

A simple algorithm to study phylogeographies & speciation patterns in space

Philippe Desjardins-Proulx,^{1,2,*} Timothée Poisot,² and Dominique Gravel²

¹*College of Engineering, University of Illinois at Chicago, USA.*

²*Université du Québec à Rimouski, Canada.*

This document is a working paper in which we present a simple theoretical framework based on network theory to study how speciation, the process by which new species appear, shape spatial patterns of diversity.

Keywords: Phylogeography; Speciation; Networks; Spatial ecology; Macroevolution.

I. MOTIVATION

The peculiar spatial relationship between closely related species was among the first pattern of diversity used to infer evolution. As early as the 1850s, Alfred Wallace noted that the closest relatives were often observed in adjacent yet non-overlapping regions [29, 30]. Wagner and Jordan later relied on a similar observation to argue for the importance of geography and isolation in the formation of new species [2]. And finally, Mayr developed a theory of allopatric speciation, a cornerstone of the modern synthesis, again using similar observations [18, 19]. The relationship between phylogeny and geography has shaped our understanding of the origin of species [2, 15]. It is also crucial to the development of a unified theory of community assembly [22, 25]. Yet, theory remains mostly silent about the subject. Few models can generate phylogeographies, and none can be used to study the effect of complex spatial structures [1]. This is surprising, not only because of the theoretical importance of phylogeography, but also because several phylogenetic methods use geography to infer patterns of speciation [1, 16, 17]. Part of the problem lies in the limitations of traditional mathematical methods: analytical solutions to spatially-explicit models are often only available for the most trivial cases [7]. Thus, we are left with no theoretical framework to study the patterns noted by Wallace, Wagner, and Mayr. In this document, we describe a very simple algorithm to generate phylogeographies in spatial networks. Our approach is inspired by metapopulation theory [9, 10, 13] although the spatio-temporal scale is different: we're interested in the dynamics of populations at the regional scale during long periods. The model will be used to study phylogeographies in various spatial contexts and to develop better tools to study the relationship between phylogeny and geography.

II. MODELING THE LANDSCAPE

We model the landscape as a spatial network of communities. A network is a flexible mathematical object

defined as a set of vertices V and a set of edges E , which are used to connect the vertices [23]. Here, the vertices represent communities and the edges denote migration [4–6]. Spatial networks are simply networks in which vertices are embedded in a known topological space [12], in our case a two-dimensional map. Thus, each community is represented by a vertex in the network and to a position on a map. Networks are increasingly common in ecology as they can be used to model complex structures and quantify the effect of clustering, connectivity, and isolation [3, 20, 21, 27]. In particular, isolation is the most important factor in many speciation events [2], making networks well-suited to study patterns of speciation in different contexts [4]. The spatial network can be built in two ways. First, random geometric networks can be generated by randomly placing the vertices on a surface, normally the unit square, and linking all communities within some threshold distance [24]. This technique is used to test network algorithms applied to maps [26]. Second, a real map can be used as a template for a spatial network [3, 20]. This method offers the exciting opportunity to generate predictions specific to a given spatial structure, and test the predictions of our algorithm against empirical data.

III. THE ALGORITHM

A species is divided in populations which are distributed in a network of communities. A species is either present or absent in a community, we do not keep track of the number of individuals. Occupancy thus follows the standard colonization/extinction dynamics of metapopulation theory [9]. For each time step, all populations have the opportunity to colonize adjacent communities (the vertices connected by an edge in the network). The probability of a successful colonization of community x by species i is

$$c_{ix} = c_{max} \exp \left(-\alpha \sum_{j \in \{S_x \setminus i\}} \delta_{ij}^{-1} \right),$$

with $\{S_x \setminus i\}$ being the set of populations present in community x minus i , δ_{ij} the time since species i and j 's most recent common ancestor (similar to the Faith index

