molecular studies to date (Farrell *et al.*, 2001; Jordal *et al.*, 2008; Jordal & Cognato, 2012). These three genera are obviously misplaced in the current classification, as they – in addition to molecular data – are supported by a range of morphological characters, particularly in the proventriculus (Nobuchi, 1969).

Tribe Bothrosternini. The species included in this study represent the full range of morphological variation for the tribe and were monophyletic. It is apparently not closely related to any other tribes, although a very distant and weakly supported relation to Scolytini has been suggested (Jordal & Kaidel, 2016).

Tribe Phloeotribini. This tribe currently contains three genera, which are defined by a characteristic lamellate antennal club. Our molecular data for two of the genera, and morphological data (Cognato & Smith, 2010), support monophyly of the tribe. Because *Dryotomicus* appeared nested in the genus, a thorough revision of *Phloeotribus* seems necessary. A revision also needs to consider the closely related phloeosinine genera *Chramesus* and *Pseudochramesus* (see Jordal & Cognato, 2012), especially because the sister group to these four genera was a Neotropical species of *Xylechinus*, belonging to a group of closely related *Xylechinus* species previously placed in the now synonymized phloetribine genus *Phtorophloeus* Rey (Wood & Bright, 1992).

Tribe Phloeosinini. The eight genera included in this study grouped into six separate lineages and the tribe therefore appears highly polyphyletic as currently classified. Only *Hyledius* Sampson was placed close to the type genus *Phloeosinus* as expected (Jordal & Cognato, 2012), and *Chramesus* and *Pseudochramesus* grouped together, but with Phloeotribini. This tribe is clearly an artificial group including species with very different morphologies, which were probably grouped together in the current classification because they had no obvious relation to other tribes (Wood, 1986; Jordal, 2010).

Tribe Hypoborini. A well-defined and nearly monophyletic tribe, except for the genera *Chaetophloeus* and *Zygophloeus* which represent independent lineages. On the other hand, two additional and putatively undescribed genera (vouchers PhCla01 and ChCh01) were unambiguously assigned to the tribe as successive sister groups to the remaining genera of Hypoborini. Desciption of new genera will be treated in a separate taxonomic publication. Furthermore, *Liparthrum* was paraphyletic with respect to *Hypoborus* Erichson, possibly also *Styracoptinus* Wood and *Dacryostactus*, which require a thorough revision of this genus. Hypoborini share several morphological traits with Micracidini (see Jordal & Kaidel, 2016), but our new molecular data suggest a sister relation to the well-resolved Dryocoetini and Ipini.

Tribe Polygraphini. The tribe Polygraphini is a highly polyphyletic assemblage of genera which encompasses at least four lineages of distantly related genera (Fig. 1), and therefore requires substantial revision. Preliminary data (B. H. Jordal, unpublished data) indicate that *Carphobius* Blackman is unrelated to the polygraphine genera included here. It seems likely that *Serrastus* Nunberg is nested in *Polygraphus* Erichson and therefore requires synonymy of this genus in *Polygraphus*. This is an unexpected outcome on the basis that very little morphological variation occurs in the large genus *Polygraphus* and its sister genus *Dolurgocleptes* Schedl (see Jordal, 2009).

Tribes Hexacolini and Scolytoplatypodini. All recent phylogenetic studies have supported a sister relationship between Scolytoplatypodini and at least two genera in Hexacolini (*Scolytodes* and *Gymnochilus*). A third genus – *Pycnarthrum* Chapuis – did not group with other Hexacolini in this study, as in other studies (Jordal *et al.*, 2011; Jordal & Kaidel, 2016); however, low data coverage for this taxon may indicate that this is a spurious result. The genus *Microborus*, on the other hand, has frequently been placed outside of the core Hexacolini in molecular analyses, usually at the base of Scolytinae (Jordal & Cognato, 2012). *Microborus* is also quite unique biologically, and uses the entrance holes of other bark beetles or cossonines (Jordal, 2017), similar to genera in Crypturgini and in *Liparthrum* (Jordal, 2006).

