



<https://doi.org/10.11646/zootaxa.4247.1.10>

<http://zoobank.org/urn:lsid:zoobank.org:pub:BA3B5F19-132A-41CB-9F67-A6B5876A0826>

LepNet: The Lepidoptera of North America Network

KATJA C. SELTMANN¹, NEIL S. COBB^{2,30}, LAWRENCE F. GALL³, CHARLES R. BARTLETT⁴, M. ANNE BASHAM²⁹, ISABELLE BETANCOURT⁵, CHRISTY BILLS⁶, BENJAMIN BRANDT², RICHARD L. BROWN⁷, CHARLES BUNDY⁸, MICHAEL S. CATERINO⁹, CAITLIN CHAPMAN², ANTHONY COGNATO¹⁰, JULIA COLBY¹¹, STEPHEN P. COOK¹², KATHRYN M. DALY¹³, LEE A. DYER¹⁴, NICO M. FRANZ²⁹, JON K. GELHAUS⁵, CHRISTOPHER C. GRINTER¹⁵, CHARLES E. HARP¹⁶, RACHEL L. HAWKINS¹⁷, STEVE L. HEYDON¹⁸, GEENA M. HILL¹⁹, STACEY HUBER¹⁹, NORMAN JOHNSON²⁰, AKITO Y. KAWAHARA¹⁹, LYNN S. KIMSEY¹⁸, BORIS C. KONDRATIEFF¹⁶, FRANK-THORSTEN KRELL²¹, LUC LEBLANC¹², SANGMI LEE²⁹, CHRISTOPHER J. MARSHALL²², LINDSIE M. MCCABE², JOSEPH V. MCHUGH²³, KATRINA L. MENARD²⁴, PAUL A. OPLER¹⁶, NICOLE PALFFY-MUHORAY³, NICK PARDIKES¹⁴, MERRILL A. PETERSON²⁶, NAOMI E. PIERCE¹⁷, ANDRE POREMSKI²⁶, DEREK S. SIKES¹³, JASON D. WEINTRAUB⁵, DAVID WIKLE², JENNIFER M. ZASPEL^{11,27} & GREGORY ZOLNEROWICH²⁸

¹*Cheadle Center for Biodiversity and Ecological Restoration, University of California Santa Barbara, Santa Barbara, CA 93106*

²*Merriam-Powell Center for Environmental Research and Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011 U.S.A.*

³*Entomology Division, Peabody Museum of Natural History, Yale University, New Haven, CT 06511 U.S.A.*

⁴*Entomology and Wildlife Ecology, University of Delaware, Newark, DE 19716 U.S.A.*

⁵*Entomology Department, Academy of Natural Sciences, Philadelphia, PA 19103 U.S.A.*

⁶*Natural History Museum of Utah, University of Utah, Salt Lake City, UT 84108 U.S.A.*

⁷*Mississippi Entomological Museum, Mississippi State University, Mississippi State, MS 39762 U.S.A.*

⁸*Department of Entomology, Plant Pathology, and Weed Science, New Mexico State University, Las Cruces, NM 88003 U.S.A.*

⁹*School of Plant and Environmental Sciences, Clemson University, Clemson, SC 29634 U.S.A.*

¹⁰*Department of Entomology, Michigan State University, East Lansing, MI 48824 U.S.A.*

¹¹*Invertebrate Zoology, Milwaukee Public Museum, Milwaukee, WI 53233 U.S.A.*

¹²*Department of Plant, Soil and Entomological Sciences, University of Idaho, Moscow, ID 83844 U.S.A.*

¹³*University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, AK 99775 U.S.A.*

¹⁴*Biology Department, University of Nevada-Reno, Reno, NV 89557 U.S.A.*

¹⁵*California Academy of Sciences, San Francisco CA 94118 USA.*

¹⁶*Department of Bioagricultural Science and Pest Management, Colorado State University, Fort Collins, CO 80521 U.S.A.*

¹⁷*Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138 U.S.A.*

¹⁸*Department of Entomology, University of California Davis, Davis, CA 95616 U.S.A.*

¹⁹*McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, FL 32611 U.S.A.*

²⁰*C.A. Triplehorn Insect Collection, The Ohio State University, Columbus, OH 43212 U.S.A.*

²¹*Department of Zoology, Denver Museum of Nature & Science, Denver, CO 80205 U.S.A.*

²²*Oregon State Arthropod Collection, Oregon State University, Corvallis, OR 97331 U.S.A.*