*E-mail: philippe.d.proulx@gmail.com

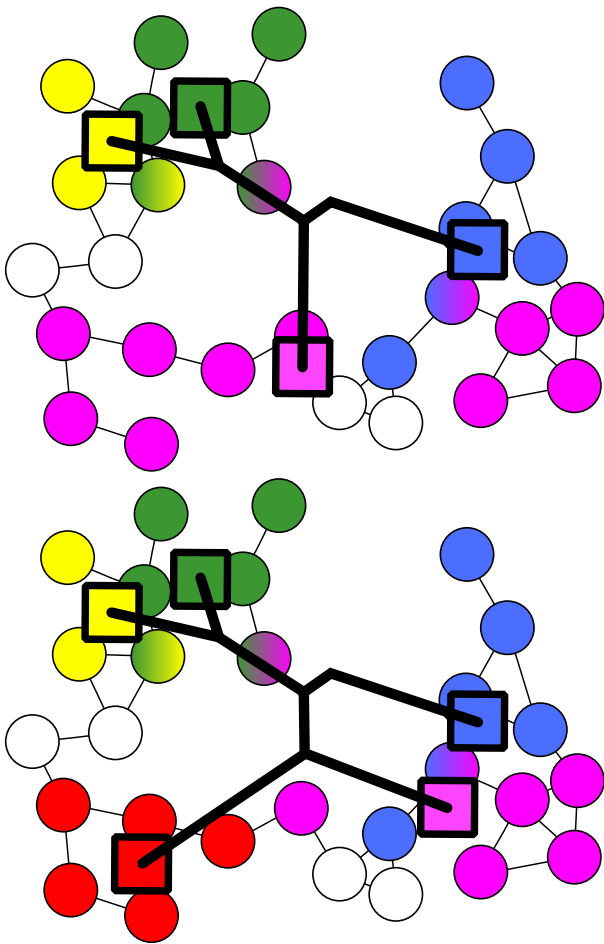


FIG. 1: Top: a phylogeography with four species (yellow, blue, green, pink). The populations are distributed in a spatial network, with each community (circles) hosting populations from 0 or more species. Empty communities are white and a gradient is used for communities with more than one species. The communities are connected by migration (thin black lines). Bottom: a speciation event. The pink species is divided in three groups of populations. Its leftmost group undergoes speciation and a connected subgroup now belongs to a new species (in red).

in groups of populations. We define a group as a set of connected populations from the same species (Fig. 1). Each group has a probability v of undergoing speciation. When speciation occurs in a group, a random subset of $[1, n]$ connected populations will speciate, with n being the number of populations in the original group (Fig. 1).

IV. IMPLEMENTATION

An implementation is available on github: <https://github.com/PhDP/wagner> and is released under the permissive MIT license.

[8]), c_{max} the highest possible colonization rate and α a constant (with $\alpha \geq 0$). α describes the decline of the intensity of interactions with phylogenetic divergence. In short, a higher α makes it difficult for closely related species to coexist. c_{ix} is a very simple function derived from exponential decay: $dm/d\delta = -\alpha\delta^{-1}$. It is based on an old hypothesis by Darwin: closely related species are more likely to compete. It has recently received experimental support [11, 28]. A strong assumption of trait conservatism underlies the model [14]. Here α is fixed, but it could vary in time and space. For example, smaller regions could have higher α to account for a lower carrying capacity. At each time step, all populations have the same probability e of extinction. Speciation occurs

