ANNUAL REVIEWS

Annual Review of Entomology

Entomological Collections in the Age of Big Data

Andrew Edward Z. Short,[1,*] Torsten Dikow,[2] and Corrie S. Moreau[3]

[1]Department of Ecology and Evolutionary Biology; and Division of Entomology, Biodiversity Institute, University of Kansas, Lawrence, Kansas 66045, USA; email: aezshort@ku.edu

[2]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA; email: dikowt@si.edu

[3]Department of Science and Education, Field Museum of Natural History, Chicago, Illinois 60605, USA; email: cmoreau@fieldmuseum.org

Annu. Rev. Entomol. 2018. 63:513–30

First published as a Review in Advance on October 20, 2017

The *Annual Review of Entomology* is online at ento.annualreviews.org

https://doi.org/10.1146/annurev-ento-031616-035536

*Corresponding author

Keywords

bioinformatics, biological specimens, digitization, genomics, natural history collections

Abstract

With a million described species and more than half a billion preserved specimens, the large scale of insect collections is unequaled by those of any other group. Advances in genomics, collection digitization, and imaging have begun to more fully harness the power that such large data stores can provide. These new approaches and technologies have transformed how entomological collections are managed and utilized. While genomic research has fundamentally changed the way many specimens are collected and curated, advances in technology have shown promise for extracting sequence data from the vast holdings already in museums. Efforts to mainstream specimen digitization have taken root and have accelerated traditional taxonomic studies as well as distribution modeling and global change research. Emerging imaging technologies such as microcomputed tomography and confocal laser scanning microscopy are changing how morphology can be investigated. This review provides an overview of how the realization of big data has transformed our field and what may lie in store.

INTRODUCTION

Natural history collections are in a state of rapid transition, perhaps more so than at any other time in their history (13, 74, 145). Once considered largely the domain of taxonomists and morphologists, biological collections have emerged as the scientific stage for a host of disciplines that rely on the breadth and depth of biodiversity captured in museums (130). Biological collections are foundational across many research fields in endeavors such as documenting and describing populations, species, and their tremendous diversity; improving public health, agricultural practices, and food security; monitoring environmental contamination; and studying the effects of biological invasions and global climate change (29, 130). These transitions are driven largely along two fronts: first, the liberation of vast troves of specimen data that were previously available only to those who physically handled specimens and, second, innovations and technological advances in genomic sequencing and morphological imaging. Altogether, these developments have unleashed a torrent of data, the magnitude of which was almost unimaginable even a decade ago.

BIOLOGICAL COLLECTIONS AND GENOMICS

Most material housed in entomological collections was collected prior to the routine generation and inclusion of molecular data in taxonomic and evolutionary studies, often before DNA itself was discovered. Consequently, most insect specimens were not collected with molecular and DNA-based studies in mind, and while many museum specimens are caught and preserved today using methods to minimize genetic degradation, this is by no means standard practice for a variety of practical, economic, and/or historical reasons. Since the majority of biodiversity represented in museum holdings is likely to be difficult to re-collect due to destruction of habitats, species and population loss, prohibitive costs of revisiting field sites, and shifting collecting and export policies, the scientific community has increasingly turned to these existing museum collections as sources of genetic material.

Specimen Preservation and Storage

Entomological collections typically consist of a dry (pinned, pointed, or enveloped) collection, a wet or alcohol collection, and a slide-mounted collection. Historically, these types of collections encompassed the vast majority of diversity contained in entomology collections and permitted a range of research techniques, including morphological examinations and dissections, even with very old samples. With a diversity of options available for preserving specimens, museum professionals are faced with often competing dilemmas on how to ensure collections have the highest scientific value now and into the future, while balancing time investment and cost-effectiveness to care for the large collections that most museums house (77, 79).

The relative success of extracting DNA from museum specimens preserved in these traditional ways and whether this genetic material is suitable for the study at hand depend on many factors, many of which remain poorly understood (64). Some molecular and genomic methods can take advantage of traditionally preserved specimens (i.e., Sanger gene-based sequencing, reduced-representation locus and genome sequencing and assembly), but new collections or alternatively preserved specimens are required for other genomic techniques and technologies (gene expression, genome annotation, whole-genome sequencing) (**Table 1**).

Recognizing the limitations of traditional preservation methods almost three decades ago, many collection curators and other scientists who vouchered their specimens in museum collections began advocating for storage techniques that more readily facilitated molecular studies (17, 79, 146). Although it may therefore seem that all future museum collections should be preserved in

Table 1 Museum storage methods and utility in DNA and genomic research

	Gene-based DNA sequencing, DNA barcoding, and ESTs	Reduced-representation genome sequencing (e.g., RAD-seq, UCEs)	Genome sequencing and assembly	Genome and gene annotation	Gene expression and transcriptomes	Host-associated microbiomes	Reference(s)
Dry/pinned storage	Some utility	Some utility	Some utility	No	No	Some utility	45, 64, 75 125, 129, 133, 136
Medium-grade (70%) ethanol	Some utility	Some utility	Some utility	No	No	Some utility	126
Propylene glycol	Some utility	Some utility	Some utility	No	No	Some utility	93, 100, 126
High-grade ethanol (95–100%)	Yes	Yes	Yes	Some utility	No	Yes	93, 107, 126
−20°C freezer	Yes	Yes	Yes	Some utility	No	Yes	2
−80°C ultracold freezer	Yes	Yes	Yes	Yes	Yes	Yes	2, 107
RNA preservative	Yes	Yes	Yes	Yes	Yes	Yes	93, 107
Cryogenics/liquid nitrogen	Yes	Yes	Yes	Yes	Yes	Yes	107

Some methods for which utility has been found may be only for short- or medium-term durations rather than permanent archival storage. Abbreviations: ESTs, expressed sequence tags; RAD-seq, restriction site–associated DNA sequencing; UCEs, ultraconserved elements.

high-grade ethanol, ultracold freezers, and/or cryogenic facilities (**Table 1**), these storage methods make the specimens more difficult to access, are more costly to implement and maintain than traditional methods, and in most cases are likely cost-prohibitive for many small to medium-sized collections (22). Several networks have recently been created for preserving biodiversity for genomic methods, including the Global Genome Initiative (**http://ggi.si.edu**) and Global Genome Biodiversity Network [**http://www.ggbn.org/ggbn_portal/** (33)], although none specifically target insects. Additionally, for groups such as Lepidoptera, new collection and storage protocols have been developed to optimize the use of specimens for both DNA and morphological work (19). This balance of benefits and costs requires museum professionals to prioritize collections for preservation media, which is the model most collections currently follow.

Many institutions have adopted policies for destructive and nondestructive sampling of accessioned specimens. Although some permit destructive methods for sampling DNA, such methods preclude preservation of a meaningful voucher and the single-use nature of such potentially irreplaceable specimens is not desirable. Fortunately, several nondestructive or minimally destructive methods for extracting genetic material have been developed for insect specimens (18, 45, 124, 134, 136). These methods do not generally require damage or trauma to the exterior of the specimen, although soft tissues may be lost in the process.

We argue here that rare specimens and nonmodel and rarely studied groups of organisms should be preserved in a manner that increases the likelihood they can be studied into the future. Regardless of storage preservative or method, tissues or whole specimens must be carefully entered into databases and labeled to ensure the highest standard for research utility.

Museum Specimens and Sanger Sequencing

Integrative taxonomy: combination of traditional morphological approaches with molecular data for species delimitation in revisionary taxonomy

Gene-based sequencing using Sanger sequencing technology has greatly advanced understanding of the tree of life during the last several decades. Several studies have considered the utility of traditional museum storage techniques for DNA-based studies of insects (31, 38, 39, 71, 93, 106–108, 110, 138) and insect-associated microbiota (47, 93) using Sanger sequencing, which has less stringent specimen preservation requirements than many modern genomic methods (**Table 1**).