Tribe Micracidini. Several new taxa have recently been added to Micracidini, including *Cactopinus* and *Dendrochilus* (see Jordal & Kaidel, 2016). This is the oldest group of scolytines that are currently defined as a tribe, about the same age as Scolytini, and begun diversifying more than 80 Ma (cf. Fig. 5). Several Afrotropical genera are almost as old as the tribe and are therefore difficult to define. A revision of *Afromicracis* Schedl, *Lanurgus* Eggers and *Pseudomicracis* Eggers is currently in progress (B. H. Jordal, unpublished data)

Tribe Ipini. With the recent addition of Premnobiina and its two containing genera (Cognato, 2013), Ipini is a well-defined sister group to Dryocoetini. The tropical genera *Acanthotomicus* Blandford, *Premnophilus* and *Premnobius* Eichhoff form a clade separate from *Pityogenes* Linnaeus, *Pityokteines* LeConte, *Orthotomicus* Wollaston and *Ips* Böerner, while *Pseudips* Cognato may belong to either clade. The tropical genus *Acanthotomicus* appears to be much older than the recently evolved boreal conifer associates of the tribe (Jordal & Cognato, 2012), and hence probably requires multiple genera.

Tribe Dryocoetini. One of the first clear results of molecular studies of Scolytinae was the recent origin of xyleborine genera inside Dryocoetini (Jordal *et al.*, 2000, 2002, 2008; Farrell *et al.*, 2001; Jordal & Cognato, 2012; Jordal & Kaidel, 2016). Xyleborini was therefore demoted by many authors to the subtribe Xyleborina (Rabaglia *et al.*, 2006; Hulcr *et al.*, 2007). The only defining feature of this subtribe is the fungus-farming behaviour which distinguishes the xyleborine genera from *Coccotrypes*, *Dryocoetiops* and *Ozopemon* in a more inclusive clade of permanently inbreeding and haplo-diploid species (Normark *et al.*, 1999). *Dryocoetes* Ratzeburg constitutes at least two successive sister groups to the inbreeding clade, and hence needs revision (Jordal & Kambestad, 2014). A second and strongly supported clade in Dryocoetini includes *Thamnurgus* Bach, *Triotemnus* Wollaston, *Lymantor*, *Xylocleptes* and *Dactylotrypes* Wollaston, and, to a lesser degree, *Taphrorychus* Dufour and *Cyrtogenius* Blandford (weakly supported). Almost all of these genera need taxonomic revision as the generic limits are uncertain in light of molecular and morphological data (B. H. Jordal, unpublished data).

Tribe Crypturgini. The monophyly of this tribe has not been disputed and is strongly supported by molecular data (see also Jordal & Cognato, 2012). *Coleobothrus* Enderlein was synonymized with *Aphanarthrum* Wollaston by Bright (2014) based on Jordal & Hewitt (2004). Several recent molecular studies furthermore indicate that *Cisurgus* Eichhoff is nested in *Crypturgus* LeConte with the potential for synonymy provided that more complete species sampling will produce the same result.

Tribe Xyloctonini. Molecular and morphological data provide very few indications as to the affinity of this tribe. Only one of the analyses resulted in monophyly, with most analyses placing *Glostatus* elsewhere. This genus is morphologically variable and includes all species described in the cryphaline genus *Stephanopodius* (B. H. Jordal, unpublished data).

Tribe Xyloterini. This is a morphologically and ecologically homogeneous group of beetles, strongly supported by molecular data. Previous analyses of molecular (Jordal & Cognato, 2012) as well as morphological data (Jordal & Kaidel, 2016) have indicated a sister relationship to *Cryphalus* and *Hypocryphalus*, but this was only supported by some of the current analyses. The presumption that Xyloterini is sister lineage to Xyleborini (Wood, 1986) is false.