-
- [1] T. G. Barraclough and A. P. Vogler. Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.*, 155(4):419–434, 2000.
 - [2] J. A. Coyne and H. A. Orr. *Speciation*. Sinauer Associates, 2004.
 - [3] M. R. T. Dale and M.-J. Fortin. From graphs to spatial graphs. *Annu. Rev. Ecol. Evol. Syst.*, 41:21–38, 2010.
 - [4] P. Desjardins-Proulx and D. Gravel. How likely is speciation in neutral ecology? *Am. Nat.*, 179(1):137–144, 2012.
 - [5] E. P. Economo and T. H. Keitt. Species diversity in neutral metacommunities: a network approach. *Ecol. Lett.*, 11:52–62, 2008.
 - [6] E. P. Economo and T. H. Keitt. Network isolation and local diversity in neutral metacommunities. *Oikos*, 10:1–9, 2010.
 - [7] B. K. Epperson, B. H. McRae, K. Scribner, S. A. Cushman, M. S. Rosenberg, M. J. Fortin, P. M. A. James, M. Murphy, S. Manel, P. Legendre, and M. R. T. Dale. Utility of computer simulations in landscape genetics. *Mol. Ecol.*, 19:3549–3564, 2010.
 - [8] D. P. Faith. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.*, 61:1–10, 1992.
 - [9] I. Hanski. *Metapopulation ecology*. Oxford University Press, 1999.
 - [10] I. Hanski. The theories of island biogeography and metapopulation dynamics. In J. B. Losos and R. E. Ricklefs, editors, *The Theory of Island Biogeography Revisited*, pages 186–213. Princeton University Press, 2009.
 - [11] L. Jiang, J. Tan, and Z. Pu. An experimental test of darwin’s naturalization hypothesis. *Am. Nat.*, 175:415–423, 2010.
 - [12] K. Kobayashi. On the spatial graph. *Kodai Math. J.*, 17(3):511–517, 1994.
 - [13] R. Levine. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.*, 15:237–240, 1969.
 - [14] J. B. Losos. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.*, 11:995–1003, 2008.
 - [15] J. B. Losos. *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, 2009.
 - [16] J. B. Losos and R. E. Glor. Phylogenetic comparative methods and the geography of speciation. *Trends Ecol. Evol.*, 18(5):220–227, 2003.
 - [17] J. D. Lynch. The gauge of speciation: on the frequency of modes of speciation. In D. Otte and J. A. Endler, editors, *Speciation and its consequences*, pages 527–553. Sinauer Associates, 1989.
 - [18] E. Mayr. *Systematics and the origin of species*. Columbia University Press, New York., 1942.
 - [19] E. Mayr. *Animal Species and Evolution*. Belknap, Cambridge, MA., 1963.
 - [20] E. S. Minor and D. L. Urban. Graph theory as a proxy for spatially explicit population models in conservation planning. *Ecol. Appl.*, 17(6):1771–1782, 2007.
 - [21] E. S. Minor and D. L. Urban. A graph theory framework for evaluating landscape connectivity and conservation planning. *Conservation Biology*, 22(2):297–307, 2008.
 - [22] N. Mouquet, V. Devictor, C. Meynard, L. F. Bersier, J. Chave, P. Couteron, A. Dalecky, C. Fontaine, D. Gravel, O. Hardy, F. Jabot, S. Lavergne, M. Leibold, T. Muenkemueller, S. Pavoine, A. Prinzing, A. Rodrigues, R. Rohr, and E. Thebault. Phylogenetic ecology: advances and perspectives. *Biol. Rev.*, 2012.
 - [23] M. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
 - [24] M. Penrose. *Random geometric graphs*. Oxford University Press, 2003.
 - [25] R. E. Ricklefs. Disintegration of the ecological community. *Am. Nat.*, 172(6):741–750, 2008.
 - [26] R. Sedgewick. *Algorithms in C++ Part 5: Graph algorithms*. Addison-Wesley Professional, 3rd edition, 2001.
 - [27] D. L. Urban, E. S. Minor, E. A. Treml, and R. S. Schick. Graph models of habitat mosaics. *Ecol. Lett.*, 12:260–273, 2009.
 - [28] C. Violle, D. R. Nemergut, Z. Pu, and L. Jiang. Phylogenetic limiting similarity and competitive exclusion. *Ecol. Lett.*, 14(8):782–787, 2005.
 - [29] A. R. Wallace. On the monkeys of the amazon. *Proc. Zool. Soc. Lond.*, 20:107–110, 1852.
 - [30] A. R. Wallace. On the law which has regulated the introduction of new species. *Ann. Magazine Nat. Hist. Ser. 2*, 16:184–196, 1855.