Mobilizing Mollusks: Status Update on Mollusk Collections in the U.S.A. and Canada

Authors: Sierwald, Petra, Bieler, Rüdiger, Shea, Elizabeth K., and Rosenberg, Gary

Source: American Malacological Bulletin, 36(2) : 177-214

Published By: American Malacological Society

URL: https://doi.org/10.4003/006.036.0202

Amer. Malac. Bull. **36**(2): 177–214 (2018)

Mobilizing mollusks: Status update on mollusk collections in the U.S.A. and Canada

Petra Sierwald[1], Rüdiger Bieler[1], Elizabeth K. Shea[2], and Gary Rosenberg[3]

[1]Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605-2496, U.S.A. psierwald@fieldmuseum.org; rbieler@fieldmuseum.org

[2]Delaware Museum of Natural History, 4840 Kennett Pike, Wilmington, Delaware 19807, U.S.A.

[3]Academy of Natural Sciences of Philadelphia, Drexel University, 1900 Benjamin Franklin Parkway, Philadelphia, Pennsylvania 19103, U.S.A.

Abstract: In 2017, a minimum of 8.5 million mollusk lots representing some 100 million specimens were held by 86 natural history collections in the U.S. (81) and Canada (5). Of these, 6.2 million lots representing 70 million specimens were cataloged (73%), another 2.3 million lots were considered quality backlog awaiting cataloguing, and 4.5 million lots (53% of the total) had undergone some form of data digitization. About 1.1 million (25%) of the digitized lots have been georeferenced, albeit with different approaches to accuracy and uncertainty. Fewer than 25% of collections, mainly larger ones, claim to be fully Darwin Core compliant. There are 35,000 primary type lots and 66,000 secondary type lots, representing 1.6% of cataloged lots. About 87% of lots are dry and 13% are fluid preserved, with less than 0.3% frozen. The majority of lots are gastropods (71%) and bivalves (26%). By habitat, 54% of lots are marine, 26% terrestrial, 19% freshwater, and 1% brackish. About 43% of marine and 57% of non-marine holdings are from North America including the Caribbean.

Solem (1975), in a previous survey of U.S. and Canadian malacological collections, reported 3.74 million lots of which 775,000 (21%) were uncataloged backlog, and suggested that backlog was growing at a faster rate than specimens were being cataloged. Since then the overall size of mollusk collections has grown by 227% and cataloged lots by 208%, but quality backlog has grown by 300%, confirming Solem's extrapolation. Solem noted that the eight largest collections held 78% of the lots, but in 2017 the eight largest (now with a slightly different composition) held only 63.5% of the lots, reflecting substantial growth of small and mid-sized collections, and the larger number of institutions that we surveyed. Solem reported a substantial gap between large collections (≥160,000 lots; AMNH, ANSP, BPBM, DMNH, FMNH, LACM, MCZ, UF, UMMZ, USNM) and mid-sized ones (35,000-75,000 lots; ChM, FWRI, Hefner, HMNS, SDNH, NCSM, SIO-BIC, UCM, UWBM, YPM), but seven collections now fall in the range of 76,000 to 160,000 (CM, BMSM, CASIZ, CMNML, INHS, OSUM, and SBMNH), and two have jumped to the large category (UF and DMNH).

Often overlooked is Solem's conclusion that mollusk collections in the United States and Canada are second only to insect collections for number of specimens, which is still true. Because there are far fewer species of mollusks than insects, mollusks have more specimens per species, averaging 1,100 in our survey, almost ten times what Solem reported for insects and approaching what he reported for fish. Bivalvia may have as many as 2,400 specimens/species, which makes them among the best-sampled classes of metazoans. The high number of specimens/species among mollusk and fish collection makes them well-suited for environmental studies that track faunal change over time and space.

Key words: biodiversity, collection management, databases, digitization, Mollusca

Mollusks represent the second largest phylum in the animal kingdom, one that contains extraordinary ecological diversity, spanning terrestrial, freshwater, and marine environments, and has a fossil record dating back to the Cambrian. Formally and permanently accessioned mollusks in institutional collections constitute a rich library of morphological and genetic diversity and provide baseline data of the group's distribution in time and space. As such, they contribute to an enormous range of research fields, from evolutionary history of life forms, to the occurrence and abundance and management needs of species, shifting of distribution ranges (including fisheries and pest species), and changing attributes (*e.g.* body size) over time. High quality molluscan specimen data contained in natural history collections provide a foundation for environmental monitoring of all human-impacted habitats.

The ecological and economic importance of North American specimen collections can only be fully assessed and harnessed if the data are accessible in meaningful and comparable ways. Traditionally, taxon-specific publications have reported on the scope of individual museum collections or type material (*e.g.* Bieler and Bradford 1991) and that practice continues today across taxa (*e.g.* Ciubuc 2017). Individual publications are an important way to annotate collections and holdings, but inefficient for providing wide access to collections' information. Ariño (2010) estimated 3% of the

possible 2.1 billion natural history collection lots of all taxa were available through GBIF. GBIF (accessed in June 2018) lists 156,000,000 specimen records (11 million of which are mollusks), so the figure has grown to perhaps 8%, but clearly there is still far to go in digitizing collections. Having a realistic sense of the scope of collections is essential to planning for efficient digitization, data and specimen management, and to promote data and specimen usage. Digitizing metadata about collections may be a global first step (Berendsohn and Seltmann 2010, Scoble 2010, Schindel *et al.* 2016), but local collections must publish their holdings as quickly as time and resources allow.

The curators and collections managers responsible for mollusk collections have made several significant attempts to understand and document the size and scope of their holdings. In response to the Association of Systematics Collection's National Plan (Irwin *et al.* 1973), Field Museum curator Alan Solem published a seminal work on the state of the U.S. and Canadian mollusk collections (Solem 1975). He surveyed 125 institutions and 100 private collectors and provided a synopsis of the data from 45 mollusk collections and 50 collectors from the U.S. and Canada. Nineteen mollusk collections with fewer than 5,000 lots were excluded from his analysis. He calculated that 78% of all molluscan holdings were contained within eight institutions and argued that supporting these collections would have the maximum benefit to molluscan research. Since then, additional compilations of institutions with type specimens (Kabat and Boss 1992, 1997) or important holdings (Sturm 2006) have been published, and an extensive list of worldwide mollusk collections with contact information and summary collections data has been maintained online by Cummings *et al.* (last updated 2009). An additional comprehensive resource was provided by Coan and Kabat (2018), who compiled biographical and bibliographical publications for more than 10,000 malacologists and other individuals with an interest in and relevant contributions to mollusks' natural history and distribution. However, there has not been another comprehensive survey of mollusk collection holdings and their scope in over 40 years.

Collection management of natural history collections has changed fundamentally over the past decades, incorporating advances in archival storage materials and techniques, digitization of text data and images, and global gathering and sharing of specimen and metadata information via the Internet. The rapid development and adoption of such approaches in mollusk collections is demonstrated in the "Standards for Malacological Collections," developed and published by Solem *et al.* (1981) for the North American Council of Systematic Malacologists. Focus therein was on the physical well-being of the collections (proper storage of material for future *morphological* study) and the *local* availability of specimen and collecting event data.

Online accessibility of specimen records now allows harvesting locality data that can be used in a growing and ever-developing range of research fields, such as biogeography, species range shifts, niche modeling, environmental monitoring, and conservation research, as well as documenting spatial, temporal, and taxonomic collecting gaps. Such data mining is greatly enhanced by data aggregators (*e.g.* GBIF) and unified collection portals, such as iDigBio (idigbio.org), InvertEBase (invertebase.org), and SCAN (scan-bugs.org). Increasing data quality (*e.g.* through improved georeferencing), data scope (*e.g.* by adding 2D and 3D images), and specimen attributes (*e.g.* documenting host-parasite associations) forms the foundation of a new range of specimen-based research activities (*e.g.* see Digital Data in Biodiversity Research Conference series, organized by iDigBio).

Here, we report on the results of a new survey of United States and Canadian mollusk collections that was conceived and initiated prior to the Molluscan Digitization Workshop at the 2017 American Malacological Society meeting (Shea *et al.* 2018, this volume). This survey revisited some of the same questions that Solem (1975) addressed and investigated new issues including georeferencing and moving collections data onto the web. Importantly, this survey also focused on finding and including smaller, lesser-known, and "hidden" collections to get a more complete understanding of the scope of molluscan holdings in the United States and Canada (documented in Appendix 2). The institutions surveyed are listed in Table 1. The results provide new insights into the complex landscape of natural history holdings and will help prioritize and maximize limited resources to improve the care of, access to, and research use of mollusk collections.

BACKGROUND ON MOLLUSK COLLECTIONS

The structure and nature of molluscan (malacological) collections reflect the specific physical attributes of the phylum Mollusca, the species-richness (Table 2) and unique characteristics of each included group, their collection-forming history, the advancement of preservation techniques (Appendix 4), and the ever-increasing research use and research techniques applied to these collections. Perhaps more so than most other groups of organisms in collections, mollusk collections have a history of contributions by amateur collectors. In addition to major collecting efforts by researchers and government agencies, Solem (1975: 223) estimated that 85% of the mollusks in major institutional collections were collected by amateurs. These specimens often were (and are) of very high quality and with good locality data but may be biased towards large and attractive shells. In addition, such material from private collections consists predominantly of dry shells, without tissues suitable for anatomical

Table 1. Surveyed mollusk collections – List of U.S. and Canadian mollusk collections, in alphabetical order of museum or collection identifier. All contacted collections are listed. Museum identifiers are those the institutions currently prefer and may differ from acronyms or identifiers used in other listings. Column 1975 shows collections surveyed by Solem (1975). Column 1996/2009: indicates collections included in Cummings *et al.* 2009 (latest partial updates are from 2009). Museum identifiers used by Solem and Cummings *et al.* 2009, if different from this list, are given in the respective columns. Column 2017: shows the current survey. Notations: $^{E\,NA}$ = data limited to eastern North America; LD = limited data provided (these collections are included in subsequent tables and appendices only when sufficient data are available); NC = no mollusk collection present; [F] = indicates fossil holdings; [R] indicates Recent holdings in largely paleontological mollusk collections.

Collection	Institution's name, city, state/province	1975	1996/2009	2017
AMNH	American Museum of Natural History, New York, NY	X	X	LD
ANSP	Academy of Natural Sciences of Philadelphia, Drexel University, Philadelphia, PA	X	X	X
ARC	Atlantic Reference Center, Huntsman Marine Science Center, St. Andrews, New Brunswick, Canada		X	X
ARK	University of Arkansas, Fayetteville, AR		X	X
AUMNH	Auburn University, Museum of Natural History, Auburn, AL			X
Beneski Museum	Beneski Museum of Natural History, Amherst College, Amherst, MA; see Appendix 2			LD
Berkshire	Berkshire Museum, Pittsfield, MA, see Appendix 2			LD
BMS	Buffalo Museum of Science, Buffalo, NY, see Appendix 2			LD
BMSM	Bailey Matthews National Shell Museum, Sanibel, FL		X	X
BPBM	Bernice P. Bishop Museum, Honolulu, HI	X	X	X
Brazosport	Brazosport Museum of Natural Science, Clute, TX; see Appendix 2			LD
CASIZ	California Academy of Sciences, San Francisco, CA	X	X	X
SU	Stanford University, Department of Geology, collection transferred to CAS	X		
CASPNNM	Chicago Academy of Sciences, Peggy Notebaert Nature Museum, Chicago, IL		X	X
ChM	Charleston Museum, Invertebrate Collection, Charleston, SC	X CMC	X	LD
CLEV	Cleveland Museum of Natural History, Cleveland, OH; mollusk collection partially transferred to FMNH in 2017		X	X
CM	Carnegie Museum of Natural History, Pittsburgh, PA	X	X	X
CMC	Cincinnati Museum Center, Museum of Natural History and Science, Cincinnati, OH		X	LD
CMNML	Canadian Museum of Nature, Mollusk Collection, Ottawa, Canada	X NMC	X	X
DMF	Daniel M. Fisk Museum of Natural History, Hillsdale College, Hillsdale, MI, see Appendix 2			LD
DMNH	Delaware Museum of Natural History, Wilmington, DE	X	X	X
DMNH-P	Perot Museum of Nature and Science; formerly Dallas Museum of Natural History, Dallas, TX	X	X	X
DMNS	Denver Museum of Nature and Science, Denver, CO		X	X
EKY	Eastern Kentucky University, Richmond, KY	X	X EKU	LD
Everhart	Everhart Museum of Natural History, Science and Art, Scranton, PA, see Appendix 2			LD
FMNH	Field Museum of Natural History, Chicago, IL	X	X	X
FWRI	Fish and Wildlife Research Institute, Specimen Information Services [formerly Florida Marine Research Institute], St. Petersburg, FL		X	X
FWM	Fort Worth Museum of Science and History, Fort Worth, TX	X	X	LD
GTMC-GMNH	University of Georgia, Museum of Natural History, Athens, GA		X	LD
HBOM $^{E\,NA}$	Harbor Branch Oceanographic Institution, Museum, Florida Atlantic University, Fort Pierce, FL; $^{E\,NA}$ data limited to eastern North America			X
Hefner	Hefner Zoology Museum, Miami University, Miami, OH		X	X

Table 1. (Continued)

Collection	Institution's name, city, state/province	1975	1996/2009	2017
HMNS	Houston Museum of Natural Sciences, Houston, TX		X	X
INHS	Illinois Natural History Survey, Campaign, IL	X	X	X
UIMNH	University of Illinois Museum of Natural History, collection merged with INHS in May 2008	X UINH	X	
INSM	Indiana State Museum, Indianapolis, IN		X	X
WMI	The New Harmony Workingmen's Institute, New Harmony, IN, curated by INSM		X	X
ISM	Illinois State Museum, Springfield, IL		X	LD
JFBM	Bell Museum of Natural History, St. Paul, MN		X	X
LACM	The Natural History Museum of Los Angeles County, Los Angeles, CA	X	X	X
LSUMG-I	Louisiana State University Museum of Natural Sciences, Invertebrate Paleontology Collection, Baton Rouge, LA			X
MCPR	James F. Matthews Center for Biodiversity Studies, Charlotte, NC			X
MCZ	Museum of Comparative Zoology, Harvard University, Cambridge, MA	X	X	X
MLBeanLSM	Monte L. Bean Life Science Museum, Brigham Young University, Provo, UT			LD
MMNHC	Frank McClung Museum of Natural History and Culture, University of Tennessee, Knoxville, TN		X UT	X
MMNS	Mississippi Museum of Natural Sciences, Jackson, MS		X	X
MPM	Milwaukee Public Museum, Milwaukee, WI	X	X	X
MUWV	Marshall University, Huntington, WV		X	NC
NCSM	North Carolina State Museum of Natural Sciences, Raleigh, NC		X	X
NFM	The Rooms Provincial Museum, St. John's, Newfoundland, Canada		X	LD
NHSM	Natural History Society of Maryland, Baltimore, MD			LD
NJSM	New Jersey State Museum, Trenton, NJ; see Appendix 2			LD
NYSM	New York State Museum, Albany, NY		X	X
OGL	Ocean Genome Legacy, Northeastern University, Boston, MA, maintains DNA extracts			X
OMNH	Sam Noble Museum, Norman, OK			X
OSUM	Ohio State University Museum, Museum of Biological Diversity, Columbus, OH	X OSU	X OSM	X
PRI	Paleontological Research Institution, Ithaca, NY		X	X
RBCM-INVZ	The Royal British Columbia Museum, Invertebrate Zoology, Victoria, BC, Canada	X UBC	X BCPM	X
RMUW	Richter Museum, University of Wisconsin-Green Bay, WI			X
ROM	Royal Ontario Museum, Toronto, Canada		X	X
RSMAS	Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL		X	X
SBMNH	Santa Barbara Museum of Natural History, Santa Barbara, CA	X	X	X
SCSM [F]	South Carolina State Museum, Columbia, SC			LD
SDNH	San Diego Natural History Museum, San Diego, CA	X	X	LD
SFCMC	South Florida Collections Management Center, Homestead, FL			LD
SIO-BIC	Scripps Institution of Oceanography, Benthic Invertebrate Collection, La Jolla, CA		X	X
SIO-PIC	Scripps Institution of Oceanography, Pelagic Invertebrate Collection, University of California, San Diego, CA			X
SMM	Science Museum of Minnesota, Saint Paul, MN, see Appendix 2		X	LD
SMNC	Schiele Museum of Natural History, Malacology Collection, Gastonia, NC			X

Table 1. (Continued)

Collection	Institution's name, city, state/province	1975	1996/2009	2017
Stevens Point	University of Wisconsin - Stevens Point, Stevens Point, WI			NC
SUI	University of Iowa Museum of Natural History and Dept. of Geology, Iowa City, IA [formerly the State University of Iowa]		X	X
TNHC	Texas Natural Science Center, Austin, TX (fish collection), see Appendix 2		X	LD
UAM	Museum of the North, University of Alaska, Fairbanks, AK		X	X
UAZ	Invertebrate Museum, University of Arizona, Tucson, AR	X UAT	X UAZ	X
UCM	University of Colorado Museum, Boulder, CO	X	X	X
UCMP	University of California, Museum of Paleontology, Berkeley, CA		X	X
UF	University of Florida Museum of Natural History, Gainesville, FL, formerly FLMNH	X FSM	X	X
UMAMZ	University of Massachusetts Museum of Zoology, Amherst, MA		X	LD
UMMZ	University of Michigan Museum of Zoology, Ann Arbor, MI	X	X	X
UMNH	Natural History Museum of Utah, Salt Lake City, UT			X
UNM(MSB)	Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM; see Appendix 2			LD
UNM(MSB) [P]	Museum of Southwestern Biology, Division of Parasites, University of New Mexico, Albuquerque, NM			X
UNSM	University of Nebraska State Museum, Division of Zoology, Lincoln, NE	X	X UN	X
UPRM (INVCOL)	University of Puerto Rico, Mayagüez, PR, see Appendix 2			LD
USDA-Aphis	USDA Aphis, at ANSP, Philadelphia, PA			X
USNM	National Museum of Natural History, Washington, DC	X NMNH	X	X
UTEP	Biodiversity Collections, University of Texas at El Paso, El Paso, TX	X	X	X
UWBM	Burke Museum of Natural History and Culture, University of Washington, Seattle, WA	X WSM	X UW	X
UWZM	University of Wisconsin, Zoological Museum, Madison, WI		X	X
VIMS	Virginia Institute of Marine Science, Gloucester Point, VA			LD
VMNH	Virginia Museum of Natural History, Martinsville, VA		X	X
YPM	Yale Peabody Museum, New Haven, CT		X	X

and molecular study. An outstanding example of a private collection absorbed by a U.S. museum is Leslie Hubricht's collection of about 500,000 specimens in 43,000 lots of eastern U.S. land snails that forms the backbone of FMNH's North American land snail collection (Solem 1986, Gerber 2010). Molluscan collections cover a wide range of specimen sizes, from microscopic snails to giant squid, and preservation types including dry shells, fluid preserved bodies, fossil material, and other derivative materials, *e.g.* dissected specimens and histological preparations on microscope slides. With increasing focus toward modern research applications, the diversity of preservation techniques and concomitant storage needs have evolved since the 1970s to include cryogenic facilities and electron microscopy mounts.

