UNIVERSIDADE DE BRASÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOLOGIA

Especiação, diversidade genética e disparidade morfológica de um grupo de espécies crípticas, *Gymnodactylus* gr. *amarali* Barbour, 1925

Izabel Cristina dos Santos Câmara Salvi

Brasília
Agosto/2024

UNIVERSIDADE DE BRASÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOLOGIA

Especiação, diversidade genética e disparidade morfológica de um grupo de espécies crípticas, *Gymnodactylus* gr. *amarali* Barbour, 1925

Izabel Cristina dos Santos Câmara Salvi

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Zoologia, Instituto de Ciências Biológicas, da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Mestre em Zoologia.

Orientadora: Julia Klaczko
Co-orientadora: Flávia Maria Darcie Marquitti

Brasília
Agosto/2024

Especiação, diversidade genética e disparidade morfológica de um grupo de espécies crípticas, *Gymnodactylus gr. amarali* Barbour, 1925

Izabel Cristina dos Santos Câmara Salvi
Julia Klaczko
Flávia Maria Darcie Marquitti

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Zoologia, Instituto de Ciências Biológicas, da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Mestre em Zoologia.

Banca Examinadora
Julia Klaczko, Presidente – UnB
Sabrina Borges Lino Araujo - UFPR
Lilian Gimenes Giugliano - UnB
Reuber de Albuquerque Brandão (Membro suplente) - UnB

Brasília
Agosto/2024

À minha mãe Roselane, meu pai André
e meus irmãos de coração Amanda, Felipe e Gabriel.

AGRADECIMENTOS

A primeira grande diferença que senti saindo da graduação para o mestrado foi a solidão. Os momentos em grupo, as aulas, as discussões diminuem, e os momentos de introspecção, leitura, pesquisa aumentam. O mestrado que fiz não precisava de coletas, de trabalho de bancada, desde início era eu e meu computador. Essa solidão poderia ter perdurado, mas com o tempo, diversas pessoas foram se acumulando no meu caminho e um período que poderia ser de apenas frustrações e dificuldades, se tornou um ano e meio que irei lembrar carinhosamente para sempre. Gostaria, então, de agradecer a todas essas pessoas que me fizeram companhia e me ajudaram durante o mestrado.

Gostaria de agradecer primeiramente meu amigo Leonardo, quem conheço desde o início da graduação e caminhou todo esse percurso comigo. Uma pessoa tão envolvida na pesquisa que, independente das nossas áreas de trabalho terem uma intersecção pequena, ele sempre estava ativamente me ajudando, me instigando, interessado no meu trabalho e perguntando do modelo “neutre”. Ter você por perto me inspira a ser uma cientista melhor.

Outra pessoa que sofreu junto comigo durante vários momentos do mestrado foi minha amiga Isabela. Passamos juntas pelo sofrimento de ter que reaprender cálculo, o sofrimento de aplicar para universidade fora e muitos outros momentos de angústia que foram mais leves pois tive você do meu lado.

Gostaria de agradecer a todos do orfanato do LIS. Julia, Veida, Rayssa, Arthur, Isabelle, Brenda e Luiz. Era uma dificuldade enorme trabalhar com todos vocês lá, mas valia sempre a pena. Acho que nunca presenciei um ambiente de trabalho que me trazia tanta felicidade (mesmo não sendo oficialmente do laboratório) e acredito que nunca mais vou vivenciar outro ambiente como o LIS. Vou sentir muita falta de todos vocês. Do Arthur escutando música alta enquanto mede peixe e briga com a Julia, da Julia que, por sua vez, estaria sofrendo por trabalhar com raias enquanto discute com o Arthur, da Veida completamente focada escrevendo a dissertação no meio da bagunça, da Rayssa com seu facão e cigarro de bolinha, do Luiz aparecendo para tomar café e só, da Isabelle fazendo palavra cruzada. Espero que meu próximo laboratório seja um por cento do que vocês são. Obrigada professora Verônica por ter permitido esse espaço, e por estar sempre de portas abertas para papos, choros, reclamações e fofocas.

Gostaria de agradecer ao Fabrícius pelos dados e por todo apoio e sugestões durante o processo. A todos da Coleção Herpetológica da UnB (CHUNB), Carol, Malu, Humberto, Lala e Ticiane, que estavam sempre dispostos a me ajudar, e quebraram meu galho muitas vezes. Ao professor Guarino, que não só abriu as portas do seu laboratório para mim, disponibilizou seus dados, mas também me deu a oportunidade de ir a campo, conhecer meu objeto de estudo e apreciar toda uma parte que eu possivelmente teria negligenciado, a parte de coleta em campo.

Minhas orientadoras, Flávia e Julia. Não poderia ter sonhado por orientadoras melhores. Fiquei muito mal-acostumada. Queria agradecer a paciência, calma e energia. Adorei todas as reuniões que tive com vocês, até as que eu levava resultados temerosos que levavam horas para desvendar. Aprendi demais com ambas e espero poder continuar trabalhando com vocês durante muito tempo.

Kauê, Amanda, Camilla, Nicole, Thales, Julia, Matheus, todas as pessoas que conheci em São Paulo nos seis meses mais intensos da minha vida. A dor da saudade dos nossos momentos juntos ainda não foi embora, mas é uma saudade boa, que vem com ela memórias maravilhosas.

Meu namorado Benjamin. Uma das pessoas que mais me incentiva na minha carreira científica. Passei por muitos momentos em que eu não saia do lugar, que eu não acreditava mais em mim, que não me sentia o suficiente. E você teve o trabalho árduo de me convencer o contrário em todos eles. Sou muito grata pelo seu apoio durante todo o mestrado. Seu envolvimento e paixão por ciência é contagiante. Nunca vi um físico tão animado aprendendo sobre nomenclatura de espécies. Falar de ciência com você, desde aquele dia na lavanderia que você me contou sobre Placozoa pela primeira vez, me faz lembrar todos os dias porque escolhi essa carreira.

Quero agradecer à minha família, Tia Lina, Dinda, Flávio, Amanda, Felipe e Gabriel, vocês são minha vida inteirinha. E meus pais, Roselane e André. Não existem palavras para expressar a gratidão que eu tenho de ter nascido nesta casa. Tenho muita sorte de estar num ambiente de tanto amor, com pessoas que me dão os melhores exemplos para tudo na vida. Tudo que eu tenho, tudo que eu fiz, tudo que eu sou é por causa de vocês. Obrigada pela paciência comigo durante meu processo de aplicação do doutorado e agora na finalização da escrita da dissertação. Amo vocês demais.

Por fim, gostaria de agradecer à Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) pela bolsa.

SUMÁRIO

RESUMO

Especiação, diversidade genética e disparidade morfológica de um grupo de espécies crípticas, *Gymnodactylus gr. amarali* Barbour, 1925

Izabel Cristina dos Santos Câmara Salvi
Julia Klaczko
Flávia Maria Darcie Marquitti

Resumo da Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Zoologia, Instituto de Ciências Biológicas, da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Mestre em Zoologia.

Delimitar e descrever os processos que levam à formação de diferentes padrões é uma das questões centrais na biologia evolutiva. Com o avanço das novas tecnologias, observamos cada vez mais a descrição de uma diversidade antes oculta devido à falta de variação morfológica, que provavelmente é resultado de processos não adaptativos. Durante meu mestrado busquei entender mecanismos que levam à formação de novas espécies sem variação morfológica evidente, bem como a formação de diversidade por processos não adaptativos, utilizando tanto dados empíricos quanto simulações. No primeiro capítulo, busquei identificar se existia alguma estruturação morfológica nas espécies crípticas de uma lagartixa do Cerrado, *Gymnodactylus gr. amarali* Barbour, 1925, que se diversificou durante o Quaternário. Para tanto, utilizei métodos de *machine learning* em dados de medidas e folidose (contagem de escamas). Por mais que as médias e medianas de algumas características morfológicas variem entre os grupos (populações e espécies) as suas distribuições se sobrepõem, impossibilitando que os caracteres sejam usados para diagnosticar as espécies. A fim de identificar os diferentes fatores que levaram ao isolamento, analisei se o clima e a distância geográfica explicavam a alta variação genética e a diversidade morfológica dentro do grupo. Os resultados apontam para uma clara relação entre variação genética e clima e distância geográfica, enquanto a variação morfológica inconspícua estaria relacionada apenas ao clima. Inspirada em *Gymnodactylus gr. amarali* e na sua diversificação no Quaternário, para o segundo capítulo, utilizei simulações baseadas no modelo de Derrida-Higgs para entender se a dinâmica de contração e expansão das áreas de distribuição, resultantes das flutuações climáticas do Quaternário, deixava marcas nos padrões macroevolutivos de diversificação. Nossos resultados mostraram que essas dinâmicas não levaram a um aumento na formação de novas espécies. No entanto, a existência de múltiplos refúgios durante períodos de contração levou a formação de filogenias balanceadas. Esses resultados não corroboram o padrão encontrado em *Gymnodacylus* gr. *amarali* o que pode ser devido a homogeneidade dos refúgios definidos em nosso modelo. Estudos futuros devem incluir refúgios de tamanho e distância diferentes.

Palavras-chave: seleção não adaptativa, Derrida-Higgs, herpetologia, padrões macroevolutivos, refúgios, quaternário.

Brasília

Agosto/2024

ABSTRACT

Speciation, genetic diversity, and morphological disparity in a group of cryptic species, *Gymnodactylus* gr. *amarali* Barbour, 1925

Izabel Cristina dos Santos Câmara Salvi
Julia Klaczko
Flávia Maria Darcie Marquitti

Abstract da Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Zoologia, Instituto de Ciências Biológicas, da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Mestre em Zoologia.

Defining and describing the processes that lead to the formation of different patterns is one of the central issues in evolutionary biology. With the advancement of new technologies, we increasingly observe the description of a diversity that was previously hidden due to a lack of morphological variation, which is likely the result of non-adaptive processes. During my master's degree, I sought to understand mechanisms that lead to the formation of new species without obvious morphological variation, as well as the formation of diversity through non-adaptive processes, using both empirical data and simulations. In the first chapter, I aimed to identify whether there was any morphological structure in the cryptic species of a lizard from the Cerrado, *Gymnodactylus* gr. *amarali* Barbour, 1925, which diversified during the Quaternary. To this end, I used machine learning methods on data of measurements and pholidosis (scale count). Although the means and medians of some morphological traits vary between groups (populations and species), their distributions overlap, making it impossible to use these traits to diagnose species. To identify the different factors leading to isolation, I analyzed whether climate and geographic distance explained the high genetic variation and morphological diversity within the group. The results indicate a clear relationship between genetic variation and climate and geographic distance, while inconspicuous morphological variation would be related only to the climate. Inspired by *Gymnodactylus* gr. *amarali* and its diversification in the Quaternary, for the second chapter, I used simulations based on the Derrida-Higgs model to understand if the dynamics of contraction and expansion of distribution areas, resulting from Quaternary climatic fluctuations, left marks on macroevolutionary diversification patterns. Our results showed that these dynamics did not lead to an increase in the formation of new species. However, the existence of multiple refugia during contraction periods led to the formation of balanced phylogenies. These results do not corroborate the pattern found in *Gymnodactylus* gr. *amarali*, which may be due to the homogeneity of the refugia defined in our model. Future studies should include refugia of different sizes and distances.

Keywords: *Non-adaptive selection, Squamata, Derrida-Higgs model, macroevolutionary patterns, Quaternary*.

Brasília
August/2024

INTRODUÇÃO GERAL

No final da *Origem da Espécies*, Darwin sintetiza o processo evolutivo e a biodiversidade como “infinitas formas de grande beleza”. Entender os precursores da visível disparidade, da diversidade de espécies, comportamentos e formas é o que impulsiona os grandes trabalhos dentro da biologia evolutiva (Grant & Grant, 1989; Losos *et al.*, 2006). Entretanto, ao olhar para o lado oposto, para falta de variação, qualquer que seja, e entender diferentes processos que podem culminar nesse padrão, ampliamos o entendimento do funcionamento da natureza em diferentes escalas de tempo e espaço (Adams *et al.*, 2009) .

A formação da biodiversidade pode se dar de duas maneiras, sendo o resultado de um processo adaptativo ou não adaptativo. Historicamente, o processo de formação de novas espécies remete a processos como radiação adaptativa (Losos & Mahler, 2010), no qual existe um claro acoplamento entre a formação de espécies e o surgimento de disparidade morfológica (Gould & Lewontin, 1979). No entanto, com o aumento e diversificação de estudos e técnicas moleculares, cada vez mais temos observado um desacoplamento tanto nos processos, com exemplos de taxas de especiação desacopladas às taxas de disparidade morfológica (Adams *et al.*, 2009; Ramírez-Reyes *et al.*, 2022; Missagia *et al.*, 2023), quanto nos padrões, com grupos apresentando alta variação genética e baixa variação morfológica (Domingos *et al.*, 2014; Olave *et al.*, 2017; Singhal *et al.*, 2018). O estudo desse desacoplamento para além dos padrões gerados (espécies crípticas ou alta variação morfológica intraespecífica), nos permite descrever novos processos que levam a formação de diversidade (Nosil, 2008; Fitzpatrick *et al.*, 2008; de Aguiar *et al.*, 2009).

Podemos nos perguntar: quando removemos adaptação e variação morfológica, o que sobra? Quais processos podem explicar uma falta de diversidade morfológica, e quais padrões são formados por processos não adaptativos? Sendo assim, meu trabalho se divide em duas linhas: a primeira consiste no estudo de caso de um grupo de espécies crípticas de lagartixas do norte do Cerrado, *Gymnodactylus* gr. *amarali* Barbour, 1925. Analisei os possíveis fatores que levam ao isolamento dessas espécies, e explicariam a falta de variação morfológica. Inspirada no possível processo de diversificação de especiação de *Gymnodactylus* gr. *amarali*, a segunda parte busca entender, por meio de simulações, o efeito de processos topográficos não adaptativos que ocorreram durante o momento das diversificações do grupo estudado no capítulo 1 - oscilações climáticas e ciclos glaciais - na formação de espécies.

CHAPTER 1: Climatic selective pressures responsible for morphological structure of crepuscular rock dwelling gecko cryptic species complex, *Gymnodactylus* gr. *amarali* Barbour, 1925 (Gekkota: Phyllodactylidae)

Submitted to the *Zoological Journal of the Linnean Society*.

Authors: Izabel Salvi, Fabricius M. C. B. Domingos, Guarino R. Colli and Julia Klaczko.

Abstract

Cryptic species can be a challenge to study. Not only finding and delimiting them can be complex, due to their inherent morphological resemblance, but also understanding the processes and evolutionary pressures that culminated in high genetic variation with low morphological disparity may not be straightforward. In our work, we aimed at understanding what could be maintaining the high genetic variation of a crepuscular gecko cryptic species group from the northern Brazilian Savanna (Cerrado), *Gymnodactylus* gr. *amarali* Barbour, 1925 (Gekkota: Phyllodactylidae), in the absence of clear allopatric barriers, and also which factors could explain their extremely conserved morphologies. We used tree-based machine learning models to identify possible morphological characteristics that vary between groups, and tested if climatic parameters or geographic distance could explain the genetic variation and morphological patterns. We found that both isolation by adaptation and isolation by distance are responsible for maintaining the isolation of the groups within *Gymnodactylus* gr. *amarali*, and there is some morphological variation between groups that is shaped by climatic factors, but the range overlap of those traits between different populations/species are high enough that we cannot use them to diagnose the different species.

Keywords: *isolation by distance, isolation by adaptation, speciation, Squamata, microevolution.*

Introduction

In the past couple of decades, with the improvement and dissemination of molecular sequencing, we have seen an increase in descriptions of a so far hidden biodiversity (Bickford *et al.*, 2007; Fontaneto *et al.*, 2009; Domingos *et al.*, 2017; Struck *et al.*, 2018; Tessler *et al.*, 2022) . Coalescent-based species delimitation methods (Domingos *et al.*, 2014, 2017; Olave *et al.*, 2016) and new genomic technologies (Domingos *et al.*, 2017; Tessler *et al.*, 2022) allowed to encounter a series of cryptic species that would not be found using traditional morphology-based methods only.

Although cryptic species have been commonly defined as a group of two or more species incorrectly assigned to one formal taxonomic name (i.e., a single species) (Bickford *et al.*, 2007), this concept can lead to different operational problems as it is not accurately based on biological parameters. Moreover, the concept of what a species actually represents is not strictly defined or agreed upon in the specialized literature (de Queiroz, 2007; Salles & Domingos, 2023). Considering species under de Queiroz's unified concept (de Queiroz, 2007), a more precise and operative delimitation of cryptic species would be groups with clear genetic separation, representing different degrees of reproductive isolation (Stankowski *et al.*, 2024), with overlapping morphological clusters of the different genotypes (Struck & Cerca, 2019).

The study of cryptic species benefits from going beyond the description of patterns to provide analysis of the underlying evolutionary processes (Struck *et al.*, 2018, p. 208; Singhal *et al.*, 2018; Marques-Souza *et al.*, 2020; Kordbacheh *et al.*, 2023). By examining both genetic structure and morphological variation among groups revealed through the use of genetic data, we can elucidate possible processes and selective pressures that culminated in the lack of external morphological diversity (Wojcieszek & Simmons, 2012, 2013; Singhal *et al.*, 2018). This can help us understand what factors contribute to the emergence of morphological diversity and what may inhibit it, even when there is a substrate of high genetic variability (Struck *et al.*, 2018). For instance, an apparent morphological stasis could result from stabilizing selection due to environmental pressures (Smith *et al.*, 2011; Olave *et al.*, 2016; Singhal *et al.*, 2018; Zhao *et al.*, 2019) or constrained adaptation caused by pleiotropic genes (Hansen & Houle, 2004). Species have also been shown to be cryptic not only as the result of stabilizing selection but of directional selection for characteristics that are imperceptible to the human eye, or at least inconspicuous (Henry, 1994; Fišer *et al.*, 2018; Zozaya *et al.*, 2019).

Understanding the processes that maintain these species genetically distinct without apparent morphological and, sometimes, even ecological differences, without clear allopatric barriers, is not straightforward. While considering the isolation by distance (IBD) and isolation by adaptation (IBA) dichotomy (Nosil, 2012), both could lead to the formation of cryptic species (Singhal *et al.*, 2018; Tonzo *et al.*, 2019). The first process, IBD, can lead to drift in the morphological characters, resulting in no apparent interspecific morphological structure. The second, IBA, has been found in different groups that lack interspecific morphological variability (Singhal *et al.*, 2018; Tonzo *et al.*, 2019), though the reasons for that are unclear. Less commonly studied processes related to post or pre-zygotic barriers, such as genitalia incompatibility (Nunes *et al.*, 2012), pheromones (Zozaya *et al.*, 2019) and allochrony (Taylor & Friesen, 2017), can maintain species' reproductive isolation even in the absence of clear spatial isolation.

