BioOne COMPLETE

Estimating Fossil ant Species Richness in Eocene Baltic Amber

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Estimating fossil ant species richness in Eocene Baltic amber

DAVID PENNEY and RICHARD F. PREZIOSI

Fossil insects in amber are often preserved with life-like fidelity and provide a unique insight to forest ecosystems of the geological past. Baltic amber has been studied for more than 300 years but despite the large number of described fossil species (ca. 3500 arthropods) and abundance of fossil material, few attempts have been made to try and quantify statistically how well we understand the palaeodiversity of this remarkable Fossil-Lagerstätte. Indeed, diversity estimation is a relatively immature field in palaeontology. Ants (Hymenoptera: Formicidae) are a common component of the amber palaeobiota, with more than 100 described species representing approximately 5% of all inclusions encountered. Here we apply quantitative statistical species richness estimation techniques to Baltic amber data for the first time. We use species level data from a sample size of 12 769 specimens and conclude that around 29% of the Baltic amber ant fauna has yet to be discovered. The species richness accumulation curve clearly reaches its asymptote at around 9650 specimens, indicating this as the minimum sample size required for a reasonable estimate of species richness for ants alone. Hence, it is hardly surprising that previous studies concerning so-called "representative" samples of the entire palaeobiota, consisting of at most a few thousand inclusions do not agree with each other. Nonetheless, we demonstrate that it is possible to apply quantitative techniques to amber derived data and this should be the preferred approach wherever possible, rather than generating qualitative conclusions of little value for comparative purposes.

Introduction

Amber is fossilized tree resin renowned for preserving an immense diversity of fossil arthropods with life-like fidelity (Penney 2010; Penney and Jepson 2014). Quantifying this palaeodiversity is as much a fundamental necessity for understanding the nature of fossil ecosystems, as is quantifying biodiversity for extant organisms to understand the world we live in today. Yet our knowledge in this regard is extremely limited and indeed, diversity estimation is a relatively immature field in palaeontology (Huang et al. 2014). It is possible to conduct large-scale quantitative palaeoecological investigations of amber faunas (e.g., Penney 2002, 2005; Penney and Langan 2006), but in many cases such investigations are generally hindered by lack of data conducive to statistical analysis. This is rather absurd given that Baltic amber has been studied for almost 300 years

(Wichard and Grevin 2009), has approximately 3500 species of arthropods described from it (Weitschat and Wichard 2010), and is still being extracted from the ground in considerable quantities. For example, it has been estimated that approximately 510 tonnes of amber were extracted in the Baltic region during the year 2000 and that approximately two million (a very crude estimate) new inclusions from Baltic amber alone should be available for study each year (e.g., Clark 2010). Indeed, Klebs recognized the need for quantifying the palaeodiversity of amber inclusions at the turn of the twentieth century (Klebs 1910).

Klebs (1910) investigated an unsorted Baltic amber sample of 200 kg directly from the mine and documented a total of 13 877 inclusions, but these were identified only to order. In a similar study, Sontag (2003) investigated an "unsorted" Baltic amber sample of 42 kg, documenting 7079 arthropods, again identified mainly to order. However, Zherikhin and Eskov (2006) considered Sontag's (2003) samples to have been "pre-selected" in favour of larger pieces, and hence not truly representative. Additional studies comparing so-called "representative" samples as baseline palaeodiversity datasets, including for other deposits such as Bitterfeld (Germany) and Rovno (Ukraine) ambers (often considered to be the same as Baltic amber, e.g., Szwedo and Sontag 2013), include Hoffeins and Hoffeins (2004) and Perkovsky et al. (2007). None of these studies produced any conclusive quantitative results.

Zherikhin and Eskov (2006) sampled unsorted amber samples in situ at a Baltic amber mine, identified 1312 inclusions to suborder/family level and compared their assemblage to known museum collections. They summarized their conclusions as three main points: (i) so-called representative (i.e., non-pre-selected, unbiased) samples can differ from one another for reasons currently unknown; (ii) generalized higher rank (e.g., order) taxonomic lists are of low comparative value; (iii) lower rank indicator taxa (e.g., genera and species) are more likely to be useful for comparative purposes.

