The chemistry of American and African amber, copal, and resin from the genus *Hymenaea*

Victoria E. McCoy*, Arnoud Boomb, Mónica M. Solórzano Kraemerc, Sarah E. Gabbotta

a Department of Geology, University of Leicester, University Road, Leicester, LE1 7RH, UK
b Department of Geography, University of Leicester, University Road, Leicester, LE1 7RH, UK
c Senckenberg Forschungsinstitut und Naturmuseum, Paleontology and Historical Geology, Senckenberganlage 25 60325 Frankfurt am Main, Germany

*Corresponding author: vem10@le.ac.uk, +44(0)116-252-3315

Abstract

The comparison of the chemical composition of fossilized amber, copal, and resin is important for determining the botanic origin and original chemical composition of fossilized amber and copal, and for understanding the ecologic role of resin. Here we use solid phase microextraction-gas chromatography-mass spectrometry (SPME-GC-MS) to investigate the volatile and semi-volatile composition of amber, copal and resin from Africa and the Americas, produced by trees from the genus *Hymenaea*. We found there are four subgroups of *Hymenaea* resin, copal, and amber, based upon age and chemical similarity: African amber, American amber, African resin/copal (which also includes Colombian copal), and American resin/copal. This analysis allows us to narrow down the potential botanic origin of amber and copal samples, and also indicates that within this genus, resin similarity does not correspond closely with phylogenetic relationships. Therefore, resin chemistry may have been controlled by ecologic pressures, such as defence against herbivores, wood borers,
humidity, diseases, etc. and the original chemical composition of amber and copal could potentially be used to understand the role of resin in plant-insect interactions through time.

**Keywords:** Solid phase microextraction-gas chromatography-mass spectrometry, *Hymenaea*, Resin, Copal, Amber
1. Introduction

Trees in the genus Hymenaea (Fabaceae (=Leguminosae): Detarieae) (Fig. 1) have been prolific and important resin producers in Africa and the Americas for millions of years, producing extensive palaeontologically significant deposits of copal (semi-fossilized tree resin) and amber (fossilized tree resin) many of which contain exceptionally preserved fossils. The wealth of fossil and subfossil Hymenaea amber and copal is due primarily to the polylabdanoid macromolecular structure of the Hymenaea resin – which is characterized by ozic acid, ozol and enantio bifromenes (Anderson et al., 1992) – that allows for rapid polymerization into amber and copal that is very durable over geologic time (Langenheim, 1995). We have investigated the volatile and semi-volatile chemical composition of known Hymenaea amber, copal and resin from a variety of sites, all described in more detail below.

The two most celebrated and best-studied deposits are the Miocene Dominican and Mexican Chiapas ambers (Penney, 2010; Solórzano Kraemer, 2010). Both contain diverse and abundant assemblages of exceptionally-preserved biological inclusions, providing a wealth of palaeobiological information about the small, soft-bodied, terrestrial fauna in a Miocene-age Neotropical forest ecosystem (Arillo and Ortuño, 2005; Penney, 2010; Solórzano Kraemer, 2010).

Dominican amber in particular is notable for its unusually high-quality preservation (Grimaldi et al., 1994; Grimaldi and Engel, 2005; Penney, 2010; McCoy et al., 2017). For example, of the approximately five percent of the amber pieces which contain biological inclusions (Lambert et al., 1985), 93 percent of these, when examined with tomography or dissection have internal soft tissues preserved (Grimaldi et al., 1994; Stankiewicz et al., 1998;...
Heethoff et al., 2009; Greco et al., 2011; Van et al., 2014; McCoy et al., 2017), including
details such as muscle fibres, myofibrils, and mitochondria (Henwood, 1992; Grimaldi et al.,
1994; Labandeira, 2014). Chemical analyses (FTIR, C\textsuperscript{13}NMR, pyrolysis GC-MS) have
highlighted similarities between Dominican amber and modern *Hymenaea* resin
(Langenheim, 1969; Cunningham et al., 1983; Anderson et al., 1992; Langenheim, 1995;
Martínez-Richa et al., 2000; Penney, 2010), and palaeobotanical investigations indicate that
Dominican amber was produced by the extinct tree *H. protera* (Hueber and Langenheim,
1986; Poinar, 1991; Langenheim, 1995; Penney, 2010), a close relative of the extant *H.
verrucosa* (Poinar, 1991). Amber attributed to *H. protera*, is also known from Cuba, Haiti,
Puerto Rico and Jamaica (Iturralde-Vinent, 2001), but with such low abundance that it has
never been studied in detail.

