

Editorial: herbarium collection-based plant evolutionary genetics and genomics

[Health & Medicine](#)



**ASSIGN
BUSTER**

Editorial on the Research Topic

Herbarium Collection-Based Plant Evolutionary Genetics and Genomics

Out of the Box: Herbarium Collections as Research Tools

Over the past centuries and especially the last decades, herbarium collections world-wide have amassed an estimated 350 million specimens of plants and fungi, deposited in 3, 400 herbaria world-wide ([Soltis, 2017](#)). Many of these specimens comprise yet-undescribed species, estimated to be around 70, 000 ([Bebber et al., 2010](#)), and therefore represent a potentially highly significant contribution to taxonomy and systematics.

These collections are a huge repository of botanical metadata, not only at the level of the specimen itself—with its associated collection locality, population, and associated pathogen data ([Yoshida et al., 2015](#) ; [James et al., 2018](#))—but also at the level of characters and traits such as leaf morphology ([Queenborough, 2017](#)), gene ([Bieker and Martin, 2018](#)), and genome (GGBN; [Seberg et al., 2016](#)). Analyzing this metadata permits us to look back—and perhaps even forward—in time, testing historical biological hypotheses, identifying extinct genotypes, as well as modeling past ecological processes or extrapolating future trends. Especially for plants with long generation times, it is often not feasible to observe genetic changes over many generations in greenhouse experiments. As time series for many plant species can be amassed from global herbaria, these collections facilitate studying plants in more detail, for example to estimate mutation rates after introduction to a novel range ([Exposito-Alonso et al., 2018](#)). In fact, collectively, herbarium collections can be seen as being part of a *Global*

Museum in which new questions can be addressed using novel combinations of disciplines that so far may not typically interact ([Bakker et al., 2020](#)), and in which digitisation of herbarium collections plays an important role ([Soltis, 2017](#)).

Herbarium DNA

Over the past three decades, a considerable body of literature has accumulated concerning herbarium DNA, mostly focusing on its properties and post-mortem damage (e. g., [Savolainen et al., 1995](#) ; [Staats et al., 2011](#) , [2013](#) ; [Weiss et al., 2016](#)) and its potential utility for biological inference (e. g., [Erkens et al., 2008](#) ; [Drábková et al., 2002](#) ; [Telle and Thines, 2008](#) ; [Särkinen et al., 2012](#) ; [Bakker et al., 2016](#) ; [Gutaker et al., 2017](#) ; [Brewer et al., 2019](#)). The overarching conclusion has been that plant archival DNA is remarkably well-preserved, in spite of the heat often applied to the specimens during collection and preservation. In contrast with animal tissues, the cell wall in plant (and fungal) material probably provides good protection against DNA damage caused by oxidative stress ([Mateiu and Rannala, 2008](#) ; [Roldán-Arjona and Ariza, 2009](#)).

Plant nuclear genomes are generally much larger than animal and fungal genomes ([Gregory et al., 2007](#)) and are characterized by abundant repeats, which can hamper the assembly of genome sequences. But provided an existing reference genome sequence exists, whole-genome studies including herbarium specimens are possible. For example, [Exposito-Alonso et al., 2018](#) re-sequenced the nuclear genomes of 36 *Arabidopsis thaliana* herbarium specimens collected between 1863 and 1993. As the number of available

reference genomes rapidly increases, in part through initiatives such as the Earth BioGenome Project ([Lewin et al., 2018](#)), this approach becomes available even for non-model plant species.

Plastomes, with their structural conservation across land plants ([Wicke and Schneeweiss, 2015](#)), modest length around 160 kbp, and high copy number in the cell, have proven to be feasible targets for herbarium DNA studies, especially through the application of “ genome skimming ” ([Straub et al., 2012](#) ; [Bakker, 2017](#)). Remarkable examples of plastome sequencing from herbarium DNA include now-extinct species, as in, for instance, the *de-novo* assembly of the complete ([Zedane et al., 2015](#)) and mitogenome ([Van de Paer et al., 2018](#)) from a 140-year-old specimen of *Hesperelaea* (Oleacea), the reference-guided assembly of the plastome from a 167-year old specimen of *Leptagrostis schimperiana* (Arundinoideae, Poaceae) to resolve its taxonomic position ([Hardion et al., 2020](#)), as well as the reconstruction of the complete plastome from the now-extinct, endemic Hawaiian mint *Stenogyne haliakalae* ([Welch et al., 2016](#)). Due to their influence on crop production and their development of herbicide resistance, weeds are also of general interest. As an example, here Sablok et al. contribute the reconstructed plastome sequence of *Ambrosia trifida* (Asteraceae) from a herbarium specimen collected in 1886, and investigate the plant's resistance to the widely used herbicide glyphosate.