- The largest and oldest component of mollusk collections are the dry shell collections, predominantly of gastropods and bivalves but also scaphopods and polyplacophorans and the occasional shelled cephalopod. Most collections hold predominantly dry material (Appendix 4), which is arranged in systematic order, according to one or more higher-level taxon treatments (*e.g.* WoRMS and MolluscaBase). Within each family, organization generally is alphabetical or geographic, but this may vary by size or local needs and interest.
- Material of shell-less or largely soft-bodied groups (*e.g.* cephalopods, aplacophorans, nudibranchs, terrestrial slugs) is usually fluid preserved and often fixed in formalin. Some specialized collection (*e.g.* ARC, UNM(MSB)[Parasites], SIO-PIC) are essentially entirely wet-preserved. The final storage solution usually is 70-80% ethanol. Fluid-preserved specimens are often stored in numerical order to save space but can be stored in systematic order. Various protocols have been followed in tissue fixation (Roper and Sweeney

Table 2. Average number of lots per molluscan species by class. The number of lots held for each class is divided by the number of species in that classes as a reflection of taxonomic coverage across the surveyed collections. Data on number of marine species is from MolluscaBase; data on non-marine is from Rosenberg (2014) updated with recently described species from MolluscaBase.

	Gastropoda	Bivalvia	Cephalopoda	Scaphopoda	Polyplacophora	Aplacophora	Total
MolluscaBase (2018)							
marine + brackish	36,744	8,435	810	576	1,023	424	48,012
terrestrial since 2011	964	-	-	-	-	-	964
freshwater since 2011	334	24	-	-	-	-	358
Rosenberg (2014)							
terrestrial to 2012	24,380	-	-	-	-	-	24,380
freshwater to 2012	3,900	1,200	-	-	-	-	5,100
Total accepted species 2018	66,322	9,659	810	576	1,023	424	78,814
Percentage of lots (from App. 5)	71.2%	26.1%	1.2%	0.4%	1.0%	0.1%	100%
Lots/species (of catalogued)	66	167	93	47	60	14	79
Lots/species (including backlog)	91	230	128	65	82	19	108

1983, see papers cited in Sturm *et al.* 2006). Material intended for anatomical study, especially of marine mollusks, often underwent fixation in buffered formalin or Bouin's solution (especially for histological investigations) before transfer to alcohol. Specialized histological techniques introduced additional fixatives (Howard *et al.* 2004). A formal record of such fixation becomes an important part of specimen metadata as it will influence the tissue selection for successful anatomical and molecular approaches. With the advent of molecular component extraction and analyses, preservation of soft tissue increased significantly, with storage in high-percentage ethanol without prior chemical fixation, in nucleic acid preservation buffer fluids, or direct freezing in ultracold freezers or liquid nitrogen. All glass- and plasticware as well as labels need to be of archival quality, selected to handle chemical and/or low-temperature exposure to ensure long-lasting preservation of the material.

- Primary types (holotypes, syntypes, lectotypes, neotypes) are concentrated in the larger, older collections, especially USNM, ANSP, and MCZ (documented in Appendix 3). These collections house material from the early phase of documenting North American molluscan diversity, going back to authors such as William Dall (1845-1927) and Henry Pilsbry (1862-1957). Large numbers of paratypes exist in other collections (*e.g.* DMNH) where acquiring specimens has been emphasized over describing new species. Accumulation of secondary types may also result from more recent international collecting practices whereby primary types are deposited in the host country and secondary types deposited in additional museums (*e.g.* Solem's extensive land snail work in Australasia, with many secondary types deposited at FMNH). Type collections are often housed separately from the main collection (*e.g.* USNM), but can be integrated with it in systematic order (*e.g.* ANSP).
- Fossil and Recent mollusks are traditionally separated into different organizational units (*e.g.* invertebrate paleontology vs. zoology) within collection-holding institutions. Pleistocene and subfossil Holocene (*e.g.* loess) material, particularly of species also known from the Recent is often included as part of the "Recent" collection unit.
- A multitude of additional preparation types exist in mollusk collections. Preserved egg masses and radula slides have a long history in the field, while scanning electron microscope mounts and frozen tissue samples are relatively new. Some institutions preserve associated parasites (*e.g.* UNM(MSB)), and others preserve hosts parasitized by mollusks (*e.g.* ANSP). Extensive holdings of field photographs of living animals (*e.g.* of deep-sea cephalopods, DMNH), digital specimen photography, x-ray, and CT scanning have added new layers of virtual collections and will likely grow in the future.

Curation Status

Recent decades have seen substantive improvements in standards for the archival care necessary to assure the long-term integrity of calcareous shells and associated soft-bodies and tissues. Molluscan shells are known to be susceptible to so-called Bynesian decay, an efflorescence triggered by acid vapors from wood or paper materials that can destroy shell surfaces (*e.g.* Tennent and Baird 1985). Acidic wood material, specimen boxes, label paper, and organic cotton should be replaced by acid-free archival-quality materials (or non-archival materials should be isolated from direct specimen

contact). Most North American collections are in various stages of this shift (*e.g.* from wooden to metal drawers and cabinets) as staffing and funding allow, and many collection staff members indicated a pressing need for such re-curation in the survey questionnaires and at the 2017 Mollusk Digitization Workshop.

Digitization

Mollusk collections today are moving rapidly from a hand-written "ledger and label" system to being digitized in a variety of data management systems. The term "digitization" as used here encompasses any specimen data capture in digital form regardless of software platform: from word processor and spreadsheet flat files to relational databases. Most mollusk collections started their transition by entering ledger and label information (specimen identification and locality data) into a spreadsheet or database. Collections that began digitizing in the 1980's (*e.g.* DMNH) may have a digital record for most of their collection, but amount and scope of data captured varied over time. Thus these records may now be considered incomplete (skeleton data) or are non-normalized due to file-size or field-length constraints of early databases. Although labor-intensive, digitization makes day-to-day operations such as loan transactions, printing labels, and updating taxonomy more efficient, and broadens their availability for research use. A few collections (*e.g.* DMF) digitize accessions, be they lots or specimens, rather than cataloguing lots. The type of captured data in such accessions may be rather variable and not comparable to digitization of cataloged lots. Recently, significant attention has been paid to mining online specimen data for occurrence records and traits, opening up a new field of data analysis based on the online accessibility of specimen data (Beaman and Cellinese 2012, Ball-Damerow *et al.* 2014).

Georeferencing

The development of the Internet provided an opportunity for serving data online and exposing individual databases to much larger potential userbases through institutional websites and data aggregators such as iDigBio and GBIF. Adding images to specimen records represented another milestone, as did the inclusion of mappable geodetic coordinates, known as georeferencing. Over time, georeferencing has evolved from adding rough map coordinates to detailed point data aided by GPS units in the field and modern online tools such as Google Maps, GEOLocate and other specialized georeferencing software.

SURVEY PROCEDURE

A survey questionnaire (see Appendix 1) was distributed (by RB) in March 2017 as a Microsoft Excel worksheet to 80 known or expected institutional collections of extant mollusks in the U.S. and Canada. We targeted collections listed by Solem (1975), those identified in Kabat and Boss (1992, 1997) as having molluscan type holdings, and those in the online listing by Cummings *et al.* (2009). The survey focused on extant holdings of formal institutional collections; exclusively paleontological collections or private collections were not included. After the initial survey, some additional collections came to our attention and were sent the survey individually. For some of these mostly smaller collections, data were collected and added as late as June 2018. In total, 60 (70%) of the 88 collections contacted provided full or partial data (Table 1); 27 collections provided limited data, and one collection (Marshall University) has been closed (V. Fet in lit.).

Intensive efforts were made to obtain comparable data for collection sizes measured in cataloged lots, digitized lots, and quality backlog. We often sent follow up questions to individual respondents to clarify ambiguities, such as reporting specimen numbers rather than lot numbers or accession numbers rather than cataloged lots; including backlog in estimates of proportion of holdings by taxonomy, geography, habitat or preservation; or providing such proportions only for digitized parts of the collection rather than for all cataloged lots. Some data in the Tables and Appendices are supplemented from other sources, especially from Cummings *et al.* (2009), and institutional websites. Nonetheless, some inconsistencies remain across the figures reported in the tables and are flagged with superscripts in Appendices 3-10.

Although we made every effort to find and include all known mollusk collections in the United States, we have undoubtedly missed some collections, misinterpreted some free-style responses, and were not able to obtain detailed records from some collections in time for this analysis. We hope that we will be able to include these additional, under-documented, or undiscovered collections in future treatments.

Data compilation

The collections surveyed were in various stages of curation and digitization (see definition of digitization above), and respondents often replied with educated guesses and estimates, frequently in narrative form (*e.g.* "at least 20% of our holdings are marine", see Appendices 3-10). To standardize the data across collections, some interpretation and recalculation was necessary to turn narrative responses and estimates into comparable numbers. While necessary, this approach may have under- or overestimated collections' holdings and there are surely errors in the tables. We maintained the data in Microsoft Excel, with two of us (PS and GR) independently compiling data in the tables. Totals and other statistics were calculated in versions of the tables that had non-numerical characters stripped out, which means that indications such as

greater than or less than signs were omitted and assumed to average out across collections. Where a range was given, we calculated based on the average of that range. We have done our best to be conservative in our estimates and minimize errors; however, we apologize for any misinterpretations of institutional data that we might have introduced and will be happy to update our dataset in response to comments. Given the large number of institutions surveyed, such errors should average out, so our overall conclusions should be reliable. The presence of backlog material and the difference between counting individuals, specimens, and lots may further affect the size estimates.

Backlog

Specimens that have not been formally evaluated or added to a collection are often referred to as backlog. Depending on individual collection management practices, this might include anything from a small quantity of research specimens of a recognized authority, to an orphaned institutional collection in need of specialist taxonomic review, to containers of mixed shells collected during a sampling expedition. These scenarios are separated by the degree to which they have been physically and academically curated.

We adopt the term "**quality backlog**" to describe specimen lots that have good locality data and confident identifications that are ready for digitization with minimal physical curation necessary.

We adopt the term "**deep backlog**" to describe materials that have locality data but are either not sorted into lots or are not (or poorly) identified, and require considerable physical curation prior to digitization.

The numbers reported here are for quality backlog; however, the deep backlog in some collections approaches or even greatly exceeds the number of previously cataloged lots in the respective collection (*e.g.* CASIZ, FWRI, SBMNH, UNSM, YPM. see Appendix 3).

Individuals, specimens, lots, and records: In principle, it is desirable to count the total number of individuals in a collection; however, the variety of preparation types employed to preserve soft bodies and hard shells makes this goal difficult to achieve. Mollusk collections manage four different unit concepts: an individual, a specimen, a lot, and a record.

- An individual. An individual is a single, whole organism. In mollusks collections, the individual may be represented by a single body (*e.g.* cephalopods, aplacophorans); a single shell with or without a body (gastropods, scaphopods); by two or more articulated shell valves, with or without a body (bivalves, chitons); or by disarticulated shell valves (bivalves, chitons). Some collections count valves rather than individuals, even for live-collected specimens. When soft bodies are removed and preserved in ethanol separate from the dry shell, the individuals may be counted twice in some collections.
- A specimen. A specimen is not equivalent to an individual because derived objects (*e.g.* microscope slides of a radula, frozen tissue samples) may also be counted as specimens. The individual that originally contained the derived object may be preserved separately and may even reside at a different institution. The number of specimens therefore is an estimate of the objects managed in a collection, not the number of individual organisms preserved.
- A lot. The commonly accepted definition of a lot is a group of individuals (n=1 to many) of the same species that were collected during a single collecting event (same locality, same date), but it is context dependent. If a lot is split and a part is sent to another institution, it then becomes two lots. Individuals of one species from a particular collecting event are generally counted as different lots if they have different preservation (dry versus alcohol), but those lots often have the same catalogue number and are tracked as a single database record. Specimens might also be cataloged as individuals if it is necessary to track information at the individual level, for example, a holotype split from paratypes, or an individual from which a DNA sequence is available. The number of lots therefore is an estimate of the number of samples (containing one of more specimens) managed by a collection.
- A record. A record in a database generally corresponds to a line in a hand-written ledger of a collection, and some collection databases originated in part from digitizing such ledgers. Like a lot, a record refers to a group of individuals of the same species (or taxon if sorting is incomplete) that came from a single collecting event. Samples with different preservation (alcohol or dry) or derived objects (slides or SEM stubs) may or may not be managed through a single database record, depending on institutional convention. The record is the usual level from which point specimen occurrences are mapped.

Specimen counting and collections size estimates

Institutional estimates of the number of individuals from the number of counted (or estimated) lots vary widely due to the nature of the specimens in a given collection (*e.g.* large-bodied marine species usually contain fewer specimens per lot than microscopic land-snails), but also due to individual collection conventions for using multipliers to estimate individuals. For example, to streamline cataloguing, large lots with many specimens may be recorded in ledgers or on labels as "many", "> 100" or even "∞". Counts of individuals are

frequently estimated assuming that each lot contains an average of 4–20 specimens depending on the collection.

Finally, collection-specific workflows may affect how easy it is to count specimens.

Accessioning is the formal process of transferring ownership of an object(s) to a museum for inclusion in a permanent, managed collection, with associated legal and ethical obligations to care for those objects (Simmons 2006). Traditionally in natural history collections, a single accession (or acquisition) number is given to an entire incoming collection (*e.g.* the Smith Collection of land snails) regardless of whether it contains one object or many. In some institutions or collections, an object is not considered accessioned until it is cataloged.

Furthermore, if a collection uses consecutive catalogue numbers, the size of the collection (in lots) is roughly the same as the latest number assigned. However, this number is often impacted by historic breaks or duplications in the catalogue numbering sequence, deaccessioning of material, the inclusion of non-molluscan taxa (*e.g.* brachiopods) or fossil taxa in the same numbering sequence, or by different practices of assigning single or multiple numbers to sublots (*e.g.* those stored in different media or sorted into age classes).

RESULTS

Size of collections

The mollusk collections of the United States and Canada are diverse in size and specialization (Tables 1, 4; Appendices 3, 5, and 6). They range from the Smithsonian's National Museum of Natural History, which is the largest general molluscan collection in the world with more than one million cataloged lots, to small collections with regional or topical holdings. It should be noted that molluscan collection size is often, but not necessarily, a reflection of overall institutional size, and holdings included here are, for instance, smaller extant molluscan holdings in a much larger predominantly fossil-oriented collection (*e.g.* PRI).

The current listing of these collections encompasses 86 institutions (Table 1). Of these, 30 were included in Solem's (1975) survey. We identified five size-categories of collections, four of which are directly comparable to Solem (1975) with a new category of large collections that occupies a space between Solem's large and medium groups (Table 3, Appendix 3; Figure 3). Solem categorized institutional collections in the following size classes: large (>160,000 lots), middle-sized (30,000 – 75,000 lots), and small (9,000 – 29,000) lots).

Cataloged lots totaled 6.2 million across the collections, with 2.3 million backlog lots for a total of 8.5 million lots in U.S. and Canadian mollusk collections (Appendix 3). Some institutions also provided estimates of the number of specimens, allowing calculation of the number of specimens per lot, which ranged from 1 to 35. The weighted average is 10.5 and the straight average is 8.3 specimens/lot (excluding three institutions with only one per lot because of their specialty). Using 10 specimens/lot as a reasonable average, we calculated a number of specimens per lot when institutions did not provide this number (marked with a double asterisk in Appendix 3; this was also used in a few cases to indicate the reverse calculation where an institution provided specimen but not lot numbers). For USNM, we used 16.2 specimens/lot based on Solem (1975), since using 10/lot would have resulted in estimating fewer specimens than reported more than 40 years ago. The total estimate of the number of mollusk specimens in the collections surveyed here is 70,500,000, of which 4,590,000 resulted from estimation (10/lot). If the straight average rather than the weighted average were used, the total would be reduced by 770,000, which still yields 70 million as an estimate for the number of cataloged specimens.

A backlog of 2.3 million lots implies 23 million backlog specimens, or a total of 93 million specimens, but that estimate is surely low. Some institutions did not provide estimates of the size of their backlog, and we asked for estimates of quality backlog, *i.e.* material ready to catalogue, not deep backlog, which would include unsorted and unidentified material. Also, the estimate of the number of cataloged specimens is probably low. Specimen counts for large lots in mollusk collections are often estimates, *e.g.* ">50" or "100+". Only the numeric portion can be summed, resulting in an underestimate of number of specimens. We therefore regard 100 million specimens in mollusk collections in the U. S. and Canada as a minimum estimate.

Type specimens

Our survey showed about 35,000 primary type lots (holotypes, lectotypes, syntypes, neotypes) among our surveyed collections, and 66,000 secondary type lots (paratypes, paralectotypes) (Appendix 3). The number of primary type lots is likely to decline – there can be multiple lots of syntypes for a given name – but a lectotype designation renders all but one specimen paralectotypes. The ten largest collections in terms of cataloged lots hold 84% of the Recent type material (USNM, ANSP, LACM, UF, FMNH, MCZ, AMNH, BPBM, UMMZ, DMNH).