Tribe Cryphalini. Genera currently classified as Cryphalini occurred in six independent lineages in our analyses,

representing a taxonomic problem that needs to be solved soon. *Trypophloeus* Hopkins is the most atypical of the cryphaline genera and is difficult to relate to any other scolytine genus (Jordal & Cognato, 2012). The name-bearing genus *Cryphalus* and the possibly synonymous genus *Hypocryphalus* were similarly isolated, although previous studies and some of our analyses related these genera to Xyloterini. Most, if not all, species of *Stephanopodius* will become synonyms of *Glostatus* in the next taxonomic revision (all types examined; B. H. Jordal, unpublished data).

The largest clade of cryphaline genera included typical Holarctic genera such as *Procryphalus* LeConte, *Ernoporus* and *Ernoporicus* Lindemann, but also some tropical genera such as *Scolytogenes*, and one species of *Dryocoetiops* which is clearly not a member of that genus (Beaver, 1990). However, it is difficult to see how *D. petioli* fits within the current circumscription of *Procryphalus* where its position is strongly indicated by molecular data (Figs 1, 2).

Many species of *Cosmoderes* and *Ptilopodius* are reciprocally misplaced in the current classification, and furthermore need many of the species assigned to these genera in Madagascar and Eastern Africa a new genus (B. H. Jordal, unpublished data). These species are particularly interesting because they form the sister lineage to the permanently inbreeding *Hypothenemus* and thereby indicate the geographical origin of sibling mating and paternal genome elimination (Brun *et al.*, 1995). *Hypothenemus* was previously (Wood, 1982, 1986) placed close to another group of species that permanently inbreed – the Neotropical genus *Cryptocarenus* (Kirkendall, 1983). However, this genus is closely related to *Araptus* in Corthylini.

Tribe Corthylini. This tribe is characteristically defined by an oblique locking suture on the mesanepisternum (Wood, 1986), except for *Dendroterus* which split early from the other members of the tribe. Morphological examination of *Cryptocarenus diadematus* and *Cryptocarenus seriatus* Eggers revealed the typical corthyline locking suture in these species. This genus and other Corthylini also share a conserved nucleotide indel motif in the second variable region of *Iap2* (Fig. 4), which together with all phylogenetic analyses strongly support a placement of *Cryptocarenus* in Corthylini, close to the genera *Araptus* and *Dacnophthorus*. It is a mystery to us why Wood (1982, 1986) excluded *Cryptocarenus* from Corthylini, particularly because, he was the one that established this character as a useful synapomorphy for corthylines. *Cryptocarenus* then represents a second origin of permanent inbreeding in Corthylini, unless the subclade of 13 inbreeding *Araptus* species (Kirkendall *et al.*, 2015) form the sister group to *Cryptocarenus*.

Conclusion

This study highlights a complex and intricate relationship among the oldest tribes of the subfamily Scolytinae. Adding new data from 13 protein-coding gene fragments (6456 bp) to a previously used molecular dataset (c. 3700 aligned nucleotides) contributed some, but still limited, resolution for ancient divergences. However, many new and interesting relationships were strongly supported, whereas several tribes – as currently classified – were significantly rejected due to poly- and paraphyly. As a consequence, a revision of these tribes is critically needed. The use of additional molecular data and a thorough survey of morphological characters will be of major importance to increase our understanding of the evolutionary history of Scolytinae.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/syen.12281

Table S1. Taxa sampled, with the country of origin, classification, and DNA voucher code.

Table S2. GenBank accession numbers for taxa included in this study (see Table S1).

Figure S1. Phylogenetic tree based on the Bayesian analysis of 18 partitions (genes). Posterior probabilities for the 18-partition analysis are reported on branches, followed by those maximum posterior probabilities (*) for the seven-partition analysis.

Figure S2. Phylogenetic tree based on the analysis of amino acids in MRBAYES. Node support (PP/BS) is reported above and below branches, respectively.

Figure S3. Phylogenetic tree based on the Bayesian analysis of 13 novel protein-coding genes. The dataset was partitioned by codon position; posterior probabilities are reported on the branches.

Figure S4. Phylogenetic tree based on the Bayesian analysis of five commonly used genes in beetle phylogeny. The dataset was partitioned by codon position; posterior probabilities are reported on the branches.

Figure S5. Phylogeny based on the Bayesian analysis of *Arrestin2*. Posterior probabilities are reported on the branches.