So-called "representative" or unbiased (through lack of pre-selection) samples reported in the literature to date no doubt differ as a result of insufficient sample sizes, so cannot really be considered as representative. However, this is difficult to assess in published studies because most have merely compared percentage values, without any data rarefaction or weighting to account for different sample sizes. Even for very large amber samples the absence of a particular taxon does not necessarily reflect its non-existence in the original amber forest (e.g., when comparing "representative" samples from contemporaneous lo-
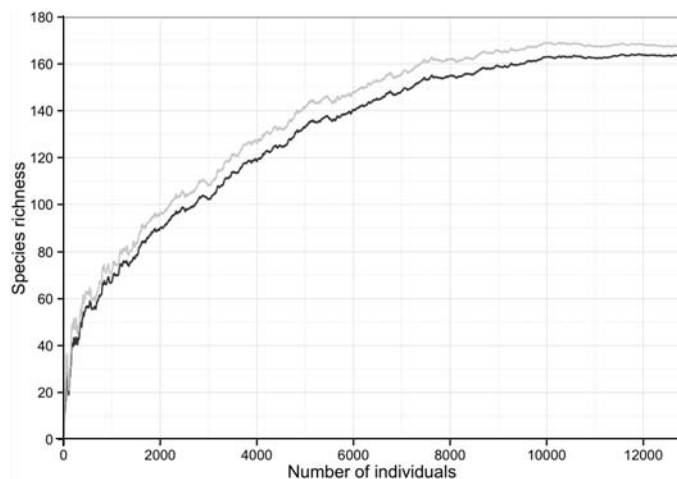
Fig. 1. The Chao 1 (top line) and ACE (bottom line) richness estimates computed using Colwell (2013); note the slightly lower ACE.
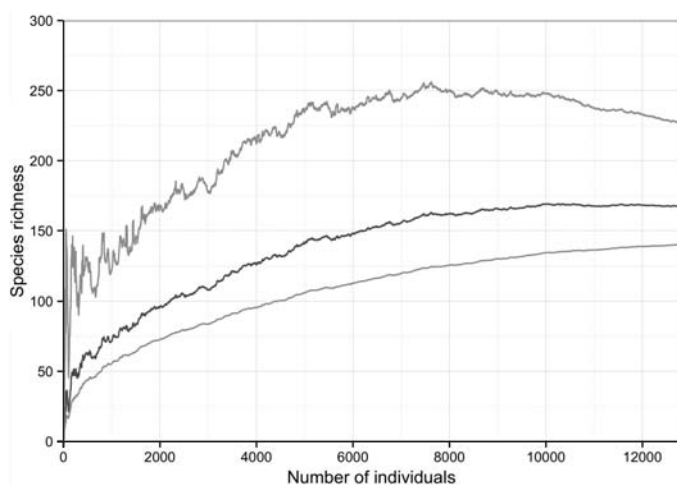

Fig. 2. The Chao1 estimate with 95% confidence intervals; asymptote value = 167.44.

calities, such as Rovno and Baltic) because it is impossible to prove such a negative and this is particularly true for rare taxa.

There is no obvious reason why it should not be possible to calculate what an appropriate "representative" sample size should be by sampling unsorted amber inclusions and generating species accumulation curves for the entire fauna and/or for individual taxonomic groups (e.g., species within families). Dlussky and Rasnitsyn (2009) compared the ant faunas in European Eocene ambers. They found a continuing linear increase in numbers of new species encountered in collections of 1500 specimens or less, concluding that the asymptote would not be approached until significantly larger collections were examined, but this issue was not considered further. Here, we reanalyse these data using extrapolative statistical techniques to predict the total species richness of the Baltic amber ant fauna and use our results to consider why previous comparisons of so-called "representative samples" have not yielded any satisfactory results to date.

Abbreviations.—ACE, abundance coverage-based richness estimate; CI, 95% confidence intervals; CV, Chao's estimated coefficient of variation for abundance distribution.