The taxon that currently comprises the name of the Neotropical lizard *Gymnodactylus* gr. *amarali* Barbour, 1925 would be considered a group of cryptic species by taxonomic and genetic concepts. Since the initial description in 1925, these animals were never split into different species until molecular evidence suggested otherwise (Domingos *et al.*, 2014), although other species have been synonymized into *G. amarali* (casemiro). They are rock-dwelling geckos with crepuscular habits and specialized insectivorous feeding, predominantly composed of termites (Vitt *et al.*, 2007). Their populations are distributed throughout the northern Brazilian Savanna (Cerrado) in regions of varying altitudes (valleys and plateaus) (Colli *et al.*, 2003). Molecular analyses using coalescent based methods revealed eight (Domingos *et al.*, 2014) (or nine, see (Domingos, 2015)) cryptic species within *G. amarali*, with the majority of these species diversifying during the Quaternary period. Despite morphological variation among different populations/species of *G. amarali*, the variability found in external morphological characters does not recover the species found by molecular analyses.

Given that *Gymnodactylus amarali* is a group of possibly up to nine cryptic species with high genetic variation, inconspicuous phenotypic variation, and apparent niche conservation among all populations, the group is a great candidate to study the mechanisms and processes that promoted genetic diversity and constrained morphological variation during speciation. In this study, we aimed at investigating the patterns of interspecific morphological variability of the *Gymnodactylus amarali* species group and identify the possible factors determining the genetic and morphological structure of the group.

Knowing that these are ectothermic animals living at different altitudes, but all different populations are located in rocky environments, we may wonder if the existing climatic variation is enough of a factor to isolate different groups, or the opposite, if the ecological similarities between localities produces some type of stabilizing selection, and isolation is due to other extrinsic factor, such as distance. We hypothesize that (i) given that this are small animals with apparent restricted niches, geographic distance is enough to isolate the different populations/species, or that (ii) the existing climatic variation is enough to limit the dispersion of different popualtions/species, and inconspicuous morphological variation between groups may exists, especially in the scale patterns, and that variation is correlated to the existing climatic variation.

With the intention of finding any trait that could possibly differentiate the analyzed species/populations of *Gymnodactylus* gr. *amarali* we conducted tree-based classification analyzes. Moreover, to contrast our two hypotheses, we tested the different populations for isolation by distance (IBD) against isolation by adaptation (IBA), while also analyzing if the morphological traits could be explained by intrinsic (genetic) or extrinsic (climate, distance) factors.

Methods

Sampling

We analyzed specimens of *Gymnodactylus* gr. *amarali* from 10 different populations (Figure 1), encompassing 8 or 7 cryptic species according to Domingos *et al.* (2014), or Domingos (2015) (Table S1). All specimens are deposited at the 'Coleção Herpetológica da Universidade de Brasília' (CHUNB) (see electronic supplementary material, Table S2 for specimen details, coordinates, and vouchers)

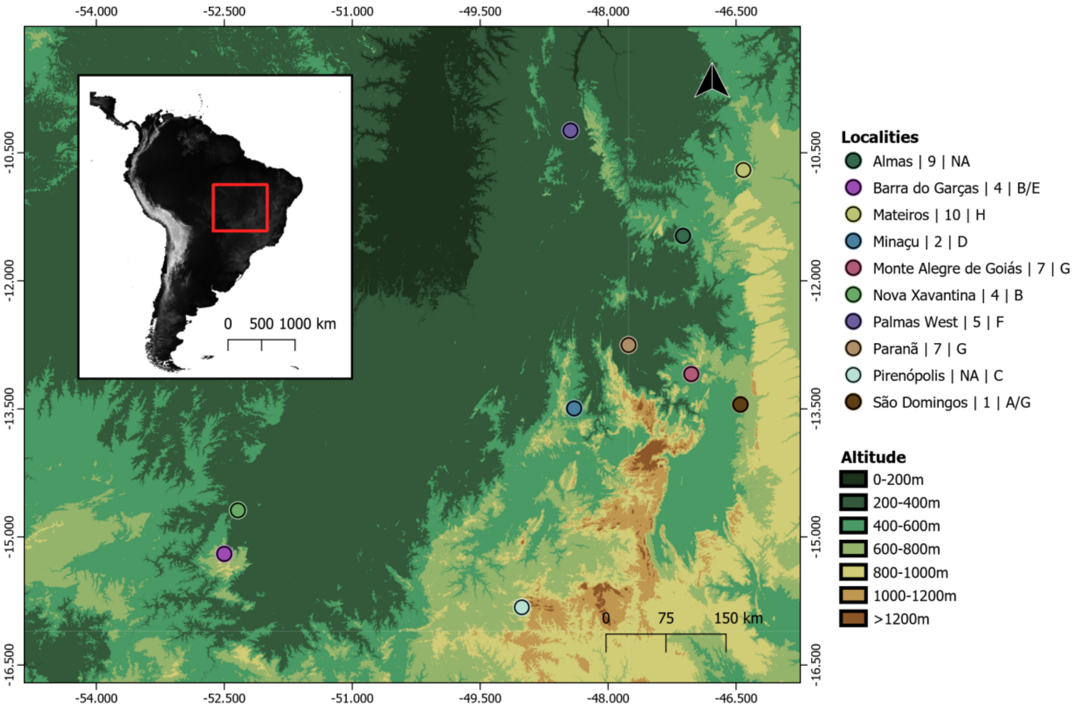


Figure 1. Map of populations analyzed in this study. On the side of each population, the name of the species present in each locality, from A – G according to Domingos *et al.* (2014), and from 1-10 according to Domingos (2015).

We will use the different populations as the unit of study, bearing in mind that they represent different species. See supplementary material for analysis using, instead of populations, the species previously delimited by Domingos *et al.* (2014) and Domingos (2015).

Morphological data

We analyzed eight morphometric and 29 pholidosis (scale) characters. Morphometric measurements were taken in the field with live specimens, except for the individuals from Monte Alegre and Almas, which were measured in the lab using electronic calipers. The measurements were snout-vent length (SVL), posterior member length (PML), anterior member length (AML), head length (HL), head height (HH), head width (HW), body height (BH), body width (BW). All measurements were log-transformed, and we used the residuals from a regression by SVL to size-correct them. All scale characters were acquired using a stereomicroscope (see Table S3, for details on scale measurements).

Molecular data

We used cytochrome b (cytb) sequences of 100 specimens (minimum 7 per population). For extraction methods, see (Domingos *et al.*, 2014). The sequences were aligned in MEGA11 using the ClustalW algorithm (Tamura *et al.*, 2021).

Climatic distance

The climatic data was extracted from the WorldClim database (Fick & Hijmans, 2017) with a 10km² resolution. Subsequently, we calculated climatic distance between populations as the Euclidean distance of the populations' scores on the first four principal components resulting from a PCA generated by all WorldClim climatic variables (Proportion of variance: PC1 = 52.64%, PC2 = 19.83%, PC3 = 14.3%, PC4 = 8.04%).

Geographical distance

We computed the shortest distance between population coordinates, considering existing topography to better reflect biological movement. We used the *topoDist* function from the R package *topoDistance* (Wang, 2020), along with a raster of Brazil's altitude from the R package raster (Hijmans, 2023).

Machine learning trait selection

To filter and identify the characteristics that best delineate groups within *G. amarali*, machine learning analyses were conducted using tree-based classification methods, *Bagging*, *Random Forest* with the R package *randomForest* (Breiman *et al.*, 2022) and *Generalized Boosted model*, an implementation of the AdaBoost algorithm, with the package *gbm* (Ridgeway & Edwards, 2024). The parameters for the three models were obtained through model tuning analyses performed with the *caret* package (Kuhn, 2008). We used 75% of the data to train the model and the rest to check the model's accuracy. We calculated the boosting accuracy via cross-validation. To confirm the non-important traits, we use a feature selection algorithm with the package *Boruta* (Kursa & Rudnicki, 2010).

F_{ST} and P_{ST}

With the cytb data, we calculated Rousset's F_{ST} (Rousset, 1997) using Weir and Cockerham pairwise θ (Weir & Cockerham, 1984). To estimate the morphological distances between populations, we separately calculated the pairwise P_{ST} for scale counts and morphometric measurements. We used the first 4 Principal Components from two PCAs, one for the measurements and another for the scale characters, and calculated the P_{ST} values with the package *Pstat* (Silva & Silva, 2018).

Isolation by distance and isolation by adaptation

To understand the factors responsible for the maintenance of genetic isolation, check if there is isolation by adaptation of isolation by distance, we did a Multiple Matrix Regression with Randomizations (MMRR) (Wang, 2013) with 10000 randomizations. Given that the different species and populations from the *Gymnodactylus* gr. *amarali* species group has the same type of habitat (Colli *et al.*, 2003), we accessed isolation by adaption using climatic distances only. For the first model (Genetic distance ~ geographic distance + climatic distance), we used F_{ST} to represent genetic distance. We also did two other models to understand the factors responsible for the observed morphological structure between populations (I. Morphological distance (morphometric measurements) ~ geographic distance + climatic distance + genetic distance, II. Morphological distance (scale counts) ~ geographic distance + climatic distance + genetic distance).

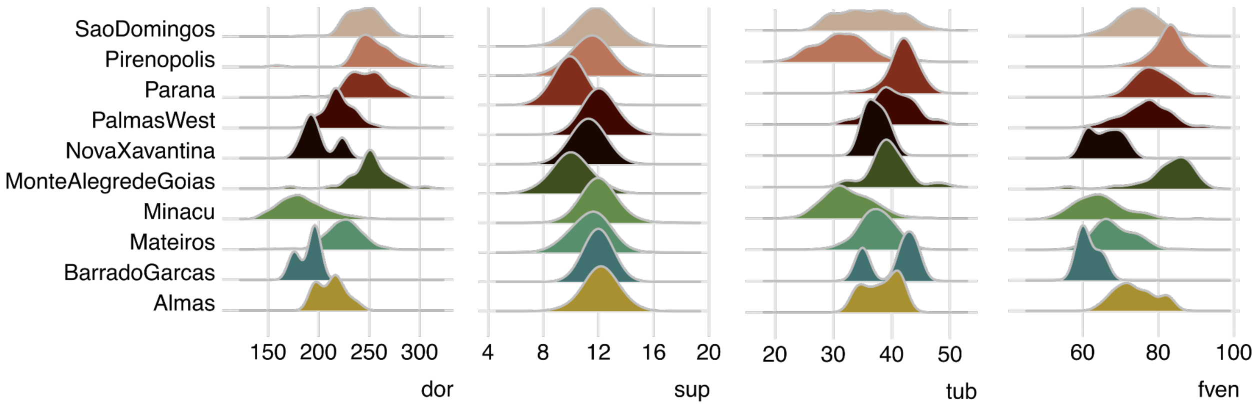
Results

Trait selection

The parameters given by the model tunning were *mtry* = 2 (random forest), *ntree* = 300 and *interaction depth* = 1 (boosting), *shrinkage* was held constant at 0.1. The models' accuracies were: Bagging = 0.8297872 (0.7574, 0.8878), RF = 0.8865 (0.8223, 0.933), Boosting = 0.9086823. The accuracy of the models was higher when we used the species as the groupings, instead of populations (see supp. material Table S4). Barra do Garças, Nova Xavantina, and Pirenópolis specimens were constantly misclassified (Figure S1). The Boruta analysis showed that all but three traits (Figure S2) were considered non-significant.

The traits that were consistently recovered as the most important to delineate the different populations were the number of dorsal tubercules (dor), number of supralabials (sup), number of longitudinal tubercles rows (ftub), number of transverse rows of ventrals (fven), HH, BW, SVL (Figure S3A and S3B). No single trait can delineate the different populations (Figure 2). Although we see different peaks for different populations/species, there is a high variation for every morphological character within the different populations/species in comparison with across populations/species. Body width was the only character that could differentiate species 2 Domingos (2015) (or D (Domingos *et al.* (2014)), from every other (Figure S4).

(A) Scale counts



(B) Morphometric measurements

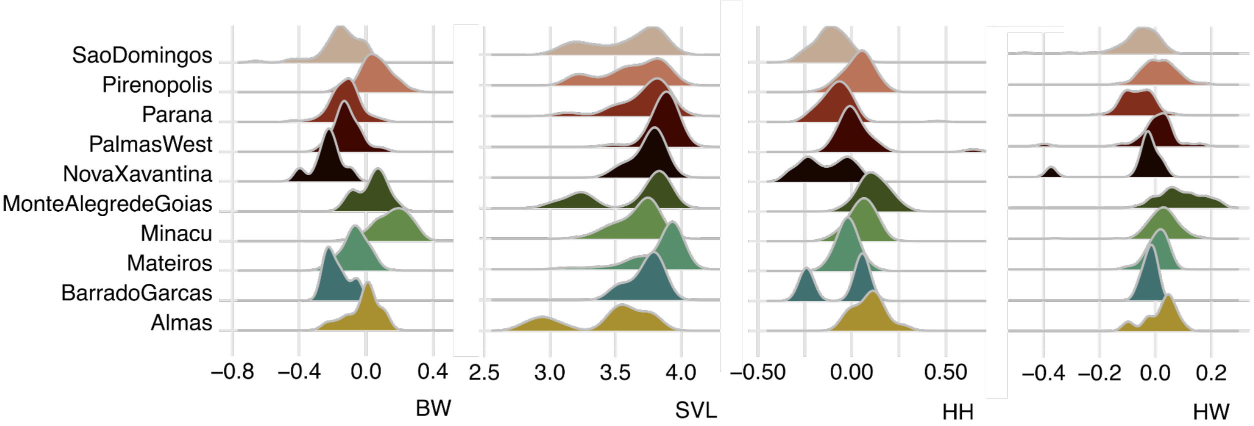


Figure 2. Distribution of the traits that were recurrently found as the most important for delimiting the populations in all three machine learning analyses for (A) scale counts (*dor*: number of dorsal

tubercules, *sup:* number of supralabials, *tub*: number of paramedian tubercules, *fven*: number of transverse rows of scales), and (B) morphometric measurements (*BW*: body width, *SVL:* snout.-vent length, *HH*: head height, *HW:* head width). For the distribution of these traits by species see supp. material Figure S4 and Figure S5.

IBD x IBA

The first four principal components of morphometric and scale count PCAs explained a total of 73.8% and 42.4% (Table S6) of the observed variation, respectively. The first model (genetic distance ~ geographical distance + climatic distance) was significant (Table 1), showing a correlation between genetic distance and both climatic and geographic distance.

Most morphological variables (PC1, PC3, and PC4 of morphometric data and PC2 of scale count data) exhibited a significant correlation with climatic distance. Specifically, the PC1 of the bioclimatic variables, encompassing overall temperature differences. The climate PC2 with most precipitation variables exhibited no significant relation with any dependent variable. The morphometric PC2 regressions and scale count PC1 and PC3 showed no significant correlation with any tested variable. Scale count PC4 were significantly correlated with genetic distance, albeit with a low R^2.

Table 1. Results of multiple regression analyses for genetic distance ~ geographical distance + climatic distance, and morphological distance (Morph.: morphometry Fol.: scale count) ~ geographical distance + climatic distance + genetic distance. Coefficients (c) and p-values for each term and R^2 of the equation are presented in the table. Significant principal components of climatic distance are indicated in parentheses.

	Geographic dist.		**Climatic dist.**		**Genetic dist.**		**R^2**
	c	***p***	**c**	***p***	**c**	***p***	
Fst	0.69	0.001	(PC1) -0.53	0.044			0.26
Morph.							
PC1	-	-	(PC1) 0.77	0.022	-	-	0.12
PC2	-	-	-	-	-	-	-
PC3	0.54	0.014	(PC1) -0.81	0.001	-0.34	0.017	0.49
			(PC4) 0.31	0.044			

PC4	-	-	(PC1) -0.42	0.025	-	-	0.17
Scales							
PC1	-	-	-	-	-	-	-
PC2	-	-	(PC1) -0.43 (PC3) 0.52	0.014 0.004	0.371	0.0003	0.54
PC3	-	-	-	-	-	-	-
PC4	-	-	-	-	0.304	0.027	0.09

Discussion

The problems in delineating some species morphologically are usually associated with some sort of conserved morphology among groups, which can be the result of a stabilizing selective pressure (Struck & Cerca, 2022). Although a lack of morphological variation can be associated with cryptic species, what we found in *Gymnodactylus* gr. *amarali* was not a lack of overall variation in the analyzed traits, but the presence of intraspecific variation. Looking at the distribution of the traits recurrently found as the most important for delimiting the populations (Figure 2) and species (Figure S4 and S5) in the machine learning analysis, we can see that in some cases, the means and medians of the distribution varies between populations/ species but with clear overlaps.

One trait that is worth taking a closer look is body width. Body width separated species 2 (Domingos, 2015) from all species, both peak and overall distribution. Species 2 (Minaçu population) contain groups that live both in mainland and in artificial islands from the creation of a dam almost 30 years ago. Eloy de Amorim *et al.* (2017) found a displacement in eating habits between individuals from island and mainland, with island individuals eating bigger prey. Difference was also found in the head length, which is correlated with the content found in their stomachs (bigger the head, bigger the stomach). Although we did not test it, the increase in body size could be related to this recently found food ecology niche for the individuals in the island groups.

The machine learning models showed that a combination of the observed characteristics is capable of predicting different species/ populations (Figure S1) with low accuracy errors associated with some groups, but not all. Some species/populations presented between 80 – 100% accuracy error across all models. Moreover, looking at the traits, even the best-defined morphological groups present fuzzy boundaries, which is congruent to our results that factors other than the genetic variation is structuring the morphology of these groups.

We found that most morphological traits varied with climate. We hypothesized that scale count traits could be structured by climatic factors. Previous studies have demonstrated a relationship between scale patterns, particularly scale size, and adaptation to different climates (Calsbeek *et al.*, 2006; Wegener *et al.*, 2014) Squamates with larger scales would have less interstitial space where water could evaporate, and the bigger scale could irradiate more heat than smaller ones. We did not see this kind of pattern in our analyses. Only PC2 showed correlation with climatic factors, and the best model showed a relation with also genetic distance. Other scale count principal components showed no correlation to any other bioclimatic variables.

Morphometric measurements, however, were also found to change with climate. Bergmann's rule states the existence of a macroevolutionary correlation between latitudinal clines and body size, with bigger endothermic animals in higher latitudes, and vice-versa (Meiri, 2011). Although heavily discussed and disproven in multiple groups (Adams & Church, 2008), when looking at intraspecific variation, or at closely related taxons, several taxa found this correlation, even in ectotherms (Partridge & French, 1996; Zamora-Camacho *et al.*, 2014). Moreover, body size could be an important factor determining the thermal limit in which a lizard could live (Claunch *et al.*, 2021). We also found that not only overall body size varied with climate, but also morphological measurements and overall body shape proportions varied, but the reason for that correlation is unclear.