Mexican Chiapas amber, like Dominican amber, has exceptional soft tissue
preservation of biological inclusions (Solórzano Kraemer, 2007, 2010) such that 55 percent
of studied specimens have internal soft tissues (McCoy et al., 2017). Mexican amber is
approximately the same age as Dominican amber, and is chemically very similar, but in some
deposits has undergone more extensive thermal degradation (Bryant, 1983; Lambert and
Poinar, 2002; Solórzano Kraemer, 2007, 2010). Chemical analyses (IR and C\textsuperscript{13}NMR) of
Mexican amber from Chiapas shows many similarities with Dominican amber and modern
*Hymenaea* resins, but, in addition, subtle differences indicate that the Chiapas amber was
produced by one or more extinct tree species in the genus *Hymenaea* (Langenheim, 1966;
Lambert et al., 1989; Lambert and Poinar, 2002), most likely *H. mexicana* (Brown, 2002) and
*H. allendis* (Calvillo-Canadell et al., 2010).
Another amber produced by *Hymenaea* is the less well known and recently described Ethiopian amber. Schmidt et al. (2010) described abundant well preserved arthropod inclusions and considered it as early Late Cretaceous in age (Cenomanian, ~93–95 Ma) although they did not comment on the botanic origin of the amber. Later Perrichot et al. (2016) re-evaluated the site based on additional amber material and associated sediment, which provided compelling evidence that Ethiopian amber is Cenozoic, likely Miocene. Moreover, these additional samples indicated that the amber was produced by the genus *Hymenaea*, similar to East African copals and Miocene ambers from Mexico or Dominican Republic.

In addition to the localities above we have also analysed samples from a recently discovered site in Venezuela which is of late Early Miocene to early Middle Miocene (Pérez et al., 2016) in age. The botanic origin of this amber has never been studied. In contrast to Dominican and Chiapas amber, no biological inclusions have been identified in Venezuelan amber (Pérez et al., 2016), and its botanic origin has not been investigated.

Copal deposits from, for example, Madagascar, Brazil, the Dominican Republic, Puerto Rico, and Colombia also offer abundant exceptionally preserved inclusions (Penney and Preziosi, 2010). The assemblages in copal have received little attention from palaeontologists because they are so young, but they are still an important resource for understanding the current biodiversity crisis and biases of preservation in amber (Penney and Preziosi, 2010). These copal samples are assumed to be produced by a species of *Hymenaea* based on chemical analyses and considerations of the major resin-producing trees in these regions (Schlüter and Von Gnielinski, 1986; Fearnside, 1989; Poinar, 1992; Langenheim, 1995; Clifford et al., 1997; Martinez-Richa et al., 2000; Lambert et al., 2002, 2005, 2014).
Hymenaea currently comprises 15 species, 14 of which are well distributed in the tropical and subtropical forests from Central America to Brazil and the West Indies. *Hymenaea verrucosa* is the sole species found in East Africa and Madagascar and is considered the most primitive species of the genus (Lee and Langenheim, 1975; Langenheim, 2003; Fougère-Danezan et al., 2010). However, a complete phylogenetic and molecular study of all 15 species has not yet been carried out.

Thus far, chemical analyses have been very successful at constraining the botanic origin of the amber and copal samples to the genus *Hymenaea*, but they have been less successful at constraining it further to the species level. For example, some chemical analyses find Dominican amber is very similar to *H. verrucosa* resin (Cunningham et al., 1983), although most highlight similarities with *H. courbaril* resin (Langenheim, 1969; Lambert et al., 1985, 2008, 2015). Others studies find very little chemical difference between any *Hymenaea* species, both extant and extinct (Lambert et al., 2014). From the chemical analyses alone, it is not clear that Dominican amber is produced by an extinct tree; chemical differences with modern resin could be due to amberization (the chemical changes during fossilization to transform resin into copal and then amber) or intraspecific resin variability rather than to a different botanic origin. Determining the botanic origin of copal or amber using chemical comparisons to modern resin requires either that the resin-producing species is still extant (and included in the analysis), or that chemical similarity indicates phylogenetic similarity. However, the chemical analyses to date do not consistently indicate whether chemical similarity indicates a close phylogenetic relationship between the botanic producers. Some analyses found that chemical similarity follows broad phylogenetic patterns (Lambert et al., 2005), but others suggest that resin chemistry is more strongly controlled by
environmental factors, and that resin chemical variation is more likely to be functionally
rather than phylogenetically controlled (Langenheim, 1995).

The ecologic role of resin is not fully understood, but it includes defence against
insect herbivores, healing wounds (Langenheim, 1990, 1995, 2003; Pichersky and Raguso,
2016) and to prevent bacterial and fungal infections, or infestations by wood-boring
arthropods (McKellar et al., 2011; Beimforde et al., 2017).

Plant-insect interactions are a product of hundreds of millions of years of antagonisms
and co-evolution, and include some of the most complicated and important interactions in
modern ecosystems (Bryant et al., 1991; Labandeira et al., 1994; Howe and Jander, 2008;
War et al., 2012; Labandeira and Currano, 2013). By characterizing the chemical
components of amber, resin, and copal, and by precisely identifying the botanic producer, we
can better understand these interactions through time from the modern day and in the fossil
record. This is particularly interesting for amber fossil sites which preserve much of the
original insect herbivore fauna as inclusions in amber (Penney, 2010; Solórzano Kraemer,
2010; Labandeira, 2014; Peris et al., 2015), and also preserve some direct evidence of plant-
insect interactions in the form of leaves with herbivore damage (Labandeira, 2014).