Although the choice of genes that can be targeted depends on the age of the group of study, in many groups of insects and other arthropods, the number of available primer sets is restricted, leaving researchers with a limited choice of markers. In fact, for most DNA barcoding studies in insects, only a single mitochondrial gene fragment, cytochrome oxidase I, is included (48, 122), an approach that has come under criticism in the scientific literature (59, 82, 94, 123). Although there are concerns about the broad utility of DNA barcoding, this method has been successfully used on DNA obtained from insect and other arthropod museum specimens (51, 88, 124). The small number of developed primer sets can be viewed as a limitation, but the limited number has also permitted the combination of independently developed data sets to produce a broader understanding of the phylogenetic relationships of groups of arthropod taxa or to facilitate integrative taxonomy; examples include beetles (58), centipedes (46), ants (12, 92), and butterflies (142). In addition, for researchers who are interested in adding one or a few new samples to an existing phylogeny, generating the Sanger gene-based data to include these new taxa in legacy data sets is still the best option based on time and cost.

Another approach that relies on Sanger sequencing or 454 pyrosequencing but produces significantly more data than the gene-based approach is the creation of an expressed sequence tag (EST) database. ESTs can be used to help annotate genomes but have also been used successfully to reconstruct the phylogenetic relationships within several arthropod groups [e.g., beetles (57), Polyneoptera and Paraneoptera (72), arthropods (83), Hymenoptera (119)]. Although ESTs have been phylogenetically informative, the much higher cost and time investment incurred to generate these data when compared with newer technologies suggests this approach may be short lived for phylogenomics.

Museum Specimens and High-Throughput Sequencing

Reduced-representation genome sequencing (RRGS) protocols using next-generation and high-throughput sequencing are providing large-scale multigene/multilocus data sets. These methods, including restriction site–associated DNA sequencing (RAD-seq), genotyping by sequencing (GBS), targeted sequence capture/hybrid enrichment, transcriptomes for gene capture, and highly conserved or ultraconserved elements (UCEs), are proving valuable for inferring insect phylogenies [e.g., beetles (64); butterflies and moths (7, 66); wingless insects (27); ants, bees, and wasps (36, 61); lower neopterans (73); insects (89); holometabolous insects (102); polyneopterans (121)]. As these methods rely on high-throughput sequencing, which targets shorter fragments than traditional Sanger sequencing, some researchers have been able to employ them to sequence museum specimens that have been traditionally preserved, resulting in DNA that is frequently highly fragmented (97, 136).

A few studies have explicitly compared RRGS data with the traditional Sanger gene-based sequencing [e.g., beetles (24), ants (11)], demonstrating the utility and power of these approaches. One concern with some of these methods is the level of evolutionary divergence between the included taxa. For example, through a simulation study based on real genome data from fruit flies and other taxa, Rubin et al. (112) demonstrated that RAD-seq and genotyping-by-sequencing methods are informative for phylogenetic inference for clades younger than 50 million years.

This drop-off in utility at deeper timescales is due to mutations in the restriction sites that result in loci not being sequenced for some taxa and uncertain homology of short fragments with many nucleotide substitutions. Targeted capture and UCEs may suffer less from this specific problem depending on the capture design and probe set developed, which can be targeted to include loci with differing evolutionary rates. Furthermore, UCE probe sets can be developed for very specific questions [e.g., from genus- or even species- to order-level phylogenies (60)] and can be resequenced to generate compatible data sets for other lineages or to add new taxa to an existing lineage, similar to Sanger-based data sets.

The expansion of molecular systematics into the genomic era (137) may broaden the usefulness of traditionally preserved material, which represents the bulk of existing museum collections (150). Although RRGS methods are proving highly informative in phylogenetic analyses and may permit inclusion of traditionally preserved museum specimens, the bottleneck for many RRGS data sets is bioinformatic in nature, requiring the ability to manipulate and analyze the ever-growing data sets efficiently. Accordingly, students of entomology, ecology, and evolution need to be well trained in bioinformatics and big data analysis techniques.

Genome Sequencing, Comparative Genomics, and Gene Expression

The number of insect genomes and molecular data sets increases each year (**Figure 1**), providing the opportunity to ask novel questions about the evolution, physiology, gene function, and development of a diversity of nonmodel organisms. Although most genome approaches require specimens that have been preserved with nontraditional methods (**Table 1**), traditionally preserved museum insect specimens have been used for mitochondrial genome sequencing. Due to

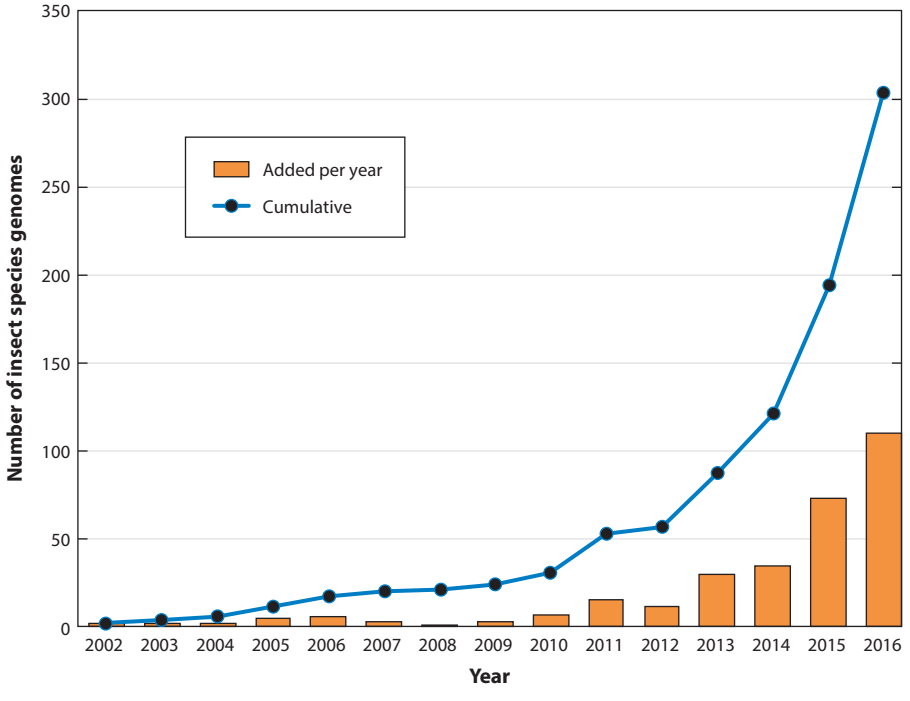

Figure 1

Number of insect species with genomes in the GenBank database from 2002 through 2016.

their smaller size and higher copy number per cell, mitochondrial genomes are ideal for capture from DNA obtained from museum specimens. Timmermans et al. (135) were able to sequence the mitochondrial genomes of 35 butterflies using museum specimens, and Staats et al. (125) sequenced the mitochondrial genome of two flies (*Aedes albopictus* and *Ceratitis capitata*) and one beetle (*Anoplophora glabripennis*) species from museum specimens. Although no entire nuclear genome has yet been sequenced from a traditionally preserved museum insect specimen due to the degraded nature of the DNA, this will likely change as sequencing technologies improve. As recovering mitochondrial genes and genomes is more likely to occur when museum specimens are used, Cameron (14) advocates for mitochondrial genome sequencing to serve as a bridge between legacy and contemporary data sets, given that these genomes are by-products of almost all high-throughput sequencing genome projects.

In fact, of the 304 insect genomes available in the National Center for Biotechnology Information (NCBI) GenBank (**http://www.ncbi.nlm.nih.gov/genbank**; accessed February 12, 2017), 46 species (20%) have at least one author with a museum institutional affiliation. Museum genomics is still in its infancy, but carefully designed practices for the preservation of genetic material have poised museum collections to take a place at the forefront of biodiversity genomics (30).

