Time and Chance Happen to Them All: A Macroevolutionary Examination of the Effects of Whole Genome Multiplications

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Biology

by

Eric Hagen
Wesleyan University
Bachelor of Arts in Biology and Earth & Environmental Sciences, 2018

August 2023
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Jeremy Beaulieu, Ph.D.
Dissertation Director

Andrew Alverson, Ph.D.
Committee Member

Adam Siepielski, Ph.D.
Committee Member

Andrew Leslie, Ph.D.
Committee Member (Ex-Officio)

ABSTRACT

Polyploidy, the state of having more than two complete sets of chromosomes, is common in plants and has been linked to several beneficial traits. On the macroevolutionary scale, the effects of polyploidy have been hotly debated for over one hundred years, being alternatively described as an "evolutionary dead end" and the most important discovery in evolutionary biology since Darwin and Wallace. This thesis aims to contribute to the debate by studying the diversification, biogeography, and ecophysiology of polyploid flowering plants with recently developed phylogenetic comparative methods. This dissertation has three chapters. In Chapter I, I review work on the so-called "dead-end hypothesis" in polyploid research, which I argue is in fact multiple hypotheses masquerading as one. I supplement the review with an analysis of tip diversification rates in Solanaceae, employing the MiSSE model. In Chapter II, I examine the "latitudinal polyploidy gradient," in which polyploid plants comprise greater proportions of the flora at higher latitudes. To compare latitudinal movement and patterns of origination between diploids and polyploids across four flowering plant clades, I use the novel machuruku model to reconstruct ancestral ranges and develop a new function for ancestral state reconstruction within the corHMM package. In Chapter III, I perform the first systematic review and meta-analysis comparing pathogen resistance in diploid and polyploid plants, incorporating phylogenetic information at the family level. Across these three analyses, I do not find support for associations between polyploidy and increased diversification or the evolution of beneficial traits. It is thus possible that any beneficial effects resulting from polyploidy can be chalked up to the "luck of the draw." Together, these chapters all present novel or under-utilized methods of studying the effects of polyploidy in phylogenetic context.

ACKNOWLEDGMENTS

In March 2021, when I had not been able to work in the lab for a year, thought that the death of comparative methods might be upon us, and was languishing in bed from a mysterious skin disease, I was sure that this dissertation would never be completed. You are reading this only because I received immense personal and professional support from so many people.

I owe much gratitude to my advisor, Jeremy Beaulieu, who has inspired me with his humor, work ethic, and encyclopedic knowledge of evolution, music, sports, and many other things since he accepted me into his lab five years ago. Jeremy is the type of advisor who puts great trust in his students to conduct independent research, and because of this he has many times had to listen to me wax prolix about dubious ideas that have popped into my head. But the latitude he gave me as a student led me to study polyploidy, an excellent medium through which to pursue my interests in classic evolutionary theory. If I ever have the pleasure of being considered a great programmer, presenter, or advisor, I owe a great debt to Jeremy.

I also need to thank my friend, colleague, roommate, and landlord James Boyko. James has listened and responded to my ideas more than anyone else, especially when we were stuck in the house together for more than a year. Whether my ideas were potentially good or downright stupid, James has always given me honest feedback. He has been a teacher, a mentor, and, at times, a much-appreciated thorn in my side. I also owe much of the knowledge that shaped this dissertation to conversations I had with my good friends Zach Zbinden, Simon Tye, and Peter Hasik, both in the lab and over beers at Smoke & Barrel.

I received invaluable instruction from the two Beaulieu lab postdocs, Daniel Caetano and Thais Vasconcelos. I benefited greatly from their thoughts and critiques, as they were like second advisors to me. We have starkly different views on many issues in biology, and I thank them for

challenging me. Thanks also to Rosana Zenil-Ferguson, who provided invaluable help on Chapter I of this dissertation, and Chase Mason, who collaborated with me on Chapter III. They enriched this work greatly with careful edits despite their immensely busy schedules.

I would not have pursued this PhD without the education and encouragement I received from my previous mentors. Nicole Justice, my high school biology teacher, first inspired my interest in and eventual love for evolution. My research mentors at Wesleyan University, particularly Michael Singer, Helen Poulos, and Dana Royer, greatly prepared me for my doctoral work. Dana advised my undergraduate thesis on the taphonomy of leaf fossils, and he probably no longer has an open-door policy because of my constant questions.

I thank my committee members, Andy Alverson, Adam Siepielski, and Andrew Leslie, who have commented on and inspired various parts of this thesis. My development as a coder was greatly accelerated by Andy's programming class, and Chapter III of this dissertation would not exist if I had not begun it as a project in Adam's meta-analysis class. All three have challenged me and inspired me to be a more serious scientist.

Finally, thank you to my parents. They have supported me tirelessly in every pursuit I've ever undertaken, especially academia. This dissertation is lovingly dedicated to them.

The title of this thesis, which comes from Ecclesiastes 9:11, is supposed to reference the fact that the "evolutionary success" often observed in polyploid plant lineages is likely due to chance events and luck over time: "The race is not to the swift, nor the battle to the strong . . . but time and chance happeneth to them all." Despite my hard work on this thesis, I also must thank time and chance for my lucky trajectory as a scientist and for having brought all the people mentioned here into my life.

DEDICATION

Dedicated to Mom and Dad, who used their knowledge of science to help people lead healthy lives. I stand on your shoulders.

TABLE OF CONTENTS

LIST OF PUBLISHED PAPERS

Chapter I:
Hagen, E.R., and J.M. Beaulieu. 2023. New beginnings for dead ends: polyploidy, SSE models, and the dead-end hypothesis. *Annals of Botany*. In prep.

Chapter II:
Hagen, E.R., T. Vasconcelos, J.D. Boyko, and J.M. Beaulieu. 2023. Historical causes for the greater proportion of polyploid plants in higher latitudes. *American Journal of Botany.* In prep.

Chapter III:
Hagen, E.R., and C.M. Mason. 2023. Differences in pathogen resistance between diploid and polyploid plants: a systematic review and meta-analysis. *Oikos.* In revision.

INTRODUCTION

[W]e now know that a new species can arise in one step by the formation of a sterile or nearly sterile hybrid followed by a doubling of chromosomes which renders it fertile but still vigorous. Such hybrids, which are called allopolyploids . . . probably [are] . . . the most important amendment to Darwin and Wallace's account of evolution as a historical fact.

- *J.B.S. Haldane (1959a)*

One of the most influential forces that has shaped the evolutionary trajectory of plants is polyploidization, in which the number of complete sets of chromosomes increases. Diploid plants generate polyploid offspring primarily through the production of unreduced gametes, which are formed due to chromosome division errors in meiosis I or II (Rothfels and Otto 2016). Unreduced gametes can either combine with other unreduced gametes to create tetraploid offspring, or they combine with diploid gametes to produce (sometimes) fertile triploids (Ramsey and Schemske 1998). Since crosses between plants of different ploidy levels often produce inviable seeds, usually due to endosperm abnormalities (Köhler et al. 2021), polyploidization usually results in instant speciation, whether the polyploid is formed by gametes from the same species (known as "autopolyploidy") or from different species (known as "allopolyploidy").

In flowering plants, about 15% of all speciation events are believed to be accompanied by polyploidization (Wood et al. 2009), and all flowering plants are theorized to descend from a common ancestor that underwent polyploidization and was followed by a massive radiation (Jiao et al. 2011; but see Ruprecht et al. 2017). Within the clade, many other radiations appear to have followed polyploidization events, which may have led to advantageous traits (Soltis et

al. 2009; Van de Peer et al. 2017). Unlike familiar mutations, in which relatively small amounts of the genetic code are changed, polyploidization can be thought of as a "macromutation" (Doyle and Coate 2020) that multiples the total number of genes in an organism possesses by two or more, which can create noticeable phenotypic changes in just a single generation.

The immediate phenotypic changes often created by polyploidization can be so rapid and striking that, when Hugo de Vries observed (unbeknownst to him) polyploid offspring of evening primrose, he used it as the basis for his "mutation theory" of evolution, which threatened to replace Darwinian gradualism in the early twentieth century (Larson 2004). After the chromosomal nature of polyploidy had been identified (Lutz 1907), it became an object of intense biological research. Early botanists identified several phenotypic changes often seen after polyploidization, including the "gigas effect," in which morphological traits and cells increase in size (Gates 1909), weakening or breakdown of self-incompatibility (Darlington 1928; Crane and Lewis 1942), and slower rates of development (Müntzing 1936). Attempting to harness possible beneficial effects of polyploidy, researchers found ways to induce polyploidy through breeding (Karpechenko 1928) and with chemicals like colchicine (Blakeslee and Avery 1937), a practice still in use by agriculturists to this day (Touchell et al. 2020). The major architects of the Modern Synthesis even extended their population genetic models to account for populations exhibiting polyploidy and polysomic inheritance (Haldane 1930; Wright 1938; Fisher 1943). J.B.S. Haldane (1959b) considered polyploidy to be a remarkable evolutionary exception because polyploidization is a major force guiding the evolution of plants that is not itself caused by natural selection.

However, one architect of the Modern Synthesis, botanist Ledyard Stebbins, set the stage for a massive reversal in opinion on polyploidy. Stebbins put forward various arguments

for polyploids being ineffectual, or even an "evolutionary dead end," including that polyploidy does not lead to morphologically distinct higher clades (Stebbins 1940; 1971), that polyploidy inhibits responses to selection through redundancy (Stebbins 1971), and that only allopolyploids show advantages due to hybrid vigor (Stebbins 1980). Current polyploidy research has refuted many of Stebbins's ideas. For example, whereas he believed that approximately 30 to 35 percent of flowering plant species formed via polyploidy (Stebbins 1971), recent estimates for the number of species with polyploidy in their histories range as high as 70% (Masterson 1994), and autopolyploidy is now believed to occur much more frequently than Stebbins believed (Soltis et al. 2007; Barker et al. 2016a). Regardless, his ideas greatly shaped the narrative around polyploidy in the late 20th century (Ramsey and Ramsey 2014; Soltis et al. 2014), arguably spearheading the dominant perception that polyploidy was "evolutionary noise" (Wagner 1970) and even a "black hole of research" (Barker et al. 2016b).

In the twenty-first century, interest in polyploidy research has resurged, largely due to the advent of modern genomic and phylogenetic tools. Whereas early botanists studied polyploidy by examining chromosomes, we are now able to sequence entire plant genomes and have determined that many "diploid" plant species have histories of polyploidy that were previously unknown (Arabidopsis Genome Initiative 2000). Whereas earlier phylogenies were largely made using molecular characters and few species, we are now able to build phylogenies from molecular data (Woese and Fox 1977; Felsenstein 1981). And since the advent of phylogenetic independent contrasts (Felsenstein 1985), there has been a veritable explosion of phylogenetic comparative methods (PCMs) that continues to this day. Modern discussions about polyploidy are largely based on results obtained using modern trait based PCMs, particularly the advent of state-dependent speciation and extinction (SSE) models (Maddison et

al. 2007) and transition models capable of reconstructing chromosome changes (e.g., chromEVOL; Glick and Mayrose 2014). PCMs are now central to debates surrounding polyploidy, as scientists have shifted focus from vaguely defined concepts like "morphological novelty" to specific evolutionary rates calculated based on molecular phylogenies and trait data.

This thesis aims to apply modern PCMs to problems related to polyploidy in several biological disciplines. The main theoretical developments I apply in this work are hidden Markov models (HMMs), which are used when modeling sequence evolution (Felsenstein and Churchill 1996), character state transitions (Beaulieu et al. 2013), and diversification of lineages (Beaulieu and O'Meara 2016). Phylogenetic HMMs allow users to incorporate the possibility that forces beyond the data included in the model drive the evolutionary process under study. For example, if you suspect that polyploidy affects diversification, you can compare a model in which polyploids and diploids possess different net diversification rates to one in which species vary not only in terms of ploidy but also in a hypothetical "hidden" character unknown to the user. This removes the risk of false positive results, especially when null models containing hidden characters are included in model selection (Caetano et al. 2018). Chapters I through III feature recently developed HMMs, including two semi-novel models developed in this dissertation. With these powerful tools, this thesis aims to cover two major topics in contemporary polyploidy research.

The first theme I address is the effect that polyploidy has on diversification over millions of years. Using various clades, character datasets, and SSE models, workers have found evidence that polyploids exhibit decreased diversification (Mayrose et al. 2011), increased diversification (Han et al. 2020), and similar rates of diversification relative to diploids (Estep et al. 2014; Landis et al. 2018; Román-Palacios et al. 2020). Variations in

findings could be due to differences in sampling, model structure, or various features unique to specific clades. Recently, the development of HiSSE (Beaulieu and O'Meara 2016), which incorporates hidden characters into SSE models, has allowed for tests of whether polyploidy itself drives diversification rates more so than hidden states. Some have found that it does not, and that linked traits like shifts to self-compatibility may instead better explain observed diversification variation (Zenil-Ferguson et al. 2019). There is also the possibility that polyploidization does not immediately lead to shifts in diversification. One difficulty is that current SSE methods will not pick up these possible diversification "lags" between the transition to a trait (in this case polyploidy) and diversification rate shifts. This could be the case if polyploid success depends on subsequent events like environmental shifts or genome reorganization (Schranz et al. 2012; Robertson et al. 2017). It is presently unclear whether lags that have been observed are real or artifactual. Available methods identify diversification shifts with trait-free tools like MEDUSA (Alfaro et al. 2009) and then associate shifts with polyploidization events post-hoc (Tank et al. 2015; Landis et al. 2018; Smith et al. 2018), as testing for lags in a single model framework has not yet been pursued.

The second topic this thesis investigates is whether polyploidy consistently confers beneficial traits over shorter time periods. The current consensus is that the effects of polyploidy are inconsistent and heavily dependent on the ecological (Segraves 2017), genomic (Otto 2007), and phylogenetic (Burleigh 2012) contexts of any organism that undergoes genome multiplication. However, some regular patterns do appear. For instance, polyploid plants comprise larger proportions of the flora at higher latitudes relative to more equatorial latitudes, possibly due to greater cold tolerance relative to diploids (Rice et al. 2019). While numerous traits have been studied in terms of their relationship to polyploidization, studies that

test for the relationship in a unified model framework are rare: frequently, as in lag studies, ploidy changes are inferred with programs like chromEVOL separately and then related to ploidy changes post-hoc (e.g., Linder and Barker 2014; Baniaga et al. 2020). Many traits, such as pathogen resistance, have not even been examined in phylogenetic context. Despite the indispensable information that phylogeny provides, trees are still insufficiently used in contemporary studies of polyploidy's effects on trait evolution.

In this thesis, I aim to fill in some gaps in phylogenetic research on polyploidy. In Chapter I, I review the diversification literature to interrogate whether polyploidy is an "evolutionary dead end," and I conduct a study on Solanaceae with the recently developed model MiSSE (Vasconcelos et al. 2022). In Chapter II, I apply the newly introduced model *machuruku* (Guillory and Brown 2021), which conducts ecological niche modeling at paleoclimatic time slices in phylogenetic context, to attempt to uncover the mechanisms behind the latitudinal gradient in polyploid frequency. In Chapter III, I conduct the first-ever meta-analysis on whether polyploid plants exhibit superior pathogen resistance relative to diploids. Finally, in the conclusion, I discuss some limitations of this thesis as well as possible directions for future polyploidy research.

References

Alfaro, M.E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D.L. Rabosky, G. Carnevale, and L.J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* 106: 13410–13414.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.

Baniaga, A.E., H.E. Marx, N. Arrigo, and M.S. Barker. 2020. Polyploid plants have faster rates of multivariate niche differentiation than their diploid relatives. *Ecology Letters* 23: 68–78.

Barker, M.S., N. Arrigo, A.E. Baniaga, Z. Li, and D.A. Levin. 2016a. On the relative abundance of autopolyploids and allopolyploids. *New Phytologist* 210: 391–398.

Barker, M.S., B.C. Husband, and J.C. Pires. 2016b. Spreading Winge and flying high: the evolutionary importance of polyploidy after a century of study. *American Journal of Botany* 103: 1139–1145.

Beaulieu, J.M., B.C. O'Meara, and M.J. Donoghue. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62(5): 725– 737.

Beaulieu, J.M., and B.C. O'Meara. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* 65: 583–601.

Blakeslee, A.F., and A.G. Avery. 1937. Methods of inducing doubling of chromosomes in plants: by treatment with colchicine. *Journal of Heredity* 28: 393–411.

Burleigh, J.G. 2012. Identifying the phylogenetic context of whole-genome duplications in plants. *In* Soltis, P.S., and D.E. Soltis (eds.) Polyploidy and genome evolution. Springer.

Caetano, D.S., B.C. O'Meara, and J.M. Beaulieu. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* 72: 2308–2324.

Crane, M.B., and D. Lewis. 1942. Genetical studies in pears III. Incompatibility and sterility. *Journal of Genetics* 43: 31–43.

Darlington, C.D. 1928. Studies in *Prunus*, I and II. *Journal of Genetics* 19: 213–264.

Doyle, J.J., and J.E. Coate. 2020. Autopolyploidy: an epigenetic macromutation. *American Journal of Botany* 107: 1097–1100.

Estep, M.C., M.R. McKain, D.V. Diaz, J. Zhong, J.G. Hodge, T.R. Hodkinson, D.J. Layton, S.T. Malcomber, R. Pasquet, and E.A. Kellogg. 2014. Allopolyploidy, diversification, and the Miocene grassland expansion. *Proceedings of the National Academy of Sciences* 111: 15149–15154.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.

Felsenstein, J., and G.A. Churchill. 1996. A Hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution* 13: 93–104.

Fisher, R.A. 1943. Allowance for double reduction in the calculation of genotype frequencies with polysomic inheritance. *Annals of Eugenics* 12: 169–171.

Gates, R.R. 1909. The stature and chromosomes of *Oenothera gigas*, DeVries. *Zeitschrift für Induktive Abstammungs-und Vererbungslehre* 3: 220.

Glick, L., and I. Mayrose. 2014. ChromEvol: assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Molecular Biology and Evolution* 31: 1914–1922.

Guillory, W.X., and J.L. Brown. 2021. A new method for integrating ecological niche modeling with phylogenetics to estimate ancestral distributions. *Systematic Biology* 70: 1033–1045.

Haldane, J.B.S. 1930 Theoretical genetics of autopolyploids. *Journal of Genetics* 22: 359–372.

Haldane, J.B.S. 1959a. The theory of natural selection to-day. *Nature* 183: 710–713.

Haldane, J.B.S. 1959b. Natural selection. *In* Bell, P.R. (ed.) Darwin's biological work: some aspects reconsidered. Cambridge University Press.

Han, T.-S., Q.-J. Zheng, R.E. Onstein, B.M. Rojas-Andrés, F. Hauenschild, A.N. Muellner-Riehl, and Y.-W. Xing. 2020. Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist* 225: 571–583.

Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph, L.P. Tomsho, Y. Hu, H. Liang, P.S. Soltis, and D.E. Soltis. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Karpechenko, G.D. 1928. Polyploid hybrids of *Raphanus sativus* L. × *Brassica oleracea* L. *Molecular and General Genetics* 48: 1–85.

Köhler, C., K. Dziasek, and G. Del Toro-De León. 2021. Postzygotic reproductive isolation established in the endosperm: mechanisms, drivers and relevance. *Philosophical*

Transactions of the Royal Society B: Biological Sciences 376: 20200118.

Landis, J.B., D.E. Soltis, Z. Li, H.E. Marx, M.S. Barker, D.C. Tank, and P.S. Soltis. 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* 105: 348–363.

Larson, E.J. 2004. Evolution: the remarkable history of a scientific theory. Random House Digital, Inc.

Linder, H.P., and N.P. Barker. 2014. Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113: 1175–1183.

Lutz, A.M. 1907. A preliminary note on the chromosomes of *Œnothera lamarckiana* and one of its mutants, *O. gigas. Science* 26: 151–152.

Maddison, W.P., P.E. Midford, and S.P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56: 701–710.

Masterson, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264: 421–424.

Mayrose, I., S.H. Zhan, C.J. Rothfels, K. Magnuson-Ford, M.S. Barker, L.H. Rieseberg, and S.P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.

Müntzing, A. 1936. The evolutionary significance of autopolyploidy. *Hereditas* 21: 263–378.

Otto, S.P. 2007. The evolutionary consequences of polyploidy. *Cell* 131: 452–462.

Ramsey, J., and D.W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.

Ramsey, J., and T.S. Ramsey. 2014. Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130352.

Rice, A., P. Šmarda, M. Novosolov, M. Drori, L. Glick, N. Sabath, S. Meiri, J. Belmaker, and I. Mayrose. 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* 3: 265–273.

Robertson, F.M., M.K. Gundappa, F. Grammes, T.R. Hvidsten, A.K. Redmond, S. Lien, S.A. Martin, P.W. Holland, S.R. Sandve, and D.J. Macqueen. 2017. Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biology* 18: 1–14.

Román-Palacios, C., Y.F. Molina-Henao, and M.S. Barker. 2020. Polyploids increase overall diversity despite higher turnover than diploids in the Brassicaceae. *Proceedings of the Royal Society B: Biological Sciences* 287: 20200962.

Rothfels, C.J., and S.P. Otto. 2016. Polyploid speciation. *In* Kliman, R.M. (ed.) Encyclopedia of evolutionary biology. Academic Press.

Ruprecht, C., R. Lohaus, K. Vanneste, M. Mutwil, Z. Nikoloski, Y. Van de Peer, and S. Persson. 2017. Revisiting ancestral polyploidy in plants. *Science Advances* 3: e1603195.

Schranz, M.E., S. Mohammadin, and P.P. Edger. 2012. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion In Plant Biology* 15: 147–153.

Segraves, K.A. 2017. The effects of genome duplications in a community context. *New Phytologist* 215: 57–69.

Smith, S.A., J.W. Brown, Y. Yang, R. Bruenn, C.P. Drummond, S.F. Brockington, J.F. Walker, N. Last, N.A. Douglas, and M.J. Moore. 2018. Disparity, diversity, and duplications in the Caryophyllales. *New Phytologist* 217: 836–854.

Soltis, D.E., P.S. Soltis, D.W. Schemske, J.F. Hancock, J.N. Thompson, B.C. Husband, and W.S. Judd. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56: 13–30.

Soltis, D.E., V.A. Albert, J. Leebens-Mack, C.D. Bell, A.H. Paterson, C. Zheng, D. Sankoff, C.W. de Pamphilis, P.K. Wall, and P.S. Soltis. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96: 336–348.

Soltis, D.E., M.C. Segovia-Salcedo, I. Jordon-Thaden, L. Majure, N.M. Miles, E.V. Mavrodiev, W. Mei, M.B. Cortez, P.S. Soltis, and M.A. Gitzendanner. 2014. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytologist* 202: 1105–1117.

Stebbins, G.L. 1940. The significance of polyploidy in plant evolution. *American Naturalist* 74: 54–66.

Stebbins, G.L. 1971. Chromosomal evolution in higher plants. Edward Arnold Ltd.

Stebbins, G.L. 1980. Polyploidy in plants: unsolved problems and prospects. *In* Lewis, W.H. (ed.) Polyploidy: biological relevance. Plenum Press.

Tank, D.C., J.M. Eastman, M.W. Pennell, P.S. Soltis, D.E. Soltis, C.E. Hinchliff, J.W. Brown, E.B. Sessa, and L.J. Harmon. 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207: 454–467.

Touchell, D.H., I.E. Palmer, and T.G. Ranney. 2020. In vitro ploidy manipulation for crop improvement. *Frontiers in Plant Science* 11: 722.

Van de Peer, Y., E. Mizrachi, and K. Marchal. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18: 411–424.

Vasconcelos, T., B.C. O'Meara, and J.M. Beaulieu. 2022. A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution* 76: 1420–1433.

Wagner, W.H. 1970. Biosystematics and evolutionary noise. *Taxon* 19: 146–151.

Woese, C.R., and G.E. Fox. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences* 74: 5088–5090.

Wood, T.E., N. Takebayashi, M.S. Barker, I. Mayrose, P.B. Greenspoon, and L.H. Rieseberg. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* 106: 13875–13879.

Wright, S. 1938. The distribution of gene frequencies in populations of polyploids. *Proceedings of the National Academy of Sciences* 24: 372–377.

Zenil-Ferguson, R., J.G. Burleigh, W.A. Freyman, B. Igić, I. Mayrose, and E.E. Goldberg. 2019. Interaction among ploidy, breeding system and lineage diversification. *New Phytologist* 224: 1252–1265.

CHAPTER I

New beginnings for dead ends: polyploidy, SSE models, and the dead-end hypothesis

Eric R. Hagen and Jeremy M. Beaulieu

Abstract

Since the mid-twentieth century, it has been argued that the transition from diploidy to polyploidy is an "evolutionary dead end" in plants. While this point has been debated ever since, multiple definitions of "dead end" have been used in the polyploidy literature without sufficient differentiation between alternative uses. In this review, we focus on the two most common conceptions of the dead-end hypothesis currently discussed: the "lowering diversification" hypothesis and the "rarely successful" hypothesis. We discuss the evidence for both hypotheses, and we employ a recently developed method of inferring tip diversification rates to examine the effect of ploidy on diversification in Solanaceae. We find that diversification rates in the family are not strongly correlated with ploidy or with the closely related trait of breeding system. We also outline recent work in the field that moves beyond the relatively simple question of whether polyploidy increases, decreases, or does not significantly affect diversification rates in plants.

Introduction

In a sense, we can assume that any species that originates will inevitably end in a blind alley: over 99% of all species that have ever lived have gone extinct (Jablonski 2004), and all extant species have a non-zero probability of going extinct (Marshall 2017). Even well-adapted species can fall victim to changes in their environment (Vrba 1993) or being out-competed (Van Valen 1973). Specific traits in such situations, which were once adaptive but became deleterious due to environmental changes, are known as "maladaptations" (Crespi 2000). However, can the evolution of an adaptive trait itself lead to the demise of a species? Questions in this vein are rarely asked by biologists, but they have been studied. Examples include "evolutionary suicide," in which traits that are adaptive at the level of the organism lead to the extinction of a population (Rankin and López-Sepulcre 2005); "macroevolutionary self-destruction," referring to traits that evolve frequently but are rapidly lost through extinction of their possessors (Bromham et al. 2016); and "evolutionary dead ends," in which traits that are adaptive at the level of the population are deleterious at the level of species through the lowering of diversification rates. Support for the dead-end phenomenon has been found for a variety of traits, including narrow host range in Tachinid flies (Day et al. 2016), fossoriality in snakes (Cyriac and Kodandaramaiah 2018), shifts to self-compatibility in Solanaceae (Goldberg et al. 2010), and shifts to polyploidy in plants (Mayrose et al. 2011).

The idea that polyploidy is a dead-end has a long history in botany. The person most responsible for the idea that polyploidy is an evolutionary dead end is G. Ledyard Stebbins, who strongly shaped the history of polyploidy research with his extensive writings (see Soltis et al. 2014a). However, Stebbins's conception of polyploidy as a dead end was very different from the definition given above. Instead of referring to evolutionary rates, he offered several different

arguments, including that selection cannot work efficiently on polyploids due to the masking of gene copies, that polyploidy is not causally related to the diversification of plants, and that polyploidization produces no subsequent increases in morphological disparity or the evolution of any traits that may be called key innovations (Stebbins 1950; 1971). In fact, Stebbins never even used the phrase "evolutionary dead end" to refer to polyploidy (Soltis et al. 2014a); he did, however, apply the similar term "blind alley" to self-fertilization, for which he offered another explanation: Stebbins (1957) defined a dead end as the result of an "unlucky accident" in which a species acquired advantageous mutations but lost its ability to adequately respond to environmental pressures, in this case through outcrossing, and has thus increased its likelihood of extinction. The connection between Stebbins's ideas about polyploidy and modern tests of the dead-end hypothesis that invoke evolutionary rates (e.g., Mayrose et al. 2011) thus appears tenuous.

In contemporary polyploidy research, the dead-end hypothesis as it relates to polyploid rates is controversial, due to both arguments about new methods of inferring diversification rates as well as the veracity of the dead-end hypothesis itself (Soltis et al. 2014b; Mayrose et al. 2015). Additionally, there are now two dead-end hypotheses often lumped together into one: a "traditional" dead-end hypothesis, in which polyploidy generally lowers diversification, and the "rarely successful" hypothesis, in which *most* polyploids are dead ends, but some occasionally establish successfully and go on to diversify, sometimes greatly (Arrigo and Barker 2012). The purpose of this review is to discuss these alternative formulations of dead-end hypotheses in polyploidy research and the evidence currently supporting each. To spur further research on polyploid diversification rates, we also perform the first study comparing tip diversification rates in polyploids and diploids, using Solanaceae as a case study. Finally, we outline emerging

directions in the field that go beyond the dead-end hypothesis, asking more complex questions about the relationship between ploidy and diversification. Ultimately, we argue that researchers should no longer be asking whether polyploidy is a dead end, as its effects on diversification, if any, depend too heavily on chance and ecological factors.

The "Traditional" Dead-End Hypothesis

Since the advent of modern phylogenetic comparative methods, arguments for and against dead-end status almost invariably involve how polyploidy affects specific evolutionary rates like speciation, extinction, net diversification, and transition rates among the ploidy states. Arguably the most common definition of an evolutionary dead end is a trait that confers a short-term selective advantage to a species but in the long run leads to decreased net diversification (e.g., Mayrose et al. 2011; 2015). The first test of this hypothesis with PCMs was conducted with sister clade comparisons in Rosaceae (Vamosi and Dickinson 2006), finding that polyploidy is associated with increased species richness in the clade. However, sister clade comparisons can be misleading, as they are unable to differentiate whether a character is common because it causes diversification or because transitions are biased toward it, and because they do not use all the information available in a phylogeny (Maddison 2006). Hence, researchers in the field now tend to favor using state-dependent speciation and extinction (-SSE) models, which allow for estimates of diversification and transition rates as well as compare them between taxa possessing different characters. In examining the "traditional" dead-end hypothesis, -SSE models can both infer how ploidy states affect diversification rates and infer rates of transition between ploidy states, as irreversibility of a state is often cited as a second criterion for a character state to be a dead end (Ng and Smith 2014).

Mayrose et al. (2011) used BiSSE (binary state speciation and extinction) models to show that polyploidy generally lowers diversification relative to diploidy. While Soltis et al. (2014b) argued against their findings on the grounds that their results were confounded by clade age and size, Mayrose et al. (2015) successfully defended their study by pointing out that BiSSE rates are scaled by time, thus avoiding these concerns. However, the authors did acknowledge that their analysis was limited to recently developed polyploids, or "neopolyploids," in young clades, thus sidestepping the question of whether polyploidy contributed to diversification in deep time (e.g., Fawcett et al. 2009). Studies have since examined this question with newly developed SSE models: for example, Landis et al. (2018) used MuSSE, which extends BiSSE to allow for multiple characters with multiple states (FitzJohn et al. 2012), to conclude that multiple rounds of polyploidization generally increases diversification rates across angiosperms.

With the development of HiSSE (Beaulieu and O'Meara 2016), which allows for models in which diversification is controlled not only by observed traits but also "hidden" traits representing hypothetical, unobserved factors controlling evolutionary rates, researchers were able to test whether ploidy controls diversification more than unconsidered influences. This was beneficial for hypothesis testing: if a model including a hidden trait is favored, and one finds that the hidden trait controls diversification rates more strongly than ploidy, one may then look to other traits often linked to ploidy, such as selfing (Barringer 2007; but see Mable 2004) or herbaceousness (Zenil-Ferguson et al. 2017). HiSSE was used by Zenil-Ferguson et al. (2019) to find that selfing does indeed explain diversification better than ploidy, and by Han et al. (2020) to find that lineages with greater proportions of polyploids exhibit higher diversification rates.

Linked Traits and the "Rarely Successful" Hypothesis

While analyses of multiple traits using models like MuSSE and HiSSE are certainly useful in that they have enabled joint analysis of ploidy and other traits, the effects of polyploidy *qua* polyploidy are often insufficiently distinguished from other evolutionary developments that frequently accompany polyploidization in formulations of the dead-end hypothesis. For example, since polyploidy is believed to occur most frequently in perennial herbs (Stebbins 1971; Zenil-Ferguson et al. 2017), and because perenniality may increase extinction rates (Soltis et al. 2013), polyploidy may thus indirectly lead to decreased diversification. Unless ploidy shifts are causally linked to concomitant shifts in other traits, or it can be shown that polyploidy causes decreased diversification in cases where such shifts do not occur, the label of "dead end" seems inapplicable in this case.

Another example of such a trait is breeding system. Shifts to self-compatibility are frequently believed to accompany ploidy shifts, and they have been cited as a reason for polyploidy being an evolutionary dead end (Stebbins 1950). What is odd is that selfing has been used as evidence both for polyploids being dead ends and for polyploids exhibiting increased diversification in deep time. While shifts to self-compatibility have been demonstrated to be linked to decreased diversification rates (Goldberg et al. 2010), selfing has also been cited as crucial polyploids to overcome minority cytotype exclusion (Levin 1975) as well as to survive mass extinctions and repopulate empty niches (Fawcett et al. 2009; Lohaus and Van de Peer 2016; Freeling 2017). The potential for such widely differing responses to environmental pressure at the species level is, in the case of polyploidy, the basis of the second dead-end hypothesis, the "rarely successful" hypothesis.

In the “rarely successful” hypothesis, most avenues lead down a blind alley, but some occasionally lead to diversification, sometimes at very high rates. Sessa (2019) uses the metaphor of the “Las Vegas strategy,” where plants that undergo polyploidization effectively “gamble” to possibly gain evolutionary advantages. While this bet most often results in extinction, plants occasionally “win big,” and their bet pays off with increased diversification. Aforementioned SSE models can hint at whether this is the case: Roman-Palacios et al. (2020) found that although polyploids exhibit similar net diversification (speciation – extinction) rates in Brassicaceae to diploids, they also exhibit higher turnover (speciation + extinction), which is a measure that suggests a higher frequency of both speciation and extinction event over evolutionary time (Vasconcelos et al. 2022a). Yet the “rarely successful” hypothesis requires the deep time perspective noted to be missing in studies like that of Mayrose et al. (2011): whereas lowered diversification rates can be detected on relatively small and shallow trees, finding “rare success” in the descendants of certain polyploids requires studying very large trees with deep roots. Candidates for “rarely successful” polyploids include, possibly, the ancestors of all flowering plants and all seed plants (Jiao et al. 2011; but see Ruprecht et al. 2017). Additionally, whereas the “traditional” dead-end hypothesis requires only present-day ploidy data, testing the “rarely successful” hypothesis necessitates data about histories of ploidy hidden in the genomes of species that have since downsized and re-organized their genomes through the process of diploidization (see Dodsworth et al. 2016). Several rounds of so-called “paleopolyploidy” have been uncovered even in plants with very small genomes (Bowers et al. 2003). Very interestingly, these are often clustered near times of major environmental stress (Vanneste et al. 2014; Novikova et al. 2018), and many bursts of diversification seem to occur after long “lags” sometimes lasting millions of years (Tank et al. 2015; Landis et al. 2018).

Testing Polyploid Diversification with Tip Rate Correlation

One limitation in testing the "rarely successful" hypothesis is that much information on extinction is lost in deep time and unavailable on a phylogeny (O'Meara and Beaulieu 2021). Despite its shallower time scale, this problem affects analyses of the "traditional" dead-end hypothesis as well. In the absence of the ability to quantify extinction risk with microevolutionary analyses, one method that could be useful for predicting which taxa are traveling down evolutionary "blind alleys" would be comparing the present-day "snapshots" of evolutionary rates of diploids and polyploids.