Methods

There are a variety of powerful techniques for modelling species accumulation curves and for estimating total species richness (e.g., Colwell et al. 2004; Chao et al. 2009; Colwell 2013), including standard software packages that also produce indicators of variance and confidence limits as measures of reliability. We used EstimateS by Colwell (2013) to generate the Chao 1 richness estimate based on the Baltic amber ant data of Dlussky and Rasnitsyn (2009: table 2); a complete user guide to this software is available online: http://viceroy. eeb.uconn.edu/estimates/EstimateSPages/EstSUsersGuide/EstimateSUsersGuide.htm. The two largest species level samples for Baltic amber ants (Wheeler 1915, corrected for synonymy [data column 3 in table 2 of Dlussky and Rasnitsyn 2009] and Dlussky 2009 [new original data, column 4 in table 2 in Dlussky and Rasnitsyn 2009]) were combined to produce a total sample size of 12 769 individuals of 118 different species. Species previously unrecognized by Wheeler (1915) (depicted with a "+" in Dlussky and Rasnitsyn [2009: table 2]), where entered in our data with a value of one; the undescribed genus in Dlussky's (2009 [see column 4 in table 2 in Dlussky and Rasnitsyn 2009]) new data was not included. Ants are useful in such analyses because the majority of individuals encountered are female workers. By contrast, in other common inclusions such as flies (Diptera), males and females are often encountered and usually cannot be matched confidently to the same species. The predicted species richness estimates were calculated using Colwell (2013) set at the default parameters and using type 3 data entry (individual-based abundance data).

In addition to the Chao 1 richness estimate with corresponding 95% confidence intervals, EstimateS concurrently produces an abundance coverage-based richness estimate (ACE). Chao 1 is a universally valid lower bound estimator of species richness and can be applied to any species abundance distribution and any sample size. In general, it is more reliable as a richness indicator if the sample size is sufficiently large. A rough guideline for "sufficient" sample size is that the estimated sample completeness should be at least 50%. For Chao 1, this means the proportion of singletons should be less than 50%, which is the case for the data analyzed here. Because Chao's CV for abundance distribution was >0.5 the analyses were re-computed with the Chao 1 parameter set at the classic, instead of bias-converted, option in the diversity settings. Using these settings the largest value (ACE vs Chao 1) is considered to be the best approximation. The data were exported and graphs generated using Microsoft Excel and R.

Results

The Chao 1 richness estimate was slightly higher than the ACE (Fig. 1). Chao 1 had an asymptote value of 167.44 and 95% confidence intervals of 140.43–226.98 with an analytical standard deviation of 20.78 (Fig. 2). The input and output data are provided as Supplementary Online Material available at http:// app.pan.pl/SOM/app59-Penney_Preziosi_SOM.pdf.

Discussion

This is the first attempt at producing a statistical richness estimate based on a very large amber dataset at the species level. Regardless of the CIs and the slightly high standard deviation, there can be no doubt that what appears to be a reliable asymptote has been reached. Hence, although our sample size could still warrant being slightly larger, this is a reasonably good approximation and suggests that around 29% of the Baltic amber ant fauna has yet to be discovered. This does not seem unreasonable given the large number of singletons (species known from one individual) in the original dataset (see Dlussky and Rasnitsyn 2009: table 2). Indeed, new Baltic amber ant species have been described since their data were compiled (e.g., Dlussky 2010; Radchenko and Dlussky 2013). Remember, this is a lower bound estimate for ant species richness in Baltic amber not ant species richness in the amber forest, which would have been significantly greater, although how much so is impossible to speculate on at present.

The species richness accumulation curve clearly reaches its asymptote at around 9650 specimens, indicating this as the minimum sample size required for a reasonable estimate of species richness in this instance. Taking ants as representing approximately 5% of all Baltic amber inclusions (LaPolla et al. 2013) and that the asymptote in the above analysis was not reached until 9650 inclusions, it follows (very crudely but based on the only reasonable analysis available) that the minimum number of inclusions for estimating the richness of total arthropod palaeodiversity in Baltic amber may be in the region of 193 000. Hence, it is hardly surprising that previous studies concerning so-called representative samples of at most a few thousand inclusions do not agree with each other.

Conclusions

Unbiased (without pre-selection) amber collections referred to in the published literature are too small to be truly representative of the palaeobiota and so should be used with caution. Just as neontologists use quantitative techniques to address large (and small) scale ecological questions with respect to the modern insect fauna, in order to address topical issues such as changes in biodiversity as a result of habitat loss, climate change, etc., we believe a similar approach using large palaeoentomological datasets, such as those found in amber, can be used to investigate changes over geological timescales. Here, we demonstrate for the first time that it is possible to apply quantitative richness estimation techniques to amber derived data and conclude that this should be the preferred approach wherever possible, rather than generating qualitative conclusions of little value for comparative purposes.

References

Chao, A., Colwell, R.K., Lin, C.-W., and Gotelli, N. 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* 90: 1125–1133.

Clark, N.D.L. 2010. *Amber: Tears of the Gods*. 118 pp. Dunedin Academic Press, Edinburgh.