Microbiome: the collective microorganisms associated with a host or environment

Metagenomics: a technique in which DNA sequencing is used to study genetic material obtained from community sampling

GBIF: Global Biodiversity Information Facility; a public repository for species occurrence data

Host-Associated Microbiomes and Metagenomics

Microbes can have profound effects on their hosts, from negative interactions in the form of pathogenesis and parasitism to highly beneficial associations in the form of nutrient provisioning and immune defense (32, 63, 81, 90, 96, 128). Identifying which arthropods harbor microbes and characterizing the diversity of microbes that are specialized partners with their hosts are still at an early stage (21, 62, 65, 80, 103, 113, 116, 144). With large numbers of identified specimens, museum collections are likely to have a significant impact on elucidating host–microbe partnerships, because most amplicon sequencing of bacterial communities targets short stretches of microbial DNA and museum collections house millions of specimens that could be explored for their host-associated microbes.

In addition to determining the diversity of microbes present in or on a host, understanding the function of these microbiomes is key to comprehending their roles in host health and productivity. Sequencing the genomes of host-associated microbes can be highly informative but has historically required the ability to isolate or culture each species or strain (16, 26, 54, 91, 131). Metagenomics can serve as a tool to sequence the genomes of microbial partners found in a host or specific location in or on a host and has been informative for elucidating the functions of microbes across diverse insects and other arthropods (3, 34, 143). University and museum collections house the broadest sampling of insect diversity and are thus certain to make significant contributions toward understanding host-associated microbial interactions, although the full utility of such collections in this field has not been realized to date.

COLLECTION DIGITIZATION AND INFORMATICS

The size and complexity of entomological collections, and consequently the magnitude of the task in digitizing them, are unrivaled in natural history collections. For example, the federated database VertNet contains approximately 18 million digitized records for vertebrate specimens from more than 300 collections. The National Museum of Natural History, Smithsonian Institution, alone is estimated to contain >34 million insect specimens, of which only 421,698 records are available online through the GBIF (**http://www.gbif.org**; see also the collections survey in **Supplemental Materials**). The challenge is matched by the opportunities that the liberation of these

Supplemental Material

data can unlock. The diversity of research questions that can be addressed is enormous, and entomologists have only begun to unleash the power of these data [e.g., The Atlas of Living Australia (**http://www.ala.org.au/**)].

Darwin Core: a set of controlled vocabularies for biodiversity data and collections

Efforts to digitize collection objects—defined here to include data capture and/or imaging of either individual specimens or management units such as drawers, vials, and slides (and not simply lists of museum holdings)—now play a major role in the curation of many natural history museums. Digitization of entomological collections is now firmly in its third decade. However, it is only in the last five to ten years that the practice has expanded from early adopting collections to become fairly widespread in the community. At the same time, what it means to digitize collections has evolved and diversified, often depending on the objectives and available resources underpinning the effort.

Specimen-Level Data Capture

Traditional specimen-level data capture (SLDC) has been and remains the gold standard for the digitization of insect specimens. We define SLDC to include at a minimum the (*a*) physical labeling of individual specimens with a human- and/or machine-readable unique specimen identifier; (*b*) transcription of the specimen labels, including taxonomic identification, into a database; and (*c*) parsing of these digital data into appropriate data standards such as Darwin Core. Georeferencing, while not a necessity of basic SLDC, significantly increases the value of specimen records. When shared via publicly accessible web portals such as the GBIF, specimens and their associated data digitized using SLDC protocols become discoverable to the global scientific community and other stakeholders. These records not only can assist traditional users of museum specimens in discovering holdings of interest but also can permit research on environmental change, public health, invasive species, conservation, and a host of other topics (4). That being said, data cleanliness remains a significant issue inhibiting the application of biodiversity data that needs to be addressed (76).

Imaging-based approaches for semiautomated digitization of individual specimens or slides are also being developed. One example, the GIGAmacro Magnify^2 system (**http://www.gigamacro.com/gigapixel-macro-imaging-system/**), includes robotic automation, image capture, postprocessing, online viewing, sharing, and annotation. This system enables the user to set up 44 pinned or 84 slide-mounted specimens to be imaged in a single session. The position of each specimen or slide is registered by the operator in the included software, and the system will proceed unattended to take images of each specimen in the *X*-, *Y*-, and *Z*-axes. These images are first stacked and then stitched together, and the final image can have gigapixel resolution. Pinned specimens can then be rotated for automated imaging of other features in different views.

Digitizing in Bulk

In recent years, whole-drawer imaging (WDI) of entomological collections has become a growing trend. Rather than imaging or data capture of individual specimens, WDI involves creating an image of the entire specimen drawer (6, 49). While not a replacement for specimen-level digitization, WDI efforts have been driven by the increasing realization that there are no shortcuts to SLDC and the relatively quick, cheap, and unique benefits WDI provides (53). While, in principle, WDI can be done with a single photograph, such images would provide scant resolution for individual specimens. Consequently, composite images are created by stitching together many (often hundreds of) image tiles, each representing a small fraction of the drawer. These scalable composite images often allow even the smallest insects to be viewed with a relatively high level of detail and any label information that may be visible from above to be read.

Early approaches to whole-drawer digitization adapted existing technologies from panoramic photography such as GigaPan to image whole drawers (8). Here, a camera is suspended above

Audubon Core: controlled vocabularies for multimedia such as images and videos associated with biodiversity data

the object and swivels to take several images in the *X*- and *Y*-axes that are stitched together for a final, ultrahigh-resolution image of the entire object. The parallax effect, unfortunately, is still apparent, and images will be distorted and less accurate toward the edges. Other recent efforts such as SatScan Collections (10, 78), DScan (118), and InvertNet (28) have worked to build more customized hardware for automated digitization of an entire drawer containing specimens or a tray of slides. In SatScan and DScan, the camera is moved horizontally and takes several images in the *X*- and *Y*-axes of the entire object that are stitched together to provide a final, ultrahigh-resolution image of the entire drawer. The parallax effect of regular photography is eliminated in both systems. The SatScan integrated software allows the user to add metadata. In the robotic system developed by InvertNet (28), the camera is situated in the center and takes several hundred images in the *X*- and *Y*-axes of a drawer that are stitched together and result in a final, ultrahigh-resolution image of the entire object. The novelty of the InvertNet system is that the camera is also tilted to take overlapping images from various oblique angles. The individual pinned specimens can then be examined in dorsal and oblique views so that the label data become more easily readable. Through online transcription by amateurs or scientists, data about individual specimens in the image can be transcribed and added to a specimen-level database (52).

Various software packages have been developed to accommodate the needs of digitizing insect drawers. Hudson et al. (56) provide a modular, cross-platform suite of open-source software tools called Inselect to automatically identify and isolate specimens from an imaged drawer. Inselect allows the user to add metadata to the single-specimen images. While obtaining specimen-level data is an important goal, the label data and unique specimen identifier label might not be visible in a SatScan Collections or DScan drawer image. InvertNet developed its hardware and software suite to take the explicit specimen-level data capture into consideration by imaging specimens and labels in several angles. GIGAmacro, GigaPan, and InvertNet provide a custom image-viewing solution on their respective platforms, while whole-drawer images from SatScan Collections and DScan can be viewed in a variety of available viewers. The Australian National Insect Collection provides access to the whole-drawer images in Morphbank (**http://www.morphbank.net**). There are other custom imaging solutions available in natural history museums, such as pinned insects photographed on a rotating platform that can be viewed from all sides [ZooSphere (**http://www.zoosphere.net**)].