These methods quantify what are known as species-specific diversification rates, or tip rates (Title and Rabosky 2019), which have recently gained interest in studies of polyploidy (e.g., Testo and Sundue 2018; Román-Palacios et al. 2020) but remain an under-utilized area with great potential (Soltis et al. 2019). Tip rates are calculated using the evolutionary history of each lineage, and it has been argued that, like a general diversification rate can be inverted for a waiting time, the reciprocal of a tip rate can be interpreted as a hypothesis for the time it will take for some speciation or extinction or other event to occur (Title and Rabosky 2019; but see Vasconcelos et al. 2022b). Many simple model-free tip rate statistics are very computationally inexpensive yet still remarkably informative and reliable for determining traits that drive diversification through tip-rate correlation (Freckleton et al. 2008; Jetz et al. 2012; Harvey and Rabosky 2018), including the *phylometrics* package (Bromham et al. 2016), which was explicitly designed to test dead-end hypotheses. In model-based comparative methods, increasing numbers of diversification models now include functions that allow the calculation of tip rates based on marginal reconstruction, including BAMM (Rabosky 2014) and MiSSE (Vasconcelos et al. 2022b).

MiSSE, which is implemented in the *hisse* R package, is structurally very similar to HiSSE except that no observed states are incorporated into models. Unlike other trait-free diversification models like MEDUSA (Alfaro et al. 2009), MiSSE allows users to design custom models, composed solely of hidden states, that can vary the number of free parameters for turnover and extinction fraction. Tip rates calculated from model-averaged reconstructions can then be flexibly statistically analyzed post-hoc, such as in regressions after they have been corrected with phylogenetic independent contrasts (Felsenstein 1985; see Vasconcelos et al. 2022b). MiSSE is still a new tool, and to our knowledge it has not yet been employed to compare diversification between diploids and polyploids. Therefore, we re-analyzed the Solanaceae phylogeny (Särkinen et al. 2013) with MiSSE to explore its potential as well as compare its results with those obtained with HiSSE in a previous study on Solanaceae (Zenil-Ferguson et al. 2019).

For our analysis, models were determined using the model set-up function *generateMiSSEGreedyCombinations*, allowing for up to 10 free turnover parameters and 3 free extinction fraction parameters within a single model. For all 30 models, we ran MiSSE using the *MiSSEGreedy* function, which employs a "greedy" hill-climbing optimization routine, and then pruned redundant models, which left us with 29 in total. We then performed model-averaged marginal reconstructions and calculated tip rates. To examine the relationships between tip rates and observed ploidy data, we performed phylogenetic ANOVA (Garland et al. 1993; Revell 2012) in addition to phylogenetic logistic regression (Ives and Garland 2010), as well as their non-phylogenetic statistical counterparts. We used these tests instead of MiSSE's TipCorrelation function, which employs phylogenetic independent contrasts (Felsenstein 1985) because our analysis required correlating discrete characters with tip rates instead of continuous ones. Prior to

all analyses, we removed “cherries” from the phylogeny. “Cherries” are sister tips that share the same branch length to their direct ancestor node, and MiSSE gives the option to remove them in its own TipCorrelation function because they inherit the same rate class probabilities and thus constitute pseudoreplicates (Vasconcelos et al. 2022b).

Ploidy states and tip turnover rates across the clade are depicted in Fig. 1. Visual inspection of turnover rates calculated at diploid and polyploid tips indicates that diploids consistently exhibit higher average estimates for all five tip rates (Fig. 2). However, while non-phylogenetic ANOVA and logistic regression indicated significant differences between the two groups (ANOVA: F=5.822, p=0.017; logit: p=0.018), incorporating phylogenetic information results in no significant differences in evolutionary rates (phylogenetic ANOVA: F=5.822, p=0.481; phylogenetic logit: p=0.957). In other words, ploidy states appear to be clumped on the Solanaceae phylogeny, suggesting that it is less labile than observed tip turnover rates. We also detected no significant correlation between tip rates and either breeding system (phylogenetic ANOVA: F=1.831, p=0.549; phylogenetic logit: p=0.291) or combinations of breeding system and ploidy states (phylogenetic ANOVA: F=6.886, p=0.458; see Fig. 3).

It is possible that, since this study examines neopolyploids instead of the influence of ploidy in deeper evolutionary history, not enough time has passed for polyploidy to significantly affect turnover rates. It is also possible that traits beyond ploidy and breeding system more strongly control tip diversification rates in Solanaceae. While our findings regarding ploidy accord with those of Zenil-Ferguson et al. (2019), which conducted a HiSSE analysis with the same data, our analysis differs in that we find no significant correlation of either breeding system or combinations of ploidy and breeding system states with tip diversification rates in Solanaceae.

Emerging Directions in Polyploid Diversification Research

In looking at the progression of polyploid diversification research alongside the development of SSE models, it appears that evidence for a significant association between ploidy and evolutionary rates wanes as models become more sophisticated. This is likely a result of both biology and modeling: the effects of ploidy are highly dependent on chance, evolutionary history, and ecological context (Segraves 2017; Meudt et al. 2021), and the addition of hidden states in SSE models removes false positive results that would otherwise be found in models like BiSSE (Beaulieu and O'Meara 2016; Caetano et al. 2018). In a recent systematic review of SSE model studies performed on angiosperms, Helmstetter et al. (2023) show that, in the three studies that have used HiSSE or its multi-state counterpart MuHiSSE to study ploidy, two analyses found that net diversification was higher for the polyploid state while two other analyses found that it was lower. This is a very small number of studies, but depending on the clade being studied, and whether ploidy state is classified based on neopolyploidy or paleopolyploidy, finding support for the dead-end hypothesis may be analogous to flipping of a coin.

What if the question is not as simple as comparing diversification rates in diploids and polyploids? It is possible that instead of comparing how polyploidy affects diversification across clades, we also should compare how its effects on diversification change over time. Huang et al. (2020) underscores the fact that diversification patterns are not cleanly related to individual polyploidization events but instead are remnants of polyploidization-diploidization cycles (Baduel et al. 2018). Due to this, they argue that whole genome multiplications are not consistently related to diversification over time but instead should be viewed as a kind of "pump" for species diversity. Another area of research that could lead to a decline in the invocation of the dead-end hypothesis is study of mixed-ploidy systems, which has gained

interest recently (Kolář et al. 2017). Most research up to this point has compared diploid vs. polyploid diversification in species that are either one cytotype or another, and they can thus be categorized with binary traits. Mixed ploidy species, in which taxa contain both diploid and polyploid individuals, have been shown to diversify faster than single polyploids or diploids, likely because polymorphic species are able to speciate rapidly into both mixed and polyploid new species (Wei et al. 2018). The rate of speciation in mixed-ploidy systems is also related to the proportion of polyploids and diploids, being higher when the proportion of polyploids is higher (Han et al. 2020). Additionally, mixed-ploidy species occupy larger geographic/latitudinal ranges than single polyploids or diploids (Lobato-de Magalhães et al. 2021). These findings suggest a role for conspecific diploids working together with polyploids in speciating and radiating geographically as opposed to merely competing.

With these directions in mind, and with a greater focus on specific hypotheses in trait-dependent diversification modeling, it seems quite possible that polyploid research will soon leave discussions of dead ends behind. The effects of polyploidy are too intertwined with other traits, too ecologically dependent, and too dependent on factors like time and clade biology to be so easily categorized. While the dead-end concept may have reached its own dead end, many roads extend ahead for future polyploid diversification research.

References

Alfaro, M.E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D.L. Rabosky, G. Carnevale, and L.J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13410–13414.

Arrigo, N., and M.S. Barker. 2012. Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15: 140–146.

Baduel, P., S. Bray, M. Vallejo-Marin, F. Kolář, and L. Yant. 2018. The "Polyploid Hop": shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers in Ecology and Evolution* 6: 117.

Barringer, B.C. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany* 94: 1527–1533.

Beaulieu, J.M., and B.C. O'Meara. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* 65: 583–601.

Bowers, J.E., B.A. Chapman, J. Rong, and A.H. Paterson. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422: 433–438.

Bromham, L., X. Hua, and M. Cardillo. 2016. Detecting macroevolutionary self-destruction from phylogenies. *Systematic Biology* 65: 109–127.

Caetano, D.S., B.C. O'Meara, and J.M. Beaulieu. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* 72: 2308–2324.

Crespi, B.J. 2000. The evolution of maladaptation. *Heredity* 84: 623–629.

Cyriac, V.P., and U. Kodandaramaiah. 2018. Digging their own macroevolutionary grave: fossoriality as an evolutionary dead end in snakes. *Journal of Evolutionary Biology* 31: 587–598.

Day, E.H., X. Hua, and L. Bromham. 2016. Is specialization an evolutionary dead end? Testing for differences in speciation, extinction and trait transition rates across diverse phylogenies of specialists and generalists. *Journal of Evolutionary Biology* 29: 1257–1267.

Dodsworth, S., M.W. Chase, and A.R. Leitch. 2016. Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Botanical Journal of the Linnean Society* 180: 1–5.

Fawcett, J.A., S. Maere, and Y. Van De Peer. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America* 106: 5737–5742.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.

FitzJohn, R.G. 2012. *Diversitree*: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* 3: 1084–1092.

Freckleton, R.P., A.B. Phillimore, and M. Pagel. 2008. Relating traits to diversification: a simple test. *American Naturalist* 172: 102–115.

Freeling, M. 2017. Picking up the ball at the K/Pg boundary: the distribution of ancient polyploidies in the plant phylogenetic tree as a spandrel of asexuality with occasional sex. *The Plant Cell* 29: 202–206.

Garland, T., A.W. Dickerman, C.M. Janis, and J.A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. *Systematic Biology* 42: 265–292.

Goldberg, E.E., J.R. Kohn, R. Lande, K.A. Robertson, S.A. Smith, and B. Igić. 2010. Species selection maintains self-incompatibility. *Science* 330: 493–495.

Han, T.S., Q.J. Zheng, R.E. Onstein, B.M. Rojas-Andrés, F. Hauenschild, A.N. Muellner-Riehl, and Y.W. Xing. 2020. Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist* 225: 571–583.

Harvey, M.G., and D.L. Rabosky. 2018. Continuous traits and speciation rates: alternatives to state-dependent diversification models. *Methods in Ecology and Evolution* 9: 984–993.

Helmstetter, A.J., R. Zenil-Ferguson, H. Sauquet, S.P. Otto, M. Méndez, M. Vallejo-Marin, J. Schönenberger, C. Burgarella, B. Anderson, H. de Boer, S. Glémin, and J. Käfer. 2023. Trait-dependent diversification in angiosperms: patterns, models and data. *Ecology Letters* 26: 640–657.

Huang, X.C., D.A. German, and M.A. Koch. 2020. Temporal patterns of diversification in Brassicaceae demonstrate decoupling of rate shifts and mesopolyploidization events. *Annals of Botany* 125: 29–47.

Ives, A.R., and T. Garland. 2010. Phylogenetic logistic regression for binary dependent variables. *Systematic Biology* 59: 9–26.

Jablonski, D. 2004. Extinction: past and present. *Nature* 427: 589.

Jetz, W., G.H. Thomas, J.B. Joy, K. Hartmann, and A.O. Mooers. 2012. The global diversity of birds in space and time. *Nature* 491: 444–448.

Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph, L.P. Tomsho, Y. Hu, H. Liang, P.S. Soltis, and D.E. Soltis. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Kolář, F., M. Čertner, J. Suda, P. Schönswetter, and B.C. Husband. 2017. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in Plant Science* 22: 1041–1055.

Levin, D.A. 1975. Minority cytotype exclusion in local plant populations. *Taxon* 24: 35–43.

Lobato-de Magalhães, T., K. Murphy, A. Efremov, V. Chepinoga, T.A. Davidson, and E. Molina-Navarro. 2021. Ploidy state of aquatic macrophytes: global distribution and drivers. *Aquatic Botany* 173: 103417.

Lohaus, R., and Y. Van de Peer. 2016. Of dups and dinos: evolution at the K/Pg boundary. *Current Opinion in Plant Biology* 30: 62–69.

Mable, B.K. 2004. Polyploidy and self-compatibility: is there an association? *New Phytologist* 162: 803–811.

Maddison, W.P. 2006. Confounding asymmetries in evolutionary diversification and character change. *Evolution* 60: 1743–1746.

Maddison, W.P., P.E. Midford, and S.P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56: 701–710.

Marshall, C.R. 2017. Five palaeobiological laws needed to understand the evolution of the living biota. *Nature Ecology & Evolution* 1: 0165.

Mayrose, I., S.H. Zhan, C.J. Rothfels, N. Arrigo, M.S. Barker, L.H. Rieseberg, and S.P. Otto. 2015. Methods for studying polyploid diversification and the dead end hypothesis: a reply to Soltis et al. (2014). *New Phytologist* 206: 27–35.

Mayrose, I., S.H. Zhan, C.J. Rothfels, K. Magnuson-Ford, M.S. Barker, L.H. Rieseberg, and S.P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.

Meudt, H.M., D.C. Albach, A.J. Tanentzap, J. Igea, S.C. Newmarch, A.J. Brandt, W.G. Lee, and J.A. Tate. 2021. Polyploidy on islands: its emergence and importance for diversification. *Frontiers in Plant Science* 12: 637214.

Ng, J., and S.D. Smith. 2014. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. *Journal of Evolution Biology* 27: 2035–2045.

Novikova, P.Y., N. Hohmann, and Y. Van de Peer. 2018. Polyploid *Arabidopsis* species originated around recent glaciation maxima. *Current Opinion in Plant Biology* 42: 8–15.

O'Meara, B.C., and J.M. Beaulieu. 2021. Potential survival of some, but not all, diversification methods. *EcoEvoRxiv* https://doi.org/10.32942/osf.io/w5nvd

Rabosky, D.L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9: e89543.

Rankin, D.J., and A. López-Sepulcre. 2005. Can adaptation lead to extinction? *Oikos* 111: 616–619.

Revell, L.J. 2012. *phytools*: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 2: 217–223.

Román-Palacios, C., Y.F. Molina-Henao, and M.S. Barker. 2020. Polyploids increase overall diversity despite higher turnover than diploids in the Brassicaceae. *Proceedings of the Royal Society B* 287: 20200962.

Ruprecht, C., R. Lohaus, K. Vanneste, M. Mutwil, Z. Nikoloski, Y. Van de Peer, and S. Persson. 2017. Revisiting ancestral polyploidy in plants. *Science Advances* 3: e1603195.

Särkinen, T., L. Bohs, R.G. Olmstead, and S. Knapp. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology* 13: 1–5.

Segraves, K.A. 2017. The effects of genome duplications in a community context. *New Phytologist* 215: 57–69.

Sessa, E.B. 2019. Polyploidy as a mechanism for surviving global change. *New Phytologist* 221: 5–6.

Soltis, P.S., R.A. Folk, and D.E. Soltis. 2019. Darwin review: angiosperm phylogeny and evolutionary radiations. *Proceedings of the Royal Society B* 286: 20190099.

Soltis, D.E., M.E. Mort, M. Latvis, E.V. Mavrodiev, B.C. O'Meara, P.S. Soltis, J.G. Burleigh, and R.R. de Casas. 2013. Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany* 100: 916–929.

Soltis, D.E., C.J. Visger, and P.S. Soltis. 2014a. The polyploidy revolution then… and now: Stebbins revisited. *American Journal of Botany* 101: 1057–1078.

Soltis, D.E., M.C. Segovia-Salcedo, I. Jordon-Thaden, L. Majure, N.M. Miles, E.V. Mavrodiev, W. Mei, M.B. Cortez, P.S. Soltis, and M.A. Gitzendanner. 2014b. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytologist* 202: 1105–1117.

Stebbins, G.L. 1950. *Variation and evolution in plants.* New York: Columbia University Press.

Stebbins, G.L. 1957. Self fertilization and population variability in the higher plants. *American Naturalist* 91: 337–354.

Stebbins, G.L. 1971. *Chromosomal evolution in higher plants.* London: Edward Arnold (Publishers) Ltd.

Tank, D.C., J.M. Eastman, M.W. Pennell, P.S. Soltis, D.E. Soltis, C.E. Hinchliff, J.W. Brown, E.B. Sessa, and L.J. Harmon. 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207: 454–467.

Testo, W.L., and M.A. Sundue. 2018. Are rates of species diversification and body size evolution coupled in the ferns? *American Journal of Botany* 105: 525–535.

Title, P.O., and D.L. Rabosky. 2019. Tip rates, phylogenies and diversification: what are we estimating, and how good are the estimates? *Methods in Ecology and Evolution* 10: 821–834.

Vamosi, J.C., and T.A. Dickinson. 2006. Polyploidy and diversification: a phylogenetic investigation in Rosaceae. *International Journal of Plant Sciences* 167: 349–358.

Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* 1: 1–30.

Vanneste, K., G. Baele, S. Maere, and Y. Van de Peer. 2014. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. *Genome Research* 24: 1334–1347.

Vasconcelos, T., B.C. O'Meara, and J.M. Beaulieu. 2022a. Retiring "cradles" and "museums" of biodiversity. *American Naturalist* 199: 194–205.

Vasconcelos, T., B.C. O'Meara, and J.M. Beaulieu. 2022b. A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution* 76: 1420–1433.

Vrba, E.S. 1993. Turnover-pulses, the Red Queen, and related topics. *American Journal of Science* 293: 418–452.

Wei, R., R.H. Ree, M.A. Sundue, and X.C. Zhang. 2018. Polyploidy and elevation contribute to opposing latitudinal gradients in diversification and species richness in lady ferns (Athyriaceae). *bioRxiv* 351080.

Zenil-Ferguson, R., J.G. Burleigh, W.A. Freyman, B. Igić, I. Mayrose, and E.E. Goldberg. 2019. Interaction among ploidy, breeding system and lineage diversification. *New Phytologist* 224: 1252–1265.

Zenil-Ferguson, R., J.M. Ponciano, and J.G. Burleigh. 2017. Testing the association of phenotypes with polyploidy: an example using herbaceous and woody eudicots. *Evolution* 71: 1138–1148.

Appendix

Figure 1 – Solanaceae phylogeny with tip states and turnover tip rates. Tip states for ploidy are shown in orange (diploids) and green (polyploids). Tip states for breeding system are shown in red (self-incompatible) and blue (self-compatible). Turnover rates for each species are shown on the right, with color varying from black to light yellow as rates increase.

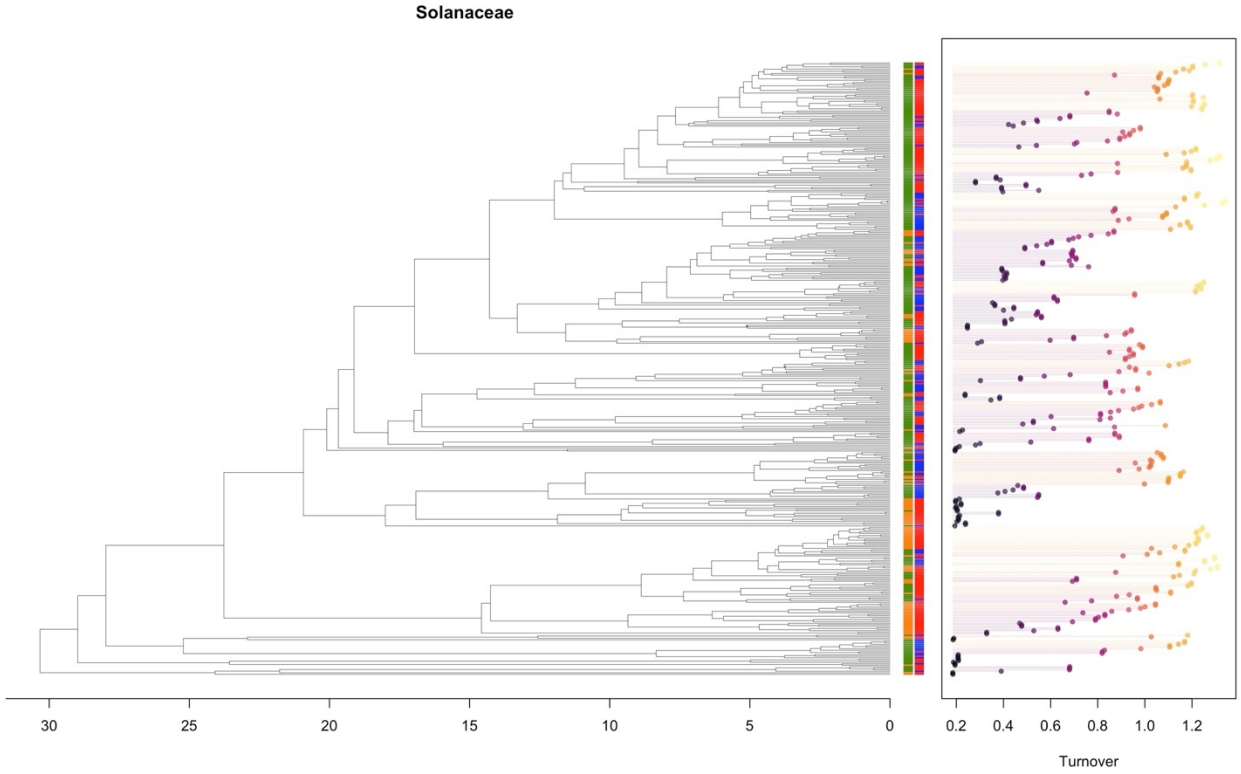


Figure 2 – Boxplots comparing tip rates across ploidy states. Diploids ("0") are shown in orange while polyploids ("1") are shown in green. While visually they are significant differences among the different metrics of diversification, incorporating phylogenetic information in the summary statistics shows these differences are not significant and may have arisen by chance.

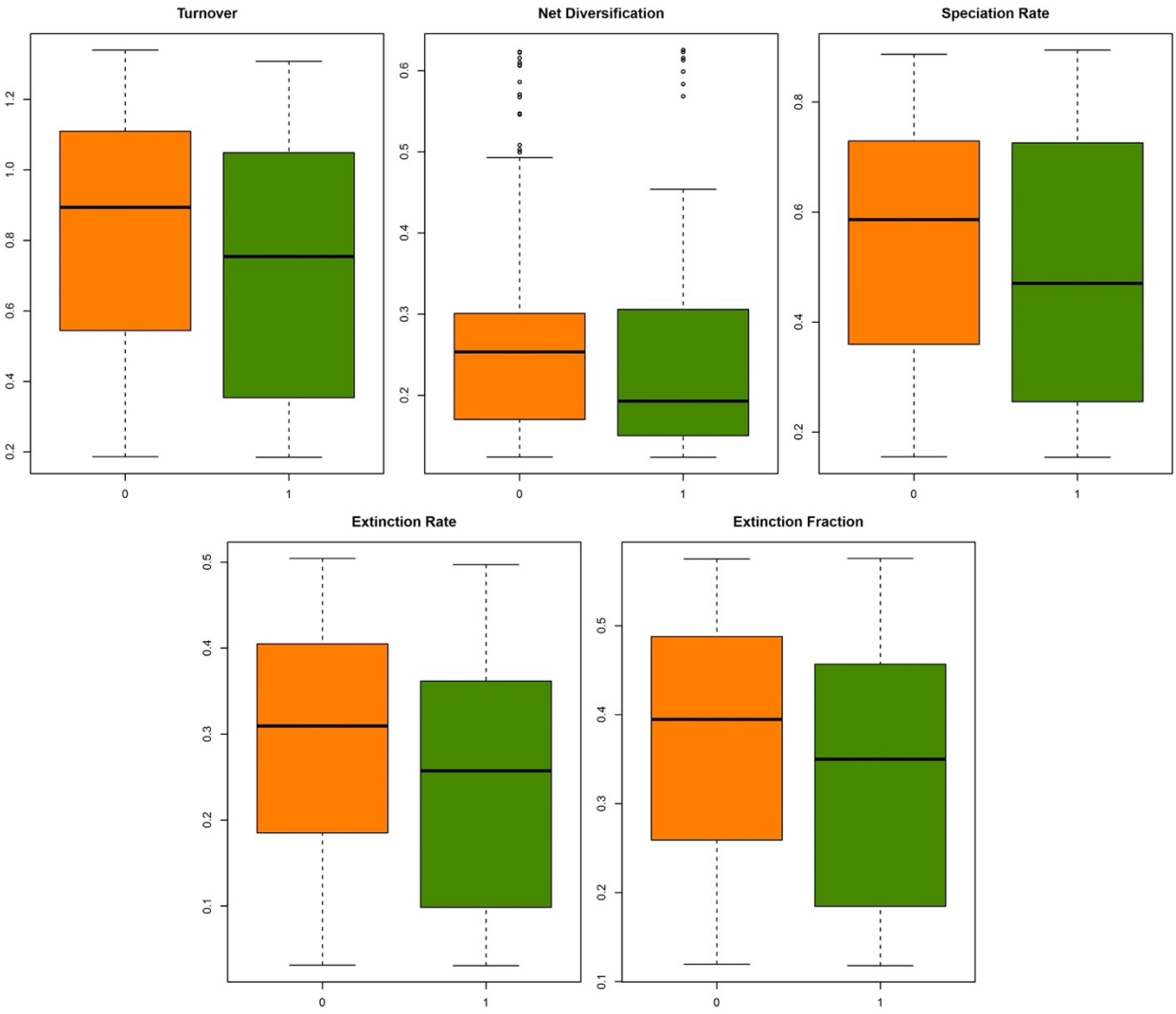
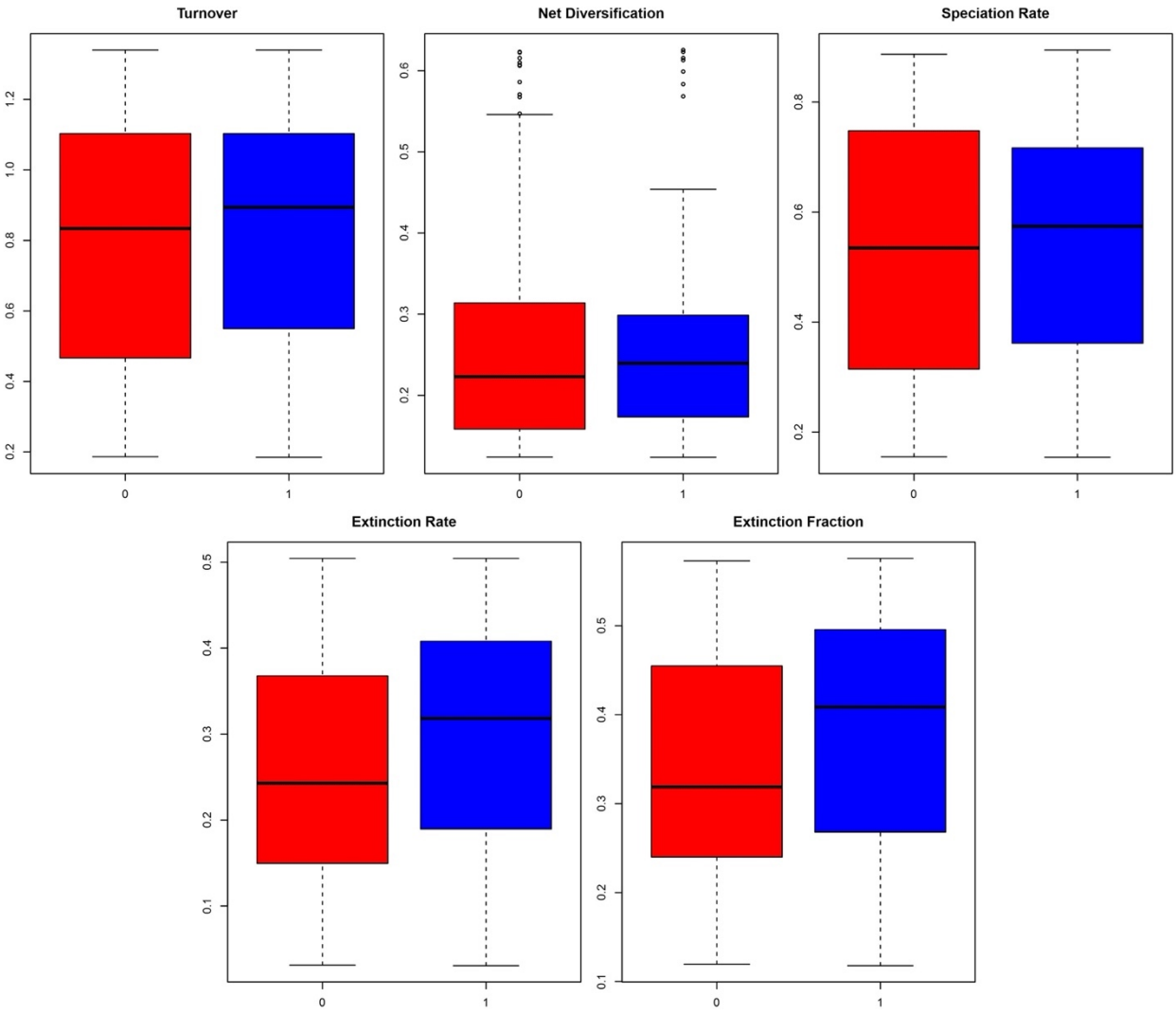


Figure 3 – Boxplots comparing tip rates across breeding system states. Self-incompatible species ("0") are shown in red while self-compatible ones ("1") are shown in blue. As with differences in ploidy state (Fig. 2), incorporating phylogenetic information in the summary statistics shows these differences are not significant and may have arisen by chance.



CHAPTER II

Historical causes for the greater proportion of polyploid plants in higher latitudes

Eric R. Hagen, Thais Vasconcelos, James D. Boyko, and Jeremy M. Beaulieu

Abstract

The proportion of polyploid plants in a community increases with latitude, and different hypotheses have been proposed about which factors drive this pattern. Here, we aim to understand the historical causes of the latitudinal polyploidy gradient using novel ancestral state reconstruction methods. Specifically, we assess whether (1) polyploidization enables movement to higher latitudes (i.e., polyploidization precedes occurrences in higher latitudes) or (2) higher latitudes facilitate polyploidization (i.e., occurrence in higher latitudes precedes polyploidization). We customize the discrete character evolution model corHMM to allow reconstruction of states at specific time slices rather than at the nodes of a phylogeny. We reconstruct the ploidy states and ancestral niches of 1,032 angiosperm species at four paleoclimatic time slices ranging from 3.3 million years ago to the present, comprising taxa from four well-represented clades: Onagraceae, Primulaceae, *Solanum* (Solanaceae), and Pooideae (Poaceae). Patterns of latitudinal movement are reconstructed and compared in relation to inferred ploidy shifts. We find that no single hypothesis applies equally well across all analyzed clades. While significant differences in median latitude and movement were detected in the largest clade, Pooideae, between species that diploidized and those that polyploidized, almost no differences were detected in our smallest clade, Primulaceae, in which few ploidy transitions were reconstructed. Given that patterns seem to be clade-specific, a higher number of clades must be analyzed in future studies for generalities be drawn.

Introduction

Polyploidy – i.e., the state of having more than two complete sets of chromosomes – has continually shaped the evolutionary history of flowering plants. Indeed, whole genome duplications are identified along the stem leading to all flowering plants as well as many events occurring all throughout some of the most diverse as well as some of the most depauperate clades nested within (Jiao et al. 2011). Through comparisons of diploid and polyploid plants, polyploidy appears linked to a variety of evolutionary changes, including novel phenotypic traits (Levin 1983), ecological relationships (Segraves 2017), and macroevolutionary patterns (e.g., Mayrose et al. 2011; Soltis et al. 2014). In biogeography, polyploidy is largely studied in the context of latitudinal and elevational gradients, in which polyploids tend to compose larger proportions of the flora at higher latitudes and elevations than at lower ones (Stebbins 1950; Brochmann et al. 2004; Rice et al. 2019). The so-called "latitudinal polyploidy gradient" (LPG) has long been observed in individual clades (e.g., Löve and Löve 1943; 1949), and recent studies incorporating large amounts of distribution data across clades have largely confirmed the generality of this pattern (Rice et al. 2019).

Proposed mechanisms responsible for the creation of the LPG can be divided into two categories. First, conditions of poleward environments lead to higher rates of polyploid formation at higher latitudes. Harsh environmental conditions like cold stress are known to induce polyploidy (De Storme and Geelen 2014; Lohaus and Van de Peer 2016), and the fragmented ranges of poleward environments could lead to allopolyploid formation via repeated contacts after range expansion (Stebbins 1985). Second, various adaptations of polyploids lead them to preferentially move into poleward environments at rates higher than those of diploids. Polyploids are believed to have generally greater colonizing ability than diploids due to higher

rates of self-compatibility (Bierzychudek 1985; Barringer 2007) and phenotypic plasticity (Price et al. 2003; Leitch and Leitch 2008). Thus, in the time since freezing conditions began to appear at northern latitudes during the Pliocene (Mudelsee and Raymo 2005), the LPG could have been caused by plant lineages generally moving to higher latitudes after polyploidization events.

These two scenarios, which we call the "centers of origin" and "centers of arrival" hypotheses, respectively, are not mutually exclusive. If polyploid plants are, in fact, generally better adapted to cold environments, one may expect to see both greater rates of polyploidization near the poles and greater movement into these environments of polyploids that originate elsewhere. It should also be noted that there remains the possibility that the LPG emerges passively. For example, Rice et al. (2019) found that global polyploid distribution is strongly correlated with climate, though they suggest that this is mainly indirect, as polyploids tend to be perennial (Van Drunen and Husband 2019), herbaceous plants (i.e., chamaephytes) that are low to the ground and able to survive the harsh conditions of poleward environments (Raunkiaer 1934). However, while present-day climatic variables do correlate with biogeographic patterns, the modern distributions of plants largely result from past climate changes (Normand et al. 2011). Additionally, correlations between specific traits and environmental variables may be shaped more by shared evolutionary history among species sharing those traits rather than functional relationships (Svenning and Skov 2007; Ma et al. 2016; Sundaram and Leslie 2021), so phylogenetic information must be considered as well. In any event, teasing apart the evidence for each scenario across flowering plants would provide invaluable clues about the historical causes for the LPG.

Such an investigation comes with some challenges. The first is the need to incorporate information about historical plant distributions, which is particularly difficult due to the large

number of biotic and abiotic factors that can potentially influence a species' geographic range. For instance, Rice et al. (2019) included paleoclimatic data from the Last Glacial Maximum (LGM; 21 kya) in their analysis, but this was only used in the context of correlating deglaciation extent with ploidy distributions, and they implicitly assumed that ranges remained unchanged to the present. While discrete-area methods of inferring ancestral ranges (e.g., Ree and Smith 2008; Matzke 2014) are in wide use, the reconstructions they create are usually coarse and contain few areas. Continuous-area methods requiring paleoclimatic data and shallower timescales are less common (Guillory and Brown 2021), but these methods are powerful for inferring latitudinal movement than tracking movement between arbitrarily designated geographic areas. The second difficulty is reconstructing ploidy changes through time. While ploidy can be reconstructed from fossils with preserved cuticle (McElwain and Steinthorsdottir 2017), fossil data is too sparse for a large-scale study. Instead, one would need to rely on reconstructions using a model of ploidy state transitions over time. One advantage of using such models is the ability to reconstruct not only polyploidization but also diploidization, which is the reorganization of the genome that returns a plant to a diploid (or "diploid-like") state after whole genome multiplication. Although we have no expectation of how species will move latitudinally following diploidization, it may be illuminating to compare movement between species that polyploidize as opposed to diploidize, as well as stay polyploid or diploid, as a kind of "control."

Here, in what we believe is the first attempt to discern the historic causes of the LPG, we analyze the distributions of plants in historical and phylogenetic context to determine how plants in specific clades move across latitudes after ploidy transitions. Specifically, by analyzing the timing of reconstructed ploidy changes and biogeographic movements, we tested the "centers of arrival" hypothesis, in which range movement towards higher latitudes happens most often after

polyploidization events (Fig. 1a), and the "centers of origin" hypothesis, in which polyploids form mostly at poleward environments and subsequently stay or move towards the equator (Fig. 1b).