The goal of this research is to use headspace solid phase microextraction-gas
chromatography-mass spectrometry (SPME-GC-MS) to elucidate the volatile and semi-
volatile components of various African and American resin, copal, and amber known to be
produced by the genus *Hymenaea*. Headspace SPME uses a coated fibre to extract
compounds from the headspace of a sealed vial containing a sample and transfer them to a
GC-MS for identification and quantification (Pawliszyn, 2011). This method has previously
been used to differentiate samples of modern resin (Hamm et al., 2003, 2005), identify small
amounts of resin in archaeological samples (Hamm et al., 2004), characterize the volatile components in Benzoin gum (Castel et al., 2006), identify two volatile degradation compounds of Baltic amber (Pastorelli, 2011), and differentiate Baltic and Romanian amber (van der Werf et al., 2014). These previous studies have also involved extensive methods testing (Hamm et al., 2003), which has informed our selection of SPME fibre, incubation temperature and time, and sampling time. Our analyses on Hymenaea resin, copal, and amber will provide a simplified, comparable chemical characterization of resin, copal, and amber that focuses on some of the most ecologically active compounds (Langenheim, 2003). This new source of chemical data from these samples will also complement previous analyses, to help better elucidate the botanic origin and ecologic role of these samples.

2. Materials and Methods

2.1. Samples

Samples of amber, copal, and resin were obtained from the Senckenberg Research Institute and Natural History Museum (SMF), from Alcaldía Municipio Urumaco, Colección Paleobotánica, Venezuela (AMU-PB) or collected from various locations around Africa and the Americas (table 1, Fig. 1): two samples of Ethiopian amber; two samples of Mexican Chiapas amber, one each from the Simojovel and Totolapa mines; one sample of Dominican amber; one sample comprised of various small pieces of Venezuelan amber, one sample of copal from the Dominican Republic; two samples of copal from Colombia; one sample of copal from either Puerto Rico or the Dominican Republic; nine samples of Hymenaea verrucosa resin, six of which were from the same tree; one sample of H. courbaril resin; one
sample of H. parvifolia resin; and two unknown samples, one from the collections of the Senckenberg Research Institute and Natural History Museum and one purchased from ebay under a listing for ‘Dominican amber.’ The pieces were selected to be homogenous and free of organic inclusions. Each sample was reduced to a fine powder using a ball mill, and 0.5 grams were sealed into a 20 ml headspace vial with a PTFE septa and magnetic screw top caps. Our sample sizes (0.5 g) are much larger than those used in previous SPME analyses of resin, copal, and amber, which range from 0.002 to 0.04 mg (Hamm et al., 2003; Hamm et al., 2004; Pastorelli, 2011). However, for headspace SPME, sample size is dependent upon headspace volume; we used 20 ml vials rather than the 2 ml vials used in the previous analyses (Hamm et al., 2003; Hamm et al., 2004; Pastorelli, 2011) and therefore increased our sample size accordingly. Moreover, we tested our method with various sample sizes of Copaifera officinalis resin purchased from an online supplier and found the results were essentially identical for samples sizes ranging from 0.01 g to 1 g (Supplementary Figs. 1 and 2).

2.2 Headspace SPME-GC-MS

The vials containing the powdered samples were randomly loaded in a Triplus RSH autosampler and placed in an agitator where they were equilibrated at 80°C for one hour. The SPME fibre, a 65 μm Polydimethylsiloxane/Divinylbenzene (PDMS/DVB) fibre initially conditioned at 250°C for 30 minutes per the manufacturer’s instructions, was automatically introduced and exposed to the head space for one hour. After sampling, the fibre was inserted into the injection port of a Thermo Scientific Trace 1310 GC, which had a Thermo TG-5MS
30 m column with 0.25 mm ID and 0.25 μm film thickness, coupled to an ISQ QD single quadrupole mass spectrometer, where it was desorbed for 2 minutes at 250°C. Splitless injection (1 minute) was used. The injector temperature and transfer line temperature were 250°C, and the GC program had an initial temperature of 40°C, held for five minutes, that ramped up to 280°C at a rate of 10°C per minute, where it was held for 5 minutes. A liquid injection of a standard mixture containing a series of \( n \)-alkanes was used to calibrate retention indices to aid in identifying the peaks.

### 2.3 Peak identification

The chromatograms for each sample were imported into the program AMDIS, which automatically deconvolutes the data to extract the pure component spectra, allowing for more accurate identification. The major peaks in each chromatogram were identified through a National Institute of Standards and Technology (NIST) MS database search, including information from both the MS fragmentation patterns and the retention indices. We found 126 compounds, which we compiled into a search library in AMDIS. We then used the analysis function in AMDIS to automatically compare all peaks in each chromatogram to this search library, so that we had comparable data for each sample.