²³*Department of Entomology, University of Georgia, Athens, GA 30602 U.S.A.*

²⁴*Sam Noble Museum, University of Oklahoma, Norman, OK 73072 U.S.A.*

²⁵*Biology Department, Western Washington University, Bellingham, WA 98225 U.S.A.*

²⁶*Change Theory, Washington, DC 20001 U.S.A.*

²⁷*Department of Entomology, Purdue University, West Lafayette, IN 47907 U.S.A.*

²⁸*Department of Entomology, Kansas State University, Manhattan, KS 66506 U.S.A.*

²⁹*School of Life Sciences, Arizona State University, Tempe AZ 85287 U.S.A.*

³⁰*Corresponding author. E-mail: neil.cobb@nau.edu*

The Lepidoptera of North America Network, or LepNet, is a digitization effort recently launched to mobilize biodiversity data from 3 million specimens of butterflies and moths in United States natural history collections (<http://www.lepnet.org/>). LepNet was initially conceived as a North American effort but the project seeks collaborations with museums and other organizations worldwide. The overall goal is to transform Lepidoptera specimen data into readily available digital formats to foster global research in taxonomy, ecology and evolutionary biology.

Why Lepidoptera of North America? Lepidoptera are disproportionately common in both large and small research collections across the continent, with an estimated total of 17 million specimens, and their popularity has made them among the best known and most collected of all insects worldwide (Kawahara & Pyle 2012). Although their abundance in natural history collections is high, the amount of digitized specimen data for Lepidoptera is proportionately low (particularly in comparison to plants and vertebrates). LepNet aims to rectify this deficiency, as butterflies and moths are an obvious group to support data-driven research.

Insect herbivores and their host plants dominate terrestrial biomes, and as both herbivores and pollinators, Lepidoptera are one of the most important insect orders linked to the radiation of flowering plants (e.g., Menken *et al.* 2009). They are also among the most diverse and publicly recognizable insects, with more than 157,000 described species globally (e.g., van Nieuwerkerken 2011) and over 14,000 species in 86 families in North America alone. Lepidoptera are among the most impactful groups of insect pests to agriculture, and are the primary source of food for many vertebrates as well as hosts of parasitoid wasps and flies (e.g., Wagner 2001). Lepidoptera have served as model systems for studies of genetics, physiology, development, and many aspects of ecology and evolutionary biology including insect/plant coevolution, conservation biology and biogeography (e.g., Monteiro & Pierce 2001, Als *et al.* 2004, Braby *et al.* 2006, Canfield *et al.* 2008, Roe *et al.* 2009, Vila *et al.* 2011, Talavera *et al.* 2013). Since the pioneering work of Ehrlich and Raven (1964) on the coevolution of butterflies and their hosts, there has been great interest in trying to detect and understand macroevolutionary patterns in insect-plant associations (review by Janz 2011). Recent studies are using morphological and/or molecular data from collections to construct phylogenies of Lepidoptera and determine the evolution of host associations, as well as examine continental scale data to understand broad patterns in herbivore community dynamics (e.g., Hamback *et al.* 2007, Ockinger *et al.* 2010).

We have the opportunity to enable data synthesis research with the largest clade of herbivores, and the data are ready and waiting on specimens and their labels in natural history collections. At present, fewer than 10 percent of the North American species have sufficiently accessible digital occurrence data to enable such research. Moreover, public policy and expenditures are currently being driven by incomplete and/or geographically constrained studies. Thus, integration of existing but currently unconnected digitization efforts in North America is urgently needed to realize the outstanding potential of Lepidoptera and translate the data into transformative research and outreach (Cobb *et al.* 2016).

Under the lead of Northern Arizona University, LepNet currently comprises 27 collaborating collections across 26 US states (Figure 1) with collective holdings of approximately 6 million Lepidoptera specimens. Funding for LepNet is being provided from 2016 to 2020 by the National Science Foundation (NSF) through its Advancing Digitization of Biodiversity Collections (ADBC) program. The ADBC funds two main initiatives: a national hub that coordinates US collections digitization efforts, and specific collaborative digitization projects called Thematic Collections Networks (TCNs). LepNet is one of 18 currently active TCNs, which range widely in research focus. Integrated Digitized Biocollections (iDigBio) is the national ADBC hub and is based at the University of Florida. iDigBio supports national digitization by fostering partnerships and integrations among TCN participants, establishing best practices and workflows, disseminating the data and products generated, and promoting use of biodiversity data by the scientific community and other stakeholders. Among the TCNs, LepNet is well positioned to engage a broadly representative community of researchers and to serve as a model for additional large arthropod-based digitization projects. Although the ADBC program has achieved considerable and rapid progress, the challenge with Lepidoptera reflects the obstacles involved in digitizing the 287 million arthropod specimens estimated to reside in North American collections. Of the 74.7 million existing iDigBio records, only 15 percent are arthropods, and Lepidoptera currently comprise 838,000 records of those, which is just 1 percent of the overall iDigBio total (data as of 15 January 2017).