Preservation

Across institutions, there were 4,677,000 cataloged dry lots and about 742,000 wet lots (Appendix 4), the total of which (5,420,000) is 771,000 lots less than the total cataloged lots reported. This difference is partly because some institutions, generally those with fewer than 40,000 lots, did not report dry versus wet lots, and partly because some institutions reported numbers only for digitized material, rather than from the whole collection. Assuming that institutions that did not report wet holding have essentially entirely dry

Table 3. Comparison of mollusk collection sizes – Categories are a combination of the Large, Medium, Small, and Very Small sizes presented in Solem (1975) plus a new, Medium-Large category based on 2017 data. [1]In Solem (1975), the Small category ranged from 9,000 (EKY) - 27,700 (ChM), but it was expanded to 9,000 –29,000 to exclude gaps between collection categories. [2]Only 8 collections are listed here because the SU collection was transferred to CAS.

Category	Size range (cataloged lots)	Solem (1975)	This paper
Large	≥160,000	8 collections: AMNH, ANSP, BPBM, FMNH, LACM, MCZ, UMMZ, USNM	Expanded to 10 collections: AMNH, ANSP, BPBM, DMNH, FMNH, LACM, MCZ, UF, UMMZ, USNM
Medium –Large	76,000 - 159,999		New Category, 7 collections: BMSM, CASIZ, CM, CMNML, INHS (includes UIMNH), OSUM, SBMNH
Medium	30,000 – 75,999	9 collections[2]: CASIZ, DMNH, INHS, CMNML, OSUM, SBMNH, SDNH, UINH	Expanded and changed to 10 collections: ChM, FWRI, Hefner, HMNS, SDNH, NCSM, SIO-BIC, UCM, UWBM, YPM
Small[1]	9,000 - 29,999	9 collections: ChM, EKY, MPM, UAT, UBC, UCM, UF, UNSM, WSM	Expanded and changed to 20 collections: AUMNH, Brazosport, CASPNNM, DMNS, GTMC-GMNH, INSM, ISM, JFBM, MMNHC, MMNS, MPM, PRI, RBCM-INVZ, ROM, RSMAS, UAM, UCMP, USDA, UWZM, UTEM, WMI, curated by INSM
Very Small	< 9,000	19 unidentified collections with < 5,000 cataloged lots were not analyzed	New category, 32 collections: ARC, ARK, Berkshire, BMS, CLEV, CMC, DMF, DMNH-P, EKY, FWM, HBOM, LSMUG-I, MCPR, NFM, OGL, NHSM, NYSM, OMNH, RMUW, SCSM, SIO-PIC, SMNC, SUI, UAZ, UMNH, UMAMZ, UNM(MSB), UNM(MSB) [P], UNSM, UPRM (INVCOL), VIMS, VMNH

collections, 194,000 should be added to the total for dry lots, which means that about 13% of lots are fluid preserved. Backlog was more than 90% dry preserved, 1,130,000 versus 106,000 lots (Appendix 4), but only 53% of the total backlog of 2.3 million is accounted for in this figure.

Only 14 institutions reported frozen holdings (tissues and whole animals), and only one (UCMP) included frozen lots in its count of cataloged lots. Frozen lots (cataloged and backlog, not counting DNA extracts) total about 15,000 (Appendix 4) and so less than 0.3% of holdings. Other holdings included slides (radular and histology), SEM stubs, egg masses, hosts and parasites, DNA extracts, and images (Appendix 4).

Taxon coverage

About 71% of lots are gastropods, 26% bivalves, 1.2% cephalopods, 1% chitons, 0.4% scaphopods, and 0.1% aplacophorans (Appendix 5). Some institutions reported on only the digitized parts of their collection, whereas others included backlog material. Despite inclusion of backlog, total lots reported across classes was only 5.2 million, 1 million less than total cataloged lots. Although some institutions might have prioritized cataloguing or digitizing certain groups of mollusks, there is no particular reason to expect percentage by class to differ substantially on average between cataloged and backlog material across institutions, so we accept these percentages as representative.

Based on these percentages, we calculated number of lots per species by class (Table 2), based on data from MolluscaBase and Rosenberg (2014) on currently accepted species of mollusks. Average across the Mollusca was 79 cataloged lots per species and 108 total lots per species. Highest coverage is for bivalves, with 230 lots per species (based on total lots), next cephalopods at 128 lots per-species, gastropods and chitons roughly even, with 91 and 82 lots per species, scaphopods at 65 lots per species, and aplacophorans substantially lower at 19 lots per species.

Habitat coverage

Across the surveyed collections, 54% of lots were marine, 19% freshwater, 26% terrestrial, and 1% brackish (Appendices 6–7). In Solem's (1975) survey, 50% were marine, 21% freshwater, and 29% terrestrial. These percentages are probably not significantly different. In Solem's survey seven out of 21 institutions that provided a habitat break down of their collection holdings assumed an even split between two habitat types, for example a predominantly marine collection, like

USNM reported 60% marine and 20% for both freshwater and terrestrial. In the 2017 survey, while the percentage may still be estimates, only three of 58 institutions made such an assumption (Appendix 7) (two others were our own assumption, PRI 50:50 freshwater and terrestrial and FWM 50:50 freshwater and marine).

Despite inclusion of backlog by some institutions (* in Appendix 6), total lots reported across habitats was only 5.6 million, 600,000 less than total cataloged lots. Some institutions did not have a mechanism for determining habitat in their database, such as a taxonomic dictionary with habitats coded for family, genus, or species. Even with this capability, if a specimen lot is not identified in the database, the habitat cannot be determined. Determining brackish status in particular was problematic as some institutions do not distinguish such specimen lots from those from marine environments, and the survey did not give a definition of the difference between these habitats. Some institutional respondents reported percentages as high as 10% brackish, which may require substantiation.

Geographic coverage

Not surprisingly, North American material was the largest component of institutional molluscan holdings at 43%. Caribbean was 6% and South American was 4%. These figures combine marine and non-marine taxa (Appendices 9–10). For marine material from the Americas, about 64% was from the western Atlantic and 36% was from the eastern Pacific. As with taxon and habitat coverage, some institutions included backlog in their figures for geographic coverage, but total lots reported across regions was 1 million less that total cataloged lots. Unlike taxonomic and habitat categories, the holdings in geographic categories do not sum to 100%, since our survey did not ask about areas outside the Americas, so the percentages for geographic coverage were calculated without data that included backlog (* in Appendices 9-10).

Digitization

Of the 6.2 million cataloged lots, 4.5 million (73%) have undergone some form of data digitization (which includes all forms of digitization, *e.g.* ledger records entered, transcribed, or imported into word processor, spread sheet, or relational database formats). About 1.1 million (25%) of digitized records have been georeferenced, which represents 18% of all cataloged lots. Only 20 collections (<25%) claim to be fully Darwin Core compliant, however, 34 of the 66 collections with some form of digitization are searchable online (Appendix 8) through iDigBio, Arctos, or other portals, or directly through institutional websites.

Mollusk Collections in the U.S. and Canada in 2017

In the United States, collections included herein were from 37 states, the District of Columbia, and the Territory of Puerto Rico; Canada was represented by five collections in four provinces (Figure 1 – Map). Collections are concentrated along both coasts of North America, with five of the ten largest in the Boston-Washington corridor of the east coast, two in the Midwest, and one each in Florida, California, and Hawaii. States with the largest populations have multiple collections (*e.g.* California, Florida, Illinois, Massachusetts, New York, Ohio, and Texas); whereas large parts of the central and northwestern United States do not have any identified mollusk collections.

Mollusk collections are maintained by a wide variety of institutions and were developed for a similarly wide variety of reasons. Collections are developed, acquired, and grown due to research, educational, avocational, or monitoring activities that are part of the overall mission of the institution. The largest mollusk collections in the U.S. and in Canada are both federally funded. Their collections are geographically global in scope, although North American materials dominate (Table 3 – Collection Size Categories, Appendices 9, 10). Many of the largest mollusk collections (>160,000 lots) are "private" museums, which, in the U.S., are tax-exempt, not-for-profit, 501c3 institutions (*e.g.* AMNH, DMNH, FMNH). The formal reports on 990 tax forms (https://irs.gov/forms-pubs/about-form-990) show a wide variety in what it means to be "private," including public-private partnerships. Large collections are also found at university-affiliated and supported museums (*e.g.* ANSP at Drexel University, MCZ at Harvard University, OSUM at Ohio State University, UF at the University of Florida), many of which have a long history of mollusk-centered research. In contrast, smaller collections are often found at institutions with a regional audience or research focus. For example, the Natural History Society of Maryland in Baltimore has a land snail collection, and state wildlife agencies such as the Fish and Wildlife Research Institute in St. Petersburg, Tampa or the Florida Fish and Wildlife Conservation Commission maintain specimen collections as a result of their mission to understand and manage local wildlife.

In addition to full-time, permanent collection staff, survey respondents included emeritus faculty, graduate students, and volunteers. The wide range of individuals who answered the survey is both a testament to the depths of dedication to the care of these important collections, and simultaneously a sign of an ongoing staffing crisis.

The number and kind of collection-related staff positions reported were highly variable (Table 4). In larger collections, there was generally one or more senior research biologists with a mollusk-specific research interest plus administrative responsibilities for the collection (*e.g.* strategic oversight, growth, funding) plus one or more collection professionals with day-to-day operational responsibility for managing and caring for the collection. Job titles for the research/administrator

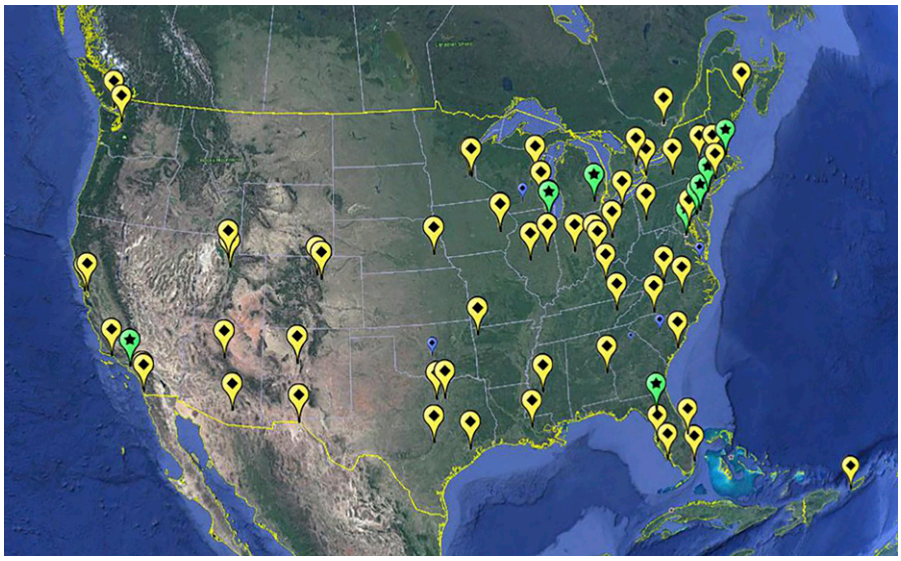

Figure 1. Distribution of surveyed molluscan collections in North America (note that two collections, in Alaska and Hawaii, are outside the map area). Green/asterisk flags mark collections that rank among the 10 largest by number of cataloged lots; yellow/diamond flags mark collections with at least 1,000 verified research-quality specimen lots of extant mollusks; smaller blue flags mark collections with smaller such holdings or for which only limited data could be obtained during the survey. Map data ©2018 Google, Image Landsat/Copernicus, Data SIO, NOAA, US Navy, NGA, NSF, GEBCO, US Dept of State Geographer.

position vary from Curator to Professor to Research Biologist, and often this person has significant additional responsibilities elsewhere such as teaching and mentoring undergraduate and graduate students. In smaller collections, this senior position may be held by a volunteer. In larger collections, operations staff may be a full or part-time Collection Manager, Collection Assistant, data entry specialist, archivist, or someone completely outside the collections' community (*e.g.* an aquarist). Especially at some of the larger university-supported institutions, Collection Managers often have advanced degrees, some with PhDs, and are expected to participate in grant writing and collection-based research (*e.g.* UMMZ).

Collection operations staff may be generalized and pooled across disciplines so that a few people are responsible for multiple extant and paleontological invertebrate collections. In these scenarios, it is not uncommon for collections to be without relevant molluscan curatorial expertise, sometimes for many years (*e.g.* AMNH, DMNH-P, SDNH, UNSM). Our data demonstrate that collections exposed to prolonged periods without dedicated staff with expertise in the field become increasing inaccessible and lag in technological and digital data management advances (see Table 1, collections indicated by LD provided only limited data in our survey and also Appendix 3; collections such as CMC and UAZ reporting little or no change of their collection data between 1975 and 2017).

Digital data in mollusk collections of the U.S. and Canada

Most North American collections nowadays have reached some level of digitization, with about 73% of total cataloged lots digitized (Appendix 8). Many collections (ANSP, ARC, CASPNNM, DMNH-P, FWM, FWRI, INHS, JFBM, LSUMG-I, MCZ, MMNHC, MMNS, OSUM, RBCM-INVZ, SMNC, SUI, UCM, UCMP, UF, VMNH) have reached complete or near-complete data entry of their cataloged collections.

Digital collections

Software platforms for the digitization efforts of these molluscan collections are remarkably diverse and a reflection of both the different directions and levels of institutional information technology development and of the individual initiatives and preferences of collections staff (Figures 2, 4). Eight collections reported no past or present digitization efforts. Of the collections with digitized data, many described using the dedicated commercial or open-source collections software solutions *Specify* (11 collections), *EMu* (8), *Arctos* (5), *PastPerfect* (3), *Mimsy XG* (2), and *Proficio* (1); others are based on generic current or legacy database (*Access* [14]; *FileMaker Pro* [11], *Paradox* [1]) and spreadsheet (*Excel* [8]) applications, sometimes combining more than one software solution. Several collections currently on generic database or spreadsheet systems indicated plans of switching to either *Specify* or *EMu* in the near future. Although most of the collections use database systems, it became clear during communications about the survey questions that many collections were not in a position to run basic queries on their database holdings. While some of this reflected a lack of authority files (*e.g.* to link to higher taxa or geographic hierarchy), in other cases it was clearly a lack of institutional staffing support. Particular data types (*e.g.* habitat coverage data) are currently unavailable for eleven collections (see Appendices 5, 6). Extraction of geographic (regional) data is currently problematic for 29 collections (Appendices 9, 10).

DISCUSSION

Collection staffers are routinely asked the deceptively simple question – how big is your collection? Most collections have settled on a standard answer to this question that is

Table 4. Mollusk Collection Metadata – Collection management metadata about mollusk collections with >85K cataloged in 2017. Abbreviations: staffing c = curator, which includes the senior administrative staff responsible for research and oversight of the collection; staffing s = staff, which includes operational staff, volunteers, and students responsible for the day to day activities within the collection; [1]Curators, Collection Managers and Collections Assistants are counted as 1 FTE Full-time) or 0.5 FTE (part-time, or emeritus). Volunteers or casual to part time employees are counted as 0.2 FTE per day of volunteering.

			Data available on					Staffing full-time equivalents (FTE)[1]							
								2017		1975		cataloged lots		Change from 1975 to 2017	
Collection acronym	Institution type	Data management system	Home website	iDigBio	InvertEBase	GBIF	OBIS	c	s	c	s	2017	1975	lots (%)	c
AMNH	private, non-profit	Filemaker Pro	X					0	0.2	1.0	3.5	319,000	175,000	182	-1.0
ANSP	university museum	Filemaker Pro	X	X		X	X	1.0	4.6	2.0	5.5	501,000	336,737	149	-1.0
BMSM	private, non-profit	MS Access	X	X		X	X	1.0	3.4	0.0		119,822			1.0
BPBM	private, non-profit	MySQL/ MS Access		X		X		0	2.3	1.0	0	300,000	160,000	188	-1.0
CASIZ	private, non-profit	Paradox & Specify	X	X	X	X		0	2.4	1.0	3.5	91,858	35,000	262	-1.0
CM	private, non-profit	EMu		X	X			1.0	1.5	1.0		152,521			0.0
CMNML	federal corporation	MIMSY XG moving to EMu	X	X		X		1.0	3.5			100,000	70,632	142	1.0
DMNH	private, non-profit	Excel and Specify		X	X			1.0	3.3	1.0	4.5	220,287	70,650	312	0.0
FMNH	private, non-profit	EMu	X	X	X	X		2.0	3.5	1.0	2.5	382,000	179,000	213	1.0
INHS	university museum	Filemaker Pro	X	X				1.0	0.7	0.5	0	86,790	75,000	116	0.5
LACM	private, non-profit	EMu	not online					1.0	1.5	1.0	1.25	500,000	160,000	313	0.0
MCZ	university museum	Arctos	X	X		X		1.0	2.5	2.0	0.25	372,056	270,000	138	-1.0
OSUM	university museum	Filemaker Pro	X	X				1.0	1.6	2.0	3	120,180	30,000	401	-1.0
SBMNH	private, non-profit	Specify	X					1.5	1.7	1.0	1.5	100,170	52,000	193	0.5
UF	university museum	Specify	X	X	X	X	X	1.0	5	1.0	0.5	497,459	22,174	2,243	0.0
UMMZ	university museum	Specify	X	X	X	X		2.0	1.6	2.5	2.5	251,000	232,373	108	-0.5
USNM	federal government	EMu	X	X		X	X	3.0	pool	4.0	10	1,000,000	740,000	135	-1.0
							Total:	18.5		22.0					

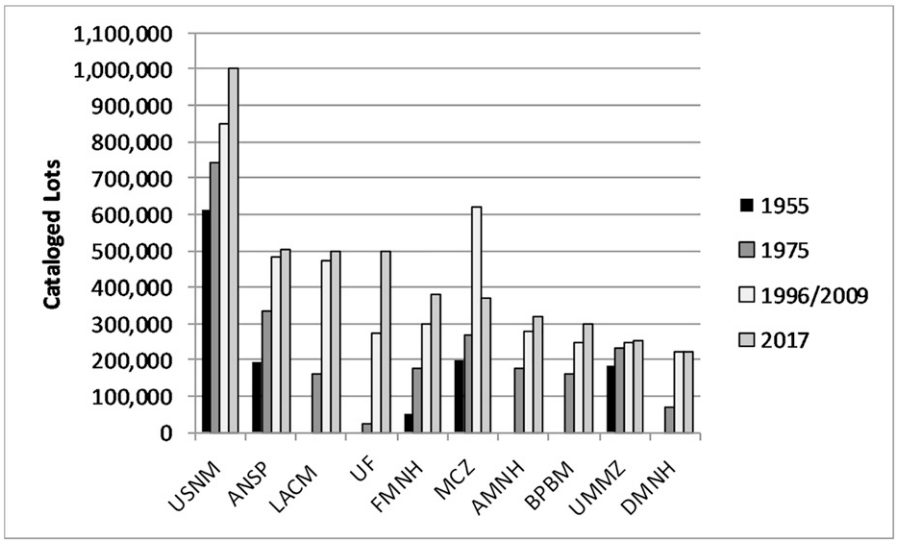

Figure 2. Large mollusk collections. The 1955 data are from Table 4 in Solem (1975).

both defensible and approximately correct using available tools. As digitization continues, collections will be able to report on their size and scope with increasing rigor and nuance. However, nationwide, collections are mid-way through this transition and are grappling individually with challenges of rapid technological changes, enormous workloads and backlogs, all contained within an environment of declining resources and reduction in staffing (Table 4).

