Received: 17 February 2023 | Accepted: 16 October 2023

DOI: 10.1111/icad.12697

Insect Conservation and Diversity

ORIGINAL ARTICLE

Shedding light on dark taxa in sky-island Appalachian leaf litter: Assessing patterns of endemicity using large-scale, voucher-based barcoding

Michael S. Caterino | **Ernesto Recuero**

Department of Plant & Environmental Sciences, 277 Poole Agricultural Center, Clemson University, Clemson, South Carolina, USA

Correspondence
Michael S. Caterino, Department of Plant & Environmental Sciences, 277 Poole Agricultural Center, Clemson University, Clemson, SC 29634-0310, USA.
Email: mcateri@clemson.edu

Funding information
Clemson University Experiment Station; National Science Foundation

Editor: Raphael K. Didham and Associate Editor: Nusha Keyghobadi

Abstract

1. Developing systematic conservation plans depends on a wealth of information on a region's biodiversity. For 'dark taxa' such as arthropods, such data are usually very incomplete and in most cases left out from assessments.
2. Sky islands are important and often fragile biodiversity hotspots. Southern Appalachian high-elevation spruce–fir forests represent a particularly threatened sky-island ecosystem, hosting numerous endemic and threatened species, but their arthropods remain understudied.
3. Here we use voucher-based megabarcoding to explore genetic differentiation among leaf-litter arthropod communities of these highlands, and to examine the extent to which they represent dispersed communities of more or less coherent species, manageable as a distributed unit. We assembled a dataset comprising more than 6000 COI sequences representing diverse arthropod groups to assess species richness and sharing across peaks and ranges. Comparisons were standardised across taxa using automated species delimitation, measuring endemism levels by putative species.
4. Species richness was high, with sites hosting from 86 to 199 litter arthropod species (not including mites or myriapods). Community profiles suggest that around one fourth of these species are unique to single sky islands and more than one third of all species are limited to a particular range. Across major taxa, endemicity was lowest in Araneae, and highest in neglected groups like Isopoda, Pseudoscorpionida, Protura and Diplura.
5. Southern Appalachian sky islands of spruce–fir habitat host significantly distinct leaf-litter arthropod communities, with high levels of local endemicity. This is the first work to provide such a clear picture of peak and range uniqueness for a taxonomically broad sample. Ensuring the protection of a sizeable fraction of high-elevation litter species richness will therefore require attention at a relatively fine spatial scale.

KEYWORDS
arthropoda, cytochrome oxidase I, endemism, leaf litter, megabarcoding, sky islands

INTRODUCTION

Developing systematic conservation plans depends on a wealth of information on a region's biodiversity (Margules & Pressey, 2000; Nielson et al., 2022). Knowing what species occur where is a prerequisite for implementing any sort of broad systematic approach. But it is rare that such data cover a significant fraction of an area's actual biodiversity, particularly when 'dark taxa' (sensu Hausmann et al., 2020) are considered, with important implications for generalising about ecological processes (Kortmann et al., 2022) and diversity patterns (Hartop et al., 2022). Whether conservation management should focus on regional, landscape or local scales depends strongly on fine-grained knowledge of species distributions and patterns of endemism (Daru et al., 2020), as do assessments of extinction threats. But until such patterns can be confidently predicted to pertain to most of an area's biodiversity, conservation units may under- or overestimate the most relevant scale for protection and management. Bringing dark taxa, such as arthropods, into the picture would greatly scale up the data on which such assessments can be based. Here we explore the scale of endemism in litter-dwelling arthropods on the southern Appalachian 'sky islands'.

Sky islands are high-elevation habitats isolated in different mountains or mountain ranges, and fragmented by the dominating habitats present at lower elevations. They are particularly important hotspots of biodiversity, though they vary considerably in many specifics, including elevation, climatic gradients and degree of habitat distinctness from the surrounding lowlands. Perhaps the most widely recognised are the relatively mesic highlands of the desert south-west of North America, with oak–pine forest rising out of a sea of aridity, to elevations of 3000 meters or more (Heald, 1951). These have received considerable study as reservoirs of plant (e.g., Bowers & McLaughlin, 1996), herpetofaunal (e.g., Bezy & Cole, 2014) and arthropod (e.g., Maddison & McMahon, 2000; Masta, 2000; Monjaraz-Ruedas et al., 2023; Ober et al., 2011) diversity. The latter, however, still lag other taxa considerably (Meyer et al., 2015). Some work in other sky-island systems has begun to demonstrate the uniqueness of elements of their insect faunas, in the Ozark Highlands of central North America (Monroe et al., 2022), in the temperate sky islands of tropical México (Uscanga et al., 2021), in the Cameroonian highlands (Grebennikov, 2021) and in the southern Appalachians (Hedin et al., 2015). While endemic species in such systems are not difficult to find, important questions remain regarding the extent of endemism among arthropods. Particularly for more recently isolated sky-island communities, are isolated endemics the rule, are they the exceptions that happen to have caught researchers' attention, or something in between? Answers to these questions translate directly into conservation decision-making. From acquisition and assigning degree of protection, to on-the-ground habitat management strategies such as invasive species management, fire prevention (or augmentation) or revegetation, at what spatial scale these actions are targeted will depend on how locally to regionally distributed unique elements of biodiversity are.

In the southern Appalachian Mountains, the highest elevations, above about 1500 m (5000 ft), reveal a sky-island forest community dominated by the conifers red spruce (*Picea rubens*) and Fraser fir (*Abies fraseri*), very different from the predominantly deciduous broadleaf forests of lower elevations. The presence of these trees is permitted by cooler, moister environmental conditions than are found elsewhere in the south-eastern United States. During cooler Pleistocene times, similar coniferous forest was widespread in the south-eastern United States, occurring over a broad area surrounding and within the Appalachian Mountains. However, it started to retreat, northward and upward, around 15,000 YBP (Boehm, 2012; Watts, 1970; Whitehead, 1981), becoming progressively more fragmented, culminating in more or less completely isolated patches on the higher peaks of the southern Appalachians by 8-9000 YBP (Delcourt, 1985), covering a total of just over 18,000 hectares across seven discrete ranges (Figure 1). This system thus clearly merits the 'sky island' categorization, hosting a recent, but distinctive and highly fragmented biota.

A large variety of organisms are known to occur only within these southern spruce–fir communities. Fraser fir is considered a distinct, endemic species, having become isolated from the more boreal balsam fir, *Abies balsamea*, during late Pleistocene times (Potter et al., 2008). Other endemics include species of salamanders (Crespi et al., 2003; Crespi et al., 2010; Moskwik, 2014), small mammals (Sipe & Browne, 2004), lichens (Allen & Lendemer, 2016), bryophytes (Anderson & Zander, 1973), land snails, (Dourson & Langdon, 2012; Slapcinsky, 2018) and numerous arthropods (Barr, 1979; DeSisto, 2014; Hedin et al., 2015; Park et al., 2010; Peck, 1973, 1978; Smolis & Bernard, 2017; Wheeler & McHugh, 1994). The spruce–fir forests may not reach the gross species richness that deciduous forests at lower elevations do, but the higher elevations appear to contain a larger proportion of short-range endemics (Barr, 1969). As is seen in other sky-island systems (Uscanga et al., 2021), these seem to represent a combination of paleoendemics (formerly widespread species now extirpated elsewhere, (e.g., Hedin et al., 2015; Keith & Hedin, 2012; Thomas & Hedin, 2008) and neoendemics, those that seem to have arisen more or less in situ, resulting in cryptic species complexes, like *Trechus* ground beetles (Barr, 1979), *Geostiba* rove beetles (Gusarov, 2002) and *Adelopsis* fungus beetles (Peck, 1973, 1978). Across all these taxa, we see a range of distributional extents from those like Fraser fir itself, occurring across essentially all the suitable high-elevation patches, to a number of beetle species that are single-peak endemics.

Given how limited and fragmented these communities are, they would naturally present an important focus for conservation efforts. However, they have also experienced numerous anthropogenic challenges. Nearly all of the patches have experienced intensive logging, often with subsequent fire (Delcourt et al., 1998) that negatively affected the forest's re-establishment (Yarnell, 1998). Airborne pollution from mining in the region generated damaging acidic precipitation through much of the early 20th century (Likens & Bormann, 1974; McLaughlin et al., 1987). The arrival of the fir-feeding balsam woolly adelgid (*Adelges piceae*) in the southern Appalachians in the 1950s (White et al., 2012) resulted in the death of a large proportion of the

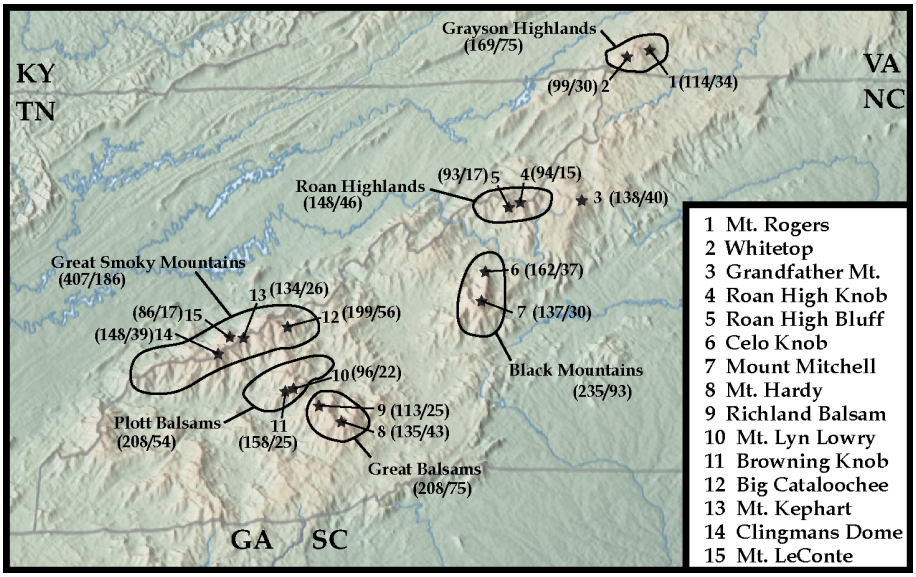

FIGURE 1 Map of the southern Appalachians, showing sampling sites (numbered) and ranges (outlined). Numbers in parentheses represent numbers of species/endemics, at both the peak and range levels, as assessed by the methods described herein.