Our results showed that both climate and geographic distance are enough to keep the different species isolated. This combination of IBA and IBD has been found in other reptile groups (Singhal *et al.*, 2018) and other animals (for examples see Nosil *et al.*, 2009)**.**

We can understand what keeps the species isolated, but we cannot know for sure what processes took place that lead to cryptic speciation. Nevertheless, knowing what maintains them isolated now, and which factors structure their morphology, we can access different possible processes that could have happened that lead to the formation of our cryptic species.

One common explanation for cryptic species is that they are young. Most cladogenesis within *Gymnodactylus* gr. *amarali* occurred in the last 2.5 million years, but studies show that this relation between speciation time and morphological distinctiveness is not always concrete, with plenty of different cryptic species having diversified at the same time scale as non-cryptic species (Fišer *et al.*, 2018). Moreover, this correlation between the emergence of morphological disparity with time, usually access traits with neutral Brownian morphological evolution. We saw that our morphological data change with climate, and possibly there is an adaptation factor that causes the morphological variation we see. Adaptive processes have shown to be relatively fast (Harris *et al.*, 1998; van Rijssel *et al.*, 2021; Crouch & Tobias, 2022), and also, Eloy de Amorim *et al.* (2017) found evidence of rapid adaptation in one of the populations we analyzed.

A stabilizing selection process could lead to conservative morphologies and could be a result of phylogenetic niche conservatism (Cerca *et al.*, 2020). All our populations are from a savanna biome, and all individuals live near rocky environments. This could indicate homogeneous environments and climate which would inflict the same selective pressures on all analyzed groups. Nonetheless, our results showed that not only were the environments not completely homogeneous, as there was a small yet significant climatic variation between populations; and this variation could be a key factor isolating the different populations and shaping the analyzed morphological structure.

Another possible cause is directional selection of an inconspicuous trait. A trait we did not analyze could vary according to the genetic variation. Traits that are associated with sexual selection or prezygotic barriers can be the definitive factor leading to the isolation of different sister species such as genitalia (Nunes *et al.*, 2012) and pheromones (Zozaya *et al.*, 2019). Coelho and Klaczko (2020) found differences in genitalia shape between *G. amarali* species, but their studies did not take into consideration intraspecific variation.

Conclusion

We saw that a combination of climatic factors and geographic distance is enough to genetically structure the different populations. Since isolation by geographic distance does not have an adaptive nature, the morphological variation not related to climate could be just the result of drift. That results in unstructured morphological variation which could lead to the formation of cryptic species. But that does not explain all genetic/morphological variation. Climate is an important factor, given that these are small ectothermic animals, and the climatic microvariation, possibly related to the verticality of the population distribution, combined with geographic distance, is enough to isolate the different populations.

CHAPTER 1: Supplementary material

Table S1. Number of individuals used for morphometric, scale counts and genetic analysis, by population and species. We assigned each specimen species based on their population, using the species delimited by Domingos *et al*. (2014) and Domingos (2015).

Species (Domingos, 2015)	**Species (Domingos, 2014)**	**Population**	**Genetic**	**Scale count**	**Morphometric measurements**
9	NA	Almas	7	16	16
4	B/E	Barra do Garça	7	3	5
10	H	Mateiros	11	179	106
2	D	Minaçu	11	182	184
7	G	Monte Alegre	10	36	35
4	B	Nova Xavantina	11	8	8
5	F	Palmas West	7	44	33
7	G	Paranã	13	85	83
NA	C	Pirenópolis	13	54	52
1	A/G	São Domingos	10	106	66

Table S1. Coordinates of different localities.

Locality	UF	Latitude	Longitude
Almas	TO	-11.4737	-47.1211
Barra do Garcas	MT	-15.2	-52.5
Mateiros	TO	-10.7022	-46.4128
Minacu	GO	-13.4958	-48.3974
Monte Alegre de Goiás	TO	-13.2	-47.1
NovaXavantina	MT	-14.69	-52.34
Palmas West	TO	-10.1891	-48.1085
Paranã	TO	-12.7531	-47.7591
Pirenópolis	GO	-15.826	-49.011
São Domingos	GO	-13.4498	-46.4481

Table S3. Scale measurements.

Abbreviation	**Description**
cantr	Number of scales in canthus rostralis, counted from post nasal to the eye
asup	Number of scales above and in contact with the supralabials, counted from frontonasal to last supralabial
ainf	Number of scales below and in contact with the infralabials, counted from mental to last infralabial
sup	Number of supralabials (sum of both sides)
inf	Number of infralabials (sum of both sides)
supra	Number of enlarged supraciliary scales
dor	Number of dorsal scales, counted from rostral scale to posterior margin of thigh (before tail)
fqui	Number of keeled scale rows in tail
qui	Number of keeled scales in one row in tail, counted in the third keeled scales row
tub	Number of paramedian tubercles, counted from tympanum to posterior margin of thigh (before tail)
ocel	Number of paramedian ocelli, counted in one row from rostral to posterior margin of thigh (before tail)
ocef	Number of longitudinal ocelli at midbody
ftub	Number of longitudinal tubercles rows at midbody
vem	Number of longitudinal rows of ventral scales at midbody
men	Number of scales between enlarged post mentals, in contact with mental
lam	Number of subdigital lamellae on fourth finger
fven	Number of transverse rows of ventral scales, counted from mental to cloaca
fem	Number of femoral and tibial ventral scale rows, counted from cloaca (start of thigh) to foot at mid part of the limb
lama	Number of subdigital lamellae on fourth toe
clo	Number of granule like scales from cloaca to first enlarged subcaudal
ane	Number of white bands in tail
nas	Relative size of post nasals in relation to supranasal
supn	Contact between supranasals
front	Alignment between frontonasals division and the incomplete suture of rostral
ftim	Ear opening shape
ptim	Ear opening position
odor	Dorsal ocelli
omem	Ocelli in limbs
anec	Bands in tail

Table S4. Parameters used for each model, resulted from the parameter tuning using the caret package. For boosting shrinkage was held constant at 0.1

	Random Forest	*Boosting*	
	mtry	**ntree**	**interaction depth**
Populations	2	300	1
Species (Domingos, 2014)	2	200	1
Species (Domingos, 2015)	4	150	5

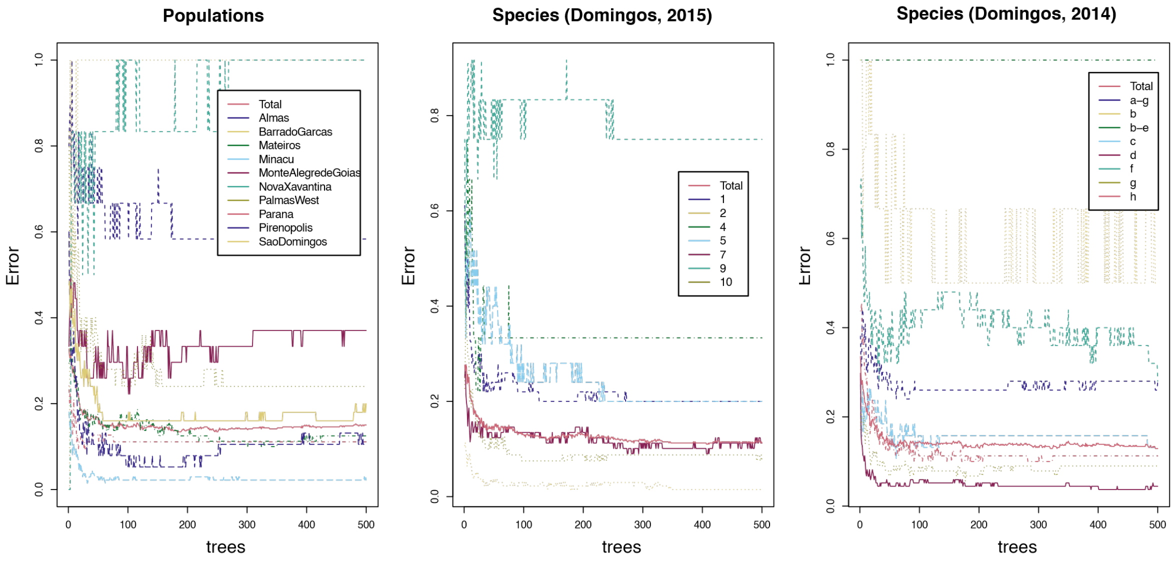
Table S5. Accuracy of each model, avg (min, max), resulted from the parameter tuning using the caret package. For boosting shrinkage was held constant at 0.1.

	Bagging	*Random Forest*	*Boosting (cv)*
Populations	0.8297872 (0.7574, 0.8878)	0.8865 (0.8223, 0.9337)	0.9086823
Species (Domingos, 2014)	0.9854 (0.9483, 0.9982)	0.9854 (0.9483, 0.9982)	0.9136446
Species (Domingos, 2015)	0.969 (0.9225, 0.9915)	0.9493 (0.9119, 0.9873)	0.9123208

Table S6. The proportion of morphological measurement variances and scale count principal components from their respective PCAs.

	PC1	**PC2**	**PC3**	**PC4**
Morphometric measurements	0.32	0.17	0.12	0.11
Scale count	0.16	0.11	0.07	0.06

(A) Bagging



(B) Random Forest

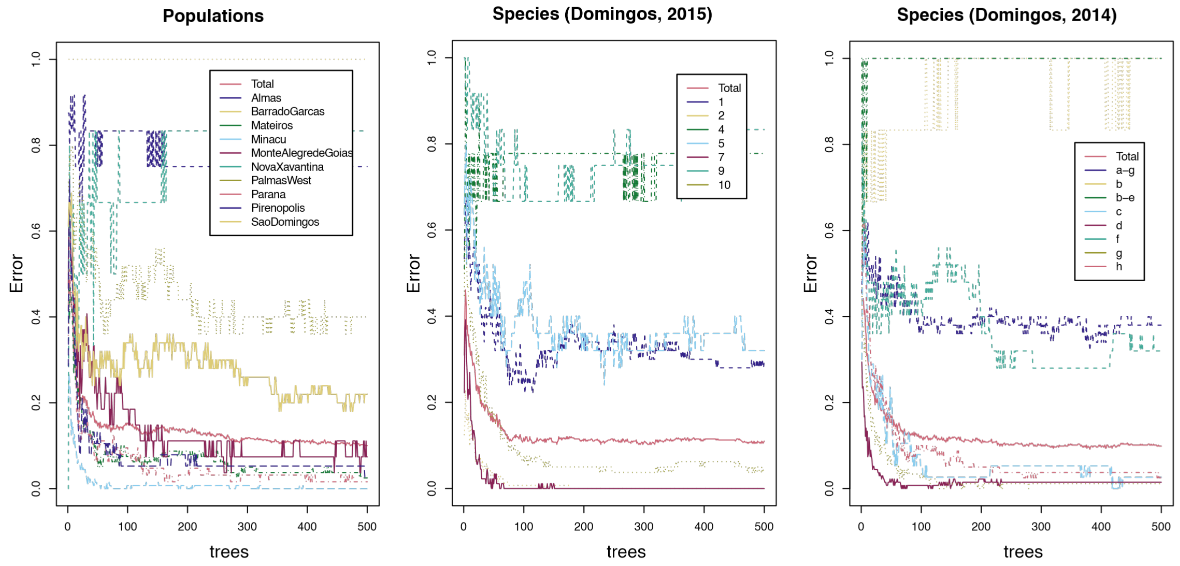
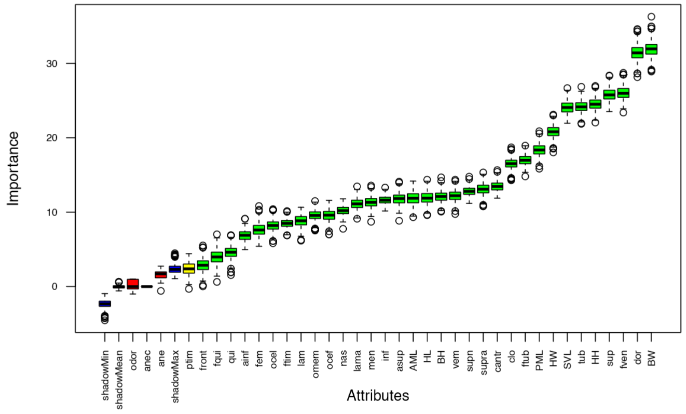
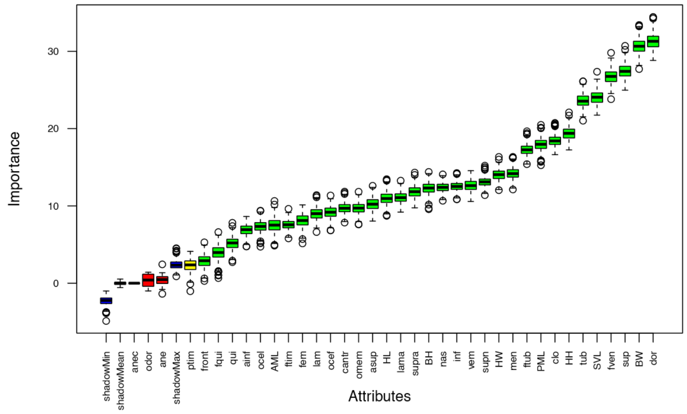


Figure S1. Out of bag error of the different machine learning models (A – Bagging, B – Random Forest) for each population or species.

(A) Populations



(B) Species (Domingos, 2014)



(C) Species (Domigos, 2015)

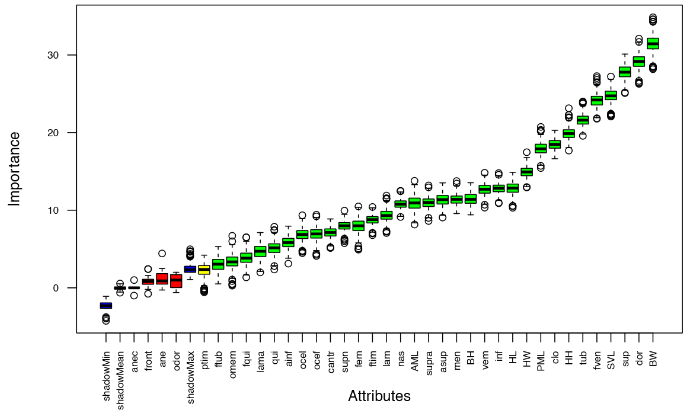
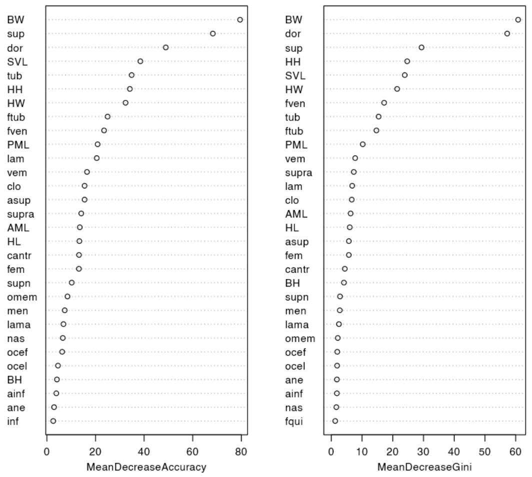


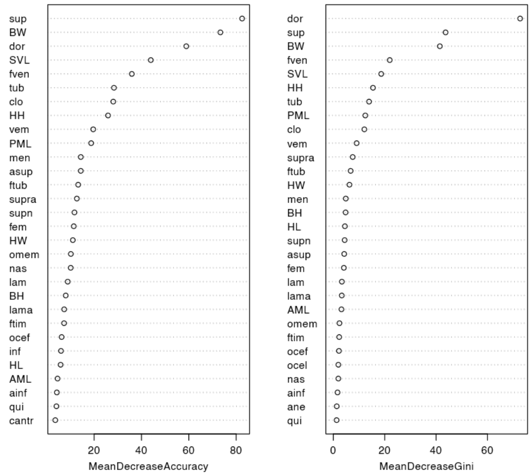
Figure S2. Results of the *boruta* analysis. Traits found significant in green, traits found not significant in red, shadow random traits in blue. Analysis done using populations (A) and species (B, C) as groups.

(A) Bagging

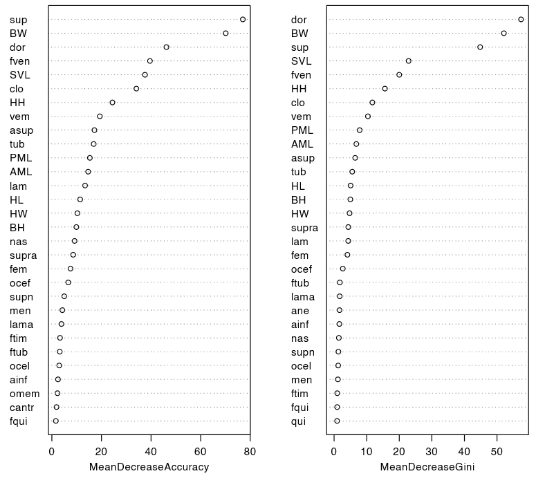
Populations



Species - Domingos, 2014

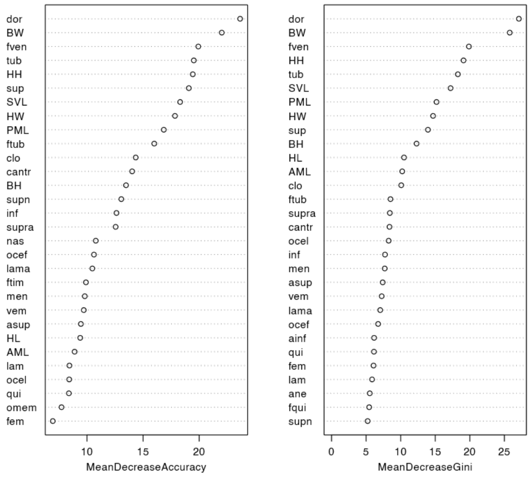


Species - Domingos, 2015

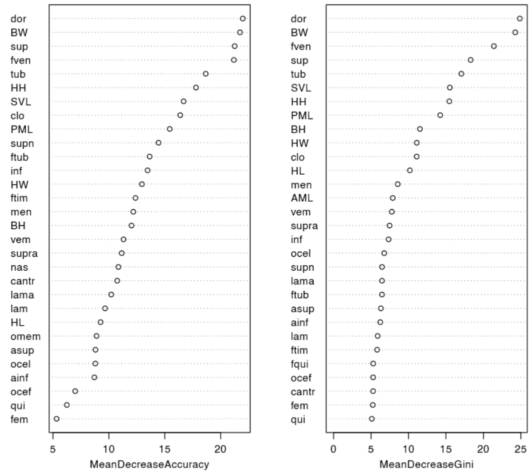


(B) Random Forest

Populations



Species - Domingos, 2014



Species - Domingos, 2015

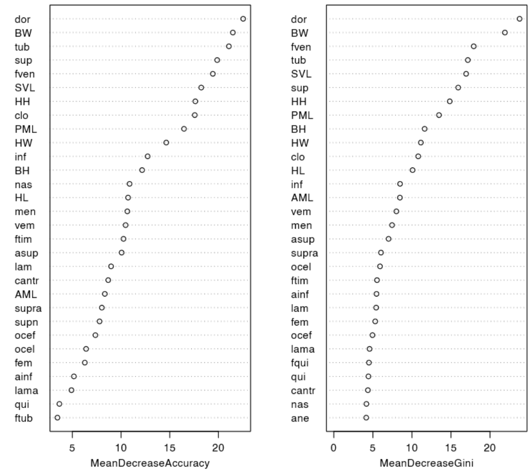


Figure S3A ***(relative to figure in page 34)*****.** Relative importance of each trait, calculated as the mean decrease accuracy/ mean decrease gini of each model when those traits are removed, for the bagging (A) and random forest analysis (B).