Materials and Methods

Phylogenetic and ploidy datasets

We opted to use a multi-clade approach for this work, which aims to discern both biological generalities as well as clade-specific patterns (e.g., Boyko et al. 2023; Vasconcelos 2023). The main reason for choosing this approach is to reduce the impact of sampling bias in subsequent analyses of ancestral state and ancestral range reconstructions by focusing on clades that are particularly well sampled, as opposed to using super matrix trees (e.g., Smith and Brown 2018) that are unevenly sampled. These biases are also caused by available ploidy data being skewed toward certain taxonomic groups, particularly those studied in the Global North (Marks et al. 2021), and the fact that available GBIF data are incomplete as well as spatially clustered (Beck et al. 2014).

Our work makes use of four well-represented clades with relatively high availabilities of ploidy data and sampling at the species level: Onagraceae (Freyman and Höhna 2019), Primulaceae (De Vos et al. 2014), Pooideae (Poaceae; Spriggs et al. 2014), and *Solanum* (Solanaceae; Särkinen et al. 2013). The Onagraceae tree contains 292 species (c. 45% sampling; 186 with ploidy data), Primulaceae contains 263 species (c. 9.4% sampling [Xu and Chang 2017]; 141 with ploidy data), Pooidae contains 1,312 species (c. 40.6% sampling [Soreng et al. 2017]; 748 with ploidy data), and *Solanum* contains 441 species (c. 33.3% sampling; 256 with ploidy data). The Pooidae and *Solanum* trees were pruned from larger phylogenies of Poaceae

and Solanaceae, respectively, because the larger Poaceae and Solanaceae trees had data coverage of less than 50% for ploidy data, and pruning to include only these lower taxonomic rankings allowed us to focus on clades that are particularly data rich. These four clades were selected for this study because they are comparatively large, are well-represented in our ploidy dataset, are geographically widespread, come from different parts of the phylogeny of angiosperms (different major clades – rosids, asterids, and monocots), and because preliminary analyses recovered a relatively large number of recent polyploidization and diploidization events in the Quaternary.

Ploidy data was extracted from the supplementary data of Rice et al. (2019), which is contained in individual ChromEvol output files separated by genus. We combined these individual files into a master table and filtered it for species represented in our four phylogenies. In our analysis, we define a "polyploid" narrowly to specifically refer to a neopolyploid (i.e., newly formed polyploids; Ramsey and Schemske 2002), following the methodology of Rice et al. (2019). Neopolyploids are cytologically distinct from their diploid progenitors, and they have undergone whole genome multiplication sufficiently recently that they retain additive genome sizes of their parents as well as distinguishable subgenomes (Mandáková and Lysak 2018). In contrast, mesopolyploids and paleopolyploids are species that underwent polyploidization further in the past and have undergone diploidization to decrease their genome size as well as genome restructuring. We use this definition for two reasons: (1) the LPG is a gradient of plants that *are* polyploid (i.e., neopolyploids) rather than of plants that *behave* like polyploids (in the sense of gaining advantageous traits rather than chromosomal behavior), and (2) because we examined latitudinal changes after inferred events of both polyploidization and diploidization, so it did not make sense to consider re-diploidized plants in our analysis as polyploids, that is,

paleopolyploids (see “Integrating trait evolution models with reconstructions of past climatic niches”).

Distribution data

We downloaded all occurrence points available on GBIF that were based on preserved specimens (i.e., excluding human observations) for the four focal clades in our study (GBIF 2022). We then removed inaccuracies following protocols similar to those of Boyko et al. (2023). Our final occurrence point database accounts for 331,434 points in total, including 43,408 for *Solanum*, 11,200 for Primulaceae, 210,461 for Pooideae, and 66,365 for Onagraceae. After filtering for only those phylogenetically represented species with ploidy data and *sufficient* occurrence points (3 or more), the following number of species was analyzed for each clade: 543 in Pooideae, 218 in Solanum, 164 in Onagraceae, and 107 in Primulaceae, for a grand total of 1,032 species.

Integrating trait evolution models with reconstructions of past climatic niches

Most models that connect biogeographic shifts with discrete trait evolution require modeling areas discretely rather than continuously (e.g., Ree and Smith 2008; Goldberg et al. 2011; Caetano et al. 2018), and while methods have recently been developed to connect the evolution of discrete traits with continuous ones (e.g., Boyko et al. 2022), modeling range evolution continuously usually involves climatic parameters like temperature and precipitation. Because it is unclear whether the LPG may be caused by climatic factors or other biogeographic causes (e.g., Stebbins 1985), we opted to instead model range evolution and ploidy evolution separately and test for connections between the two post-hoc.

We began by modeling ploidy shifts along the phylogeny of each clade during the past c. 3.3 million years. To that end, we used corHMM (Beaulieu et al. 2013; Boyko and Beaulieu 2021) with modified functions that allow for ancestral state reconstruction at specific time slices

rather than at nodes. We designed this because we were interested in inferring ploidy states at shared time slices, specifically those with climatic data available from the Paleoclim database (Brown et al. 2018), rather than at asynchronous branching points (i.e., the nodes of a phylogeny) as is the default of the software (see Fig. 2). In our corHMM models, we calculate the conditional likelihood of an ancestral node, *k*, by multiplying the probabilities of observing the given character states of its descendants (see Felsenstein 1981). The likelihood function is given below, following Beaulieu et al. (2013),

$$L_k(c_k) = \left(\sum_{c_i} c_k c_i(t_i) L_i(c_i)\right)\left(\sum_{c_j} c_k c_j(t_j) L_j(c_j)\right)$$

where the conditional likelihood of an anagenetic taxon at a given time slice, *k*, is the product of the probability of observing all descendent character states from its descendant(s), node *i* and node *j*, given that *k* is in character state c_k. For each phylogeny, we tested three different model structures, none of which utilized hidden states: ER (equal transition rates between diploid state and polyploid state), ARD (transition rates between diploidy and polyploidy are allowed to vary), and a custom-made unidirectional structure where reversal to diploidy was disallowed after polyploidization. Some models of ploidy evolution (e.g., Robertson et al. 2011) disallow reversals to diploidy based on arguments that ploidy evolution is significantly asymmetrical (e.g., Stebbins 1971; Meyers and Levin 2006). However, much research suggests that reversals to diploidy are prevalent and potentially evolutionarily significant in flowering plants (Dodsworth et al. 2016), and other models of ploidy evolution reflect this (Zenil-Ferguson et al. 2019). Given our interest in both polyploidization and diploidization, we opted for the latter

strategy. We evaluated support for each model using AIC (Akaike 1974) and a significance criterion of a difference of at least 2 AIC units (Burnham and Anderson 2002).

Once corHMM models were run, we reconstructed ploidy states using the marginal reconstruction method given by Yang (2006, p. 121),

$$f(x_0|x_h;\theta) = \frac{f(x_0|\theta)f(x_h|x_0;\theta)}{f(x_h|\theta)} = \frac{\pi_{x_0}L_0(x_0)}{\Sigma_{x_0}\pi_{x_0}L_0(x_0)}$$

where $\boldsymbol{x}_h$ is the vector of tip states in the phylogeny, the numerator of the final equation is the joint probability of those tip states, x_0 is the ancestral state we are trying to estimate, π_{x0} is the prior probability that the state of the given ancestor is x_0, θ is the collection of parameters estimated by the corHMM model (transition rates), and $L_0(x_0)$ is the probability of observing the descendant tip(s) of the ancestor under study given the ancestral state. We used a novel marginal reconstruction function to calculate the marginal probabilities of anagenetic taxa occurring at the time slices for which we had paleoclimatic data.

Once ploidy shifts had been modeled, we reconstructed the range evolution of lineages in each tree using the recently developed package machuruku (Guillory and Brown 2021), a tool for phylogenetic niche modeling that allows for continuous reconstruction of ranges at time slices with paleoclimatic data as well as visualization of inferred spatial distributions. We reconstructed ranges at 4 time slices based on data from Paleoclim (Brown et al. 2018): the Last Interglacial (LIG, c. 130 ka), Marine Isotope Stage 19 (MIS19, c. 787 ka), the mid-Pliocene Warm Period (mPWP, c. 3.205 Ma), and Marine Isotope Stage M2 (M2, c. 3.3 Ma), all using the spatial resolution of 10 arc-minutes (~20 km). For each time slice, we first estimated tip response curves to each climatic variable using the function “machu.1.tip.resp,” then estimated the ancestral

niches of each taxon extant at each time slice with “machu.2.ace,” and finally projected the ancestral climatic niche models for each slice onto maps containing paleoclimatic variables with “machu.3.anc.niche.” We ran the “machu.3.anc.niche” function with the “clip.Q” option set to False, which produces models including less suitable areas but which prevented the function from returning NA results for some lineages.

Biogeographic analyses

To examine biogeographic movements through time, we parsed latitudinal changes between time slices concurrent with different ploidy transitions, characterizing species ranges by their median latitudes. We divided possible ploidy transitions into four ploidy status categories: (1) staying diploid, (2) staying polyploid, (3) diploidization, and (4) polyploidization. We retained for analysis species that did not change ploidy for three reasons: (1) to use as controls against which we could compare species that did change ploidy; (2) in order to see if we could recover the LPG in our data; and (3) because in some clades, no species underwent ploidy change in some time slices. The “centers of arrival” hypothesis would be supported when movement towards higher latitudes occurs more frequently after polyploidization than any other category of ploidy change. On the other hand, the “centers of origin” hypothesis will be supported where starting latitudes at the time slice when polyploidization occurs is significantly higher than for the other ploidy change categories. For each category, we tested for significant trends in movement (absolute latitudinal change) using a simple sign test (Conover 1971), employing the “binom.test” function in R (R Core Team 2022) to compare median latitudes at the beginning and end of each time slice. To account for the magnitude of change in addition to whether movement was generally equatorial or antiequatorial, we also conducted Wilcoxon signed-rank tests (Wilcoxon 1945) on the same data, both with and without phylogenetic weights incorporated.

To quantitatively compare whether latitudinal movements across all time slices significantly differed between species that polyploidized and those that diploidized, we conducted two-sided Kolmogorov-Smirnov tests (Smirnov 1939). General latitudinal movements in each category across time slices were also visualized using a density plot. Finally, to determine whether species that polyploidize possess ranges at significantly different latitudes relative to species in the other three ploidy status categories, we used two-sample t-tests (Student 1908) as well as phylogenetic paired t-tests (Revell 2012) to compare reconstructed median starting latitudes and latitudinal change across species. Comparisons were restricted to be conducted within clades and within the same time slices.

Results

General trends

We first examined the extent to which the latitudinal polyploidy gradient (LPG) is present among the four separate clades we examined. Except for Onagraceae, in all clades the present-day absolute latitudes of polyploid plants were, on average, significantly higher than those of diploids (Fig. 3; phylogenetic paired t-tests: Onagraceae p=0.690; p<0.0001 for Primulaceae, *Solanum*, and Pooideae). We next modeled ploidy shifts along the phylogeny of each clade, comparing a set of models that made different assumptions about transitions between each ploidy state. For example, based on AIC, within Onagraceae, the "all rates different" model (ARD), which assumed a unique rate for transitions from diploid to polyploid and from polyploid to diploid (AIC_{ARD}=82.58), was not supported over a model that assumed a single rate for both transition types ("equal rates," or ER; AIC_{ER}=80.63). When we assume unidirectional transitions, where only allowing transitions from diploid to polyploid are allowed, the fit was substantially

worse than ER and ARD (AIC_{uni}= 105.48). Results within the remaining three clades were mixed (Primulaceae, AIC_{ER} = 87.89, AIC_{ARD}=82.96, AIC_{uni}=92.74; Solanum, AIC_{ER} = 103.31, AIC_{ARD}=103.31, AIC_{uni}=103.0; Pooideae, AIC_{ER} = 503.09, AIC_{ARD}=498.46, AIC_{uni}=586.63), but always allowed some transitions between the two states. In other words, the unidirectional model was never favored. Inferred rates of polyploidization were 0.019 transitions Myr^{-1} in Onagraceae, 0.09 transitions Myr^{-1} in Primulaceae, 0.02 transitions Myr^{-1} in *Solanum*, and 0.21 transitions Myr^{-1} in Pooideae. Rates of diploidization were 0.019 transitions Myr^{-1} in Onagraceae, 0.23 transitions Myr^{-1} in Primulaceae, 0.02 transitions Myr^{-1} in *Solanum*, and 0.14 transitions Myr^{-1} in Pooideae. The four phylogenies, with marginal reconstructions of ploidy states at nodes rather than time slices, are depicted in Fig. 4. Ploidy was reconstructed using ARD in Primulaceae and Pooideae, as it was favored by >2 AIC units, and with ER in Onagraceae and *Solanum*, because we defaulted to the model with fewest parameters since no model was favored by AIC comparison.

Trends by Time Slice

Overall, when we correlated inferred ploidy shifts from the above model fits at particular time slices with estimated latitudinal changes, we found mixed support for the "centers of origin" and "centers of arrival" hypotheses, varying across clades and time slices. Across all clades, phylogenetic paired t-tests detected significant differences between the starting latitudes of lineages that polyploidized and lineages in the other ploidy status categories. The only non-significant comparison was between species that polyploidized vs. stayed polyploid in Onagraceae (p=0.890). Since corHMM reconstructed only one diploidization event in Onagraceae, and zero in *Solanum*, comparisons between the polyploidized and diploidized ploidy status groups could not be conducted in these clades. However, lineages that

polyploidized did not exhibit consistent patterns in starting latitudes relative to those that diploidized, and overall, latitude of origination does not seem to vary much among groups. In all clades except Pooideae the mean starting absolute latitude was higher in lineages that diploidized as opposed to polyploidized, and in all clades except Onagraceae, the outer-quartile range of lineages that polyploidized is contained within the ranges of lineages that stayed diploid or polyploid (see Figs. 5 and 6).

Regarding support for "centers of arrival," t-tests did not support a significant difference in latitudinal movement or starting latitude between lineages that diploidized as opposed to polyploidized. While we were able to recover the LPG, as in polyploids generally are located at higher latitudes than diploids (Figs. 3 and 4), visual inspection of the density plot (Fig. 4) shows that most species do not exhibit much movement, especially when they do not undergo ploidy shifts. For those that did, our general results are largely driven by Pooideae, for which corHMM recovered the largest number of ploidy shifts by far (Fig. 7). Pooideae was estimated to have undergone 43 polyploidizations and 53 diploidizations since the M2; the next largest number of events was in Primulaceae, with 7 polyploidizations and 5 diploidizations. Additionally, we have very little data about polyploidization and diploidization during the M2 and LIG slices because so few events were recovered, likely due to the small size of those slices (about 100,000 years each). Despite this, we were able to compare species across ploidy status categories in terms of their latitudinal movement.

A binomial sign test detected marginally significant directional movement only in species that polyploidized in Primulaceae (p=0.0625), which moved on average 9.11 degrees latitude antiequatorially, as expected under the "center of arrival" hypothesis. However, a Wilcoxon signed-rank test, which accounts for both direction and magnitude of movement, was not

significant for species that either diploidized (p=0.8125) or polyploidized (p=0.375) in the family. The only significant Wilcoxon test was found for species that polyploidized in Pooideae, both with (p=0.0358) and without (p=0.019) phylogenetic correction. In this case, lineages that polyploidized tended to move equatorially, rather than antiequatorially as expected under the "centers of arrival" hypothesis. Kolmogorov-Smirnov tests for differences between the changes in median latitudes of species that polyploidized vs. diploidized were all insignificant, and t-tests only found significant differences in latitudinal change between species that polyploidized and diploidized in Pooideae with phylogenetic correction (p=0.0006). Visual inspection of Fig. 8 suggests that diploidization generally leads to equatorial movement in Onagraceae, while all other groups appear to remain around 0, as in no latitudinal movement. This may be due to low sampling.

Discussion

corHMM Results

The lack of support for the unidirectional model for all clades was surprising given the assumption in many models of ploidy evolution that polyploidization is "irreversible" (Meyers and Levin 2006) and the fact that all polyploid tips were neopolyploids. In studies of neopolyploid evolution over short timescales, models frequently assume or show support for unidirectional ploidy evolution (e.g., Mayrose et al. 2011). Despite our machuruku analysis over the relatively short time scale of 3.3 million years, our corHMM findings are complicated by using phylogenies with root ages extending far beyond this boundary (c. 100 my in Onagraceae). Thus, although polyploid tips are strictly neopolyploids, as are the polyploids that compose the LPG in the present day, the lack of support for the unidirectional model makes sense in the light

of our deep time scale of study. Most analyses of ploidy evolution over deeper time scales allow for dual transitions between diploidy and polyploidy (e.g., Zenil-Ferguson et al. 2019). Regardless, it is interesting that the unidirectional model was not supported in Onagraceae or *Solanum* given the fact that corHMM reconstructed only one diploidization event in the former and none in the latter.

Causes of the LPG

Our study aimed to determine whether the LPG was better explained by greater rates of origination in or movement into poleward environments by polyploid flowering plants relative to diploid ones. At this point, given the low amount of support for either hypothesis, we do not favor one hypothesis over the other when it comes to general patterns. Species that polyploidize do show noticeable spikes in northward movement relative to other groups in some clades and time slices (Figs. 7 and 8), but these findings are countered by the mostly insignificant Wilcoxon and Kolmogorov-Smirnov test results.

The general lack of support for either greater rates of movement or origination at high latitudes relative to diploids may accord with a third hypothesis to explain the LPG that we did not test, that of higher latitude environments being “centers of survival” for polyploid plants. In this scenario, polyploids do not originate or move to poleward latitudes at higher rates relative to diploids. Rather, polyploids and diploids originate at the same rates in high latitude environments, but diploids go extinct more frequently than polyploids. In this case, the harsh environmental conditions hypothesized under “centers of origin” to create the conditions for higher rates of unreduced gamete formation, and thus polyploidization, instead filter out diploids in favor of polyploids, perhaps due to polyploidy conferring beneficial traits to tolerate abiotic stresses (Tossi et al. 2022). Our current analysis is unable to detect this possible pattern because

we do not examine diversification rates, and strong support for “centers of survival” would involve finding higher rates of extinction in diploids at high latitudes relative to polyploids. We hope that future work will address this and other possible mechanisms for the LPG.

The study conducted here is very much a preliminary one: despite our unclear results, we hope that phylogenetic-informed ecological niche modeling will continue to be used to study both the LPG and other biogeographic patterns. Such methods would be improved by the introduction of more sophisticated ancestral state reconstruction. In machuruku, ancestral characters are estimated assuming a simple Brownian motion model of evolution, and parameters underlying the evolutionary model are not free for the user to adjust or conduct model selection procedures. Fortunately, new work is being conducted studying whether bioclimatic variables are correlated with diversification rate changes (Zhang et al. 2021) and allowing for selective models of climatic evolution like Ornstein-Uhlenbeck models with hidden states (Boyko et al. 2023).

Clade-specific patterns

We were surprised to find little difference between movements in polyploidized vs. diploidized species in *Solanum*, as this is the only one of our groups that is distributed primarily in the southern rather than northern hemisphere (Olmstead and Palmer 1997). It is possible that their Andean center of richness causes species to move elevationally rather than latitudinally, though there is a noticeable spike in antiequatorial movement in lineages that polyploidized during the MIS19 (Fig. 7). Their Andean distribution may also explain the equatorial movement seen in *Solanum* during the mPWP. In the temperate clades, latitudinal differences are difficult to decipher, possibly due to the narrow and biased GBIF ranges centered on Europe. The largest group with the most reconstructed ploidy shifts, Pooideae, showed the most significant results by far, with the most tests showing significant differences in latitudinal movement among groups,

though movement is most clearly observed in species that diploidized during the LIG (Fig. 7). It is possible that the other, smaller clades with few reconstructed ploidy shifts leave us with little statistical power to detect associations between ploidy and latitudinal movement.

Alternatively, it is possible that the biogeographic patterns displayed by species in each ploidy status category, which compose the LPG, arose prior to the time scale we studied, such as during one of the Pliocene glaciation events in northern latitudes prior to the M2 (De Schepper et al. 2014). In this scenario, species may not exhibit significant movement in the present day or recent geologic past due to niches already being filled in polar environments. Additionally, there is the possibility that the LPG is created via polyploid formation due to secondary contacts of previously isolated populations confined to glacial refugia (Stebbins 1984; 1985). Testing this hypothesis would require comparing polyploid frequencies in deglaciated areas to non-deglaciated areas rather than a simple latitudinal comparison, and the pattern would likely emerge largely between the LGM (c. 21 ka) and the present. If this is the case, it would explain the lack of movement mostly observed in temperate clades, which possess ranges that overlap with potential glacial refugia (Comes and Kadereit 1998).

While we did find support for antiequatorial movement in Pooideae in species that polyploidized relative to those that diploidized, and the opposite pattern in Primulaceae, these findings may be better explained by methodological limitations rather than clade-specific traits. While rates of diploidization vary across species (Li et al. 2021), it is likely that full genome reorganization requires much more time than was included in in the 3.3 million years for which we possessed paleoclimatic data (Landis et al. 2018; Lynch and Conery 2003). Future studies may benefit from examining longer time scales than we considered here.

Caveats

Our study is not without important caveats. First, ploidy levels of tropically distributed plant species remain largely uncharacterized relative to those with temperate distributions (Husband et al. 2013; Vasconcelos 2023). This pattern is reflected in a large European bias in the distributions of plants included in this study. Additionally, our interpretations of how ploidy changes relate to subsequent latitudinal movements are limited by the available resolution of paleoclimatic data through time. For example, species that exhibit small amounts of latitudinal change after ploidy change may have transitioned soon before the end of the time slice, and movement in the subsequent time slice that may be caused by the ploidy change would not be detected by our methods. In other words, it is possible that latitudinal movement may occur after a "lag" (Schranz et al. 2012). While the lag hypothesis focuses on gaps between polyploidy and diversification, if lags are often required for the "success" of polyploids, this may also explain delayed ecological shifts or phenotypic shifts that enable range expansion and alteration. Our finding that lineages that stayed polyploid frequently occurred at higher latitudes relative to those that polyploidized (Figs. 5 and 6), which was not detected in our analyses of movement, may be evidence of such a lag.

Finally, the unevenness of our historical data makes it difficult to solidly connect ploidy shifts to subsequent latitudinal changes. Our time slices range widely in size: the gap between the M2 and the mPWP is smaller than 100,000 years, while our largest gap between the mPWP and MIS19 is almost 2.5 million years. The large number of ploidy shifts detected between the mPWP and the MIS19 could be attributed to climatic changes, as mean annual temperature declines during this period (Lisiecki and Raymo 2005), or to the relatively long period between these time slices. Additionally, while the inclusion of phylogeny in reconstructing ancestral

ranges will, in theory, produce better predictions, estimates can be spurious in cases where closely related species on the phylogeny exhibit widely disjunctive ranges. In the case of Pooideae, species in the genus *Aciachne* were reconstructed to have a very high median latitude around 50 degrees north in the MIS19 and prior. However, all three species of the genus included in our study have present-day median latitudes around -10 degrees south of the equator, making such large shifts suspect. This is likely driven by the biogeographic influence of closely related genera like *Oryzopsis*, in which all three of the species included in our dataset retain median latitudes around the range of 40 to 50 degrees north from the M2 to the present, and *Piptocheatium*, in which species ranges vary widely. As examples, *P. lasianthum* currently occurs in southern Brazil and northeast Argentina, while *P. avenaceum* occurs from Mexico to southeast Canada (POWO 2023).

Conclusions

Our first examination of the historical causes of the latitudinal polyploidy gradient found clade-specific differences in support for whether the pattern is driven more by polyploid origination at higher latitudes or polyploid movement to higher latitudes. When comparing the median latitudes and latitudinal movement across species that stayed polyploid, stayed diploid, polyploidized, and diploidized in individual time slices, we found significant differences in our largest clade, Pooideae. We also found significant differences in starting latitudes across clades, though the latitudinal relationship between species that polyploidized vs. diploidized varied. While we were able to detect the LPG in differences in median latitudes occupied by species that stay polyploid as opposed to stay diploid, we likely lack sufficient data to detect differences

between species that polyploidize as opposed to diploidize. We hope that further studies using similar methods will re-investigate this question with different, larger clades.

References

Akaike, H. 1974. A new look at statistical model identification. *IEEE Transactions on Automatic Control* AU-19: 716–722.

Barringer, B.C. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany* 94: 1527–1533.

Beaulieu, J.M., B.C. O'Meara, and M.J. Donoghue. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62: 725–737.

Beck, J., M. Böller, A. Erhardt, and W. Schwanghart. 2014. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics* 19: 10–15.

Bierzychudek, P. 1985. Patterns in plant parthenogenesis. *Experientia* 41: 1255–1264.

Brochmann, C., A.K. Brysting, I.G. Alsos, L. Borgen, H.H. Grundt, A.-C. Sheen, and R. Elven. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* 82: 521–536.

Boyko, J.D., and J.M. Beaulieu. 2021. Generalized hidden Markov models for phylogenetic comparative datasets. *Methods in Ecology and Evolution* 12: 468–478.

Boyko, J.D., E.R. Hagen, J.M. Beaulieu, and T. Vasconcelos. 2023. The evolutionary responses of life-history strategies to climatic variability in flowering plants. *New Phytologist.*

Brown, J.L., D.J. Hill, A.M. Dolan, A.C. Carnaval, and A.M. Haywood. 2018. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data* 5: 180254.

Burnham, K.P., and D.R. Anderson. 2002. Model selection and inference: a practical information-theoretic approach, 2nd ed. Springer-Verlag: New York, New York, USA.

Caetano, D.S., B.C. O'Meara, and J.M. Beaulieu. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* 72: 2308–2324.

Comes, H.P., and J.W. Kadereit. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* 3: 432–438.

Conover, W.J. 1971. Practical nonparametric statistics. John Wiley & Sons: New York, New York, USA.

De Schepper, S., P.L. Gibbard, U. Salzmann, and J. Ehlers. 2014. A global synthesis of the marine and terrestrial evidence for glaciation during the Pliocene Epoch. *Earth-Science Reviews* 135: 83–102.

De Storme, N., and D. Geelen. 2014. The impact of environmental stress on male reproductive development in plants: biological processes and molecular mechanisms. *Plant, Cell & Environment* 37: 1–18.

De Vos, J.M., R.O. Wüest, and E. Conti. 2014. Small and ugly? Phylogenetic analyses of the "selfing syndrome" reveal complex evolutionary fates of monomorphic primrose flowers. *Evolution* 68: 1042–1057.

Dodsworth, S., M.W. Chase, and A.R. Leitch. Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Botanical Journal of the Linnean Society* 180: 1–5.

Felsenstein, J. 1981. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biological Journal of the Linnean Society* 16: 183–196.

Freyman, W.A., and S. Höhna. 2019. Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. *Systematic Biology* 68: 505–519.

Goldberg, E.E., L.T. Lancaster, and R.H. Ree. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* 60: 451–465.

Guillory, W.X., and J.L. Brown. 2021. A new method for integrating ecological niche modeling with phylogenetics to estimate ancestral distributions. *Systematic Biology* 70: 1033–1045.

Husband, B.C., S.J. Baldwin, and J. Suda. 2013. The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. *In* Leitch, I.J., J. Greilhuber, J. Dolezel, and J. Wendel [eds.], Plant genome diversity, vol. 2, 255–276. Springer, Vienna, Austria.

Jiao, Y, N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph, L.P. Tomsho, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Landis, J.B., D.E. Soltis, Z. Li, H.E. Marx, M.S. Barker, D.C. Tank, and P.S. Soltis. 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* 105: 348-363.

Leitch, A.R., and I.J. Leitch. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* 320: 481–483.

Levin, D.A. 1983. Polyploidy and novelty in flowering plants. *The American Naturalist* 122: 1–25.

Li, Z., M.T.W. McKibben, G.S. Finch, P.D. Blischak, B.L. Sutherland, and M.S. Barker. 2021. Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology* 72: 387–410.

Lisiecki, L.E., and M.E. Raymo. 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* 20: PA1003.

Lohaus, R., and Y. Van de Peer. 2016. Of dups and dinos: evolution at the K/Pg boundary. *Current Opinion in Plant Biology* 30: 62–69.

Löve, A., and D. Löve. 1943. The significance of differences in the distribution of diploids and polyploids. *Hereditas* 29: 145–163.

Löve, A., and D. Löve. 1949. The geobotanical significance of polyploidy. I. Polyploidy and latitude. *In* Goldschmidt, R.B. [ed.], Portugaliae Acta Biologica, Série A special volume, p. 273–352. Instituto Botânico de Lisboa, Lisbon, Portugal.

Lynch, M., and J.S. Conery. 2003. The evolutionary demography of duplicate genes. *In* Meyer, A., and Y. Van de Peer [eds.], Genome evolution: gene and genome duplications and the origin of novel gene functions, 35–44. Springer, Dordrecht.

Ma, Z., B. Sandel, and J.-C. Svenning. 2016. Phylogenetic assemblage structure of North American trees is more strongly shaped by glacial-interglacial climate variability in gymnosperms than in angiosperms. *Ecology and Evolution* 6: 3092–3106.

Mandáková, T., and M.A. Lysak. 2018. Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology* 42: 55–65.

Marks, R.A., S. Hotaling, P.B. Frandsen, and R. VanBuren. 2021. Representation and participation across 20 years of plant genome sequencing. *Nature Plants* 7: 1571–1578.

Matzke, N.J. 2014. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts.

Mayrose, I., S.H. Zhan, C.J. Rothfels, K. Magnuson-Ford, M.S. Barker, L.H. Rieseberg, and S.P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.

McElwain, J.C., and M. Steinthorsdottir. 2017. Paleoecology, ploidy, paleoatmospheric composition, and developmental biology: a review of the multiple uses of fossil stomata. *Plant Physiology* 174: 650–664.

Meyers, L.A., and D.A. Levin. 2006. On the abundance of polyploids in flowering plants. *Evolution* 60: 1198–1206.

Mudelsee, M., and Raymo, M.E. 2005. Slow dynamics of the Northern Hemisphere glaciation. *Paleoceanography* 20: PA4022.

Normand, S., R.E. Ricklefs, F. Skov, J. Bladt, O. Tackenberg, and J.-C. Svenning. 2011. Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B* 278: 3644–3653.

Olmstead, R.G., and J.D. Palmer. 1997. Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Systematic Botany* 22: 19–29.

POWO. 2023. Plants of the World Online. Royal Botanic Gardens, Kew. http://www.plantsoftheworldonline.org/

Price, T.D., A. Qvarnström, and D.E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society B* 270: 1433–1440.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. https://www.R-project.org/.

Ramsey, J., and D.W. Schemske. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589–639.

Raunkiaer, C. 1934. The life forms of plants and statistical plant geography; being the collected papers of C. Raunkiaer. Clarendon Press, Oxford, UK.

Ree, R.H., and S.A. Smith. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.

Revell, L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Rice, A., P. Šmarda, M. Novosolov, M. Drori, L. Glick, N. Sabath, S. Meiri, et al. 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* 3: 265–273.

Robertson, K., E.E. Goldberg, and B. Igić. 2011. Comparative evidence for the correlated evolution of polyploidy and self-compatibility in Solanaceae. *Evolution* 65: 139–155

Särkinen, T., L. Bohs, R.G. Olmstead, and S. Knapp. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology* 13: 1–15.

Schranz, M.E., S. Mohammadin, and P.P. Edger. 2012. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology* 15: 147–153.

Segraves, K.A. 2017. The effects of genome duplications in a community context. *New Phytologist* 215: 57–69.

Smirnov, N.V. 1939. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Moscow University Mathematics Bulletin* 2: 3–26.

Soltis, D.E., M.C. Segovia-Salcedo, I. Jordon-Thaden, L. Majure, N.M. Miles, E.V. Mavrodiev, W. Mei, et al. 2014. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytologist* 202: 1105–1117.

Soreng, R.J., P.M. Peterson, K. Romaschenko, G. Davidse, J.K. Teisher, L.G. Clark, P. Barberá, et al. 2017. A worldwide phylogenetic classification of the Poaceae (Gramineae) II: an update and a comparison of two 2015 classifications. *Journal of Systematics and Evolution* 55: 259–290.

Spriggs, E.L., P.-A. Christin, and E.J. Edwards. 2014. C4 photosynthesis promoted species diversification during the Miocene grassland expansion. *PloS ONE* 9: e97722.

Stebbins, G.L. 1950. Variation and evolution in plants. Columbia University Press, New York, New York, USA.

Stebbins, G.L. 1971. Chromosomal evolution in higher plants. Addison-Wesley, London, UK.

Stebbins, G.L. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. *Botanica Helvetica* 94: 1–13.

Stebbins, G.L. 1985. Polyploidy, hybridization, and the invasion of new habitats. *Annals of the Missouri Botanical Garden* 72: 824–832.

Student 1908. The probable error of a mean. *Biometrika* 6: 1–25.

Sundaram, M., and A.B. Leslie. 2021. The influence of climate and palaeoclimate on distributions of global conifer clades depends on geographical range size. *Journal of Biogeography* 48: 2286–2297.

Svenning, J.-C., and F. Skov. 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters* 10: 453–460.

Tossi, V.E., L.J. Martinez Tosar, L.E. Laino, J. Iannicelli, J.J. Regalado, A.S. Escandón, I. Baroli, H.F. Causin, and S.I. Pitta-Álvarez. 2022. Impact of polyploidy on plant tolerance to abiotic and biotic stresses. *Frontiers in Plant Science* 13: 869423.

Van Drunen, W.E., and B.C. Husband. 2019. Evolutionary associations between polyploidy, clonal reproduction, and perenniality in the angiosperms. *New Phytologist* 224: 1266–1277.

Vasconcelos, T. 2023. A trait-based approach to determining principles of plant biogeography. *American Journal of Botany* 110: e16127.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* 1: 80–83.

Xu, Z., and L. Chang. 2017. Primulaceae. *In* Xu, Z., and L. Chang [eds.], Identification and control of common weeds, vol. 3, 51–81. Springer, Singapore.

Yang, Z. 2006. Computational molecular evolution. Oxford University Press, London, UK.

Zenil-Ferguson, R., J.G. Burleigh, W.A. Freyman, B. Igić, I. Mayrose, and E.E. Goldberg. 2019. Interaction among ploidy, breeding system and lineage diversification. *New Phytologist* 224: 1252–1265.

Zhang, X., J.B. Landis, Y. Sun, H. Zhang, N. Lin, T. Kuang, X. Huang, T. Deng, H. Wang, and H. Sun. 2021. Macroevolutionary pattern of Saussurea (Asteraceae) provides insights into the drivers of radiating diversification. *Proceedings of the Royal Society B* 288: 20211575.

Appendix

Figure 1. Conceptual diagram of historical biogeographic patterns expected to be observed under the "centers of arrival" hypothesis (Fig. 1a) and the "centers of origin" hypothesis (Fig. 1b). Under the "centers of arrival" scenario, polyploidization occurs across the globe but is followed by higher rates of antiequatorial movement relative to diploids, thus creating the LPG. Under the "centers of origin" scenario, the LPG is created by higher rates of polyploidization in poleward environments.