### 2.4. Semi-quantitative analysis

The relative amounts of each compound were calculated as the percent of the total peak area of the 126 selected compounds, and these data were analysed with principal components analysis (PCA) (following van der Werf et al., 2014) using the program R. These data quantitatively represent the SPME chromatograms, and therefore provide a way to
quantitatively compare the chromatograms. However, the SPME chromatograms do not
quantitatively represent the samples, because SPME extracts different compounds with
differing levels of completeness (Hamm et al., 2003). Nonetheless, the results for identical
samples are completely reproducible because the use of an autosampler results in exactly
comparable timings and temperatures; these are two factors that play a major role in the
efficiency of the equilibrium-based extraction. Therefore, these analyses allow reproducible,
semi-quantitative comparisons of the samples (van der Werf et al., 2014).

3. Results

Components 1 and 2 of the PCA (which encompass ~ 49% of the variation in the
dataset (Fig. 2)) indicate that the volatile and semi-volatile chemical constituents of resin,
copal, and amber vary based on age and location (Fig. 3). Components 3-6, although they
together encompass another 31% of the variation, do not separate the samples in any
meaningful way (Supplementary Fig. 3) and so are not considered. Component 1 (~38% of
variation (Fig. 2)) separates the ancient amber from the recent resin/copal (Fig. 3).
Component 2 (~11% of variation (Fig. 2)) separates the African samples from the American
samples (Fig. 3). The exception to this is Colombian copal, both samples of which group
with the African resins, rather than with the other American samples (Fig. 3). Within these
groups, we see large variation within species, and overlap between species. The nine samples
of *H. verrucosa* resin all cluster within the same group (the African resin/copal group) in the
PCA, but they span the entire range of variation of that group, overlapping with the
Colombian copal samples, which were almost certainly produced by one of the American
*Hymenaea* species, rather than *H. verrucosa* which is restricted to Africa (Fig. 3). Multiple
samples from one *H. verrucosa* tree have a restricted chemical composition relative to the entire range of *H. verrucosa*, but they still show some variation (Fig. 3). Similarly, the two samples of Mexican amber from Chiapas (both produced by *H. mexicana* or *H. allendis*) fall within the American amber group in the PCA, but span the entire range of variation of the group and overlap with the Dominican amber sample (produced by *H. protera*).

The four groups in the PCA (Fig. 3) are primarily determined on the basis of 12 of the 126 chemical compounds (Fig. 4). All of the ambers (in contrast to the resin and copal samples) have high amounts of 1,1,4,5,6-pentamethyl-2,3-dihydro-1H-indene, 4,8,11,11-tetramethyl-tricyclo[7.2.0.0(3,8)]undec-4-ene, trimethylphenyl- butanone, caryophyllene isomer, and tetrahydro-tetramethyl-naphthalene (Figs. 4 and 5). American amber (see Supplementary Fig. 4 for chromatograms) is distinguished from African amber (see Supplementary Fig. 5 for chromatograms) by the relative proportions of these compounds (Figs 4 and 5): American amber is dominated by 1,1,4,5,6-Pentamethyl-2,3-dihydro-1H-indene and African amber by the other four compounds. All of the resin and copal samples have high abundances of 13-epimanool, caryophyllene, biformalene, and α-curcumene.

However, American resin/copal (Figs. 4 and 5, see Supplementary Fig. 6 for chromatograms) also has high amounts of humulene-1,2-epoxide, α-humulene, and β-bisabolene. Colombian copal (see Supplementary Fig. 7 for chromatograms) has relatively small amounts of these three compounds (Fig. 6) and so it clusters with the African resin/copal (see Supplementary Fig. 8 for chromatograms) in the PCA (Fig. 3).

The analysis also includes two samples of unknown geographic and botanic origin (see Supplementary Fig. 9 for chromatograms). The first unknown was suspected to be *H. courbaril* resin from Mexico or Colombia. In our analysis, this sample (labelled ‘unknown’)
groups with the recent American resin/copal samples but not with the Colombian copal samples, suggesting it was not from Colombia (Fig. 3). However, there is not sufficient differentiation between different American resins and copals to discriminate any further the exact provenance of this sample. The second unknown sample was purchased from ebay and was sold as Dominican Amber (labelled ‘Dominican amber?’) (Fig. 3). Our analysis strongly suggests that this sample is from the Americas, however, it is more likely to be resin or copal than amber.

4. Discussion

Our results support four conclusions regarding the volatile/semi-volatile composition of known *Hymenaea* resin/copal/amber: (i) there is extensive intraspecific variation and interspecific overlap in resin chemistry; (ii) chemical similarity does not correspond to phylogenetic similarity at the species level, which has implications for using resin chemistry to determine the botanic origin of amber, copal, or an unknown resin sample; (iii) environmental factors might be important for controlling chemical composition; and (iv) resin chemistry changes during the process of amberization, but despite this it is may still be possible to gain some understanding of the original resin chemistry of an amber sample.