Fraser firs that had survived or repopulated in the wake of the previous stressors. The high Appalachian spruce–fir forest now represents one of the most endangered ecosystems in the United States (Hamilton et al., 2022; Noss & Peters, 1995). While few direct effects of climate change on spruce–fir communities to date have yet been documented, niche modelling for some restricted taxa predicts significant reduction in habitable area (Erlandson, 2018; Ulrey et al., 2016), and some climate change scenarios predict complete extirpation of these communities before 2100 (Delcourt & Delcourt, 1998).

Understanding the distributions of the species living in these communities, as well as their particular environmental requirements, is critical to assessing threats to their persistence and developing conservation strategies for their preservation. One sub-community that has received particularly meagre attention is the diverse arthropod fauna of the forest floor. Arthropods living in leaf litter represent critical players in nutrient cycling, assisting with mechanical and chemical breakdown of plant and fungal tissues, directly and indirectly (via subsequent fungal and bacterial digestion) replenishing soils (Kampichler & Bruckner, 2009; Lawrence & Samways, 2003; Pramanik et al., 2001), and they also represent a wealth of unknown and underappreciated biodiversity, with new species and major range extensions discovered regularly, even in relatively well-studied areas like the south-eastern United States (Caterino, 2022; Caterino & Vasquez-Velez, 2017; Caterino & Vásquez-Vélez, 2022; Draney et al., 2019; Ferro, 2010; Hedin & Milne, 2023; Marek, 2010; Marek et al., 2018; Means et al., 2021; Owens & Carlton, 2016; Park et al., 2010; Platnick, 1999; Sierwald et al., 2019; Smolis & Bernard, 2017; Sokolov et al., 2004, 2007). Arthropods residing in the leaf litter tend to be of small body size, and many of the insects have evolved to be flightless (Anderson & Ashe, 2000). So, they frequently exhibit small overall range sizes, and have limited dispersal abilities. Together these factors have helped propel their diversification, frequently resulting in high local endemism (Barr, 1974, 1979; Gusarov, 2002; Hedin & Milne, 2023), but also making them particularly vulnerable to habitat alteration, as they cannot easily migrate from a suddenly unsuitable patch to one that is relatively intact.

What studies there have been on the litter fauna of the southern Appalachians have tended to emphasise particular taxa, resulting in many of the taxonomic papers cited above, or attempting to characterise faunas within limited groups (e.g., Bernard & Felderhoff, 2007; Caterino et al., 2017; Lamoncha & Crossley, 1998; Rieske & Buss, 2001). Others have examined intraspecific phylogeographic patterns within single lineages (Caterino & Langton-Myers, 2018, 2019; Hedin et al., 2015; Thomas & Hedin, 2008). Very few have attempted to characterise the fauna more generally (e.g., Gist & Crossley, 1975), and only one, that we are aware of, has approached such work specifically in the spruce–fir forest (Hughes, 1993).

One of the main challenges in characterising broader litter arthropod communities is the inadequate state of the taxonomy of many members. Even where modern revisions exist, identification is extremely challenging, relying on microdissections of genitalia and slide preparations, which hinder even arthropod specialists attempting to work outside their own particular group. Recent advances in next-generation sequencing (NGS)-based DNA taxonomy hold great promise for addressing these challenges, and it is perhaps the most reasonable and effective way to measure diversity in such hyperdiverse communities (Porter & Hajibabaei, 2018; Yeo et al., 2020). Combined with some standardised approaches for species delimitation, large-scale DNA-based characterizations of communities allow tabulation of overall species richness, and may readily permit comparisons of community overlap and complementarity among sites. Furthermore, such approaches are not hindered by the difficulty of identifying immature stages or sexes that do not exhibit species-specific taxonomic characters (Yeo et al., 2018, 2021). Molecular taxonomic

approaches are still challenged to an extent by the (limited) availability of reference sequences in public databases to make accurate identifications (Recuero et al., in press). But for the quantification of community richness and objective comparison of species similarity, where taxonomic identifications are not critical, they provide an exceptionally useful tool.

In the course of more completely documenting the leaf-litter arthropods living in the spruce–fir forests at the highest elevations of the Appalachian Mountains, we are working to examine broad patterns of species richness across ranges, and to assess levels of local endemicity across spatially isolated peak populations. In this article, we explore the potential of NGS barcoding to more fully document these diverse and taxonomically challenging high-elevation arthropod communities. Specifically, we employ voucher-based mtDNA megabarcoding (sensu Chua et al., 2023) on exemplars of morphospecies of most arthropods collected from leaf litter at a network of 15 high-elevation sites spanning the north–south range of spruce–fir distribution in the southern Appalachians. We compare the overall richness of these sites, explore patterns of diversity and endemicity, and examine the degree to which sites host unique species versus populations of more widely distributed ones. We predict that the highest species richness will occur in sites to the southern part of the range, where there are larger spruce–fir patches and especially because these areas will have experienced the lowest severity of Pleistocene cooling, and that northern peaks will have lower proportions of endemism, having been established (becoming habitable) and isolated more recently. We also predict that many species will be restricted to one or the other sides of the Asheville Depression, formed by the French Broad River (FBR) valley, as this feature has emerged as the region's predominant biogeographic barrier in numerous studies to date (Barr, 1969; Caterino & Langton-Myers, 2018, 2019; Crespi et al., 2003; Garrick et al., 2017; Gusarov, 2002; Hedin & McCormack, 2017; Hedin & Thomas, 2010; Herman & Bouzat, 2016; Newton et al., 2020; Thomas & Hedin, 2008). It is likely that during post-Pleistocene migrations this lowland feature served as a filter for some number of boreal taxa as they attempted to retreat northward. The degree of endemism versus similarity among these sky-island litter arthropod communities will provide an unprecedented look at the uniqueness and conservation value of high Appalachian habitats, and will allow more informed approaches to their future management.

METHODS

Sites and sampling

We sampled litter arthropods from 15 high Appalachian peaks within 7 more or less well-defined ranges, from (south-west to north-east): Clingmans Dome, Mount Kephart, Mount LeConte and Big Cataloochee (Great Smoky Mountains), Browning Knob and Mt. Lyn Lowry (Plott Balsams), Richland Balsam and Mt. Hardy (Great Balsams), Mount Mitchell and Celo Knob (Black Mountains), Roan High Knob and Roan High Bluff (Roan Highlands), Grandfather Mountain, and Whitetop and Mount Rogers (Grayson Highlands of Virginia). Some specific details on each of these sites are provided in Table 1, and their geography is shown in Figure 1. Each site was visited twice, once in late spring/early summer and once in fall. Beyond these 'primary sites' (listed in green in Supporting Document 1), analyses include a comparable number of barcode sequences from what will be referred to as secondary sites. The majority of these sequences

TABLE 1 Sampling sites, listed roughly north to south.

	Range	Elevation (m)	Lat	Long	Approximate area of contiguous conifer cover (nearest 10^2 ha)	Abbrev.
Mount Rogers	Grayson Highlands	1746	36.66	−81.55	400	MRg
Whitetop	Grayson Highlands	1682	36.64	−81.61	100	WT
Grandfather Mountain	N/A	1812	36.11	−81.81	300	GrM
Roan High Bluff	Roan Highlands	1910	36.09	−82.15	400	RHB
Roan High Knob	Roan Highlands	1916	36.10	−82.12	[contiguous with above]	RHK
Celo Knob	Black Mountains	1928	35.85	−82.25	100	CK
Mount Mitchell	Black Mountains	2037	35.76	−82.26	200	MM
Mount Hardy	Great Balsams	1865	35.30	−82.93	400	MHy
Richland Balsam	Great Balsams	1954	35.37	−82.99	1000	RB
Mount Lyn Lowry	Plott Balsams	1902	35.46	−83.11	400	LL
Browning Knob	Plott Balsams	1918	35.46	−83.14	[contiguous with above]	BK
Big Cataloochee	Great Smokies	1876	35.67	−83.18	500	BCat
Mount Kephart	Great Smokies	1895	35.63	−83.39	700	MK
Clingmans Dome	Great Smokies	2025	35.56	−83.50	3200	CD
Mount LeConte	Great Smokies	2010	35.65	−83.44	1700	MLc

represent other higher elevation sites (>1000 m) in the region that do not host spruce–fir communities. A smaller number from a wider selection of sites increases representation within particular genera of primary focus in our lab.

Each visit we took three leaf-litter samples by sifting. Litter in most spruce–fir sites consists of deep needle litter, with minor components of deciduous leaves and fine woody debris. Litter was sifted down to the soil surface (or to a depth where litter was so decayed as to be indistinguishable from soil, where the interface was not a hard boundary), over an area of approximately one square meter, through an 8-mm mesh, until a bag of approximately 6 litres was filled. Specific GPS coordinates were captured for each sample. Samples were processed in the laboratory using Berlese–Tullgren funnels, running subsamples until thoroughly dry, approximately 12 h. Specimens were collected directly into 100% ethanol, and moved to −20°C storage after each sub-sample was complete. We sorted each to major arthropod group (roughly Linnean Order) prior to sorting each to morphospecies. All arthropods, except Acari (mites; they have not yet been fully processed) and Myriapoda (they will be treated separately), represented by more than 20 samples (which excludes a few incidentals like Blattodea, Orthoptera and Mecoptera) were included in these analyses. Each set of three samples from a given date were considered together in circumscribing morphospecies—individual representatives of every distinct morphospecies were pulled from sample 1, then samples 2 and 3 were examined for additional morphospecies not represented in the others. Morphospecies sorting erred on the side of caution, presuming any slightly different morphotype to represent a distinct OTU. Thus, sexually dimorphic males and females were both included as separate morphospecies, immature or larval forms were considered separate, etc. Spring and fall samples were considered separately, such that morphospecies common to both may be represented by two specimens for each site. Samples were sorted to morphospecies by multiple workers over a couple of years, so for multiple reasons should not be taken as directly comparable. They are only meant as an indication of number of samples processed. Each morphospecies was assigned a unique code based on an abbreviation for the locality, 'A' or 'B' for spring or fall, respectively, and a number (e.g., the first morphospecies from spring Mount Mitchell samples was named MM.A.001).