References

Beaver, R.A. (1990) New records and new species of bark and ambrosia beetles from Thailand (Coleoptera; Scolytidae and Platypodidae). *Deutsche Entomologische Zeitschrift*, **37**, 279–284.

Beaver, R.A. & Gebhardt, H. (2004) Notes on the tribe Hyorrhynchini (Col., Curculionidae, Scolytinae). *Serangga*, **91**, 91–102.

Bright, D.E. (2014) A catalog of Scolytidae and Platypodidae (Coleoptera). Supplement 3 (2000-2010), with notes on subfamily and tribal reclassifications. *Insecta Mundi*, **356**, 1–336.

Brun, L.O., Stuart, J., Gaudichon, V., Aronstein, K. & French-Constant, R.H. (1995) Functional haplodiploidy: a mechanism for the spread of insecticide resistance in an important international insect pest.

Proceedings of the National Academy of Sciences of the United States of America, **92**, 9861–9865.

Cameron, S.A. & Mardulyn, P. (2001) Multiple molecular data sets suggest independent origins of highly eusocial behavior in bees (Hymenoptera: Apinae). *Systematic Biology*, **50**, 194–214.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540–552.

Cognato, A.I. (2013) Molecular phylogeny and taxonomic review of Premnobiini Browne, 1962 (Coleoptera: Curculionidae: Scolytinae). *Frontiers in Ecology and Evolution*, **1**, 1.

Cognato, A.I. & Grimaldi, D. (2009) 100 million years of morphological conservation in bark beetles (Coleoptera: Curculionidae: Scolytinae). *Systematic Entomology*, **34**, 93–100.

Cognato, A.I. & Smith, S.M. (2010) Resurrection of *Dryotomicus* Wood and description of two new species from the Amazon River Basin (Coleoptera, Curculionidae, Scolytinae, Phloeotribini). *ZooKeys*, **56**, 49–64.

Dohet, L., Grégoire, J.C., Berasategui, A., Kaltenpoth, M. & Biedermann, P.H. (2016) Bacterial and fungal symbionts of parasitic *Dendroctonus* bark beetles. *FEMS Microbiology Ecology*, **92**, pii: fiw129. https://doi.org/10.1093/femsec/fiw129.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Farrell, B.D., Sequeira, A.S., O'Meara, B.C., Normark, B.B., Chung, J.H. & Jordal, B.H. (2001) The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution*, **55**, 2011–2027.

Fulton, T.L. & Strobeck, C. (2006) Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. *Molecular Phylogenetics and Evolution*, **41**, 165–181.

Gillett, C.P., Crampton-Platt, A., Timmermans, M.J., Jordal, B.H., Emerson, B.C. & Vogler, A.P. (2014) Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). *Molecular Biology and Evolution*, **31**, 2223–2237.

Gohli, J. & Jordal, B.H. (2017) Explaining biogeographical range size and measuring its effect on species diversification in bark beetles. *Journal of Biogeography*, **44**, 2132–2144. https://doi.org/10.1111/jbi.13005.

Gohli, J., Kirkendall, L.R., Smith, S.M., Cognato, A.I., Hulcr, J. & Jordal, B.H. (2017) Biological factors contributing to bark and ambrosia beetle species diversification. *Evolution*, **71**, 1258–1272.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Research*, **41**, 95–98.

Hulcr, J. & Dunn, R.R. (2011) The sudden emergence of pathogenicity in insect-fungus symbioses threatens naive forest ecosystems. *Proceedings of the Royal Society Biological Sciences Series B*, **278**, 2866–2873.

Hulcr, J., Dole, S.A., Beaver, R.A. & Cognato, A.I. (2007) Cladistic review of generic taxonomic characters in Xyleborina (Coleoptera: Curculionidae: Scolytinae). *Systematic Entomology*, **32**, 568–584.

Hulcr, J., Atkinson, T.H., Cognato, A.I., Jordal, B.H. & McKenna, D.D. (2015) Morphology, taxonomy and phylogenetics of bark beetles. *Bark Beetles: Biology and Ecology of Native and Invasive Species* (ed. by F.E. Vega and R.W. Hofstetter), pp. 41–84. Elsevier Inc, New York, New York.