4.1. Intra- and inter-specific variation

Previous analyses of *Hymenaea* resin chemistry have encompassed both leaf/primary stem resin (Martin et al., 1971, 1974; Langenheim et al. 1978) and trunk resin (Lambert et al., 1985, 1989; Cunningham et al., 1983; the current study). These two sources of resin vary because trunk resin has a polymeric macromolecular structure as well as volatile and semi-
volatile components whereas leaf/primary stem resin lacks the polymeric structure (Langenheim, 1995).

Previous studies of the macromolecular chemical composition of *Hymenaea* resin using C\(^{13}\) NMR and infrared (IR) spectroscopy have highlighted intraspecific resin variation (Martin et al., 1971, 1974; Langenheim et al., 1978). These studies primarily focused on *H. courbaril*, which is a highly variable species including multiple subspecies, and which may actually represent up to three species (Souza et al., 2014). Our analyses focused on the variation within *H. verrucosa* resin, and yet still found extensive chemical variation (Fig. 3), suggesting that, regardless of whether a consensus phylogenetic grouping for *H. courbaril* is met, variable resin chemistry does occur within species and individuals of *Hymenaea*.

Analyses of Dominican (Lambert et al., 1985) and Mexican amber (Lambert et al., 1989) also show significant chemical variation, suggesting intraspecific chemical variation within *H. protera* and *H. mexicana/H. allendis* resin.

Previous studies comparing the resin chemistry of different species of *Hymenaea* have reached mixed conclusions. The most common comparison is between *H. verrucosa* and *H. courbaril*, which typically can be distinguished using C\(^{13}\) NMR and IR spectroscopy, which elucidates the macromolecular structure (Cunningham et al., 1983). In contrast, analyses of the macromolecular structure of resin from multiple species of *Hymenaea* (including both *H. verrucosa* and *H. courbaril*), using C\(^{13}\), H\(^{1}\) and COSY NMR and IR find that they are chemically very similar (Martin et al., 1976; Lambert et al., 2014).

Our results suggest that the volatile and semi-volatile compositions of *Hymenaea* resins follow the same broad patterns as the macromolecular structures of these resins: *H. verrucosa* and *H. courbaril* have very different volatile and semi-volatile compositions; but
there is also significant overlap between different *Hymenaea* resins. *H. courbaril* and *H. parvifolia* (as well as the tree(s) which produced most of our American copal samples) cannot be distinguished, and *H. verrucosa* and the Colombian copal tree cannot be distinguished.

4.2. Phylogenetic similarity and botanic origin

Large scale chemical analyses suggest that resin chemical groups do follow broad taxonomic patterns: resin chemistry can often be used to identify families, sometimes to identify genera, but rarely to identify species (Lambert et al., 2005; Sonibare et al., 2012). As such, chemical analyses are useful for identifying the botanic origin of an unknown sample (including amber samples) to higher taxonomic levels, but species-level identification of an amber-producing tree requires palaeobotanical investigation to supplement the chemical analyses.

Our analyses, which are restricted to one genus, cannot address the utility of SPME-GC-MS for family or genus level identification of the botanic source of an unknown sample. We have found that resin chemical similarity for species within the genus *Hymenaea* does not correlate to close phylogenetic similarity, and therefore, as with other analyses, is not sufficient for species-level identification of the botanic origin of a fossil sample or an unknown recent sample. As previous chemical analyses have suggested (Lambert et al., 2015), the volatile and semi-volatile composition of Dominican amber is more similar to American resins and copal, such as *H. courbaril* resin, than to the more closely related *H. verrucosa* resin. Both samples of Colombian copal (which were almost certainly produced by one of the American *Hymenaea* species) fall within the range of variation of the less closely related African *H. verrucosa* resin. However, Martínez-Richa et al. (2000) have also
previously noted that Colombian samples were very similar to African samples. Moreover, some modern species (\textit{H. courbaril} and \textit{H. parvifolia}; and \textit{H. verrucosa} and the Colombian copal tree) overlap in the SPME-GC-MS PCA and therefore cannot be distinguished at all. These analyses can be used to rule out potential botanic producers of an unknown sample (for example if it clusters with the American resin/copal samples it was not produced by \textit{H. verrucosa}), but cannot be used to identify it definitively to the species level.

The sample of amber from Venezuela clusters nicely with the Chiapas and Dominican amber, confirming that the botanic origin of this newly discovered amber is very likely also a species of \textit{Hymenaea}. An alternative hypothesis is that the Venezuelan amber samples are produced by a species of \textit{Copaifera}. This genus also includes prolific resin producing trees, has a very similar distribution as \textit{Hymenaea} in the Americas, and \textit{Copaifera} resin has previously been very difficult to distinguish from \textit{Hymenaea} resin using $^{13}$C, $^1$H, and COSY NMR spectroscopy (Lambert et al., 2009; Lambert et al., 2014). However, all the \textit{Hymenaea} samples in this analysis, and the Venezuelan amber samples, were clearly distinct from the \textit{Copaifera officinalis} samples used for methods testing (Supplementary Fig. 2), suggesting that the Venezuelan amber samples were more likely produced by a species of \textit{Hymenaea}.