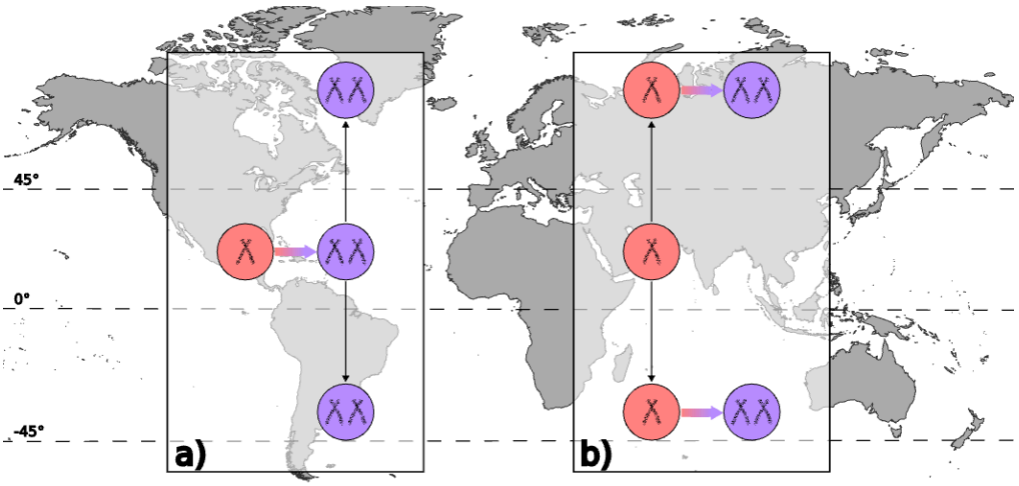


Figure 2. Conceptual figure showing our method of correlating inferred ploidy shifts at paleoclimatic time slices with estimated latitudinal changes, allowing for the connection of ploidy shifts to biogeographic movement. This scenario depicts the expectation under the "centers of arrival" hypothesis, in which shifts in ploidy (Fig. 2a) are followed by antiequatorial latitudinal movement (Fig. 2b).

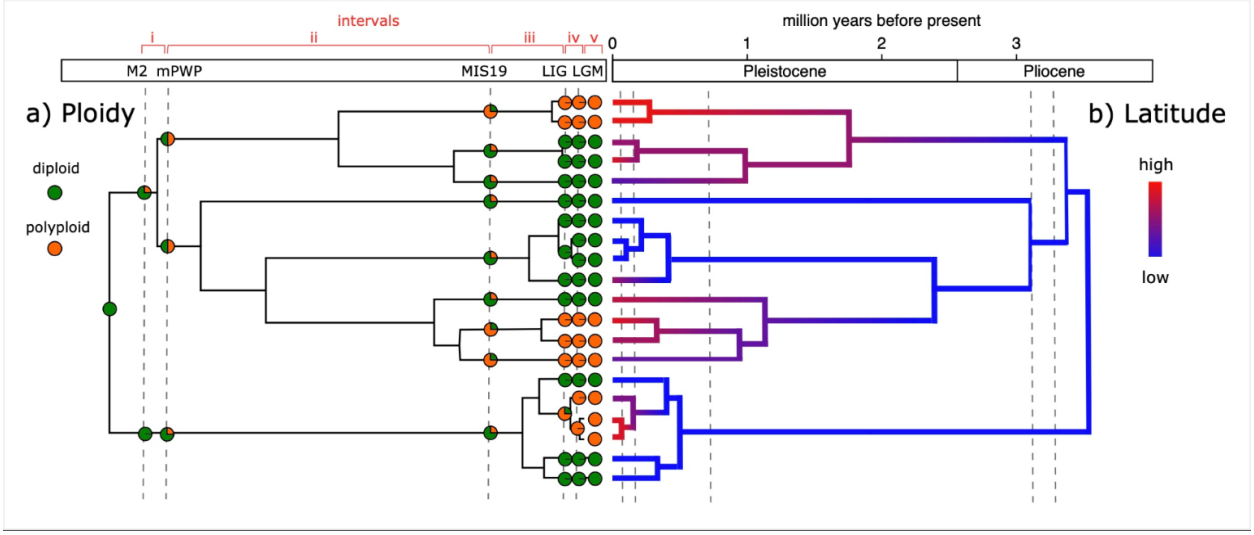


Figure 3. Boxplot showing present-day absolute latitudes of all plants in our dataset by ploidy and by clade.

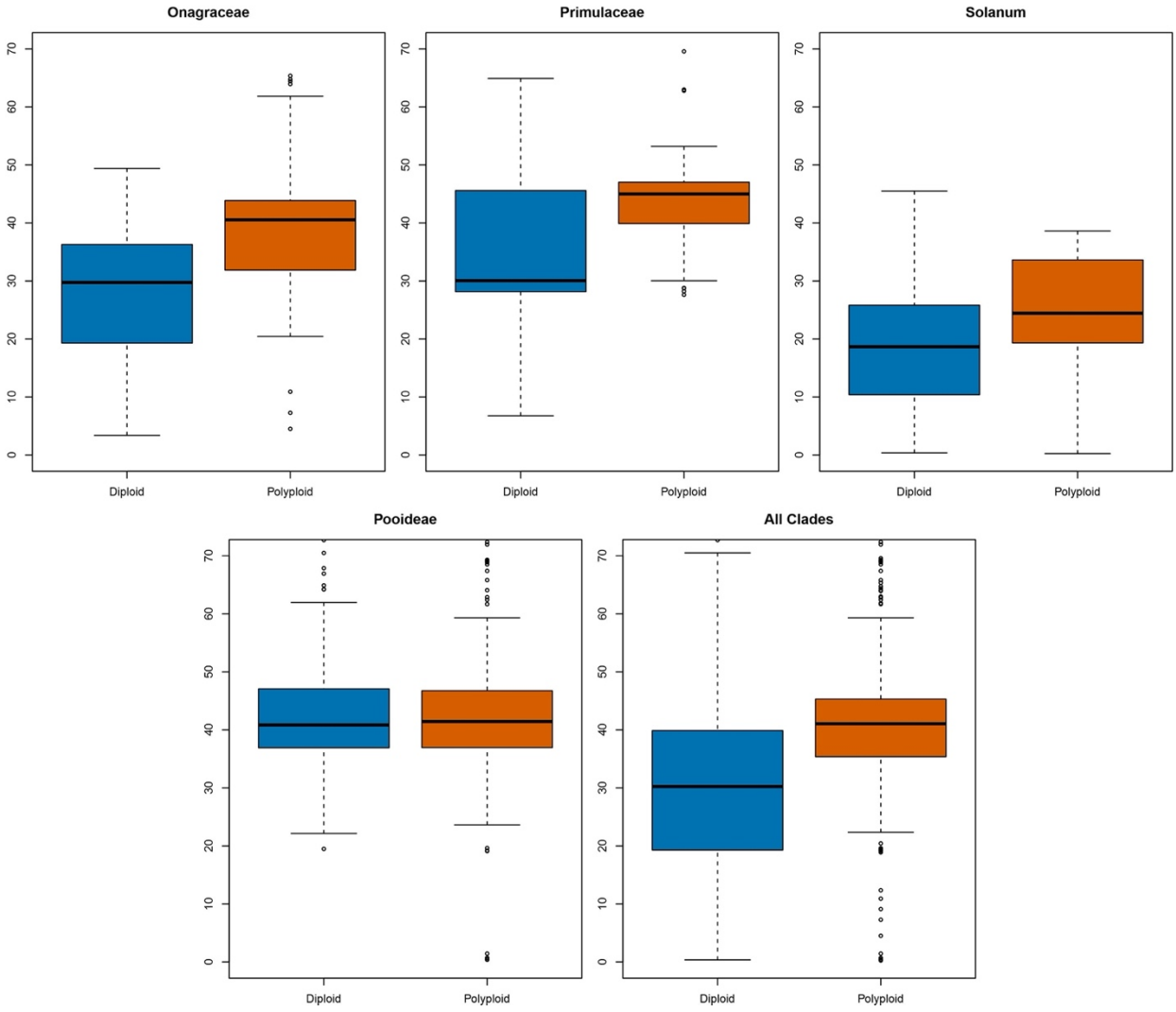


Figure 4. Plot showing the marginal reconstruction of ploidy states with traditional corHMM in, from top to bottom, Onagraceae, Pooideae, Primulaceae, and *Solanum*. Dotted lines near the tips represent the four time slices for which we possessed paleoclimatic data. Time in millions of years is shown at the top. Blue colors represent diploids, red colors represent polyploids, and gray colors represent tips with no ploidy data. Inset is a plot depicting latitudinal movement between time slices, averaged across all time slices but separated by clade.

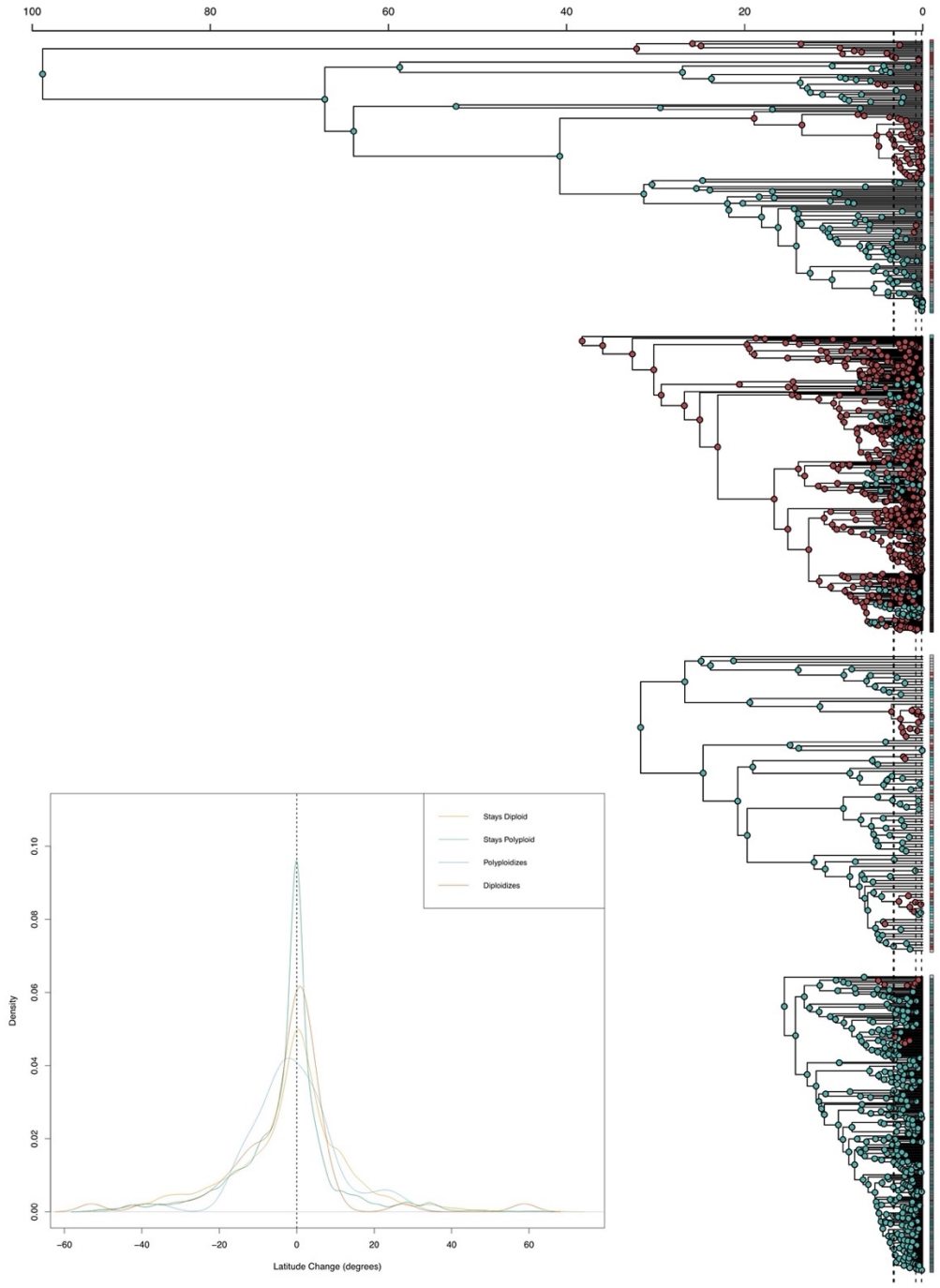


Figure 5. Boxplots comparing the starting absolute latitudes across each of the four ploidy groups, across all time slices, divided by clade. Ploidy status categories from left to right in each plot are: diploidized ("D"), polyploidized ("P"), stayed diploid ("SD"), and stayed polyploid ("SP").

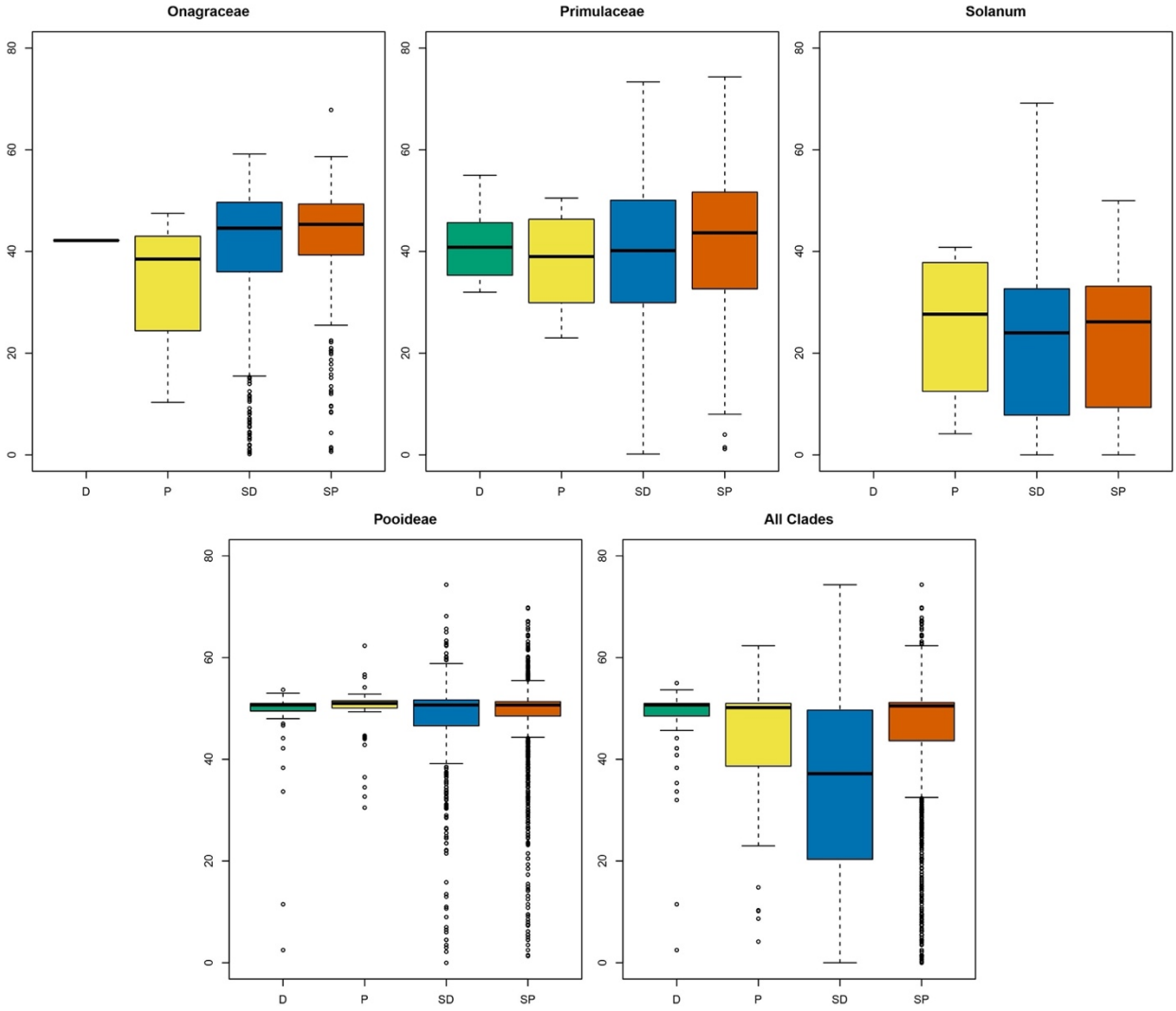


Figure 6. Boxplots comparing the starting absolute latitudes across each of the four ploidy groups, separated by clade and time slice. Ploidy status categories from left to right in each plot are: diploidized ("D"), polyploidized ("P"), stayed diploid ("SD"), and stayed polyploid ("SP"). Each column corresponds to a time slice (labeled at top with the beginning slice) and each row corresponds to a clade (with movement across all clades for each time slice in the bottom row).

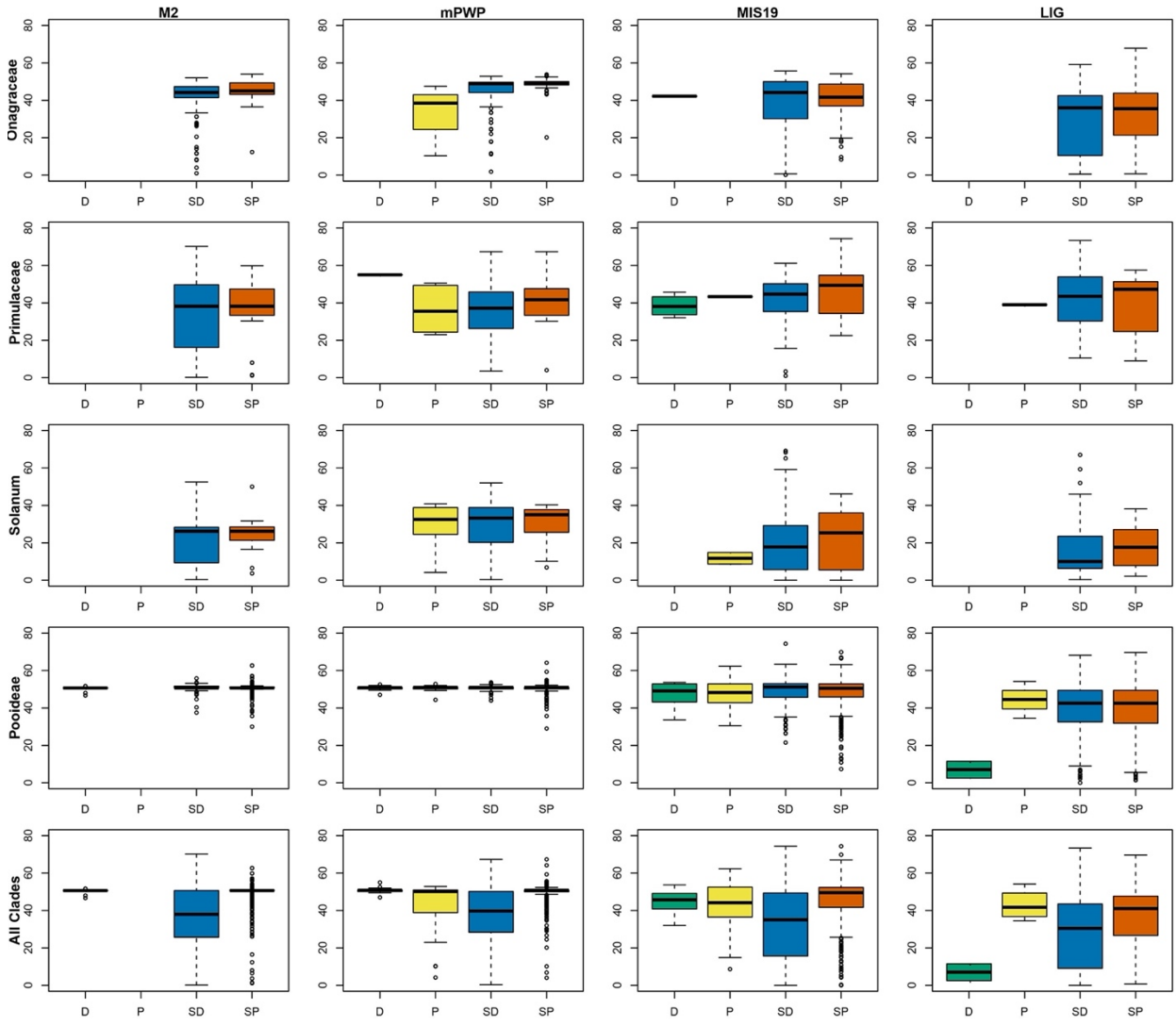


Figure 7. Boxplots comparing the change in median latitude within individual time slices and clades, divided by ploidy status group. Ploidy status categories from left to right in each plot are: diploidized ("D"), polyploidized ("P"), stayed diploid ("SD"), and stayed polyploid ("SP"). Each column corresponds to a time slice (labeled at top with the beginning slice) and each row corresponds to a clade (with movement across all clades for each time slice in the bottom row).

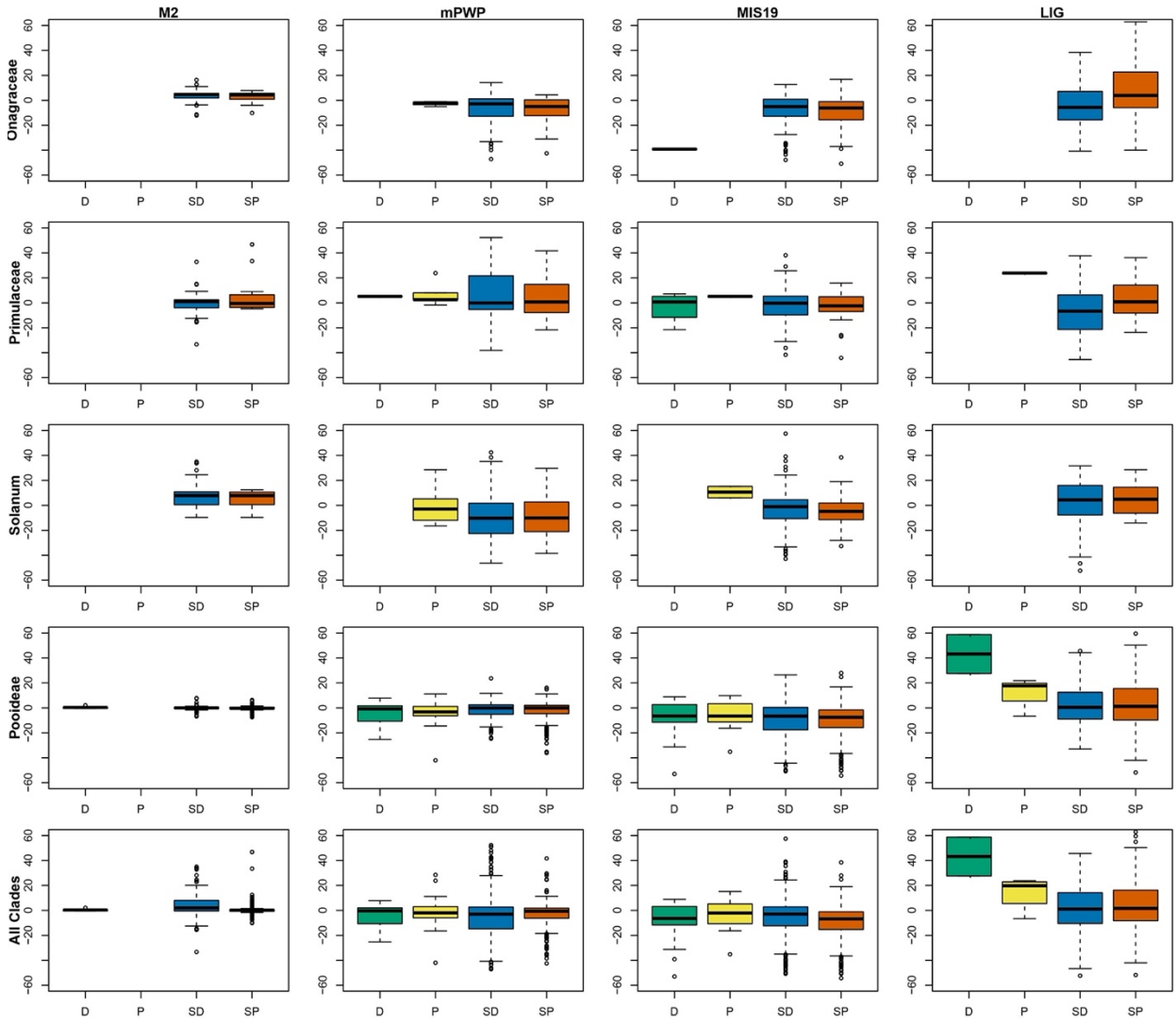
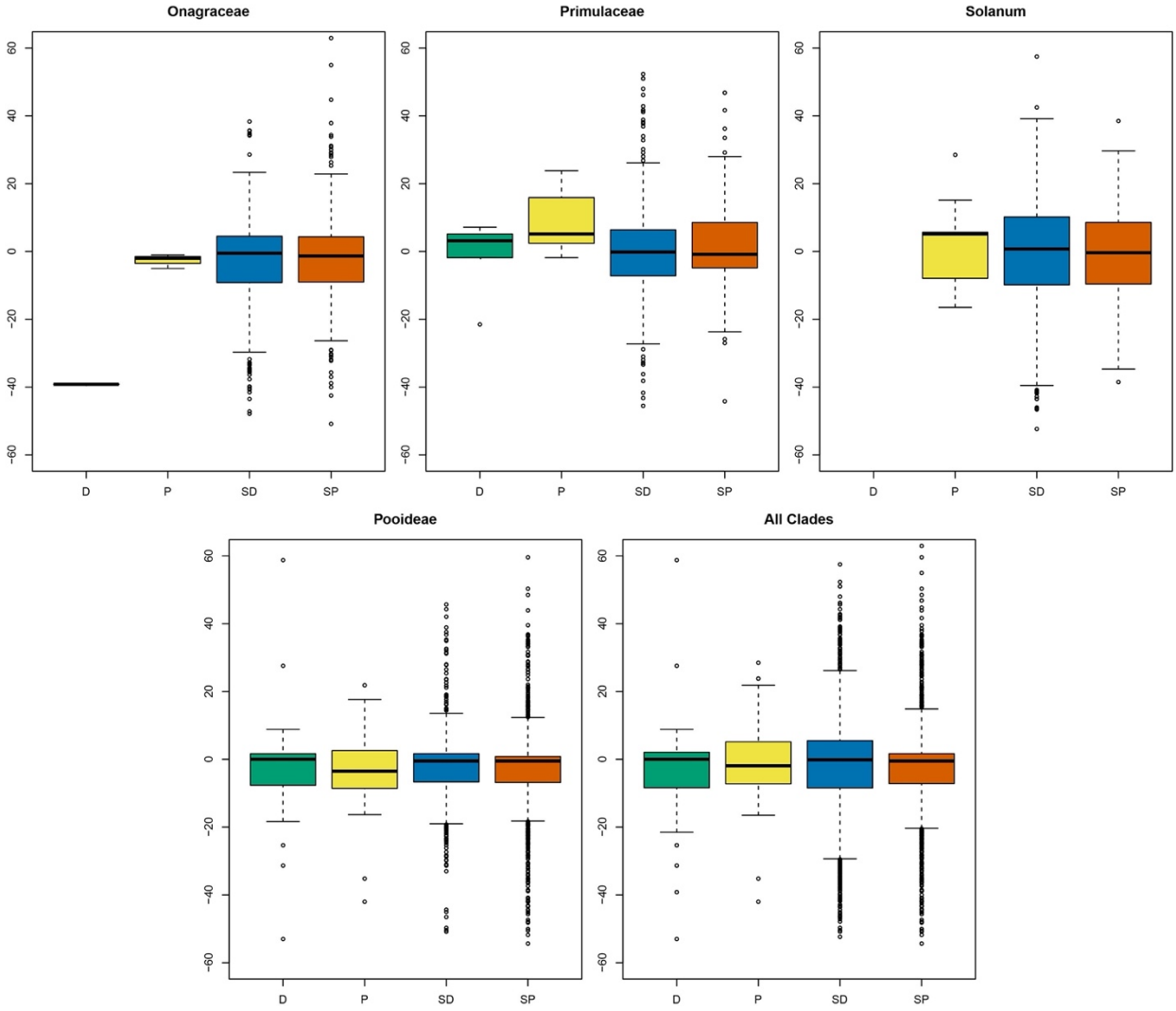


Figure 8. Boxplots comparing the change in median latitude across clades, averaged across all time slices and separated by ploidy status category. Ploidy status categories from left to right in each plot are: diploidized ("D"), polyploidized ("P"), stayed diploid ("SD"), and stayed polyploid ("SP").



CHAPTER III

Differences in pathogen resistance between diploid and polyploid plants: a systematic review and meta-analysis

Eric R. Hagen and Chase M. Mason

Abstract

Polyploidy has been hypothesized to provide increased ability to resist pathogens and parasites. However, studies comparing pathogen resistance in conspecific and congeneric diploids and polyploids have produced mixed results, and the supposed relationship between polyploidy and pathogen resistance has never been subjected to a systematic meta-analysis. We examined the effect of polyploidy on pathogen resistance by synthesizing 214 effect sizes from 128 studies. We find that, overall, there is no consistent effect of polyploidy on pathogen resistance. Subgroup analyses suggest that polyploids perform significantly better than diploids only in resisting hemibiotrophic pathogens, and autopolyploids tend show greater resistance than allopolyploids. This is surprising given the fact that polyploids possess extra allele copies of R-gene alleles that provide resistance to biotrophic pathogens, and this pattern may indicate that signaling cascades needed to elicit hypersensitive responses are disrupted by polyploidy. Disruption is supported by the observation that, across all pathogens, autopolyploids show significantly greater resistance compared to diploids, whereas allopolyploids do not. This is corroborated by the observation that synthetic autopolyploids perform significantly better than their allopolyploid and established counterparts. Analyses of publication bias indicate little to no bias, and analyses of heterogeneity indicate that phylogeny explains almost none of the observed heterogeneity. These results underscore the importance of not only systematic review but also the strong degree to which the effects of polyploidy depend on ecological context.

Introduction

Polyploidy, the state of having more than two full sets of chromosomes, has been studied in plants for over one hundred years by scientists in several biological fields (Ramsey and Ramsey 2014). This immediate doubling, tripling, or further multiplication of genetic variation on which evolution can act can lead to immediate speciation and considerable phenotypic changes, sometimes being called a "macromutation" (Goldschmidt 1940; Doyle and Coate 2020). Numerous reviews discuss the supposed benefits and disadvantages that polyploidy provides plants, arguing, for example, that it could increase growth rates (Udall and Wendel 2006), can enhance tolerance of various environmental stresses (Stebbins 1950), is associated with phenotypic "key innovations" that increase diversification rates (Soltis and Soltis 2016), and is associated with increased resistance to pathogens (e.g., Levin 1983).

In the case of pathogen resistance specifically, polyploidization can increase the production of existing secondary defense compounds (Levin 1976; Dhawan and Lavania 1996) or lead to the creation of novel metabolites (Schranz et al. 2011; Su et al. 2021). For these reasons, among others, polyploidy is commonly employed to improve cultivated plants (Touchell et al. 2020), including in some of the most common groups of crops in the world, such as wheat, bananas, brassicas, potatoes, and coffee (Kyriakidou et al. 2018). Polyploid cultivars in major crops often show the greatest tolerance to biotic stresses like infections, leading some to argue that artificially inducing polyploidy could be effective in mitigating the increased susceptibility of crops to pathogens that is expected under future climate change (Ruiz et al. 2020).

However, despite the purported positive relationship between polyploidy and pathogen resistance, empirical support for this association is lacking. The physiological effects of polyploidization are, in general, poorly understood (Soltis et al. 2010), and few non-agricultural

studies have examined the effect of polyploidy on tolerance of pathogens and parasites (Segraves and Anneberg 2016). While ecological modeling studies have shown support for a positive effect (Oswald and Nuismer 2007), narrative reviews of the literature suggest mixed results, indicating that the relationship is complex and dependent on ecological context (King et al. 2012; Segraves 2017). Additionally, in a recent systematic review and meta-analysis of the effect that plant polyploidy has on secondary metabolite composition, Gaynor et al. (2020) found no support for a consistent relationship. Despite the wide use of induced polyploidy as a method of crop improvement, there appears to be no existing meta-analysis that explicitly quantifies pathogen resistance in diploid and polyploid plants.

One difficulty with such an undertaking is that most appropriate studies that could be included in a meta-analysis examine human-bred plant cultivars in agricultural settings (Segraves and Anneberg 2016). If one finds a positive influence of polyploidy on pathogen resistance, it may be because the polyploids under study were not only induced but also subsequently selectively bred for favorable traits. Conversely, if one finds a negative relationship, this could be the result of comparing newly induced polyploids to diploid crops that have been selectively bred for specific resistances. Additionally, some cultivars used in such studies have been genetically modified for pathogen resistance and other traits, making it difficult to determine the specific contribution of ploidy to observed differences. To control for such factors, meta-analysis would need to not only include publications that studied both cultivated and wild species, but it would also need to filter out studies where genetic editing was used.

Here we report the results of the first such meta-analysis conducted to date. In our analysis, we included studies of cultivated species (those with a history of human cultivation and thus subject to either methodical or unconscious artificial selection; Darwin 1868) only if they

met several criteria, including that the aim of the paper must not involve active breeding of more resistant plants. Overall, we were able to calculate 214 effect sizes from 128 different studies. We find that current evidence supports no consistent effect of polyploidy on pathogen resistance in flowering plants, and any observed improvement in resistance that coincides with polyploidy is likely contingent on random chance and biological context.

Methods

Literature Search and Selection

The following literature search is briefly summarized in a PRISMA flow diagram (Fig. 1; O'Dea et al. 2021). Searches were performed in March 2022 with Google Scholar, employing individual searches for studies comparing diploids and triploids, tetraploids, pentaploids, hexaploids, octaploids, and decaploids (septaploids, dodecaploids, and others were excluded due to their rarity in the literature). The query terms used for each search, the number of papers that each returned, and the number of papers that remained after screening are shown in Table 1. The queries were highly specific due to the necessities of removing studies that examined plants bred or genetically modified to be pathogen resistant, removing ploidy comparison studies not focused on pathogen resistance, and capturing the many different types of plant pathogens under study.

In total, our searches returned 1,602 studies. During the abstract screening process, papers were removed from consideration if they involved breeding for superior traits, focused on genetic or biochemical underpinnings rather than pathogen resistance, had appeared in a previous search, included confounding variables, focused on non-pathogen organisms like aphids, were not written fully or partially in English, or generally studied irrelevant subject matter.

Additionally, 53 papers were removed from screening because, while Google Scholar showed full abstracts for these studies, Internet searches and/or interlibrary loan requests turned up no results for existing full texts. In many cases, these may have been conference proceedings or other publications of sets of abstracts without the publication of full texts. In the end, our search produced 100 articles that were determined to be eligible for meta-analysis, of which 73 were appropriate to be analyzed. In addition to articles identified through our systematic literature search, we also included 55 papers found through other means which met our inclusion/exclusion criteria, mainly previous ad hoc non-systematic searches on Google Scholar with variable Boolean language (see "Group" column in Table 2). The data is available in Tables 2–4 and is briefly summarized in Fig. 2.

Data Extraction

For both sets of articles, we extracted data from each paper. In addition to the means, standard deviations, and sample sizes (number of genotypes/varieties in each category) for both the diploid and polyploid groups for each effect size entry, the following items were recorded: paper authors, plant family, polyploid plant family, ploidy level of the polyploid, whether or not the diploids under study were hybrids (or a mix of hybrids and non-hybrids), whether the polyploids under study were autopolyploids (non-hybrids) or allopolyploids (hybrids) or a mix of both, whether the species under study were wild or cultivated (or a mix of both), whether the polyploids under study were synthetic (i.e., anthropogenically induced) or established (i.e., have undergone significant genome reorganization since polyploidization), the type of pathogen with which plants were infected (e.g., fungus, virus, etc.), and the effect direction (i.e., whether a higher value indicates greater or lesser pathogen resistance).