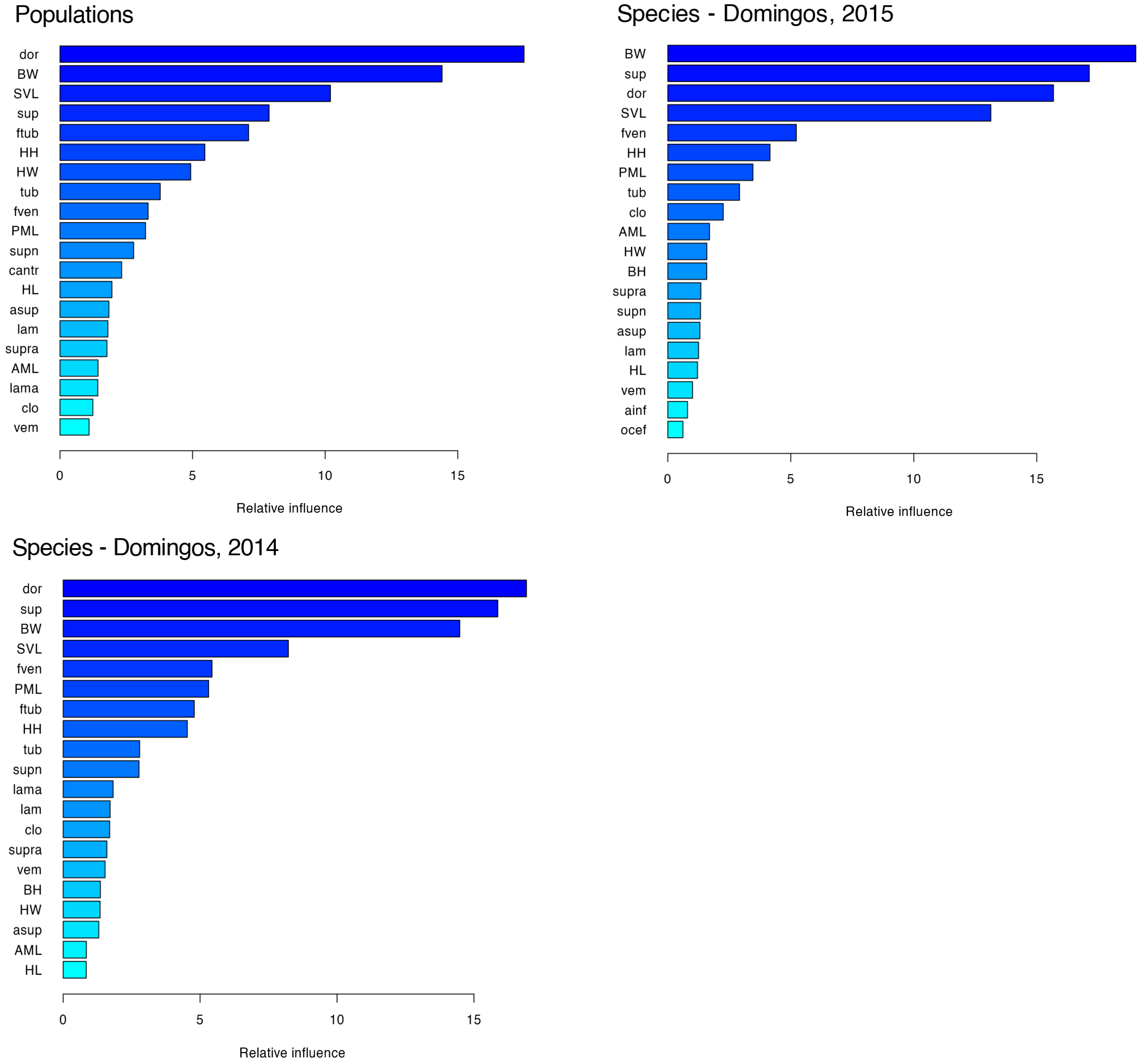


Figure S3B. Relative influence of each variable in reducing the loss function for the boosting analysis.

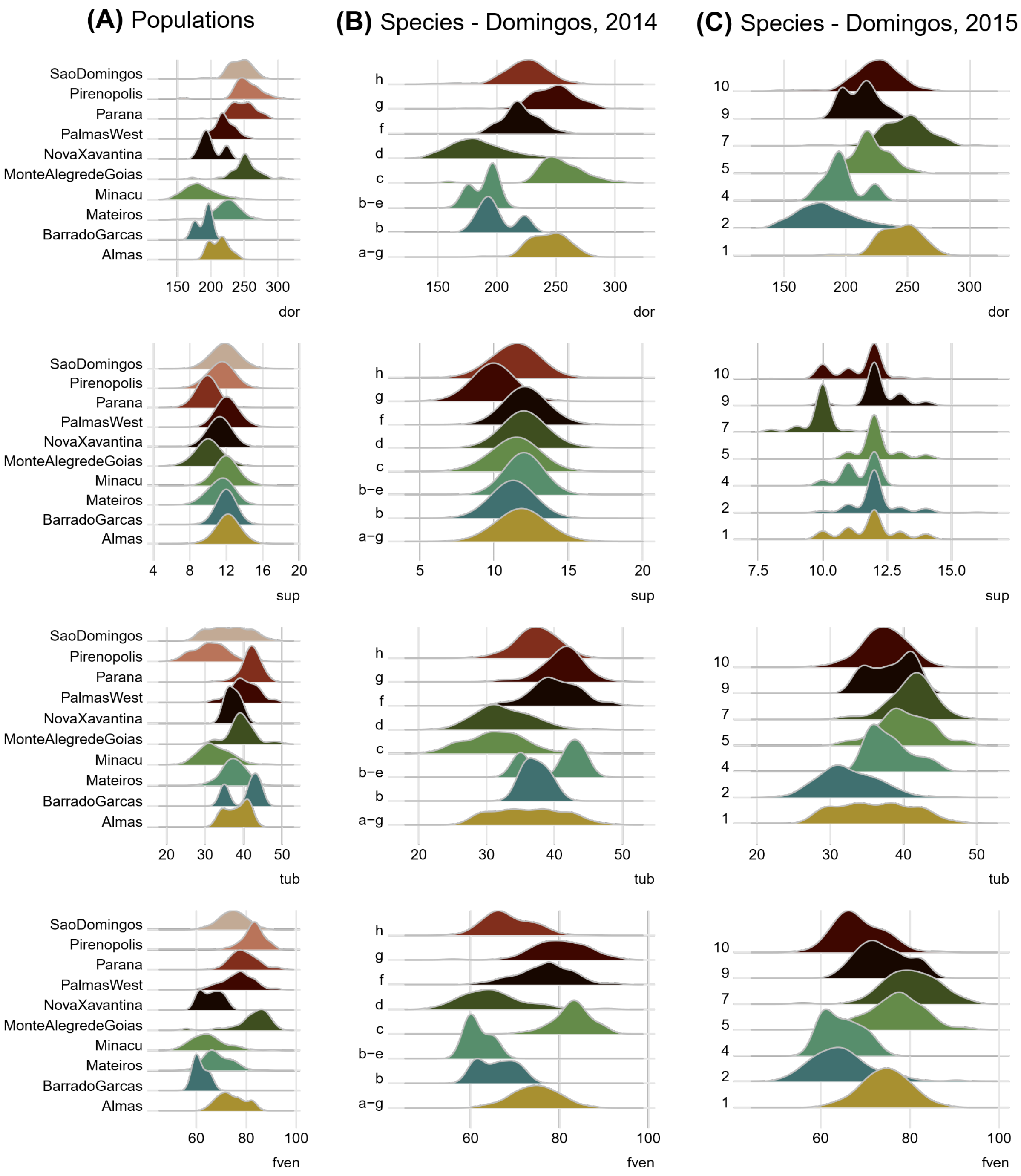


Figure S4. Most important scale count characters in the machine learning analysis are distributed by population (A), or species (B, C).

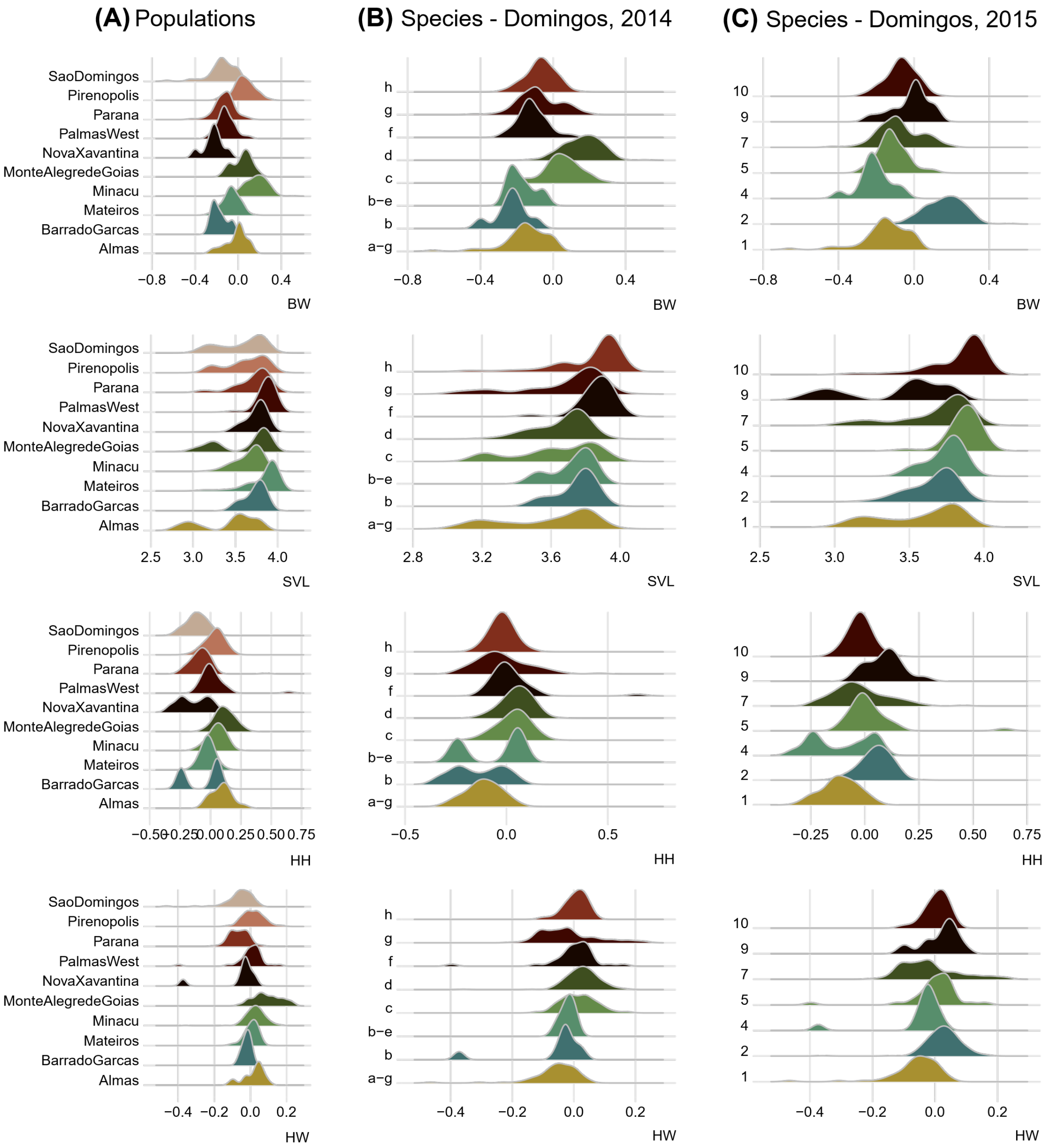


Figure S5. The most important morphometric measurement characters recuperated in the machine learning analysis are distributed by population (A) or species (B, C).

CHAPTER 2: Effects of Quaternary's cycles of range expansion and contraction on macroevolutionary patterns.

Abstract

The Quaternary is a geological period marked by glacial cycles and climatic fluctuations that led to expansions and contractions of different biomes, affecting the local diversity and population structure of various groups. In several Brazilian biomes, this movement of expansion and contraction created refuges that acted as islands of one biome within another. Just as *Gymnodactylus* gr. *amarali*, various groups also diversified during this period, and it is hypothesized that the expansion and contraction of areas could be a mechanism leading to species formation. In this chapter, I used a theoretical approach based on the Derrida-Higgs model, with individual-based simulations to understand the macroevolutionary patterns generated by the contraction and expansion of an area, combined with the formation of refuges. We simulated a complete cycle of contraction and expansion, and the reverse movement of expansion and contraction. During contraction phases, we simulated the area with one refuge, or subdivided it into two or four refuges. We also conducted control simulations with areas that were constantly contracted or expanded. From the simulation results, we constructed phylogenetic trees and calculated diversification rates over time and the balance of the trees. We also observed the number of species generated in each scenario. Contraction dynamics did not drive speciation, with the highest number of species produced in our simulations being equal to or fewer than the number of species in simulations with constantly expanding areas. In large genomes, we also did not observe the effect of allopatry on increasing the number of species. The presence of refuges altered the balance of the trees, making them more balanced. All of our simulations showed high diversification rates, both in tree dynamics scenarios and in constant scenarios. Our results help to better understand the relationship between geographic processes and macroevolutionary patterns.

Keywords: *Speciation, neutral model, changing landscapes, phylogeny, micro-macro gap.*

Introduction

Speciation

What defines a speciation process depends on what we believe a species is. Following the definition we stated in Chapter 1, understanding species as metapopulation lineages evolving separately (de Queiroz, 2007), speciation can be defined as simply the origin of new lineages (Wiens, 2004). Speciation processes can be divided into two types, ecological or geographic (Nosil, 2012). The existence of physical barriers or the effects of landscape change as the sole responsible for cease in gene flow between populations, either being allopatric (non-overlapping ranges between incipient species), parapatric (some overlapping range) or sympatric (overlapping ranges) (Phillimore *et al.*, 2014) is defined as geographic speciation. On the other hand, ecological speciation is observed when speciation results from a combination of ecological processes, mainly adaptive processes (Nosil, 2012). In this scenario, a cease in gene flow happens due to an accumulation of differences that are the outcome of adaptive processes (e.g. Darwin Finches, Grant and Grant (1989), *Anolis* lizards, Losos *et al.* (2006)).

Ecological and geographic speciation are not mutually exclusive, given that isolation by distance and adaptation can be found in groups concurrently (Nosil *et al.*, 2009). Multiple theoretical models study the effects of a combination of both in the formation of species (Aguilée *et al.*, 2013; Gascuel *et al.*, 2015; Alencar & Quental, 2021), but we are still looking to understand which processes are responsible for different patterns of speciation, and which processes lead to the cladogenesis of different groups. Moreover, there is still an existing debate over the possibility and importance of some speciation processes, such as sympatric speciation (Fitzpatrick *et al.*, 2008), speciation with gene flow (Nosil, 2008) and non-adaptive radiations (Gittenberger, 1991).

A widespread challenge in evolutionary biology is understanding how microevolutionary mechanisms affect the speciation dynamics at the macroevolutionary scale of a focus study group. This problem is coined the micro-macro gap, and it occurs when we cannot link microevolutionary processes, such as range expansion, competition and adaptation, to macroevolutionary patterns, which involve patterns above the species level, such as the tempo and mode diversification (Alencar & Quental, 2021; Rolland *et al.*, 2023).

Bridging the micro-macro gap is a challenging task due to their inherent different timescales. Although we have advanced in the genomic area, which allows us to pinpoint with more accuracy diversification times, ancestral locations and bottleneck events (Gehara *et al.*, 2017; Burbrink *et al.*, 2022), and thus infer possible diversification scenarios, theoretical models have been utilized through time to explain the effects of ecology (Aguilée *et al.*, 2013; Gavrilets, 2014) and geography (Manzo & Peliti, 1994; Marquitti *et al.*, 2020; Princepe *et al.*, 2022, 2024) on speciation, while also widening our understanding of possible speciation causes.

Quaternary

The Quaternary period spans approximately the last 2.6 million years and is defined by climatic oscillations, notably glacial cycles, that usually spanned between 40 - 100ky (de (Hammem & Hooghiemstra, 2000; Snyder, 2016). These cycles were characterized by significant climate change, both in temperature and in precipitation (Snyder, 2016). As a result of those climatic fluctuations, different biome areas are predicted to have contracted and expanded, following the cycles of interglacial and glacial maximum (Werneck *et al.*, 2012; Costa *et al.*, 2018a), affecting the local diversity and population structure of different groups (Gehara *et al.*, 2017).

During a biome's contraction phase, peripheral areas often persist, surrounded by newly dominant biomes. For instance, when the Brazilian savanna (Cerrado) contracted, the areas that were all part of the savanna may now be surrounded by rainforests. These remaining patches are known as refugia. Refugia act as crucial areas where species can persist during periods of biome contraction, providing shelter and resources that support species survival throughout the various cycles of expansion and contraction (Werneck *et al.*, 2012).

Vanzolini & Williams (1981) hypothesized the effects of refugia on the formation of different species. Their vanishing refugee model had diversification as a byproduct of adaptation to different habitats due to the landscape changes throughout the Quaternary. However, the area dynamics alone could affect diversity patterns. The changes in area range can induce isolation due to the pulses of allopatry when the area contracts and multiple refugia are created.

There is still some uncertainty on the effects of the Quaternary climatic cycles on speciation (Zink *et al.*, 2004), but the expansion and contractions of different biomes clearly affected the genetic structure of different populations (Gehara *et al.*, 2017), species adaptation (Lister, 2004) and diversity patterns seen today (Werneck *et al.*, 2012). We do know that different groups diversified in the Quaternary (Domingos *et al.*, 2014; Singhal *et al.*, 2018; Kadereit & Abbott, 2021), and the Quaternary landscape dynamics have been proposed to be the trigger of some radiations (Singhal *et al.*, 2018). One example is the gecko species group from the Brazilian Cerrado, *Gymnodactylus* gr. *amarali* Barbour, 1925*,* in which most speciation's happened during the Quaternary. They are endemic to a single biome, and present the same trophic habits and live in the same habitat (Colli *et al.*, 2003; Domingos *et al.*, 2014), which could indicate that their speciation was not based in the adaptation to new niches. Moreover, geographic distance is one of the main factors isolating the different species/populations (Chapter 1), which, again, could be a result of geographic mode of speciation.