LepNet has four interrelated goals for digitization. The first goal is databasing and mobilizing label information from over 1.7 million specimens in the 27 partnering collections, and integrate these with more than 1 million existing records. Databasing in LepNet consists of digitally capturing the information from specimen labels and georeferencing as many as possible of the geographic localities represented, using a combination of automated and manual pipelines. The second goal is databasing 35,000 larval samples with associated host plant data, which will mark the first significant digitization of Lepidoptera immatures from North American collections. The third goal is generating 214,000 images of Lepidoptera specimens. Approximately 80,000 will be high-resolution (digital single-lens reflex) camera images, at a species-exemplar level, that will document over half of the Lepidoptera species diversity in North America. The remainder will be smartphone-quality images for selected lepidopteran groups, which will be represented by larger numbers of images per species; these will include butterflies, distinctive moths, and species exhibiting geographic variation. Collectively these digitization targets are aimed at elevating a projected 5,000 North American Lepidoptera species to a "research ready" status suitable for complex, data-driven analyses. The fourth goal is making all images and data available online and available as downloadable datasets through the LepNet, iDigBio, and GBIF data portals. All three portals will frequently be updated with new data from collections to keep the public datasets current.

LepNet Research Collections

www.lep-net.org

1,740,047 Total Specimen Records
214,705 Total Specimens Imaged

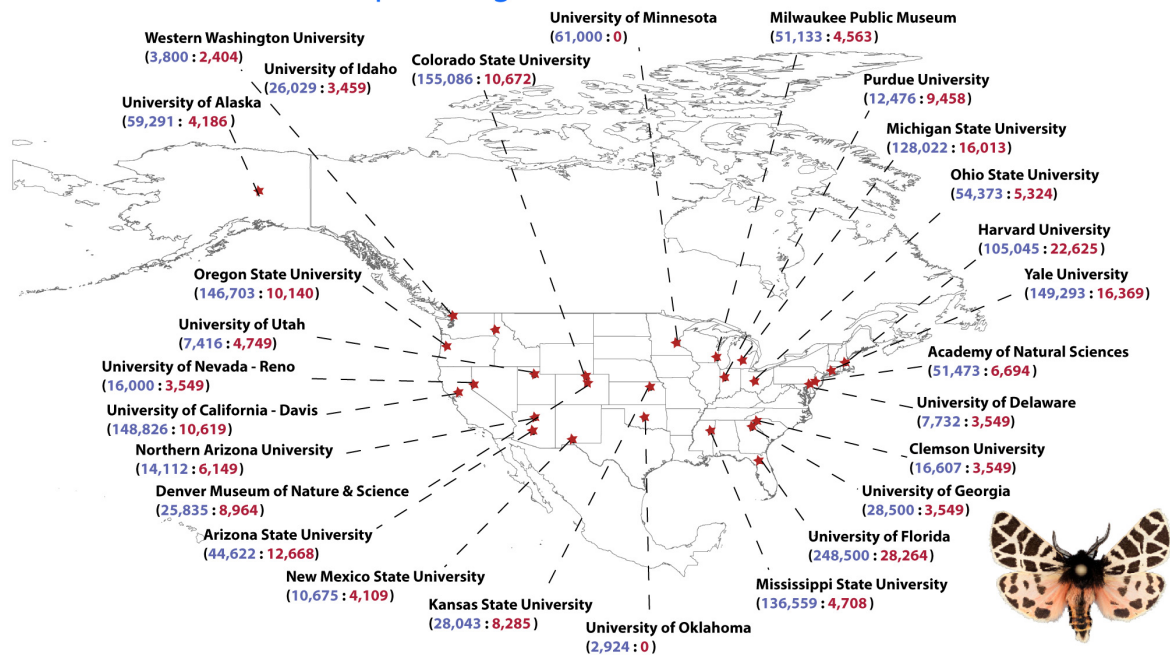


FIGURE 1. Geographic distribution of the 27 partner institutions in LepNet and their digitization targets for specimen label data (blue) and specimen images (red).