Overall, we estimate that in 2017 there are at least 100 million mollusks in 8.5 million lots held by about 90 mollusk collections across the U.S. and Canada. The Smithsonian's Natural History Museum has the single largest cataloged collection in North America with more than 1 million cataloged lots; however, there are also > 20 mollusk collections with between 1,000 – 8,000 cataloged mollusk lots and > 20 mollusk collections with between 9,000 – 30,000 lots (Appendix 3). Some of the smaller collections have important regional and taxonomic holdings. For example, SIO-PIC has the fourth largest cephalopod collection, and RBCM has the sixth largest chiton collection among surveyed institutions (Appendix 5). The number of specimens per lot reported in collections varies widely, from 1 to 35, depending on the focus of an institution. OGL is a collection of DNA extracts from frozen tissues wherein every extract is a specimen. UNM is a parasite collection, with individual mollusks preserved as the source of parasites, so it also has a 1:1 ratio of specimens to lots. Brazosport has a synoptic display collection, usually with only one specimen per lot. On the other end of the scale, institutions that have largely expedition and survey material tend to have larger number of specimens per lot.

U.S. and Canadian collections hold about 101,000 type lots (Appendix 3), but the actual number of taxon names typified by these is unknown (*e.g.* multiple institutions may hold syntype lots for a single nominal species). The number of names typified can eventually be determined once most collections are represented in online portals such as GBIF and iDigBio. Unrecognized type material might also be found through these portals, by searching for specimens collected, donated, or formerly owned by the authors of taxa. Images of primary type specimens collected through these general natural history portals, and taxon-specific ones such as MolluscaBase, will be an enormous resource for the scientific community, as they are the standards for authoritative identification.

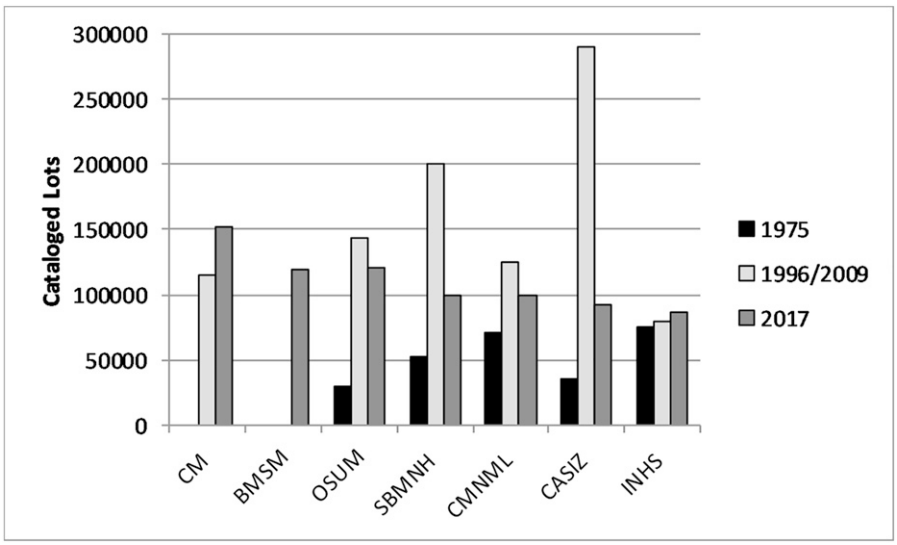

Figure 3. Mid-sized plus mollusk collections. This category is defined as having 85,000 – 120,000 cataloged lots and did not exist in Solem (1975).

Type specimens represent less than 2% of the holdings of the surveyed institutions, and so present an obvious, achievable target for digitization and

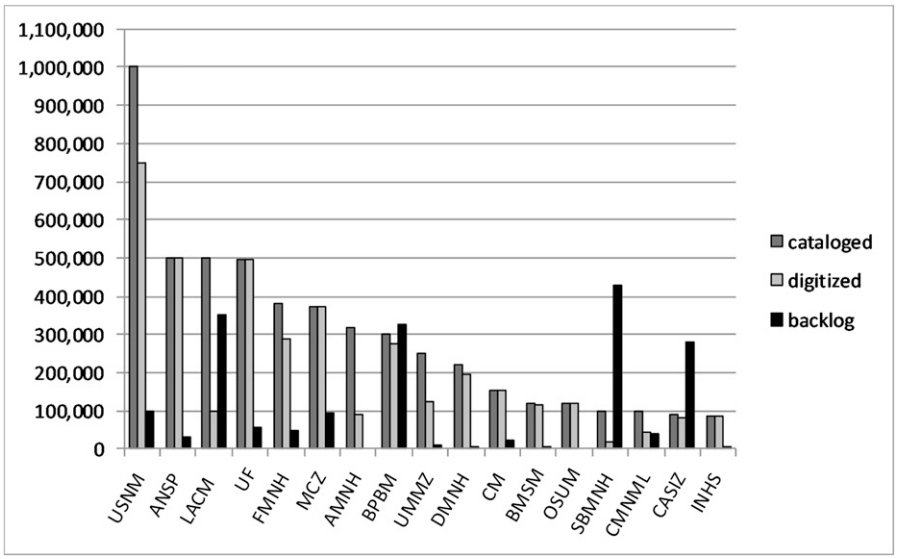

Figure 4. A snapshot of curation status in large and mid-sized plus collections. All data from the 2017 survey. Backlog numbers are estimates of quality backlog.

imaging efforts. However, mobilizing data on a much larger scale is needed to understand changes in distribution patterns of species over time. Across the surveyed collections, 73% of cataloged material is digitized in the broadest sense (Appendix 8), but when quality backlog material is included, this drops below 53%.

Computerization or digitization of collections, then referred to as electronic data processing (EDP), was in its infancy when Solem prepared his 1975 report, which was preparatory to large scale efforts to digitize mollusk collections. Digitization of 4.5 million molluscan lots since then is an impressive achievement, but, including backlog, 4 million more already in collections remain to be digitized, in addition to new acquisitions resulting from ongoing field collecting, donations of collections from private individuals, and acceptance of orphaned collections from other institutions (*e.g.* universities that reduce their organismal research programs).

The greatest need for digitization in mollusks collections is probably georeferencing, which is essential for mapping and visualizing distributions of species. Only 25% of digitized mollusk lots and only 18% of cataloged lots are georeferenced (Appendix 8), which means that most institutions are not yet in position to automate sharing of their spatial data through standard portals. Our survey asked only if coordinate data were available, not whether the georeferencing met modern standards. For example, fields for source of coordinates, and error radius (Shea *et al.* 2018), which are currently part of Darwin Core (see at: http://rs.tdwg.org/dwc/terms/, Biodiversity Information Standards TDWG) were not initially part of the Darwin Core standard (Wieczorek *et al.* 2012). Only 20 collections claimed that all their records were fully Darwin Core compliant, but 66 collections reported having some level of digitization. Of these, 34 have their data online through searchable databases. This suggests that more than 20 collections have databases that are compliant.

Our survey asked for the number of records captured with skeleton data (minimal or incomplete label data, *e.g.* only identification and country) or with fields not standardized, versus records that were Darwin Core compliant (Appendix 1, question 5), but at least one institution replied that all of its records were skeleton data only, but that all were Darwin Core compliant. This is technically correct—the database is compliant, and the data that are present are compliant, but a large amount of data remain to be input or retrofitted. Clearly, the responses to this part of our survey were heterogeneous.

Solem (1975) emphasized that mollusk collections are exceeded in number of specimens only by entomological collections. Several other collection communities produced surveys similar to Solem's in the 1970s, from which he noted 120 million insects, 72 million mollusks, and 46 million plants. The U.S. National Herbarium currently reports that its 5 million specimens represent about 8% of the holdings in the United States, implying that there are about 63 million total specimens (http://botany.si.edu/colls/collections_overview.htm), so the relative rankings of insects, mollusks and plants are unchanged since the 1970s.

Solem also estimated that mollusks, with about 850 specimens per species across the surveyed collections were second only to fish, with about 1,750 collection specimens per species in U.S. and Canadian collections. These high numbers of specimens available per species for study and comparison make these two groups extremely well-suited to environmental monitoring, allowing measurement of attribute changes (*e.g.* body size) through space and time. No other taxon group passed 500 specimens per species in the nation's collections (Solem 1975: 231). Since Solem's time, the estimated number of described mollusk species have declined from 85,000 to about 78,800 (Table 2), whereas the number of fish species has increased from the 20,000 Solem used for calculations to 34,852 (Eschmeyer and Fong 2018). We found an average of 108 lots per species for mollusks (Table 2), or about 1,100 specimens per species, using the

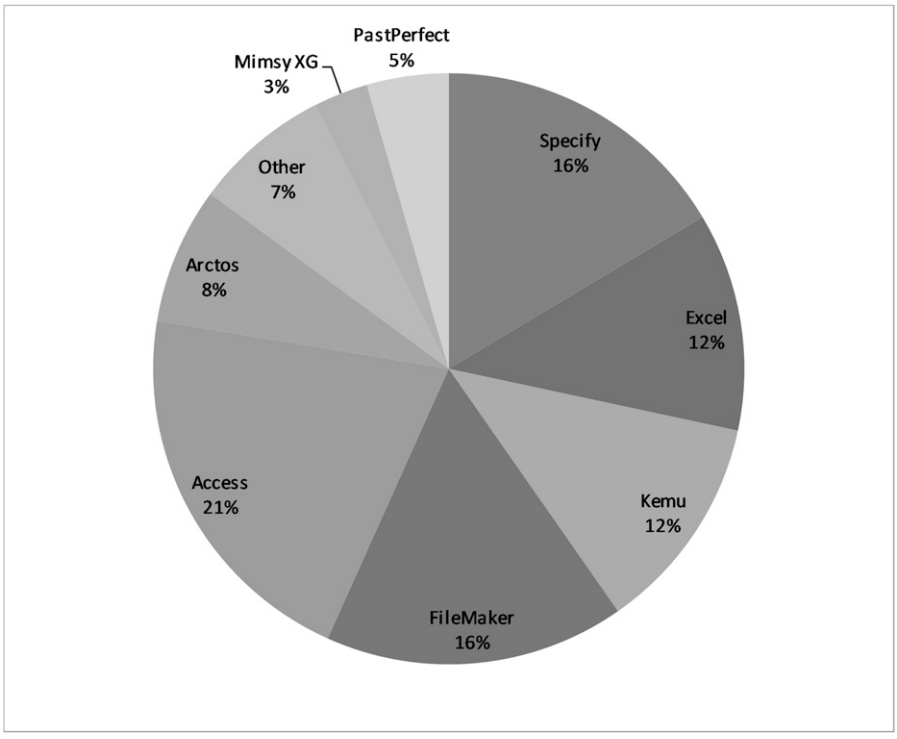

Figure 5. Software applications used to manage mollusk collection data in the U.S. and Canada in 2017.

weighted average of 10.5 specimens per lot (Appendix 3). This is almost 10 times what Solem reported for insects, 120 specimens per species, a figure that is probably high, given the explosive increase in estimates for diversity of insects since that time (Stork *et al.* 2015).

The estimate of 1,100 specimens per species might also be considered high, since mollusk collections contain undescribed species and unidentified specimens. Many mollusks species are rare, however, being known only from the type locality, and so may not even be present in North American collections. We expect that many common species will be represented by hundreds of samples and thousands of specimens that will allow construction of time series for studying changes in distribution patterns.

The pattern of lots per species among the molluscan classes in U.S. and Canadian collections is interesting. (We switch here to using lots rather than specimens per species since it is not clear that the average number of specimens per lot will be the same across classes and our survey data do not allow us to calculate it). Bivalves (230 lots per species) and cephalopods (128 lots per species) have the highest values, which perhaps reflects that some species have high enough abundance that they can sustain fisheries, but they are collected with different methods, since virtually all bivalves are benthic whereas many cephalopods are pelagic. Gastropods and chitons have similar, intermediate values, 91 and 82 lots per species, which is probably coincidental since gastropods occupy terrestrial and freshwater habitats in addition to the marine habitat of chitons. Scaphopods are somewhat lower at 65 lots per species, which might reflect their exclusively infaunal habitat. Aplacophorans are undersampled relative to other mollusks at 19 lots per species, reflecting that they lack shells and therefore are not collected post-mortem like shell-bearing mollusks; also, their small worm-like bodies may not be recognized as molluscan by the non-specialist. If the average value of specimens per lot (10.5, Table 3) is applied across the molluscan classes, the aplacophorans end up in Solem's "low" category for specimens per species, whereas the bivalves have the highest value, at 2,400 specimens per species, which suggests that bivalves are one of the best sampled classes of metazoans. Only among microorganisms such as diatoms might higher number of specimens per species be expected in natural history collections.

Fish collections have grown more rapidly than mollusk collections, increasing from 35 million to 64 million specimens by the early 1990s (Poss and Collette 1995). Much of this growth however has been in larval fish, which are difficult to identify by morphological means (Ko *et al.* 2013). Also, unlike mollusk collections, where all the large collections reported an increased number of lots between Solem's 1975 survey and our 2017 survey, some fish collections show a decrease in number of lots from 1995 to today, judging from collection websites (LACM, 7 million to 4 million; CAS: 2.16 million to 1.2 million; USNM, 5 million to 4 million, MPM, 1.5 million to 685,000). It is therefore difficult to judge whether mollusk or fish collections have more material per species that is relevant for assessing environmental change. What is more important is the similarity of fish and mollusk collections in being lot based, which means they can be more rapidly and effectively digitized than taxa in which cataloguing and labelling is individual based. A single lot in a mollusk or fish collection can contain hundreds of specimens, which allows studies of environmentally mediated change in morphological and genetic variation over time, in addition to changes in distribution patterns.

Over the past 15 years, the value of natural history collections (National Science and Technology Council 2009), their use in formal and informal education (Cook *et al.* 2014, Ellwood *et al.* 2015, Hiller *et al.* 2017), and their increased use

in research due to digitization efforts has been the subject of much discussion. There are many challenges to digitization that still must be addressed, including a need for coordination of activities across all natural history collections, and means of finding efficiencies in similar tasks (Vollmar *et al.* 2010).

The scope of the challenge is staggering, and global efforts to consider the best approach have focused on digitizing metadata from all collections across the world as a starting point (Berendsohn and Seltmann 2010, Scoble 2010, Page *et al.* 2015). As part of the U.S. response to this enormous task, the National Science Foundation Program "Advancing Digitization in Biological Collections" has been funding digitization and web-publishing of data in non-federal collections since 2011. The iDigBio program (Integrated Digitized Biocollections), headquartered at the Florida Museum of Natural History and currently supported by NSF, has become an important collaborator and coordinator of collections information and resources for collections trying to improve their digitization efforts.

Sustained funding and coordinated collaboration have brought millions of new specimen records online, and natural history collections are slowly being recognized as the important source of biodiversity data that they are; however, the usefulness of natural history collections in the 21st century is directly tied to their availability online and the ability of researchers to use those data quickly and with confidence. The rate at which mollusk collection data are annotated, contextualized with provenance data, and published online needs to be accelerated. Planning for this challenge requires an understanding of the amount of data available and its state of preparedness for digitizing.