Ancient Alleles

Herbarium specimens contain not only DNA from the specimen itself, but also from associated microbes and pathogens that can be exploited for

evolutionary and ecological inference ([Bieker et al., 2020](#)). Indeed, examples in which herbarium DNA contributed to evolutionary genetic inference have focused on historical pathogens ([Martin et al., 2013](#) ; [Yoshida et al., 2014](#) , [2015](#)), determining the genotype of the oomycete plant pathogen *Phytophthora infestans* that caused the nineteenth-century Irish potato famine. Ristaino summarizes emerging patterns in global *Phytophthora* distribution and how mycological and plant herbaria have played an important role in reconstructing pathways of plant pathogen movement. Ancient alleles in *Alopecurus myosuroides* Huds., relevant to herbicide resistance but pre-dating human influence, were detected from herbarium DNA by [Délye et al. \(2013\)](#) . Likewise, [Besnard et al. \(2014\)](#) , using DNA from a 100-year-old Madagascan herbarium specimen, reconstructed the shift in underlying genetics from C₃ to C₄ in grass photosynthesis. These examples demonstrate that we are currently at the dawn of an era of historical herbarium genomics ([Buerki and Baker, 2015](#) ; [Bieker and Martin, 2018](#)), and it is likely that a large body of plant archival genomic data will be generated in the years to come.

That clade- or phylum-specific challenges remain in sequencing herbarium DNA is illustrated by Forin et al. , who describes the case of the Saccardo mycological herbarium and how ribosomal DNA sequencing of 100-year old fungal specimens was feasible. Forin et al. point out the potential of mycological herbarium genomics, as only an estimated 1% of total known fungal species currently have associated DNA sequences in public databases.

DNA Barcoding and Metabarcoding

DNA barcoding is a commonly used method for species identification and phylogenetic analysis. It relies on the amplification of short, conserved genomic loci that show enough variation to separate species but have low intraspecific variation. Due to the high levels of DNA fragmentation ([Weiss et al., 2016](#)), it is challenging to obtain barcode sequencing for herbarium specimens ([Prosser et al., 2016](#)). Kistenich et al. explore DNA sequencing of a ca. 900-bp portion of the long mitochondrial ribosomal small subunit (mtSSU) from historical lichen specimens using a two-step PCR approach followed by Ion Torrent sequencing. Their approach demonstrated a high success rate compared to traditional Sanger sequencing, providing enough sequencing information for species identification even in a 150-year old specimen. Moreover, no significant correlation between sequencing success and habitat ecology of the investigated specimens was found, which provides confidence for future lichen herbarium sequencing projects. These and other methodological advancements suggest that barcode information can also be more easily retrieved from herbarium and fungarium type specimens, which are often more than 100 years old.

Barcoding studies, and especially those employing metabarcoding, rely heavily on the completeness of databases for species identification. In order to complete the databases, which currently only contain an estimated 20% of described land plants ([Wilkinson et al., 2017](#)), herbarium collections offer a valuable and largely untapped resource. Focusing on Australian plant biodiversity, Dormontt et al. make a strong case for how herbarium collections should be systematically analyzed to capitalize on their scientific

potential, and that the curation of specimen reference data is paramount in this.

Genomic Sequencing and Other Data From Herbarium DNA

The increasingly common retrieval of genome-wide SNP data from herbarium DNA presents an enormous potential for future phylogenetic, population genomic, and molecular-based ecological studies that include these valuable specimens. For instance, genotyping by sequencing (GBS) of *Solidago* species from relatively young herbarium specimens collected between 1970 and 2010 was successful in 98% of samples ([Beck and Semple, 2015](#)). More recently, [Lang et al. \(2020\)](#) described a very promising reduced-representation sequencing method designed for hybridization-capture of ddRAD loci from historical plant specimens, demonstrating its utility on *Arabidopsis thaliana* and *Cardamine bulbifera* . At a lower-throughput genomic scale, targeted enrichment approaches aim to obtain sequence data for hundreds of nuclear-encoded loci from herbarium samples ([Hart et al., 2016](#) ; [Brewer et al., 2019](#) ; [Viruel et al., 2019](#)). Forrest et al. test the limits of the Hyb-Seq approach for sequencing herbarium specimens, especially with regards to historic and contemporary techniques of specimen preservation. They show that although data could also be obtained from poor quality samples, preservation methods like heat and alcohol treatment yielded greater DNA degradation and poorer DNA retrieval. Finally, Dodsworth et al. explore the angiosperm non-coding genome, showing that herbarium DNA is virtually indistinguishable from fresh DNA in analyses of nuclear genomic repeat clusters and their abundancies. Thus, they argue

that herbarium collections can facilitate further genomic exploration of the repetitive content of plant genomes, which yields additional (nuclear) phylogenetic markers useful at the (sub)species-level (see also [Dodsworth et al., 2015](#)).