Laboratory methods

One individual of each morphospecies was prepared for DNA extraction. Each was imaged, subdivided or punctured to permit tissue digestion, and placed in a separate well in a 96-well plate. Images of morphospecies are archived on our lab Flickr page, named by morphospecies code (https://www.flickr.com/photos/183480085@N02/albums). Tissues were digested with lysis buffer and proteinase K (Omega BioTek, Norcross, GA), and then the liquid fraction was removed to a new plate, with the voucher remains saved for archiving. The digested tissue mixture was extracted using Omega BioTek's Mag-Bind HDQ Blood and Tissue kit on a Hamilton Microlab Star automated liquid handling system, eluting with 150 µL elution buffer.

These analyses include sequences from three separate sequencing approaches. For some Collembola, we amplified a 658 base pair region of the cytochrome oxidase one (COI) mitochondrial 'barcoding' gene using primers LCO1490 and HCO2198 (GGTCAACAAATCATAAAG ATATTGG and TAAACTTCAGGGTGACCAAAAAATCA, respectively; Folmer et al., 1994). These PCR products were run on an agarose gel to assess amplification success and sent for clean-up and Sanger sequencing to Psomagen (Rockville, MD); amplicons were sequenced in both directions. This produced 64 of the sequences used here. The other specimens were sequenced using next-generation platforms as 'mini-barcodes', a 421 bp fragment of the mitochondrial COI gene using the primers BF2-BR2 (GCHCCHGAYATRGCHTTYCC and TCDGGRTGNCCRAARAAYCA, respectively; Elbrecht & Leese, 2017), corresponding to the downstream two thirds of the standard barcoding region. Each well was tagged with a unique combination of forward and reverse 9 bp indexes, synthesised as part of the primer by Eurofins Genomics (Louisville, KY). These indexes were derived from a list provided by Meier et al. (2016), to allow multiplexed NGS. All PCRs were conducted in 12.5 µL volumes (5.6 µL water, 1.25 µL Taq buffer, 1.25 µL dNTP mix [2.5 mM each], 0.4 µL MgCl [50 mM], 1.5 µL each primer, 0.05 µL Platinum Taq polymerase, 1 µL DNA template, with a 95C initial denaturation for 5 min, followed by 35 cycles of 94C (30 s), 50C (30 s), 72C (30 s) and a 5-min 72C final extension on an Eppendorf Gradient Mastercycler.

For library preparation, PCR products were combined and purified using Omega Bio-Tek's Mag-Bind Total Pure NGS Kit, in a ratio of 0.7:1 (enriching for fragments >300 bp). Illumina adapters and sequencing primers were ligated to PCR products using New England BioLab's Blunt/TA Ligase Master Mix. The amplicon+adapter library was again purified using Mag-Bind Total Pure NGS and subsequently quantified using a Qubit fluorometer. Final libraries were sequenced on an Illumina MiSeq using a v.3 2 × 300 paired-end kit or on a Nanopore MinION using a v10.4 flowcell and the ligation sequencing kit LSK-112 (Oxford Nanopore Technologies, Oxford, UK).

Data analysis

Sanger sequences were edited in Geneious (v8.1.8) by combining forward and reverse reads, confirming basecalls and exporting as text. Illumina reads were processed with bbtools software package (https://jgi.doe.gov/data-and-tools/bbtools/; v38.87 Bushnell et al., 2017) to merge paired read ends, remove PhiX reads, trim Illumina adapters, filter reads for the correct size, remove reads with quality score <30, cluster sequences by similarity allowing five mismatches (~1%) and generate a final matrix in FASTA format. Nanopore reads were basecalled using the 'super-accurate' algorithm of Guppy (v6.1.2), then demultiplexed using ONTbarcoder v0.1.9 (Srivathsan et al., 2021), with minimum coverage set at 5. FASTA files from all sequencing methods were trimmed to match the shorter 421 bp BF2-BR2 fragment, combined and aligned

with the online version of Mafft v7 (Katoh et al., 2017) using the auto strategy. All barcodes that did not meet the necessary quality standards were removed from the analyses, including sequences with high levels of ambiguous positions and barcodes that did not match with our coarse morphological identifications (to the order or family level). We carefully reviewed both the alignments and the phylogenetic reconstructions to identify additional, potential cases of contamination, which were in most cases the results of bacterial DNA amplification. Final barcode sequences have been deposited in GenBank under accession nos.: OR169027–OR174759.

Species delimitation analyses used all available sequences, including those from secondary sites, under the assumption that the most accurate delimitations will be found using the largest available data set. These were carried out one major monophyletic taxon (as listed in Table 3) at a time to facilitate thorough computation, and to allow the barcoding gap to be optimised to particular lineages. Results reported only include those OTUs represented with at least one occurrence at a primary focal site. We used the ASAP (Assemble Species by Automatic Partitioning; Puillandre et al., 2021) algorithm for species delimitation, as distance-based methods, and ASAP, in particular, have been found to most reasonably approximate 'true' species boundaries in other arthropod taxa delimitation studies (Copilaş-Ciocianu et al., 2022; Guo & Kong, 2022; Magoga et al., 2021). Using a single, consistent method across comparisons should permit meaningful comparisons of lineage composition among sites, even if the genetic units compared do not strictly correspond to biological species. Kimura 2-parameter distances were used due to generally high inter-population genetic distances, and the top three delimitations were saved, though only the best (lowest ASAP) score results are reported here. We acknowledge the arguments in favour of p-distances in barcode-based identification (e.g., Srivathsan & Meier, 2011) but inferred barcoding gaps nearing and exceeding 10% suggested that correction would be appropriate. Delimited species within each taxon were combined to a single list of all species and all primary, high-elevation sites. This species list was converted to presence–absence table using the splist2presabs function of the fuzzysim package in R (Barbosa, 2015; R Core Team, 2022).

Site endemicity and range endemicity (lumping peaks within ranges as in Table 1) and proportions of shared species among sites were assessed using EstimateS (Colwell, 2005). To analyse similarity in species composition among high-elevation sites only, we used non-metric multidimensional scaling (NMDS) of Sørensen among-site similarity. Similarity scores and NMDS scores were calculated in R using the Vegan package (Oksanen, 2017). The plot was created in R using the ggplot2 package (Wickham, 2016). We also built a maximum parsimony area cladogram with PAUP v4.0a (Swofford, 2002) using species as characters with presence–absence at a peak as character states, to reflect the diversity affinities among sites and ranges. The tree was rooted with a hypothetical "all species absent" locality, so that sharing any species is considered synapomorphic. We ran the heuristic search algorithm with TBR branch swapping and no MaxTrees limit. We assessed the stability of internal branches with non-parametric bootstrapping (1000 pseudoreplicates). We tested the possible effect of "isolation-by-distance" (IBD) using Mantel's tests to determine the correlation of pairwise similarities and geographic distances. Mantel's test is frequently used to test IBD, although results may be affected by autocorrelation of data. IBD analyses were performed with GENALEX v6.503 (Peakall & Smouse, 2006) with 99,999 permutations, comparing all localities, groups of localities north and south the FBR barrier, and within the Great Smoky Mountains range.

Vouchering

Following digestions, remains of extracted specimens were recombined with any non-extracted body parts, labelled, assigned unique CUAC (Clemson University Arthropod Collection) identifiers (but retaining also initial sample IDs) and curated into the CUAC. Unextracted representatives of morphospecies, if any, remain in bulk order-level samples, and are also permanently vouchered in the CUAC, as are unsorted residues (containing additional representatives of hyperabundant taxa, principally Acari and Collembola). A complete list of all specimens extracted, with collecting data, DNA extraction codes and voucher codes is available as a Supplemental Document 1.

RESULTS

The analyses for this article involved 6293 individual specimens, of which 2664 were from our primary high-elevation sites. The latter were resolved by ASAP analyses into 993 hypothesised species. This final number varied somewhat under alternative delimitation methods/metrics, with uncorrected distance-based ASAP reducing total species numbers by up to 5%, and mPTP reducing them by a third or more. Re-analysis under these scenarios would reduce inferred endemicity, but would also assert some species to contain much higher levels of polymorphism. Looking at these first by locality (Table 2), high-elevation communities hosted between 86 and 199 estimated species in the focal groups of litter arthropod (excluding mites and myriapods). These numbers, even aside from uncertainties arising from possible ASAP/'true' species mismatch, can only be considered provisional, as sequencing success rates were only approximately 65%, due to a combination of PCR failure, incomplete representation due to outcompetition in sequencing libraries, and contamination (and thus de facto exclusion from the final data set). There is some taxonomic signal in success rates, with Hymenoptera at the low end of only around 30%, while Coleoptera, Collembola and Araneae were near or over 80% (for more details, see Recuero et al., in press). Assuming some consistency across taxa in primer match and PCR success rates, at least, there should be more consistent representation in final species pools than these numbers might otherwise indicate.

The average endemicity by peak was 24%; of the species found in a given community, approximately one fourth were found only at that site. There is no obvious north–south trend in these numbers, with the higher endemicities scattered among northern (Whitetop: 30%,

TABLE 2 Species delimited by site, with endemicity by peak and by range.

Peak	Range	Total spp.	Site endemic	Site endemic (%)	Range spp.	Range endemic	Range endemic (%)
MRg	Grayson Highlands	114	34	30%	169	75	44%
WT	Grayson Highlands	99	30	30%			
GrM	Grandfather Mountain	138	40	29%	138	40	29%
RHB	Roan Highlands	93	17	18%	148	46	31%
RHK	Roan Highlands	94	15	16%			
CK	Black Mountains	162	37	23%	235	93	40%
MM	Black Mountains	137	30	22%			
MHy	Great Balsams	135	43	32%	208	75	36%
RB	Great Balsams	113	25	22%			
LL	Plott Balsams	96	22	23%	208	54	26%
BrK	Plott Balsams	158	25	16%			
BCat	Great Smokies	199	56	28%	407	186	46%
MK	Great Smokies	134	26	19%			
CD	Great Smokies	148	39	26%			
MLc	Great Smokies	86	17	20%			
Averages		127.1	30.4	24%	216.1	81.3	36%

TABLE 3 Species delimited by taxon and endemicity by peak and by range.