Modeling speciation

Speciation can be simplified in a way that can facilitate mathematical modeling, it can be defined as the formation of two distinct peaks in a previously unimodal distribution (Gavrilets, 2014). Adaptive peaks, and adaptation in general are the basis of many speciation models (Gavrilets, 1997, 2003), in which the divergence of a genotype to different adaptation optima leads to isolation and speciation.

Other models portray the accumulation of genetic incompatibilities and reproductive isolation as the result of a non-adaptive process. The neutral theory of biodiversity, showed that, in different ecological contexts, we can find the exact patterns seen in nature modeling them in a flat fitness landscape (Bell, 2000). In an evolutionary context, one of the first models to tackle non-adaptive speciation is the Derrida-Higgs model (Higgs & Derrida, 1991, 1992).

The Derrida-Higgs model is a model of evolving populations with sexual reproduction, no space and no fitness, and it is based on an infinite genome. The result of reproduction is an offspring with a genome that is the recombination of both parents' genomes, and an intrinsic mutation rate. Through a combination of large genomes, mutation and sexual reproduction, a pairwise similarity coefficient threshold arises, and, if reproduction is limited to only pairs with similarities above the threshold value, species emerge. The Derrida Higgs model may be oversimplified, but different more complex models were built on top of it. Manzo & Peliti (1994) introduced space in the model, with allopatry and migration being the triggers for speciation. Moreover, de Aguiar *et al.* (2009, 2017) also spatialized it but now with finite genomes and a reproductive radius.

Besides showing us the prerequisites necessary for speciation, models can help us bridge the aforementioned micro-macro gap. The macroevolutionary parameters that can be analyzed and estimated from real phylogenies are the diversification rates, through time, and between lineages. Heterogeneity in the diversification rates through time produces trees with early burst or late burst patterns, while heterogeneity in diversification rates yields unbalanced trees, with some clades having more species than others (Costa *et al.*, 2019; Harmon *et al.*, 2019). Prolonged speciation times (Etienne & Rosindell, 2012), allopatric barriers (Marquitti *et al.*, 2020) and range expansions (Alencar & Quental, 2021) have shown to affect speciation rates through time and between clades.

In our work, inspired by the diversification processes of *G. amarali* and other radiations during the Quaternary, we used a theoretical approach, with a spatially explicit individual based model, to understand the speciation process, describing what are the macroevolutionary patterns resulting from an expansion and contraction movement of an area, combined with the formation of refugia. We analyzed the balance and tempo of the resulting phylogenies, and examined if those area movements can propel speciation.

Material and methods

The model

We utilized an individual based model (IBM) of speciation based on the Higgs and Derrida (1991) neutral model (i.e., with no fitness) and adapted from the spatially explicit simulations proposed by de Aguiar et al. (2009), with dynamical landscapes, which represents different biome area fluctuations over time during the Quaternary.

Individuals

The model starts with *M* individuals defined by their x_i, y_i position in an *L* = 100x100 lattice. Their genome is finite and is represented by a binary string $\sigma_i = (\sigma_{i1}, \sigma_{i2},...,\sigma_{iB})$ of size *B* in which each position σ_{ik} can assume the values 0 or 1 (see Table 1 for all parameters utilized). At the beginning of each simulation, the population is homogeneous, with all individuals having identical genomes.

Reproduction

Throughout the simulation, the individuals reproduce sexually, generating offspring, and there are no overlapping generations. In each generation, all individuals have a chance to reproduce with an arbitrary partner that is within distance of *S* from them. For reproduction to occur, individuals *i* and *j* must be compatible. In our definition, compatibility occurs when *i* and *j* have less than $G = 0.05*B$ differences in their genomes, which is calculated based on the Hamming distance: $d^{i,j} = \sum_{k=1}^{B} |\sigma\mathrm{ik} - \sigma\mathrm{jk}|$.

If an individual is unable to find a compatible partner, it has three other chances to find another present in a radius that increases by one each trial. After the third attempt, if the individual does not succeed it won't reproduce and is the equivalent of dying. Another individual from the neighborhood will be chosen to reproduce in its place, to keep the population size constant.

The result of reproduction is an offspring that inherits both parent alleles, with a 50% chance of having the first or second parent allele in each loci. The alleles from each individual also have a probability μ of mutation. Offspring will be located in the position of the focal parent, with a probability *D* of dispersing to one of the closest 20 positions.

Species in this model are defined based on the compatibility distance. We build a network of individuals who are possible sexual mates (i.e $d^{i,j} \leq G$). In this compatibility network, those individuals who form a group of individuals reproductively isolated from all others are considered a species. In the network sense, species is a connected component, with the link between them being gene flow. In this manner, two individuals that are not genetically compatible to reproduce could be from the same species if a third individual could reproduce with both.

Table 1. Parameters utilized in the simulations.

B *(genome size)*	1500, 150
G *(reproductive genetic threshold)*	75, 7
M *(number of individuals)*	1000
μ *(loci mutation probability)*	0.00025
S *(reproductive radius)*	5
D *(dispersion probability)*	0.01, 0.05, 0.08, 0.1
N *(grid length x width)*	100x100

Changing landscapes

To define the overall conformation of the area through time and the amount of total contraction, we used real projections of the climate in the last 2 Myr (Snyder, 2016) and predictions of the effect of the change in temperature in the area of different biomes (Werneck *et al.*, 2012; Costa *et al.*, 2018a). Specifically, we defined that the most contracted area would be 60% of the size of the completely expanded area, given the results of the predictions made by Costa et al. (2018), of the Brazilian Savanna region during the last glacial and interglacial periods.

We modeled three different scenarios. The first scenario, the area did not change through time (the CTE scenario), it remained either completely expanded (Figure 1E), or completely contracted (Figure 1C 1 ref, C 2 ref, C 4 ref). The other two scenarios represented a single full cycle of contraction and expansion (or vice versa). Starting with the area fully expanded, then contracted, then expanded again (the ECE scenario) or the reverse, contracted - expanded - contracted (the CEC scenario). Between being completely contracted and expanded, the area goes through transition configurations (Figure 1T1, T2, T3). In the latter two scenarios, we did simulations with different contracted area configurations, to represent the formation of allopatric barriers, and refugia. There was a total of 3 contracted landscape configurations: (a) with only one refugium (Figure 1C 1 ref), (b) two refugia (Figure 1C 2 ref), or (c) four refugia (Figure 1C 4 ref). For the changing landscape scenarios, the number of generations in the extended/contracted configurations was defined as the amount of time that a simulation with constant area would take to reach the mean number of species found in equilibrium (Figure 2A). The times defined were 1000 generations for fully contracted/expanded, and 500 generations for each transition area (Table 2). After each contraction, individuals located in sites that are now unavailable would have seven chances to find an available location within the distance S (S = 5 in all simulations). After the seventh trial, if the individual failed to find a permitted position it was excluded (same as dying) from the time series. We ran 20 iterations for each different scenario, using the coefficient diffusion = 0.01: (i) The CTE scenario x completely expanded, contracted with 1 refugium, contracted with 2 refugia, contracted with 4 refugia; (ii) the ECE scenario x 1 refugium, 2 refugia, and 4 refugia; and (iii) the CEC x 1 refugium, 2 refugia, and 4 refugia. We also did 10 simulations of each of the scenarios with different diffusion coeficients = 0.05, 0.08 and 0.1.

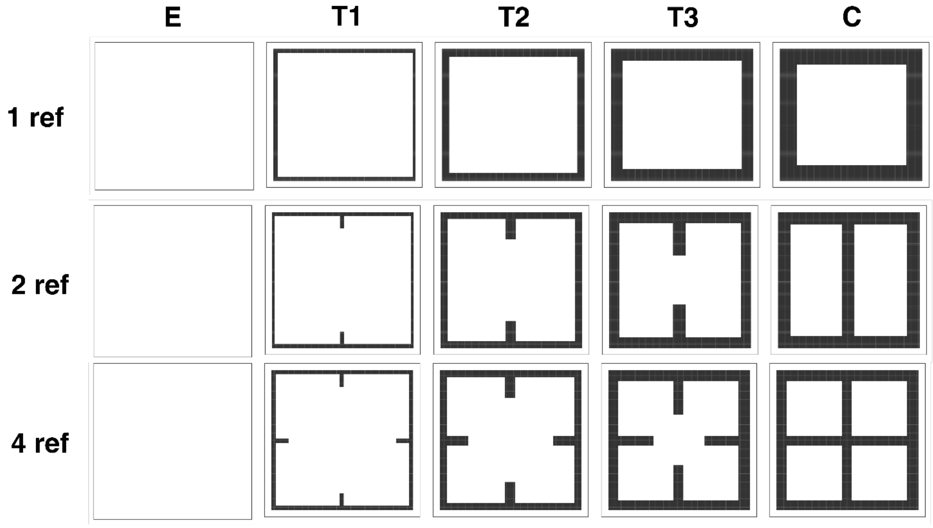


Figure 1. Different landscapes conformations utilized in the simulations. E: expanded, T: transitions, C: contracted, for scenarios with one (1 ref), two (2 ref) or four (4 ref) refugia. The contracted configurations are 60% the size of the completely expanded area.

Table 2. Generation time in each of the areas. E: expanded, T: transition, C: contracted. Scenario that starts contracted in parentheses, and scenario that starts expanded without parentheses. Illustration of each moment in Figure 1.

E (C)	T1 (T3)	T2 (T2)	T3 (T1)	C (E)	T3 (T1)	T2 (T2)	T1 (T3)	E (C)
0 - 1000	1000 - 1500	1500 - 2000	2000 - 2500	2500 - 3500	3500 - 4000	4000 - 4500	4500 - 5000	5000 - 8000

Tree metrics

Following Costa *et al.* (2018b, 2019), we built the phylogenetic by recording parenthood for the entire population and registering the time to the Most Recent Common Ancestor (MRCA) between each pair of individuals. Then, we built their genealogical relationships and we had the MRCA time between all individuals of the extant population, which evolved from a single ancestor. We define the branch lengths and the structure of the phylogenetic trees using only one individual of each species.

Indexes were calculated from a resulting phylogenetic tree in different moments of the evolutionary process. The chosen indexes take the diversification rates through time and among lineages into account.

The Sackin index, I(N), shows the heterogeneity in speciation rates through lineages. It quantifies the balance of the trees (the larger the index, the less balanced the tree is) and is defined as: $I(N) = \sum_{j=1}^{N} d_j$, in which N equals number of species (leaves) and d_j is the number of nodes between each leaf and the root of the tree.

The gamma-statistics, γ, represents diversification rates through time. High gamma-statistics values mean an acceleration in diversification rates and "tippy" trees, while low gamma values represent trees with most speciation events happening at the beginning of diversification, with a deceleration of that rate through time, resulting in more "steamy trees". It is calculated as:

$$\gamma = \frac{1}{D}\left[\frac{1}{R-2}\sum_{k=2}^{R-1}\Theta(k) - \frac{\Theta(R)}{2}\right]$$

in which:

$$\Theta(k) = \sum_{j=2}^{k} j g_j \ ,$$

$$D = \frac{\Theta(R)}{\sqrt{12(R-2)}},$$

g_k is the time interval between the *k* and *k-1* speciation events and $g_k = \frac{1}{bk}$, where *b* is the fixed rate in which species bifurcate and *R* is the number of species.

Both vary due to the number of leaves present in the analyzed tree. Given that, we also calculated a normalized version of the Sackin index (Cardona *et al.*, 2012) under the Yule model. For the diversification acceleration, we calculated the alpha value (Costa *et al.*, 2019) which is invariant to the number of leaves in the phylogenetic tree, making comparison between different trees of different sizes possible. The alpha value can be calculated as: $g_k(\alpha) = \frac{1}{bk^{\alpha}}$, with b and k being the same parameters from the gamma-statistics.

Since we are looking to understand the dynamics through time, and the changing nature of our model prevents the number of species from reaching an equilibrium (Figure 2B, C), we collected information on the number of species at every 25 generations and tree parameters at every 500 generations (right before the transition to a new landscape conformation).

We calculated Cardona's (2012) normalized Sackin index, alpha value, and gamma statistics for the empirical tree of *Gymnodactylus* gr. *amarali* from Domingos (2015) and compared it with the simulation results.

Results

Speciation

All results shown have a dispersion parameter (*D*) of 0.01, unless otherwise specified. The constant landscape scenarios with the largest genome (*B* = 1500) showed similar speciation patterns for all contracted landscape scenarios (Figure 2A), reaching the same mean number of species. The expanded constant scenario reached approximately 10 more species than the contracted one. The speciation bursts around the same number of generations just after initial time (~ 300 generations). However, after initialized the expansion of the spatial configuration, it took as many generations to increase the number of species as it took to decrease during the contraction (time ~ 2500 generations).

When we used the smallest genome size (*B* = 150) in an area of 60% of the original spatial availability (in the constantly contracted landscape configuration), no species formed in most simulations (Figure 3A). When there were barriers splitting the spatial configuration, we observed the formation of two species and four species in the scenario (Figure 3A, Ref 2, and Ref 4, respectively). The completely expanded scenario reached approximately 10 fewer species than the same scenario with a bigger genome.

In the simulations with the biggest genome (*B* = 1500) and changing landscapes (Figure 2B, C), there were no differences in the mean number of species given by the different contracted area configurations. In the most expanded moments, we observed the biggest number of species, reaching the same amount as the constantly expanded simulations. The same happened in the most contracted moments, the number of species diminished, reaching the mean found in the constantly contracted scenario. Both ECE and CEC scenarios produced the same results; that is, the number of species in a given time depends on the area configuration at the same time.

Simulations with the smaller genome and changing landscapes, produces a similar pattern for the scenarios starting already expanded (ECE) (Figure 3B), with species number reaching the observed for constant scenarios when maximally expanded and when maximally contracted. However, in the scenarios in which the area starts contracted (CEC) (Figure 3C) there is an effect of the initial condition in the number of species reached while the area is completely expanded. With the simulations starting with already four refugia reaching more than the double of species while completely expanded than the simulations starting with a single refugium.

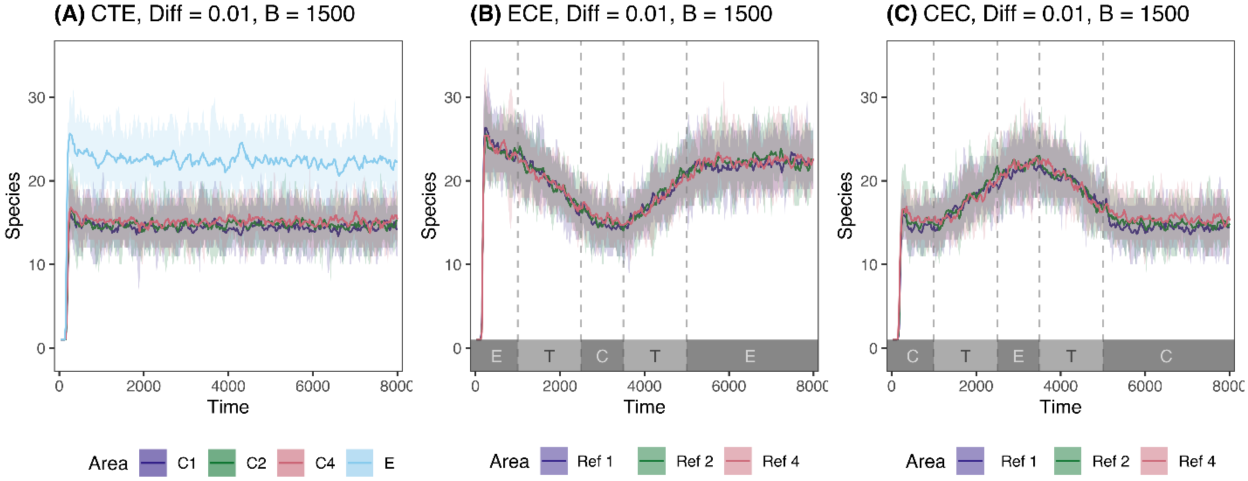


Figure 2. Number of species through time (number of generations). (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four

refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 1500 and diffusion = 0.01.

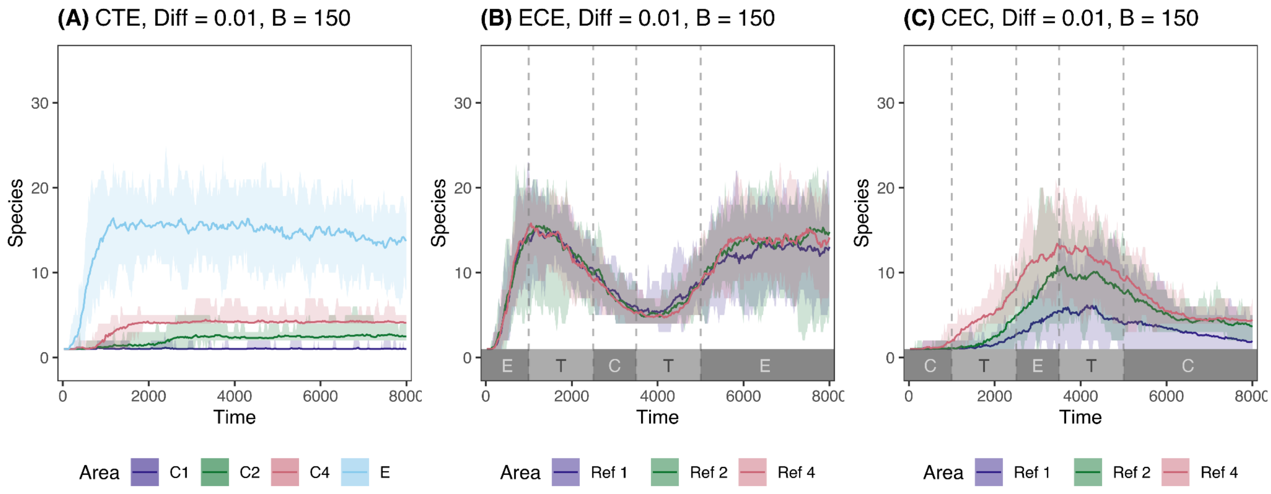


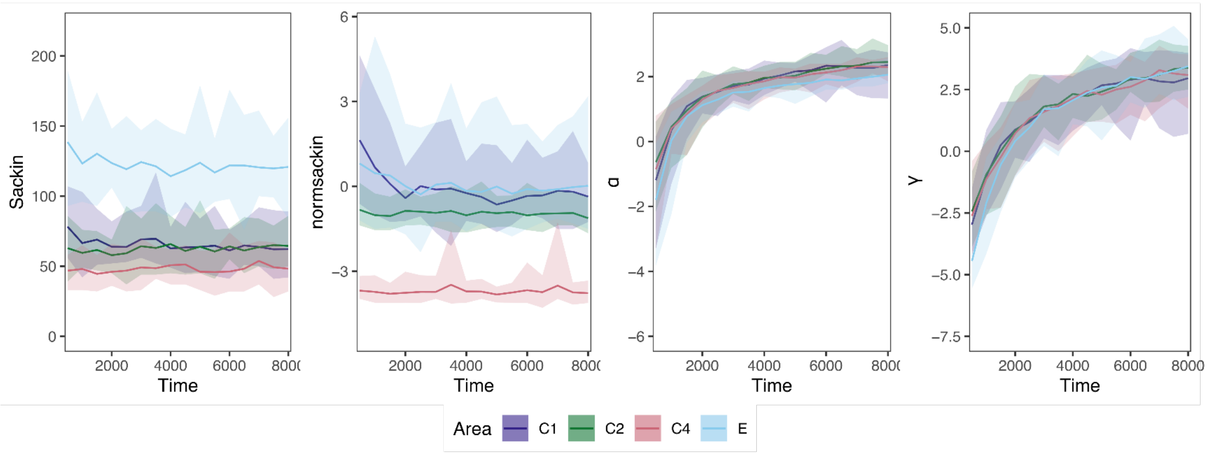
Figure 3. Number of species through time (number of generations). (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 150 and diffusion = 0.01.