4.3. Environmental factors

SPME-GC-MS analysis separates known \textit{Hymenaea} resin/copal/amber into four distinct subgroups on the basis of chemical similarity. As discussed above, these groups are not based upon phylogenetic similarity, which suggests chemical variation in \textit{Hymenaea} resin is more strongly influenced by environmental variation (e.g. biotic factors such as herbivore pressures and abiotic factors such as temperature and aridity) than by phylogenetic
constraints. Most of the research on the ecologic role of resin (including for *Hymenaea* resin) focuses on leaf resins, which may have very different composition than the trunk resins, even from the same tree (Langenheim, 1995, 2003). However, some of the general conclusions about specific compounds are still applicable to trunk resin. The non-volatile compounds (which are not considered in this analysis) generally affect the viscosity and the polymerization of the resin, and provide physical defences such as trapping attackers and coating and sealing wounds (Langenheim, 2003; Martínez-Delclòs et al., 2004). The volatile compounds (which include those measured in this analysis) generally provide chemical defences (Langenheim, 2003). Some are directly toxic to herbivores or fungi (Langenheim et al., 1980; Arrhenius and Langenheim, 1983; Welker et al., 2007), and others attract predators or parasites of attacking herbivores (Dicke et al., 1990; Langenheim, 1994). Compositional variation in *Hymenaea* resin has been linked to selection in response to the types and quantities of attacking pests (Langenheim, 2003). However, the efficacy of a resin chemical compound against a specific attacking organism varies based on abiotic environmental factors suggesting that abiotic factors may have an indirect effect on resin chemical composition (Langenheim, 1995).

The two resin/copal subgroups defined in our study (the American resin group and the African resin group, Fig. 4) therefore most likely indicate two distinct biotic environmental pressure regimes, including attacks on the trees by herbivores, wood-infesting arthropods, pathogens, and fungi. Most of the seven compounds that distinguish these two groups have been linked to defensive functions, although they have only been investigated in a few tree species and against a few types of attackers (Table 3). African resins are characterized by four chemical compounds, of which three have been subject to an investigation of their
13-epimanool is associated with resistance to vole browsing in the bark of larch trees (Sato et al., 2009; Seki et al., 2012); caryophyllene in pine trees has been shown to inhibit complete needle destruction by caterpillars (although it is also associated with a higher frequency of caterpillar attacks) (Petrakis et al., 2005), and to discourage attacks by ants and fungi (Barnola et al., 1997), in *Hymenaea* to discourage ant, and caterpillar attacks (Langenheim et al., 1980; Hubbell et al., 1983), and in Dipterocarps to discourage termite attacks (Messer et al., 1990); α-curcumene has been found to repel whiteflies in tomatoes (Bleeker et al., 2011). Finally, the effects of biformene on attackers has not been investigated. The American resins and copals are defined on three compounds, two of which have been studied: α-humulene discourages termites in dipterocarps (Messer et al., 1990) and insect herbivores in *Hymenaea* (Langenheim et al., 1980, 1986); humulene-1,2-epoxide deters caterpillar herbivores in *Hymenaea* (Langenheim et al., 1980); and β-bisabolene has not been studied. More research on the effects of these chemicals against a wider range of attackers, and which organisms typically attack the different *Hymenaea* species, is necessary to determine what selective pressures promote the production of one chemical compound over another, and therefore influence the chemical composition of *Hymenaea* resin. However, it is interesting to note that the African resins have compounds that deter mammalian attackers and fungi as well as arthropods. Moreover, in many analyses, caryophyllene (characterizing the African resins and copals) is identified as one of the most prominent and effective anti-herbivore defence chemicals (Langenheim et al., 1980, 1986; Welker et al., 2007). It may be that the African (and Colombian) species of *Hymenaea* are subject to attack by a more diverse and persistent fauna than the American species.
4.4. *Amberization and original volatiles*