	Total individuals	Individuals (focal sites only)	ASAP spp	No. peak endemics	% peak end.	No. range endemics	% range end.
Isopoda	124	9	3	1	33%	2	67%
Araneae	558	237	51	10	20%	17	33%
Pseudoscorpionida	135	64	53	47	89%	53	100%
Protura	116	51	32	26	81%	32	100%
Diplura	67	23	11	6	55%	9	82%
Collembola	1517	690	294	134	46%	175	60%
Thysanoptera	24	8	3	1	33%	1	33%
Hemiptera	84	46	25	15	60%	16	64%
Coleoptera	2615	999	276	98	36%	145	53%
Lepidoptera	53	29	10	3	30%	3	30%
Diptera	655	405	185	88	48%	98	53%
Hymenoptera	343	111	50	28	56%	29	58%
Total	6291	2672	993	457	49% (AVG)	580	61% (AVG)

Grandfather Mt.: 29%) and southern (Mt. Hardy: 32%, Big Cataloochee: 28%) sites. Sites with higher species richness tended to harbour greater numbers of endemics ($R^2 = 0.68$), but there was only a weak indication that richer sites had higher proportions of endemics ($R^2 = 0.10$). Considering broader geographic scales, endemicity by mountain range averaged 36%, with highs at both the north-eastern (44%) and south-western (46%) extremes. So approximately half of the species occurring within a contiguous range occur only in that range, but without distinct broader geographic trends. Very few species were found over much broader ranges (Figure 2), with only small numbers occurring at more than a half-dozen sites.

Considering endemicity by taxon, there were some very pronounced differences (Table 3). The lowest average endemicities among taxa were in the Thysanoptera (33% [1 of 3] reported from a single site), Araneae (20%), Lepidoptera (30%) and Coleoptera (36%). The highest site endemicities were found in Hymenoptera (56%, though with the preceding caveats on spotty sequencing success), Protura (81%) and Pseudoscorpionida (89%). Range endemicities of taxa are, again, slightly higher, with more than half of all species of most orders endemic to just a single range. In terms of actual numbers, about half (457 of 993) were found on only a single peak, while 580 were found only within a single mountain range.

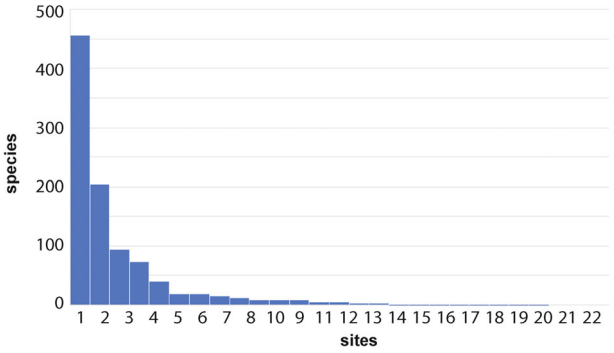

FIGURE 2 How widespread are most species? The vast majority of species (*y*-axis) were found at only 1 or 2 sites (*x*-axis).

Similarities across peaks (Table 4) largely follow expected geographic patterns, with Sørensen similarities (shared species) highest for within-range comparisons (e.g., Celo Knob vs. Mt. Mitchell, Mt. Rogers vs. Whitetop, Roan High Knob vs. Roan High Bluff, Clingmans Dome vs. Mt. Kephart, etc.). A few exceptions are noteworthy. For instance, Big Cataloochee's (Smokies) highest similarity (0.318) is to Browning Knob (Plott Balsams), and its closest within-range similarity, to Mt. Kephart, is considerably lower (0.270). However, Big Cataloochee is one of the most isolated high peaks in the Smokies, and its 'as-the-crow-flies' distance to Browning Knob is only slightly greater than to its nearest sampled neighbours within the Smokies (24 vs. 20 km). So, some IBD signal is probably reflected in these within- and among-range comparisons as well as environmental discontinuity. In fact, we observed a significant negative correlation between geographic distance and similarity considering all sites (Figure 3) and also for

TABLE 4 Sørensen similarity among sites (above diagonal) and actual number of shared species between sites (below diagonal).

Locality	MRg	WT	GrM	RHB	RHK	CK	MM	MHy	RB	LL	BrK	BCat	MK	CD	MLc
MRg		0.413	0.251	0.251	0.192	0.216	0.175	0.104	0.15	0.181	0.183	0.121	0.113	0.122	0.11
WT	44		0.192	0.219	0.166	0.137	0.16	0.137	0.142	0.144	0.163	0.114	0.129	0.146	0.108
GrM	32	23		0.265	0.221	0.256	0.222	0.145	0.181	0.169	0.187	0.118	0.167	0.166	0.106
RHB	26	21	31		0.396	0.272	0.251	0.123	0.146	0.148	0.159	0.13	0.123	0.149	0.145
RHK	20	16	26	37		0.248	0.233	0.114	0.126	0.105	0.111	0.109	0.114	0.132	0.111
CK	30	18	39	35	32		0.43	0.14	0.188	0.146	0.211	0.143	0.174	0.141	0.16
MM	22	19	31	29	27	65		0.103	0.159	0.111	0.202	0.107	0.154	0.182	0.17
MHy	13	16	20	14	13	21	14		0.323	0.26	0.224	0.192	0.149	0.163	0.154
RB	17	15	23	15	13	26	20	40		0.306	0.279	0.231	0.219	0.192	0.191
LL	19	14	20	14	10	19	13	30	32		0.345	0.231	0.235	0.23	0.209
BrK	25	21	28	20	14	34	30	33	38	44		0.318	0.314	0.293	0.229
BCat	19	17	20	19	16	26	18	32	36	34	57		0.27	0.248	0.175
MK	14	15	23	14	13	26	21	20	27	27	46	45		0.34	0.245
CD	16	18	24	18	16	22	26	23	25	28	45	43	48		0.274
MLc	11	10	12	13	10	20	19	17	19	19	28	25	27	32	

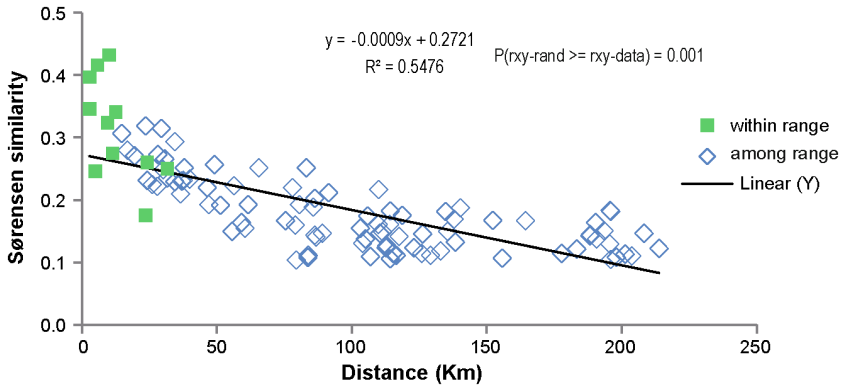

FIGURE 3 Isolation-by-distance of communities, plotting linear geographic distance in km (x-axis) against Sørensen between-population similarity (*y*-axis).

populations north or south of the Asheville Depression. However, within-range sampling was not dense enough to assess IBD at a finer scale. Lowest among-peak similarities all span the Asheville Depression, and this divide is seen very clearly in the NMDS plot (Figure 4), with all north-eastern peaks found on the left side of the plot and all south-western peaks on the right-hand side. Both groups are also clearly supported by the maximum parsimony area cladogram (Figures SS1 and S2). Range affiliations among peaks are generally evident in NMDS results, though clustering among them is very weak. Distances among range groups are very similar to actual geographic distances, at least to the extent of nearest neighbours (e.g., Smokies closest to Plott Balsams, then to Great Balsams). The cladogram provides strong support for most within-range relationships; only the Smokies and Plott Balsams appear intermixed.

DISCUSSION

The analyses carried out in this study reveal a highly detailed picture of high Appalachian arthropod endemism. Approximately one fourth of the taxa recognised here (by ASAP, more on which later) are found only on a single peak, and a significant additional number are found only in neighbouring communities, that is, other instances of spruce–fir habitat in the same mountain range. Looking at distributions slightly more broadly, more than one third of the species were found only within a particular mountain range. Truly widespread taxa were very much the exception, with only about 50 out of 993 species being found at more than five sites. It is expected that further sampling in these sites will reveal still more species in the more diverse taxa. We would expect that ratios of endemics to widespread species would likely remain similar as more species are found, although it is possible that levels of endemicity may be moderately affected.

Absolute diversity is very high, with a per-site average of 127 species for just our focal groups (that is, all arthropods excluding

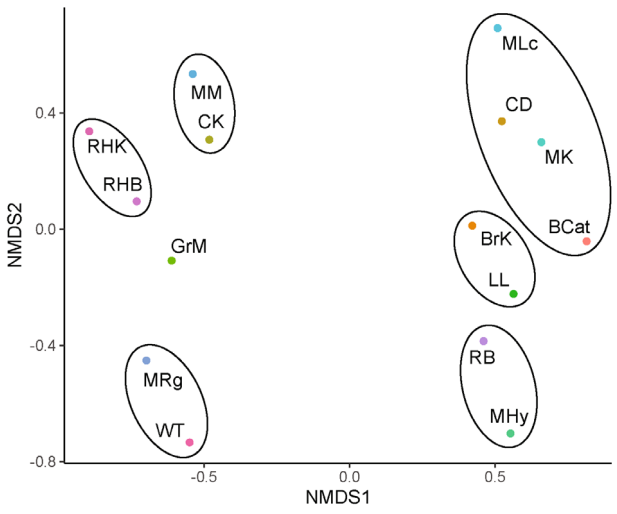

FIGURE 4 NMDS plot of among-site similarity, with ranges enclosed by ellipses.