Field to Database: Digitization Beyond Specimens

Data associated with museum specimens are often not limited to what is printed on the labels. Original field notes, images of habitat, or recordings of sound or behavior are just a few of the rich resources that may be included in research collections. The existence of these materials is often unknown to other researchers, and they are typically less discoverable than the specimens themselves, if at all. The inclusion of these materials in digitization efforts was virtually unheard of a decade ago but is now increasing in prevalence. Some efforts, such as the Field Book Project (98), have focused on the digitization and annotation of original field notes, sketches, maps, and images. Others, such as the Collection Resources for Aquatic Coleoptera project (**http://creac.kubiodiversityinstitute.org/collections/**), emphasize not just the digitization of these materials but also their association with the specimens themselves (if they exist) in a single integrated platform. The unification of specimens with their original field notes and images facilitates the detection of labeling errors, more accurate georeferencing, and insights into the taxon's biology. The data standard Audubon Core (95) is available to facilitate the sharing of many kinds of these data.

Uses of Digitized Collections Data

Despite considerable SLDC efforts in the last 20 years, we estimate that fewer than 2% of insect specimens in museum collections have been digitized (see the collections survey in **Supplemental Materials**), highlighting both the magnitude of the task ahead and the enormous amount of data yet to be made publicly available. Despite the long road ahead toward complete digitization, the increase in quantity, quality, and availability of data from SLDC in the last decade has already substantially expanded the reach of collections data. Ecological niche modeling in particular has emerged as a powerful tool capable of utilizing the large troves of distributional data that specimens provide to detect past changes and predict future trends in the ranges of insects (e.g., 19). Bees have been a premier exemplar of how concerted efforts to digitize specimen data for a particular taxon can yield powerful insights into species distributions (20, 67), the effects of environmental change (15, 69), and evolution (120). Digitized field notes have even facilitated the discovery of insects once thought to be extinct (148).

Supplemental Material

Large-scale imaging of collections, most notably with WDI, have afforded taxonomists and the general public the ability to browse museum collections from their offices or homes. It allows taxonomists to discover holdings of interest that previously would have never been studied. For some taxa, it can facilitate preliminary sorting or identifications without the need to transport the material, reducing the risk to the specimens.

MORPHOLOGY: THE CUTICLE AND BEYOND

The morphological study of insects, which relies heavily on specimens in entomological collections, has seen dramatic changes over its more than 260-year history (9, 44). Probably the biggest boost to morphology outside of improved microscopes was the advent of digital, three-dimensional (3D) reconstructions of morphological features (25, 44). Such 3D data have been used for deciphering the phylogenetic relationships of insect taxa (e.g., 9, 41, 42, 43, 55, 102), obtaining additional data from amber and compression fossils (e.g., 1, 50, 101, 104, 117, 132), examining internal anatomy (e.g., 35, 40, 68, 86, 87, 105, 124, 152), investigating internal musculature used for power-amplified movements (e.g., 149), and describing new species in taxonomic revisions (e.g., 37, 84, 85, 127). However, Deans et al. (25, p. 328) highlight that "internal anatomy remains a largely untapped resource for evidence of taxonomic association and evolutionary history" and call for expanded use of modern techniques that provide access to internal features.

The raw data for 3D reconstruction and visualization originate from a diversity of methods such as X-ray computed tomography (CT), micro-CT (μCT), and CT using synchrotron devices, laser ablation tomography, confocal laser scanning microscopy (CLSM), and traditional semithin cross-sectioning of embedded objects followed by digitization and stacking of the individual slices (for an overview, see 44). While μCT, synchrotron devices, and CLSM are nondestructive methods, laser ablation tomography and thin sectioning are destructive methods. Specimens with thorough 3D reconstructions can be virtually dissected on a computer screen to reveal minute and internal structures, which is especially useful for the study of very small insects that cannot be dissected with traditional methods (e.g., 149). All of these new imaging advances capitalize on the vast existing holdings of collections, providing the ability to examine larger numbers of characters at levels of detail that were previously intractable. While specimens should be preserved in special ways for some applications (44), pinned and ethanol-preserved specimens that make up the vast majority of existing insect collections can be used to gather novel morphological data with the above techniques.

Three-dimensional reconstructions can play an important role in teaching and outreach by bringing the study of insects and detailed morphological analysis to students and the public at large through classroom and museum exhibits with interactive displays.

Microcomputed Tomography

In 2002, Hörnschemeyer et al. (55) published one of the first morphological phylogenetic studies using μCT data to elucidate the position of the cupedid beetle *Priacma serrata* within Archostemata (Coleoptera). Since then, the available resolution has been enhanced to 0.1 μm [nano-CT (9)]. With μCT and, to a lesser degree, nano-CT scanners becoming more easily accessible for research purposes, 3D reconstructions will become commonplace in taxonomic, systematic, and evolutionary studies in entomology in the near future. Because of its nondestructive nature, μCT scanning can advance the study of rarely collected or unique type specimens (9). It also works well for fossils and in particular for amber fossils because, while fossils can be scanned with synchrotron sources (132), the high energy used in synchrotrons can damage fossilized resin.

Confocal Laser Scanning Microscopy

For the study of external morphology, CLSM takes advantage of the autofluorescence of the cuticle (9, 70). Small insects or parts thereof can be studied easily, but internal structures of larger insects can be examined only after clearing the cuticle (9). In contrast to scanning electron microscopy, CLSM provides the opportunity for 3D reconstruction and visualization of small structures. While there is potential for signal loss artifacts, Klaus et al. (70) provide details on how to overcome them.

Managing and Sharing Data

Data standards are increasingly important for making the sharing and reuse of gathered morphological data a reality (140, 141). Media files in particular require associated metadata so that they can be stored in databases and made accessible to the community. Ontologies, defined and formalized vocabularies of terms and relationships (139, 151), similarly enhance the ability to apply previously gathered morphological data to questions in both entomological science and research fields outside of entomology. Balhoff et al. (5) provide an example of combining cybertaxonomic tools with annotating character and character-state combinations with the Hymenoptera Anatomy Ontology (151).

New imaging methods have necessitated the development of new and novel publishing tools. Authors are now able to not only include static representations of 3D reconstructions such as images in manuscripts but also embed videos and other media that can be interactively viewed by the reader in the article PDF (114, 115). In the published PDF, Faulwetter et al. (37) include videos derived from μCT data of the volume renderings of a polychaete worm for the virtual dissection and surface models for viewing a structure from all angles. The reader can interact with the video and study the virtual type material, called a cybertype by Faulwetter et al. (37), in the desired position. Ernst et al. (35), Mikó et al. (87), and Popovici et al. (105) provide access to the interactive volume renderings and surface models derived from CLSM through a data depository, and the individual videos are identified by digital object identifiers (DOIs) in article PDFs for easy retrieval.

Big data gathered through μCT and CLSM highlight the critical need for mechanisms for utilizing and sharing scientific data. Data depositories from which information can be retrieved and reanalyzed have a long history in the biological sciences; NCBI's GenBank is probably the most common depository used by entomologists. Guidance for locating a depository can be found in Whyte (147). Cranston et al. (23) provide ten simple rules on best practices for data sharing in

taxonomic and phylogenetic research. Just like other raw data, 2D (pixel) and 3D (voxel) imaging data should be made accessible and uploaded to dedicated, open-access depositories. Rowe & Frank (111) emphasize that 3D image data drastically lag behind in archiving and that at the time only two depositories, DigiMorph (**http://www.digimorph.org**) and Digital Fish Library (DFL), were suitable for archiving 3D voxels as raw data. This situation has not changed, unfortunately. Dedicated 2D image depositories such as Morphbank enable the user to add metadata, and several others such as MorphoBank (99) also enable the use of ontologies. General depositories such as Dryad (**http://www.datadryad.org**), Figshare, or Zenodo, all of which provide DOIs for each data set, can be used for archiving raw 3D data, volume renderings, and surface models so that they can be cited in a publication.