As resin fossilizes into copal or amber (‘amberization’), it undergoes a complex process of maturation including oxidation, oligomerization, and cross linking that changes the molecular structure (Grimalt et al., 1988; Anderson and Winans, 1991; Anderson et al., 1992; Anderson and Crelling, 1995; Tonidandel et al., 2008). The SPME method used in this research only captures small molecular weight compounds and so we observe two particular consequences of amberization: a decrease in the original low molecular weight volatile and semi-volatile compounds (Tonidandel et al., 2008), and an increase in low molecular weight degradative compounds (Pastorelli, 2011). Although we did observe fewer peaks (which in this analysis all represent low molecular weight volatile and semi-volatile compounds) in the amber chromatograms than in the resin and copal chromatograms (Fig. 5), the amber was separated from the resin/copal in the PCA primarily by the presence of degradative compounds, which have aromatic rings and only very short side chains (Fig. 4). These variations allow us to distinguish amber samples from more recent resin/copal samples, which is useful for determining if unidentified samples are amber or not (Fig. 3). However, this analysis is not very precise, and could not be used to get relative ages for two samples unless they are very different in age: e.g. Chiapas amber and Dominican amber are very close in age (Penney, 2010; Solórzano Kraemer, 2010) but cover a wide range in component 1 in the PCA (Fig. 3), and we cannot distinguish between resins and copals (Fig. 3). Previous attempts to use NMR, FT-Raman spectroscopy, thermogravimetric analyses, and atmospheric pressure photoionization (Brody et al., 2001; Ragazzi et al., 2003; Kimura et al., 2006; Tonidandel et al., 2008; Lambert et al., 2015) have been more effective at determining the
age of a resin, copal or amber sample; for more recent samples, $^{14}$C dating can also be
effective (Burleigh and Whalley, 1983).

In order to understand the role of resin in plant insect interactions in the fossil record,
it is necessary to know the chemical composition of the original resin, rather than the
fossilized amber, which is often obscured by the amberization process. Based on our
analyses, the amber does retain some of the original volatile compounds (e.g. Ethiopian
amber 2 contains $\alpha$-pinene). However, we can also make assumptions about the original
volatile and semi-volatile composition based on similarities to modern resins. For example,
the Dominican and Chiapas amber samples clustered with the American resins/copals, and
were therefore most likely originally characterized by 13-epimanool, $\alpha$-curcumene,
biformene, and caryophyllene. In contrast, the Ethiopian amber is most similar to the African
resin and copal group, and therefore may have been originally characterized by $\alpha$-humulene,
humulene-1,2-epoxide, and $\beta$-bisabolene isomer.

5. Conclusions

SPME-GC-MS can distinguish four chemical subgroups within *Hymenaea* resin,
copal and amber: American amber; African amber; American resin/copal; and African
resin/copal (which also includes Colombian resin). Both amber groups are defined based on
degradative chemical compounds produced during the process of amberization, and the
differences between the groups can be explained by different original chemical compounds
(which are then influenced by amberization). The resin/copal groups are defined on the basis
of original volatile and semi-volatile compounds that all play a role in defence against
herbivores, fungi, and pathogens. Variations in the chemical composition of different
Hymenaea species do not follow phylogenetic patterns, and are most likely due to selection pressures from different herbivore fungi, and pathogen assemblages. More research is required to determine which herbivores, fungi or pathogens exert most selective pressure on the chemical composition of resin. However, the current knowledge of the defensive role of the key volatile compounds suggests that the African resin/copal group is defined by more effective and more broadly applicable defensive chemicals, and therefore they may need to defend against a more diverse fauna. The chemical similarity between the amber and resin, in combination with some remnants of original volatile compounds, may help infer the original volatile and semi-volatile composition of the amber samples. This, in combination with more research on the defensive role of specific resin chemicals and the preserved arthropod herbivore fauna in the amber fossil assemblages, may provide insights on the role of resin in plant-insect interactions through geologic time.

6. Acknowledgements

Thanks to Armando Arroyo Gomez for H. parvifolia sample Jorge Domingo Carrillo Briceño, and Marcelo Sánchez-Villagra (Paleontological Institute and Museum, Switzerland) and Carlos Jaramillo (Smithsonian Tropical Research Institute) for Venezuela amber samples and Instituto del Patrimonio Cultural de Venezuela to allow the collection of them, Ravaka Ravelomanana for the collection of some Madagascar resin, Claudia Franz (Senckenberg Research Institute, Germany) for preparing samples of copal and amber, and Enrique Pañalver (Instituto Geológico y Minero de España, Spain) and Xavier Delclòs (Barcelona University, Spain) for important discussion. This work was supported by a Royal Society Newton International Fellowship to VEM, and AMBERIA 2015-2017 [grant number
CGL2014-52163] and VolkswagenStiftung [grant number 90 946] to MSK. The manuscript was also improved by the comments of two anonymous reviewers.
Figure Captions

Figure 1: Resin *in situ* on a *Hymenaea verrucosa* tree in the northwest of Madagascar.

Figure 2: The percent of variation that is controlled by each component of the PCA. Components 1 and 2 control ~ 49% of the variation.

Figure 3: PCA of all samples of resin, copal, amber, including two unknowns. Red squares are amber samples, orange circles are copal samples, yellow stars (and the yellow polygon) are resin samples, and black plus signs are unknown samples. Component 1 separates the ancient amber samples (on the negative side of dimension 1) from the recent resin and copal samples (positive side of dimension 1). Component 2 separates the American samples (positive side) from the African samples (negative side), with the exception of Colombian copal which clusters with the African resins. Each of the four groups is delineated by a convex hull polygon.