Colwell, R.K. 2013. *EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples. Version 9*. User's Guide and application published at: http://purl.oclc.org/estimates.

Colwell, R.K., Mao, C.X., and Chang, J. 2004. Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology* 85: 2717–2727.

Dlussky, G.M. 2010. Ants of the genus *Plagiolepis* Mayr (Hymenoptera, Formicidae) from Late Eocene ambers of Europe. *Paleontological Journal* 44: 546–555.

Dlussky, G.M. and Rasnitsyn, A.P. 2009. Ants (Insecta: Vespida: Formicidae) in the Upper Eocene amber of Central and Eastern Europe. *Paleontological Journal* 43: 1024–1042.

Hoffeins, C. and Hoffeins, H.W. 2004. Untersuchungen über die Häufigkeit von Inklusen in Baltischem und Bitterfelder Bernstein (Tertiär, Eozän) aus unselektierten Aufsammlungen unter besonderer Berücksichtigung der Ordnung Diptera. *Studia Dipterologica* 10: 381–392.

Huang, B., Taylor Harper, D.A., and Zhan, R. 2014. Test of sampling sufficiency in palaeontology. *Journal of the Geological Society of Sweden* 136: 105–109.

Klebs, R. 1910. Über Bersteineinschlüsse im allgemeinen und die Coleopteren meiner Bernsteinsammlung. *Schriften der Physikalisch-Ökonomischen Gesellschaft zu Königsberg* 51: 217–242.

LaPolla, J.S., Dlussky, G.M., and Perrichot, V. 2013. Ants and the fossil record. *Annual Review of Entomology* 58: 609–630.

Penney, D. 2002. Paleoecology of Dominican amber preservation—spider (Araneae) inclusions demonstrate a bias for active, trunk-dwelling faunas. *Paleobiology* 28: 389–398.

Penney, D. 2005. Importance of Dominican Republic Amber for determining taxonomic bias of fossil resin preservation—a case study of spiders. *Palaeogeography, Palaeoclimatology, Palaeoecology* 223: 1–8.

Penney, D. (ed.) 2010. *Biodiversity of Fossils in Amber from the Major World Deposits*. 304 pp. Siri Scientific Press, Manchester.

Penney, D. and Jepson, J.E. 2014. *Fossil Insects: An Introduction to Palaeoentomology*. 224 pp. Siri Scientific Press, Manchester.

Penney, D. and Langan, A.M. 2006. Comparing amber fossils across the Cenozoic. *Biology Letters* 2: 266–270.

Perkovsky, E.E., Rasnitsyn, A.P., Vlaskin, A.P., and Taraschuk, M.V. 2007. A comparative analysis of the Baltic and Rovno amber arthropod faunas: Representative samples. *African Invertebrates* 47: 229–245.

Radchenko, A. and Dlussky, G.M. 2013. *Bilobomyrma* new genus, a new extinct ant genus (Hymenoptera, Formicidae) from the late Eocene European ambers. *Journal of Paleontology* 87: 1060–1066.

Sontag, E. 2003. Animal inclusions in a sample of unselected Baltic amber. *Acta Zoologica Cracoviensia* 46 (Supplement: Fossil Insects): 431–440.

Szwedo, J. and Sontag, E. 2013. The flies (Diptera) say that amber from the Gulf of Gdańsk, Bitterfeld and Rovno is the same Baltic amber. *Polskie Pismo Entomologiczne* 82: 379–388.

Weitschat, W. and Wichard, W. 2010 Baltic amber. *In*: D. Penney (ed.), *Biodiversity of Fossils in Amber from the Major World Deposits*, 80–115. Siri Scientific Press, Manchester.

Wheeler, W.M. 1915. The ants of the Baltic amber. *Schriften der Physikalisch-Ökonomischen Gesellschaft zu Königsberg* 55: 1–142.

Wichard, W. and Grevin, H. 2009. Über Tausendfüßler, spanische Fliegen und Heuschrecken – zur "Historia Succinorum" des Nathanael Sendel von 1742. *Denisia* 26: 267–294.

Zherikhin, V.V. and Eskov, K.Y. 2006. On the real proportions of the main arthropod groups in the Baltic amber fauna, based on representative sampling. *Arthropoda Selecta* 15: 173–179.

David Penney [david.penney@manchester.ac.uk] and Richard F. Preziosi [richard.f.preziosi@manchester.ac.uk], Faculty of Life Sciences, the University of Manchester, Manchester, M13 9PT, UK.