With the tools available to gather and store data, it is now incumbent upon the community to adopt the sharing of raw and edited data to stimulate vigorous scientific debate through testing previous hypotheses and erecting new hypotheses (109). GenBank has been successful for storing and providing access to molecular data for decades because it was a standard that was adopted by the community, and the morphology community needs to follow this example.

FUTURE ISSUES

1. In addition to traditional preservation methods, more museum collections are archiving entomological specimens with future genetic and genomic uses in mind. However, inconsistent data sharing currently inhibits the discovery and optimal utilization of these resources. Efforts to digitize and network these genomic resources, such as the Global Genome Biodiversity Network, would greatly benefit the community.
2. Recent and emerging genomic approaches are increasingly able to utilize historical museum specimens whose DNA may have been too fragmented for traditional Sanger sequencing methods in the past. While there have been many successes, there remains a need for deeper understanding and more accurate quantification of the factors that affect the suitability of genomic DNA from historical specimens.
3. With only 2% of entomological specimens estimated to have been individually digitized, new workflows for specimen-level data capture (SLDC) must be developed and/or additional resources must be allocated to increase the rate of data dissemination from entomological collections.
4. Digitization and integration of specimen-associated collecting data such as field notes, habitat images, and vocal recordings are likely to expand rapidly as these practices are adopted by more entomological collections. New protocols are needed to ensure these data are fully integrated with SLDC practices.
5. Current and emerging methods to digitally capture insect morphology both internally and externally provide the opportunity to establish new data sources for taxonomy and phylogeny. Making these data accessible through open depositories and using character ontologies can provide the basis for comparative evolutionary studies across diverse taxa that are not currently possible.
6. Because museum collections may in some cases contain the only record of a species or the specific geographic location of a species/population and new technologies are permitting broader uses of these collections, it is imperative that institutions and governments continue to invest in and support these invaluable archives of biodiversity.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank museum professionals, past, present, and future, around the world who oversee the care, maintenance, and growth of these invaluable scientific collections. This work was supported in part by National Science Foundation awards EF-1115051 and DEB-1453452 to A.E.Z.S. and awards DEB-1050243, DEB-1442316, and IOS-1354193 to C.S.M.

LITERATURE CITED

1. Arillo A, Peñalver E, Pérez-Delafuente R, Delclòs X, Criscione J, et al. 2015. Long-proboscid brachyceran flies in Cretaceous amber (Diptera: Stratiomyomorpha: Zhangsolvidae). *Syst. Entomol.* 40(1):242–67
2. Astrid T, Margit E, Leopold F. 2016. Ethanol: a simple and effective RNA-preservation for freshwater insects living in remote habitats. *Limnol. Oceanogr. Methods* 14:186–95
3. Aylward FO, Burnum KE, Scott JJ, Suen G, Tringe SG, et al. 2012. Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. *ISME J.* 6(9):1688–701
4. Baird R. 2010. Leveraging the fullest potential of scientific collections through digitization. *Biodivers. Inform.* 7:130–36
5. Balhoff JP, Mikó I, Yoder MJ, Mullins PL, Deans AR. 2013. A semantic model for species description applied to the ensign wasps (Hymenoptera: Evaniidae) of New Caledonia. *Syst. Biol.* 62(5):639–59
6. Balke M, Schmidt S, Hausmann A, Toussaint E, Bergsten J, et al. 2013. Biodiversity into your hands—a call for a virtual global natural history 'metacollection.' *Front. Zool.* 10:55
7. Bazinet AL, Cummings MP, Mitter KT, Mitter CW. 2013. Can RNA-seq resolve the rapid radiation of advanced moths and butterflies (Hexapoda: Lepidoptera: Apoditrysia)? An exploratory study. *PLOS ONE* 8:e82615
8. Bertone M, Blinn R, Stanfield T, Dew K, Seltmann K, Deans A. 2012. Results and insights from the NCSU Insect Museum GigaPan project. *ZooKeys* 209:115–32
9. Beutel RG, Friedrich F, Ge SQ, Yang XK. 2014. *Insect Morphology and Phylogeny*. Berlin: De Gruyter
10. Blagoderov V, Kitching I, Livermore L, Simonsen T, Smith V. 2012. No specimen left behind: industrial scale digitization of natural history collections. *ZooKeys* 209:133–46
11. Blaimer BB, Brady SG, Schultz TR, Lloyd MW, Fisher BL, Ward PS. 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evol. Biol.* 15:e271
12. Blanchard BD, Moreau CS. 2017. Defensive traits exhibit an evolutionary trade-off and drive diversification in ants. *Evolution* 71(2):315–28
13. Buerki S, Baker WJ. 2016. Collections-based research in the genomic era. *Biol. J. Linn. Soc.* 117:5–10
14. Cameron SL. 2014. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Syst. Entomol.* 39:400–11
15. Cameron SL, Lozier JD, Strange JP, Koch JB, Cordes N, et al. 2011. Patterns of widespread decline in North American bumble bees. *PNAS* 108:662–67
16. Campbell MA, Van Leuven JT, Meister RC, Carey KM, Simon C, McCutcheon JP. 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. *PNAS* 112:10192–99
17. Catzeflis F. 1991. Animal tissue collections for molecular genetics and systematics. *Trends Ecol. Evol.* 6:168
18. Chapco W, Litzenberger G. 2004. A DNA investigation into the mysterious disappearance of the Rocky Mountain grasshopper, mega-pest of the 1800s. *Mol. Phylogenetics Evol.* 30:810–14

19. Cho S, Epstein SW, Mitter K, Hamilton CA, Plotkin D, et al. 2016. Preserving and vouchering butterflies and moths for large-scale museum-based molecular research. *PeerJ* 4:e2160
20. Collins SD, McIntyre NE. 2015. Modeling and distribution of odonates: a review. *Freshw. Sci.* 34(3):1144–58
21. Colman DR, Toolson EC, Takacs-Vesbach CD. 2012. Do diet and taxonomy influence insect gut bacterial communities? *Mol. Ecol.* 21:5124–37
22. Corthals A, Desalle A. 2005. An application of tissue and DNA banking for genomics and conservation: the Ambrose Monell Cryo-Collection (AMCC). *Syst. Biol.* 54(5):819–23
23. Cranston K, Harmon LJ, O'Leary MA, Lisle C. 2014. Best practices for data sharing in phylogenetic research. *PLOS Curr.* 6:ecurrents.tol.bf01eff4a6b60ca4825c69293dc59645
24. Cruaud A, Gautier M, Galan M, Foucaud J, Sauné L, et al. 2014. Empirical assessment of RAD sequencing for interspecific phylogeny. *Mol. Biol. Evol.* 31:1272–74
25. Deans AR, Mikó I, Wipfler B, Friedrich F. 2012. Evolutionary phenomics and the emerging enlightenment of arthropod systematics. *Invertebr. Syst.* 26:323–30
26. Degnan PH, Lazarus AB, Wernegreen JJ. 2005. Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects. *Genome Res.* 15:1023–33
27. Dell'Ampio E, Meusemann K, Szucsich NU, Peters RS, Meyer B, et al. 2014. Decisive data sets in phylogenomics: lessons from studies on the phylogenetic relationships of primarily wingless insects. *Mol. Biol. Evol.* 31:239–49
28. Dietrich C, Hart J, Raila D, Ravaioli U, Sobh N, et al. 2012. InvertNet: a new paradigm for digital access to invertebrate collections. *ZooKeys* 209:165–81
29. DiEuliis D, Johnson KR, Morse SS, Schindel DE. 2016. Opinion: Specimen collections should have a much bigger role in infectious disease research and response. *PNAS* 113:4–7
30. Dikow RB, Frandsen PB, Turcatel M, Dikow T. 2017. Genomic and transcriptomic resources for assassin flies including the complete genome sequence of *Proctacanthus coquilletti* (Insecta: Diptera: Asilidae) and 16 representative transcriptomes. *PeerJ* 5:e2951
31. Dillon N, Austin AD, Bartowsky E. 1996. Comparison of preservation techniques for DNA extraction from hymenopterous insects. *Insect Mol. Biol.* 5:21–24
32. Douglas AE. 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Annu. Rev. Entomol.* 60:17–34
33. Droege G, Barker K, Seberg O, Coddington J, Benson E, et al. 2016. The Global Genome Biodiversity Network (GGBN) Data Standard specification. *Database* 2016:baw125
34. Engel P, Martinson VG, Moran NA. 2012. Functional diversity within the simple gut microbiota of the honey bee. *PNAS* 109:11002–7
35. Ernst A, Mikó I, Deans A. 2013. Morphology and function of the ovipositor mechanism in Ceraphronoidea (Hymenoptera, Apocrita). *J. Hymenopt. Res.* 33:25–61
36. Faircloth BC, Branstetter MG, White ND, Brady SG. 2014. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Mol. Ecol. Resour.* 15:489–501
37. Faulwetter S, Vasileiadou A, Kouratoras M, Dailianis T, Arvanitidis C. 2013. Micro-computed tomography: introducing new dimensions to taxonomy. *ZooKeys* 263:1–45
38. Ferro ML, Park JS. 2013. Effect of propylene glycol concentration on mid-term DNA preservation of Coleoptera. *Coleopt. Bull.* 67:581–86
39. Frampton M, Conrad S, Prager T, Richards MH. 2008. Evaluation of specimen preservatives for DNA analyses of bees. *J. Hymenopt. Res.* 17:195–200
40. Friedrich F, Beutel RG. 2008. The thorax of *Zorotypus* (Hexapoda, Zoraptera) and a new nomenclature for the musculature of Neoptera. *Arthropod Struct. Dev.* 37(1):29–54
41. Friedrich F, Beutel RG. 2010. Goodbye Halteria? The thoracic morphology of Endopterygota (Insecta) and its phylogenetic implications. *Cladistics* 26:579–612
42. Friedrich F, Beutel RG. 2010. The thoracic morphology of *Nannochorista* (Nannochoristidae) and its implications for the phylogeny of Mecoptera and Antliophora. *J. Zool. Syst. Evol. Res.* 48(1):50–74
43. Friedrich F, Farrell BD, Beutel RG. 2009. The thoracic morphology of Archostemata and the relationships of the extant suborders of Coleoptera (Hexapoda). *Cladistics* 25(1):1–37