To achieve the LepNet goals, the 27 partnering institutions are using a variety of existing successful databasing and imaging solutions in their respective collections, while adhering to a set of common standards and principles developed through iDigBio and prior TCN experience. In addition, LepNet is using an integrated Symbiota software portal (Gries *et al.* 2014; <http://symbiota4.acis.ufl.edu/scan/lepnet/portal/index.php>) that allows direct input of specimen data and images by those partner institutions lacking suitable digital infrastructure. LepNet data from all 27 partners are aggregated in the Symbiota portal, and this portal in turn provisions iDigBio and GBIF.

LepNet is also collaborating with Visipedia (<http://www.visipedia.org>) and Fieldguide (<https://fieldguide.net>) in developing LepSnap, a computer vision project that includes a smartphone app and an identification widget incorporated into websites (van Horn *et al.* 2015). LepSnap will facilitate taxonomic identifications of images and provide tools for improved image searching. LepSnap is an insect-based extension of the Visipedia collaboration with the Cornell Laboratory of Ornithology (Merlin). In a pilot project with 62 species of tiger moths (Arctiinae) from the Pacific Northwest, over 75 percent of the tested images were correctly identified to the respective species with 80-100 percent probability based solely on a photo. Accuracy is expected to improve as LepNet produces more training images per species and data driven ecological filters are assembled (e.g., phenology, geography). The capacity to identify specimens quickly and efficiently by simply taking a photo or dragging an image file onto the website widget and submitting it to LepSnap for identification, will greatly aid the LepNet contributor and user communities in providing determinations and improving data usability.

We estimate that there are approximately 6 million fully curated and ready-to-database Lepidoptera specimens housed among the 27 LepNet partner collections, about 4 million of which are from North America. With the seed funding provided by the NSF-ADBC program we expect most of these specimens to be digitized in the next 10 years. LepNet is therefore proposing to database a significant amount of the estimated 17 million specimens held in North American collections. The ADBC effort will cover 90 percent of the Lepidoptera specimens in the 22 smaller partner collections, and up to 50 percent of the five largest partner collections. We expect to sponsor at least four additional proposals from museums that want to become funded LepNet partners through the NSF ADBC "Partners to Existing Networks" (PEN) program. The PEN projects should yield an additional 500,000 records by 2020. LepNet will be digitizing Lepidoptera in proportion to the number of specimens housed in the partner collections. In order to obtain sufficient data across all 86 Lepidoptera families, our approach is skewed towards moths, which comprise 94 percent of

North American lepidopteran species-level diversity—butterflies comprise just 6 percent of our fauna, but represent 40 percent of the existing records in iDigBio. Despite this imbalance, we will digitize butterflies in somewhat greater proportion than their diversity, in part because the life histories of butterflies are better known than any other group of arthropods and they are exceedingly popular in outreach, citizen science and the public. Overall, we are targeting approximately 1.2 million records for moths and 500,000 records for butterflies (Figure 2).

The expected impacts of LepNet beyond enabling biodiversity research will be substantial. The charisma of butterflies and moths profoundly inspires children and adults, and promotes public understanding of their relevance like no other arthropod group. The visibility and beauty of many moths and butterflies have captivated amateur collectors and professional entomologists for centuries, thus providing a unique foundation for influencing education, public awareness, and conservation (Kawahara & Pyle 2012).

Collectively, the LepNet partner institutions operate 67 existing outreach and education programs, including summer camps, annual events, and workshops that reach a diverse nationwide audience of 2.5 million people per year. Many of these programs target non-traditional and underserved audiences beyond the university setting. During the grant period, LepNet's educational program “Explore more with LepXplor” will connect marginalized learners in formal/informal settings to Lepidoptera data, collections and related key concepts (e.g., host plant associations). Participants will learn about Lepidoptera diversity via an augmented reality (AR) application that overlays digital information about specimen collections via tangible AR cards. This specimen-based learning tool aims to promote both STEM literacy and English acquisition for English language learners in adult literacy programs and K-12 classrooms. English language learners often struggle with academic language, and the integration of immersive technologies can directly impact academic achievement and motivation in learning (Solak & Cakir 2015). It is anticipated that more than 400 students and 3,500 volunteers will be involved directly with LepXplor, with public engagement channeled through the respective institutional outreach programs (e.g., collaborating with educators to integrate curriculum with United States federal and state-level education standards). These efforts will simultaneously emphasize the importance of scientific natural history collections in supporting collection-based learning experiences, and foster understanding of environmental impacts, climate change and biodiversity.