Scoring the moderators was straightforward except for polyploid type (allopolyploid or autopolyploid), cultivation status (cultivated or wild), and whether polyploids were synthetic vs. established. Polyploid type is difficult to categorize in binary form due to the many different, and often controversial, definitions of polyploidy that exist (Parisod et al. 2010). Additionally, cultivation histories of plants can be complicated or unclear, such as in einkorn wheat (Zaharieva and Monneveux 2014), and because of this it can also be difficult to determine whether a polyploid cultivar is man-made or naturally established. Therefore, in cases where studies did not explicitly label plants with regard to these variables, we defined inclusion criteria for these three moderators.

For polyploid type, we followed the simple definitions of Ramsey and Ramsey (2014), who designate as autopolyploids any polyploids arising from parents that are members of the same single species and define as allopolyploids any polyploids that derive from interspecific hybridization. Effect sizes with the label "both" come from papers where the polyploid group contained both allopolyploids and autopolyploids that were not separable. Studies for which it was still unclear what type of polyploid was under study were labeled as "unknown." For cultivation status, species were defined as "wild" very narrowly, following the definition in Gaynor et al. (2020) as having no history of anthropogenic manipulation whatsoever (whether through cultivation, induced polyploidy, or manual hybridization). Otherwise, the species were labeled as "cultivated," or "both" if both wild and anthropogenically manipulated species were not separable and contained in a single effect size. Studies for which we could not gather information about the presence or absence of anthropogenic manipulation were marked as "unknown." To remove the potential bias of studies where researchers bred plants for improved polyploid crops, publications were excluded from consideration if one group (diploids or

polyploids) was anthropogenically improved while the other was not. Improvement does not encompass papers that merely crossed species rather than used species which had undergone artificial selection for trait improvement or were otherwise explicitly identified as “improved.” To score polyploids as synthetic or established, we labeled as “synthetic” any paper which explicitly stated that the polyploid was developed anthropogenically during the study or soon before. Wild polyploids were automatically labeled as “established,” and all others were labeled as “unsure”. Because rates of genome reorganization vary widely (Li et al. 2021), and because it is unclear for many older polyploid cultivars whether polyploidization occurred naturally or anthropogenically, we erred on the side of caution in labeling most effect sizes as “unsure”.

We used WebPlotDigitizer Version 4.5 (Rohatgi 2021) to extract values for some articles that only included bar graphs instead of tables. Papers for which both mean and standard deviation could not be calculated were removed. In total, 214 effect sizes from 128 articles were recorded and able to be meta-analyzed.

Meta-analyses

We used the R Statistical Software (R Core Team 2016) versions 4.0.3 and 4.2.1 to perform all the following analyses. Effect sizes were calculated using standardized mean difference (SMD, i.e., Hedges’s g; Hedges 1981) with the *escalc* function from the R package metafor (Viechtbauer 2010). We used this metric because all studies measured pathogen resistance, but different studies used a variety of metrics to compare diploids and polyploids, from the proportion of surviving plants after infection to the Area Under Disease Progress Curve (AUDPC; Van der Plank 1963). Seven effect sizes were missing standard deviations for the diploid group, and five had no standard deviations for the polyploid group. So, following the method of Bracken (1992), these values were imputed prior to effect size calculation by multiplying the mean of the entry by

the quotient of the sum of all standard deviations from entries with complete information in the dataset, divided by the sum of all means.

After effect size directions were standardized based on the "Effect Direction" column in the dataset, we used multi-level meta-analytic models to systematically assess the data. This was done using the *rma* and *rma.mv* functions from metafor. The initial *rma.mv* model included four random effects: average infection time in days before disease incidence was calculated, the between-study effect (variation among effect sizes from different studies), the within-study effect (variation among effect sizes from the same study), and a *phylo* variable calculated using a family-level phylogeny of angiosperms (Qian and Zhang 2014). We were unable to examine the effect of phylogeny in the pathogen column due to the diversity of included organisms as well as the lack of robust phylogenies for pathogens like fungi, bacteria, and viruses (Gani et al. 2019). It also included six categorical moderator variables: different diploids (to account for effect sizes from the same publication comparing different sets of polyploids to the same set of diploids), ploidy level (triploid, tetraploid, etc.), polyploid type (autopolyploid, allopolyploid, or both included in the study), cultivation status (cultivated, wild, or both), whether polyploids were synthetic or established, and pathogen type (fungus, oomycete, bacterium, virus, or nematode). Since phylogeny explained almost none of the observed heterogeneity, we analyzed a second model in which between- and within-study effects were the only included random effects. We examined the amount of heterogeneity explained by each random effect using the I^2 statistic (Higgins and Thompson 2002) calculated using the *i2_ml* function from the package orchaRd (Nakagawa et al. 2020), and we determined whether there were significant differences in mean effect size between subgroups (i.e., moderators) by looking at the *p*-value of the "test of moderators" (Q_M statistic; Deeks et al. 2001) provided in the *rma.mv* output.

We examined the influence of publication bias on our results by creating a funnel plot (Egger et al. 1997) of all SMD values against their respective standard errors. Using that same plot, we tested for asymmetry using the trim-and-fill method (Duval and Tweedie 2000). Since funnel plot asymmetry can be caused by things other than publication bias (Nakagawa et al. 2021), we also used the following regression-based tests of publication bias: Egger's test of the relationship between residual effect size and study precision (Egger et al. 1997) and a test for time lag bias (Jennions and Møller 2002). Finally, even though fail-safe numbers do not adequately control for heterogeneity and non-independence (Nakagawa et al. 2021), we calculated Rosenthal's (1979), Orwin's (1983), and Rosenberg's (2005) fail-safe *N* statistics.

Results

Multi-level Modeling

Overall, we found no difference between diploids and polyploids in their abilities to resist pathogens ($Q_M = 25.645$, $p = 0.267$). Effect sizes in the full dataset ranged from a standardized mean difference of -6.73 to 4.75, with a mean effect size of -0.028. A caterpillars plot of these results can be seen in Fig. 3.

The first iteration of the mixed-effects model indicated that the random effect of phylogeny explained essentially none of the observed heterogeneity ($I^2 = 5.8 \times 10^{-8}$), so the mixed-effects model was run again with only between-study and within-study heterogeneity included as random effects. The results of this model suggest that none of the included moderators have significant influences on the degree to which polyploidy affects pathogen resistance, with no significant *p*-values inferred for any moderators. Additionally, the confidence intervals for all moderators overlap 0. Of the total observed heterogeneity ($I^2 = 93.4\%$), between-

study heterogeneity was larger than within-study heterogeneity (I^2 = 74.4% and 19%, respectively). The insignificant value of the test of moderators (Q_M) also indicates little variation between subgroups.

Subgroup Analysis

Despite the lack of significant moderators in our full multi-level model, single-moderator models shed light on interesting patterns of resistance in subgroups. Diploids seem to slightly outperform polyploids overall when all families are examined, but the performance is about equal when the two largest ones, Musaceae (n = 55) and Poaceae (n = 46), are removed (Q_M = 17.749, p = 0.34). While no family showed statistically significant resistance in either direction, Asparagaceae (n = 1) shows the strongest signal for diploid resistance over polyploid resistance (estimate = 2.163, p = 0.233), while Apocynaceae (n = 1) shows the strongest pattern in the opposite direction, though it is not significant (estimate = -1.941, p = 0.27; see Fig. 4).

Diploids do not exhibit greater pathogen resistance when compared against allopolyploids than against autopolyploids (Fig. 5). A simple *rma* model with "both" and "unknown" values removed shows that autopolyploids exhibit slightly greater resistance than diploids (estimate = -0.313, p = 0.073) while allopolyploid resistance is not significantly different from that of diploids (estimate = 0.126, p = 0.467). This pattern holds when pathogens are broken down by lifestyle, though diploids do show slightly greater resistance to biotrophic pathogens compared to autopolyploids (Fig. 6). When polyploids are divided into synthetic vs. established, synthetic autopolyploids tend to outperform diploids (estimate = -0.637, p = 0.068; see Fig. 7), and synthetic polyploids perform better relative to diploids than their counterparts in the "established" and "unknown" categories (estimate = -0.409, p = 0.05).

Diploids show significantly greater resistance to fungal (estimate = 0.911, $p = 0.047$) and nematode (estimate = 1.185, $p = 0.018$) pathogens (Fig. 8), while polyploids outperformed diploids in resisting hemibiotrophic pathogens (estimate = -0.895, $p = 0.043$). In individual subgroup analyses, no significant differences in resistance were observed on the basis of cultivation status.

Publication Bias

Across all included effect sizes, there is little evidence that publication bias significantly affects our meta-analysis. Visual inspection of our funnel plot (Fig. 9) shows a symmetrical distribution of SMD and standard error values. This was corroborated by trim-and-fill analysis, which produced no imputed studies and showed no evidence of significant bias ($p = 0.92$). However, Egger's regression test does suggest significant funnel plot asymmetry ($p = 0.02$), and individual trim-and-fill analyses of effect sizes found through systematic search and those found from other sources each showed evidence of significant bias ($p < 0.0001$ for both). While fail-safe N values varied widely (3,408 for Rosenthal's, 0 for Orwin's, and 33,141 for Rosenberg's), they generally suggest little bias. We also found no influence of publication year on our results ($p = 0.38$).

The standardized mean differences of the in-search effect sizes are significantly different from those found outside the systematic search ($Q_M = 8.22$, $p = 0.042$; see Fig. 10), with studies found outside our search showing greater pathogen resistance in diploids relative to polyploids. Trim-and-fill analysis of each group showed bias in opposite directions, but when these are combined, the total data shows little bias.

Discussion

Based on our analyses there is no evidence that polyploidy is consistently associated with overall increased (or decreased) resistance to pathogens and parasites in flowering plants. While the association has been suggested in previous narrative reviews (e.g., Levin 1983; Van de Peer et al. 2017), many have been cautious about proposing a general effect (e.g., King et al. 2012; Segraves and Anneberg 2016). Given the lack of any significant moderators in our general multi-level model, as well as the lack of any effect of phylogeny, our results support this caution. The effect of polyploidy on pathogen resistance likely depends greatly on factors like ecological context, time since polyploid formation and the degree of subsequent genome rearrangement, and the luck of the genomic draw.

We expected that polyploids would exhibit superior resistance than diploids because R-gene alleles, which mediate resistance to biotrophic and hemibiotrophic pathogens, might be present in double their quantity in polyploids relative to diploids. Instead, we found no significant differences in resistance to biotrophs between diploids and polyploids. While this may be due to chance alone, especially since polyploids show significantly greater resistance to hemibiotrophs, these results may instead indicate that polyploidy causes breakdown in R-gene signaling pathways, or that doubled R-genes are lost during diploidization (Innes et al. 2008; Soltis et al. 2010). This seems to be especially the case in allopolyploids, in which one would expect to see higher allelic diversity of R-genes, yet which consistently show decreased resistance relative to diploids. While the degree to which allopolyploidy disrupts proper genomic functioning is still uncertain (Parisod et al. 2010), it is plausible that signaling pathways are generally disrupted after allopolyploidization, and heterosis effects often seen in allopolyploids

may require processes like diploidization to reorganize genomes before beneficial traits can appear (Dodsworth et al. 2016).

The apparent superiority of synthetic polyploids, particularly synthetic autopolyploids, over established ones in resisting pathogens relative to their diploid counterparts may also be evidence that genome reorganization leads to a loss of R-gene alleles, though this finding is open to interpretation. For example, Clo and Kolář (2022) found that younger, synthetic polyploids exhibit lower amounts of inbreeding depression relative to ones that are older and/or established. Our finding suggests that, in crop improvement efforts, any initial advantages of polyploidy may be short-lived (without, perhaps, subsequent breeding efforts). However, we are cautious about these findings due to the difficulty in demarcating "synthetic" from "established" polyploids (Tayalé and Parisod 2013), especially in crop plants with unclear histories of anthropogenic intervention.

The large amount of heterogeneity indicated by the I^2 statistics (93.4% for the robust model with all effect sizes included) suggests that other factors besides those examined in our study may shed further light on the differences in pathogen resistance between polyploids and diploids. Within-study effects (19%) are present, but not terribly large, which bodes well for the ability of future researchers to identify other explanatory moderators, especially given the complexity of the effects of polyploidy and the difficulty of generalizing them across clades and ecological conditions (Stebbins 1950; Soltis and Burleigh 2009). This randomness as well as dependence on ecological context are likely large parts of what is being captured by the within-study effect. Polyploid success is highly contingent, depending on being at the "right place at the right time" (Oswald and Nuismer 2011). Regarding diversification over long periods, Sessa (2019) calls polyploidy a "Las Vegas strategy," where genome multiplications usually end in

"losses" (i.e., extinction), but on rare occasions cause plants to "win big" (i.e., succeed and diversify). When it comes to pathogen resistance, this randomness seems very explanatory, but in most cases, polyploidy appears to lead to only small losses and wins relative to diploids, being more of a "Reno strategy."

As anthropogenic climate change continues to raise global temperatures and atmospheric carbon dioxide concentrations, it is possible that plants could become more susceptible to pathogen attacks, raising the specter of massive crop losses (Lake and Wade 2009; Velásquez et al. 2018). While it has been proposed that experimentation with polyploidy may improve crops in the face of nutritional and growth losses expected under future climate change (Cheng et al. 2022), our results indicate that polyploidy is not a reliable path forward for increasing pathogen resistance in crops.

Conclusion

We found that there are no consistent overall differences between diploids and polyploids in their abilities to resist pathogens. None of the moderators included in our multi-level model showed significant effects. The overall similarity in resistance to biotrophic and hemibiotrophic pathogens between diploids and polyploids suggests that increased numbers of R-gene alleles do not lead to decreased infections in polyploids, and the lack of difference between diploids and polyploids in cultivated plants calls into question the utility of using polyploid crops for decreasing susceptibility to disease. Given the need for crop breeding strategies that can address the likely increase in disease susceptibility that will accompany future climate change, our results are disconcerting, but they may guide agriculturists toward other strategies for increasing crop resistance to infections.

References

Bracken, M.B. 1992. Statistical methods for analysis of effects of treatment in overviews of randomized trials. *In* Sinclair, J.C., and M.B. Bracken (eds.) Effecive care of the newborn infant. Oxford University Press.

Cheng, A., N.M. Hanafiah, J.A. Harikrishna, L.P. Eem, N. Baisakh, and M. Mispan. 2022. A reappraisal of polyploidy events in grasses (Poaceae) in a rapidly changing world. *Biology* 11: 636.

Clo, J., and F. Kolář. 2022. Inbreeding depression in polyploid species: a meta-analysis. *Biology Letters* 18: 20220477.

Darwin, C. 1868. The variation of animals and plants under domestication. John Murray.

Deeks, J.J., D.G. Altman, and M.J. Bradburn. 2001. Statistical methods for examining heterogeneity and combining results from several studies in meta-analyses. *In* Egger, M., G.D. Smith, and D. Altman (eds.) Systematic reviews in healthcare: meta-analysis in context (2nd edition). BMJ Books.

Dhawan, O.P. and U.C. Lavania. 1996. Enhancing the productivity of secondary metabolites via induced polyploidy: a review. *Euphytica* 87: 81–89.

Dodsworth, S., M.W. Chase, and A.R. Leitch. 2016. Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Botanical Journal of the Linnean Society* 180: 1–5.

Doyle, J.J., and J.E. Coate. 2020. Autopolyploidy: an epigenetic macromutation. *American Journal of Botany* 107: 1097–1100.

Duval, S., and R. Tweedie. 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56: 455–463.

Egger, M., G.D. Smith, M. Schneider, and C. Minder. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315: 629–634.

Gani, M., T. Hassan, P. Saini, R.K. Gupta, and K. Bali. 2019. Molecular phylogeny of entomopathogens. *In* Khan, M.A., and W. Ahmad (eds.) Microbes for sustainable insect pest management: an eco-friendly approach (volume 1). Springer Nature.

Gaynor, M.L., S. Lim-Hing, and C.M. Mason. 2020. Impact of genome duplication on secondary metabolite composition in non-cultivated species: a systematic meta-analysis. *Annals of Botany* 126: 363–376.

Goldschmidt, R.B. 1940. The material basis of evolution. Yale University Press.

Hedges, L.V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6: 107–128.

Higgins, J.P.T., and S.G. Thompson. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 21: 1539–1558.

Innes, R.W., C. Ameline-Torregrosa, T. Ashfield, E. Cannon, S.B. Cannon, B. Chacko, N.W. Chen, A. Couloux, A. Dalwani, R. Denny, and S. Deshpande. 2008. Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiology* 148: 1740–1759.

Jennions, M.D., and A.P. Møller. 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences* 269: 43–48.

King, K.C., O. Seppälä, and M. Neiman. 2012. Is more better? Polyploidy and parasite resistance. *Biology Letters* 8: 598–600.

Kyriakidou, M., H.H. Tai, N.L. Anglin, D. Ellis, and M.V. Strömvik. 2018. Current strategies of polyploid plant genome sequence assembly. *Frontiers in Plant Science* 9: 1660.

Lake, J.A., and R.N. Wade. 2009. Plant-pathogen interactions and elevated CO2: morphological changes in favour of pathogens. *Journal of Experimental Botany* 60: 3123–3131.

Levin, D.A. 1976. Chemical defenses of plants to pathogens and herbivores. *Annual Review of Ecology, Evolution, and Systematics* 4: 249–259.

Levin, D.A. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* 122: 1–25.

Li, Z., M.T.W. McKibben, G.S. Finch, P.D. Blischak, B.L. Sutherland, and M.S. Barker. 2021. Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology* 72: 387-410.

Nakagawa, S., M. Lagisz, R.E. O'Dea, J. Rutkowska, Y. Yang, D.W.A. Noble, and A.M. Senior. 2021. The orchard plot: cultivating a forest plot for use in ecology, evolution, and beyond. *Research Synthesis Methods* 12: 4–12.

Nakagawa, S., M. Lagisz, M.D. Jennions, J. Koricheva, D.W.A. Noble, T.H. Parker, A. Sánchez-Tójar, Y. Yang, and R.E. O'Dea. 2022. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution* 13: 4–21.

O'Dea, R.E., M. Lagisz, M.D. Jennions, J. Koricheva, D.W. Noble, T.H. Parker, J. Gurevitch, M.J. Page, G. Stewart, D. Moher, and S. Nakagawa. 2021. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biological Reviews* 96: 1695–1722.

Orwin, R.G. 1983. A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics* 8: 157–159.

Oswald, B.P., and S.L. Nuismer. 2007. Neopolyploidy and pathogen resistance. *Proceedings of the Royal Society B: Biological Sciences* 274: 2393–2397.

Oswald, B.P., and S.L. Nuismer. 2011. A unified model of autopolyploid establishment and evolution. *American Naturalist* 178: 687–700.

Parisod, C., R. Holderegger, and C. Brochmann. 2010. Evolutionary consequences of autopolyploidy. *New Phytologist* 186: 5–17.

Qian, H., and J. Zhang. 2014. Using an updated time-calibrated family-level phylogeny of seed plants to test for non-random patterns of life forms across the phylogeny. *Journal of Systematics and Evolution* 52: 423–430.

Qiu, T., Z. Liu, and B. Liu. 2020. The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports* 47: 5549–5558.

R Core Team. 2023. R: a language and environment for statistical computing. R Foundation for Statistical Computing.

Ramsey, J., and T.S. Ramsey. 2014. Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society B* 369: 20130352.

Rohatgi, A. 2021. WebPlotDigitizer Version 4.5.

Rosenberg, M.S. 2005. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 59: 464–468.

Rosenthal, R. 1979. The “file drawer problem” and tolerance for null results. *Psychological Bulletin* 85: 638–641.

Rothfels, C.J., and S.P. Otto. 2016. Polyploid speciation. *In* Kliman, R.M. (ed.) Encyclopedia of evolutionary biology. Academic Press.

Ruiz, M., J. Oustric, J. Santini, and R. Morillon. 2020. Synthetic polyploidy in grafted crops. *Frontiers in Plant Science* 11: 540894.

Schranz, M.E., P.P. Edger, J.C. Pires, N.M. van Dam, and C.W. Wheat. 2011. Comparative genomics in the Brassicales: ancient genome duplications, glucosinolate diversification and Pierinae herbivore radiation. *In* Edwards, D., J. Batley, I. Parkin, and C. Kole (eds.) Genetics, genomics and breeding of oilseed brassicas. CRC Press.

Segraves, K.A. 2017. The effects of genome duplications in a community context. *New Phytologist* 215: 57–69.

Segraves, K.A., and T.J. Anneberg. 2016. Species interactions and plant polyploidy. *American Journal of Botany* 103: 1326–1335.

Senior, A.M., C.E. Grueber, T. Kamiya, M. Lagisz, K. O'Dwyer, E.S.A. Santos, and S. Nakagawa. 2016. Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology* 97: 3293–3299.

Sessa, E.B. 2019. Polyploidy as a mechanism for surviving global change. *New Phytologist* 221: 5–6.

Soltis, D.E., R.J.A. Buggs, J.J. Doyle, and P.S. Soltis. 2010. What we still don't know about polyploidy. *Taxon* 59: 1387–1403.

Soltis, D.E., and J.G. Burleigh. 2009. Surviving the KT mass extinction: new perspectives of polyploidization in angiosperms. *Proceedings of the National Academy of Sciences* 106: 5455–5456.

Soltis, P.S., and D.E. Soltis. 2016. Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology* 30: 159–165.

Stebbins, G.L. 1950. Variation and evolution in plants. Columbia University Press.

Su, W., Y. Jing, S. Lin, Z. Yue, X. Yang, J. Xu, J. Wu, Z. Zhang, R. Xia, J. Zhu, and N. An. 2021. Polyploidy underlies co-option and diversification of biosynthetic triterpene pathways in the apple tribe. *Proceedings of the National Academy of Sciences* 118: e2101767118.

Tayalé, A., and C. Parisod. 2013. Natural pathways to polyploidy in plants and consequences for genome reorganization. *Cytogenetic and Genome Research* 140: 79–96.

Touchell, D.H., I.E. Palmer, and T.G. Ranney. 2020. *In vitro* ploidy manipulation for crop improvement. *Frontiers in Plant Science* 11: 722.

Udall, J.A., and J.F. Wendel. 2006. Polyploidy and crop improvement. *Crop Science* 46: S3–S14.

Van de Peer, Y., E. Mizrachi, and K. Marchal. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18: 411–424.

Van der Plank, J.E. 1963. Plant disease epidemics and control. Academic Press.

Velásquez, A.C., C.D.M. Castroverde, and S.Y. He. 2018. Plant-pathogen warfare under changing climate conditions. *Current Biology* 28: R619–R634.

Viechtbauer, W. 2010. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 36: 1–48.

Zaharieva, M., and P. Monneveux. 2014. Cultivated einkorn wheat (*Triticum monococcum* L. subsp. *monococcum*): the long life of a founder crop of agriculture. *Genetic Resources and Crop Evolution* 61: 677–706.

Appendix

Figure 1 – PRISMA flow diagram depicting our systematic literature search and application of inclusion-exclusion criteria.

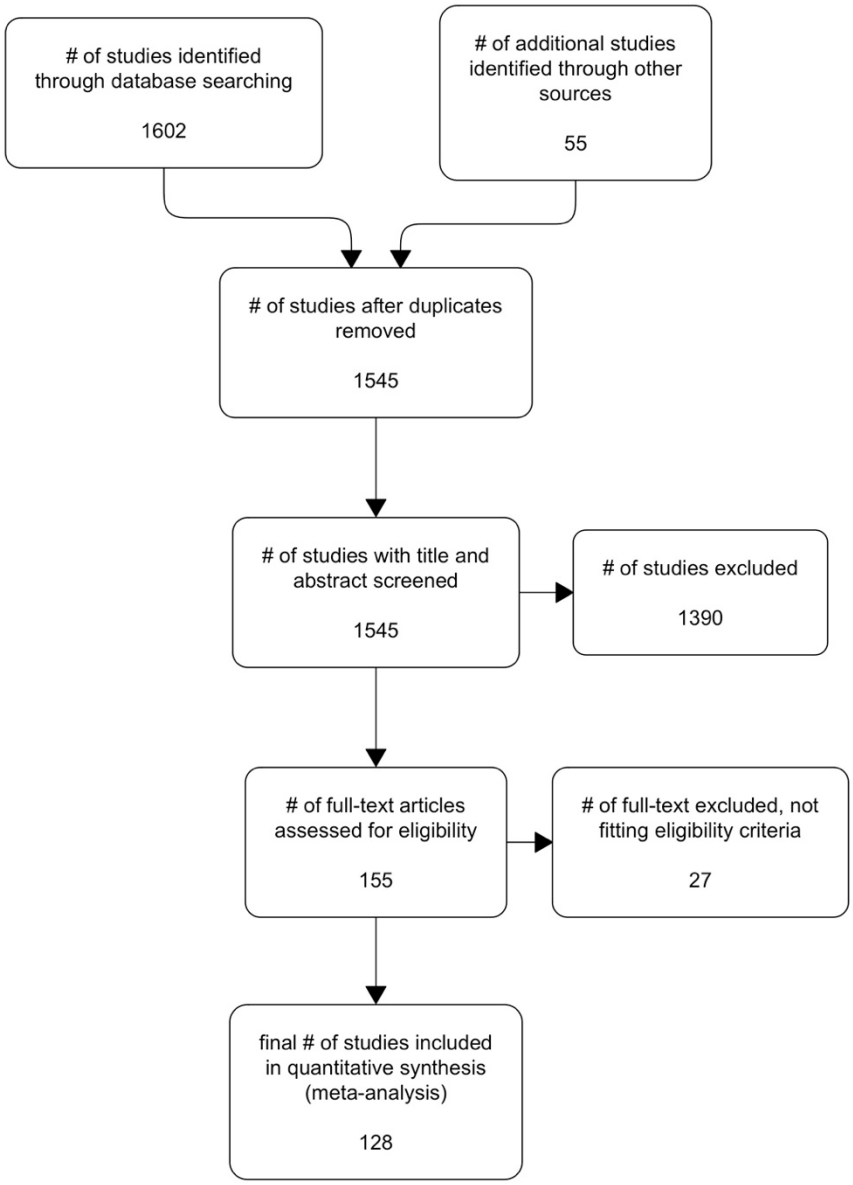


Figure 2 – Plots summarizing the studies included in our meta-analysis. (A) Histogram showing the number of papers published per 8-year period between 1957, the year of the earliest included publication, and 2021, when the most included recent papers were published. (B) Pie chart displaying the proportions of pathogens studied across publications. (C) Pie chart displaying the proportions of families studied across publications.

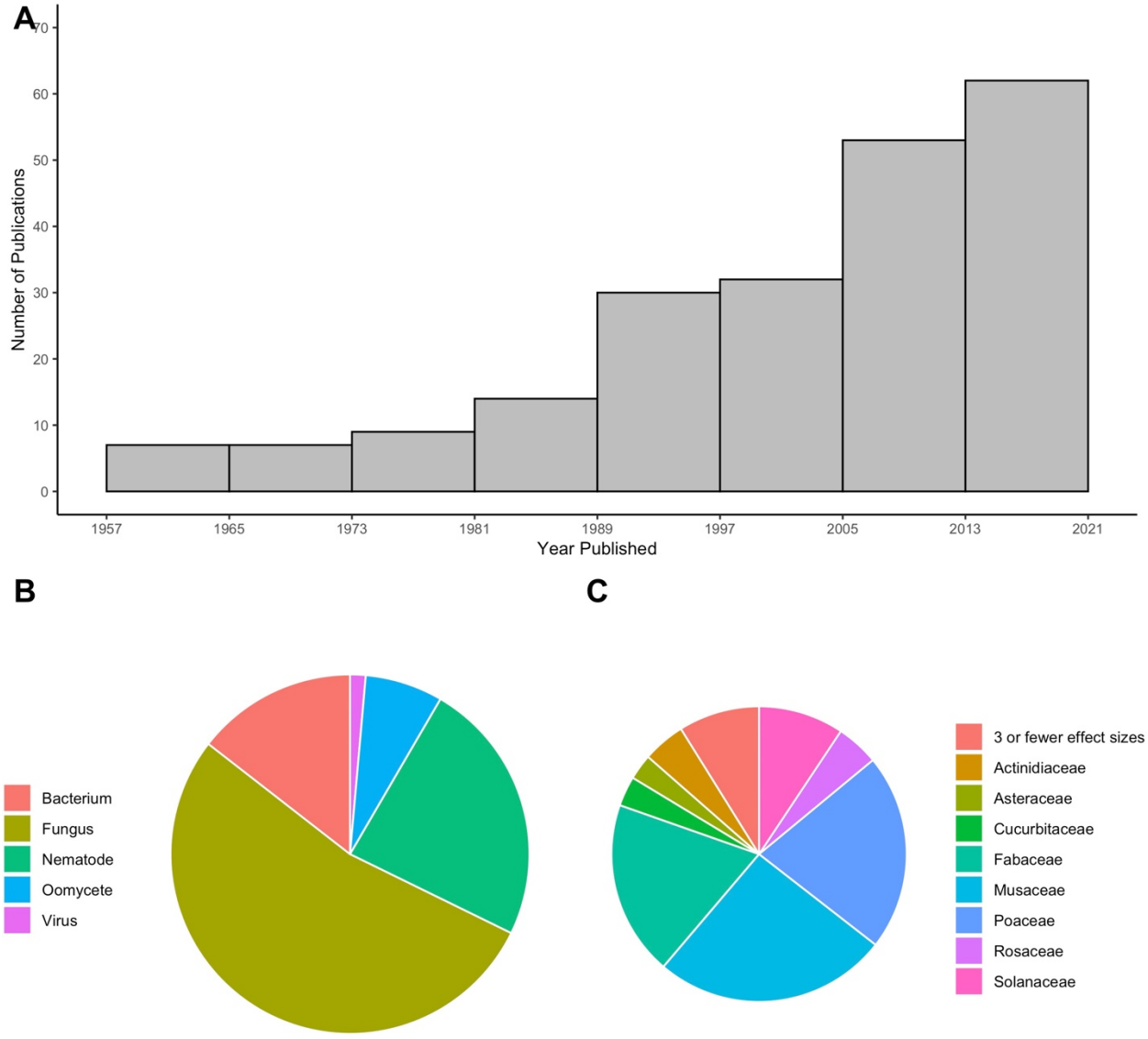


Figure 3 – A caterpillars plot showing the effect sizes (yellow circles) and 95% confidence intervals (green bars) of all 214 effect sizes included in the meta-analysis. The total I2 value is shown in the lower left, and the overall point estimate and confidence interval are displayed in red at the bottom. This plot was made using the *caterpillars* function in the R package orchaRd (Nakagawa et al. 2020).

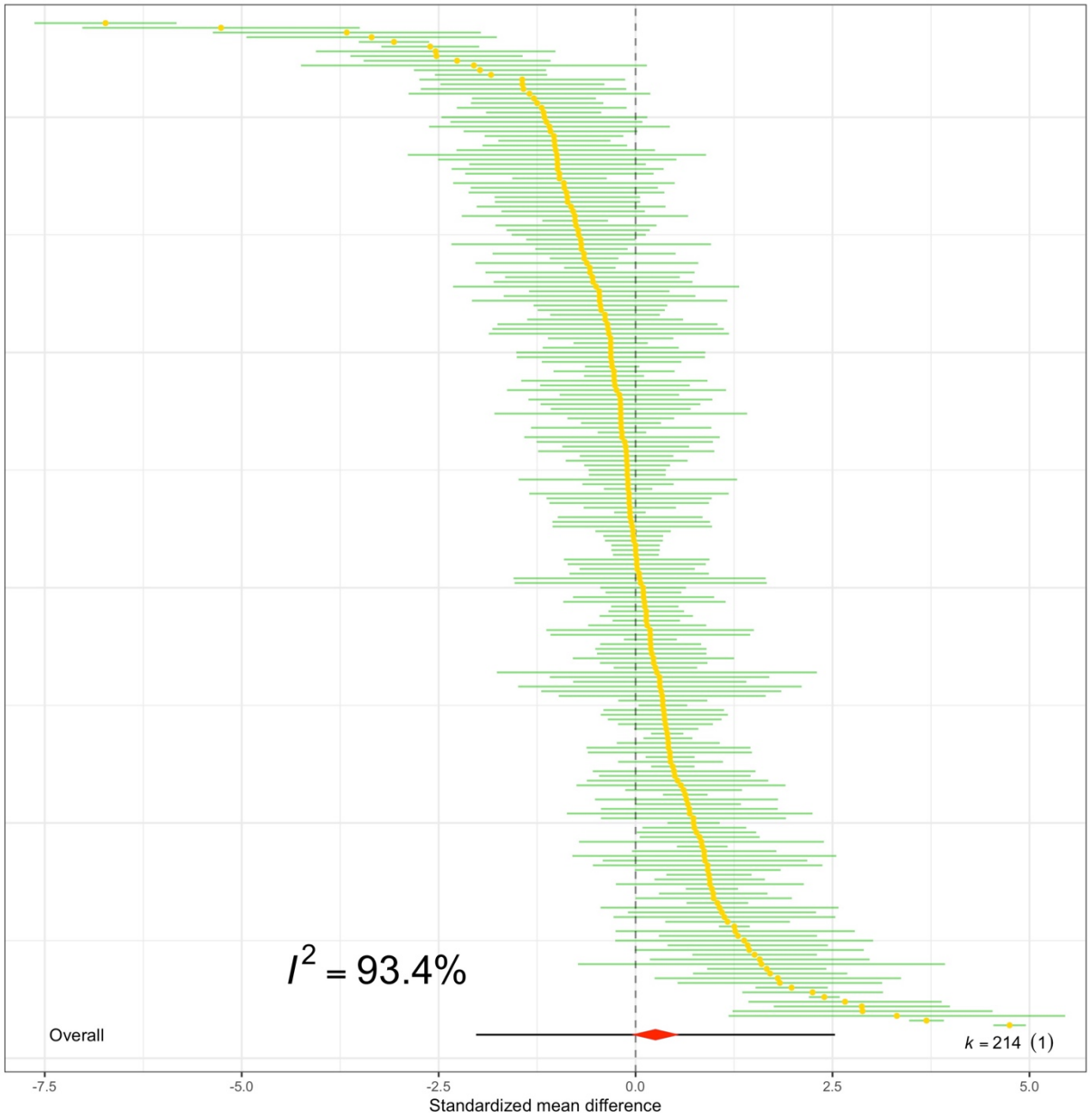


Figure 4 – An orchard plot showing the distribution of effect sizes across plant families. *k* is the number of effect sizes, and numbers in parentheses are the number of studies. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles) while 95% prediction intervals are shown with thinner lines. Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. Poaceae shows the strongest advantage of diploids over polyploids in pathogen resistance, while Apocynaceae shows the strongest difference in resistance in favor of polyploids.