ACKNOWLEDGEMENTS

A big thank you goes, posthumously, to Alan Solem (FMNH) who had the foresight and stamina to request and compile the original collection survey results in the early 1970s. He likely suspected how valuable the baseline that he generated would become. For this new round of questionnaires, we had the support of a very large number of colleagues at numerous collections across North America who responded to our (RB and PS) requests for survey data and/or helped answering various associated questions (here arranged by U.S. states and Canadian provinces):

J. Andrés Lopez, UAM (AK); Jason Bond, Melissa Callahan, AUMNH (AL); Nancy G. McCartney, ARK (AR); Peter N. Reinthal, UAZ (AZ); Margaret Dykens, SDNH (CA); Christina Piotrowski, Elizabeth Kools, CASIZ (CA); Erica Clites, UCMP (CA); Lindsey T. Groves, Jann Vendetti, LACM (CA); Charlotte Seid, SIO-BIC (CA); Linsey Sala, SIO-PIC (CA); Paul Valentich-Scott, SBMNH (CA); CalCOFI (CA);Jingchun Li, Kelly Martin, UCM (CO); Paula Cushing, Phyllis Sharp, DMNS (CO); Eric A. Lazo-Wasem, YPM (CT); Ellen Strong, USNM (DC); Alex Kittle, DMNH (DE); Dennis Hanisak, HBOM (FL); Paul Larson, Laura Wiggins, FWRI (FL); José Leal, BMSM (FL); Nancy A. Voss, RSMAS (FL); John Slapcinsky, Gustav Paulay, UF (FL); Byron J. Freeman GTMC-GMNH (GA); Norine W. Yeung, Richard Pyle, Jaynee R. Kim, BPBM (HI); Cindy Opitz, SUI (IA); Dawn Roberts, CASPNNM (IL); Kevin S. Cummings, INHS (IL); Meredith Mahoney, ISM (IL); Jochen Gerber, Janeen Jones, FMNH (IL); Ronald L. Richards, Ryan Rokicki, Randy and Deborah Patrick, INSM and WMI (IN); David M. Hayes, EKY (KY); Lorene E. Smith, LSUMG-I (LA); Adam Baldinger, MCZ (MA); Scott Jervas, Berkshire Museum (MA); Akiko Okusu, UMass Amherst (MA); Hayley Singleton, Beneski Museum, Amherst College (MA); Hannah Appiah-Madson, OGL (MA); James Young, NHSM (MD); Anthony L. Swinehart, DMF (MI); Taehwan Lee, UMMZ (MI); Sean Keogh, Andrew M. Simons, JFBM (MN); Richard J. Oehlenschlager, SMM (MN); Robert L. Jones, MMNS (MS); Lenny Lampel, MCPR (NC); Arthur E. Bogan, Jamie Smith, NCMNS (NC); Denise Furr, SMNC (NC); Patricia W. Freeman, Thomas E. Labedz UNSM (NE); David Parris, NJSM (NJ); Sara V. Brant, Sandra L. Brantley, MSB (NM); Christine Johnson, AMNH (NY); Isabel P. Hannes, Kathryn Leacock, BMS (NY); Denise A. Mayer, NYSM (NY); Greg Dietl, Leslie L. Skibinski, PRI (NY); Nicole Gunter, Gavin Svenson, CLEV (OH); Francisco Borrero, Emily Imhoff, CMC (OH); Steven Sullivan, Hefner (OH); G. Thomas Watters, OSUM (OH); Katrina Menard, OMNH (OK); Christopher Marshall, OSAC (OR); Nezka Pfeifer, Everhart (PA); Timothy Pearce, CM (PA); Paul Callomon, ANSP (PA); David Robinson, USDA (PA); Alex Van Dam, UPRM (PR); Dave Cicimurri, SCSM (SC); Matthew Gibson, ChM (SC); Gerald R. Dinkins, MMNHC (TN); Teresa Mayfield, UTEP (TX); Melissa Casarez, TNHC (Austin, TX); Karen Morton, DMNH-P (TX); Lacie Ballinger, FWM (TX); Tina Petway, HMNS (TX); Christy Bills, UMNH (UT); Wesley Skidmore, MLBeanLSM (UT); Jennifer C. Dreyer, VIMS (VA); Haley Cartmell, VMNH (VA); Melissa Frey, UWBM (WA); Emily Halverson, Laura A. Monahan, UWZM (WI); Daniel J. Meinhardt, Richter (WI); Julia Colby, MPM (WI); Claire Goodwin, Rebecca Milne, ARC (New Brunswick, Canada); Henry Choong, RBCM-INVZ (British Columbia, Canada); Jean-Marc Gagnon, CMNML (Ontario, Canada); Sebastian Kvist, Maureen Zubowski; ROM (Ontario, Canada).

Many other colleagues helped with additional data, among them Jay Codeiro, Victor Fet, Daniel L. Graf, and Timothy Rawlings, and we apologize for any inadvertent omissions from this list. The work on this project was partly supported by NSF award EF-Digitization TCN 14-02667 to P. Sierwald and R. Bieler, and by a mollusk collection digitization workshop grant through iDigBio to P. Sierwald and E. Shea (https://idigbio.org/wiki/index.php/Digitizing_the_2nd_largest_Invertebrate_Phylum:_Mollusks).

LITERATURE CITED

Arctos. 2018. Collection Management Solution. Available at: https://arctosdb.org/about/

Ariño, A. H. 2010. Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics* **7:** 81–92.

Ball-Damerow, J. E., L. K. M'Gonigle, and V. H. Resh. 2014. Changes in occurrence, richness, and biological traits of dragonflies and damselflies (Odonata) in California and Nevada over the past century. *Biodiversity and Conservation* **23:** 2107–2126.

Beaman, R. and N. Cellinese. 2012. Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science. *ZooKeys* **209:** 7–17.

Berendsohn, W. G. and P. Seltmann. 2010. Using geographical and taxonomic metadata to set priorities in specimen digitization. *Biodiversity Informatics* **7:** 120–129.

Bieler, R. and A. Bradford. 1991. Annotated catalogue of type specimens in the malacological collection of the Delaware Museum of Natural History. Gastropoda (Prosobranchia and Opisthobranchia). *Nemouria*, *Occasional Papers of the Delaware Museum of Natural History* **36:** 1–48.

Ciubuc, C. 2017. The Catalogue of "Constantin Ciubuc" Trichoptera (Insecta) collection of "Grigore Antipa" National museum of natural history. *Travaux du Muséum National d'Histoire Naturelle,* **60**(1)**:** 7–371.

Coan, E. V. and A. R. Kabat. 2018. *2,400 years of Malacology* [15th Edition – February 12, 2018]. The American Malacological Society, on-line publication available at: http://www.malacological.org/2004_malacology.html.

Cook, J. A., S. V. Edwards, E. A. Lacey, R. P. Guralnick, P. S. Soltis, D. E. Soltis, C. K. Welch, K. C. Bell, K. E. Galbreath, C. Himes, J. M. Allen, T. A. Heath, A. C. Carnaval, K. L. Cooper, M. Liu, J. Hanken and S. Ickert-Bond. 2014. Natural history collections as emerging resources for innovative education. *BioScience* **64:** 725–734.

Cummings, K. S., A. Oleinik, and J. H. Slapcinsky. 2009. Systematic Research Collections (Recent and Fossil Mollusca). Online publication available at: http://wwx.inhs.illinois.edu/collections/mollusk/links/list [file update date given as 1 October 2009; collection descriptions often individually dated]

Darwin Core. 2018. Darwin Core quick reference guide. Biodiversity Information Standards TDWG; Available at: http://rs.tdwg.org/dwc/terms/. 8 November 18

Digital Data in Biodiversity Research Conference. 2018. Digital Data in Biodiversity Research Conference. Available at: https://www.idigbio.org/wiki/index.php/Digital_Data_in_Biodiversity_Research_Conference,_Berkeley. 8 November 18

Ellwood, E. R., B. A. Dunckel, P. Flemons, R. Guralnick, G. Nelson, G. Newman, S. Newman, D. Paul, G. Riccardi, N. Rios, K.C. Seltmann, and A. R. Mast. 2015. Accelerating the digitization of biodiversity research specimens through on-line public participation. *BioScience* **65:** 383–396.

Eschmeyer, W. N. and J. D. Fong. 2018. Catalog of Fishes, version of 2. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp 2 August 2018.

GBIF: The Global Biodiversity Information Facility. 2018. What is GBIF? Available at: https://www.gbif.org/what-is-gbif 13 August 2018.

GEOLocate. 2018. *A Platform for Georeferencing Natural History Collection Data*. Available at http://www.geo-locate.org/

Gerber, J. 2010. Leslie Hubricht (1908-2005), his publications and new taxa. *American Malacological Bulletin* **28:** 15–27. doi: 10.4003/006.028.0208

Hiller A. E., C. Cicero, M. J. Albe, T. L. W. Barclay, C. L. Spencer, M. S. Koo, R. C. K. Bowie, and E. A. Lacey. 2017. Mutualism in museums: A model for engaging undergraduates in biodiversity science. *PLoS Biology* **15:** 1-11.e2003318.

Howard, D. W., E. J. Lewis, B. J. Keller, and C. S. Smith. 2004. Histological techniques for marine bivalve mollusks and crustaceans. NOAA Technical Memorandum NOS NCCOS **5**, 218 pp.

iDigBio. 2017. iDigBio Portal. Available at: https://www.idigbio.org/portal 21 August 2018.

Irwin, H. S., W. W. Payen, D. M. Bates and P. S. Humphrey, eds. 1973. *America's Systematics Collections: a National Plan.* [Washington, D. C.], Association of Systematics Collections. xiii, 63 pp.

InvertEBase. 2017. InvertEBase Data Portal Home. Available at: http://www.invertebase.org/portal/index.php 21 August 2018.

Kabat, A. R. and K. J. Boss. 1992. An indexed catalogue of publications on molluscan type specimens. *Occasional Papers on Mollusks* **5:** 157–336.

Kabat, A. R. and K. J. Boss. 1997. Addendum to "An indexed catalogue of publications on molluscan type specimens" (1992). *Occasional Papers on Mollusks* **5:** 337–370.

Ko, H.-L., Y.T. Wang, T.-S. Chiu, M.-A. Lee, M.y.-Leu, K.-Z. Chen, W.-Y. Chen and K.-T. Shao. 2013. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One* **8:** e53451.

MolluscaBase. 2017. Mollusc Base. Available at: http://www.molluscabase.org. 21 August 2018.

National Science and Technology Council, Committee on Science, Interagency Working Group on Scientific Collection. 2009. *Scientific Collection: Mission-Critical Infrastructure of Federal Science Agencies.* Office of Science and Technology Policy, Washington D.C. 47 pp.

OBIS. 2018. Ocean Biogeographic Information System. Intergovernmental Oceanographic Commission of UNESCO. Available at: www.iobis.org. 8 November 18.

Page, L. M., B. J. Macfadden, J. A. Fortes, P. S. Soltis and G. Riccardi. 2015. Digitization of biodiversity collections reveals biggest data on biodiversity. *Bioscience* **65:** 841–842.

Poss, S. G. and B. B. Collette. 1995. Second survey of fish collections in the United States and Canada. *Copeia* **1995(1):** 48–70.

Roper, C. F. E. R. and M. J. Sweeney. 1983. Techniques for fixation, preservation, and curation of cephalopods. *Memoirs of the National Museum of Victoria* **44:** 29–47.

Rosenberg, G. 2014. A new critical estimate of named species-level diversity of the Recent Mollusca. *American Malacological Bulletin* **32:** 308-322.

Rosenberg, G. and M. Khoo. 2018. Improving provenance data in natural history collection databases. *American Malacological Bulletin*, **36(2):** 215–231.

Schindel D, S. Miller, M. Trizna, E. Graham, A. Crane. 2016. The global registry of biodiversity repositories: a call for community curation. *Biodiversity Data Journal* **4:** e10293.

Scoble, M. J. 2010. Natural history collections digitization: Rationale and value. *Biodiversity Informatics* **7:** 77–80.

Shea, E. K., P. Sierwald, R. Bieler, and G. Rosenberg. 2018. Priorities and opportunities for digitizing mollusk collections**.** *American Malacological Bulletin.* **36(2):** 171–176.

Simmons, J. E. 2006. *Things Great and Small: Collections Management Policies*. American Association of Museums, Washington, D.C. Pp. 208.

Solem, A. 1975. The Recent mollusk collections resources of North America. *The Veliger* **18:** 222–236.

Solem, A. 1986. A collector's tale. *Field Museum of Natural History Bulletin* **57:** 22–25.

Solem, A., W. K. Emerson, B. Roth, and F. G. Thompson. 1981. Standards for malacological collections. *Curator* **24:** 19–28.

Stork, N. E., J. McBrooma, C. Gelyb, and A. J. Hamilton. 2015. New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. *Proceedings of the National Academy of Sciences* **112:** 7519–7523.

Sturm C. F. 2006. Chapter 13: Museums and Malacology. In: C. F. Sturm, T. A. Pearce and A. Valdés (eds). *The Mollusks: A Guide to their Study, Collection, and Preservation*. American Malacological Society, Pittsburgh, PA, USA. Pp. 181–187.

Sturm, C.F., R. Mayhew, and B. R. Bales. 2006. Chapter 2: Field and laboratory methods in malacology. In: C. F. Sturm, T. A. Pearce and A. Valdés (eds). *The Mollusks: A Guide to their Study, Collection, and Preservation*. American Malacological Society, Pittsburgh, PA, USA. Pp. 9–31.

Tennent, N. H. and T. Baird. 1985. The deterioration of Mollusca collections: identification of shell efflorescence. *Studies in Conservation* **30:** 73–85.

Vollmar A., J. A. Macklin and L. S. Ford. 2010. Natural history specimen digitization: challenges and concerns. *Biodiversity Informatics* **7:** 93–112.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson and D. Vieglais. 2012. Darwin Core: An evolving community-developed biodiversity data standard. *PloS One* **7:** e29715.

WoRMS. 2004. World Register of Marine Species: Mollusca. Available at: http://marinespecies.org/aphia.php?p=taxdetails&id=51. 21 August 2018.

Submitted: 27 August 2018; **accepted:** 9 November 2018;
final revisions received: 12 November 2018

APPENDICES

Appendix 1. Questionnaire – In the questionnaire, requested data were grouped by topics. The questionnaire was organized in a spread sheet format, prompting entries into cells next to the item in question. # = the questionnaire requested entering a number; several items requested a write-in reply or write-in narrative, here indicated by a colon:_____; comments in () or [] indicate explanations of the requested data type, *e.g.* a) skeleton (= partial) data only (*e.g.* name and basic locale only).

1) Name of collection and preferred acronym:
1a) Name of person completing the survey:
1b) Type of institution [*e.g.* private non-profit, State agency, university museum]:

2) Nature of Recent molluscan collection
a) Origins of collection, major highlights:
b) Other major collections absorbed:
c) Focus of collection:
d) Main users of collection:
e) Do you include fossils [*e.g.* Pleistocene material] in your holdings? [if so, explain and give numbers if possible]:

3) Size of collection
a) Cataloged material total
total # of specimen series/lots:
total # of specimens:
b) Cataloged material digitization [Digitization here refers to text data, not imaging]
total # of specimen series/lots databased:
total # of specimen series/lots not yet databased:
c) Uncataloged but sorted material awaiting cataloguing (aka "quality backlog")
total # of specimen series/lots:
total # of specimens:

4) Level of identification
a) Lots already cataloged
total # of specimen series/lots identified to species (or genus):
total # of specimen series/lots identified only to family level
total # of specimen series/lots unidentified
b) Uncataloged but sorted material awaiting cataloguing (aka "quality backlog")
total # of specimen series/lots identified to species (or genus):
total # of specimen series/lots identified only to family level:
total # of specimen series/lots unidentified:

5) Extent of data capture in previously digitized lots (see #3b)
a) skeleton (= partial) data only (*e.g.* name and basic locale only), # of records:
b) data captured but fields not standardized, # of records:
c) Darwin Core compliant), # of records:
d) georeferenced (with latitude & longitude), # of records:
f) database software used: write in_____

6) Data currently served online
a) URL:_____
b) How served:_____
c) Institutional ipt address:_____

7) Taxonomic break down
total lots of Gastropoda:
total lots of Bivalvia:
total lots of Cephalopoda:
total lots of Scaphopoda:
total lots of Polyplacophora:
total lots of Aplacophora:

8) Habitat break down
total # of lots of marine:
total # of lots of freshwater:
total # of lots of estuarine/brackish:
total # of lots of terrestrial:

9) Geographic composition of holdings (total lots)
a) North America: **marine** holdings
North Atlantic, total # of lots:
Gulf of Mexico, total # of lots:
Caribbean, total # of lots:
North Pacific, total # of lots:
b) North America: **freshwater and land**
Continental North America, total # of lots:
Caribbean Islands, total # of lots:
Central America, total # of lots:
c) South America, **marine holdings**
South Atlantic, total # of lots:
South Pacific, total # of lots:
d) South America: **freshwater and land**
total # of lots:
e) Other geographic strengths of the collection, please write in geographic regions or country names and give approximate number of lots

10) Major time periods of collecting (total lots)
<1850: total # of lots:__; 1850-1950: total # of lots:__; 1950-present: total # of lots:__;
Use acquisition date if collecting date unknown
Is the collection actively growing at present: YES/NO

11) Water depth (total lots)
0-50m: total # of lots:__; 50-100m: total # of lots:__; 100-500m: total # of lots:__; >500m: total # of lots:__;

12) Preservation
a) # of cataloged lots dry:
total # of lots:__;
fluid: total # of lots:__;
frozen: total # of lots:__;

DNA extracts: total # of lots:__;
b) # of quality backlog lots
dry: total # of lots:__;
fluid: total # of lots:__;
frozen: total # of lots:__;
DNA extracts: total # of lots:__;
Associated specialty collections:
c) # of cataloged lots: slides: # of slides; SEM stubs: # of stubs; Host-parasites: # of lots; Egg masses: # of egg masses
other (specify)
d) # of quality backlog lots
slides: # of slides; SEM stubs: # of stubs; Host-parasites: # of lots; Egg masses: # of egg masses
other (specify)
e) Describe associated images:
f) Describe associated additional metadata:

13) Type specimens
a) Number of primary type lots (holotypes, neotypes, lectotypes, syntypes)
b) Total number of type lots
c) Number of types imaged
d) Are images online? YES? NO?

14) Collection management
a) Describe collection management staff:
b) Describe data entry staff (*e.g.* resident fulltime, students, volunteers):
c) Describe resident taxonomic expertise in mollusks:

15) Most urgent needs?
a) What are your greatest curatorial needs?
16) How were the lot/specimens numbers generated? Mark all that apply: estimated; hand-counted; from digitized records

Appendix 2. Collection details and finding aids to mollusk collections. Collections were asked to identify significant holdings and donations and provide additional narratives to document the scope of holdings. These responses are excerpted here, arranged in alphabetical order.