Scaling-Up the Use of Collections

This topic has collected a variety of outstanding recent successes in genetic analysis of herbarium specimens, inspiring our prediction of the great potential of these approaches for evolutionary genomic, population genetic, phylogenetic, and biosystematic discovery in archival plants and their associated microbial communities. For instance, future studies of the evolution of photosynthesis, or the evolutionary ecology of adaptive and functional traits, will undoubtedly benefit from the extensive samples offered by herbarium collections. We posit that we are at the dawn of an era of herbarium genomics in which, augmenting traditional phylogenetic studies that have become the standard for herbarium collection-based genetic work, soon will come a flood of ecological and evolutionary investigations utilizing large quantities of genomic and genetic data gleaned directly from voucher specimens.

Author Contributions

MM conceived of the Research Topic. FB drafted the editorial with significant contributions from VB and MM. All authors contributed to the article and approved the submitted version.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We thank the many authors who contributed their manuscripts on this research topic. Funding was provided by Norwegian Research Council Young Research Talents grant 287327 to MM.

References

- Bakker, F. T. (2017). Herbarium genomics: skimming and plastomics from archival specimens. *Webbia J. Plant Taxon. Geogr.* 72, 35–45. doi: 10.1080/00837792.2017.1313383
- Bakker, F. T., Antonelli, A., Clarke, J. A., Cook, J. A., Edwards, S. V., Ericson, P. G. P., et al. (2020). The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ*. 8: e8225. doi: 10.7717/peerj.8225
- Bakker, F. T., Lei, D., Yu, J., Mohammadin, S., Wei, Z., Van de Kerke, S., et al. (2016). Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an iterative organelle genome assembly (IOGA) pipeline. *Biol. J. Linnean Soc.* 117: 33–43. doi: 10.1111/bij.12642
- Bebber, D. P., Carine, M. A., Wood, J. R. I., Wortley, A. H., Harris, D. J., Prance, G. T., et al. (2010). Herbaria are a major frontier for species discovery. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22169–22171. doi: 10.1073/pnas.1011841108
- <https://assignbuster.com/editorial-herbarium-collection-based-plant-evolutionary-genetics-and-genomics/>

Beck, J. B., and Semple, J. C. (2015). Next-generation sampling: pairing genomics with herbarium specimens provides species-level signal in *Solidago* (Asteraceae). *Appl. Plant Sci.* 3: 1500014. doi: 10. 3732/apps. 1500014

Besnard, G., Christin, P.-A., Malé, P.-J. G., L'Huillier, E., Lauzeral, C., Coissac, E., et al. (2014). From museums to genomics: old herbarium specimens shed light on a C3 to C4 transition. *J. Exp. Bot.* 65, 6711–621. doi: 10. 1093/jxb/eru395

Bieker, V. C., and Martin, M. D. (2018). Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Bot. Lett.* 165, 409–418. doi: 10. 1080/23818107. 2018. 1458651

Bieker, V. C., Sánchez Barreiro, F., Rasmussen, J. A., Brunier, M., Wales, N., and Martin, M. D. (2020). Metagenomic analysis of historical herbarium specimens reveals a postmortem microbial community. *Mol. Ecol. Resour.* 20, 1206–1219. doi: 10. 1111/1755-0998. 13174

Brewer, G. E., Clarkson, J. J., Maurin, O., Zuntini, A. R., Barber, V., Bellot, S., et al. (2019). Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of Angiosperms. *Front. Plant Sci.* 10: 1102. doi: 10. 3389/fpls. 2019. 01102

Buerki, S., and Baker, W. J. (2015). Collections-based research in the genomic era. *Biol. J. Linn. Soc.* 117, 5–10. doi: 10. 1111/bij. 12721