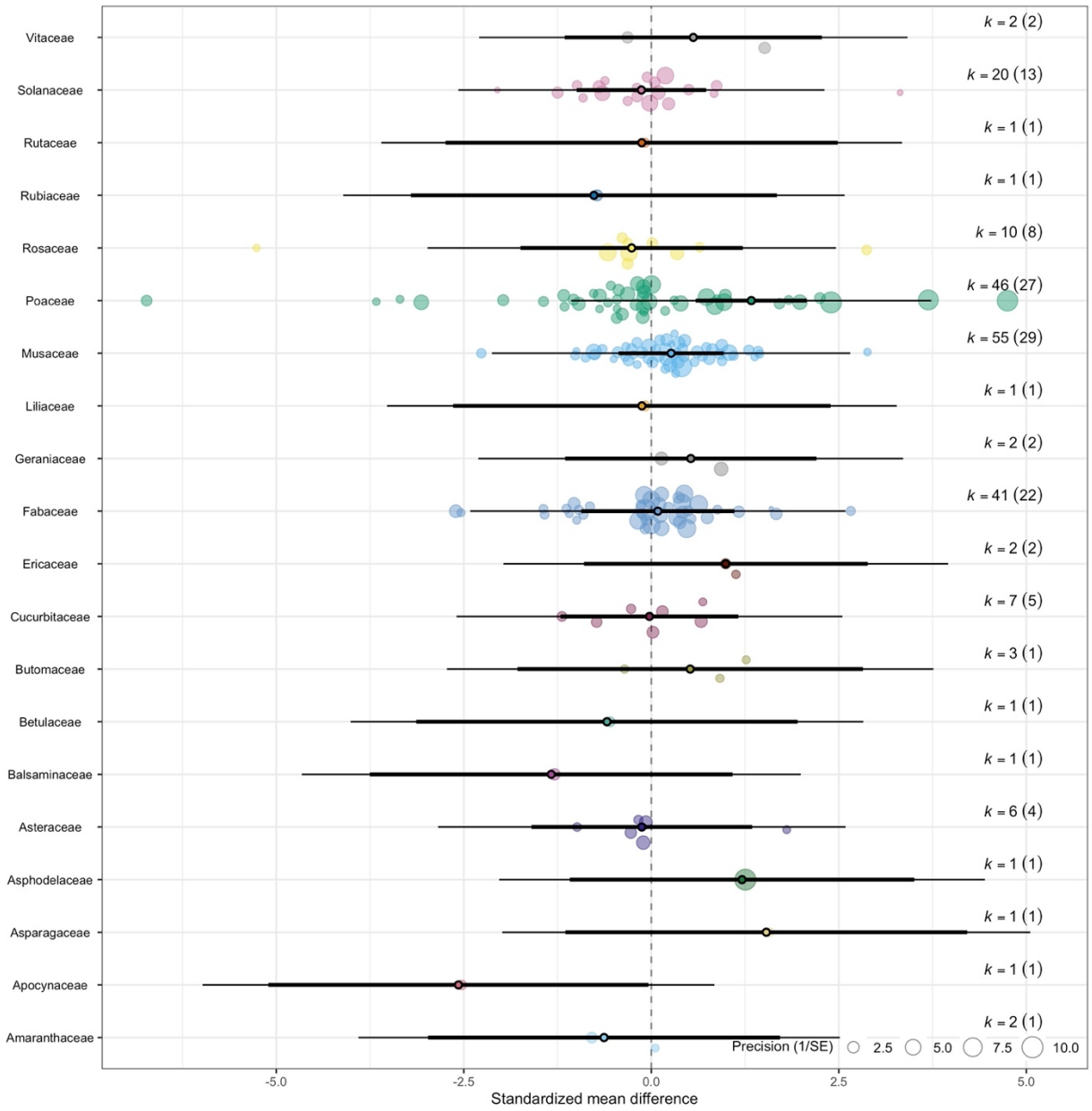


Figure 5 – An orchard plot showing the distribution of effect sizes across polyploid types (autopolyploid vs. allopolyploid). *k* is the number of effect sizes, and numbers in parentheses are the number of studies. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles). Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. Effect sizes with polyploid types "Both" and "Unknown" are not shown.

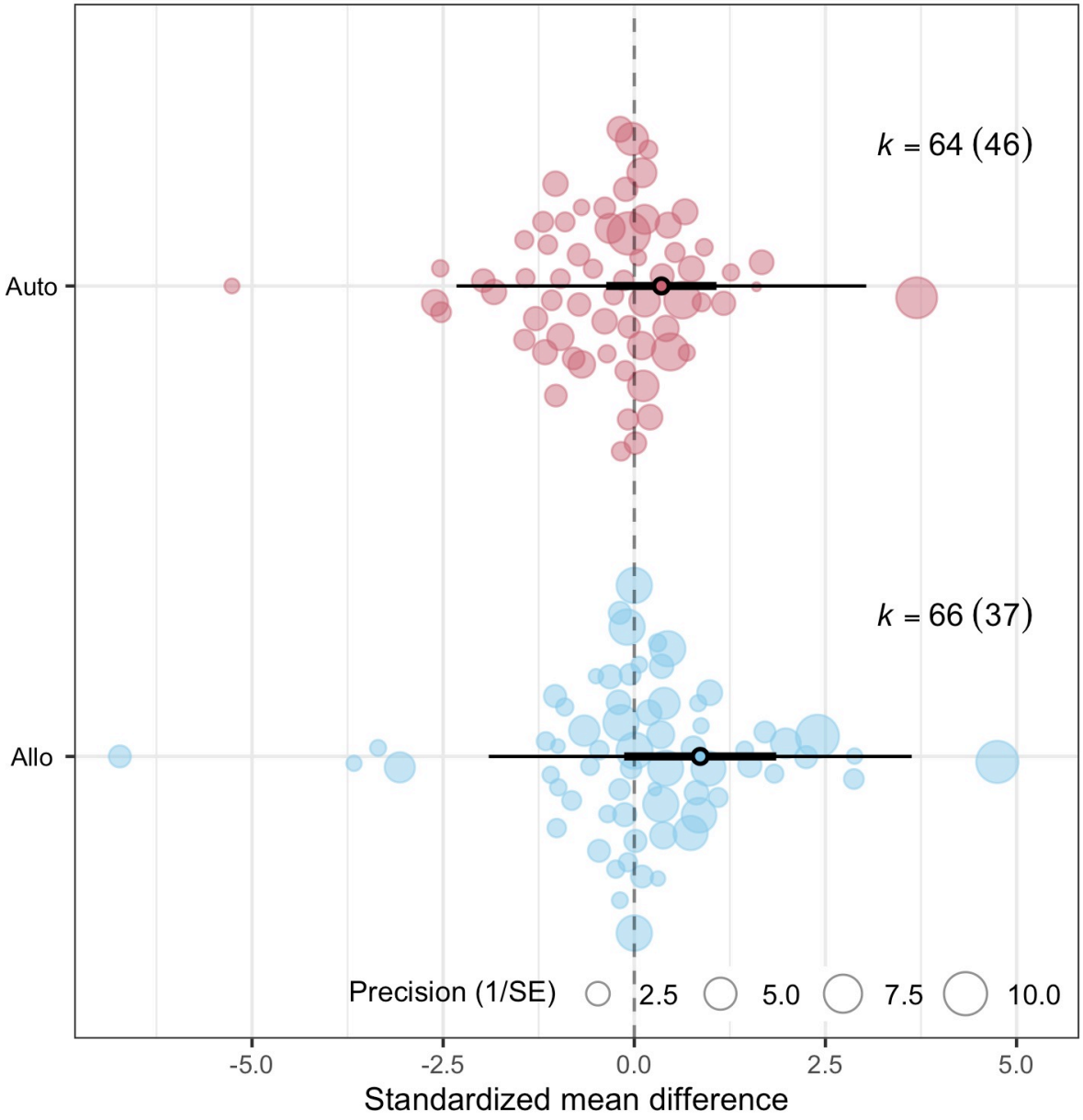


Figure 6 – An orchard plot showing the distribution of effect sizes across combinations of polyploid types ("Auto" for autopolyploids and "Allo" for allopolyploids) and pathogen lifestyles. *k* is the number of effect sizes, and numbers in parentheses are the number of studies. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles) while prediction intervals are shown with thinner lines. Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. Effect sizes for groups with polyploid types or pathogen lifestyles labeled "Both" or "Unknown" are not shown.

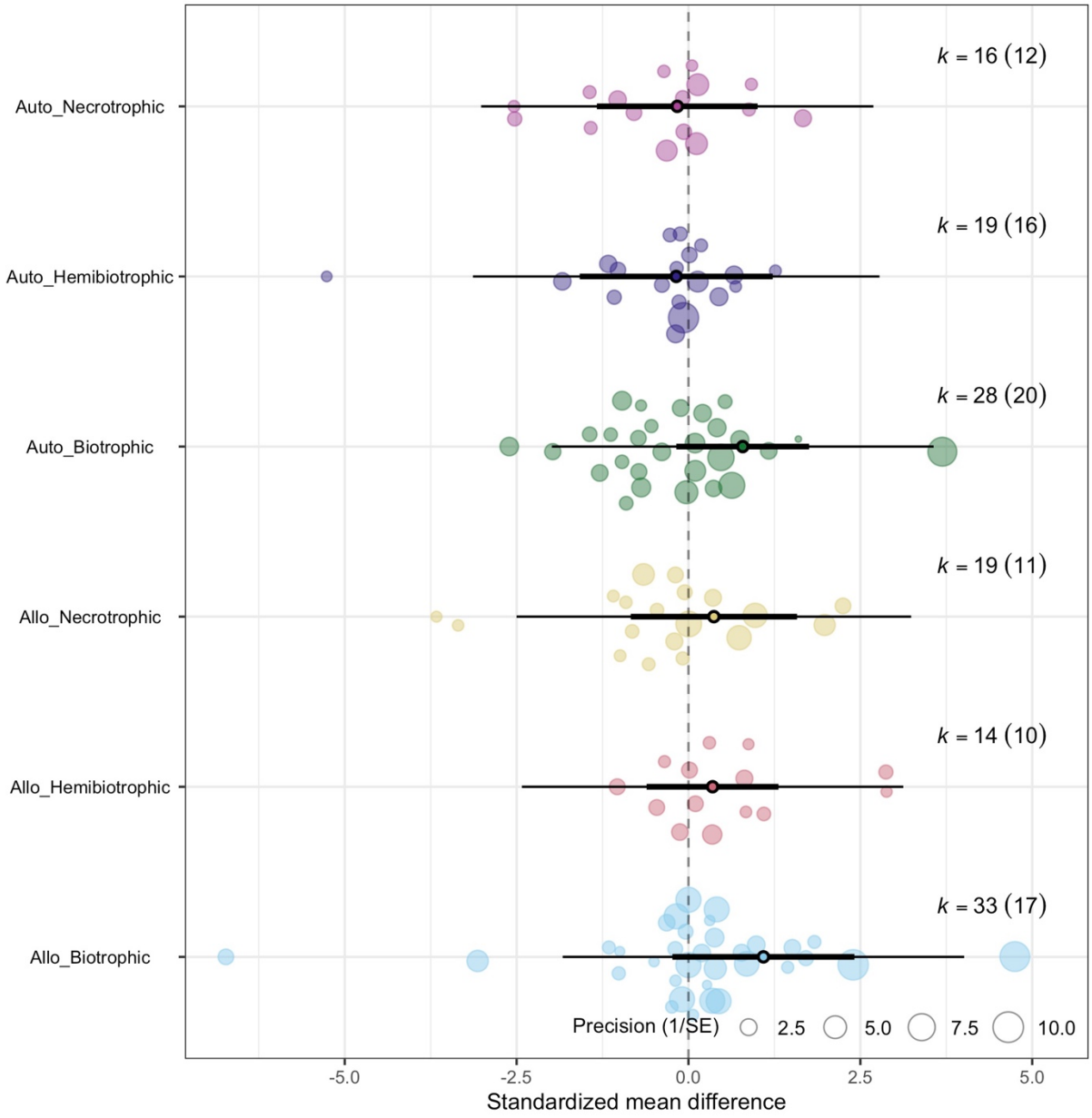


Figure 7 – An orchard plot showing the distribution of effect sizes across combinations of whether polyploids were labeled as autopolyploid ("Auto") or allopolyploid ("Allo") as well as synthetic ("Synth") or established ("Est"). All other effect sizes, where these designations were unable to be made with certainty, fall into the "Other" category. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles) while prediction intervals are shown with thinner lines. Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. Synthetic autopolyploids show the greatest resistance relative to diploids.

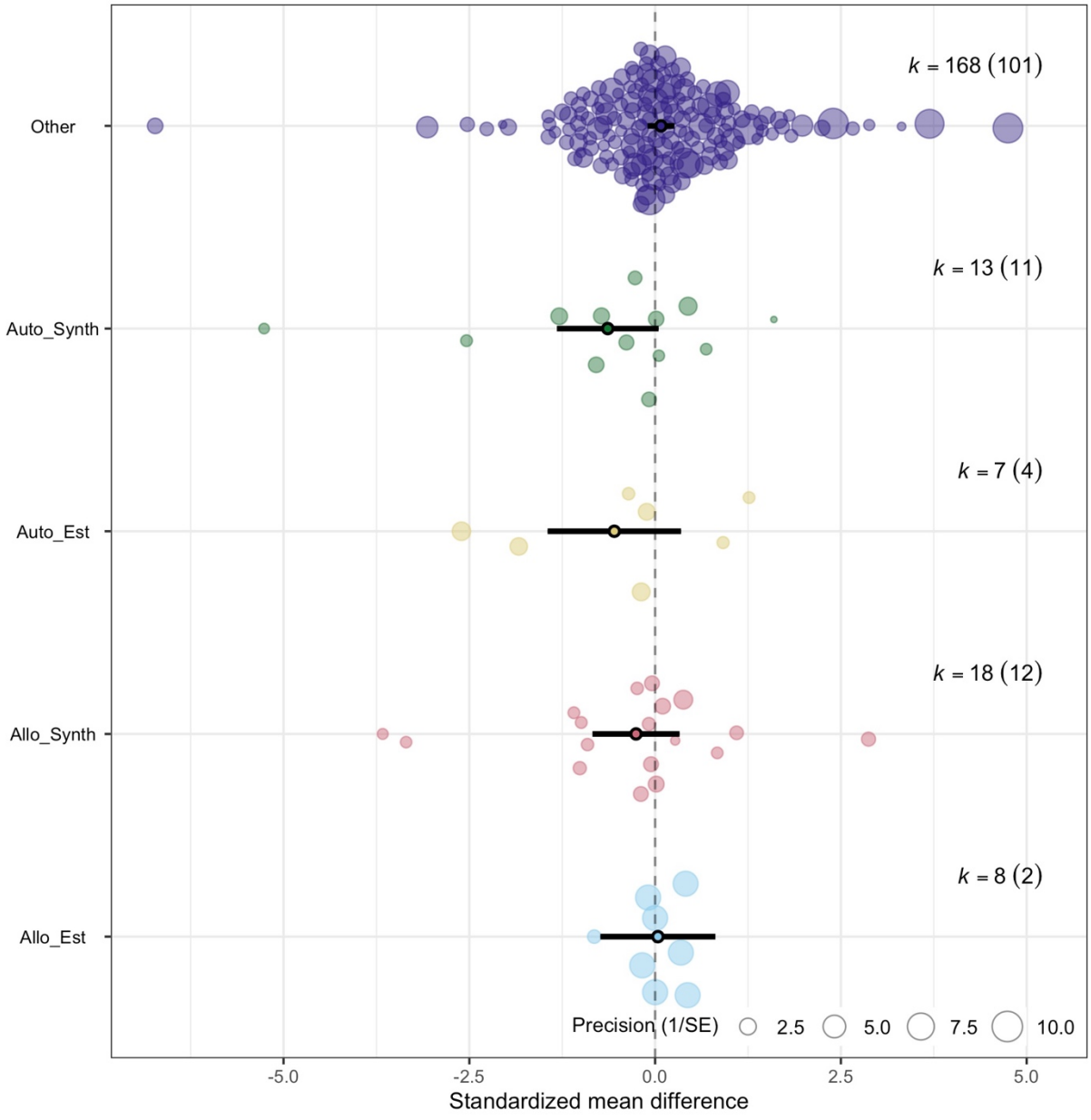


Figure 8 – An orchard plot showing the distribution of effect sizes across pathogen types. *k* is the number of effect sizes, and numbers in parentheses are the number of studies. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles) while prediction intervals are shown with thinner lines. Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. Diploids outperform polyploids in resistance to pathogens that are fungi or nematodes.

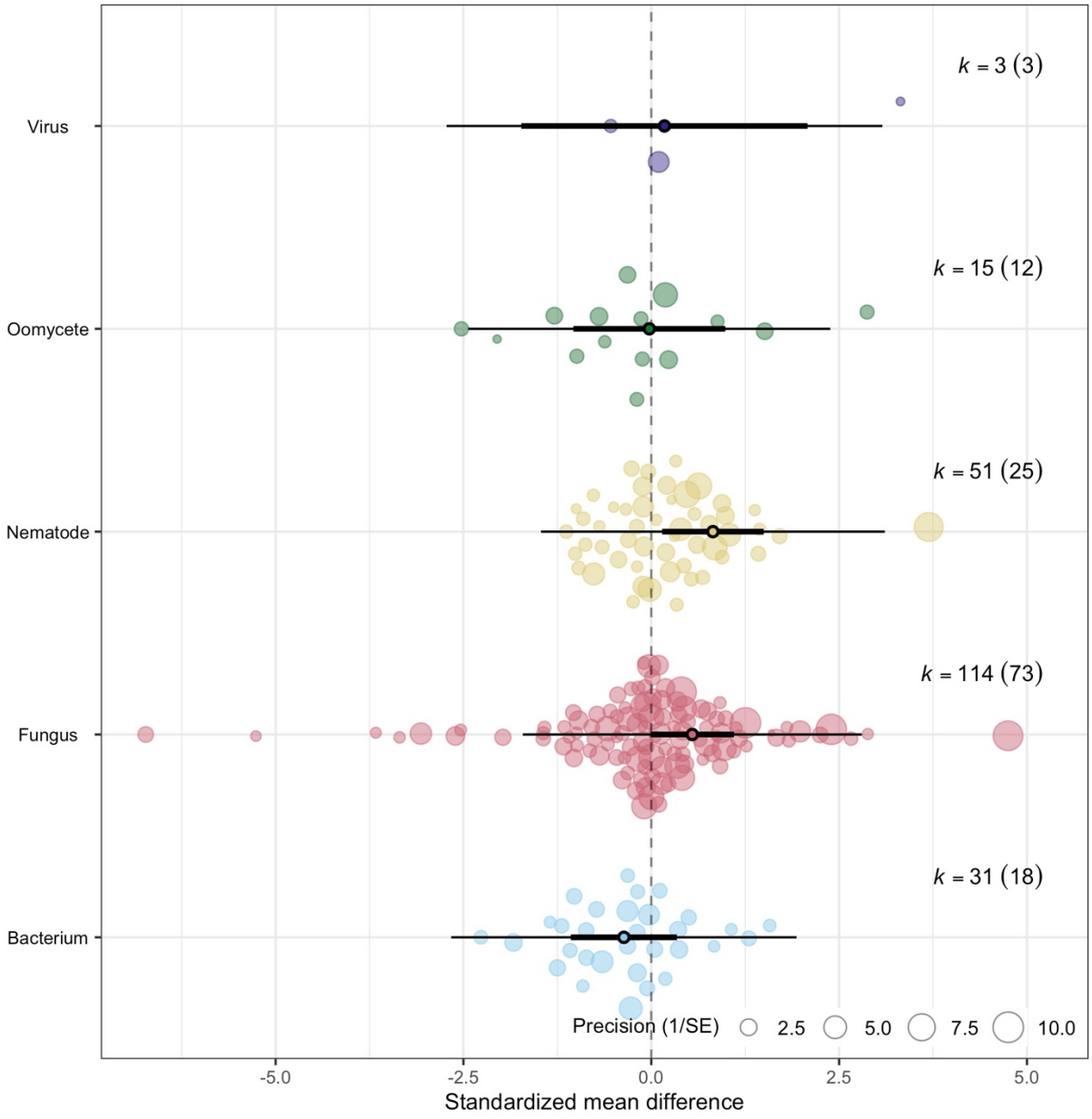


Figure 9 – Funnel plot displaying the relationship between standardized mean difference (SMD; Hedges's g) values and their respective standard errors. A trim-and-fill analysis for funnel plot asymmetry shows no filled "missing" studies (white dots).

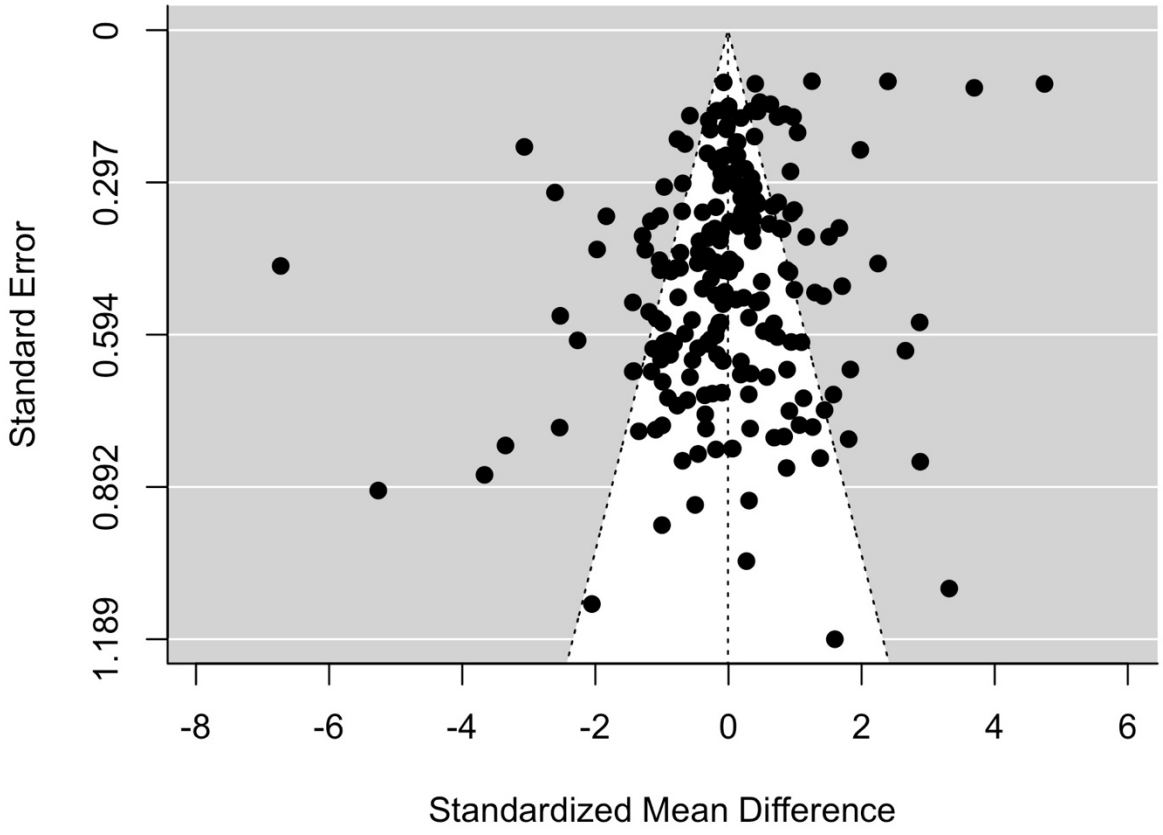


Figure 10 – Orchard plot displaying the clustering of standardized mean difference values for effect sizes in our systematic search ("In") and outside of it ("Not"). k is the number of effect sizes, and numbers in parentheses are the number of studies. 95% confidence intervals are displayed as bold lines around the overall estimates (bold circles) while prediction intervals are shown with thinner lines. Positive standardized mean difference values indicate greater pathogen resistance in diploids than in polyploids. The overall point estimates show significant difference, but the effect sizes cluster around very similar values, suggesting that this difference is driven by influential outlier effect sizes for those outside our Boolean search.

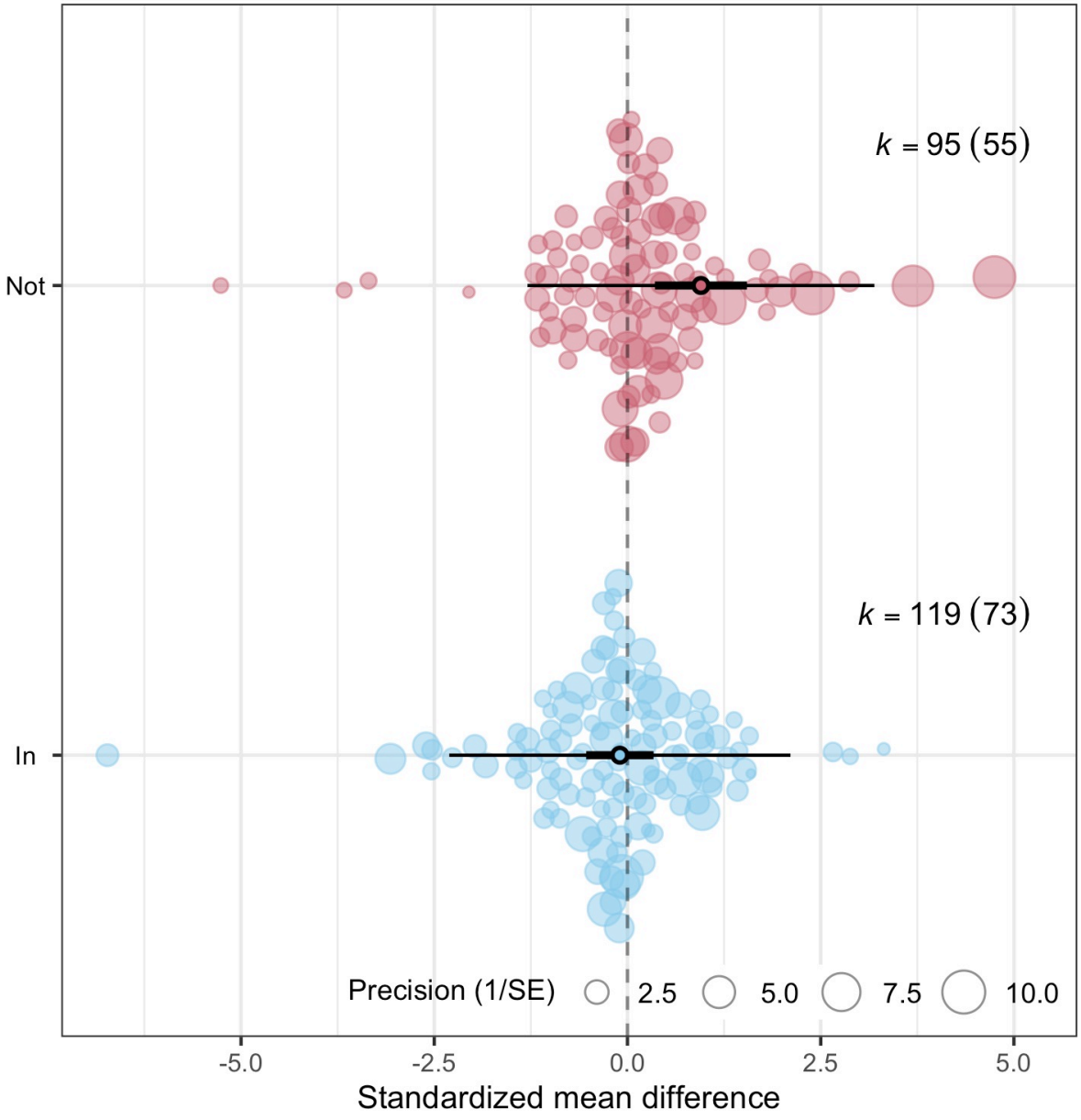


Table 1 – Information about individual Google Scholar searches during our literature search.

Query Terms	**Number of Papers Returned**	**Number Included in Meta-analysis**
"diploid*" + "**triploid***" + inoculat* + (disease* OR nematode* OR wilt* OR rot* OR fungus* OR oomyc* OR rust* OR smut* OR virus* OR bacteri*) -transfer* + (plant* OR crop* OR tree* OR shrub* OR herb*) -graft* -drought* -QTL* -protein* -radiation	434	19
tetraploid*	822	52
pentaploid*	32	2
hexaploid*	248	0
octaploid*	11	0
octoploid*	50	0
decaploid*	5	0

Table 2 – Description of data obtained from each paper included in our meta-analysis. The "Group" column denotes whether the paper came from our systematic search ("0") or from other sources ("1"). The "Synth or Estab." column denotes whether the polyploids in the effect size were synthetic ("Synth"), naturally established ("Est"), or if there was too little information to conclude one way or the other ("Unsure").

Reference	Group	Family	Polyploid type	Auto or allo	Cultivation status	Synth or Estab.
Abdelhalim et al. 2016	1	Poaceae	Tetraploid	Both	Cultivated	Unsure
Abdelhalim et al. 2016	1	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Alam & Gustafson 1988	1	Poaceae	NA	Allo	NA	Unsure
Arrivillaga et al. 2004	0	Fabaceae	Tetraploid	Auto	Cultivated	Synth
Babiker et al. 2018	0	Ericaceae	Tetraploid	NA	Both	Unsure
Barekye et al. 2009	1	Musaceae	Tetraploid	Auto	NA	Synth
Bekal et al. 1998	1	Poaceae	Hexaploid	Allo	Both	Unsure
Bekal et al. 1998	1	Poaceae	Tetraploid	Allo	Both	Unsure
Bekal et al. 1998	1	Poaceae	Hexaploid	Allo	Both	Unsure
Bekal et al. 1998	1	Poaceae	Tetraploid	Allo	Both	Unsure
Blythe et al. 2015	1	Asphodelaceae	Tetraploid	NA	Cultivated	Unsure
Bon et al. 2020	0	Fabaceae	Triploid	Allo	Cultivated	Synth
Bon et al. 2020	0	Fabaceae	Triploid	Allo	Cultivated	Synth
Borner et al. 2006	1	Poaceae	Tetraploid	Allo	Both	Unsure
Borner et al. 2006	1	Poaceae	Hexaploid	Allo	Both	Unsure
Bradshaw et al. 2021	1	Asteraceae	Tetraploid	NA	Wild	Est
Bradshaw et al. 2021	1	Asteraceae	Tetraploid	Both	Wild	Est
Bradshaw et al. 2021	1	Asteraceae	Hexaploid	Both	Wild	Est
Burdon & Marshall 1981	0	Fabaceae	Tetraploid	Auto	Wild	Est
Busch & Smith 1981	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Busey et al. 1993	1	Poaceae	Various_polyploids	NA	Cultivated	Unsure
Carbajal et al. 2021	1	Poaceae	Triploid	Auto	Both	Unsure
Carputo et al. 1997	0	Solanaceae	Tetraploid	Allo	Cultivated	Synth
Carputo et al. 1997	0	Solanaceae	Tetraploid	Allo	Cultivated	Synth
Celebi et al. 1998	0	Solanaceae	Tetraploid	NA	Cultivated	Unsure
Cheo & Beaupre 1981	0	Asteraceae	Tetraploid	NA	Cultivated	Unsure
Chung et al. 2011	1	Solanaceae	Tetraploid	NA	Wild	Est
Chung et al. 2011	1	Solanaceae	Hexaploid	NA	Wild	Est
Costa et al. 2008	0	Musaceae	Triploid	Auto	Cultivated	Unsure
Cotrut et al. 2013	0	Actinidiaceae	Tetraploid	Auto	NA	Unsure
Cotrut et al. 2013	0	Actinidiaceae	Hexaploid	Auto	NA	Unsure
Craenen et al. 1997	0	Musaceae	Triploid	Allo	Cultivated	Synth
Craenen et al. 1997	0	Musaceae	Tetraploid	Allo	Cultivated	Synth
Das et al. 2010	0	Musaceae	Tetraploid	Allo	Cultivated	Synth
Das et al. 2010	0	Musaceae	Triploid	Both	Cultivated	Synth
Das et al. 2013	1	Musaceae	Tetraploid	Allo	Cultivated	Synth
Das et al. 2013	1	Musaceae	Triploid	Allo	Cultivated	Synth
Das et al. 2013	1	Musaceae	Pentaploid	Allo	Cultivated	Synth
Das et al. 2014 (a)	0	Musaceae	Tetraploid	Allo	Cultivated	Synth
Das et al. 2014 (b)	0	Musaceae	Tetraploid	Both	Cultivated	Synth
Das et al. 2014 (b)	0	Musaceae	Triploid	Both	Cultivated	Synth
Datson et al. 2015	0	Actinidiaceae	Tetraploid	NA	NA	Unsure
Datson et al. 2015	0	Actinidiaceae	Hexaploid	NA	NA	Unsure

De Matos et al. 2009	1	Musaceae	Tetraploid	Allo	Cultivated	Synth
Devi et al. 2021	0	Musaceae	Triploid	Both	Both	Unsure
Dijkstra 1964	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Dijkstra 1964	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Dochez et al. 2006	0	Musaceae	Triploid	Allo	Both	Unsure
Dochez et al. 2013	0	Musaceae	Triploid	Both	Cultivated	Unsure
Dochez et al. 2013	0	Musaceae	Triploid	Both	Cultivated	Unsure
Duan et al. 2021	0	Solanaceae	Tetraploid	NA	Both	Unsure
Duan et al. 2021	0	Solanaceae	Hexaploid	NA	Both	Unsure
Ehlenfeldt & Stretch 2001	1	Ericaceae	Tetraploid	NA	Both	Unsure
Felber 1987	1	Poaceae	Tetraploid	Auto	Wild	Est
Fock et al. 2005	1	Solanaceae	Tetraploid	Allo	Cultivated	Synth
Fogain 2000	0	Musaceae	Triploid	Both	Both	Unsure
Franco et al. 2015	1	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Franco et al. 2015	1	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Geise 1957	0	Fabaceae	Tetraploid	Auto	Both	Unsure
Giblin-Davis et al. 1995	1	Poaceae	Various	Auto	Cultivated	Unsure
Goncalves et al. 2019	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Goncalves et al. 2019	0	Musaceae	Triploid	Allo	Cultivated	Unsure
Gooding et al. 1981	1	Rosaceae	Octoploid	Allo	Cultivated	Synth
Green 1959	0	Cucurbitaceae	Tetraploid	Auto	Cultivated	Synth
Green 1959	0	Cucurbitaceae	Triploid	Auto	Cultivated	Synth
Gultyaeva et al. 2016	0	Poaceae	Tetraploid	Allo	Both	Unsure
Gultyaeva et al. 2016	0	Poaceae	Hexaploid	Allo	Both	Unsure
Gunavathi 2000	0	Musaceae	Triploid	Both	Cultivated	Unsure
Gunavathi 2000	0	Musaceae	Triploid	Both	Cultivated	Unsure
Gunter & Egel 2012	0	Cucurbitaceae	Triploid	Auto	Cultivated	Unsure
Hadi et al. 2012	0	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Harding 1971	1	Poaceae	Tetraploid	Allo	Cultivated	Synth
Harding 1971	1	Poaceae	Hexaploid	Allo	Cultivated	Synth
Harms et al. 2020	1	Butomaceae	Triploid	Auto	Wild	Est
Harms et al. 2020	1	Butomaceae	Triploid	Auto	Wild	Est
Harms et al. 2020	1	Butomaceae	Triploid	Auto	Wild	Est
Hartman et al. 2000	1	Fabaceae	Tetraploid	Allo	Both	Est
Hecker & Ruppel 1976	1	Amaranthaceae	Tetraploid	Auto	Cultivated	Synth
Hecker & Ruppel 1976	1	Amaranthaceae	Triploid	Auto	Cultivated	Synth
Henderson & Jenkins 1977	1	Cucurbitaceae	Tetraploid	Both	Cultivated	Synth
Henderson & Jenkins 1977	1	Cucurbitaceae	Triploid	Both	Cultivated	Synth
Hias et al. 2018	1	Rosaceae	Tetraploid	Auto	Cultivated	Synth
Irwin 1981	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Irwin et al. 1997	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Jacob et al. 2010	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Jansky et al. 2006	1	Solanaceae	Tetraploid	NA	Wild	Est
Jansky et al. 2006	1	Solanaceae	Hexaploid	NA	Wild	Est
Julier et al. 1996	0	Fabaceae	Tetraploid	Both	Both	Unsure
Julier et al. 1996	0	Fabaceae	Tetraploid	Both	Both	Unsure
Julier et al. 1996	0	Fabaceae	Tetraploid	Both	Both	Unsure
Julier et al. 1996	0	Fabaceae	Tetraploid	Both	Both	Unsure