AMNH: Grace G. Eddison, Helen DuShane, Katherine C. Vaught, Walter E. Sage collections. Primarily N & S Atlantic & Pacific Marine and North American Terrestrial; **ANSP**: *Manual of Conchology*; Patrick Center for Environmental Research; past and current curators, H.B. Baker, B. R. Bales, A.D. Brown, M. Cahill, A. R. Cahn, H. Crampton, R. Dillon, J.T. Gulick, S. S. Haldeman, A. Heilprin, J. & C. Hemmen, P. Hesse, S.D. Kaicher, H. Katori, I. Lea, H. Lewis, A. Morelet, M. Naide, O. Oswald, R. E. Petit, J. H. Redfield, S.R. Roberts, T. Say, J. Schwengel. M. A. Snyder, R. Swift, Lowe-Wollaston, C. M. Wheatley, J. Worsfold, C. B. Wurtz, Alfred University, Brooklyn Museum, Princeton University, Syracuse University, Trinity University, and the University of Pennsylvania collections. The oldest mollusk collection in US with extensive type material and a worldwide focus. **ARC:** Canadian North Atlantic Fishes and Invertebrates. Department of Fisheries and Oceans Canada collections, especially from Newfoundland branch. Material from Huntsman Marine Science Center projects (including Census of Marine Life and Marine Protected Area assessments). Significant collection of cephalopods. **ARK**: A. J. Brown collection; terrestrial snails from Arkansas. **AUMNH**: collections of faculty and students, Buck, G.I.K. Davis, Wittig-Skinner, Samford U., U. South Alabama. Focus on unionids and Indo-Pacific gastropods. **Beneski MNH**: recent mollusks reside in the Amherst Bunker and some were transferred to Harvard Museum of Natural History. **Berkshire**: Z. Crane, A.E. Craven, C.E. Bidwell, F. Lyman, E.S. Clark, A.L. Sosa, B. Proctor and B. Kellam collections. Specimens from southeast Asia, Cuban tree snails, mid-19thCentury western European land and freshwater shells and marine Mollusks of the North Atlantic. **BMS**: Imogene Robertson Collection of Marine, freshwater and terrestrial mollusks with a historical collection of freshwater mussels from the lower Great Lakes. **BMSM**: Colin Redfern Collection including material illustrated in Bahamian Seashells; Southwest Florida, Florida, Gulf of Mexico, Southeastern US, Caribbean, Western Atlantic, worldwide. **BPBM**: Ditlev D. Thaanum, Mangarevan and Micronesian Expedition, D'Alte A. Welch, William H. Meinecke Collections; Pacific focus with over 164,000 lots from Hawaii and over 65,000 lots from South Pacific islands. **Brazosport**: Mildred Tate, Alan Might and DeWitt Collections; display collection of 16,400 records, cataloged and digitized in Excel. **CASIZ:** Collections include historical and recent collections made from 1833 to present, rebuilt following 1906 earthquake. Expeditions focus on the Eastern Pacific, Alaska to the Galapagos and include the John Steinbeck and Ricketts Expedition. **CASPNNM**: William Stimpson, Frank C. Baker, Elizabeth Emerson Atwater, William Wirt Calkins, Ruthven Deane, J. H. Handwerk, R. L. Lea, Howard N. Lyon, E. N. King, J. W. Velie, Frank Morley Woodruff, and Mrs. E. C. Wiswall. **ChM**: Collections of E. Ravenel, Mazyck, Gibbs, D. Holt. Cosmopolitan collection with a recent shift to low country South Carolina species. **CLEV**: North American unionid bivalves. **CM**: George Clapp, Arnold Ortmann, Victor Sterki, S.T. Brooks, Jay Tripp, Fred Aldrich collections; strengths are eastern North America land snails, North America Unionidae, Sphaeriidae. **CMC**: L. Brand, T. Say, the Cincinnati Society of Natural History, U.P. and J.F. James and E.D. Cope Collections. Approximately 16,400 cataloged lots, especially unionid bivalves, with mid-West focus. **CMNML**: Frederick W. Schueler and F. Wayne Grimm, Eastern Ontario Biodiversity Museum collections; focus on Canadian and Arctic marine, terrestrial and freshwater mollusks. **DMF**: Anson A. Hinkley, Anthony L. Swinehart and Hillsdale College affiliated collectors. Marine gastropods (especially small-sized) from Long Key, FL, freshwater gastropods of Hillsdale County, MI and unionid bivalves. **DMNH:** J.E. du Pont, R.T. Abbott, N.M. Hepler, R.W. Jackson, E. Doremus, J. E. Norton, N.W. Lermond. Worldwide collection, with strengths in Indo-West Pacific, FL, and Delmarva landsnails. **DMNH-P**: E.P. Cheatum, J.H. McLellan collections. Texas landsnails. **DMNS**: worldwide marine, freshwater and terrestrial mollusks, recent large donations. **EKY**: focus on KY land and freshwater mollusks. **Everhart**: small uncataloged collection. **FMNH**: G. Arnemann, D. Beetle-Pillmore, W. Biese, C.F. Billups, T. & B. Burch, F. Button, P.P. Carpenter, P. Champe, W.J. Eyerdam, J. Ferriss, D.S. Franzen, A.L. Goodwin, E. Hall, E.E. Hand, L. Hubricht, A.J. Kohn, A.S. & A. Koto, F.F. Laidlaw, C & A. Lindar, Mizpah de Boe, C.D. Nelson, O. Park, F. Schilling, A. Stein, M. Teskey, H.J. Walter, G.R. Webb, W.F. Webb, d'A. Welch, W. Weyrauch, J. Zetek, past and present curators, University of Utah, Oberlin College, Cleveland Museum of Natural History, and World's Columbian Exposition collection. Worldwide in scope, all taxa and habitats. **FWM:** terrestrial gastropods from Southwest US and numerous Cretaceous fossils. **FWRI**: North Atlantic, Gulf of Mexico marine species. **GTMC-GMNH**: E. Keferl, J. Avise, Sapelo Island and Grays Reef collections. Freshwater and marine focus from 37 states and 42 countries. **HBOM**: largely HBOI research project collections, especially from submersibles. **Hefner**: Marine specimens from the Great Barrier Reef, Australia; Ohio unionids. **HMNS**: Pisor, Northwest Gulf of Mexico survey, Guido Poppe collections. Holdings are 40% Gulf of Mexico, 60% worldwide, with dry shells from Texas. **INHS**: University of Illinois Museum of Natural History, Frank Collins Baker, Anson A. Hinkley, Richard Ellsworth Call, Lorenzo E. Daniels, John Wesley Powell, Robert Kennicott, Max R. Matteson collections. Terrestrial and freshwater mollusks of the Midwest, especially freshwater bivalves, Venezuela specimens. **INSM**: Charles E. Russell mollusk collection. Eastern North American terrestrial and aquatic, worldwide marine. **ISM**: freshwater bivalves from the Midwestern United States. **JFBM**: D.W. Taylor collection. Freshwater bivalves and snails of Western North America. **LACM**: Alan Hancock, UCLA, CA Tech., Rae Baxter collections. Eastern Pacific, micro-gastropods, land snails. **LSUMG-I**: B.L.C. Wailes Collection, Stanford, Marco Taviani, H.V. Howe, LSU field trips and dissertation collections. Marine gastropods and bivalves primarily from the Gulf of Mexico. **MCPR**: UNC-Charlotte collection. Freshwater bivalves and land snails. **MCZ**: J.G. Anthony, S. Putzey, Alexander Agassiz and curator field collections. West Atlantic marine, Teredinidae/Pholadidae, deep-sea mollusks, Unionidae, Pleuroceridae, Achatinellidae, Cerionidae, and other pulmonates. **MLBeanLSM**: avocational collectors. 21,000 shells, mostly purchased. **MMNHC**: University of

Tennessee, Paul Parmalee, Bishop Stephen Elliott, Tennessee Technological University, Paul J. Adams collections. Freshwater bivalves and gastropods, land gastropods. **MMNS**: Mississippi Dept. of Wildlife, Fisheries, and Parks, Tulane University collections; US freshwater mollusks. **MPM**: Approx. 20,000 marine mollusks. **NCSM**: Institute of Marine Sciences, Hugh Porter, Charleston Museum, Duke Marine Lab, Appalachian State University, Shell Club collections. Southeast US, especially freshwater bivalves and gastropods. **NFM**: Approx. 7,000 wet and dry recent mollusk and Pleistocene subfossil lots. Newfoundland and Labrador with marine species from 0-4000 m (data from Cummings *et al.* 2009). **NHSM**: C.F Reed collection. Terrestrial snails from the mid-Atlantic, some Atlantic and South Pacific marine. **NJSM**: H. Richards Pleistocene fossils, some from Hudson River tunnel excavations. 5,000 cataloged recent mollusks. **NYSM**: Latham collection. New York species and North Atlantic marine specimens; freshwater collection growing fastest. **OGL**: mainly DNA extracts, additional holdings from Antarctica, Australia, New Zealand and Philippines. **OMNH**: focus on documenting Oklahoma freshwater mollusk diversity, with some 5,000-year-old specimens recovered from an archaeology survey. **OSAC:** Oregon State Arthropod collection houses a substantial terrestrial mollusk collection (C. Marshall pers. comm.). **OSUM**: Ohio Historical Museum, University of Wisconsin Stevens Point and other private collections. Mostly of freshwater specimens, especially bivalves. **PRI**: Gilbert D. Harris, University of Rochester, Syracuse University, Cornell University collections. Recent collection managed separately from fossil and includes Southeast US, Caribbean and Central New York. **RBCM**: Dr Frank Bernard, Dr. Ian McTaggart-Cowan, Robert Forsyth, Kristiina Ovaska collections. Focus on mollusks of British Columbia, Canada. **RMUW**: freshwater mollusks from Wisconsin. **ROM**: worldwide in scope, with strong holdings from North America especially Canada. **RSMAS**: RSMAS research vessels, C.T. Simpson, A. Olsson and D.R. Moore collections. Strong cephalopod holdings from tropical Atlantic, Caribbean, eastern tropical Pacific. **SBMNH:** S.S. Berry, Walter Miller, LeRoy Poorman, Kirstie Kaiser, Carol Skoglund collections. Vouchers for several books on marine bivalves and marine gastropods, strong cephalopod holdings. Eastern Pacific marine. Western North American terrestrial and freshwater. **SCSM:** York County Museum, private donors and field collections. Focus is Southeast U.S. **SDNHM**: H.N. Lowe collection. Focus on southern California and Baja Mexico, plus other international. **SIO-BIC** Verena Tunnicliffe, Bob Vrijenhoek, Scripps scientist collections. Focus on deep Pacific (>1,000 m) benthic invertebrates, chemosynthetic environments, and Antarctic expeditions. **SIO-PIC** John McGowan, John Wormuth, CalCOFI collections. Strong holdings of zooplankton including pteropods, heteropods, and cephalopods. **SMNC**: UNCC Marine Mollusk Collection. Focus of collection on Southeastern US marine mollusks and North Carolina terrestrial mollusks. **SUI**: Charles Cleveland Nutting, Bohumil Shimek, R.E. Call collections, plus materials from university and federal marine expeditions including Albatross, Bahama, Bay of Fundy, Barbados-Antigua, Fiji-New Zealand. **SFCMC**: focusing on material from Big Cypress National Preserve, De Soto National Memorial, and Biscayne, Dry Tortugas, and Everglades National Parks, with more than 73,000 biological specimens including mollusks. **SMM**: focus on freshwater mussels from Minnesota, with additional worldwide marine collection. **TNHC**: freshwater bivalves from Texas. **UAM:** University of Alaska Fairbanks, Auke Bay laboratories collection. Focus on marine invertebrates of Alaska. **UAZ**: Benton collection. Focus of on southwest United States, including the Sonoran Desert. **UCM**: Focus on marine and terrestrial gastropods and bivalves. **UCMP**: Carole Hickman, Jere Lipps, David Lindberg, Rudy Shohler, Ray Smith, and USGS Menlo Park collections. Focus on marine and non-marine taxa of the eastern Pacific from Chile to northern Alaska and east to the continental divide. **UF**: Fred Thompson, Gustav Paulay, Ken Emberton, Thomas Van Hyning and John Slapcinsky, Tulane University, University of Alabama, University of Miami, Rollins College, Florida Geological Survey, Beal-Maltbie, Weber, Charles Torrey Simpson, McGinty, and H. G. Lee collections. Especially strong in non-marine North America, Neotropics, Madagascar, and SE Asia as well as Florida and Pacific marine species. **UMAMZ**: UMass Amherst Natural History Collections. Focus on freshwater bivalves. **UMMZ**: Mina Winslow, Bryant Walker, Royal Ontario Museum, Stelfox, F.C. Baker collections. Focus on freshwater and land mollusks. **UMNH**: Janke Kolffe and Peter Hovingh collections. Focus on international, marine, and springs of intermountain west. **UNM(MSB),** UNM(MSB) [Parasites]: 50,000 lots from New Mexico and West Texas. Uncurated, but with field notes. Parasites: pulmonate mollusks collected during parasite surveys of Kenya and North America, plus some worldwide; 35% have parasites associated with them. **UNSM**: K. Perkins, K. Lingle S. Aughey, R.H. Wolcott, E.H. Barbour, C.C. Engberg, and Herman Miller collections. Land snails from Cuba 1940s-1950s. **University of Arkansas:** Dr. Seth Meek, A. J. Brown, collections. **UPRM**: Puerto Rican and Caribbean land snails. **USDA**: Collections from port interceptions, domestic & international surveys plus the Smithsonian Agricultural Collection. Focus of collection on invasive species, agricultural pests, disease vectors, domestic and international survey samples. **USNM**: Official repository for all U.S. Government expeditions and research. Curator collections including Dall, Bartsch, Rehder, Rosewater, Morrison, Roper, A. Binney, W.G. Binney, P.P. Carpenter, J.B. Henderson, P.M. Heude, J.G. Jeffreys, C.J. Maynard, I. Lea, R.E.C. Stearns, A.E. Verrill, R.S Houbrick. Global scope. **UTEP**: Faculty and student collections. Focus on freshwater mussels and terrestrial gastropods of the Chihuahuan desert and surrounding areas. **UWBM**: Young Naturalists Society, Trevor Kincaid, Kenneth and Lorraine Rhein, Phil Nudelman, Museum of Vancouver collections. Focus of collection on Northeast Pacific, Pacific Northwest, and Pacific Rim. **UWZM**: Brother Dutton, David A. Baerreis, Harold A. Mathiak collections. Collection focus on Wisconsin freshwater taxa. **VIMS**: Collections developed from historical trawl surveys, expeditions and research. Focus on marine mollusks from the Chesapeake Bay, East Coast and Atlantic Ocean. Wet collection. **VMNH**: Freshwater and land mollusks from Virginia. **WMI** (curated by INSM, see above): E. T. Cox, James Sampson, Caldwell plus exchange and purchase from L. G. Yates of California. Focus on eastern North American terrestrial, aquatic and marine coastal. **YPM**: U.S. Exploring Expedition, U. S. Fish Commission, A.E. Verrill, Heathcote Woolsey, C.E. Beecher, P.S. Remington, W. Clench; USFC, G.E. Pickford, K.J. Bush collections, Gray Museum/MBL, and Dartmouth Shell collection. Worldwide in scope with strengths in western North Atlantic marine taxa.

Appendix 3. Collection Sizes – U.S. and Canadian mollusk collections listed by size, based on number of cataloged lots reported in the 2017 survey. Numbers of cataloged lots reported by Solem (1975) and Cummings *et al.* (2009) provided for comparison. Some of Cummings *et al.*'s numbers come from 1996; "listed" means they included the collection but did not provide its size. Collection acronyms as in Table 1. Superscript notations: sp = number of specimens reported (instead of lots); 1 = no catalogue numbers assigned; bl = quality backlog included; d = digitized; U = Unionida only; LD = limited data; R&F = Recent and fossil; F = fossil only; E NA, only Eastern North American holdings shown; unknown = 'unknown' reported by the collection or data not available; * rounded calculation assuming same ratio of lots to specimens as in Solem (1975); ** rounded calculation assuming 10 specimens / lot on average across institutions, usually to convert lots to specimens, but sometime vice versa. The figure under total in the "Specimens per lot" column is the weighted average += collection incorporated by institution listed above. PRI's number for total types included fossils and so was excluded in the column total.

Collection	Solem 1975	Cummings *et al.* 1996/ 2009	2017 cataloged lots	2017 cataloged specimens	Specimens per lot	2017 backlog lots	2017 primary types	2017 total types
USNM	740,000	~850,000	1,081,000	*17,500,000	*16.2	100,000	12,200	23,000
ANSP	336,737	480,000	501,000	9,000,000	18.0	36,000	8,000	18,000
LACM	~160,000 [1]	470,000	500,000	4,500,000	9.0	~350,000	740	1448
UF	22,174	275,000	497,459	2,500,000	5.0	55,000	1,236	8,400
FMNH	179,000	300,000	382,000	4,368,000	11.4	47,000	434	6,000
MCZ	270,000	620,000	372,056	2,677,173	7.2	100,000	6,500	14,074
AMNH	175,000	278,000	319,000	2,141,051	6.7	12 cabinets	831	8,000
BPBM	160,000	248,000	300,000	6,000,000	20.0	~325,000	460	3,040
UMMZ	232,373	250,000	251,000	5,000,000	19.9	~10,000	>150	>2,000
DMNH	70,650	220,000	220,287	1,617,095	7.3	4,200	93	1,200
CM	unknown[1]	115,000	152,521	1,290,481	8.5	24,000	1,416	2,015
OSUM	30,000	143,000	120,180	1,442,160	12.0	20-30,000[sp]	3	150
BMSM		listed	119,822	480,000	4.0	3,178	18	210
SBMNH	52,000	200,000	100,170	**1,000,000		430,000	823	2,623
CMNML	70,632	~125,000	100,000	1,000,000	10.0	40,000	65	332
CASIZ	35,000	290,000	91,858	276,000	3.0	~280,000	1,876	4,618
+SU	38,000							
INHS	75,000	50,000	86,790	483,084	5.6	7,000	48	267
+UIMNH	32,216	32,000						
HMNS		~45,000	~69,000	500,000	7.2	~33,000	unknown	3200
SDNH	63,500	listed	~63,000	**630,000		unknown	134	850
NCSM		55,000	59,005	1,094,962	18.6	35,000	35	92
UWBM	11,000	29,321	55,000	220,000	4.0	20,000	1	1
FWRI		24,000	51,551	**520,000		~87,000	~50	not stated
YPM		50,000	50,000	**500,000		~60,000	not stated	not stated
UCM	26,800	45,000	~48,000	798,097	16.6	12 cabinets	not stated	1 cabinet
SIO-BIC			40,000	750,000	18.8	~5,000	100	600
ChM	27,700	listed	33,322	106,000 [bl]	3.2	not stated	not stated	not stated
Hefner			~30,000 [1]	**300,000		21,000	unknown	unknown
Brazosport			28,939	30,000	1.0	7,000	not stated	not stated
UWZM		listed	27,650	150,000	5.4	~2,000	2	44
RBCM-INVZ	10,244	listed	27,353	113,009	4.1	~5,500	7	75
PRI		23,000	26,867	88,293	3.3	27,236	27	2,251 [R&F]
ROM		50,000	23,800	**240,000		~16,000	4	13
UAM		listed	23,000	~140,000	6.1	~1,000	0	9

Appendix 3. (Continued)