Acari and Myriapoda), ranging from 86 to 199. Given that most of the oversplitting by automated species delimitation methods subdivides widely distributed but weakly divergent lineages (Mason et al., 2020; Talavera et al., 2013), we would not expect single-site species counts to be particularly inflated, though instances of oversplitting across sites, where multiple populations were sampled, might be expected. So, lacking detailed taxonomic information, what sort of confidence might we have that these ASAP-delimited species represent meaningful units? There are two potential issues here. First, we are hypothesising species-level diversity based on only a single mitochondrial marker. This comes with a list of standard caveats, mainly surrounding issues of smaller effective population sizes biasing the results towards more apparent divergence and isolation where there could be some level of gene flow. We are well aware of the many studies that have demonstrated the shortcomings of species circumscription based on COI by itself (e.g., Dupuis et al., 2012; Ranasinghe et al., 2022). Certainly additional exploration of uniqueness and levels of divergence based on a broader selection of genomic markers would test and strengthen the hypotheses suggested here.

The other question is, assuming that mitochondrial sequences reveal something meaningful about evolutionary history, do the 'species' as we report them here represent species in the biological (or any other) sense? Numerous studies have supported that species delimited on COI alone correspond to biologically meaningful units in a significant majority of cases and constitute a valuable first approximation (Chroni et al., 2016; Gomez et al., 2018; Magoga et al., 2021; Timm et al., 2022). While such reliability is not within the power of these data alone to assess, at a minimum, monophyletic groups exceeding a detected barcoding gap do represent divergent lineages, isolated for some span of time, and may be considered as potentially 'evolutionarily significant units'. For those exhibiting deeper divergences (for several higher taxa the dataset-wide 'barcoding gap' exceeds 5%), this genetic divergence might be associated with meaningful adaptive divergence to local conditions, although this remains controversial (Kohn et al., 2006). Despite coarsely similar floristic conditions among these sites, there are likely to be important differences in summer and winter temperature extremes, including frost-free days, differences in precipitation and its seasonality, and differences in co-occurring fungi and other ecological players in the communities.

Similarly, we can discount, for the majority of species occurring in the spruce–fir litter communities, the possibility that they are functioning as some sort of broader scale metapopulation, experiencing recolonization over ecological time from other patches of similar habitat. This may operate over very local scales, as treefalls, canopy openings, fire and succession modify microhabitat availabilities. But for most species, there appears to be little significant gene flow beyond the limits of each peak; mixed haplotype clusters from multiple localities are mainly seen in those relatively few widespread species.

Towards a 'reality check' on ASAP species, we can make some rough comparisons of these numbers with our pre-sequencing morphospecies counts. As described under *Materials and Methods*, these numbers are clearly not strictly comparable by design—we intentionally included immatures of unknown identity as distinct morphospecies in

hopes of making some associations through matching barcodes. Also, where taxa exhibit significant sexual dimorphism males and females are likely to have been included as separate morphospecies. On the other hand, where genitalic or other very fine morphological characters, such as chaetotaxy, would have been necessary to distinguish closely related taxa, our morphospecies counts may have undercounted actual diversity. Nonetheless, it is interesting to note that the numbers are in a similar range, with morphospecies numbers for focal groups averaging 124 species (vs. 127 for ASAP species) over all sites, ranging from 85 to 178 (vs. 86–199). If we were to add morphospecies numbers for mites and myriapods, these site totals would average 198 (147–267 total species per site). So if we may hypothesize similar levels of endemism (~1/4) across all these morphospecies at each site, most of the high peaks of southern Appalachia may be inferred to support 40–50 arthropod species found nowhere else.

While we contend that these results highlight real and important local-scale differences in the litter communities across these dispersed instances of sky-island habitat, one further consideration remains to be fully explored. This is the extent to which most of these species are in fact limited to the highest elevations versus just limited to particular mountains, with distributions extending to lower elevations but remaining somewhat isolated by valleys or other topographic or habitat features, or more generally by distance in these mostly dispersal-limited organisms. This has important implications for management, which needs to consider the actual area occupied and available to a species, and what the particular habitat associations of significance are, whether leaf litter in generic terms, or specific litters of high-elevation conifers. Work to date has tended to assume that higher elevation species were largely distinct (e.g., Hughes, 1993), but it could be that the proportions are as much influenced by drop-out of taxa restricted to lower elevations as to those unique. Our samples here have allowed a preliminary look at this question, with a selection of slightly lower and non-spruce–fir sites included in the endemicity calculations, but a broader sampling of lower elevation sites will be necessary to address this properly. A casual look at the data from this perspective reveal a mix, with some species wholly limited to the highest, spruce–fir sites, others with scattered low-elevation site records, or on lesser peaks lacking conifer cover, and yet others that are predominantly found at lower elevations with rare records in the spruce–fir. The balance of each remains to be determined.

Finally, we regret that some taxa were not able to be included in these analyses. The most significant omission is the Acari, which frequently represent the largest number of species in litter samples, high elevation and otherwise. Sadly, their diversity has been their downfall, as we were not able to fully process all mites for all sites examined here. We plan to continue processing these important samples and explore their diversities and distributions separately. Myriapod diversity is also being examined separately by collaborators. A larger proportion of myriapod (particularly Diplopoda) morphospecies are identified to named species, and these represent a similar mix of widespread and more narrowly restricted taxa. But few of the widespread ones have yet been studied in phylogeographic detail. So finer, intraspecific or cryptic, geographic differentiation remains a distinct possibility.

The overall failure rates of metabarcoding also remain a shortcoming of the present study, and one that must be kept in mind during any attempts to compare communities via metabarcoding. There are not only distinct morphospecies that failed to amplify, which would represent a more systematic bias, reducing overall numbers of taxa, but doing so more or less uniformly across sites. But there are also a substantial number of random failures, species successfully sequenced from some localities but not all. These will have some more meaningful impact on our inferences, generally making species look more narrowly endemic than they are. Because our initial morphospecies assessments could only be considered tentative, it could still be misleading to assume that because one individual of a morphospecies amplified and sequenced, they all should—there could be significant differences within some of our initial sorting units. But we did not attempt to apply a unified morphospecies taxonomy across samples—that is, morphospecies 5 from one site is not hypothesised as conspecific with morphospecies 5 from another site. They are simply all the recognisably different morphs within each sample, considered in isolation from all others.

All the above concerns notwithstanding, the overwhelming preponderance of evidence indicates that genetic diversity across these high-elevation sites is high and highly partitioned, leading to high rates of local-scale endemism. All sites host numerous unique gene pools, whether we call them species or not. In terms of landscape-level conservation management, it is clear that the local scale must be given strong consideration. Each site, as fuzzy as its boundaries for now remain, is home to numerous lineages not found anywhere else. This has further implications, considering the population-level histories that likely led to such isolation. There is little indication of haplotype sharing and therefore contemporary gene flow among sites, and, therefore, little expectation that migration corridors, if established, would facilitate movement from, for example, southern to more northerly sites as temperatures to the south warmed more rapidly. Translocation would furthermore have to be approached very cautiously, because gene pools do differ significantly among sites. The potential for site-specific adaptations would suggest that, for at least some movements, success might be low. Further, receiving sites mostly host their own instances of congeners, and there would be significant risks of genetic incompatibilities or disruption of local adaptations. The management and maintenance of each community in its original home must be very strongly preferred to any more radical approaches.

This study has greatly expanded the taxonomic and geographic scope of work on high-elevation endemism in the southern Appalachians, strongly supporting the previous suspicions that genetic uniqueness of these communities would be high. Along with prominent species of concern, such as the endangered spruce-fir moss spider (*Microhexura montivaga*), the litter community as a whole should be considered highly restricted and worthy of focused protection. We would hazard to predict from these results that litter arthropod communities in sky-island systems more generally would conform

to similar patterns and benefit from comparable conservation considerations. Such protections would include protection from invasive plants and animals, careful management of (and perhaps with) fire, and regular monitoring, now that some baseline of diversities have been established. Continued work to model the specific environmental dependencies of these taxa will further help to guide management efforts.

AUTHOR CONTRIBUTIONS

Michael S. Caterino: Conceptualization; funding acquisition; writing – original draft; data curation; formal analysis; investigation; writing – review and editing; project administration; supervision; methodology. **Ernesto Recuero:** Conceptualization; investigation; writing – original draft; writing – review and editing; methodology; formal analysis; data curation.

ACKNOWLEDGEMENTS

This study was funded by the U.S. National Science Foundation (Award DEB-1916263 to MSC) and the Clemson University Experiment Station (SC-1700596). We also acknowledge the support of the John and Suzanne Morse Endowment for Arthropod Biodiversity. For permissions and assistance with field work, we are grateful to the North Carolina State Parks, Great Smoky Mountains National Park, Blue Ridge Parkway National Park, Monica Martin, Frank Etzler, Curt Harden, Anthony Deczynski, Patricia Wooden, Adam Haberski, Roy Kucuk, Laura Vásquez-Vélez, Laary Cushman, Paul Marek, Michael Ferro and Will Kuhn. For assistance in the laboratory, we thank Grace Arnold, Mary Atieh, Caroline Dukes, Kathleen Jordan, Steven Kresovich, Caroline McCluskey, Grace Holliday, Lauren O'Connell, Ashley Phillips, Vince Richards, Craig Robertson, Hannah Skinner, Anthony Villanueva and Alejandra Carranza. Rudolf Meier and Amrita Srivathsan also provided invaluable advice in applying ONT Barcoder. Finally, we appreciate two reviewers' comments that helped improve the manuscript. This paper represents Technical Contribution No. 7189 of the Clemson University Experiment Station.