Khiutti et al. 2012	1	Solanaceae	Tetraploid	Auto	Cultivated	Unsure
Khiutti et al. 2015	1	Solanaceae	Hexaploid	NA	Wild	Est
Khiutti et al. 2015	1	Solanaceae	Tetraploid	NA	Wild	Est
Kono et al. 2014	0	Vitaceae	Tetraploid	Allo	Cultivated	Unsure
Kriel et al. 1995	0	Solanaceae	Tetraploid	Both	Both	Unsure
Kulkarni & Ravindra 1988	0	Apocynaceae	Tetraploid	Auto	Cultivated	Unsure
Kumar et al. 2009	1	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Kumar et al. 2009	1	Musaceae	Triploid	Both	Cultivated	Unsure
Kumar et al. 2009	1	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Kumar et al. 2009	1	Musaceae	Triploid	Both	Cultivated	Unsure
Lamari & Bernier 1989 (a)	0	Poaceae	Tetraploid	Allo	Cultivated	Unsure
Lamari & Bernier 1989 (a)	0	Poaceae	Hexaploid	Allo	Cultivated	Unsure
Lamari & Bernier 1989 (a)	0	Poaceae	Octoploid	Allo	Cultivated	Unsure
Lamari & Bernier 1989 (b)	0	Poaceae	Hexaploid	Allo	Cultivated	Unsure
Lamari & Bernier 1989 (b)	0	Poaceae	Tetraploid	Allo	Cultivated	Unsure
Levinson et al. 2021	1	Fabaceae	Tetraploid	Allo	Both	Synth
Limantseva et al. 2014	1	Solanaceae	Tetraploid	NA	Cultivated	Unsure
Mazzafera et al. 1993	1	Rubiaceae	Tetraploid	Auto	NA	Synth
Mikaliuniene et al. 2015	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Mikaliuniene et al. 2015	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Mudonyi et al. 2019	0	Musaceae	Triploid	Both	Both	Unsure
Mudonyi et al. 2019	0	Musaceae	Tetraploid	Both	Both	Unsure
Nakato et al. 2019	0	Musaceae	Triploid	Both	Both	Unsure
Nakato et al. 2019	0	Musaceae	Tetraploid	Both	Both	Unsure
Nardozza et al. 2015	0	Actinidiaceae	Tetraploid	NA	NA	Unsure
Nardozza et al. 2015	0	Actinidiaceae	Hexaploid	NA	NA	Unsure
Naydenova & Aleksieva 2017	1	Fabaceae	Tetraploid	Auto	NA	Unsure
Nguyet et al. 2002	0	Musaceae	Triploid	Both	Cultivated	Unsure
Ohberg et al. 2005	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Oliveira et al. 2018	1	Poaceae	Tetraploid	Allo	Cultivated	Unsure
Pair & Bruton 1998	1	Cucurbitaceae	Triploid	Auto	Cultivated	Unsure
Pang 2010	1	Poaceae	Tetraploid	NA	Both	Unsure
Pang 2010	1	Poaceae	Triploid	NA	Both	Unsure
Pang et al. 2011	0	Poaceae	Tetraploid	NA	Both	Unsure
Pang et al. 2011	0	Poaceae	Triploid	NA	Both	Unsure
Pang et al. 2011	0	Poaceae	Hexaploid	NA	Both	Unsure
Paul & Freudenstein 1989	0	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Pederson & Windham 1989	1	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Pederson & Windham 1989	1	Fabaceae	Hexaploid	Auto	Cultivated	Unsure
Perez et al. 2014	1	Solanaceae	Tetraploid	NA	Cultivated	Unsure
Perez et al. 2014	1	Solanaceae	Triploid	NA	Cultivated	Unsure
Pinochet et al. 1998	0	Musaceae	Triploid	Both	Cultivated	Unsure

Pinochet et al. 1998	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Pinochet et al. 1998	0	Musaceae	Triploid	Both	Cultivated	Unsure
Pinochet et al. 1998	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Pinochet et al. 1998	0	Musaceae	Triploid	Both	Cultivated	Unsure
Pinochet et al. 1998	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Podwyszynska et al. 2021	1	Rosaceae	Tetraploid	Auto	Cultivated	Synth
Poteri et al. 1997	1	Betulaceae	Tetraploid	NA	Both	Unsure
Prasad et al. 2009	0	Asteraceae	Various	NA	Wild	Est
Queneherve et al. 2009	0	Musaceae	Triploid	Both	Both	Unsure
Queneherve et al. 2009	0	Musaceae	Tetraploid	Allo	Both	Unsure
Ray et al. 1995	0	Asteraceae	NA	Auto	Cultivated	Unsure
Reboucas et al. 2018	1	Musaceae	Triploid	Both	Cultivated	Unsure
Rhodes et al. 1996	0	Cucurbitaceae	Triploid	Auto	Cultivated	Unsure
Ribeiro et al. 2018	0	Musaceae	Triploid	Both	Cultivated	Unsure
Ribeiro et al. 2018	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Rothleutner 2012	0	Rosaceae	Tetraploid	NA	NA	Unsure
Rothleutner 2012	0	Rosaceae	Triploid	NA	NA	Unsure
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schoen et al. 1992	1	Fabaceae	Tetraploid	Allo	Wild	Est
Schubiger et al. 2010	0	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Schubiger et al. 2010	0	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Schubiger et al. 2010	0	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Schuster 1991	1	Poaceae	Tetraploid	Allo	Cultivated	Unsure
Seenivasan 2017	0	Musaceae	Tetraploid	Both	Cultivated	Unsure
Seenivasan 2017	0	Musaceae	Triploid	Both	Cultivated	Unsure
Singh et al. 2006	1	Poaceae	Tetraploid	Allo	Both	Unsure
Singh et al. 2006	1	Poaceae	Hexaploid	Allo	Both	Unsure
Ssekiwoko et al. 2006	0	Musaceae	Triploid	Both	Cultivated	Unsure
Ssekiwoko et al. 2006	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Stoffelen et al. 2000	0	Musaceae	Triploid	Both	Cultivated	Unsure
Stoffelen et al. 2000	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Stoffelen et al. 2000	0	Musaceae	Triploid	Both	Cultivated	Unsure
Stoffelen et al. 2000	0	Musaceae	Tetraploid	Allo	Cultivated	Unsure
Stover & Waite 1960	0	Musaceae	Triploid	Both	Cultivated	Unsure
Svara et al. 2021	1	Rosaceae	Tetraploid	Auto	Cultivated	Synth
Swiezynski et al. 1991	0	Solanaceae	Tetraploid	NA	Cultivated	Unsure
Thangavelu et al. 2021	0	Musaceae	Triploid	Both	Cultivated	Unsure
Tofte 1990	0	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Tomlinson et al. 1987	0	Musaceae	Triploid	Both	Cultivated	Unsure
Tusa & Del Bosco et al. 2000	0	Rutaceae	Tetraploid	Allo	Cultivated	Synth
Uchneat 1997	0	Geraniaceae	Tetraploid	Both	Cultivated	Unsure
Uchneat et al. 1999	0	Geraniaceae	Tetraploid	Both	Cultivated	Unsure
Van Tuyl 1982	0	Asparagaceae	Triploid	NA	Both	Unsure
Vestad 1960	0	Fabaceae	Tetraploid	Auto	Cultivated	Synth
Viaene et al. 2003	0	Musaceae	Tetraploid	Both	Cultivated	Unsure

Viaene et al. 2003	0	Musaceae	Triploid	Both	Cultivated	Unsure
Vincelli et al. 2008	0	Poaceae	Tetraploid	NA	Cultivated	Unsure
Vining et al. 2015	1	Rosaceae	Octoploid	Allo	Both	Unsure
Vleugels et al. 2013	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Vleugels et al. 2013	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Vleugels et al. 2013	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Vymyslicky et al. 2012	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Vymyslicky et al. 2012	1	Fabaceae	Tetraploid	Auto	Both	Unsure
Wang et al. 2018	0	Balsaminaceae	Tetraploid	Auto	Cultivated	Synth
Wang et al. 2020	0	Actinidiaceae	Tetraploid	NA	Wild	Est
Wang et al. 2020	0	Actinidiaceae	Hexaploid	NA	Wild	Est
Wang et al. 2020	0	Actinidiaceae	Tetraploid	Auto	Wild	Est
Wang et al. 2020	0	Actinidiaceae	Hexaploid	Auto	Wild	Est
Wang et al. 2021	1	Liliaceae	Tetraploid	Auto	Cultivated	Synth
Whitaker & Hokanson 2009	1	Rosaceae	Tetraploid	NA	Cultivated	Unsure
Wilkins 1973	0	Poaceae	Tetraploid	NA	Cultivated	Unsure
Wise & Gobelman-Werner 1993	1	Poaceae	Hexaploid	Auto	Cultivated	Unsure
Wiwart et al. 2016	1	Poaceae	Tetraploid	Allo	Cultivated	Unsure
Wiwart et al. 2016	1	Poaceae	Hexaploid	Allo	Cultivated	Unsure
Yeates et al. 1973	1	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Yeates et al. 1973	1	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Yeates et al. 1973	1	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Yeates et al. 1973	1	Fabaceae	Tetraploid	Auto	Cultivated	Unsure
Yong-Fang et al. 1997	1	Poaceae	Tetraploid	Allo	Both	Unsure
Yong-Fang et al. 1997	1	Poaceae	Hexaploid	Allo	Both	Unsure
York 1989	1	Poaceae	Tetraploid	Auto	Cultivated	Unsure
Yun et al. 2001	0	Vitaceae	Tetraploid	Allo	Cultivated	Unsure
Zimnoch-Guzowska et al. 1999	0	Solanaceae	Hexaploid	Allo	Cultivated	Unsure
Zimnoch-Guzowska et al. 1999	0	Solanaceae	Pentaploid	Allo	Cultivated	Unsure
Zlesak et al. 2010	0	Rosaceae	Triploid	NA	Cultivated	Unsure
Zlesak et al. 2010	0	Rosaceae	Tetraploid	NA	Cultivated	Unsure

Table 3 – Further description of data obtained from each paper included in our meta-analysis. The "Assessment" column denotes the metric used in the effect size to measure pathogen resistance, and the "Notes" column contains information about where in the paper to find the data used to calculate the effect size as well as information about how it was calculated.

Reference	Pathogen type	Pathogen lifestyle	Assessment	Notes
Abdelhalim et al. 2016	Fungus	Hemibiotrophic	Dry_weight	See Table 3; averaged across weeks & experiments; higher = more resistant
Abdelhalim et al. 2016	Fungus	Hemibiotrophic	Dry_weight	See Table 3; averaged across weeks & experiments; higher = more resistant
Alam & Gustafson 1988	Fungus	Necrotrophic	Score	See Tables 1 & 2; R=3, MR=2, MS=1, S=0
Arrivillaga et al. 2004	Fungus	Biotrophic	Score	See Table 1; higher = less resistant
Babiker et al. 2018	Fungus	Biotrophic	Score	See Tables 1 & 2; ignored SE in Table 2
Barekye et al. 2009	Fungus	Hemibiotrophic	Area Under Disease Progress Curve (AUDPC)	See Table 4
Bekal et al. 1998	Nematode	Biotrophic	Number_per_plant	See Tables 3 & 5; used all values in table and ignored stdevs
Bekal et al. 1998	Nematode	Biotrophic	Number_per_plant	See Tables 3 & 4; used all values in table and ignored stdevs
Bekal et al. 1998	Nematode	Biotrophic	Number_per_plant	See Tables 3 & 5; used all values in table and ignored stdevs
Bekal et al. 1998	Nematode	Biotrophic	Number_per_plant	See Tables 3 & 4; used all values in table and ignored stdevs
Blythe et al. 2015	Fungus	Biotrophic	Score	See Table 1
Bon et al. 2020	Fungus	Necrotrophic	Score	See Table 6
Bon et al. 2020	Fungus	Necrotrophic	Score	See Table 6
Borner et al. 2006	Fungus	Biotrophic	Score	WebPlotDigitizer used on bar graph in Figure 10
Borner et al. 2006	Fungus	Biotrophic	Score	WebPlotDigitizer used on bar graph in Figure 10
Bradshaw et al. 2021	Fungus	Biotrophic	AUDPC	Data provided by Bradshaw et al.; ignored variable ploidy species
Bradshaw et al. 2021	Fungus	Biotrophic	AUDPC	Data provided by Bradshaw et al.; ignored variable ploidy species
Bradshaw et al. 2021	Fungus	Biotrophic	AUDPC	Data provided by Bradshaw et al.; ignored variable ploidy species
Burdon & Marshall 1981	Fungus	Biotrophic	Score	See Table 1
Busch & Smith 1981	Fungus	Hemibiotrophic	Score	See Tables 4 & 5; only used 2n=16 diploids and 2n=32

				tetraploids; used weighted average & stdev
Busey et al. 1993	Nematode	Biotrophic	Number_per_plant	See Table 1; used absolute number of nematodes per pot
Carbajal et al. 2021	Fungus	Hemibiotrophic	AUDPC	See Table 3 sqrt AUDPC values; averaged across inoculum values (GA, LW, & SRS); ignored all but diploid and triploid
Carputo et al. 1997	Bacterium	Necrotrophic	Percent_diseased	See Table 1; R=2, I=1, S=0; included values for both Eca & Ecc
Carputo et al. 1997	Bacterium	Necrotrophic	Percent_diseased	See Table 1; R=2, I=1, S=0; included values for both Eca & Ecc
Celebi et al. 1998	Virus	Biotrophic	Percent_diseased	See Table 2 (no variation in Table 1); used Systematic Infection % infected column; ignored pentaploid
Cheo & Beaupre 1981	Fungus	Hemibiotrophic	Wilt index	See Table 2; ignored standard deviations; used 13-27 day values for 2N & 27-40 day values for 4N
Chung et al. 2011	Bacterium	Necrotrophic	Score	WebPlotDigitizer performed on Figure 3
Chung et al. 2011	Bacterium	Necrotrophic	Score	WebPlotDigitizer performed on Figure 3
Costa et al. 2008	Nematode	Biotrophic	Reproductive factor	See Table 4 RF columns (used only 12 rows for nematode populations)
Cotrut et al. 2013	Bacterium	Hemibiotrophic	Spots per cm^2	See Table 3 (plants without injury); used values from all 3 no. spots per cm^2 columns
Cotrut et al. 2013	Bacterium	Hemibiotrophic	Spots per cm^2	See Table 3 (plants without injury); used values from all 3 no. spots per cm^2 columns
Craenen et al. 1997	Fungus	Hemibiotrophic	Score	See Table 1 Black Sigatoka reaction column (S=1, LS=2, PR=3, HR=4, ER=5)
Craenen et al. 1997	Fungus	Hemibiotrophic	Score	See Table 1 Black Sigatoka reaction column (S=1, LS=2, PR=3, HR=4, ER=5)
Das et al. 2010	Nematode	Biotrophic	Root lesion index	See Table 1 Root Lesion Index (%)
Das et al. 2010	Nematode	Biotrophic	Root lesion index	See Table 1 Root Lesion Index (%)
Das et al. 2013	Nematode	Biotrophic	Score	See Table 1; averaged root lesion index and corm grade for each entry
Das et al. 2013	Nematode	Biotrophic	Score	See Table 1; averaged root lesion index and corm grade for each entry
Das et al. 2013	Nematode	Biotrophic	Score	See Table 1; averaged root lesion index and corm grade for each entry

Das et al. 2014 (a)	Nematode	Biotrophic	Root lesion index	See Table 1
Das et al. 2014 (b)	Fungus	Hemibiotrophic	Score	See Table 1 Wilt Score for "Reaction to FOC"
Das et al. 2014 (b)	Fungus	Hemibiotrophic	Score	See Table 1 Wilt Score for "Reaction to FOC"
Datson et al. 2015	Bacterium	Hemibiotrophic	Number_infected	Data extracted from Figure 1 with WebPlotDigitizer; weighted average
Datson et al. 2015	Bacterium	Hemibiotrophic	Number_infected	Data extracted from Figure 1 with WebPlotDigitizer; weighted average
De Matos et al. 2009	Fungus	Hemibiotrophic	Score	Averaged first and second cycles to get one mean from each table (2&4)
Devi et al. 2021	Nematode	Biotrophic	Percent dead roots	See Tables 2 & 3, Dead Roots (%); ignored reference cultivars
Dijkstra 1964	Fungus	Necrotrophic	Score	Mycelial inoculation; see Table 2
Dijkstra 1964	Fungus	Necrotrophic	Score	Ascopore inoculation; see Table 2
Dochez et al. 2006	Nematode	Biotrophic	Score	See Table 3; used "percentage root necrosis" column; included values across all 4 experiments
Dochez et al. 2013	Nematode	Biotrophic	Percent root necrosis	See Table 3; used % root necrosis values from all 3 nematode populations
Dochez et al. 2013	Nematode	Biotrophic	Percent root necrosis	See Table 3; used % root necrosis values from all 3 nematode populations
Duan et al. 2021	Oomycete	Hemibiotrophic	Score	See Supplemental Table S1; included only species with ploidy values listed in Supplemental Table S5
Duan et al. 2021	Oomycete	Hemibiotrophic	Score	See Supplemental Table S1; included only species with ploidy values listed in Supplemental Table S5
Ehlenfeldt & Stretch 2001	Fungus	Necrotrophic	Number_blighted	See Table 2
Felber 1987	Fungus	Biotrophic	Score	Only used alpinum values; see Table 1
Fock et al. 2005	Bacterium	Hemibiotrophic	Disease index	See Table 1; averaged across Race 1 & 3 disease index scores; ignored BF15 & BP9
Fogain 2000	Nematode	Biotrophic	Count	See Tables 1-3
Franco et al. 2015	Fungus	Biotrophic	Number_infected	See Tables 3 & 4
Franco et al. 2015	Fungus	Biotrophic	Infection Rate	See Tables 3 & 4
Geise 1957	Fungus	Necrotrophic	Score	Only blackstem scores from Table 2 (1-10, 10 being worst infection)
Giblin-Davis et al. 1995	Nematode	Biotrophic	Count	See Table 1; used relative # nematodes per gram column
Goncalves et al. 2019	Fungus	Hemibiotrophic	Score	See Table 2; averaged internal & external disease values for each entry; S (susceptible)

				coded 0, MR (moderately resistant) coded 1, & R (resistant) coded 2
Goncalves et al. 2019	Fungus	Hemibiotrophic	Score	See Table 2; averaged internal & external disease values for each entry; S (susceptible) coded 0, MR (moderately resistant) coded 1, & R (resistant) coded 2
Gooding et al. 1981	Oomycete	Hemibiotrophic	Percent_dead	See Table 4; used % column values under "No. of Deaths"
Green 1959	Fungus	Hemibiotrophic	Number_dead	See Table 16; did not incude diploid control
Green 1959	Fungus	Hemibiotrophic	Number_dead	See Table 16; did not incude diploid control
Gultyaeva et al. 2016	Fungus	Biotrophic	Average Virulence Complexity	See Table 4; converted SE to SD; calculated combined SD values
Gultyaeva et al. 2016	Fungus	Biotrophic	Average Virulence Complexity	See Table 4; converted SE to SD; calculated combined SD values
Gunavathi 2000	Fungus	Hemibiotrophic	Root lesion index	See Table 2
Gunavathi 2000	Nematode	Biotrophic	Population per roots	See Table 3
Gunter & Egel 2012	Fungus	Hemibiotrophic	AUDPC	See Tables 2 & 4; used all 5 AUDPC values for those cultivars listed as "diploid cultivar" & "triploid cultivar" in Table 1
Hadi et al. 2012	Virus	Biotrophic	Percent_infected	See Table 2 Bahiagrass values
Harding 1971	Fungus	Necrotrophic	Number_dead	See Table 1; used % survival after inoculation
Harding 1971	Fungus	Necrotrophic	Number_dead	See Table 1; used % survival after inoculation
Harms et al. 2020	Fungus	Necrotrophic	Lesion Area	See Figure 3A
Harms et al. 2020	Fungus	Hemibiotrophic	Lesion Area	See Figure 3B
Harms et al. 2020	Fungus	Necrotrophic	Lesion Area	See Figure 3C
Hartman et al. 2000	Fungus	Necrotrophic	Number_dead	See Table 2; only used Screen-2 survival values
Hecker & Ruppel 1976	Fungus	Necrotrophic	Score	Took diploid & tetraploid means from Table 1 & averaged across the 2 years; used % healthy column
Hecker & Ruppel 1976	Fungus	Necrotrophic	Score	Used % healthy column in Table 2
Henderson & Jenkins 1977	Fungus	Hemibiotrophic	Score	See Table 1
Henderson & Jenkins 1977	Fungus	Hemibiotrophic	Score	See Table 1
Hias et al. 2018	Fungus	Hemibiotrophic	ng/ng plant DNA	WebPlotDigitizer used on Figure 4
Irwin 1981	Oomycete	Hemibiotrophic	Score	See p. 22 (Table 3); averaged across taproot & propagule DSI values & ignored SE values
Irwin et al. 1997	Oomycete	Hemibiotrophic	Disease severity index	See Table 1; used Mature Root DSI column

Jacob et al. 2010	Fungus	Biotrophic	Number_dead	See Table 1
Jansky et al. 2006	Fungus	Necrotrophic	Score	See Supplemental Table 1; used Mean Score column & averaged per species
Jansky et al. 2006	Fungus	Necrotrophic	Score	See Supplemental Table 1; used Mean Score column & averaged per species
Julier et al. 1996	Fungus	Hemibiotrophic	Score	See Table 2
Julier et al. 1996	Fungus	Hemibiotrophic	Score	See Table 2
Julier et al. 1996	Fungus	Necrotrophic	Score	See Table 2
Julier et al. 1996	Fungus	Necrotrophic	Score	See Table 2
Khiutti et al. 2012	Fungus	Biotrophic	Score	See Table 1
Khiutti et al. 2015	Oomycete	Hemibiotrophic	Score	WebPlotDigitizer used on Figure 2; included values for both trials
Khiutti et al. 2015	Oomycete	Hemibiotrophic	Score	WebPlotDigitizer used on Figure 2; included values for both trials
Kono et al. 2014	Oomycete	Biotrophic	Score	Used WebPlotDigitizer on Fig. 2
Kriel et al. 1995	Bacterium	Hemibiotrophic	Immunofluorescence	See Table 1; >100 input in calculations as 100
Kulkarni & Ravindra 1988	Oomycete	Necrotrophic	Score	See Table 2; used values from all 3 years
Kumar et al. 2009	Fungus	Hemibiotrophic	Score	See Table 1
Kumar et al. 2009	Fungus	Hemibiotrophic	Score	See Table 1
Kumar et al. 2009	Nematode	Biotrophic	Score	See Table 1; averaged pot & field values; R through HS recorded as scores 1-4
Kumar et al. 2009	Nematode	Biotrophic	Score	See Table 1; averaged pot & field values; R through HS recorded as scores 1-4
Lamari & Bernier 1989 (a)	Fungus	Necrotrophic	Score	See Table 1
Lamari & Bernier 1989 (a)	Fungus	Necrotrophic	Score	See Table 1
Lamari & Bernier 1989 (a)	Fungus	Necrotrophic	Score	See Table 1
Lamari & Bernier 1989 (b)	Fungus	Necrotrophic	Score	See Table 1; intermediate values like 1-2 were averaged; chl- and chl+ coded as 1 and 5 respectively
Lamari & Bernier 1989 (b)	Fungus	Necrotrophic	Score	See Table 1; intermediate values like 1-2 were averaged; chl- and chl+ coded as 1 and 5 respectively
Levinson et al. 2021	Fungus	Biotrophic	Score	See Table 3; used susceptibility index per leaf area (IA); combined 2017 & 2020
Limantseva et al. 2014	Nematode	Biotrophic	Score	See Supplementary Table 3; averaged all values in "score of reaction in different replications" column
Mazzafera et al. 1993	Fungus	Biotrophic	Score	Used PDL column in Table 1; ignored hexaploid

Mikaliuniene et al. 2015	Fungus	Necrotrophic	Score	Results from uninfected field experiment (Table 2); only used 2014 results
Mikaliuniene et al. 2015	Fungus	Necrotrophic	Score	Results from infected field experiment (Table 3); only used 2014 results
Mudonyi et al. 2019	Bacterium	Necrotrophic	Disease Index	See Table 3; used disease index 26 dai column
Mudonyi et al. 2019	Bacterium	Necrotrophic	Disease Index	See Table 3; used disease index 26 dai column
Nakato et al. 2019	Bacterium	Necrotrophic	Score	See Table 1; used observed ploidy values when available, expected when not; used disease index column from 1st evaluation
Nakato et al. 2019	Bacterium	Necrotrophic	Score	See Table 1; used observed ploidy values when available, expected when not; used disease index column from 1st evaluation
Nardozza et al. 2015	Bacterium	Hemibiotrophic	Percent removed	WebPlotDigitizer performed on Figure 1; conducted weighted average & Stdev on percent removed using no. Genotypes
Nardozza et al. 2015	Bacterium	Hemibiotrophic	Percent removed	WebPlotDigitizer performed on Figure 1; conducted weighted average & Stdev on percent removed using no. Genotypes
Naydenova & Aleksieva 2017	Fungus	Biotrophic	Score	See Table 2; averaged across August 2014 & 2015 mean reactions
Nguyet et al. 2002	Nematode	Biotrophic	Percent root necrosis	See Tables 1 through 3, Root Necrosis (%)
Ohberg et al. 2005	Fungus	Necrotrophic	Number dead	See Table 2
Oliveira et al. 2018	Fungus	Biotrophic	Proportion resistant	See Table 6
Pair & Bruton 1998	Bacterium	NA	Number infected	See Table 2
Pang 2010	Nematode	Biotrophic	Reproductive factor	See Table 7-2 reproductive factor columns for both trials
Pang 2010	Nematode	Biotrophic	Reproductive factor	See Table 7-2 reproductive factor columns for both trials
Pang et al. 2011	Nematode	Biotrophic	Count	See Table 2; included values from both trial columns
Pang et al. 2011	Nematode	Biotrophic	Count	See Table 2; included values from both trial columns
Pang et al. 2011	Nematode	Biotrophic	Count	See Table 2; included values from both trial columns
Paul & Freudenstein 1989	Bacterium	Necrotrophic	Score	See Table 3; included greenhouse & field trial values in calculations
Pederson & Windham 1989	Nematode	Biotrophic	Score	See Table 1
Pederson & Windham 1989	Nematode	Biotrophic	Score	See Table 1
Perez et al. 2014	Oomycete	Hemibiotrophic	Score	See Table 1
Perez et al. 2014	Oomycete	Hemibiotrophic	Score	See Table 1

Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 2; just used "percentage of galled roots" column
Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 2; just used "percentage of galled roots" column
Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 3; just used "percentage of galled roots" column
Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 3; just used "percentage of galled roots" column
Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 4; just used "root lesion index" column
Pinochet et al. 1998	Nematode	Biotrophic	Score	See Table 4; just used "root lesion index" column
Podwyszynska et al. 2021	Fungus	Hemibiotrophic	Score	See Table 2; averaged across years & each ploidy value
Poteri et al. 1997	Fungus	Biotrophic	AUDPC	See Table 1; averaged means of greenhouse & outdoors values for each species, as well as the 2 different rust sources; ignored pentaploid
Prasad et al. 2009	Fungus	Necrotrophic	Score	See Table 4
Queneherve et al. 2009	Nematode	Biotrophic	Multiplication_rate	Used multiplication rate (MR) column in Tables 2-4
Queneherve et al. 2009	Nematode	Biotrophic	Multiplication_rate	Used multiplication rate (MR) column in Tables 2-4
Ray et al. 1995	Fungus	Hemibiotrophic	Score	See Table 5
Reboucas et al. 2018	Fungus	Hemibiotrophic	Score	See Table 3; ignored tetraploid
Rhodes et al. 1996	Bacterium	Biotrophic	Score	Ignored individual standard deviations; see Table 1, column "Severity"
Ribeiro et al. 2018	Fungus	Hemibiotrophic	Score	See Table 1; ignored ES hybrid
Ribeiro et al. 2018	Fungus	Hemibiotrophic	Score	See Table 1
Rothleutner 2012	Bacterium	Necrotrophic	Percent_diseased	See Tables 4 & 5 on p. 41-44; only included values for species with ploidy values included in thesis
Rothleutner 2012	Bacterium	Necrotrophic	Percent_diseased	See Tables 4 & 5 on p. 41-44; only included values for species with ploidy values included in thesis
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted

Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schoen et al. 1992	Fungus	Biotrophic	Score	See Table 1; mean & stdev weighted
Schubiger et al. 2010	Fungus	Biotrophic	Score	See Table 2; used Mean column
Schubiger et al. 2010	Fungus	Biotrophic	Score	See Table 3; used Mean column
Schubiger et al. 2010	Fungus	Biotrophic	Score	See Table 7; used Mean column
Schuster 1991	Fungus	Biotrophic	Score	See Table 1 (diploids on top half, tetraploids on bottom half); calculated mean infection across 12 fungal races
Seenivasan 2017	Nematode	Biotrophic	Root lesion index	See Tables 7, 8, 9, & 10 root lesion index column
Seenivasan 2017	Nematode	Biotrophic	Root lesion index	See Tables 4-9 root lesion index column
Singh et al. 2006	Fungus	Necrotrophic	Score	See Table 1; only used values for wheat relative species
Singh et al. 2006	Fungus	Necrotrophic	Score	See Table 1; only used values for wheat relative species
Ssekiwoko et al. 2006	Bacterium	Necrotrophic	Score	See Table 1; used mean disease indices column & ignored SEs
Ssekiwoko et al. 2006	Bacterium	Necrotrophic	Score	See Table 1; used mean disease indices column & ignored SEs
Stoffelen et al. 2000	Nematode	Biotrophic	Nematodes per gram	See Tables 2 & 3; used nematodes per gram roots column
Stoffelen et al. 2000	Nematode	Biotrophic	Nematodes per gram	See Tables 2 & 3; used nematodes per gram roots column
Stoffelen et al. 2000	Nematode	Biotrophic	Nematodes per gram	See Tables 2 & 3; used nematodes per gram roots column
Stoffelen et al. 2000	Nematode	Biotrophic	Nematodes per gram	See Tables 2 & 3; used nematodes per gram roots column
Stover & Waite 1960	Fungus	Hemibiotrophic	Score	See Table 6; HS=1, S=2, SR=3, R=4, HR=5, I=6
Svara et al. 2021	Fungus	Hemibiotrophic	pg/ng plant DNA	WebPlotDigitizer used on Figure 3
Swiezynski et al. 1991	Oomycete	Hemibiotrophic	Score	See Table 1
Thangavelu et al. 2021	Fungus	Hemibiotrophic	Percent discolored	See Table 1; HS through I coded as 1-6; averaged both glasshouse & field values
Tofte 1990	Oomycete	Necrotrophic	Score	See p. 28 (Table 4)
Tomlinson et al. 1987	Bacterium	Necrotrophic	Percent_diseased	See Table 2; averaged sucker and bunching de-transformed percentages for each cultivar
Tusa & Del Bosco et al. 2000	Fungus	Necrotrophic	Score	See Table 1; used severity of disease rows
Uchneat 1997	Fungus	Necrotrophic	Lesion diameter	See Ch. 2 Table 2.2 Standardized (Std X) column

Uchneat et al. 1999	Fungus	Necrotrophic	Score	Averaged standardized values in Table 1
Van Tuyl 1982	Bacterium	Necrotrophic	Score	See Table 6; only used diploids (16 chromosomes) and triploids (24 chromosomes) listed in Table 1
Vestad 1960	Fungus	Necrotrophic	Number_dead	See Table 1
Viaene et al. 2003	Nematode	Biotrophic	Percent root bases with lesions	See Tables 5 & 8; included values from both tissue culture & corm tests at 16 weeks
Viaene et al. 2003	Nematode	Biotrophic	Percent root bases with lesions	See Tables 5 & 8; included values from both tissue culture & corm tests at 16 weeks
Vincelli et al. 2008	Fungus	Hemibiotrophic	Disease_severity	See Table 1; used values from all 5 experiments only for those entries where ploidy was included
Vining et al. 2015	Fungus	Hemibiotrophic	Score	See Tables 1 & 2; only used octoploids in Table 2; used Mean column
Vleugels et al. 2013	Fungus	Necrotrophic	Score	Removed NAs from Table 1 & SE values; did not include virus because there were multiple lumped together
Vleugels et al. 2013	Fungus	Biotrophic	Score	Removed NAs from Table 1
Vleugels et al. 2013	Fungus	Biotrophic	Score	Removed NAs from Table 1
Vymyslicky et al. 2012	Fungus	Hemibiotrophic	Score	See Table 3; used AGD (average grade of disease) column
Vymyslicky et al. 2012	Virus	Biotrophic	Score	See Table 3; used IPP (infected plant percentage) column
Wang et al. 2018	Oomycete	Biotrophic	Score	See Table 3; averaged across values 6-10 days & across each 3 of same ploidy; ignored SE
Wang et al. 2020	Bacterium	Hemibiotrophic	Lesion size	See Table 1; included means from all 3 years
Wang et al. 2020	Bacterium	Hemibiotrophic	Lesion size	See Table 1
Wang et al. 2020	Bacterium	Hemibiotrophic	Lesion size	See Table 2; included means from all 3 years
Wang et al. 2020	Bacterium	Hemibiotrophic	Lesion size	See Table 2
Wang et al. 2021	Fungus	Necrotrophic	Lesion size	WebPlotDigitizer used on Figure 3
Whitaker & Hokanson 2009	Fungus	Hemibiotrophic	Score	See Table 1; ignored SEs
Wilkins 1973	Fungus	Necrotrophic	Number_lesions	See Table 1
Wise & Gobelman-Werner 1993	Fungus	Biotrophic	Score	See Table 1
Wiwart et al. 2016	Fungus	Hemibiotrophic	Kernels_per_spike	See Table 1 (kernels per spike); averaged 2010-2012 & across tetraploids
Wiwart et al. 2016	Fungus	Hemibiotrophic	Kernels_per_spike	See Table 1 (kernels per spike); averaged 2010-2012
Yeates et al. 1973	Nematode	Biotrophic	Larvae_per_plant	See Table 2; averaged across 20 & 33 days
Yeates et al. 1973	Nematode	Biotrophic	Larvae_per_plant	See Table 4; averaged across 20 & 33 days

Yeates et al. 1973	Nematode	Biotrophic	Larvae_per_plant	See Table 2; averaged across 20 & 33 days
Yeates et al. 1973	Nematode	Biotrophic	Larvae_per_plant	See Table 4; averaged across 20 & 33 days
Yong-Fang et al. 1997	Fungus	Hemibiotrophic	Score	See Table 3; coded HR-HS as 1-5
Yong-Fang et al. 1997	Fungus	Hemibiotrophic	Score	See Table 3; coded HR-HS as 1-5
York 1989	Nematode	Biotrophic	Eggs_per_plant	See Table 1
Yun et al. 2001	Oomycete	Biotrophic	Score	See Table 4; used data only for Vitis vinifera-labrusca hybrids
Zimnoch-Guzowska et al. 1999	Bacterium	Necrotrophic	Score	See Table 2; included values from both resistant & susceptible diploids; used only values from tuber reaction columns
Zimnoch-Guzowska et al. 1999	Bacterium	Necrotrophic	Score	See Table 2; included values from both resistant & susceptible diploids; used only values from tuber reaction columns
Zlesak et al. 2010	Fungus	Hemibiotrophic	Score	See Table 1; S (susceptible) coded as 0, S* coded as 1, and R (resistant) coded as 2; included values from all 3 race columns
Zlesak et al. 2010	Fungus	Hemibiotrophic	Score	See Table 1; S (susceptible) coded as 0, S* coded as 1, and R (resistant) coded as 2; included values from all 3 race columns

Table 4 – Further description of data obtained from each paper included in our meta-analysis. The "m1i" column denotes the mean of the diploid group, the "sd1i" column denotes the standard deviation of the diploid group, the "n1i" column denotes the sample size of the diploid group, the "m2i" column denotes the mean of the polyploid group, the "sd2i" column denotes the standard deviation of the polyploid group, the "n2i" column denotes the sample size of the polyploid group, and the "Effect Direction" column denotes whether higher mean values indicate greater pathogen resistance ("0") or lesser pathogen resistance ("1").