Collection	Solem 1975	Cummings *et al.* 1996/ 2009	2017 cataloged lots	2017 cataloged specimens	Specimens per lot	2017 backlog lots	2017 primary types	2017 total types
AUMNH			20,500	50,000	2.4	200	not stated	not stated
MPM	18,390	**19,000	20,000	**200,000		0	0	2
RSMAS			19,649	**200,000		>2,000	0	270
JFBM		10,438	19,000	**190,000		~200	1	1
UTEP	3,650	13,000 R&F	18,165	363,300	20.0	7,000	0	6
DMNS		16,000	17,500	50,000	2.9	26,800	0	14
CASPNNM		20,000	15,369	116,642	7.6	0	7	57
UCMP			14,370	500,000	34.8	0	42	513
INSM		2,400 SP	8,191	**82,000		~5,400	0	not stated
WMI curated by INSM		40,000 SP	13,921	**140,000		~4-7,000	not stated	not stated
MMNS		50,000 SP	13,553	82,033		400	0	1
GTMC-GMNH		listed	~12,000	>50,000	4.2	10,000 SP	0	0
ISM		~10,000	11,736	14,700	1.3	~50,000 SP	not stated	not stated
USDA			11,500	~57,500	5.0	12,000	0	0
MMNHC		4,000	11,282	141,120	12.5	10,000	0	0
NYSM		23,500	10,000	30,000	3.0	2,500	not stated	not stated
EKY	~9,000	5,000	~10,000	**100,000		~3,000	not stated	not stated
UMNH			8,400	44,500	5.3	600	0	not stated
OMNH			8,057	**80,000		~1,000	0	0
ARK		listed	7,711	**77,000		unknown	not stated	not stated
DMNH-P	3,600	28,000	7,501	**75,000		~2,300	3	28
ARC		5,000	7,319	42,359	5.8	~300	0	0
SUI		2,000	5,000	**50,000		35,000	unknown	unknown
UNSM	22,000 bl	10,000	4,500	**45,000		~38,000	unknown	unknown
OGL			4,454	4,454	1.0	21	0	0
HBOM E NA			3,712	17,617	4.6	not stated	not stated	not stated
SMNC			4,067	18,700	4.6	76	0	0
SIO-PIC			~4,000	**40,000		~2,500	not stated	not stated
FWM	3,100	3,500	3,500	20,000	5.7	not stated	0	0
UAZ	12,758	listed	~3,500	**35,000		10,000	0	0
LSUMG-I			2,511	20,295	8.1	369	19 F	48 F
Berkshire			~2,300	11,000	4.8	~5,000	not stated	not stated
CLEV		6,300	2,280 U	7,217	3.2	200	0	0
VMNH		2,050	2,204	6,434	2.9	~800	1	7
NFM		7,000 bl					0	
UPRM			**1,907	19,070		not stated	not stated	not stated
CMC		16,400	1,712	**17,000		30,000	0	~33
RMUW			1,647	**16,000		not stated	not stated	not stated
BMS			~1,600	**16,000		~320	0	0
UNM(MSB) [P]			1,440 SP	1,440	1.0	280 SP	0	0
UNM(MSB)			0	0		50,000 SP	5	not stated

Appendix 3. (Continued)

Collection	Solem 1975	Cummings *et al.* 1996/ 2009	2017 cataloged lots	2017 cataloged specimens	Specimens per lot	2017 backlog lots	2017 primary types	2017 total types
SCSM			**1000 F SP	10,000 F SP		5,000 R&F SP	0	0
MCPR			558	**5,600		0	0	0
VIMS			~300	~1000	3.3	not stated	not stated	not stated
UMAMZ		1,130	280	**2,800		not stated	0	0
DMF			150-300	**2300		250-500	not stated	not stated
NHSM			0			~10,000 SP	0	not stated
Totals	**2,890,524**	**5,562,039**	**6,190,694**	**70,510,869**	**10.5**	**2,349, 055**	**35,312**	**101,220**

Appendix 4. Preservation Type. – Arranged by size as in Appendix 3. Main collection types are dry, wet (fluid preserved), and frozen tissue. Others include eggs, microscope slides of radula or histology, SEM stubs, DNA extracts and images. Abbreviations: bl = figure includes backlog lots; D = digitized lots; sp = number of specimens reported (instead of lots); "0" = collection stated that none are presently in the collection; blank: no data provided by institution.

	Cataloged			Backlog		Frozen	
Collection	**lots**	**dry lots**	**wet lots**	**dry lot**	**wet lots**	**lots**	**Other**
USNM	1,081,000	956,000	125,000	70,000	30,000	1000[bl]	DNA extracts, slides, SEM stubs, egg masses, images
ANSP	501,000	458,000	43,000	30,000	6,000	1000[bl]	images, slides, SEM stubs, egg masses, hosts
LACM	500,000	450,000	50,000				
UF	497,459	414,148	83,311	55,000	0	8,467	2,000 DNA extracts, slides, SEM stubs, hosts, egg masses
FMNH	382,000	345,000	37,000	45,000	2,000	2,000	323 DNA extracts, slides, SEM stubs, egg masses
MCZ	372,056	332,579	32,275	95,000	5,000	624	534 DNA extracts, slides, SEM stubs
AMNH	319,000	unknown	>2688	12 cabinets			AMNH
BPBM	300,000	198,300	61,300				
UMMZ	251,000	>124,000	4,554				15,000 specimen images
DMNH	220,287	219,829	458	4200			deep sea cephalopod images
CM	152,521	116,134	5,791	24,000			slides
OSUM	120,180	98,980	21,200				
BMSM	119,822	117,684	2,138	3,178			120 egg masses; 1000 images
SBMNH	100,170	79,668	7,971	420,000	10,000	10	30,000 images, slides, hosts, egg masses
CMNML	100,000	95,000	15,000	15,000	5,000		
CASIZ	91,858	26,146	65,712	190,714	17,428		slides, SEM stubs
INHS	86,790	72,484	14,306	4,000	2,000		>3,500 images
HMNS	~69,000	69,000					20 egg masses
SDNH	~63,000						
NCSM	59,005	43,945	15,660	20,000	15,000	75[bl]	25 DNA extracts, 162 host-parasites; 179 egg masses
UWBM	55,000	55,000	0	20,000	0	0	
FWRI	51,551	29,459	18,331			30	
YPM	50,000	45,000	5,000				100 egg masses ; 600 slides
UCM	~48,000	8,260	9,423				
SIO-BIC	40,000	1,416	~37,000		~5000	>100	20 egg masses
ChM	33,322[D]	~33,000					[collection largely dry lots]
Hefner	~30,000						
Brazosport	28,939	28,939		7000			
UWZM	27,650	27,650		~1500	~500	0	530 images
RBCM-INVZ	27,353	13,012	13,672				26 DNA extracts
PRI	26,867	26,867					images
ROM	23,800	10,000	3,800	5,000	2,000	6	
UAM	23,000	3,000	20,000		~1,000	500	
AUMNH	20,500	16,500	4,000	50	150		
MPM	20,000	18,000	2,000				
RSMAS	19,649	majority	few				slides, egg masses
JFBM	19,000	14,660	4,401	25	25		

Appendix 4. (Continued)

	Cataloged			Backlog		Frozen	
Collection	**lots**	**dry lots**	**wet lots**	**dry lot**	**wet lots**	**lots**	**Other**
UTEP	18,165	~16,349	~1,816	7,000			images
DMNS	17,500	17,098	15	26,800	63	0	
CASPNNM	15,369	15,370	25			0	egg masses
UCMP	14,370	13,973	376	0	unknown	21	images
INSM	8,191	8,191		~5400		0	
WMI, curated by INSM	13,921	13,921					
MMNS	13,553	9,778	3,775	100	300		
GTMC-GMNH	~12,000	9,000	3,000				
ISM	11,736	most					
USDA	11,500	~2875	~8625				
MMNHC	11,282	11,182	100	9500	500		
NYSM	10,000	7,000	1,000			1,000[bl]	2,000 DNA extracts
EKY	~10,000	most	< 50				
UMNH	8,400	6,400	2,000				
OMNH	8,057	most					
ARK	7,711	most	~200				
DMNH-P	7,501	7,501	0	unknown	0	0	
ARC	7,319	5	7,300		~300		
SUI	5,000	most	few	most	few	0	
UNSM	4,500	2087	6	38,000	0	0	
OGL	4,454 [SP]	5	299			2	4,379 DNA extracts
HBOM	4,355	1,925	1,647				
SMNC	4,067	3,500	750	60	16	0	
SIO-PIC	~4,000		~4000	0	~2,500		Some DNA extracts
FWM	3,500	most					
UAZ	3,500	2,500	1,000	1,000			
LSUMG-I	2,511	2510	1	366	3		
Berkshire	2,300	2,300		~5000			
CLEV	2,280	2,280		200			
VMNH	2,204	560	1501	~600	~100		
UPRM	1,907	1907					
CMC	1,712	~1700		~30,000		0	
RMUW	1,647						
BMS	~1,600	~1,600	0	320	0		
UNM(MSB) [P]	1,440 [SP]	5 [D]	1,186 [D]		100	0	associated parasites; DNA extracts
UNM(MSB)	0	0	0	30000 [SP]	20000 [SP]		
SCSM	1,000	1,000	0	500 [R&F]	0	0	
MCPR	558	~223	~335	0	0		
VIMS	~300	few	~300				
UMAMZ	280		280				
DMF	150-300						
NHSM	0	0	0	800	10		
Totals	**6,190,694**	**4,677,405**	**741,890**	**1,130,813**	**106,495**	**14,835**	

Appendix 5. Taxonomic composition. – Arranged by collection size as in Appendix 3. "0" means an institution reported not having (cataloged) specimens of the taxon; "–" and "[0]" mean the respondent left the field blank but the latter means that the count could be inferred to be zero as other columns sum to the expected number of lots; "?" means we could not fit the response to the table. Institutions with fewer than 40,000 lots that did not report a taxonomic break down are omitted from the table.

Collection	Cataloged lots	Gastropoda	Bivalvia	Cephalopoda	Scaphopoda	Polyplacophora	Aplacophora	Comment
USNM	1,081,000	784,000	198,000	27,000	9,000	9,500	3,000	
ANSP	501,000	391,024	70,184	775	1,670	4,255	0	
LACM	500,000	88,000	11,000	<1,000	<1,000	<1,000	<500	digitized only
UF	497,459	397,742	94,886	668	1,673	2,486	4	
FMNH	382,000	308,400	52,800	1,246	411	1,491	86	
MCZ	372,056	249,148	101,134	1,596	2,428	4,842	3	
UMMZ	251,000	78,000	57,079	–	–	137	–	
DMNH	220,287	162,805	56,660	458	166	198	[0]	
CM	152,521	122,682	28,505	104	139	601	1	
OSUM	120,180	28,190	92,000	[0]	[0]	[0]	[0]	
BMSM	199,822	95,176	22,596	598	681	771	–	
SBMNH	100,170	62,820	27,167	5,248	460	4,612	50	
CMNML	100,000	73,900	73,900	1,000	100	1,000	100	
CASIZ	91,858	299,286	61,000	1,685	1,103	11,276	650	includes backlog
INHS	86,790	36,595	49,982	11	34	148	0	
HMNS	~69,000	20,000	16,000	–	1,000	–	–	digitized only
NCSM	59,005	30,089	57,298	83	177	85	2	includes backlog
UWBM	55,000	62,000	12,000	40	130	830	0	includes backlog
FWRI	51,551	29,355	17,984	2,082	527	1,571	32	
YPM	50,000	65,000	30,000	5,000	nominal	nominal	nominal	includes backlog
UCM	~48,000	35,215	6,282	31	24	346	0	
SIO-BIC	40,000	7,154	2,561	529	159	201	68	digitized only
UWZM	27,650	17,901	8,671	–	–	–	–	
RBCM-INVZ	27,353	13,092	9,671	1,033	163	3,130	18	
PRI	26,867	22,900	5,600	50	100	200	<50	converted from specimens
ROM	23,800	25,000	12,000	175	225	250	50	includes backlog
UAM	23,000	5,817	3,677	139	85	195	14	
AUMNH	20,500	3,382	16,873	170	8	9	–	
MPM	20,000	12,000	8,000	[0]	[0]	[0]	[0]	
RSMAS	19,649	8,746	2,844	3,564	465	240	130	
JFBM	19,000	3,808	15,753	0	0	40	0	
UTEP	18,165	~16357	~7550	5	"1.44%"	"2.1%"	"1.44%"	includes backlog
DMNS	17,500	13,940	3,051	53	34	66	0	
CASPNNM	15,369	10,205	4,475	9	12	52	0	
UCMP	14,370	11,296	2,667	9	57	306	27	
INSM	8,191	3,547	4,190	7	70	377	[0]	
WMI, curated by INSM	13,921	3,025	13,415	8	20	30	–	includes backlog
MMNS	13,533	284	13,269	[0]	[0]	[0]	[0]	
GTMC-GMNH	~12,000	>599	>333	>7	>7	>13	–	

Appendix 5. (Continued)

Collection	Cataloged lots	Gastropoda	Bivalvia	Cephalopoda	Scaphopoda	Polyplacophora	Aplacophora	Comment
ISM	11,736	8,000[sp]	30,000[sp]	–	–	–	–	includes backlog
USDA	11,500	11,500	4	0	0	0	0	
MMNHC	11,282	3,958	7,324	[0]	[0]	[0]	[0]	
NYSM	10,000	5,604	2,900	24	6	57	0	
EKY	~10,000	3,000	7,000	–	–	–	–	
OMNH	8,057	4,800	3,080	22	26	121	0	
ARK	7,711	3,025	2,295	13	11	61	–	
DMNH-P	7,501	6,953	521	15	2	3	0	
ARC	7,319	1,748	2,233	3,118	56	118	44	
OGL	4,454	2,209	1,262	769	102	55	2	
HBOM	4,355	2,460	1,117	26	44	65	0	
SMNC	4,067	2,999	1,044	1	12	7	0	
SIO-PIC	~4,000	2,500	–	~4,000	–	–	–	backlog = gastropods
LSUMG-I	2,511	1,494	989	2	8	18	0	
CLEV	2,280	–	2,280	–	–	–	–	Unionida only
VMNH	2,204	332	1,584	0	0	0	0	
UPRM	1907	~1900	–	–	–	–	–	land snails
CMC	1,712	~14000	~14000	~300	~300	<300	0	includes backlog
RMUW	1,647	740	904	0	3	0	0	
BMS	~1,600	~1,300	~300	1	0	15	0	
UNM(MSB) [P]	1,440	1,186	5	–	–	–	–	
MCPR	558	324	234	0	0	0	0	
VIMS	300	150	125	15	?	?	?	
UMAMZ	280	50	200	30	0	0	0	
NHSM	0	8,000	2,000	1	50	20	0	backlog specimens
Totals	**5,388,055**	**3,686,714**	**1,352,724**	**63,046**	**22,748**	**51,092**	**4,794**	
Percentage		**71.2%**	**26.1%**	**1.2%**	**0.4%**	**1.0%**	**0.1%**	**100.0%**

Appendix 6. Number of lots by habitat. – Arranged by collection size as in Appendix 3. For Solem (1975) lots by habitat was determined by applying the percentages in his tables 1-3. This revealed a few discrepancies with the numbers he used for rankings on p. 229: FMNH was 27,000 not 29,000 for freshwater, BPBM was 112,000 not 120,000 for terrestrial and OSUM was 18,000 not 39,000 for freshwater. * indicates institution included backlog in their calculations for 2017.

	Cataloged lots		Marine		Freshwater		Terrestrial		Brackish
Collection	2017	1975	2017	1975	2017	1975	2017	1975	2017
USNM	1,081,000	740,000	766,000	444,000	95,000	148,000	190,000	148,000	
ANSP	501,000	336,737	225,094	192,000	76,110	54,000	169,784	91,000	1,280
LACM	500,000	160,000	80,000	152,000	4,000	1,600	15,000	6,400	1,000
UF	497,459	22,174	198,983	6,650	94,518	6,650	203,958	8,870	
FMNH	382,000	179,000	152,800	54,000	57,300	27,000	171,900	98,000	
MCZ	372,056	270,000	233,295	121,500	50,359	67,500	83,297	81,000	5,000
AMNH	319,000	175,000		131,000		9,000		35,000	
BPBM	300,000	160,000	68,000	40,000	3,300	8,000	180,000	112,000	1,900
UMMZ	251,000	232,373	25,000	11,600	138,000	139,000	88,000	81,500	
DMNH	220,287	70,650	132,200	42,400	33,000	14,100	44,000	14,100	11,000
CM *	152,521		27,069		92,350		56,265		0
OSUM	120,180	30,000	2,000	6,000	104,280	18,000	13,900	6,000	
BMSM	119,822		110,173		1,959		7,690		
SBMNH	100,170	52,000	91,474	39,000	615	2,600	8,827	10,400	
CMNML *	100,000	70,632	55,000	28,300	47,500	28,300	47,500	14,000	
CASIZ *	91,858	35,000	281,744	21,000	11,375	3,500	45,397	10,500	36,484
INHS	86,790	75,000	14,385	0	58,827	37,500	10,814	37,500	2,748
HMNS	69,000		50,000		10,000		5,000		
SDNH	63,000	63,500	47,500	48,000	3,000	3,000	12,500	12,500	
NCSM *	59,005		25,949		57,769		3,994		22
UWBM *	55,000	11,000	42,000	8,800	4,000	1,100	9,000	1,100	
FWRI	51,551		51,551						
YPM	50,000		42,500		5,000		2,500		
UCM	48,000	26,800		13,400		6,700		6,700	
SIO-BIC	40,000		40,000		0		0		0
Hefner	30,000		21,000		6,000				
RBCM-INVZ	27,353	10,244	22,609	10,244	2,362	0	2,279	0	62
PRI *	26,867		47,980		3,062		3,062		
ROM *	23,800		20,000		8,000		10,000		175
UAM	23,000		23,000						
AUMNH	20,500		8,000		12,500				
MPM	20,000	18,390	11,000		3,000		6,000		
RSMAS *	19,649		23,527		0		0		
UTEP *	18,165	3,650	7,046		5,033		13,086	3,650	
DMNS	17,500		15,739		143		1,050		0
INSM *	8,191		6,000		3,188		2,000		
WMI, curated by INSM	13,921		948		11,445		1,495		33
MMNS	13,553				13,547		6		
USDA	11,500		170		120		11,160		50
MMNHC	11,282		70		9,126		2,086		
NYSM	10,000		6,500		1,000		100		
EKY	10,000	9,000		0	2,000	1,800	8,000	7,200	
UMNH	8,400		5,880		1,680		840		
DMNH-P	7,501	3,600	1,641		1,670		4,179		
ARC	7,319		7,200						
OGL	4,454		4,454						
HBOM	4,355		4,355						

Appendix 6. (Continued)

	Cataloged lots		Marine		Freshwater		Terrestrial		Brackish
Collection	**2017**	**1975**	**2017**	**1975**	**2017**	**1975**	**2017**	**1975**	**2017**
SMNC	4,067		2,353		19		1,674		2
SIO-PIC	4,000		4,000						
FWM	3,500	3,100	700		700	few	2,100	3,100	
UAZ	3,500	12,758	2,500		500		500		
LSUMG-I	2,511		2,041		216		247		7
CLEV	2,280				2,280				
VMNH	2,204		18		1,808		90		
UPRM	1,907		0		0		1,907		0
CMC	1,712		340		685		685		
RMUW	1,647		241		1,405		0		
BMS	1,600		916		200		500		
UNM(MSB) [P]	1,440		76		1,160		76		144
SCSM	1,000		850		20		50		80
UMAMZ	280		100		180				
Totals	**5,999,657**	2,770,608	**3,013,971**	1,369,894	**1,041,310**	577,350	**1,442,497**	788,520	**59,987**
Percentage			**54.2%**	50.1%	**18.7%**	21.1%	**26.0%**	28.8%	**1.1%**

Appendix 7. Percentage of lots by habitat. – Under "Total of habitat", the 2017 column shows the sum of the marine, freshwater and terrestrial columns from App. 6, whereas the 1975 column repeats the cataloged lots column from App. 6.