Tree metrics

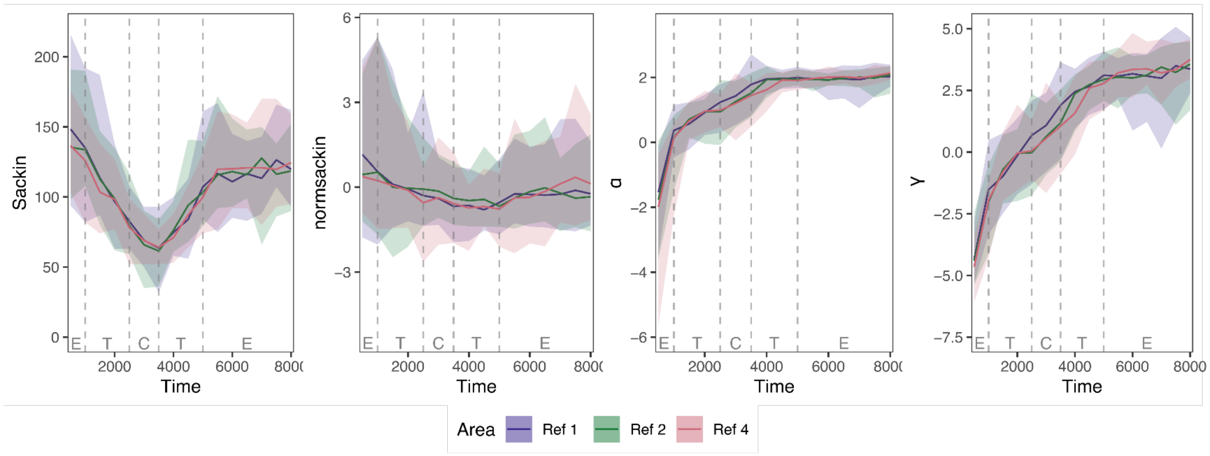
Since the different scenarios resulted in different numbers of species, we opted for the normalized Sackin index to analyze the variation in tree balance. At the constant landscape for the larger genome (*B* = 1500), the phylogenetic trees of the scenario with constant allopatry with four different areas (Figure 4A) have a normalized Sackin index lower than all other scenarios, which represents a more balanced tree. That balance persists throughout all 8000 generations, with final normalized Sackin indexes varying between four refugia configurations and others (Figure 6A). The other area configurations in the constant landscape simulations present a normalized Sackin value close to zero (Figure 4A). The changing landscape scenarios show different patterns of tree balance. In the ECE scenario (expanded-contracted-expanded) we see in the beginning a high variance of normalized Sackin values that rapidly reduces and remains also around zero, for all types of contracted configuration (one refugium, two refugia, four refugia) (Figure 4B). In the CEC scenario (contracted-expanded-contracted), simulations that started with four refugia showed smaller values of normalized Sackin in the beginning, with the index gradually increasing, reaching values close to the other two contraction scenarios (with one refugium and two refugia) (Figure 4C). However, at the end of the simulation we can still detect that the normalized Sackin index for four refugia is different, with a lower index (Figure 6C).

The simulations with genome size $B = 150$ often did not produce enough species to analyze tree balance and tempo (Figure 5). The only consistent scenario that resulted in enough species in all simulations was the ECE scenario (Figure 5), in which the normalized Sackin index showed similar values and patterns with the ones found in the simulations with $B = 1500$, with a high variance at the start, and a reduction of that variance with time, with values close to zero. In the ECE scenario, the gamma-statistics index has a very sharp change right after the contracted phase, resembling a logistic curve (Figure 5).

(A) Constant landscape



(B) Expanded – Contracted – Expanded



(C) Contracted – Expanded – Contracted

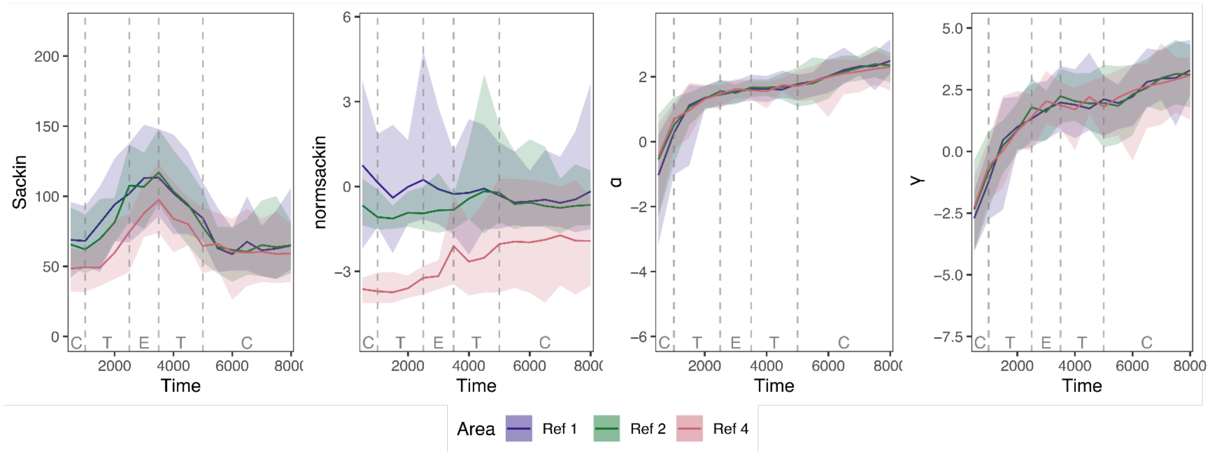


Figure 4. The tree metrics, Sackin, normalized Sackin (normsackin), alpha and gamma, through time. (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red -

four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 1500 and diffusion = 0.01.

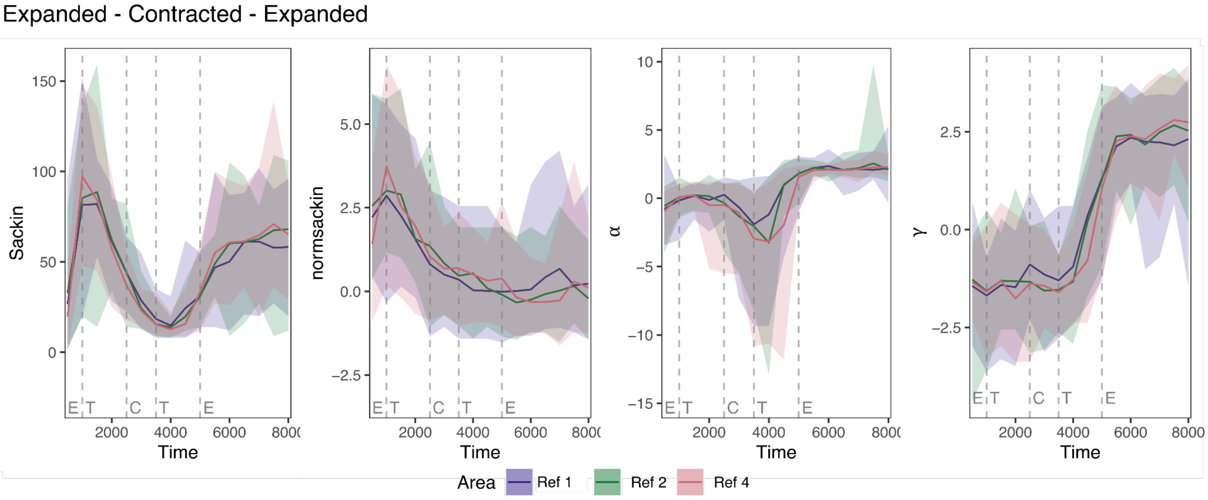


Figure 5. Tree metrics, Sackin, normalized Sackin (normsackin), alpha and gamma, through time for the ECE scenario for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations with genome size = 150 and diffusion = 0.01.

Although we see a pattern for the tempo of the trees (alpha and gamma) in all types of simulations (constant landscape, ECE or CEC) and number of refugia during contraction, with those values increasing through time, with in the end, positive values, representing tippy trees, we can still see a variation in the slopes of both indexes for the ECE and CEC scenarios (Figure 4). For *B* = 1500, in the constant scenario the gamma statistics varies with no major slope changes. In the ECE scenario, when the area starts contracting for the first time (around *T* = 1000), there is a slight change of the slope, an acceleration of the gamma-statistics index. In the CEC scenario we see a more abrupt deceleration of the gamma index right after the area is completely expanded and starts contracting (around *T* = 3500). Those differences can be seen with less intensity in the alpha values (Figure 4). None of those variations through time inflicts any change at the finishing resulting tree, as both gamma-statistics and alpha-value do not differ between scenarios at the end of the simulations (Figure 7, 8).

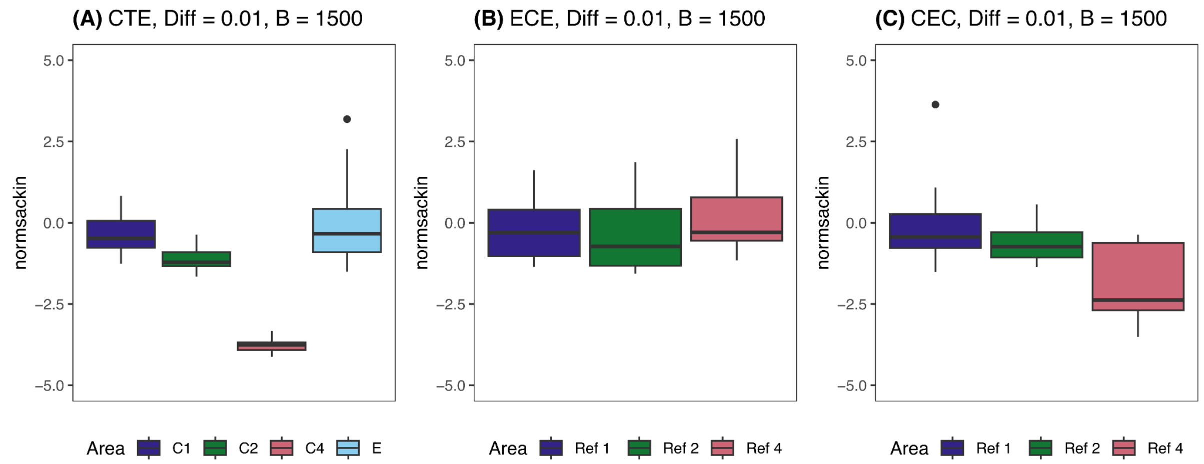


Figure 6. Normalized Sackin index at the end of every simulation. (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 1500 and diffusion = 0.01.

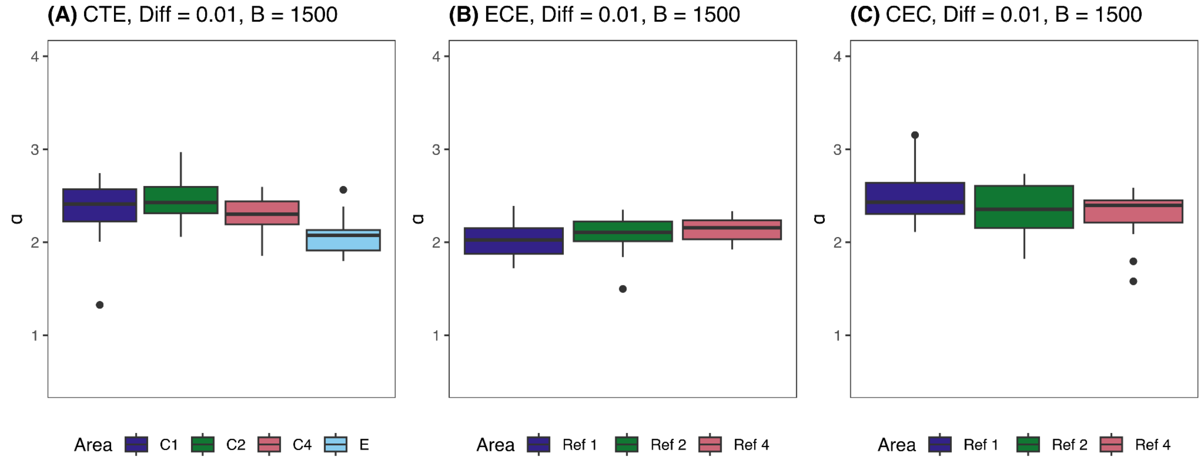


Figure 7. Alpha values at the end of every simulation. (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one

refugium, green - two refugia and red - four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 1500 and diffusion = 0.01.

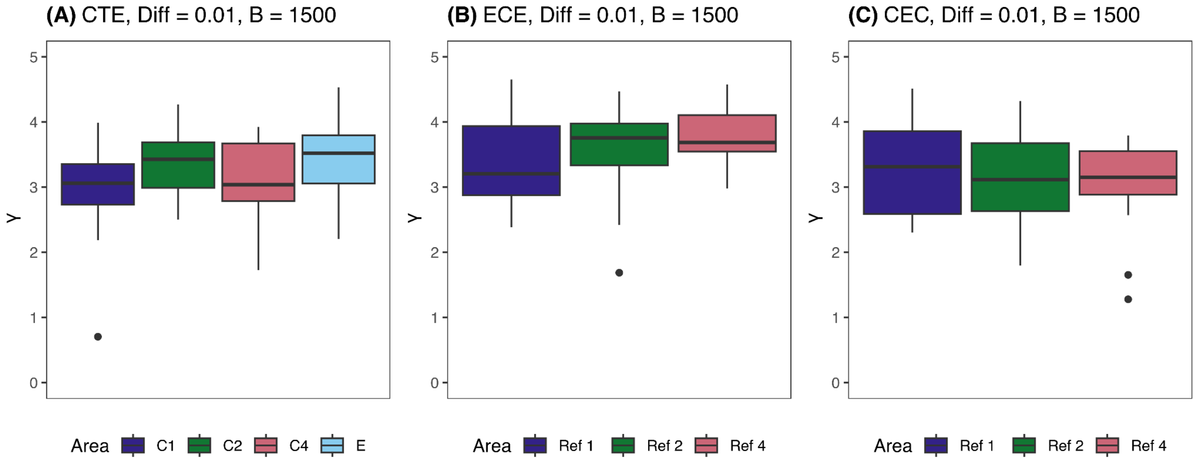


Figure 8. Gamma-statistics at the end of every simulation. (A) Simulations with constant areas: purple - contracted one refugium, green - contracted two refugia, red - contracted four refugia, and blue - constantly expanded area. (B) Simulations with changing landscapes, ECE scenario (extended – contracted – expanded), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. (C) Simulations with changing landscapes, CEC scenario (contracted – expanded - contracted), for different configurations during the contracted phases: purple - one refugium, green - two refugia and red - four refugia. Simulations done with genome size = 1500 and diffusion = 0.01.

The simulations with diffusions higher than 0.01, reached a smaller number of species, higher the diffusion, smaller the number of species (Figure S1). There are no major changes in the tree parameters for different diffusions (Figure S2, S3), only in CEC scenarios with 4 refugia, we see an increase of the normalized Sackin value from diffusions larger than 0.01, compared to the results when diffusion is equal to 0.01 (Figure S1C). The other differences in the Sackin index for different diffusions are due to the variation of species number.

Gymnodactylus gr. *amarali* empirical tree presented a high positive value of normalized Sackin, 4.450157, and positive values for the gamma and alpha indexes, 1.277366 and 1.54762, respectively.

Discussion

Quaternary's glacial dynamics could possibly affect macroevolutionary patterns and impulse species formations as a consequence of isolations of groups through time due to the changing landscape dynamics. Those dynamics could result in speciation due to allopatry or vicariance, or due to adaptation to different biomes. The vanishing refugee model (Vanzolini & Williams, 1981; Damasceno *et al.*, 2014) suggests that individuals from a refugium that disappeared could adapt to a new ecological area, producing an "ecological variant" of that species. Moreover, the contraction of an area could induce moments of sympatry for incipient species. Pulses of sympatry, with the "merge and diverge" dynamics, could increase genetic variation and accelerate the speciation rate (Grant & Grant, 1996, 2006). Here, we studied how the changes in landscapes can affect speciation dynamics by varying the level of parapatry-sympatry through time. Also, inducing isolation between individuals with pulses of allopatry during moments of spatial configuration of contraction with different numbers of refugia. For this, we developed an IBM genetically and spatially explicit with a dynamical space configuration.

Examining species formation and number of species over time, our contraction and expansion dynamics do not have a long-term effect on species number. Our results show that, given enough time, regardless of whether the area was previously contracted or expanded, the system reaches the same number of species that it would also reach in an area of the same size if it were in a constant scenario. For large genomes, the different barriers in both constant and changing landscapes also did not increase the number of species. Other than genome size, what determined the species number at a given time was the size of the area. De Aguiar *et al.* (2009, 2017) found that the genome size and compatibility threshold (*B* and *G* parameters, respectively), combined with the reproductive radius (*S*), was enough to produce species. This type of speciation, which they called topopatric, is the one we see in our simulations, with geographic isolation happening with no need for barriers. In their case, the smaller the *S, the* more species are formed and more rapidly. When diminishing the area and increasing the density, we observed similar effects that de Aguiar *et al.* (2009, 2017) and Costa (2018) found by increasing the reproductive radius *S*. This acts as a change in the level of parapatry.

Moreover, the bigger the genome, the smaller the ratio between the reproductive radius (*S*) and the area size (*L*) that can yield species. This is observed in the case of small genome size (*B = 150*), in which small areas do not induce speciation in the absence of allopatric barriers – similar to what is observed when the reproductive radius is too large (De Aguar *et al*., 2017). When the dynamics reach most contracted areas – which are not large enough for speciation – we observe the effect of allopatric barriers, with the simulations with two areas, in the constant landscape scenario, producing an average of two species, and with four isolated areas, resulting in four species (Figure 2A).