Jordal, B.H. (2006) Community structure and reproductive biology of bark beetles (Coleoptera: Scolytinae) associated with Macaronesian Euphorbia shrubs. *European Journal of Entomology*, **103**, 71–80.

Jordal, B.H. (2009) The Madagascan genus Dolurgocleptes Schedl (Coleoptera: Curculionidae, Scolytinae): description of a new species and transfer to the tribe Polygraphini. *Zootaxa*, **2014**, 41–50.

Jordal, B.H. (2010) Revision of the genus Phloeoditica Schedl – with description of two new genera and two new species in Phloeosinini (Coleoptera, Curculionidae, Scolytinae). *Zookeys*, **56**, 141–156.

Jordal, B.H. (2012) Phrixosoma concavifrons – a sexually dimorphic Phrixosomatini (Coleoptera: Curculionidae) from the Udzungwa mountains in Tanzania. *Zootaxa*, **3255**, 52–56.

Jordal, B.H. (2013) Deep phylogenetic divergence between Scolytoplatypus and Remansus, a new genus of Scolytoplatypodini from Madagascar (Coleoptera, Curculionidae, Scolytinae). *ZooKeys*, **352**, 9–33.

Jordal, B.H. (2015) Molecular phylogeny and biogeography of the weevil subfamily Platypodinae reveals evolutionarily conserved range patterns. *Molecular Phylogenetics and Evolution*, **92**, 294–307.

Jordal, B.H. (2017) Ancient diversity of Afrotropical *Microborus*: three endemic species – not one widespread. *Zookeys*, **710**, 33–42.

Jordal, B.H. & Cognato, A.I. (2012) Molecular phylogeny of bark and ambrosia beetles reveals multiple origins of fungus farming during periods of global warming. *BMC Evolutionary Biology*, **12**, 133.

Jordal, B.H. & Hewitt, G.M. (2004) The origin and radiation of Macaronesian beetles breeding in Euphorbia: the relative importance of multiple data partitions and population sampling. *Systematic Biology*, **53**, 711–734.

Jordal, B.H. & Kaidel, J. (2016) Phylogenetic analysis of Micracidini bark beetles demonstrates a single trans-Atlantic disjunction and inclusion of Cactopinus in the American clade. *The Canadian Entomologist*, **149**, 8–25.

Jordal, B.H. & Kambestad, M. (2014) DNA barcoding of bark and ambrosia beetles reveals excessive NUMTs and consistent east-west divergence across Palearctic forests. *Molecular Ecologt Resources*, **14**, 7–17.

Jordal, B.H. & Kirkendall, L.R. (1998) Ecological relationships of a guild of tropical beetles brreding in Cecropia petioles in Costa Rica. *Journal of Tropical Ecology*, **14**, 153–176.

Jordal, B.H., Normark, B.B. & Farrell, B.D. (2000) Evolutionary radiation of an inbreeding haplodiploid beetle lineage (Curculionidae, Scolytinae). *Biological Journal of the Linnean Society*, **71**, 483–499.

Jordal, B.H., Beaver, R.A., Normark, B.B. & Farrell, B.D. (2002) Extraordinary sex ratios, and the evolution of male neoteny in sib-mating Ozopemon beetles. *Biological Journal of the Linnean Society*, **75**, 353–360.

Jordal, B.H., Gillespie, J.J. & Cognato, A.I. (2008) Secondary structure alignment and direct optimization of 28S rDNA sequences provide limited phylogenetic resolution in bark and ambrosia beetles (Curculionidae: Scolytinae). *Zoologica Scripta*, **37**, 43–56.

Jordal, B.H., Sequeira, A.S. & Cognato, A.I. (2011) The age and phylogeny of wood boring weevils and the origin of subsociality. *Molecular Phylogenetics and Evolution*, **59**, 708–724.