44. Friedrich F, Matsumura Y, Pohl H, Bai M, Hörnschemeyer T, Beutel RG. 2014. Insect morphology in the age of phylogenomics: innovative techniques and its future role in systematics. *Entomol. Sci.* 17(1):1–24
45. Gilbert MTP, Moore W, Melchior L, Worobey M. 2007. DNA extraction from dry museum beetles without conferring external morphological damage. *PLOS ONE* 2:e272
46. Giribet G, Edgecombe GD. 2006. Conflict between data sets and phylogeny of centipedes: an analysis based on seven genes and morphology. *Proc. R. Soc. B* 273:531–38
47. Hammer TJ, Dickerson JC, Fierer N. 2015. Evidence-based recommendations on storing and handling specimens for analyses of insect microbiota. *PeerJ* 3:e1190
48. Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. B* 270:313–21
49. Heerlien M, van Leusen J, Schnorr S, de Jong-Kole S, Raes N, van Hulsen K. 2015. The natural history production line: an industrial approach to the digitization of scientific collections. *ACM J. Comput. Cult. Herit.* 8(1):3
50. Hendrickx H, Cnudde V, Masschaele B, Dierick M, Vlassenbroeck J, van Hoorebeke L. 2006. Description of a new fossil *Pseudogarypus* (Pseudoscorpiones: Pseudogarypidae) with the use of X-ray micro-CT to penetrate opaque amber. *Zootaxa* 1305:41–50
51. Hernández-Triana LM, Prosser SW, Rodríguez-Perez MA, Chaverri LG, Hebert PDN, Gregory TR. 2013. Recovery of DNA barcodes from blackfly museum specimens (Diptera: Simuliidae) using primer sets that target a variety of sequence lengths. *Mol. Ecol. Resour.* 14:508–18
52. Hill A, Guralnick R, Smith A, Sallans A, Gillespie R, et al. 2012. The notes from nature tool for unlocking biodiversity records from museum records through citizen science. *ZooKeys* 209:219–33
53. Holovachov O, Zatushevsky A, Shydlovsky I. 2014. Whole-drawer imaging of entomological collections: benefits, limitations and alternative applications. *J. Conserv. Mus. Stud.* 12(1):9
54. Hongoh Y, Sharma VK, Prakash T, Noda S, Toh H, et al. 2008. Genome of an endosymbiont coupling N_2 fixation to cellulolysis within protist cells in termite gut. *Science* 322:1108–9
55. Hörnschemeyer T, Beutel RG, Pasop F. 2002. Head structures of *Priacma serrata* Leconte (Coleptera, Archostemata) inferred from X-ray tomography. *J. Morphol.* 252(3):298–314
56. Hudson LN, Blagoderov V, Heaton A, Holtzhausen P, Livermore L, et al. 2015. Inselect: automating the digitization of natural history collections. *PLOS ONE* 10(11):e0143402
57. Hughes J, Longhorn SJ, Papadopoulou A, Theodorides K, de Riva A, et al. 2006. Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (beetles). *Mol. Biol. Evol.* 23:268–78
58. Hunt T, Bergsten J, Levkanicova Z, Papadopoulou A, John OS, et al. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* 318:1913–16
59. Hurst GDD, Jiggins FM. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. B* 272:1525–34
60. Ješovnik A, Sosa-Clavo J, Lloyd MW, Branstetter MG, Fernández F, Schultz TR. 2017. Phylogenomic species delimitation and host-symbiont coevolution in the fungus-farming ant genus *Sericomyrmex* Mayr (Hymenoptera: Formicidae): ultraconserved elements (UCEs) resolve a recent radiation. *Syst. Entomol.* 42(3):523–42
61. Johnson BR, Borowiec ML, Chiu JC, Lee EK, Atallah J, Ward PS. 2013. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. *Curr. Biol.* 23:2058–62
62. Jones RT, Sanchez LG, Fierer N. 2013. A cross-taxon analysis of insect-associated bacterial diversity. *PLOS ONE* 8:e61218
63. Kaltenpoth M, Engl T. 2013. Defensive microbial symbionts in Hymenoptera. *Funct. Ecol.* 28:315–27
64. Kanda K, Pflug JM, Sproul JS, Dasenko MA, Maddison DR. 2015. Successful recovery of nuclear protein-coding genes from small insects in museums using illumina sequencing. *PLOS ONE* 10:e0143929
65. Kautz S, Rubin BER, Russell JA, Moreau CS. 2013. Surveying the microbiome of ants: Comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. *Appl. Environ. Microbiol.* 79:525–34
66. Kawahara AY, Breinholt JW. 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proc. R. Soc. B* 281:20140970