Délye, C., Deulvot, C., and Chauvel, B. (2013). DNA analysis of herbarium specimens of the grass weed *Alopecurus myosuroides* reveals herbicide resistance pre-dated herbicides. *PLoS ONE* 8: e75117. doi: 10.1371/journal.pone.0075117

Dodsworth, S., Chase, M. W., Kelly, L. J., Leitch, I. J., Macas, J., Novák, P., et al. (2015). Genomic repeat abundances contain phylogenetic signal. *Syst. Biol.* 64, 112–126. doi: 10.1093/sysbio/syu080

Drábková, L., Kirschner, J., and Vlcek, C. (2002). Comparison of seven DNA extraction and amplification protocols in historic herbarium specimens of Juncaceae. *Plant Mol. Biol. Rep.* 20, 161–175. doi: 10.1007/BF02799431

Erkens, R. H. J., Cross, H., Maas, J. W., Hoenselaar, K., and Chatrou, L. W. (2008). Age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. *Blumea* 53, 407–428. doi: 10.3767/000651908X608052

Exposito-Alonso, M., Becker, C., Schuenemann, V. J., Reiter, E., Setzer, C., Slovak, R., et al. (2018). The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* 14: e1007155. doi: 10.1371/journal.pgen.1007155

Gregory, T. R., Nicol, J. A., Tamm, H., Kullman, B., Kullman, K., Leitch, I. J., et al. (2007). Eukaryotic genome size databases. *Nucleic Acids Research* , 35, D332–D338. doi: 10.1093/nar/gkl828

Gutaker, R. M., Reiter, E., Furtwängler, A., Schuenemann, V. J., and Burbano, H. A. (2017). Extraction of ultrashort DNA molecules from herbarium specimens. *Biotechniques* 62, 76–79. doi: 10. 2144/000114517

Hardion, L., Verlaque, R., Kaymak, E., Vila, B., Haan-Archipoff, G., Martin, M. M., et al. (2020). Plastome sequencing of a 167-year-old herbarium specimen and classical morphology resolve the systematics of two potentially extinct grass species. *Bot. J. Linn. Soc.* boaa065. doi: 10. 1093/botlinnean/boaa065

Hart, M. L., Forrest, L. L., Nicholls, J. A., and Kidner, C. A. (2016). Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65, 1081–1092. doi: 10. 12705/655. 9

James, S. A., Soltis, P. S., Belbin, L., Chapman, A. D., Nelson, G., Paul, D. L., et al. (2018). Herbarium data: global biodiversity and societal botanical needs for novel research: global. *Appl. Plant Sci.* 6: e1024. doi: 10. 1002/aps3. 1024

Lang, P. L. M., Weiß, C. L., Kersten, S., Latorre, S. M., Nagel, S., Nickel, B., et al. (2020). Hybridization ddRAD-sequencing for population genomics of nonmodel plants using highly degraded historical specimen DNA. *Mol. Ecol. Resour.* doi: 10. 1111/1755-0998. 13168

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., et al. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4325–4333. doi: 10. 1073/pnas. 1720115115

Martin, M. D., Cappellini, E., Samaniego, J. A., Zepeda, M. L., Campos, P. F., Seguin-Orlando, A., et al. (2013). Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nat. Commun.* 4, 1–7. doi: 10.1038/ncomms3172

Mateiu, L. M., and Rannala, B. H. (2008). Bayesian inference of errors in ancient DNA caused by postmortem degradation. *Mol. Biol. Evol.* 25, 1503–1511. doi: 10.1093/molbev/msn095

Prosser, S. W. J., deWaard, J. R., Miller, S. E., and Hebert, P. D. N. (2016). DNA barcodes from century-old type specimens using next-generation sequencing. *Mol. Ecol. Resour.* 16, 487–497. doi: 10.1111/1755-0998.12474

Queenborough, S. (2017). Collections-based studies of plant functional traits. in Pp 223-236 in I. Friis and H. Balslev (eds), *Tropical Plant Collections: Legacies from the Past? Essential Tools for the Future?* *Sci. Danica B* 6, 15–38. Available online at: <http://84.19.174.124/books/692/4979?lang=en>

Roldán-Arjona, T., and Ariza, R. R. (2009). Repair and tolerance of oxidative DNA damage in plants. *Mut. Res.* 681, 169–179. doi: 10.1016/j.mrrev.2008.07.003

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., and Bakker, F. T. (2012). How to open the treasure chest? optimising DNA extraction from herbarium specimens. *PLoS ONE* 7: e43808. doi: 10.1371/journal.pone.0043808