Jordal, B.H., Smith, S.M. & Cognato, A.I. (2014) Classification of weevils as a data-driven science: leaving opinion behind. *Zookeys*, **439**, 1–18.

Kambestad, M., Kirkendall, L.R., Knutsen, I.L. & Jordal, B.H. (2017) Cryptic and pseudo-cryptic diversity in the world's most common bark beetle – *Hypothenemus eruditus*. *Organisms Diversity & Evolution*, **17**, 633–652. https://doi.org/10.1007/s13127-13017-10334-13126.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059–3066.

Kawasaki, Y., Schuler, H., Stauffer, C., Lakatos, F. & Kajimura, H. (2016) *Wolbachia* endosymbionts in haplodiploid and diploid

scolytine beetles (Coleoptera: Curculionidae: Scolytinae). *Environmental Microbiology Reports*, **8**, 680–688.

Kirejtshuk, A.G., Azar, D., Beaver, R.A., Mandelshtam, M.Y. & Nel, A. (2009) The most ancient bark beetle known: a new tribe, genus and species from Lebanese amber (Coleoptera, Curculionidae, Scolytinae). *Systematic Entomology*, **34**, 101–112.

Kirkendall, L.R. (1983) The evolution of mating systems in bark and ambrosia beetles (Coleoptera: Scolytidae and Platypodidae). *Zoological Journal of the Linnean Society*, **77**, 293–352.

Kirkendall, L.R., Kant, D.S. & Raffa, K.F. (1997) Interactions among males, females and offspring in bark and ambrosia beetles: the significance of living in tunnels for the evolution of social behavior. *The Evolution of Social Behavior in Insects and Arachnids* (ed. by J.C. Choe and B.J. Crespi), pp. 181–215. Cambridge University Press, Cambridge.

Kirkendall, L.R., Biedermann, P.H.W. & Jordal, B.H. (2015) Evolution and diversity of bark and ambrosia beetles. *Bark Beetles: Biology and Ecology of Native and Invasive Species* (ed. by F.E. Vega and R.W. Hofstetter), pp. 85–136. Elsevier Inc, New York, New York.

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

Lartillot, N., Lepage, T. & Blanquart, S. (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, **25**, 2286–2288.

Li, Y., Simmons, D.R., Bateman, C.C., Short, D.P., Kasson, M.T., Rabaglia, R.J. & Hulcr, J. (2015) New fungus-insect symbiosis: culturing, molecular, and istological methods determine saprophytic polyporales mutualists of *Ambrosiodmus* ambrosia beetles. *PLoS One*, **10**, e0137689.

Mariño, Y.A., Verle Rodrigues, J.C. & Bayman, P. (2017) *Wolbachia* affects reproduction and population dynamics of the coffee berry borer (*Hypothenemus hampei*): implications for biological control. *Insects*, **8**, 8. http://doi.org/10.3390/insects8010008.

Massoumi Alamouti, S., Tsui, C.K. & Breuil, C. (2009) Multigene phylogeny of filamentous ambrosia fungi associated with ambrosia and bark beetles. *Mycological Research*, **113**, 822–835.

McKenna, D.D., Sequeira, A.S., Marvaldi, A.E. & Farrell, B.D. (2009) Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 7083–7088.

Mecke, R. (2004) New species of *Pachycotes* sharp and *Xylechinus* Chapuis (Coleoptera: Curculionidae: Scolytinae, Tomicini) from new Caledonian *Araucaria* spp. (Coniferales: Araucariaceae). *New Zealand Journal of Zoology*, **31**, 343–349.

Moulton, J.K. (2003) Can the current molecular arsenal adequately track rapid divergence events within Simuliidae (Diptera)? *Molecular Phylogenetics and Evolution*, **27**, 45–57.

Mugu, S., Pistone, D. & Jordal, B.H. (2018) New molecular markers resolve the phylogenetic position of the enigmatic wood-boring weevils Platypodinae (Coleoptera: Curculionidae). *Arthropod Systematics and Phylogeny*, in press.

Nobuchi, A. (1969) A comparative morphological study of the proventriculus in the adult of the superfamily Scolytoidea (Coleoptera). *Bulletin of the Government Forest Experiment Station*, **224**, 39–110.