CONFLICT OF INTEREST STATEMENT

We declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Our primary data are cytochrome oxidase I sequences associated with specimen vouchers. Locality, collection event and taxonomic data are associated with each sequence, and unique voucher codes (CUACXXXXXXXXX) tie the sequences to specimens housed in the Clemson University Arthropod Collection. Our supplementary data files contain all these basic data, and our formatted nexus files (partitioned by major arthropod taxon) have been placed on Dryad (https://doi.org/10.5061/dryad.x0k6djhq0) for future use by other researchers. All sequences have also been uploaded to GenBank. Images of every specimen extracted are archived on Flickr (https://www.flickr.com/photos/183480085@N02/albums) and can be found by searching the morphospecies codes found in Table S1 (e.g., 'BgBld.B.349').

ORCID

Michael S. Caterino https://orcid.org/0000-0002-2597-5707

REFERENCES

Allen, J.L. & Lendemer, J.C. (2016) Climate change impacts on endemic, high-elevation lichens in a biodiversity hotspot. *Biodiversity and Conservation*, 25(3), 555–568. Available from: https://doi.org/10.1007/s10531-016-1071-4

Anderson, L.E. & Zander, R.H. (1973) The mosses of the southern blue Ridge Province and their phytogeographic relationships. *Journal of the Elisha Mitchell Scientific Society*, 89, 15–60.

Anderson, R.S. & Ashe, J.S. (2000) Leaf litter inhabiting beetles as surrogates for establishing priorities for conservation of selected tropical montane cloud forests in Honduras, Central America (Coleoptera; Staphylinidae, Curculionidae). *Biodiversity and Conservation*, 9(5), 617–653.

Barbosa, A.M. (2015) fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*, 6(7), 853–858. Available from: https://doi.org/10.1111/2041-210X.12372

Barr, T.C. (1969) Evolution of the Carabidae (Coleoptera) of the southern Appalachians. In: Hoffman, R.L. & Hart, C.W. (Eds.) *P.C. Holt*. Virginia Polytechnic Institute: The Distributional History of the Biota of the Southern Appalachians. Part I. Invertebrates, pp. 67–92.

Barr, T.C. (1974) The eyeless beetles of the genus *Arianops* Brendel (Coleoptera, Pselaphidae). *Bulletin of the American Museum of Natural History*, 154(1), 1–52.

Barr, T.C. (1979) Revision of Appalachian *Trechus* (Coleoptera: Carabidae). *Brimleyana*, 2, 29–75.

Bernard, E.C. & Felderhoff, K.L. (2007) Biodiversity explosion: Collembola (springtails) of Great Smoky Mountains National Park. *Southeastern Naturalist*, 1, 175–182.

Bezy, R.L. & Cole, C.J. (2014) Amphibians and reptiles of the Madrean archipelago of Arizona and New Mexico. *American Museum Novitates*, 2014(3810), 1–24.

Boehm, M.S. (2012) *Late Pleistocene climate, vegetation, and fire history from a southern Appalachian bog, Whiteoak bottoms, Nantahala National Forest, North Carolina*. Knoxville, USA: University of Tennessee.

Bowers, J.E. & McLaughlin, S.P. (1996) Flora of the Huachuca Mountains, a botanically rich and historically significant sky Island in Cochise County, Arizona. *Journal of the Arizona-Nevada Academy of Science*, 29(2), 66–107.

Bushnell, B., Rood, J. & Singer, E. (2017) BBMerge—accurate paired shotgun read merging via overlap. *PLOS One*, 12(10), e0185056. Available from: https://doi.org/10.1371/journal.pone.0185056

Caterino, M.S. (2022) First report of the *Euconnus* Thomson subgenus *Cladoconnus* Reitter in the New World, represented by thirteen new Appalachian species (Coleoptera, Staphylinidae, Scydmaeninae). *ZooKeys*, 1137, 133–175. Available from: https://doi.org/10.3897/ZOOKEYS.1137.97068

Caterino, M.S. & Langton-Myers, S.S. (2018) Long-term population persistence of flightless weevils (*Eurhoptus pyriformis*) across old- and second-growth forests patches in southern Appalachia. *BMC Evolutionary Biology*, 18(165), 1–16.

Caterino, M.S. & Langton-Myers, S.S. (2019) Intraspecific diversity and phylogeography in southern Appalachian *Dasycerus carolinensis* (Coleoptera: Staphylinidae: Dasycerinae). *Insect Systematics and Diversity*, 3(6), 1–12. Available from: https://doi.org/10.1093/isd/ixz022

Caterino, M.S. & Vasquez-Velez, L.M. (2017) A revision of *Prespelea* Park (Staphylinidae, Pselaphinae). *ZooKeys*, 685, 105–130. Available from: https://doi.org/10.3897/zookeys.685.13811

Caterino, M.S. & Vásquez-Vélez, L.M. (2022) A new species of the southern Appalachian genus *Prespelea* Park (Coleoptera: Staphylinidae: Pselaphinae). *The Coleopterists Bulletin*, 76(4), 589–594. Available from: https://doi.org/10.1649/0010-065X-76.4.589

Caterino, M.S., Vasquez-Velez, L.M. & Myers, S.S. (2017) On the conservation value of second-growth forests for leaf-litter inhabiting beetles. *Insect Systematics and Diversity*, 1(1), 20–28.

Chroni, A., Djan, M., Obreht Vidaković, D., Petanidou, T. & Vujić, A. (2016) Molecular species delimitation in the genus *Eumerus*. *Bulletin of Entomological Research*, 107(1), 126–138.

Chua, P.Y.S., Bourlat, S.J., Ferguson, C., Korlevic, P., Zhao, L., Ekrem, T. et al. (2023) Future of DNA-based insect monitoring. *Trends in Genetics*, 39(7), 531–544. Available from: https://doi.org/10.1016/j.tig.2023.02.012

Colwell, R.K. (2005) EstimateS: Statistical estimation of species richness and shared species from samples. Freeware published at: http://viceroy.eeb.uconn.edu/EstimateS

Copilaş-Ciocianu, D., Rewicz, T., Sands, A.F., Palatov, D., Marin, I., Arbačiauskas, K. et al. (2022) A DNA barcode reference library for endemic Ponto-Caspian amphipods. *Scientific Reports*, 12(1), 1–14. Available from: https://doi.org/10.1038/s41598-022-15442-w

Crespi, E.J., Browne, R.A. & Rissler, L.J. (2010) Taxonomic revision of *Desmognathus wrighti* (Caudata: Plethodontidae). *Herpetologica*, 66(3), 283–295. Available from: https://doi.org/10.1655/HERPETOLOGICA-D-09-00002.1

Crespi, E.J., Rissler, L.J. & Browne, R.A. (2003) Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Molecular Ecology*, 12(4), 969–984. Available from: https://doi.org/10.1046/j.1365-294X.2003.01797.x

Daru, B.H., Farooq, H., Antonelli, A. & Faurby, S. (2020) Endemism patterns are scale dependent. *Nature Communications*, 11(2115), 1–11.

Delcourt, H.R. (1985) Holocene vegetational changes in the southern Appalachian mountains, USA. In: *Ecologia Mediterranea Tome XI (Fascicule 1)*, Marseille, France, pp. 9–16.

Delcourt, P.A. & Delcourt, H.R. (1998) Paleoecological insights on conservation of biodiversity: a focus on species, ecosystems, and landscapes. *Ecological Applications*, 8(4), 921–934.

Delcourt, P.A., Delcourt, H.R., Ison, C.R., Sharp, W.E. & Gremillion, K.J. (1998) Prehistoric human use of fire, the eastern agricultural complex, and Appalachian oak-chestnut forests: paleoecology of cliff palace pond, Kentucky. *American Antiquity*, 63(2), 263–278. Available from: https://doi.org/10.2307/2694697

DeSisto, J. (2014) *A first survey of the centipedes of Great Smoky Mountains National Park*. University of Connecticut (honors thesis). 9pp.

Dourson, D.C. & Langdon, K. (2012) Land snails of selected rare high elevation forests and heath balds of the Great Smoky Mountains National Park. *Journal of the North Carolina Academy of Science*, 128(2), 27–32. Available from: https://doi.org/10.2307/24336636

Draney, M.L., Milne, M.A., Ulyshen, M. & Madriz, G. (2019) Revision of the spider genus *Epiceraticelus* (Araneae, Linyphiidae) with a description of a new species. *Zootaxa*, 4646(1), 189–200. Available from: https://doi.org/10.11646/zootaxa.4646.1.11

Dupuis, J.R., Roe, A.D. & Sperling, F.A.H. (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology*, 21(18), 4422–4436.

Elbrecht, V. & Leese, F. (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, 5, 11. Available from: https://doi.org/10.3389/fenvs.2017.00011

Erlandson, S.K. (2018) *History, Dispersal Limitation, and Environment Shape the Current and Future Ranges of Forest Herbs of the Southern Appalachians*. M.S. Thesis. University of Minnesota. 53pp.

Ferro, M.L. (2010) Fifteen new species of *Sonoma* Casey from the eastern United States and a description of the male of *Sonoma tolulae* (LeConte) (Coleoptera: Staphylinidae: Pselaphinae). *Insecta Mundi*, 137, 1–44.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.

Garrick, R.C., Sabree, Z.L., Jahnes, B.C. & Oliver, J.C. (2017) Strong spatial-genetic congruence between a wood-feeding cockroach and its bacterial endosymbiont, across a topographically complex landscape. *Journal of Biogeography*, 44(7), 1500–1511. Available from: https://doi.org/10.1111/jbi.12992

Gist, C.S. & Crossley, D.A. (1975) The litter arthropod community in a southern Appalachian hardwood forest: numbers, biomass and mineral element content. *American Midland Naturalist*, 93(1), 107–122.

Gomez, D.F., Skelton, J., Sedonia Steninger, M., Stouthamer, R., Rugman-Jones, P., Sittichaya, W. et al. (2018) Species delineation within the *Euwallacea fornicatus* (Coleoptera: Curculionidae) complex revealed by morphometric and phylogenetic analyses. *Insect Systematics and Diversity*, 2(6), 1–11.

Grebennikov, V. (2021) Sky islands of the Cameroon volcanic line support the westernmost clade of five new *Typoderus* weevils (Coleoptera: Curculionidae: Molytinae). *Arthropod Systematics & Phylogeny*, 79, 57–74.