- Savolainen, V., Cuénoud, P., Spichiger, R., Martinez, M. D. P., Crèvecoeur, M., and Manen, J.-F. (1995). The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Syst. Evol.* 197, 87–98.
- Seberg, O., Droege, G., Barker, K., Coddington, J. A., Funk, A., Gostel, M., et al. (2016). Global genome biodiversity network: saving a blueprint of the Tree of Life – A botanical perspective. *Ann. Bot.* 118, 393–399. doi: 10.1093/aob/mcw121
- Soltis, P. S. (2017). Digitization of herbaria enables novel research. *Am J Bot.* 104: 1–4. doi: 10.3732/ajb.1700281
- Staats, M., Cuence, A., Richardson, J. E., Vrielink-van Ginkel, R., Petersen, G., Seberg, O., et al. (2011). DNA damage in plant herbarium tissue. *PLoS ONE* 6: e28448. doi: 10.1371/journal.pone.0028448
- Staats, M., Erkens, R. H. J., van de Vossenberg, B., Wieringa, J. J., Kraaijeveld, K., Stielow, B., et al. (2013). Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLoS ONE* 8: e69189. doi: 10.1371/journal.pone.0069189
- Straub, S. C. K., Parks, M., Weitemeir, K., Fishbein, M., Cronn, R., Liston, A., et al. (2012). Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364. doi: 10.3732/ajb.1100335
- Telle, S., and Thines, M. (2008). Amplification of *cox2* (~620 bp) from 2 mg of up to 129 years old herbarium specimens, comparing 19 extraction

methods and 15 polymerases. *PLoS ONE* 3: e3584. doi: 10. 1371/journal.pone. 0003584

Van de Paer, C., Bouchez, O., and Besnard, G. (2018). Prospects on the evolutionary mitogenomics of plants: a case study on the olive family (Oleaceae). *Mol. Ecol. Resour.* 18, 407–423. doi: 10. 1111/1755-0998. 12742

Viruel, J., Conejero, M., Hidalgo, O., Pokorny, L., Powell, R. F., Forest, F., et al. (2019). A Target capture-based method to estimate ploidy from herbarium specimens. *Front. Plant Sci.* 10: 937. doi: 10. 3389/fpls. 2019. 00937

Weiss, C. L., Schuenemann, V. J., Devos, J., Shirsekar, G., Reiter, E., Gould, B. A., et al. (2016). Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *R. Soc. Open Sci.* 3: 160239. doi: 10. 1098/rsos. 160239

Welch, A. J., Collins, K., Ratan, A., Drautz-Moses, D. I., Schuster, S. C., and Lindqvist, C. (2016). The quest to resolve recent radiations: plastid phylogenomics of extinct and endangered Hawaiian endemic mints (Lamiaceae). *Mol. Phylogenet. Evol.* 99, 16–33. doi: 10. 1016/j. ympev. 2016. 02. 024

Wicke, S., and Schneeweiss, G. M. (2015). “ Next-generation organellar genomics: potentials and pitfalls of high-throughput technologies for molecular evolutionary studies and plant systematics,” in *Next Generation Sequencing in Plant Systematics* , eds E. Hörandl and M. S. Appelhans [Hessen: International Association for Plant Taxonomy (IAPT)], 9–50.

Wilkinson, M. J., Szabo, C., Ford, C. S., Yarom, Y., Croxford, A. E., Camp, A., et al. (2017). Replacing sanger with next generation sequencing to improve coverage and quality of reference DNA barcodes for plants. *Sci. Rep.* 7: 46040. doi: 10. 1038/srep46040

Yoshida, K., Burbano, H. A., Krause, J., Thines, M., Weigel, D., and Kamoun, S. (2014). Mining herbaria for plant pathogen genomes: back to the future. *PLoS Pathog.* 10: e1004028. doi: 10. 1371/journal. ppat. 1004028

Yoshida, K., Sasaki, E., and Kamoun, S. (2015). Computational analyses of ancient pathogen DNA from herbarium samples: challenges and prospects. *Front. Plant Sci.* 6: 771. doi: 10. 3389/fpls. 2015. 00771

Zedane, L., Hong-Wa, C., Murienne, J., Jeziorsky, C., Baldwin, B. G., and Besnard, G. (2015). Museomics illuminate the history of an extinct, paleoendemic plant lineage (*Hesperelaea*, *Oleaceae*) known from an 1875 collection from Guadalupe Island, Mexico. *Biol. J. Linnean Soc.* 117, 44–57. doi: 10. 1111/bij. 12509