67. Koch JB, Lozier J, Strange JP, Ikerd H, Griswold T, et al. 2015. US*Bombus*, a database of contemporary survey data for North American bumble bees (Hymenoptera, Apidae, *Bombus*) distributed in the United States. *Biodivers. Data J.* 2015(3):e6833
68. Koeth M, Friedrich F, Pohl H, Beutel RG. 2012. The thoracic skeleto-muscular system of *Mengenilla* (Strepsiptera: Mengenillidae) and its phylogenetic implications. *Arthropod Struct. Dev.* 41(4):323–35
69. Koh I, Lonsdorf EV, Williams NM, Brittain C, Isaacs R, et al. 2016. Modeling the status, trends, and impacts of wild bee abundance in the United States. *PNAS* 113:140–45
70. Klaus AV, Kulasekera VL, Schawaroch V. 2003. Three-dimensional visualization of insect morphology using confocal laser scanning microscopy. *J. Microsc.* 212(2):107–21
71. Krehenwinkel H, Pekar S. 2015. An analysis of factors affecting genotyping success from museum specimens reveals an increase of genetic and morphological variation during a historical range expansion of a European spider. *PLOS ONE* 10:e0136337
72. Letsch HO, Meusemann K, Wipfler B, Schütte K, Beutel R, Misof B. 2012. Insect phylogenomics: results, problems and the impact of matrix composition. *Proc. R. Soc. B* 279:3282–90
73. Letsch HO, Simon S. 2013. Insect phylogenomics: new insights on the relationships of lower neopteran orders (Polyneoptera). *Syst. Entomol.* 38:783–93
74. Maddison DR. 2016. The rapidly changing landscape of insect phylogenetics. *Curr. Opin. Insect Sci.* 18:77–82
75. Maddison DR, Cooper KW. 2014. Species delimitation in the ground beetle subgenus *Liocosmius* (Coleoptera: Carabidae: Bembidion), including standard and next-generation sequencing of museum specimens. *Zoolog. J. Linn. Soc.* 172:741–70
76. Maldonado C, Molina CI, Zizka A, Persson C, Taylor CM, et al. 2015. Estimating species diversity and distribution in the era of big data: To what extent can we trust public databases? *Glob. Ecol. Biogeogr.* 24:973–84
77. Mandrioli M. 2008. Insect collections and DNA analyses: how to manage collections? *Mus. Manag. Curatorship* 23:193–99
78. Mantle B, LaSalle J, Fisher N. 2012. Whole-drawer imaging for digital management and curation of a large entomological collection. *ZooKeys* 209:147–63
79. Martin G. 2006. The impact of frozen tissue and molecular collections on natural history museum collections. *NatSCA News* 10:31–47
80. Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. 2010. A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol. Ecol.* 20:619–28
81. McFall-Ngai M, Hadfield MG, Bosch TC, Carey HV, Domazet-Lošo T, et al. 2013. Animals in a bacterial world, a new imperative for the life sciences. *PNAS* 110:3229–36
82. Meier R, Shiyang K, Vaidya G, Ng PKL. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* 55(5):715–28
83. Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, et al. 2010. A phylogenomic approach to resolve the arthropod tree of life. *Mol. Biol. Evol.* 27:2451–64
84. Michalik P, Piacentini L, Lipke E, Ramírez M. 2013. The enigmatic Otway odd-clawed spider (Progradungula otwayensis Milledge, 1997, Gradungulidae, Araneae): natural history, first description of the female and micro-computed tomography of the male palpal organ. *ZooKeys* 335:101–12
85. Michalik P, Ramírez MJ. 2013. First description of the male of *Thaida chepu* Platnick, 1987 (Araneae, Austrochilidae) with micro-computed tomography of the palpal organ. *ZooKeys* 352:117–25
86. Mikó I, Friedrich F, Yoder MJ, Hines HM, Deitz LL, et al. 2012. On dorsal prothoracic appendages in treehoppers (Hemiptera: Membracidae) and the nature of morphological evidence. *PLOS ONE* 7(1):e30137
87. Mikó I, Masner L, Johannes E, Yoder MJ, Deans AR. 2013. Male terminalia of Ceraphronoidea: morphological diversity in an otherwise monotonous taxon. *Insect Syst. Evol.* 44(3–4):261–347
88. Miller J, Beentjes K, van Helsdingen P, IJland S. 2013. Which specimens from a museum collection will yield DNA barcodes? A time series study of spiders in alcohol. *ZooKeys* 365:245–61
89. Misof B, Liu S, Meusemann K, Peters RS, Donath A, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763–67

90. Moran NA. 2007. Symbiosis as an adaptive process and source of phenotypic complexity. *PNAS* 104(Suppl. 1):8627–33
91. Moran NA, Tran P, Gerardo NM. 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. *App. Environ. Microbiol.* 71:8802–10
92. Moreau CS, Bell CD. 2013. Testing the museum versus cradle biological diversity hypothesis: phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution* 67(8):2240–57
93. Moreau CS, Wray BD, Czekanski-Moir JE, Rubin BER. 2013. DNA preservation: a test of commonly used preservatives for insects. *Invertebr. Syst.* 27:81–86
94. Moritz C, Cicero C. 2004. DNA barcoding: promise and pitfalls. *PLOS Biol.* 2:e354
95. Morris RA, Barve V, Carausu M, Chavan V, Cuadra J, et al. 2013. Discovery and publishing of primary biodiversity data associated with multimedia resources: the Audubon Core strategies and approaches. *Biodivers. Inform.* 8(2):185–97
96. Moya A, Peretó J, Gil R, Latorre A. 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat. Rev. Genet.* 9:218–29
97. Nachman MW. 2013. Genomics and museum specimens. *Mol. Ecol.* 22:5966–68
98. Nakasone S, Sheffield C. 2013. Descriptive metadata for field books: methods and practices of the Field Book Project. *D-Lib Mag.* 19(11–12). **https://doi.org/10.1045/november2013-nakasone**
99. O'Leary MA, Kaufmann SG. 2011. MorphoBank: phylophenomics in the 'cloud.' *Cladistics* 27(5):529–37
100. Patrick HJH, Chomič A, Armstrong KF. 2016. Cooled propylene glycol as a pragmatic choice for preservation of DNA from remote field-collected Diptera for next-generation sequence analysis. *J. Econ. Entomol.* 109(3):1469–73
101. Penney D, Dierick M, Cnudde V, Masschaele B, Vlassenbroeck J, van Hoorebeke L. 2007. First fossil Micropholcommatidae (Araneae), imaged in Eocene Paris amber using X-ray computed tomography. *Zootaxa* 1623:47–53
102. Peters RS, Meusemann K, Petersen M, Mayer C, Wilbrandt J, et al. 2014. The evolutionary history of holometabolous insects inferred from transcriptome-based phylogeny and comprehensive morphological data. *BMC Evol. Biol.* 14(1):52
103. Pinto-Tomás AA, Sittenfeld A, Uribe-Lorío L, Chavarría F, Mora M, et al. 2011. Comparison of midgut bacterial diversity in tropical caterpillars (Lepidoptera: Saturniidae) fed on different diets. *Environ. Entomol.* 40:1111–22
104. Pohl H, Wipfler B, Grimaldi DA, Beckmann F, Beutel RG. 2010. Reconstructing the anatomy of the 42-million-year-old fossil *Mengea tertiaria* (Insecta, Strepsiptera). *Naturwissenschaften* 97:855–59
105. Popovici O, Mikó I, Seltmann K, Deans A. 2014. The maxillo-labial complex of *Sparasion* (Hymenoptera, Platygastroidea). *J. Hymenopt. Res.* 37:77–111
106. Post RJ, Flook PK, Millest AL. 1993. Methods for the preservation of insects for DNA studies. *Biochem. Syst. Ecol.* 21:85–92
107. Prendini L, Hanner R, DeSalle R. 2002. Obtaining, storing and archiving specimens and tissue samples for use in molecular studies. In *Techniques in Molecular Systematics and Evolution*, ed. R DeSalle, G Giribet, W Wheeler, pp. 176–248. Basel, Switz.: Springer
108. Quicke DLJ, Belshaw R, Lopez-Vaamonde C. 1999. Preservation of hymenopteran specimens for subsequent molecular and morphological study. *Zool. Scr.* 28:261–67
109. Raupach MJ, Amann R, Wheeler DQ, Roos C. 2016. The application of "-omics" technologies for the classification and identification of animals. *Org. Divers. Evol.* 16(1):1–12
110. Reiss RA, Schwert DP, Ashworth AC. 1995. Field preservation of Coleoptera for molecular genetic analyses. *Environ. Entomol.* 24:716–19
111. Rowe T, Frank LR. 2011. The disappearing third dimension. *Science* 331(6018):712–14
112. Rubin BER, Ree RH, Moreau CS. 2012. Inferring phylogenies from RAD sequence data. *PLOS ONE* 7(4):e33394
113. Russell JA, Moreau CS, Goldman-Huertas B, Fujiwara M, Lohman DJ, Pierce NE. 2009. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *PNAS* 106(50):21236–41
114. Ruthensteiner B, Baeumler N, Barnes DG. 2010. Interactive 3D volume rendering in biomedical publications. *Micron* 41(7):886.e17