The biggest effect our simulations with changing landscapes had on the overall shape of the phylogenetic tree was in the tree balance in the scenarios where the area started contracted with four refugia (Figure 6). This scenario, just as with the one with the area constantly with four refugia, produced more balanced trees (lower normalized Sackin values), that perdures through time (Figure 3). These results corroborate with Marquitti *et al.* (2019) findings on allopatry producing more balanced trees. The *Gymnodactylus* gr. *amarali* tree presents a high normalized Sackin index (4.45), representative of an unbalanced tree. This imbalance is found in our simulations with fewer refugia (two or one refugia). In both simulations, ours and Marquitti *et al.* (2019), the refugia had very symmetrical areas. We know that the refugia, through time, were heterogeneous in size and distance (Werneck *et al.*, 2012; Costa *et al.*, 2018a). Because we observe a stronger effect on the tree balancing when spatial configuration is split, we expect that this heterogeneity in refugia could be an important factor leaving perhaps a different signature in the macroevolutionary pattern than the one we found with homogenous refugia. Testing, in future works, if the tree balance would perdure with multiple refugia of different sizes and distances, as it was possibly the scenario during *G. amarali*'s, and other groups (Singhal *et al.*, 2018) speciation, would give us a bigger insight into the effects of the Quaternary landscape dynamics on tree balance.

At the end of all of our simulations, the rates of diversification through time show values related to an accelerated diversification (alpha and gamma>0), which culminates with trees with a more “tippy” format. No single simulation produced a slowdown in rates of diversification through time. Diversification slowdown is commonly found in empirical trees (Moen & Morlon, 2014), and different mechanisms are hypothesized to explain this deceleration, such as niche filling (Gascuel et al., 2015) and prolonged speciation time (Etienne & Rosindell, 2012). Another discussed and modeled mechanism that results in deceleration is the formation of geographic barriers and reduction of ranges (Pigot *et al.*, 2010; Moen & Morlon, 2014). Our results go against those findings as there was no difference in the final alpha values and gamma statistics for the different scenarios and contraction configurations (Figure 7, 8). Rates of diversification in the Quaternary are not congruent throughout different groups, with studies showing a deceleration of diversification in some taxa (Zink *et al.*, 2004; Barnosky, 2005) and an acceleration in others (Kadereit & Abbott, 2021). One example is the *G. amarali* group with gamma and alpha values > 1. This variation could be due to methodological incongruencies for calculating those rates (Barnosky, 2005). Nonetheless, the contradiction in rates of diversification lies if the quaternary cycles are long enough to impulse speciation (Kadereit & Abbott, 2021) or too short, which would make the effects of the climatic cycles a background noise in the evolution of different groups (Barnosky, 2005).

With the gamma statistics and alpha values, we access diversification through time, which is a combination of both speciation and extinction rates. In many cases, the diversification rate can be a proxy for speciation rates when extinction is constant or negligible (Costa *et al.*, 2019). On the other hand, in our simulations, we observe a drastic reduction of species number, which could be the result of hybridization and extinction events. Further analysis differentiating the rates that compose the diversification, i.e, speciation, extinction and hybridization, could improve our understanding about the importance of each event to the different aspects of the macroevolutionary pattern.

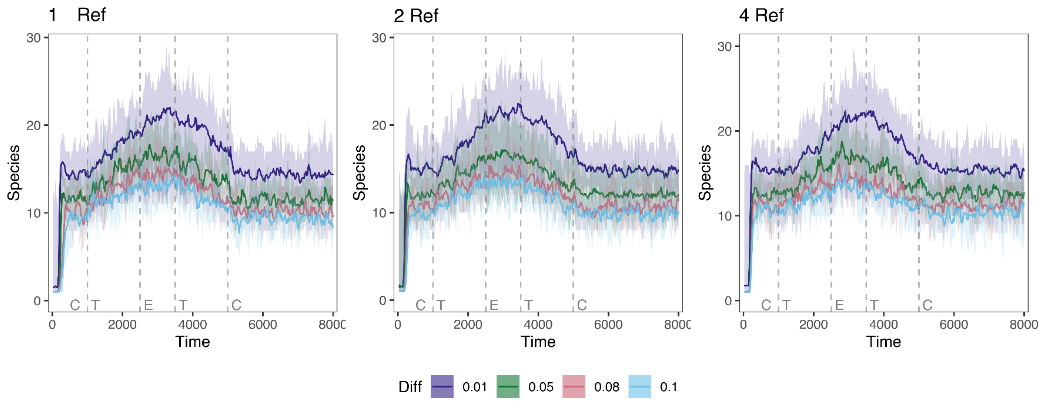
At last, in our study, our simulations were with constant population sizes. Future works could focus on varying population size through time, while maintaining the density constant, since we know that some groups have signs of bottleneck events happening during contraction periods (Gehara *et al.*, 2017). The drastic changes in population size could generate different macroevolutionary patterns than the ones found in our simulations.

Conclusion

The neutral landscape dynamics of contraction and expansion did not result in an increase of diversification, as only area size defined the number of species reached. Not only that, when our simulations started with more species than the number of isolated areas, there was no effect of allopatry in diversification. However, allopatry affects the balance of resulting phylogenies, if our simulations started in a contracted scenario. We also found that the diversification rates changed in slope after the area started expanding or contracting, accelerating and slowing down respectively.

CHAPTER 2: Supplementary material

(A) Contracted - Expanded - Contracted



(B) Expanded - Contracted - Expanded

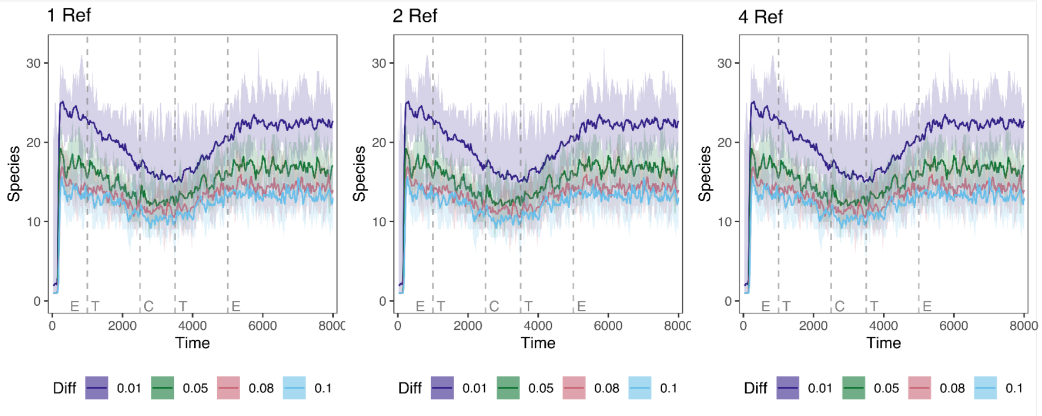
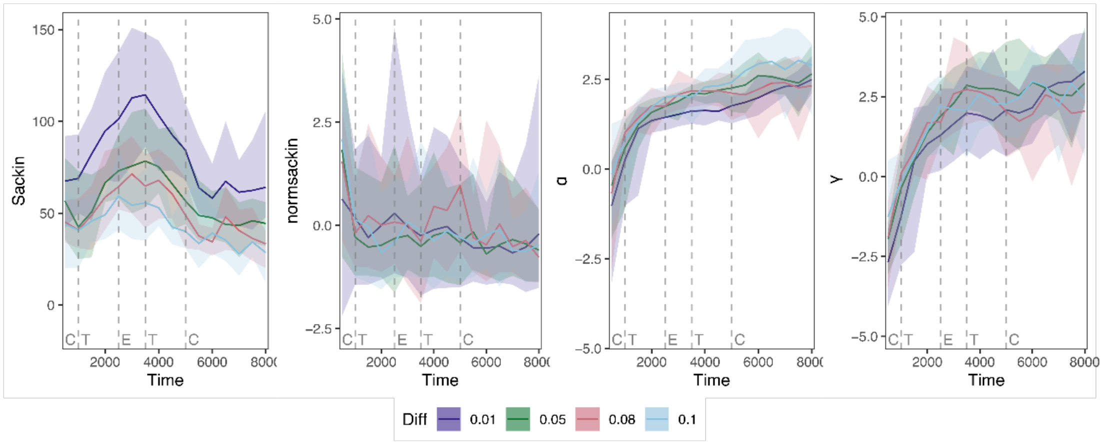
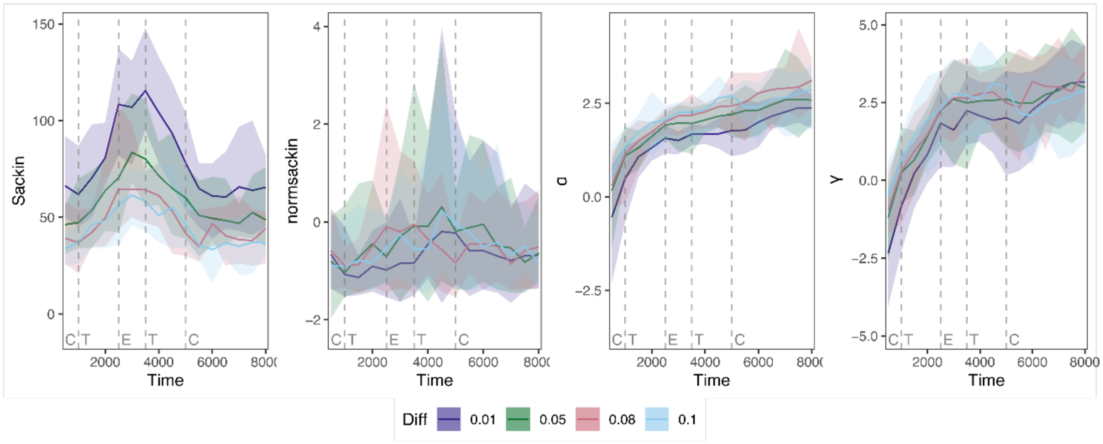


Figure S1. Number of species through time for simulations with different diffusion values (purple = 0.01, green = 0.05, red = 0.08, blue = 0.1.) (A) ECE scenario (extended – contracted – expanded). (B) CEC scenario (contracted – expanded - contracted).

(A) 1 Ref



(B) 2 Ref



(C) 4 Ref

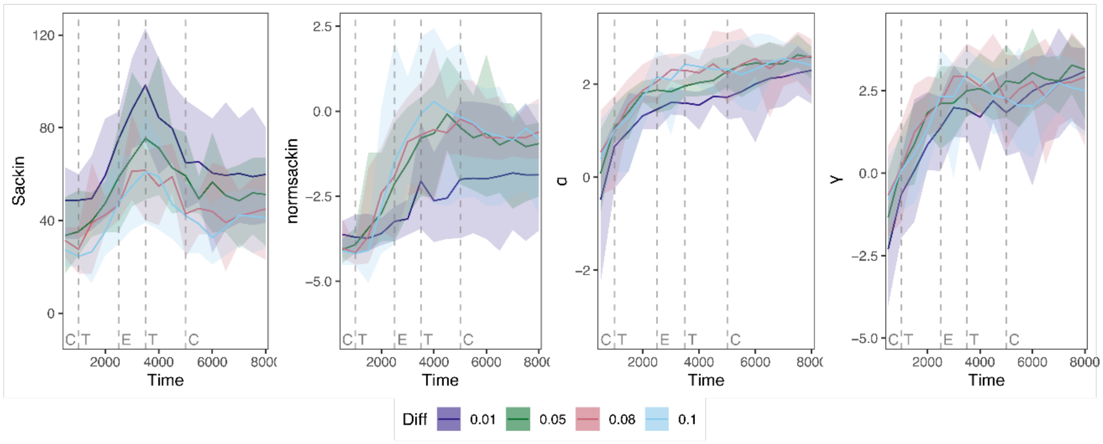
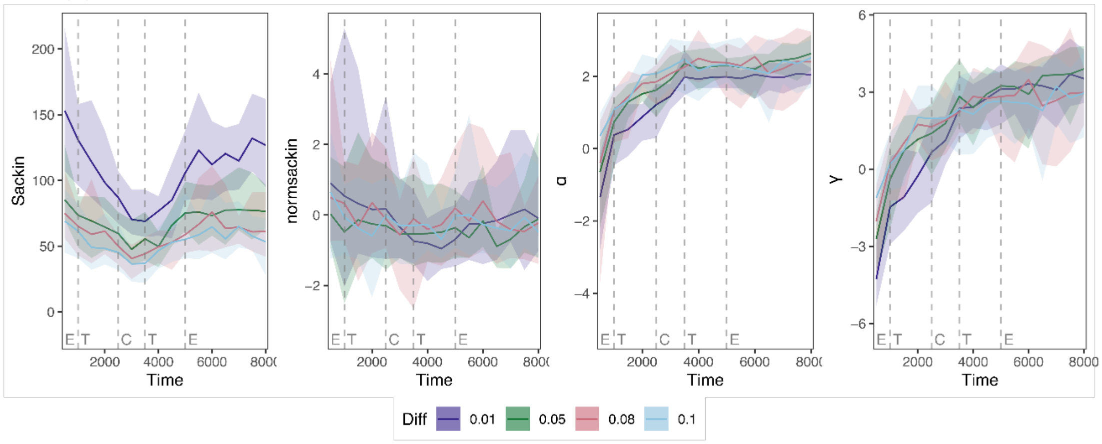
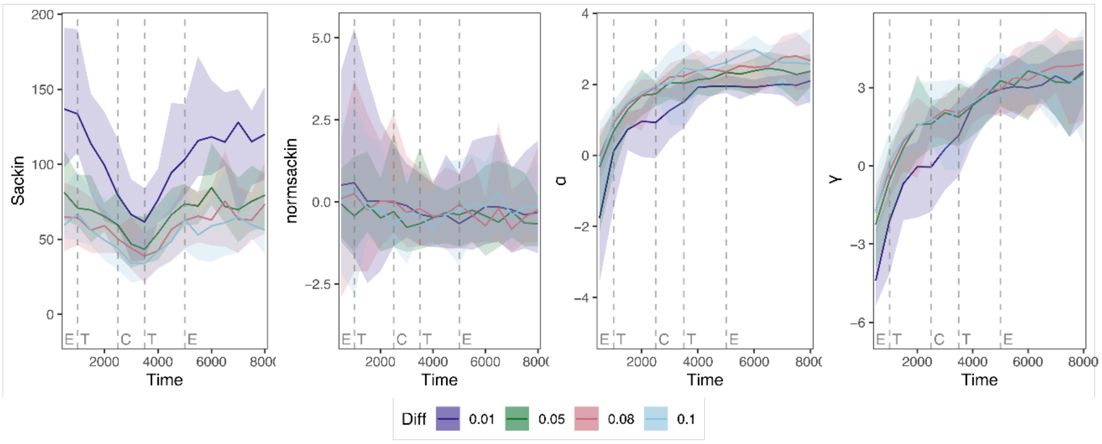


Figure S2. Tree metrics, Sackin, normalized Sackin (normsackin), alpha and gamma, through time for different diffusions in the CEC scenario, with (A) one refugia, (B) two refugia, (C) 4 refugia for different diffusion values. Simulations done with genome size = 1500.

(A) 1 Ref



(B) Ref 2



(C) Ref 4

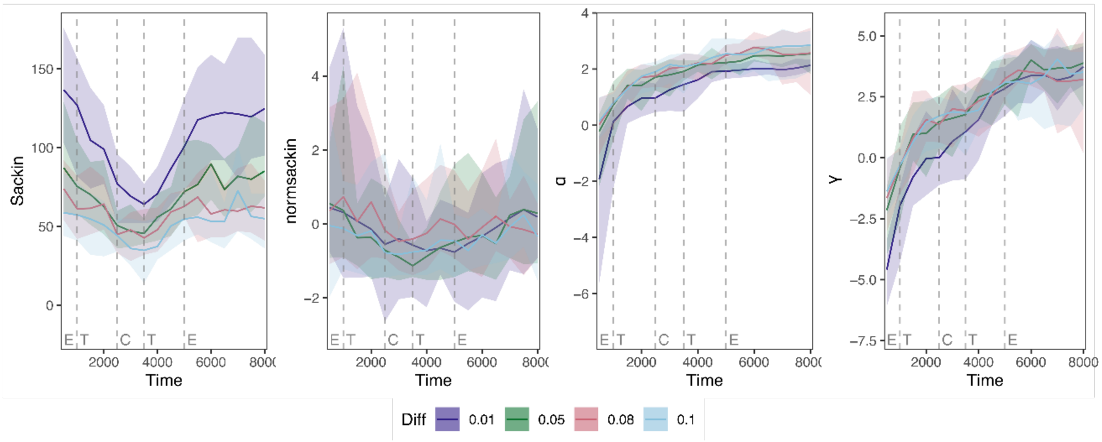


Figure S3. Tree metrics, Sackin, normalized Sackin (normsackin), alpha and gamma, through time for different diffusions in the ECE scenario, with (A) one refugia, (B) two refugia, (C) 4 refugia for different diffusion values. Simulations done with genome size = 1500.

DISCUSSÃO GERAL E CONSIDERAÇÕES FINAIS

Sistemas biológicos são complexos, confusos e contraintuitivos. É impossível construir um modelo que leve em consideração toda a variação, nuances e possibilidades existentes dentro de um processo evolutivo. Não obstante, modelos são imprescindíveis para entender causas e consequências de diferentes processos, mesmo que de uma maneira simplificada (Servedio *et al.*, 2014), e a combinação de teoria com dados nos permite uma visão mais robusta dos processos evolutivos.

No meu trabalho busquei entender o que mantém e o que pode ter promovido o isolamento de um grupo de espécies de lagartixas crípticas do Cerrado. Utilizei dados morfológicos, genéticos, climáticos e geográficos para esclarecer o que atualmente está mantendo a estruturação das populações/espécies dentro de *Gymnodactylus* gr. *Amarali*. Nossos resultados apontaram para fatores externos relacionados com a estruturação genética e morfológica do grupo. No entanto, apesar dos nossos resultados apontarem um isolamento entre as populações/ espécies por distância geográfica e adaptação ao clima, não podemos afirmar quais foram os processos que geraram os padrões observados.

A maioria das espécies de *Gymnodactylus* gr. *amarali* surgiram no Quaternário, período marcado por ciclos glaciais e alterações de área e distribuição dos biomas, que afetaram diretamente o Cerrado. Assim, surge a pergunta: os movimentos de expansão e contração do Cerrado podem ter induzido a diversificação no grupo? Decidimos então buscar outros métodos e expandimos a pergunta: será que esses movimentos, desacoplados de qualquer processo adaptativo, impulsionam a formação de espécies?