Figure 4: Variable loadings for components 1 and 2 of the PCA. The 12 specifically identified variables (chemical compounds) are most important for defining components 1 and 2; the other 114 variables cluster near the origin. The one variable identified with a grey star defines the American amber group, the four variables identified with green squares define the African amber group, the four variables identified with blue circles define the African resin/copal group, and the three variables identified with purple hexagons define the American amber/copal group. Key variable groupings are also indicated by convex hull polygons.

Figure 5: Representative chromatograms from the African amber group (A,B), the African resin/copal group (C,D), the American amber group (E,F) and the American resin/copal group (G,H). (A,C,E,G) Full chromatographs. (B,D,F,H) Proportionate peak area for selected
compounds that are most important for defining the groups; colours and shapes are as in Fig. 4. Peaks of selected compounds are labelled to identify the compounds, and correspond to the numbers in Fig. 4: 1 is 1,1,4,5,6-pentamethyl-2,3-dihydro-1H-indene; 2 is 4,8,11,11-tetramethyl-tricyclo[7.2.0.0(3,8)]undec-4-ene; 3 is trimethylphenyl- butanone; 4 is caryophyllene isomer; 5 is tetrahydro-tetramethyl-naphthalene; 6 is 13-epimanool; 7 is caryophyllene; 8 is biformene; 9 is α-curcumene.; 10 is α-humulene-1,2-epoxide; 11 is humulene; and 12 is β-bisabolene isomer.

Figure 6: Colombian copal samples illustrated as in figure 5. Full chromatograms (A,C), and proportionate peak area for selected compounds that are most important for defining the groups (B,D); colours and shapes are as in Fig. 4. Peaks of selected compounds are labelled to identify the compounds, and correspond to the numbers in Fig. 4: 1 is 1,1,4,5,6-Pentamethyl-2,3-dihydro-1H-indene; 2 is 4,8,11,11-tetramethyl-tricyclo[7.2.0.0(3,8)]undec-4-ene; 3 is trimethylphenyl- butanone; 4 is caryophyllene isomer; 5 is tetrahydro-tetramethyl-naphthalene; 6 is 13-epimanool; 7 is caryophyllene; 8 is biformene; 9 is α-curcumene.; 10 is α-humulene-1,2-epoxide; 11 is β-bisabolene isomer. Notice how the highest peaks of the selected compounds (B, D) correspond to those that are enriched in African resins.
Figures

Figure 1
Figure 2
Figure 4

1. 1,1,4,5,6-Pentamethyl-2,3-dihydro-1H-indene
2. 4,8,11,11-Tetramethyl-tricyclo[7.2.0.0(3,8)]undec-4-ene
3.Trimethylphenyl-butanone
4. Caryophyllene isomer
5. Tetrahydro-tetramethyl-naphthalene
6. 13-Epimanool
7. Caryophyllene
8. Biformene
9. α-Curcumene
10. Humulene-1,2-epoxide
11. α-Humulene
12. β-Bisabolene isomer

AMERICAN AMBER
AFRICAN AMBER
AFRICAN RESIN AND COPAL
AMERICAN RESIN AND COPAL
Figure 6

(A) Colombian copal 1

(B) Colombian copal 1

Compounds defining:
- African amber
- African resin
- American amber
- American resin

(C) Colombian copal 2

(D) Colombian copal 2

Time (minutes)

Intensity (x10^6)

Peak area
References

Chemical Society, Washington, D.C.

Anderson, K.B., Winans, R., Botto, R., 1992. The nature and fate of natural resins in the
geosphere—II. Identification, classification and nomenclature of resinites. Organic
Geochemistry 18, 829-841.

Evaluation of pyrolysis-gas chromatography mass spectrometry for the analysis of
natural resins and resinites. Analytical Chemistry 63, 2901-2908.

amber (Miocene). Stuttgarter Beitrage zur Naturkunde. Serie B (Geologie und
Palaontologie) 352, 1-68.

Arrhenius, S.P., Langenheim, J.H., 1983. Inhibitory effects of *Hymenaea* and *Copaifera* leaf
resins on the leaf fungus, *Pestalotia subcuticularis*. Biochemical Systematics and

contents in *Pinus caribaea* needles and its possible relationship to *Atta laevigata*

Resin exudation and resinicolous communities on *Araucaria humboldtensis* in New
Caledonia. Arthropod-Plant Interactions, 1-11.

Bleeker, P.M., Diergaarde, P.J., Ament, K., Schütz, S., Johne, B., Dijkink, J., Hiemstra, H.,
de Gelder, R., de Both, M.T., Sabelis, M.W., 2011. Tomato-produced 7-
epizingiberene and R-curcumene act as repellents to whiteflies. Phytochemistry 72, 68-73.


characterization of soluble labdanoid polymers, isolated from Holocene class I resins.