115. Ruthensteiner B, Heß M. 2008. Embedding 3D models of biological specimens in PDF publications. *Microsc. Res. Tech.* 71(11):778–86
116. Sanders JG, Powell S, Kronauer DJC, Vasconcelos HL, Frederickson ME, Pierce NE. 2014. Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Mol. Ecol.* 23:1268–83
117. Schmidt J, Hoffmann H, Michalik P. 2016. Blind life in the Baltic amber forests: description of an eyeless species of the ground beetle genus *Trechus* Clairville, 1806 (Coleoptera: Carabidae: Trechini). *Zootaxa* 4083(3):431–43
118. Schmidt S, Balke M, Lafogler S. 2012. DScan—a high-performance digital scanning system for entomological collections. *ZooKeys* 209:183–91
119. Sharanowski BJ, Robbertse B, Walker J, Voss SR, Yoder R, et al. 2010. Expressed sequence tags reveal Proctotrupomorpha (minus Chalcidoidea) as sister to Aculeata (Hymenoptera: Insecta). *Mol. Phylogenetics Evol.* 57:101–12
120. Silva DP, Vilela B, De Marco P Jr., Nemésio A. 2014. Using ecological niche models and niche analyses to understand speciation patterns: the case of sister neotropical orchid bees. *PLOS ONE* 9(11):e113246
121. Simon S, Narechania A, DeSalle R, Hadrys H. 2012. Insect phylogenomics: exploring the source of incongruence using new transcriptomic data. *Genome Biol. Evol.* 4:1295–309
122. Smith M, Fisher B, Hebert P. 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philos. Trans. R. Soc. B* 360:1825
123. Song H, Buhay J, Whiting M, Crandall K. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *PNAS* 105:13486
124. Spangenberg R, Hünefeld F, Schneeberg K, Beutel RG. 2012. The male postabdomen and reproductive system of *Bibio marci* Linnaeus, 1758 (Hexapoda: Diptera: Bibionidae). *J. Zool. Syst. Evol. Res.* 50(4):264–88
125. Staats M, Erkens RHJ, van de Vossenberg B, Wieringa JJ, Kraaijeveld K, et al. 2013. Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLOS ONE* 8:e69189
126. Stein ED, White BP, Mazor RD, Miller PE, Pilgrim EM. 2013. Evaluating ethanol-based sample preservation to facilitate use of DNA barcoding in routine freshwater biomonitoring programs using benthic macroinvertebrates. *PLOS ONE* 8(1):e51273
127. Stoev P, Komerički A, Akkari N, Liu S, Xin Z, et al. 2013. *Eupolybothrus cavernicolus* Komerički & Stoev sp. n. (*Chilopoda*: *Lithobiomorpha*: *Lithobiidae*): the first eukaryotic species description combining transcriptomic, DNA barcoding and micro-CT imaging data. *Biodivers. Data J.* 1:e1013
128. Stouthamer R, Breeuwer JA, Hurst GD. 1999. *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* 53:71–102
129. Strutzenberger P, Brehm G, Fiedler K. 2012. DNA barcode sequencing from old type specimens as a tool in taxonomy: a case study in the diverse genus *Eois* (Lepidoptera: Geometridae). *PLOS ONE* 7:e49710
130. Suarez A, Tsutsui N. 2004. The value of museum collections for research and society. *BioScience* 54:66–74
131. Suen G, Scott JJ, Aylward FO, Adams SM, Tringe SG, et al. 2010. An insect herbivore microbiome with high plant biomass-degrading capacity. *PLOS Genet.* 6:e1001129
132. Tafforeau P, Boistel R, Boller E, Bravin A, Brunet M, et al. 2006. Applications of X-ray synchrotron microtomography for non-destructive 3D studies of paleontological specimens. *Appl. Phys. A* 83(2):195–202
133. Tagliavia M, Massa B, Albanese I, La Farina M. 2011. DNA extraction from Orthoptera museum specimens. *Anal. Lett.* 44:1058–62
134. Thomsen PF, Elias S, Gilbert M, Haile J, Munch K, et al. 2009. Non-destructive sampling of ancient insect DNA. *PLOS ONE* 4:e5048
135. Timmermans M, Viberg C, Martin G. 2016. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. *Biol. J. Linn. Soc.* 117:83–95
136. Tin MM-Y, Economo EP, Mikheyev AS. 2014. Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. *PLOS ONE* 9:e96793
137. Trautwein MD, Wiegmann BM, Beutel RG, Kjer KM, Yeates DK. 2012. Advances in insect phylogeny at the dawn of the postgenomic era. *Annu. Rev. Entomol.* 57:449–68

138. Vink CJ, Thomas SM, Paquin P, Hayashi CY, Hedin M. 2005. The effects of preservatives and temperatures on arachnid DNA. *Invertebr. Syst.* 19:99–104
139. Vogt L. 2009. The future role of bio-ontologies for developing a general data standard in biology: chance and challenge for zoo-morphology. *Zoomorphology* 128(3):201–17
140. Vogt L. 2013. eScience and the need for data standards in the life sciences: in pursuit of objectivity rather than truth. *Syst. Biodivers.* 11(3):257–70
141. Vogt L, Nickel M, Jenner RA, Deans AR. 2013. The need for data standards in zoomorphology. *J. Morphol.* 274(7):793–808
142. Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, et al. 2005. Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proc. R. Soc. B* 272:1577–86
143. Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, et al. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450:560–65
144. Weeks AR, Velten R, Stouthamer R. 2003. Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. *Proc. R. Soc. B* 270:1857–65
145. Wen J, Ickert-Bond SM, Appelhans MS, Dorr LJ, Funk VA. 2015. Collections-based systematics: opportunities and outlook for 2050. *J. Syst. Evol.* 53:477–88
146. Whitfield JB, Cameron S. 1994. Museum policies concerning specimen loans for molecular systematic research. *Mol. Phylogenet. Evol.* 3:268–78
147. Whyte A. 2015. *Where to keep research data: DCC checklist for evaluating data repositories*, version 1.1. Edinburgh, Scotl.: Digital Curation Centre. **http://www.dcc.ac.uk/resources/how-guides-checklists/where-keep-research-data**
148. Woller DA, Hill JG. 2015. *Melanoplus foxi* Hebard, 1923 (Orthoptera: Acrididae: Melanoplinae): rediscovered after almost 60 years using historical field notes connected to curated specimens. *Trans. Am. Entomol. Soc.* 141:545–74
149. Wood HM, Parkinson DY, Griswold CE, Gillespie RG, Elias DO. 2016. Repeated evolution of power-amplified predatory strikes in trap-jaw spiders. *Curr. Biol.* 26:1057–61
150. Yeates DK, Zwick A, Mikheyev AS. 2016. Museums are biobanks: unlocking the genetic potential of the three billion specimens in the world's biological collections. *Curr. Opin. Insect Sci.* 18:83–88
151. Yoder MJ, Mikó I, Seltmann KC, Bertone MA, Deans AR. 2010. A gross anatomy ontology for Hymenoptera. *PLOS ONE* 5(12):e15991
152. Zimmermann D, Randolf S, Metscher BD, Aspöck U. 2011. The function and phylogenetic implications of the tentorium in adult Neuroptera (Insecta). *Arthropod Struct. Dev.* 40(6):571–82