Guo, B. & Kong, L. (2022) Comparing the efficiency of single-locus species delimitation methods within Trochoidea (Gastropoda: Vetigastropoda). *Genes*, 13(2273), 1–17.

Gusarov, V.I. (2002) A revision of nearctic species of the genus *Geostiba* Thomson, 1858. *Zootaxa*, 81, 1–88.

Hamilton, H., Smyth, R.L., Young, B.E., Howard, T.G., Tracey, C., Breyer, S. et al. (2022) Increasing taxonomic diversity and spatial resolution clarifies opportunities for protecting US imperiled species. *Ecological Applications*, 32(3), e2534. Available from: https://doi.org/10.1002/EAP.2534

Hartop, E., Srivathsan, A., Ronquist, F. & Meier, R. (2022) Towards large-scale integrative taxonomy (LIT): resolving the data conundrum for dark taxa. *Systematic Biology*, 71, 1404–1422. Available from: https://doi.org/10.1093/sysbio/syac033

Hausmann, A., Krogmann, L., Peters, R.S., Rduch, V. & Schmidt, S. (2020) GBOL III: Dark Taxa. *iBOL Barcode Bulletin*, 10(1), 1–4. Available from: https://doi.org/10.21083/ibol.v10i1.6242

Heald, W.F. (1951) Sky Islands of Arizona. *Natural History*, 60(56–63), 95–96.

Hedin, M., Carlson, D. & Coyle, F. (2015) Sky Island diversification meets the multispecies coalescent—divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. *Molecular Ecology*, 24(13), 3467–3484. Available from: https://doi.org/10.1111/mec.13248

Hedin, M. & McCormack, M. (2017) Biogeographical evidence for common vicariance and rare dispersal in a southern Appalachian harvestman (Sabaconidae, *Sabacon cavicolens*). *Journal of Biogeography*, 44(7), 1665–1678. Available from: https://doi.org/10.1111/jbi.12973

Hedin, M. & Milne, M.A. (2023) New species in old mountains: integrative taxonomy reveals ten new species and extensive short-range endemism in *Nesticus* spiders (Araneae, Nesticidae) from the southern Appalachian Mountains. *ZooKeys*, 1145, 1–130. Available from: https://doi.org/10.3897/zookeys.1145.96724

Hedin, M. & Thomas, S.M. (2010) Molecular systematics of eastern north American Phalangodidae (Arachnida: Opiliones: Laniatores), demonstrating

convergent morphological evolution in caves. *Molecular Phylogenetics and Evolution*, 54(1), 107–121. Available from: https://doi.org/10.1016/j.ympev.2009.08.020

Herman, T.A. & Bouzat, J.L. (2016) Range-wide phylogeography of the four-toed salamander: out of Appalachia and into the glacial aftermath. *Journal of Biogeography*, 43(4), 666–678. Available from: https://doi.org/10.1111/jbi.12679

Hughes, D.N. (1993) *Arthropods Associated with Leaf Litter of Fraser Fir in the Great Smoky Mountains National Park*. Ph.D. Thesis. University of Tennessee, Knoxville.

Kampichler, C. & Bruckner, A. (2009) The role of microarthropods in terrestrial decomposition: a meta-analysis of 40 years of litterbag studies. *Biological Reviews*, 84(3), 375–389. Available from: https://doi.org/10.1111/j.1469-185X.2009.00078.x

Katoh, K., Rozewicki, J. & Yamada, K.D. (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20, 1160–1166. Available from: https://doi.org/10.1093/bib/bbx108

Keith, R. & Hedin, M. (2012) Extreme mitochondrial population subdivision in southern Appalachian paleoendemic spiders (Araneae: Hypochilidae: *Hypochilus*), with implications for species delimitation. *Journal of Arachnology*, 40(2), 167–181. Available from: https://doi.org/10.1636/A11-49.1

Kohn, M.H., Murphy, W.J., Ostrander, E.A. & Wayne, R.K. (2006) Genomics and conservation genetics. *Trends in Ecology & Evolution*, 21(11), 629–637. Available from: https://doi.org/10.1016/J.TREE.2006.08.001

Kortmann, M., Roth, N., Buse, J., Hilszczański, J., Jaworski, T., Morinière, J. et al. (2022) Arthropod dark taxa provide new insights into diversity responses to bark beetle infestations. *Ecological Applications*, 32(2), e2516. Available from: https://doi.org/10.1002/EAP.2516

Lamoncha, K.L. & Crossley, D.A. (1998) Oribatid mite diversity along an elevation gradient in a southeastern appalachian forest. *Pedobiologia*, 42, 43–55.

Lawrence, J.M. & Samways, M.J. (2003) Litter breakdown by the Seychelles giant millipede and the conservation of soil processes on Cousine Island, Seychelles. *Biological Conservation*, 113(1), 125–132. Available from: https://doi.org/10.1016/S0006-3207(02)00355-5

Likens, G.E. & Bormann, F.H. (1974) Acid rain: a serious regional environmental problem. *Science (New York, NY)*, 184(4142), 1176–1179. Available from: https://doi.org/10.1126/science.184.4142.1176

Maddison, W. & McMahon, M. (2000) Divergence and reticulation among montane populations of a jumping spider (*Habronattus pugillis* Griswold). *Systematic Biology*, 49(3), 400–421. Available from: https://doi.org/10.1080/10635159950127312

Magoga, G., Fontaneto, D. & Montagna, M. (2021) Factors affecting the efficiency of molecular species delimitation in a species-rich insect family. *Molecular Ecology Resources*, 21(5), 1475–1489.

Marek, P.E. (2010) A revision of the Appalachian millipede genus *Brachoria* Chamberlin, 1939 (Polydesmida: Xystodesmidae: Apheloriini). *Zoological Journal of the Linnean Society*, 159(4), 817–889. Available from: https://doi.org/10.1111/j.1096-3642.2010.00633.x

Marek, P.E., Means, J.C. & Hennen, D.A. (2018) *Apheloria polychroma*, a new species of millipede from the Cumberland Mountains (Polydesmida: Xystodesmidae). *Zootaxa*, 4375, 409–425.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature (London)*, 405(6783), 243–253.

Mason, N.A., Fletcher, N.K., Gill, B.A., Funk, W.C. & Zamudio, K.R. (2020) Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. *Systematics and Biodiversity*, 18(3), 269–280.

Masta, S.E. (2000) Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): recent vicariance of sky Island populations? *Evolution*, 54(5), 1699–1711. Available from: 10.1111/J.0014-3820.2000.TB00714.X

McLaughlin, S.B., Downing, D.J., Blasing, T.J., Cook, E.R. & Adams, H.S. (1987) An analysis of climate and competition as contributors to decline of red spruce in high elevation Appalachian forests of the eastern United States. *Oecologia*, 72(4), 487–501. Available from: https://doi.org/10.1007/BF00378973

Means, J.C., Hennen, D.A. & Marek, P.E. (2021) A revision of the *minor* species group in the millipede genus *Nannaria* Chamberlin, 1918 (Diplopoda, Polydesmida, Xystodesmidae). *ZooKeys*, 1030, 1–180. Available from: https://doi.org/10.3897/zookeys.1030.62544

Meier, R., Wong, W., Srivathsan, A. & Foo, M. (2016) $1 DNA barcodes for reconstructing complex phenomes and finding rare species in specimen-rich samples. *Cladistics*, 32(1), 100–110. Available from: https://doi.org/10.1111/cla.12115

Meyer, W.M., Eble, J.A., Franklin, K., McManus, R.B., Brantley, S.L., Henkel, J. et al. (2015) Ground-dwelling arthropod communities of a sky Island mountain range in southeastern Arizona, USA: obtaining a baseline for assessing the effects of climate change. *PLOS One*, 10(9), e0135210. Available from: https://doi.org/10.1371/journal.pone.0135210

Monjaraz-Ruedas, R., Mendez, R.W. & Hedin, M. (2023) Species delimitation, biogeography, and natural history of dwarf funnel web spiders (Mygalomorphae, Hexurellidae, *Hexurella*) from the United States/Mexico borderlands. *ZooKeys*, 1167, 109–157. Available from: https://doi.org/10.3897/zookeys.1167.103463

Monroe, A.F., Hubbard, S.C., Rodriguez, R., Kuhns, F.M., McClung, M.R., Moran, M.D. et al. (2022) Biodiversity of parasitic Hymenoptera across sky islands of Arkansas, United States. *Environmental Entomology*, 51(5), 922–926. Available from: https://doi.org/10.1093/EE/NVAC059

Moskwik, M. (2014) Recent elevational range expansions in plethodontid salamanders (Amphibia: Plethodontidae) in the southern Appalachian Mountains. *Journal of Biogeography*, 41(10), 1957–1966. Available from: https://doi.org/10.1111/jbi.12352

Newton, L.G., Starrett, J., Hendrixson, B.E., Derkarabetian, S. & Bond, J.E. (2020) Integrative species delimitation reveals cryptic diversity in the southern Appalachian *Antrodiaetus unicolor* (Araneae: Antrodiaetidae) species complex. *Molecular Ecology*, 29(12), 2269–2287. Available from: https://doi.org/10.1111/mec.15483

Nielson, E.S., Hanson, J.O., Carvalho, S.B., Beger, M., Henriques, R., Kershaw, F. et al. (2022) Molecular ecology meets systematic conservation planning. *Trends in Ecology & Evolution*, 38(2), 143–155.

Noss, R.F. & Peters, R.L. (1995) *Endangered ecosystems: a status report on America's vanishing habitat and wildlife*. Washington, DC: Defenders of Wildlife.

Ober, K., Matthews, B., Ferrieri, A. & Kuhn, S. (2011) The evolution and age of populations of *Scaphinotus petersi* Roeschke on Arizona sky islands (Coleoptera, Carabidae, Cychrini). *ZooKeys*, 147, 183–197. Available from: https://doi.org/10.3897/ZOOKEYS.147.2024

Oksanen, J. (2017) Vegan: Community Ecology Package (R package version 2.6-4). Available from: https://github.com/vegandevs/vegan

Owens, B.E. & Carlton, C.E. (2016) Collecting in the museum: new species, taxonomic changes, and range extensions of euplectite Pselaphinae (Coleoptera: Staphylinidae) of the southeastern United States. *The Coleopterists Bulletin*, 70(3), 611–633.