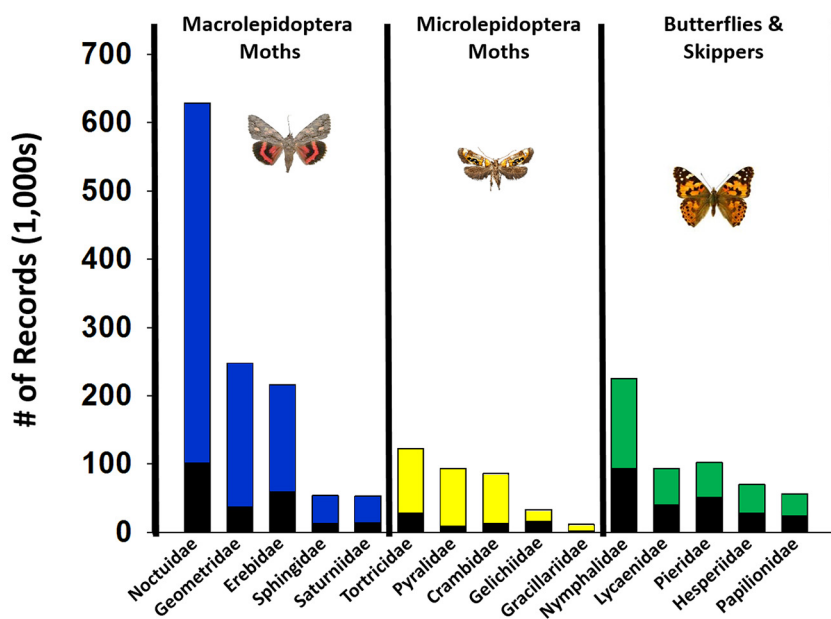


FIGURE 2. Existing (black) and expected (blue, yellow, and green) number of records to be produced by LepNet for the 3 major groups and 15 families of Lepidoptera that are most common in the collections of the partner institutions. Existing records were obtained from the SCAN and iDigBio data portals.

Acknowledgements

Many people provided advice and support in the development of LepNet, especially Ed Gilbert, Christine Johnson, Scott Miller, John Pickering, Greg Pohl, Taylan Prince, Alma Solis, Randall T. Schuh and Richard Zak. We are also grateful to the Visipedia team of Grant van Horn, R. J. Antonello, Serge Belongie, and Pietro Perona, without whose collaboration we could not have developed LepSnap. Lastly, we thank the National Science Foundation for providing financial support

for the LepNet project through the following DBI awards: 1602081, 1600937, 1601888, 1601369, 1601002, 1600774, 1600616, 1601957, 1601275, 1601164, 1601124, 1601659, 1600556, 1600824, 1601443, and 1601461.