Reference	m1i	sd1i	n1i	m2i	sd2i	n2i	Effect Direction
Abdelhalim et al. 2016	0.4775	0.31116716	4	0.5125	0.30793668	4	0
Abdelhalim et al. 2016	0.535	0.31214313	4	0.48	0.22315914	5	0
Alam & Gustafson 1988	2.10606061	0.74686602	66	2.10273973	0.81983558	146	0
Arrivillaga et al. 2004	2.78	NA	1	3.42	0.31945266	5	1
Babiker et al. 2018	2.195	1.56022755	6	3.36666667	0.93782931	15	1
Barekye et al. 2009	478	NA	20	588	NA	16	1
Bekal et al. 1998	4.5	4.51097427	54	6.85238095	7.46547298	42	1
Bekal et al. 1998	4.5	4.51097427	54	9.89072848	6.87275152	151	1
Bekal et al. 1998	2.75833333	3.27343532	12	11.47	6.35453469	10	1
Bekal et al. 1998	2.75833333	3.27343532	12	6.97647059	4.47125898	34	1
Blythe et al. 2015	1.6196319	0.62071831	163	2.40776699	0.63071262	412	1
Bon et al. 2020	0.10466667	0.14408447	3	0	0	5	1
Bon et al. 2020	0.03633333	0.02853653	3	0.0132	0.01434225	5	1
Borner et al. 2006	1.37735849	1.25320792	106	7.73514852	1.33996728	7676	1
Borner et al. 2006	1.37735849	1.25320792	106	7.10958296	2.4056151	6762	1
Bradshaw et al. 2021	107.083333	323.929652	18	789.25	756.957809	2	1
Bradshaw et al. 2021	724.776596	764.663654	94	642.266667	761.32239	15	1
Bradshaw et al. 2021	724.776596	764.663654	94	519.428571	315.445905	7	1
Burdon & Marshall 1981	0.14	NA	29	0.42	NA	47	0
Busch & Smith 1981	2.8258348	1.55611596	569	2.71724138	1.28687528	116	1
Busey et al. 1993	690	203.141987	4	470	286.298213	4	1
Carbajal et al. 2021	7.47157895	1.6690703	57	5.43222222	2.08752112	9	1
Carputo et al. 1997	1.20212766	0.87473237	94	2	0	2	0
Carputo et al. 1997	1.20212766	0.87473237	94	1.25	0.5	4	0
Celebi et al. 1998	0.5275	0.175	4	1	0	4	1
Cheo & Beaupre 1981	3.44238095	0.79965558	21	3.3625	1.25933471	24	1
Chung et al. 2011	61	30.5	27	62.5	38	6	1
Chung et al. 2011	61	30.5	27	76	20.2	5	1
Costa et al. 2008	30.4916667	18.4781669	12	35.1375	23.6606408	24	1
Cotrut et al. 2013	0.22444444	0.22820556	9	0.06083333	0.05124954	12	1
Cotrut et al. 2013	0.22444444	0.22820556	9	0.01833333	0.02401389	6	1
Craenen et al. 1997	2.25	1.21543109	12	1	0	4	0
Craenen et al. 1997	2.25	1.21543109	12	2.125	1.12599163	8	0
Das et al. 2010	0.22166667	0.13377842	6	0.216	0.12411464	10	1
Das et al. 2010	0.22166667	0.13377842	6	0.116	0.07021396	5	1
Das et al. 2013	2.91666667	1.03749163	6	2.7	1.08525471	10	1
Das et al. 2013	2.91666667	1.03749163	6	1.9	0.73786479	5	1
Das et al. 2013	2.91666667	1.03749163	6	2.66666667	0.51639778	3	1
Das et al. 2014 (a)	0.22	NA	1	0.26	0.1393391	14	1
Das et al. 2014 (b)	2.42857143	0.78679579	7	2.04166667	0.85867272	24	1
Das et al. 2014 (b)	2.42857143	0.78679579	7	1.77777778	0.83333333	9	1
Datson et al. 2015	0.44535926	0.37542	16	0.73959499	0.15089327	7	0

Datson et al. 2015	0.44535926	0.37542	16	0.96	0.05656854	2	0
De Matos et al. 2009	1.256111	2.06683372	10	1.28667	1.84610183	10	1
Devi et al. 2021	0.13188	0.06881587	10	0.19472857	0.06536609	49	1
Dijkstra 1964	7.43529412	2.10622327	17	10.9	3.67151195	3	0
Dijkstra 1964	7.13529412	2.17138814	17	10.5333333	2.89194283	3	0
Dochez et al. 2006	0.2965625	0.18456904	16	0.33166667	0.16594606	15	1
Dochez et al. 2013	15.8	5.81343272	6	13.644	8.51539178	15	1
Dochez et al. 2013	15.8	5.81343272	6	11.2166667	7.06637578	6	1
Duan et al. 2021	3.21276596	1.02413633	47	3.039	0.8809441	125	0
Duan et al. 2021	3.21276596	1.02413633	47	3.41666667	1.42156018	3	0
Ehlenfeldt & Stretch 2001	0.769375	0.17718589	8	0.548667	0.18447854	3	0
Felber 1987	0.44	0.51130999	18	0.5	0.52704628	10	0
Fock et al. 2005	0.61	0.35355339	2	0.817	0.2100291	10	1
Fogain 2000	3.70933333	0.79598754	60	2.93512821	1.25502332	39	1
Franco et al. 2015	0.89176471	0.15918681	17	0.95631579	0.03632342	38	0
Franco et al. 2015	0.10823529	0.15918681	17	0.01736842	0.036811	38	1
Geise 1957	5.590909	3.6076518	22	2.52857	1.02988317	14	1
Giblin-Davis et al. 1995	6223.33333	3933.24717	3	3786.66667	645.31646	3	1
Goncalves et al. 2019	1.91304348	0.19377669	23	1.9375	0.1767767	8	0
Goncalves et al. 2019	1.91304348	0.19377669	23	1	1.15470054	2	0
Gooding et al. 1981	0.099	0.23149357	12	0.61846154	0.09745479	13	1
Green 1959	0.47244506	0.20861828	9	0.53836923	0.27119744	4	0
Green 1959	0.47244506	0.20861828	9	0.32146429	0.12189354	2	0
Gultyaeva et al. 2016	14.6	1.24237675	35	10.1791667	0.1643	96	1
Gultyaeva et al. 2016	14.6	1.24237675	35	12.7981482	0.3916	216	1
Gunavathi 2000	0.18181818	0.10332649	11	0.31333333	0.17080691	9	1
Gunavathi 2000	718.818182	146.969261	11	1075.875	329.377091	8	1
Gunter & Egel 2012	242.945455	291.33234	22	429.62	249.365727	15	1
Hadi et al. 2012	0.336	0.29983699	5	0.16	0.28905978	5	1
Harding 1971	0.185	0.05802298	4	0.59153846	0.12542236	13	0
Harding 1971	0.185	0.05802298	4	69.1666667	20.067311	12	0
Harms et al. 2020	16.275	3.031	4	21.175	5.84	4	1
Harms et al. 2020	11.336	7.704	4	20.382	4.227	4	1
Harms et al. 2020	14.655	2.007	4	13.678	2.697	4	1
Hartman et al. 2000	0.687	0.24594715	10	0.875	0.05744563	4	0
Hecker & Ruppel 1976	0.515	0.03055051	3	0.51	0.10598742	3	0
Hecker & Ruppel 1976	0.185	0.09466315	10	0.342	0.25094488	10	0
Henderson & Jenkins 1977	10.35	11.4719059	16	12.1666667	12.528392	12	1
Henderson & Jenkins 1977	10.35	11.4719059	16	10.6	12.1029028	13	1
Hias et al. 2018	0.234	0.396	9	0.24	0.407	9	1
Irwin 1981	2.9	0.94140852	5	2.69375	1.89744895	8	1
Irwin et al. 1997	3	1	5	2.75	1.98206242	8	1
Jacob et al. 2010	0.57125	0.13361013	16	0.42076923	0.11390504	13	0
Jansky et al. 2006	0.3825433	0.1620961	25	0.2418801	0.1317962	6	0
Jansky et al. 2006	0.3825433	0.1620961	25	0.434558	0.1414256	3	0
Julier et al. 1996	4.16	0.54295488	5	4.354	0.31121615	14	1
Julier et al. 1996	3.282	0.3361101	5	3.251	0.37477358	15	1
Julier et al. 1996	4.184	0.51954788	5	4.341	0.69716853	14	1
Julier et al. 1996	0.85	0.66475559	5	3.3535	0.95404003	20	1
Khiutti et al. 2012	1.92307692	1.19743315	26	2.03846154	1.18256566	26	1
Khiutti et al. 2015	7.407	1.562	48	8.425	0.47	10	0

Khiutti et al. 2015	7.407	1.562	48	7.025	1.9817739	10	0
Kono et al. 2014	2.99492017	1.41228458	26	2.56839623	0.89239143	8	1
Kriel et al. 1995	51.1776923	37.3691002	13	8.82	27.5333474	13	1
Kulkarni & Ravindra 1988	0.545214	0.20721456	15	0.979333	0.02217544	9	0
Kumar et al. 2009	1.92307692	1.25575598	13	3.3125	1.92245503	16	1
Kumar et al. 2009	1.92307692	1.25575598	13	2.6	2.19089023	5	1
Kumar et al. 2009	2.26923077	0.85672304	13	3	0.96609178	16	1
Kumar et al. 2009	2.26923077	0.85672304	13	2.7	1.15950181	5	1
Lamari & Bernier 1989 (a)	2.6097561	1.15926806	41	3.49652778	1.20682803	288	1
Lamari & Bernier 1989 (a)	2.6097561	1.15926806	41	3.75075988	1.17593024	329	1
Lamari & Bernier 1989 (a)	2.6097561	1.15926806	41	2.375	0.91612538	8	1
Lamari & Bernier 1989 (b)	3.66666667	2.30940108	3	2.80952381	1.74982992	21	1
Lamari & Bernier 1989 (b)	3.66666667	2.30940108	3	2.38888889	1.96497102	9	1
Levinson et al. 2021	0.1376	0.47050753	25	0.36736842	0.72825561	19	1
Limantseva et al. 2014	6.70689655	2.22433606	58	6.75	1.75032465	56	0
Mazzafera et al. 1993	0.85266667	0.341749	15	0.54333333	0.51594573	9	1
Mikaliuniene et al. 2015	0.18541861	0.05600733	43	0.19619512	0.09660932	41	1
Mikaliuniene et al. 2015	0.30244186	0.0747191	43	0.31204878	0.08707355	41	1
Mudonyi et al. 2019	0.46833333	0.24194352	6	0.665	0.09804336	17	1
Mudonyi et al. 2019	0.46833333	0.24194352	6	0.51	0.10165301	4	1
Nakato et al. 2019	433.81579	182.985431	38	428.766667	119.005269	30	1
Nakato et al. 2019	433.81579	182.985431	38	455	201.623411	4	1
Nardozza et al. 2015	0.55317757	0.37507467	16	0.26031042	0.14967785	7	1
Nardozza et al. 2015	0.55317757	0.37507467	16	0.96	0.05656854	2	1
Naydenova & Aleksieva 2017	2.75	0.7594356	12	3	0.54751505	12	1
Nguyet et al. 2002	0.33361539	0.24277374	13	0.26875	0.11125742	8	1
Ohberg et al. 2005	0.20792308	0.07058619	13	0.21285714	0.06519312	7	0
Oliveira et al. 2018	0.50843265	0.27261956	3	0.73617766	0.16975573	14	0
Pair & Bruton 1998	0.7834	0.10784367	15	0.9034	0.03386444	5	0
Pang 2010	0.67083333	0.63209738	24	0.61458333	0.46447847	48	1
Pang 2010	0.67083333	0.63209738	24	0.61363636	0.52761409	22	1
Pang et al. 2011	33.8333333	31.294765	24	31.04	23.3333955	50	1
Pang et al. 2011	33.8333333	31.294765	24	30.45	26.3008405	20	1
Pang et al. 2011	33.8333333	31.294765	24	21.3333333	11.1115556	8	1
Paul & Freudenstein 1989	5	1.18292449	30	4.6602439	0.96896978	42	1
Pederson & Windham 1989	1.60576923	0.8748733	104	1.98076923	0.69646499	104	1
Pederson & Windham 1989	1.60576923	0.8748733	104	2.15463918	0.85805262	97	1
Perez et al. 2014	7.57	0.56	3	6.23	2.24859067	6	1
Perez et al. 2014	7.57	0.56	3	6.15	0.35355339	2	1
Pinochet et al. 1998	0.74	0.18384776	2	0.694	0.11711343	10	1
Pinochet et al. 1998	0.74	0.18384776	2	0.65333333	0.08144528	3	1
Pinochet et al. 1998	0.63	0.16970563	2	0.85	0.14256577	9	1
Pinochet et al. 1998	0.63	0.16970563	2	0.69333333	0.13613719	3	1
Pinochet et al. 1998	0.435	0.04949748	2	0.515	0.23847898	10	1

Pinochet et al. 1998	0.435	0.04949748	2	0.31333333	0.10214369	3	1
Podwyszynska et al. 2021	3.07142857	1.04676372	7	0.0125	0.0341565	16	1
Poteri et al. 1997	23.39875	16.7655401	16	13.74	17.5319024	4	1
Prasad et al. 2009	0.4048	0.20471487	10	0.202	0.10582533	3	1
Queneherve et al. 2009	32.8775	38.6587473	40	42.9142857	40.635573	21	1
Queneherve et al. 2009	32.8775	38.6587473	40	90.1	46.9518903	2	1
Ray et al. 1995	1.3	0.78102497	3	1.38666667	0.41895562	15	0
Reboucas et al. 2018	0.5	1	4	27.9090909	40.0261278	11	1
Rhodes et al. 1996	1.572727	1.35874273	11	0.72222	0.71200031	9	1
Ribeiro et al. 2018	0.88235294	0.33210558	17	0.80952381	0.40237391	21	0
Ribeiro et al. 2018	0.88235294	0.33210558	17	1	0	2	0
Rothleutner 2012	0.2674	0.33455777	6	0.17965263	0.26380167	38	1
Rothleutner 2012	0.2674	0.33455777	6	0.17414	0.16207086	5	1
Schoen et al. 1992	0.83	0.07265185	78	0.78	0.141	85	0
Schoen et al. 1992	0.87	0.16018409	78	0.81	0.18269058	85	0
Schoen et al. 1992	0.77	0.14703904	78	0.78	0.039428	85	0
Schoen et al. 1992	0.79	0.12499778	78	0.73	0.16195871	85	0
Schoen et al. 1992	0.68	0.1735978	78	0.68	0.1291149	85	0
Schoen et al. 1992	0.71	0.2246906	78	0.71	0.10932286	85	0
Schoen et al. 1992	0.65	0.20957701	78	0.68	0.1291149	85	0
Schubiger et al. 2010	5.125	1.33068189	8	3.35	1.04056609	10	1
Schubiger et al. 2010	4.37368421	0.93917346	19	3.99285714	0.98329453	14	1
Schubiger et al. 2010	3.9631579	0.63875793	19	2.85714286	0.38373526	14	1
Schuster 1991	0.13888894	0.26176056	6	0.77976191	0.37011814	7	1
Seenivasan 2017	18.2527778	9.33235497	36	24.2222222	10.7618281	9	1
Seenivasan 2017	18.2527778	9.33235497	36	27.162069	8.27884318	116	1
Singh et al. 2006	1.28571429	0.56061191	21	3.20098039	0.98079796	408	1
Singh et al. 2006	1.28571429	0.56061191	21	3.25	1.21543109	12	1
Ssekiwoko et al. 2006	82.7533333	21.8405869	15	88.8666667	6.44112531	15	1
Ssekiwoko et al. 2006	82.7533333	21.8405869	15	88.8166667	3.09422727	12	1
Stoffelen et al. 2000	163.666667	234.983687	3	307.222222	228.448998	9	1
Stoffelen et al. 2000	163.666667	234.983687	3	124.333333	22.2336082	3	1
Stoffelen et al. 2000	102.666667	126.847678	3	161.555556	167.824396	9	1
Stoffelen et al. 2000	102.666667	126.847678	3	111.333333	82.0507973	3	1
Stover & Waite 1960	4.25	2.06155281	4	3.75	1.43759058	16	0
Svara et al. 2021	22.2670251	43.4156727	8	8.871	16.338	8	1
Swiezynski et al. 1991	0.45166667	0.32664456	6	0.80625	0.34083668	8	0
Thangavelu et al. 2021	2.80487805	1.86274316	123	2.16402116	1.49431344	378	0
Tofte 1990	4.642	0.20608251	5	4.796	0.08561542	5	1
Tomlinson et al. 1987	0.28833333	0.14932068	6	0.04785714	0.07566918	14	1
Tusa & Del Bosco et al. 2000	2.44	0.92095603	6	2.3525	0.97769712	4	1
Uchneat 1997	2.65714286	0.65414939	21	2.7426087	0.58426502	23	1
Uchneat et al. 1999	2.51769231	0.80469171	39	3.23869565	0.68909232	23	1
Van Tuyl 1982	1.91	0.41992063	7	2.76	0.61291653	4	1
Vestad 1960	0.539	0.04455053	6	0.67	0.05053657	6	0
Viaene et al. 2003	15.1666667	14.9052563	6	25.8571429	14.2645081	7	1
Viaene et al. 2003	15.1666667	14.9052563	6	34.1666667	21.6556382	6	1
Vincelli et al. 2008	57.1766667	19.9983738	30	53.73	16.6165507	30	1
Vining et al. 2015	2.17826087	0.64240777	23	2.4	0.62225397	26	1
Vleugels et al. 2013	0.30812632	0.07287402	95	0.4315	0.08186574	8	1
Vleugels et al. 2013	2.20412371	0.51253405	97	2.42	0.552811	10	1
Vleugels et al. 2013	3.12268041	0.59169692	97	3.56	0.46303348	10	1

Vymyslicky et al. 2012	2.45638298	0.36454513	47	2.50192308	0.26588748	26	1
Vymyslicky et al. 2012	0.71652174	0.14629069	46	0.73115385	0.14339671	26	1
Wang et al. 2018	2.286	0.85966605	15	1.34066667	0.52624637	15	1
Wang et al. 2020	7.18992754	7.45186072	138	4.9869697	10.1493271	33	1
Wang et al. 2020	7.18992754	7.45186072	138	5.83333333	0.4441096	3	1
Wang et al. 2020	22.706748	11.6178253	123	1.93666667	0.31866911	9	1
Wang et al. 2020	22.706748	11.6178253	123	20.4844444	14.2734352	9	1
Wang et al. 2021	152	108.324	7	143	96.108	7	1
Whitaker & Hokanson 2009	1.48166667	1.238635	6	2.31166667	1.12847537	6	1
Wilkins 1973	4.13	NA	3	3.06	NA	3	1
Wise & Gobelman-Werner 1993	0.475	4.9244289	40	0.5875	2.99702233	80	0
Wiwart et al. 2016	0.131	0.13304511	3	0.0888333	0.11864134	6	0
Wiwart et al. 2016	0.131	0.13304511	3	0.016	0.06630234	3	0
Yeates et al. 1973	5.5	3.8340579	6	2	1.26491106	6	1
Yeates et al. 1973	0	0	6	0.16666667	0.40824829	6	1
Yeates et al. 1973	10.5	11.5195486	6	2.33333333	2.42212028	6	1
Yeates et al. 1973	9.5	10.6536379	6	1.5	1.76068169	6	1
Yong-Fang et al. 1997	5	0	5	4.72327044	0.60495789	159	1
Yong-Fang et al. 1997	5	0	5	3.775578	1.18369	909	1
York 1989	1055	NA	361	6116	NA	502	1
Yun et al. 2001	1.44	NA	18	2.72	NA	14	1
Zimnoch-Guzowska et al. 1999	5.70617647	1.82899959	204	4.54608696	1.07134469	23	1
Zimnoch-Guzowska et al. 1999	5.70617647	1.82899959	204	5.362	0.67395846	5	1
Zlesak et al. 2010	0.35	0.70890223	60	0.85858586	0.9584224	99	0
Zlesak et al. 2010	0.35	0.70890223	60	0.59722222	0.91404876	72	0

CONCLUSION

In every example in which its immediate effects have been analyzable, polyploidy has appeared as a complicating force, producing innumerable variations on old themes, but not originating any major new departures.

- *G. Ledyard Stebbins (1950)*

Summary

Polyploidy has been intensely studied by biologists for over 100 years because its effects are highly context dependent. This fact, along with the resurgence of research on polyploidy in recent decades, are responsible for the large number of reviews in the field (Soltis et al. 2010) that attempt to make sense of this exciting discipline. This thesis has added two more reviews to the pile: one narrative (Chapter I) and one systematic (Chapter III), examining persistent problems in polyploidy research with modern phylogenetic comparative methods (PCMs). They, along with Chapter II, support the context-dependent view in that polyploids do not appear to behave much differently from diploids.

In Chapter I, I critically reviewed work on the dead-end hypothesis in polyploidy research. This project developed from my initial readings of the polyploidy literature, and it required consultation of more than 250 papers, which laid the groundwork for the other two chapters in this dissertation. I argued that a supposedly singular "dead-end hypothesis" has referred to several, widely varying hypotheses since the work of G. Ledyard Stebbins (1950; 1971), and that since the advent of modern PCMs, two distinct hypotheses are actively being studied: the "traditional" dead-end hypothesis, and the "rarely successful" hypothesis. Beyond this review, I also conducted the first comparison of tip diversification rates in diploid and

polyploid plants, employing the recently developed MiSSE (Vasconcelos et al. 2022) to study Solanaceae. I found no significant differences between tip rates based on either ploidy or the closely related trait of breeding system.

Next, in Chapter II, I moved from polyploid diversification to polyploid biogeography, studying the mechanisms behind the latitudinal polyploidy gradient (LPG), in which polyploids are proportionally more frequent in plant communities at higher latitudes. Using corHMM (Beaulieu et al. 2013; Boyko and Beaulieu 2021), which I modified to perform ancestral state reconstruction at selected time slices, I found widely varying histories of ploidy transitions across four flowering plant clades. Using machuruku (Guillory and Brown 2021), I found mixed support for the "centers of origin" hypothesis, in which polyploids originate at higher rates in poleward environments, and the "centers of arrival" hypothesis, where the LPG is created by antiequatorial movement by plants post-polyploidization. In this first test of the mechanisms behind the LPG using a global phylogeny, I did not detect strong differences in geographic patterns of origination or movement between diploids and polyploids.

Finally, in Chapter III, I took a macroevolutionary perspective on the microevolutionary problem of comparing pathogen resistance between diploid and polyploid plants. To do so, I synthesized 214 effect sizes from 128 studies in the first-ever meta-analysis on the subject, incorporating a family-level phylogeny among other moderators into a multi-level model. I found no evidence of phylogeny on observed patterns of pathogen resistance, and I detected no significant advantages of polyploidy for resisting infections.

What Do We Do About Lags?

A persistent question throughout all three chapters was how results may or may not be explained

by “lags” in time between polyploidization and the emergence of beneficial traits or subsequent diversification events. Despite the macroevolutionary perspective of my work, the chapters presented consider relatively brief periods in geological time, the most ancient being the 3.3 million years of ploidy evolution considered in Chapter II. They consider only recently developed neopolyploids, but it may require millions of years post-polyploidization for effects to be observed (Schranz et al. 2012). How might the findings of each chapter change if the perspective were shifted from neopolyploids to paleopolyploids, lineages which underwent polyploidization millions of years ago and have since significantly re-organized their genomes and even re-diploidized (Dodsworth et al. 2016)?

I considered this question most intensely while working on Chapter I. While one of the advantages of tip rate studies is that most information on a phylogeny is clustered near the present (O’Meara and Beaulieu 2021), studies of ancient whole genome multiplications and their subsequent effects on diversification must examine patterns in the nodes, especially to detect a lag. Using the HiSSE framework (Beaulieu and O’Meara 2016), one can model lags using hidden states, specifically by characterizing observed ploidy states as 0 and 1 (diploid and polyploid) and hidden states as A and B (low diversification rate class and high diversification rate class). Species start in combined state 0A, and can polyploidize to transition to state 1A, but they must undergo an additional transition from state 1A to state 1B in order to diversify, with the time taken to undertake this transition simulating a lag. The lag model can be compared with such models as one with diversification but no lag (transitions to polyploidy immediately cause diversification shifts, with no hidden states) and one where polyploidy is unlinked with diversification (shifts between states 0A, 0B, 1A, and 1B are all possible and equally likely).

While the work is ongoing and was thus not included in this dissertation, I aimed to

study lags in exactly this way by developing the model PolySSE within HiSSE. PolySSE differs from HiSSE in how sampling fractions are treated (PolySSE requires two separate sampling fractions, for the diploid and polyploid tips in the tree), in that state transitions (in this case polyploidization events) are fixed along branches in the phylogeny, and in its calculation of time lags between polyploidization and diversification events (which cannot be simply derived by inverting transition rates). It provides the first direct, model-based test of the lag hypothesis where other workers have compared diversification rates and ploidy transitions post-hoc (Tank et al. 2015; Landis et al. 2018; Smith et al. 2018).

Preliminary tests using the Caryophyllales phylogeny from Smith et al. (2018) currently indicate that character-independent models, in which transitions freely occur between observed and hidden states, are supported above character-dependent lag models and models in which diversification shifts occur concomitantly with ploidy shifts. However, further work, including simulation testing and sister group comparisons, are necessary prior to publication.

Possible Extensions

In Chapter II, I discussed the possibility that a third hypothesis may explain the data: the "centers of survival" hypothesis, in which the LPG is created by greater survival of polyploids at higher latitudes relative to diploids. I would like to test this hypothesis using a diversification rate analysis, particularly whether diploids at higher latitudes show higher extinction rates than either polyploids at those same latitudes or diploids in more temperate environments. Additionally, I plan to explore new ways of examining movement, or the potential to do so, in lineages beyond comparing median latitudes reconstructed with machuruku between time slices. Ranges reconstructed with machuruku appear quite vulnerable to bias when closely related clades differ widely in their present-day latitudes: in a hypothetical three-taxon tree, if two species occur near

the equator while a third species occurs near the north pole, it seems likely that machuruku will reconstruct their common ancestor near the equator rather than at a more intermediate latitude. While this issue may be ameliorated by altering the program's ancestral state reconstruction algorithm, machuruku also appears to reconstruct consistently smaller ranges as analyses proceed further back in time, beyond what one would expect from the loss of recently diverged lineages. For these reasons, I intend to examine the LPG further using the newly developed hOUwie (Boyko et al. 2023). This program allows for jointly modeling discrete and continuous traits to determine whether discrete character transitions are correlated with the evolutionary trajectory of the continuous trait. I hope to apply this to study whether shifts to polyploidy are correlated with either latitudinal movement toward higher latitudes or climatic niche evolution toward colder climates.

For Chapter III, I do not plan to extend the scope of the literature search or the number of moderators included in the multi-level model. However, meta-analysis appears under-utilized in polyploidy research, especially considering the popularity of narrative reviews in the field. I hope to continue applying meta-analytic methods to several problems in experimental polyploidy research, particularly comparisons of diploids and polyploids regarding drought tolerance, endurance of high soil salinity, and photosynthetic rates.

Future Directions in Model-Based Polyploidy Research

Future polyploidy research will increasingly adopt the multi-clade approach (see Vasconcelos 2023), in which researchers combine the advantages of close, detailed study of individual clades and their idiosyncratic biological attributes with those of broad-scale analyses of evolutionary rates on robust phylogenies. Because the effects of polyploidy are heavily context-dependent (e.g., Segraves 2017), and because I expect, based on my results, that many clades will continue

to show little connection between ploidy and evolutionary events, future research will dedicate greater focus to the possible clade-specific causes governing ploidy's effects, or lack thereof, on diversification, trait evolution, and other evolutionary phenomena. Analyses of remarkable clades are already underway. Han et al. (2020) recently compared diversification rates across ploidy states in the genus *Allium*, which contains many species with mixed ploidy levels. They found that diversification generally correlates positively with the ratio of polyploids to diploids, possibly because polyploidy allows radiations into drier, drought-prone habitats. Additionally, they suggest that mixed ploidy systems in the genus may be advantageous for colonizing environments with varying ecological conditions, constituting a kind of "genomic plasticity." This phenomenon has, so far, received little attention in the comparative literature, but will likely be studied widely in the near future.

Concluding Remarks

Taken together, the findings of the three chapters of this thesis accord with the view that polyploidy does not consistently confer beneficial traits or lead to bursts of diversification in flowering plants. When I began this research, I was convinced that the advancement of PCMs, which were unavailable to Stebbins during most of his research career, would succeed in proving the Stebbinsian view of polyploidy wrong. However, considering only this research, I must side with the view of Stebbins over that of Haldane.

While I hope that the PCMs used here will be viewed by other researchers with great interest, it is likely that some will view them with skepticism, whether one believes that diversification analyses suffer from identifiability issues, (Louca and Pennell 2020), that rate estimates on "large" phylogenies provide insufficient detail about biological patterns (Donoghue and Edwards 2019), that phylogenetic "correction" is likely irrelevant in ecological studies

(Westoby et al. 1995), or other criticisms. While the results of any comparative study are heavily influenced by the model and clade of choice, I do not agree with these critiques. Beyond the fact that SSE-class models are likely robust to the identifiability issues identified in other diversification models (O'Meara and Beaulieu 2021), new methods in comparative biology often come under attack soon after their development. In many ways this is the scientific process working healthily: soon after the introduction of BiSSE (Maddison et al. 2007), it was discovered that the model had a high rate of false positives, even finding support for nonsensical traits driving diversification rates (Rabosky and Goldberg 2015). This, in turn, spurred the introduction of hidden states for decreasing false positive rates in -SSE analyses (Beaulieu and O'Meara 2016). Yet, for example, when Louca and Pennell (2020) published their proof that an infinite number of diversification histories may explain a given lineage-through-time plot, some questioned whether we should even continue to estimate diversification at all (Helmstetter et al. 2022). I believe much of this stems from the pugnacious history of comparative biology and systematics, once described as having more infighting than any other scientific field (Felsenstein 1986). Old habits also die hard in comparative biology: some old-school cladists still refer to phylogenies as "metaphysical" (Brower 2023). What I have learned working in this field, as well as in the elusive area of polyploidy research, is to always keep an open mind: biology thrives off new approaches, bold ideas, and examinations of new clades, scales, phenomena, etc. The strangest ideas sometimes prove to be correct (e.g., Sagan 1967), and the fields labeled "black holes" (Barker et al. 2016) today may thrive tomorrow.

References

Barker, M.S., B.C. Husband, and J.C. Pires. 2016. Spreading Winge and flying high: the evolutionary importance of polyploidy after a century of study. *American Journal of Botany* 103: 1139–1145.

Beaulieu, J.M., and B.C. O'Meara. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* 65: 583–601.

Beaulieu, J.M., B.C. O'Meara, and M.J. Donoghue. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62(5): 725–737.

Boyko, J.D., and J.M. Beaulieu. 2021. Generalized hidden Markov models for phylogenetic comparative datasets. *Methods in Ecology and Evolution* 12: 468–478.

Boyko, J.D., B.C. O'Meara, and J.M. Beaulieu. 2023. A novel method for jointly modeling the evolution of discrete and continuous traits. *Evolution* 77: 836–851.

Brower, A.V.Z. 2023. Hierarchies, classifications, cladograms and phylogeny. *Cladistics* 39: 229–239.

Dodsworth, S., M.W. Chase, and A.R. Leitch. 2016. Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Botanical Journal of the Linnean Society* 180: 1–5.

Donoghue, M.J., and E.J. Edwards. 2019. Model clades are vital for comparative biology, and ascertainment bias is not a problem in practice: a response to Beaulieu and O'Meara (2018). *American Journal of Botany* 106: 327–330.

Felsenstein, J. 1986. Waiting for Post-Neo-Darwin. *Evolution* 40: 883–889.

Guillory, W.X., and J.L. Brown. 2021. A new method for integrating ecological niche modeling with phylogenetics to estimate ancestral distributions. *Systematic Biology* 70: 1033–1045.

Han, T.-S., Q.-J. Zheng, R.E. Onstein, B.M. Rojas-Andrés, F. Hauenschild, A.N. Muellner-Riehl, and Y.-W. Xing. 2020. Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist* 225: 571–583.

Helmstetter, A.J., S. Glemin, J. Käfer, R. Zenil-Ferguson, H. Sauquet, H. de Boer, L.P.M. Dagallier, N. Mazet, E.L. Reboud, T.L. Couvreur, and F.L. Condamine. 2022. Pulled diversification rates, lineages-through-time plots, and modern macroevolutionary modeling. *Systematic Biology* 71: 758–773.

Landis, J.B., D.E. Soltis, Z. Li, H.E. Marx, M.S. Barker, D.C. Tank, and P.S. Soltis. 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* 105: 348–363.

Louca, S., and M.W. Pennell. 2020. Extant timetrees are consistent with a myriad of diversification histories. *Nature* 580: 502–505.

Maddison, W.P., P.E. Midford, and S.P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56: 701–710.

O'Meara, B.C., and J.M. Beaulieu. 2021. Potential survival of some, but not all, diversification methods. *EcoEvoRxiv* https://doi.org/10.32942/osf.io/w5nvd

Rabosky, D.L., and E.E. Goldberg. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* 64: 340–355.

Sagan, L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14: IN1–IN6.

Schranz, M.E., S. Mohammadin, and P.P. Edger. 2012. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology* 15: 147–153.

Segraves, K.A. 2017. The effects of genome duplications in a community context. *New Phytologist* 215: 57–69.

Smith, S.A., J.W. Brown, Y. Yang, R. Bruenn, C.P. Drummond, S.F. Brockington, J.F. Walker, N. Last, N.A. Douglas, and M.J. Moore. 2018. Disparity, diversity, and duplications in the Caryophyllales. *New Phytologist* 217: 836–854.

Soltis, D.E., R.J.A. Buggs, J.J. Doyle, and P.S. Soltis. 2010. What we still don't know about polyploidy. *Taxon* 59: 1387–1403.

Stebbins, G.L. 1950. Variation and evolution in plants. Columbia University Press.

Stebbins, G.L. 1971. Chromosomal evolution in higher plants. Addison-Wesley, London, UK.

Tank, D.C., J.M. Eastman, M.W. Pennell, P.S. Soltis, D.E. Soltis, C.E. Hinchliff, J.W. Brown, E.B. Sessa, and L.J. Harmon. 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207: 454–467.

Vasconcelos, T. 2023. A trait-based approach to determining principles of plant biogeography. *American Journal of Botany* 110: e16127.

Vasconcelos, T., B.C. O'Meara, and J.M. Beaulieu. 2022. A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution* 76: 1420–1433.

Westoby, M., M.R. Leishman, and J.M. Lord. 1995. On misinterpreting the 'phylogenetic correction'. *Journal of Ecology* 83: 531–534.