	Cataloged	Total of habitat		Marine + brackish		Freshwater		Terrestrial	
Collection	2017	2017	1975	2017	1975	2017	1975	2017	1975
USNM	1,081,000	1,051,000	740,000	73%	60%	9%	20%	18%	20%
ANSP	501,000	472,268	336,737	48%	57%	16%	16%	36%	27%
LACM	500,000	100,000	160,000	81%	95%	4%	1%	15%	4%
UF	497,459	497,459	22,174	40%	30%	19%	30%	41%	40%
FMNH	382,000	382,000	179,000	40%	30%	15%	15%	45%	55%
MCZ	372,056	371,951	270,000	64%	45%	14%	25%	22%	30%
AMNH	319,000		175,000		75%		5%		20%
BPBM	300,000	253,200	160,000	28%	25%	1%	5%	71%	70%
UMMZ	251,000	251,000	232,373	10%	5%	55%	60%	35%	35%
DMNH	220,287	220,200	70,650	65%	60%	15%	20%	20%	20%
CM	152,521	175,684		15%		53%		32%	
OSUM	120,180	120,180	30,000	2%	20%	87%	60%	12%	20%
BMSM	119,822	119,822		92%		2%		6%	
SBMNH	100,170	100,916	52,000	91%	75%	1%	5%	9%	20%
CMNML	100,000	150,000	70,632	37%	40%	32%	40%	32%	20%
CASIZ	91,858	375,000	35,000	85%	60%	3%	10%	12%	30%
INHS	86,790	86,774	75,000	20%	0%	68%	50%	12%	50%
HMNS	69,000	65,000		77%		15%		8%	
SDNH	63,000	63,000	63,500	75%	75%	5%	5%	20%	20%
NCSM	59,005	87,734		30%		66%		5%	
UWBM	55,000	55,000	11,000	76%	80%	7%	10%	16%	10%
FWRI	51,551	51,551		100%		0%		0%	
YPM	50,000	50,000		85%		10%		5%	
UCM	48,000		26,800		50%		25%		25%
SIO-BIC	40,000	40,000		100%		0%		0%	
Hefner	30,000	27,000		70%		20%			
RBCM-INVZ	27,353	27,312	10,244	83%	100%	9%	0%	8%	0%
PRI	26,867	54,103		89%		6%		6%	
ROM	23,800	38,175		53%		21%		26%	
UAM	23,000	23,000		100%		0%		0%	
AUMNH	20,500	20,500		39%		61%		0%	
MPM	20,000	20,000	18,390	55%		15%		30%	
RSMAS	19,649	23,527		100%		0%		0%	
UTEP	18,165	25,165	3,650	28%		20%		52%	
DMNS	17,500	16,932		93%		1%		6%	
INSM	8,191	11,188		54%		28%		18%	
WMI, curated by INSM	13,921	13,921		7%		82%		11%	
MMNS	13,553	13,553		0%		100%		0.04%	
USDA	11,500	11,500		2%		1%		97%	
MMNHC	11,282	11,282		1%		81%		18%	
NYSM	10,000	7,600		86%		13%		1%	
EKY	10,000	10,000	9,000	0%	0%	20%	20%	80%	80%
UMNH	8,400	8,400		70%		20%		10%	
DMNH-P	7,501	7,490	3,600	22%		22%		56%	
ARC	7,319	7,200		100%		0%		0%	
OGL	4,454	4,454		100%		0%		0%	
HBOM	4,355	4,355		100%		0%		0%	
SMNC	4,067	4,048		58%		0%		41%	
SIO-PIC	4,000	4,000		100%		0%		0%	

Appendix 7. (Continued)

	Cataloged	Total of habitat		Marine + brackish		Freshwater		Terrestrial	
Collection	**2017**	**2017**	**1975**	**2017**	**1975**	**2017**	**1975**	**2017**	**1975**
FWM	3,500	3,500	3,100	20%	0%	20%	0%	60%	100%
UAZ	3,500	3,500	12,758	71%		14%		14%	
LSUMG-I	2,511	2,511		82%		9%		10%	
CLEV	2,280	2,280		0%		100%		0%	
VMNH	2,204	1,916		1%		94%		5%	
UPRM	1,907	1,907		0%		0%		100%	
CMC	1,712	1,710		20%		40%		40%	
RMUW	1,647	1,646		15%		85%		0%	
BMS	1,600	1,616		57%		12%		31%	
UNM(MSB) [P]	1,440	1,456		15%		80%		5%	
SCSM	1,000	1,000		93%		2%		5%	
UMAMZ	280	280		36%		64%			
Totals	**5,999,657**	**5,557,766**	**2,770,608**						

Appendix 8. Digitization of collections. – Arranged by collection size as in Appendix 3. If an institution reported only a percentage for georeferenced lots, a rounded number of lots calculated from this percentage is shown in square brackets. "Darwin core compliance" is reported only as a percentage since some institutions answered in terms of compliance by field rather than by lot or record.

Collection	2017 cataloged lots	2017 digitized lots	Digitized & georeferenced lots	% digitized lots georeferenced	Darwin core compliance	Data searchable online
USNM	1,081,000	~750,000	120,000	16%	100%	yes
ANSP	501,000	498,000	114,058	23%	100%	yes
LACM	500,000	100,000	[75,000]	75%	0%	no
UF	497,459	497,459	[234,000]	47%	100%	yes
FMNH	382,000	290,000	57,000	20%	100%	yes
MCZ	372,056	372,056	37,673	10%	100%	yes
AMNH	319,000	92,258	unknown		0%	types
BPBM	300,000	274,330	7,831	3%	100%	yes
UMMZ	251,000	~125,000	68,403	55%	not stated	yes
DMNH	220,287	198,000	1,014	0.5%	10%	yes
CM	152,521	152,521	29,246	19%	100%	no
OSUM	120,180	120,180		"all where possible for North America"	100%	yes
BMSM	119,822	~116,400	[10,600]	9%	100%	yes
SBMNH	100,170	~20,000	[19,600]	98%	most fields	yes
CMNML	100,000	42,663	31,079	73%	95%	yes
CASIZ	91,858	81,858	23,183	28%	100%	yes
INHS	86,790	86,790	77,715	91%	75% of fields	yes
HMNS	~69,000	~30,000	18,000	60%	0%	no
SDNH	~63,000	~30,000	0	0%	not stated	[no]
NCSM	59,005	32,938	[20,800]	63%	100%	yes
UWBM	55,000	3,300	0	0%	0%	yes
FWRI	51,551	51,551	most		not stated	yes
YPM	50,000	44,000	[11,000]	25% verified	100%	yes
UCM	~48,000	47,953	8,217	17%	many fields	[no]
SIO-BIC	40,000	38,642	27,236	71%	~75%	yes
ChM	33,322	33,322	not stated	not stated	0%	[no]
Hefner	~30,000	0	0	0%	0%	no
Brazosport	28,939	17,849	not stated	not stated	0%	[no]
UWZM	27,650	27,650	0	0%	0%	no
RBCM-INVZ	27,353	27,353	17,111	63%	0%	yes
PRI	26,867	3,288	950	29%	100%	yes
ROM	23,800	23,800	11,400	48%	7%	no
UAM	23,000	23,000	[18,400]	80%	100%	yes
AUMNH	20,500	20,500	2,280	11%	100%	[no]
MPM	20,000	0	0	0%	0%	yes
RSMAS	19,649	0	0	0%	not databased	[no]
JFBM	19,000	19,000	9,142	48%	not stated	yes
UTEP	18,165	11,576	3,026	26%	10%	yes
DMNS	17,500	17,500	1,200	7%	100%	yes
CASPNNM	15,369	15,369	0	0%	0%	yes
UCMP	14,370	14,370	2,012	14%	100%	yes
INSM	8,191	8,191	not stated	not stated	0%	no
WMI, curated by INSM	13,921	13,921	not stated	not stated	7%	no
MMNS	13,553	13,553	[12,200]	90%	0%	no
GTMC-GMNH	~12,000	6,000	[2,400]	~40%	0%	no
ISM	11,736	11,736	not stated	not stated	0%	yes
USDA	11,500	11,500	4500	39%	0%	no
MMNHC	11,282	11,282	6,258	55%	not stated	[no]

Appendix 8. (Continued)

Collection	2017 cataloged lots	2017 digitized lots	Digitized & georeferenced lots	% digitized lots georeferenced	Darwin core compliance	Data searchable online
NYSM	10,000	8,000	[2,000]	25%	not stated	no
EKY	~10,000	6,000	not stated	not stated	0%	[no]
UMNH	8,400	7,800	not stated	not stated	0%	no
OMNH	8,057	8,057	5,476	68%	unknown	yes
ARK	7,711	unknown	not stated	not stated	unknown	[no]
DMNH-P	7,501	7,501	0	0%	0%	no
ARC	7,319	7,319	5,315	73%	73%	yes
SUI	5,000	5,000	0	0%	0%	no
UNSM	4,500	2,131	0	0%	0%	no
OGL	4,454	4,443	4200	95%	100%	yes
HBOM	4,355	4,355	not stated	not stated	0%	no
SMNC	4,067	4,067	1,675	41%	0%	no
SIO-PIC	~4,000	~4,000	[2,400]	~60%	0%	[no]
FWM	3,500	3,500	not stated	not stated	unknown	[no]
UAZ	~3,500	0	0	0%	not databased	no
LSUMG-I	2,511	2,511	not stated	not stated	0%	yes
Berkshire	~2,300	~2,300	not stated	not stated	unknown	[no]
CLEV	2,280	2,280	not stated	not stated	0%	no
VMNH	2,204	2,204	not stated	not stated	0%	[no]
NFM				#N/A	#N/A	
UPRM	1,907	unknown	~1500	unknown	not stated	[no]
CMC	1,712	1,712		"minimal to none"	0%	no
RMUW	1,647	1,647	not stated	not stated	unknown	[no]
BMS	~1,600	0	0	0%	not databased	no
UNM(MSB) [P]	1,440 SP	1,191 SP	1130	95%	100%	yes
UNM(MSB)	0	0	0	0%	#N/A	no
SCSM	1000	0	0	0%	not databased	no
MCPR	558	234	0	0%	100%	no
VIMS	~300	unknown	not stated	not stated	unknown	[no]
UMAMZ	280	0	0	0%	not databased	no
DMF	150-300	150-300	not stated	not stated	0%	[no]
NHSM	0	0	0	0%	#N/A	[no]
Totals	**6,190,741**	**4,513,136**	**1,105,977**	**24.6%**	**20**	**34**

Appendix 9. Marine holdings by geographic regions. –Several collections could not provide regional data in the form requested for the survey. These collections, some of which undoubtedly hold material from these regions, are omitted from this table. * means backlog included; percentages were calculated excluding backlog.

		North America			South America			
Collection	**Marine cataloged**	**North Atlantic**	**Gulf of Mexico**	**Caribbean**	**North Pacific**	**South Atlantic**	**South Pacific**	**Comments**
USNM	766,000	128,000	33,000	20,000	158,000	11,000	70,000	
ANSP	225,094	22,867	10,431	16,865	19,227	4,606	3,823	
UF	198,983	70,655	5,995	16,001	10,595	2,270	3,496	
FMNH	152,800	25,003	12,149	452	4,883	1,200	1,700	
MCZ	233,295	43,293	5,808	14,045	14,908	3,061	1,745	
AMNH	69,000	10,300	3,099	1,062	3,716	7,800	207	Assumed 75% of digitized = marine
BPBM	68,000	6	67	2	1,723	4	22	
DMNH	132,200	31,510	13,000	15,119	1,450	2,250	1,700	
CM	26,569	4,709	5,064	3,350	7,156	827	352	*excludes* backlog
OSUM	2,000	0	350	450	0	0	0	
BMSM	110,173	10,862	14,104	15,211	9,797	2,826	845	
SBMNH	91,474	8,478	1,000	3,000	43,000	900	3,127	
CMNML *	55,000	15,000	100	10,000	2,000	10	0	
CASIZ *	281,744	500	1,500	4,000	144,000	100	4,000	
INHS	14,385	1,070	763	1,082	1,526	108	67	
HMNS	50,000	8,000	25,000	10,000	5,500	6,000	4,000	
NCSM *	25,949	25,069	376	75	424	3	2	
UWBM *	42,000	2,500	3,000	3,000	20,000	500	2,500	
FWRI	51,551	14,247	27,426	0	0	0	0	2,918 estuarine removed
SIO-BIC	40,000	45	41	1	10,903	34	0	
RBCM-INVZ	22,609	11	1	1	9,425	10	51	
PRI *	47,980	680	1,437	2,057	1,215	154	277	
ROM *	20,000	3,000	1,000	2,000	4,000	90	250	
AUMNH	8,000		5,000					
RSMAS *	23,527	11,000	2,000	8,000	1000		?	
UCM	unknown	7	8	?	26	1	71	1,006 lots Caribbean habitat not stated
UCMP	mostly	73	717	261	2,445	271	328	
JFBM	not stated	0	1	0	2	0	3	
UTEP *	7,046	705	1,057	1,057	705	1,409	2,114	
USDA	220			50	120			
DMNS	15,739	362	1,884	530	745	66	62	
NYSM	6,500	6,500	0	0	0	0	0	
NHSM *	8,000	4,000	500	2,000	50	100	1,000	specimens, not lots
DMNH-P	1,641	9	189	20	5	?	?	
ARC	7,200	6,302	0	2	19	0	0	
OGL	4,454	251	183	6	1,040	223	0	
SMNC	2,353	1,829	249	160	25	11	58	
SIO-PIC	4,000		some	0	4,100	0	0	
HBOM	4,355	3685	31	639				
LSUMG-I	2,041	68	451	57	421	1	4	
VMNH	18	18	0	0	0	0	0	
RMUW	241	50	0	0	0	0	0	
Totals	2,822,141	460,664	176,981	150,555	484,151	45,835	101,804	
Without backlog	2,310,895	398,210	166,011	118,366	310,757	43,469	91,661	
Percentage		17.2%	7.2%	5.1%	13.4%	1.9%	4.0%	48.8%

Appendix 10. Non-marine holdings by geographic regions. – Several collections could not provide regional data in the form requested for the survey. These collections, some of which undoubtedly hold material from these regions, are omitted from this table. * means backlog included; percentages were calculated excluding backlog.

		North America				
Collection	**Non-marine cataloged**	**U.S., Canada & Mexico**	**Caribbean Islands**	**Central America**	**South America**	**Comments**
USNM	285,000	129,000	50,000	3,600	8,000	
ANSP	245,894	92,334	36,168	2,423	8,577	
UF	298,476	93,273	19,504	18,381	3,286	
FMNH	229,200	92,000	12,000	7,000	10,000	
MCZ	133,656	47,728	13,445	725	3,095	
BPBM	183,300	6	170	77	5	
UMMZ	226,000	107,599	477	3,137	1,036	
DMNH	77,000	17,414	4,825	300	2,608	
CM	125,115	83,374	3,439	474	2,680	*excludes* backlog
OSUM	118,180	117,830	95	120	135	
BMSM	9,649	4,319	1,211	930	1,267	
SBMNH	9,442	7,200	590	515	228	
CMNML *	95,000	88,900	1,000	10	500	
CASIZ *	56,772	47,406	750	100	3,850	
INHS	69,641	65,381	1,206	1,243	646	
HMNS	15,000	13,500	4,500	1,000	1,000	
NCSM *	61,763	61,506	100	3	154	
UWBM *	13,000	2,000	500	500	2,500	
UCM	unknown	36922	?	526		1,006 lots Caribbean habitat not stated
RBCM-INVZ	4,641	4,318	35	72	0	
PRI *	6,123	705	140	11	54	
ROM *	18,000	14,000	200	30	125	
AUMNH	12,500	10,000				
JFBM	not stated	15,700	0	0	0	
UTEP *	18,119	17,756	362	0	0	
DMNS	1,193	444	509	49	60	
UCMP	not stated	31	2	0	1	
WMI, curated by INSM	12,940	12,428				
MMNS	13,553	13,553	0	2	9	
USDA	11,280	450	700	300		
MMNHC	11,212	11,000			20	
NYSM	1,100	1,000	0	0	0	
EKY	10,000	10,000				
DMNH-P	5,849	5,514	41	?	?	
SMNC	1,693	1,694	1	0	0	
LSUMG-I	463	148	12	0	3	
CLEV	2,280	2,221	0	7	0	
VMNH	1,898	1,894	0	0	0	
UPRM	1,907	50	1800	100		converted from specimens, 10:1
RMUW	1,405	1,405	0	0	0	
UNM(MSB) [P]	1,236	609	68		111	
NHSM *	1,800	300	50	100	50	specimens, not lots
Totals	2,391,280	1,232,912	153,900	41,735	50,000	
Without backlog	2,120,703	1,000,339	150,798	40,981	42,767	
Percentage		47.6%	7.1%	1.9%	2.0%	58.6%