Literature cited

- Als, T.D., Vila, R., Kandul, N.P., Nash, D.R., Yen, S.-H., Hsu, Y.-F., Mignault, A.A., Boomsma, J.J. & Pierce, N.E. (2004) The evolution of alternative parasitic life histories in large blue butterflies. *Nature*, 432, 386–390.
<https://doi.org/10.1038/nature03020>
- Braby, M.F., Vila, R. & Pierce, N.E. (2006) Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zoological Journal of the Linnean Society*, 147, 239–275.
<https://doi.org/10.1111/j.1096-3642.2006.00264.x>
- Canfield, M.R., Green, E., Moreau, C.S., Chen, N. & Pierce, N.E. (2008) Exploring phenotypic plasticity and biogeography in emerald moths: a phylogeny of the genus *Nemoria* (Lepidoptera: Geometridae). *Molecular Phylogenetics and Evolution*, 49, 477–487.
<https://doi.org/10.1016/j.ympev.2008.07.003>
- Cobb, N.E., Seltmann, K.C. & Franz, N.M. (2016) The current state of arthropod biodiversity data: addressing impacts of global change. Available from: http://www.lep-net.org/wp-content/uploads/2017/01/Cobb-Seltmann-Franz_-ICE_2016-sept2-1.pdf (accessed 15 January 2017)
- Ehrlich, P.R. & Raven, P.H. (1964) Butterflies and plants: a study in coevolution. *Evolution*, 18, 586–608.
<https://doi.org/10.2307/2406212>
- Gries, C., Gilbert, E.E. & Franz, N.M. (2014) Symbiota – a virtual platform for creating voucher-based biodiversity information communities. *Biodiversity Data Journal*, 2, e1114.
<https://doi.org/10.3897/bdj.2.e1114>
- Janz, N. (2011) Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annual Review of Ecology, Evolution and Systematics*, 42, 1–512.
<https://doi.org/10.1146/annurev-ecolsys-102710-145024>
- Hambäck, P.A., Summerville, K.S., Steffan-Dewenter, I., Krauss, J., Englund, G. & Crist, T.O. (2007) Habitat specialization, body size, and family identity explain lepidopteran density-area relationships in a cross-continental comparison. *Proceedings of the National Academy of Sciences*, 104, 8368–8373.
<https://doi.org/10.1073/pnas.0611462104>
- Kawahara, A.Y. & Pyle, R.M. (2012) An appreciation for the natural world through collecting, owning and observing insects. In: Lemelin, R.H. (Ed.), *The Management of Insects in Recreation and Tourism*. Cambridge University Press, Cambridge, pp. 138–152.
<https://doi.org/10.1017/cbo9781139003339.011>
- Menken, S.B.J., Boomsma, J.J. & van Nieukerken, E.J. (2009) Large-scale evolutionary patterns of host plant associations in Lepidoptera. *Evolution*, 64, 1098–1119.
<https://doi.org/10.1111/j.1558-5646.2009.00889.x>
- Monteiro, A. & Pierce, N.E. (2001) Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1alpha gene sequences. *Molecular Phylogenetics and Evolution*, 18, 264–281.
<https://doi.org/10.1006/mpev.2000.0872>
- Ockinger, E., Schweiger, O., Crist, T.O. & Bommarco, R. (2010) Life-history traits predict species responses to habitat area and isolation: a cross-continental synthesis. *Ecology Letters*, 13, 969–979.
<https://doi.org/10.1111/j.1461-0248.2010.01487.x>
- Roe, A.D., Weller, S.J., Baixeras, J., Brown, J., Cummings, M.P., Davis, D.R., Kawahara, A.Y., Parr, C.S., Regier, J.C., Rubinoff, D., Simonsen, T.J., Wahlberg, N. & Zwick, A. (2009) Evolutionary framework for Lepidoptera model systems. In: Goldsmith, M.R. (Ed.), *Genetics and Molecular Biology of Lepidoptera*. CRC Press, Boca Raton, pp. 1–24.
<https://doi.org/10.1201/9781420060201-c1>
- Solak, E. & Cakir, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *The Journal of Educators Online*, 13 (2), 50–72.
<https://doi.org/10.9743/jeo.2015.2.5>
- Talavera, G., Lukhtanov, V.A., Pierce, N.E. & Vila, R. (2013) Establishing criteria for higher-level classification using molecular data: the systematics of *Polyommatus* blue butterflies (Lepidoptera, Lycaenidae). *Cladistics*, 29, 166–192.
<https://doi.org/10.1111/j.1096-0031.2012.00421.x>
- van Horn, G., Branson, S., Farrell, R., Haber, S., Barry, J., Ipeiros, P., Perona, P. & Belongie, S. (2015) Building a bird recognition app and large scale dataset with citizen scientists: the fine print in fine-grained dataset collection. Computer Vision and Pattern Recognition (CVPR). Available from: <https://vision.cornell.edu/se3/projects/visipedia/> (accessed 15 January 2017)
- van Nieukerken, E.J. (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang, Z.-Q. (Ed.), *Animal biodiversity: an outline of higher classification and survey of taxonomic richness*. *Zootaxa*, 3148, 212–221.
- Vila, R., Bell, C.D., Macniven, R., Goldman-Huertas, B., Ree, R.H., Marshall, C.R., Balint, Z., Johnson, K., Benyamini, D. & Pierce, N.E. (2011) Phylogeny and palaeoecology of *Polyommatus* blue butterflies show Beringia was a climate-regulated gateway to the New World. *Proceedings of the Royal Society B*, 278, 2737–2744.
<https://doi.org/10.1098/rspb.2010.2213>
- Wagner, D.L. (2001) Moths. In: Levin, S.A. (Ed.), *Encyclopedia of Biodiversity*. Academic Press, San Diego, pp. 249–270.
<https://doi.org/10.1016/b0-12-226865-2/00201-7>