Park, J.-S., Carlton, C.E. & Ferro, M.L. (2010) Diversity and taxonomic review of *Leptusa* Kraatz (Coleoptera: Staphylinidae: Aleocharinae) from Great Smoky Mountains National Park, U.S.A., with descriptions of four new species. *Zootaxa*, 2662, 1–27.

Peakall, R. & Smouse, P.E. (2006) Genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.

Peck, S.B. (1973) A systematic revision and the evolutionary biology of the *Ptomaphagus* (*Adelops*) beetles of North America with emphasis on cave-inhabiting species (Leiodidae). *Bulletin of the Museum of Comparative Zoology*, 145, 29–162.

Peck, S.B. (1978) Systematics and evolution of forest litter *Adelopsis* in the southern Appalachians (Coleoptera: Leiodidae; Catopinae). *Psyche: A Journal of Entomology*, 85(4), 355–382. Available from: https://doi.org/10.1155/1978/73825

Platnick, N.I. (1999) A revision of the Appalachian spider genus *Liocranoides* (Araneae, Tengellidae). *American Museum Novitates*, 3285, 1–13.

Porter, T.M. & Hajibabaei, M. (2018) Scaling up: a guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology*, 27(2), 313–338. Available from: https://doi.org/10.1111/mec.14478

Potter, K.M., Frampton, J., Josserand, S.A. & Nelson, C.D. (2008) Genetic variation and population structure in Fraser fir (*Abies fraseri*): a microsatellite assessment of young trees. *Canadian Journal of Forest Research*, 38(8), 2128–2137. Available from: https://doi.org/10.1139/X08-064

Pramanik, R., Sarkar, K. & Joy, V.C. (2001) Efficiency of detritivore soil arthropods in mobilizing nutrients from leaf litter. *Tropical Ecology*, 42(1), 51–58.

Puillandre, N., Brouillet, S. & Achaz, G. (2021) ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources*, 21(2), 609–620. Available from: https://doi.org/10.1111/1755-0998.13281

R Core Team. (2022) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ranasinghe, U.G.S.L., Eberle, J., Thormann, J., Bohacz, C., Benjamin, S.P. & Ahrens, D. (2022) Multiple species delimitation approaches with COI barcodes poorly fit each other and morphospecies – an integrative taxonomy case of Sri Lankan Sericini chafers (Coleoptera: Scarabaeidae). *Ecology and Evolution*, 12(5), e8942. Available from: https://doi.org/10.1002/ECE3.8942

Recuero, E., Etzler, F.E. & Caterino, M.S. (in press) Most soil and litter arthropods are unidentifiable based on current DNA barcode reference libraries. *Current Zoology*.

Rieske, L.K. & Buss, L.J. (2001) Influence of site on diversity and abundance of ground- and litter-dwelling Coleoptera in Appalachian oak-Hickory forests. *Environmental Entomology*, 30(3), 484–494.

Sierwald, P., Hennen, D.A., Zahnle, X.J., Ware, S. & Marek, P.E. (2019) Taxonomic synthesis of the eastern north American millipede genus *Pseudopolydesmus* (Diplopoda: Polydesmida: Polydesmidae), utilizing high-detail ultraviolet fluorescence imaging. *Zoological Journal of the Linnean Society*, 187(1), 117–142. Available from: https://doi.org/10.1093/zoolinnean/zlz020

Sipe, T.W. & Browne, R.A. (2004) Phylogeography of masked (*Sorex cinereus*) and smoky shrews (*Sorex cumeus*) in the southern Appalachians. *Journal of Mammalogy*, 85(5), 875–885. Available from: https://doi.org/10.1644/214

Slapcinsky, J. (2018) *Vitrea clingmani* Dall in Pilsbry, 1900, a snail endemic to the summits of the Black Mountains and great Craggy Mountains of North Carolina is now assigned to the genus *Pilsbryna* (Gastropoda: Stylommatophora: Oxychilidae). *The Nautilus*, 132, 1–12.

Smolis, A. & Bernard, E.C. (2017) *Intricatonura fjellbergi*, a new peculiar genus and species of Neanurini (Collembola: Neanuridae: Neanurinae) from Great Smoky Mountains National Park. *Florida Entomologist*, 100(4), 725–730. Available from: https://doi.org/10.1653/024.100.0419

Sokolov, I., Carlton, C.E. & Cornell, J. (2004) Review of *Anillinus*, with descriptions of 17 new species and a key to soil and litter species (Coleoptera: Carabidae: Trechinae: Bembidiini). *The Coleopterists Bulletin*, 58(2), 185–233.

Sokolov, I., Sokolova, Y. & Carlton, C.E. (2007) New species of *Anillinus* Casey (Carabidae: Trechinae: Bembidiini) from Great Smoky Mountains National Park, U.S.A. and phylogeography of the a. langdoni species group. *Zootaxa*, 1542, 1–20.

Srivathsan, A., Lee, L., Katoh, K., Hartop, E., Kutty, S.N., Wong, J. et al. (2021) ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone. *BMC Biology*, 19(217), 1–21. Available from: https://doi.org/10.1186/s12915-021-01141-x

Srivathsan, A. & Meier, R. (2011) On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. *Cladistics*, 28(2), 190–194.

Swofford, D.L. (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer & Associates.

Talavera, G., Dincă, V. & Vila, R. (2013) Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution*, 4(12), 1101–1110. Available from: https://doi.org/10.1111/2041-210X.12107

Thomas, S.M. & Hedin, M. (2008) Multigenic phylogeographic divergence in the paleoendemic southern Appalachian opilionid *Fumontana deprehendor* Shear (Opiliones, Laniatores, Triaenonychidae). *Molecular Phylogenetics and Evolution*, 46(2), 645–658. Available from: https://doi.org/10.1016/j.ympev.2007.10.013

Timm, V.F., Gonçalves, L.T., Valente, V.L.D.S. & Deprá, M. (2022) The efficiency of the COI gene as a DNA barcode and an overview of Orthoptera (Caelifera and Ensifera) sequences in the BOLD system. *Canadian Journal of Zoology*, 100(11), 710–718. Available from: 10.1139/CJZ-2022-0041/SUPPL_FILE/CJZ-2022-0041_SUPPLA.DOCX

Ulrey, C., Quintana-Ascencio, P.F., Kauffman, G., Smith, A.B. & Menges, E.S. (2016) Life at the top: long-term demography, microclimatic refugia, and responses to climate change for a high-elevation southern Appalachian endemic plant. *Biological Conservation*, 200, 80–92. Available from: https://doi.org/10.1016/J.BIOCON.2016.05.028

Uscanga, A., López, H., Piñero, D., Emerson, B.C., Mastretta-Yanes, A. & Parmakelis, A. (2021) Evaluating species origins within tropical sky-islands arthropod communities. *Journal of Biogeography*, 48(9), 2199–2210. Available from: https://doi.org/10.1111/jbi.14144

Watts, W.A. (1970) The full glacial vegetation of northwestern Georgia. *Ecology*, 51(17–33), 17–33.

Wheeler, Q.D. & McHugh, J.V. (1994) A new southern Appalachian species, *Dasycerus bicolor* (Coleoptera: Staphylinidae: Dasycerinae), from declining endemic fir forests. *The Coleopterists Bulletin*, 48, 265–271. Available from: https://doi.org/10.2307/4009111

White, P.B., Van de Gevel, S.L. & Soule, P.T. (2012) Succession and disturbance in an endangered red spruce–Fraser fir forest in the southern Appalachian Mountains, North Carolina, USA. *Endangered Species Research*, 18(1), 17–25.

Whitehead, D.R. (1981) Late-Pleistocene vegetation changes in northeastern North Carolina. *Ecological Monographs*, 51, 451–471.

Wickham, H. (2016) *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.

Yarnell, S.L. (1998) *The southern Appalachians: a history of the landscape*. Asheville, NC: USDA Forest Service southern Research Station.

Yeo, D., Puniamoorthy, J., Ngiam, R.W.J. & Meier, R. (2018) Towards holomorphology in entomology: rapid and cost-effective adult-larva matching using NGS barcodes. *Systematic Entomology*, 43(4), 678–691. Available from: https://doi.org/10.1111/syen.12296

Yeo, D., Srivathsan, A. & Meier, R. (2020) Longer is not always better: optimizing barcode length for large-scale species discovery and identification. *Systematic Biology*, 69(5), 999–1015. Available from: https://doi.org/10.1093/sysbio/syaa014

Yeo, H., Yeoh, T.X., Ding, H., Lee, T.T.M. & Puniamoorthy, N. (2021) Morphology and mini-barcodes: the inclusion of larval sampling and NGS-based barcoding improves robustness of ecological analyses of mosquito communities. *Journal of Applied Ecology*, 58(10), 2087–2100. Available from: https://doi.org/10.1111/1365-2664.13966

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Area cladogram created by parsimony analysis of species presence/absence by peak. Abbreviations for sites refer to those in Table 1.

Figure S2. Bootstrap consensus of area cladogram, showing bootstrap support where >50%.

Data S1. Table of all morphospecies included, with 'species' as delimited by ASAP, internal morphospecies codes, DNA extraction codes, locality and date information, unique identifiers for specimen vouchers (where preserved), GenBank accession numbers, sequencing run (1, 2, 3 on Illumina, 2b, 3b, and 4 on Nanopore), and taxonomy as far as determined. Those individuals from 'primary sites' for analysis are listed in green. Excel file.

Data S2. Delimited species presence (1)/absence (0) by primary site/peak. Excel file.

How to cite this article: Caterino, M.S. & Recuero, E. (2024) Shedding light on dark taxa in sky-island Appalachian leaf litter: Assessing patterns of endemicity using large-scale, voucher-based barcoding. *Insect Conservation and Diversity*, 17(1), 16–30. Available from: https://doi.org/10.1111/icad.12697