Normark, B.B., Jordal, B.J. & Farrell, B.D. (1999) Origin of a haplodiploid beetle lineage. *Proceedings of the Royal Society Biological Sciences Series B*, **266**, 2253–2259.

Nylander, J.A.A. (2004) *MrModeltest v2*. Computer program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala.

Pistone, D., Mugu, S. & Jordal, B.H. (2016) Genomic mining of phylogenetically informative nuclear markers in bark and ambrosia beetles. *PLoS One*, **11**, e0163529. https://doi.org/10.1371/journal.pone.0163529.

Rabaglia, R.J., Dole, S.A. & Cognato, A.I. (2006) Review of American Xyleborina (Coleoptera: Curculionidae: Scolytinae) occurring North of Mexico, with an illustrated key. *Annals of the Entomological Society of America*, **99**, 1034–1056.

Rambaut, A. & Drummond, A. (2007) *Tracer Version 1.4*. Computer program and documentation distributed by the author. URL http://beast.bio.ed.ac.uk/Tracer [accessed on 30 August 2016].

Ranger, C.M., Schultz, P.B., Frank, S.D., Chong, J.H. & Reding, M.E. (2015) Non-native ambrosia beetles as opportunistic exploiters of living but weakened trees. *PLoS One*, **10**, e0131496.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Schedl, K.E. (1977) Die Scolytidae und Platypodidae Madagaskars und einiger naheliegender Inselgruppen. *Mitteilungen der Forstlichen Bundes-Versuchsanstalt Wien*, **119**, 1–326.

Sequeira, A.S., Normark, B.B. & Farrell, B.D. (2000) Evolutionary assembly of the conifer fauna: distinguishing ancient from recent associations in bark beetles. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 2359–2366.

Six, D.L. (2012) Ecological and evolutionary determinants of bark beetle – fungus symbioses. *Insects*, **3**, 339–366. http://doi.org/10.3390/insects3010339.

Smith, S.M. & Cognato, A.I. (2014) A taxonomic monograph of Nearctic *Scolytus* Geoffroy (Coleoptera, Curculionidae, Scolytinae). *ZooKeys*, **450**, 1–182.

Swofford, D.L. (2011) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*. Sinauer Associates, Sunderland, Massachusetts.

Westerhold, T., Röhl, U., McCarren, H.K. & Zachos, J.C. (2009) Latest on the absolute age of the Paleocene–Eocene Thermal Maximum (PETM): new insights from exact stratigraphic position of key ash layers +19 and −17. *Earth and Planetary Science Letters*, **287**, 412–419.

Waterhouse, R.M., Tegenfeldt, F., Li, J., Zdobnov, E.M. & Kriventseva, E.V. (2013) OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Research*, **41**, 358–365.

Wiens, J.J. & Morrill, M.C. (2011) Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Systematic Biology*, **60**, 719–731.

Wood, S.L. (1957) New species of bark beetles (Coleoptera: Scolytidae), mostly Mexican. *Great Basin Naturalist*, **17**, 105–110.

Wood, S.L. (1982) The bark and ambrosia beetles of North and Central America (Coleoptera: Scolytidae), a taxonomic monograph. *Great Basin Naturalist Memoirs*, **6**, 1–1359.

Wood, S.L. (1986) A reclassification of the genera of Scolytidae (Coleoptera). *Great Basin Naturalist Memoirs*, **10**, 1–110.

Wood, S.L. & Bright, D. (1992) A catalog of Scolytidae and Platypodidae (Coleoptera). Part 2: taxonomic index. *Great Basin Naturalist Memoirs*, **13**, 1–1553.

Yu, Y., Harris, A.J., Blair, C. & He, X. (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, **87**, 46–49.

Zachos, J.C., Wara, M.W., Bohaty, S. *et al.* (2003) A transient rise in tropical sea surface temperature during the Paleocene-Eocene thermal maximum. *Science*, **302**, 1551–1554.

Accepted 8 November 2017

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 978-82-308-3809-9