Utilizando modelos teóricos a partir de simulações neutras e espaciais, conseguimos responder que esses movimentos de contração e expansão, que levam a variação de simpatria-parapatria com pulsos de alopatria das populações ao longo do tempo, não impulsiona a formação de novas espécies. Futuros estudos combinando adaptação com as flutuações ambientais podem trazer resultados diferentes e gerar um aumento do número de espécies.

Ao fazer várias perguntas relacionadas à trajetória evolutiva de um grupo de espécies crípticas, buscando entender o que promove a estruturação no grupo, e descrever diferentes padrões macroevolutivos, conseguimos lançar luz sobre os processos que geram e mantem a diversificação e alteram as taxas de especiação.

REFERÊNCIAS

Adams DC., Berns CM., Kozak KH., Wiens JJ. Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences* 2009; **276**: 2729–2738.

Adams DC., Church JO. Amphibians do not follow bergmann's rule. *Evolution* 2008; **62**: 413–420.

de Aguiar MAM. Speciation in the Derrida–Higgs model with finite genomes and spatial populations. *Journal of Physics A: Mathematical and Theoretical* 2017; **50**: 085602.

de Aguiar M a. M., Baranger M., Baptestini EM., Kaufman L., Bar-Yam Y. Global patterns of speciation and diversity. *Nature* 2009; **460**: 384–387.

Aguilée R., Claessen D., Lambert A. Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. *Evolution* 2013; **67**: 1291–1306.

Alencar LRVD., Quental TB. Linking population-level and microevolutionary processes to understand speciation dynamics at the macroevolutionary scale. *Ecology and Evolution* 2021; **11**: 5828–5843.

Barnosky AD. Effects of Quaternary Climatic Change on Speciation in Mammals. *Journal of Mammalian Evolution* 2005; **12**: 247–264.

Bell G. The Distribution of Abundance in Neutral Communities. *The American Naturalist* 2000; **155**: 606–617.

Bickford D., Lohman DJ., Sodhi NS., Ng PKL., Meier R., Winker K., Ingram KK., Das I. Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* 2007; **22**: 148–155.

Breiman L., Cutler A., Liaw A., Wiener M. randomForest: Breiman and Cutler's Random Forests for Classification and Regression. 2022;

Burbrink FT., Bernstein JM., Kuhn A., Gehara M., Ruane S. Ecological Divergence and the History of Gene Flow in the Nearctic Milksnakes (***Lampropeltis triangulum*** Complex) (D Warren, Ed.). *Systematic Biology* 2022; **71**: 839–858.

Calsbeek R., Knouft JH., Smith TB. Variation in scale numbers is consistent with ecologically based natural selection acting within and between lizard species. *Evolutionary Ecology* 2006; **20**: 377–394.

Cardona G., Mir A., Rosselló F. Exact formulas for the variance of several balance indices under the Yule model. *Journal of Mathematical Biology* 2012; **67**: 1833–1846.

Cerca J., Meyer C., Stateczny D., Siemon D., Wegbrod J., Purschke G., Dimitrov D., Struck TH. Deceleration of morphological evolution in a cryptic species complex and its link to paleontological stasis. *Evolution* 2020; **74**: 116–131.

Claunch NM., Nix E., Royal AE., Burgos LP., Corn M., DuBois PM., Ivey KN., King EC., Rucker KA., Shea TK., Stepanek J., Vansdadia S., Taylor EN. Body size impacts critical thermal maximum measurements in lizards. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 2021; **335**: 96–107.

Colli GR., Mesquita DO., Rodrigues PVV., Kitayama K. Ecology of the Gecko Gymnodactylus geckoides amarali in a Neotropical Savanna. *Journal of Herpetology* 2003; **37**: 694–706.

Costa GC., Hampe A., Ledru MP., Martinez PA., Mazzochini GG., Shepard DB., Werneck FP., Moritz C., Carnaval AC. Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. *Global Ecology and Biogeography* 2018a; **27**: 285–297.

Costa CLN., Lemos-Costa P., Marquitti FMD., Fernandes LD., Ramos MF., Schneider DM., Martins AB., de Aguiar MAM. Signatures of Microevolutionary Processes in Phylogenetic Patterns. *Systematic Biology* 2019; **68**: 131–144.

Costa CLN., Marquitti FMD., Perez SI., Schneider DM., Ramos MF., de Aguiar MAM. Registering the evolutionary history in individual-based models of speciation. *Physica A: Statistical Mechanics and its Applications* 2018b; **510**: 1–14.

Crouch NMA., Tobias JA. The causes and ecological context of rapid morphological evolution in birds. *Ecology Letters* 2022; **25**: 611–623.

Damasceno R., Strangas ML., Carnaval AC., Rodrigues MT., Moritz C. Revisiting the vanishing refuge model of diversification. *Frontiers in Genetics* 2014; **5**.

Domingos FMCB. Comparative phylogenomics, cryptic species and evolution of lizards in the cerrado biodiversity hotspot. Thesis, Flinders University. 2015.

Domingos FMCB., Bosque RJ., Cassimiro J., Colli GR., Rodrigues MT., Santos MG., Beheregaray LB. Out of the deep: Cryptic speciation in a Neotropical gecko (Squamata, Phyllodactylidae) revealed by species delimitation methods. *Molecular Phylogenetics and Evolution* 2014; **80**: 113–124.

Domingos FMCB., Colli GR., Lemmon A., Lemmon EM., Beheregaray LB. In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: Tropidurus). *Molecular Phylogenetics and Evolution* 2017; **107**: 455–465.

Eloy de Amorim M., Schoener TW., Santoro GRCC., Lins ACR., Piovia-Scott J., Brandão RA. Lizards on newly created islands independently and rapidly adapt in morphology and diet. *Proceedings of the National Academy of Sciences* 2017; **114**: 8812–8816.

Etienne RS., Rosindell J. Prolonging the Past Counteracts the Pull of the Present: Protracted Speciation Can Explain Observed Slowdowns in Diversification. *Systematic Biology* 2012; **61**: 204.

Fick SE., Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 2017; **37**: 4302–4315.

Fišer C., Robinson CT., Malard F. Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* 2018; **27**: 613–635.

Fitzpatrick BM., Fordyce JA., Gavrilets S. What, if anything, is sympatric speciation? *Journal of Evolutionary Biology* 2008; **21**: 1452–1459.

Fontaneto D., Kaya M., Herniou EA., Barraclough TG. Extreme levels of hidden diversity in microscopic animals (Rotifera) revealed by DNA taxonomy. *Molecular Phylogenetics and Evolution* 2009; **53**: 182–189.

Gascuel F., Ferrière R., Aguilée R., Lambert A. How Ecology and Landscape Dynamics Shape Phylogenetic Trees. *Systematic Biology* 2015; **64**: 590–607.

Gavrilets S. Evolution and speciation on holey adaptive landscapes. *Trends in Ecology & Evolution* 1997; **12**: 307–312.

Gavrilets S. Perspective: models of speciation: what have we learned in 40 years? *Evolution* 2003; **57**: 2197–2215.

Gavrilets S. Models of Speciation: Where Are We Now? *Journal of Heredity* 2014; **105**: 743–755.

Gehara M., Garda AA., Werneck FP., Oliveira EF., Da Fonseca EM., Camurugi F., Magalhães FDM., Lanna FM., Sites JW., Marques R., Silveira-Filho R., São Pedro VA., Colli GR., Costa GC., Burbrink FT. Estimating synchronous demographic changes across populations using HABC and its application for a herpetological community from northeastern Brazil. *Molecular Ecology* 2017; **26**: 4756–4771.

Gittenberger E. What about non-adaptive radiation? *Biological Journal of the Linnean Society* 1991; **43**: 263–272.

Gould SJ., Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 1979; **205**: 581–598.

Grant BR., Grant PR. Natural Selection in a Population of Darwin's Finches. *The American Naturalist* 1989;

Grant PR., Grant BR. Speciation and hybridization in island birds. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 1996; **351**: 765–772.

Grant PR., Grant BR. SPECIES BEFORE SPECIATION IS COMPLETE1. *Annals of the Missouri Botanical Garden* 2006; **93**: 94–102.

Hammem T., Hooghiemstra H. Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Reviews* 2000; **19**: 725–742.

Hansen TF., Houle D. Evolvability, stabilizing selection, and the problem of stasis. In: Pigliucci M, In: Preston KA eds. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford University Press 2004; 130–152.

Harmon LJ., Andreazzi CS., Débarre F., Drury J., Goldberg EE., Martins AB., Melián CJ., Narwani A., Nuismer SL., Pennell MW., Rudman SM., Seehausen O., Silvestro D., Weber M., Matthews B. Detecting the macroevolutionary signal of species interactions. *Journal of Evolutionary Biology* 2019; **32**: 769–782.

Harris DJ., Arnold EN., Thomas RH. Rapid Speciation, Morphological Evolution, and Adaptation to Extreme Environments in South African Sand Lizards (*Meroles*) as Revealed by Mitochondrial Gene Sequences. *Molecular Phylogenetics and Evolution* 1998; **10**: 37–48.

Henry CS. Singing and cryptic speciation insects. *Trends in Ecology & Evolution* 1994; **9**: 388–392.

Higgs PG., Derrida B. Stochastic models for species formation in evolving populations. *Journal of Physics A: Mathematical and General* 1991; **24**: L985–L991.

Higgs PaulG., Derrida B. Genetic distance and species formation in evolving populations. *Journal of Molecular Evolution* 1992; **35**: 454–465.

Hijmans RJ. The raster package. 2023;

Kadereit JW., Abbott RJ. Plant speciation in the Quaternary. *Plant Ecology & Diversity* 2021;

Kordbacheh A., Rahimian H., Fontaneto D. Mechanisms of reproductive isolation among cryptic species in monogonont rotifers. *Hydrobiologia* 2023; **850**: 4705–4718.

Kuhn M. Building Predictive Models in *R* Using the **caret** Package. *Journal of Statistical Software* 2008; **28**.

Kursa MB., Rudnicki WR. Feature Selection with the **Boruta** Package. *Journal of Statistical Software* 2010; **36**.

Lister AM. The impact of Quaternary Ice Ages on mammalian evolution (KJ Willis, KD Bennett, and D Walker, Eds.). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 2004; **359**: 221–241.

Losos JB., Glor RE., Kolbe JJ., Nicholson K. Adaptation, speciation, and convergence: a hierarchical analysis of adaptive radiation in caribbean anolis lizards. *Annals of the Missouri Botanical Garden* 2006; **93**: 24–33.

Losos JB., Mahler DL. Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. *Evolution since Darwin: the first 150*. 2010; 381–420.

Manzo F., Peliti L. Geographic speciation in the Derrida-Higgs model of species formation. *Journal of Physics A: Mathematical and General* 1994; **27**: 7079–7086.

Marques-Souza S., Machado Pellegrino KC., Oliveira Brunes T., Carnaval AC., Pacheco Damasceno R., Lima De Oliveira Borges M., Candia Gallardo C., Rodrigues MT. Hidden in the DNA: How multiple historical processes and natural history traits shaped patterns of cryptic diversity in an Amazon leaf-litter lizard *Loxopholis osvaldoi* (Squamata: Gymnophthalmidae). *Journal of Biogeography* 2020; **47**: 501–515.

Marquitti FMD., Fernandes LD., de Aguiar MAM. Allopatry increases the balance of phylogenetic trees during radiation under neutral speciation. *Ecography* 2020; **43**: 1487–1498.

Meiri S. Bergmann’s Rule – what’s in a name? *Global Ecology and Biogeography* 2011; **20**: 203–207.

Missagia RV., Casali DM., Patterson BD., Perini FA. Decoupled Patterns of Diversity and Disparity Characterize an Ecologically Specialized Lineage of Neotropical Cricetids. *Evolutionary Biology* 2023;

Moen D., Morlon H. Why does diversification slow down? *Trends in Ecology & Evolution* 2014; **29**: 190–197.

Nosil P. Speciation with gene flow could be common. *Molecular Ecology* 2008; **17**: 2103–2106.

Nosil P. *Ecological Speciation*. 2012. Oxford University Press.

Nosil P., Funk DJ., Ortiz-Barrientos D. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 2009; **18**: 375–402.

Nunes PMS., Fouquet A., Curcio FF., Kok PJR., Rodrigues MT. Cryptic species in *Iphisa elegans* Gray, 1851 (Squamata: Gymnophthalmidae) revealed by hemipenial morphology and molecular data: CRYPTIC SPECIES IN *IPHISA ELEGANS* GRAY, 1851. *Zoological Journal of the Linnean Society* 2012; **166**: 361–376.

Olave M., Avila LJ., Sites Jr JW., Morando M. Hidden diversity within the lizard genus Liolaemus: Genetic vs morphological divergence in the L. rothi complex (Squamata: Liolaeminae). *Molecular Phylogenetics and Evolution* 2016; **9**: 388–392.

Olave M., Avila LJ., Sites JW., Morando M. Hidden diversity within the lizard genus Liolaemus: Genetic vs morphological divergence in the L. rothi complex (Squamata:Liolaeminae). *Molecular Phylogenetics and Evolution* 2017; **107**: 56–63.

Partridge L., French V. Thermal evolution of ectotherm body size: why get big in the cold? In: Bennett AF, In: Johnston IA eds. *Society for Experimental Biology Seminar Series. Animals and Temperature: Phenotypic and Evolutionary Adaptation*. Cambridge: Cambridge University Press 1996; 265–292.

Phillimore A., Donoghue MJ., Levin SA., Mackay TFC., Rieseberg L., Travis J., Wray GA. Geography, Range Evolution, and Speciation. In: Losos JB, In: Baum DA, In: Futuyma DJ, In: Hoekstra HE, In: Lenski RE, In: Moore AJ, In: Peichel CL, In: Schluter D, In: Whitlock MJ eds. *The Princeton Guide to Evolution*. Princeton University Press 2014; 504–511.

Pigot AL., Phillimore AB., Owens IPF., Orme CDL. The Shape and Temporal Dynamics of Phylogenetic Trees Arising from Geographic Speciation. *Systematic Biology* 2010; **59**: 660–673.

Princepe D., Czarnobai S., Caetano RA., Marquitti FMD., de Aguiar MAM., Araujo SBL. Intermittent migration can induce pulses of speciation in a two-island system. *Evolution* 2024; 758–767.

Princepe D., Czarnobai S., Pradella TM., Caetano RA., Marquitti FMD., de Aguiar MAM., Araujo SBL. Diversity patterns and speciation processes in a two-island system with continuous migration. *Evolution* 2022; 2260–2271.

de Queiroz K. Species Concepts and Species Delimitation. *Systematic Biology* 2007; **56**: 879–886.

Ramírez-Reyes T., Velasco JA., Flores-Villela O., Piñero D. Decoupling in Diversification and Body Size Rates During the Radiation of Phyllodactylus: Evidence Suggests Minor Role of Ecology in Shaping Phenotypes. *Evolutionary Biology* 2022; **49**: 373–387.

Ridgeway G., Edwards D. gbm: Generalized Boosted Regression Models. 2024; 2.2.2.

van Rijssel JC., de Jong RCM., Kishe MA., Witte F. Rapid Evolutionary Responses in Cichlids: Genetics of Adaptation, Morphology and Taxonomic Implications. In: Abate ME, In: Noakes DLG eds. *The Behavior, Ecology and Evolution of Cichlid Fishes*. Dordrecht: Springer Netherlands 2021; 247–283.

Rolland J., Henao-Diaz LF., Doebeli M., Germain R., Harmon LJ., Knowles LL., Liow LH., Mank JE., Machac A., Otto SP., Pennell M., Salamin N., Silvestro D., Sugawara M., Uyeda J., Wagner CE., Schluter D. Conceptual and empirical bridges between micro- and macroevolution. *Nature Ecology & Evolution* 2023; **7**: 1181–1193.

Rousset F. Genetic Differentiation and Estimation of Gene Flow from *F* -Statistics Under Isolation by Distance. *Genetics* 1997; **145**: 1219–1228.

Salles M., Domingos F. Towards the next generation of species delimitation methods: an overview of Machine Learning applications. 2023;

Servedio MR., Brandvain Y., Dhole S., Fitzpatrick CL., Goldberg EE., Stern CA., Cleve JV., Yeh DJ. Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. *PLOS Biology* 2014; **12**: e1002017.

Silva S. Blondeau,Da, Silva A. Da. Pstat: An R Package to Assess Population Differentiation in Phenotypic Traits. *The R Journal* 2018; **10**: 447.

Singhal S., Hoskin CJ., Couper P., Potter S., Moritz C. A Framework for Resolving Cryptic Species: A Case Study from the Lizards of the Australian Wet Tropics. *Systematic Biology* 2018; **67**: 1061–1075.

Smith KL., Harmon LJ., Shoo LP., Melville J. Evidence of constrained phenotypic evolution in a cryptic species complex of agamid lizards: phenotypic stability in a cryptic radiation of australian lizards. *Evolution* 2011; **65**: 976–992.

Snyder CW. Evolution of global temperature over the past two million years. *Nature* 2016; **538**: 226–228.

Stankowski S., Cutter AD., Satokangas I., Lerch BA., Rolland J., Smadja CM., Segami Marzal JC., Cooney CR., Feulner PGD., Domingos FMCB., North HL., Yamaguchi R., Butlin RK., Wolf JBW., Coughlan J., Heidbreder P., Hernández-Gutiérrez R., Barnard-Kubow KB., Peede D., Rancilhac L., Salvador RB., Thompson KA., Stacy EA., Moyle LC., Garlovsky MD., Maulana A., Kantelinen A., Cacho NI., Schneemann H., Domínguez M., Dopman EB., Lohse K., Rometsch SJ., Comeault AA., Merrill RM., Scordato ESC., Singhal S., Pärssinen V., Lackey ACR., Kumar S., Meier JI., Barton N., Fraïsse C., Ravinet M., Kulmuni J. Toward the integration of speciation research. *Evolutionary Journal of the Linnean Society* 2024; **3**: kzae001.

Struck TH., Cerca J. Cryptic Species and Their Evolutionary Significance.John Wiley & Sons, Ltd ed. *eLS*. Wiley 2019; 1–9.

Struck TH., Cerca J. Connecting Micro- and Macro-Evolutionary Research: Extant Cryptic Species as Systems to Understand Macro-Evolutionary Stasis. In: Monro AK, In: Mayo SJ eds. *Systematics Association Special Volume Series*. *Cryptic Species: Morphological Stasis, Circumscription, and Hidden Diversity*. Cambridge: Cambridge

University Press 2022; 143–168.

Struck TH., Feder JL., Bendiksby M., Birkeland S., Cerca J., Gusarov VI., Kistenich S., Larsson KH., Liow LH., Nowak MD., Stedje B., Bachmann L., Dimitrov D. Finding Evolutionary Processes Hidden in Cryptic Species. *Trends in Ecology & Evolution* 2018; **33**: 153–163.

Tamura K., Stecher G., Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* 2021; **38**: 3022–3027.

Taylor RS., Friesen VL. The role of allochrony in speciation. *Molecular Ecology* 2017; **26**: 3330–3342.

Tessler M., Neumann JS., Kamm K., Osigus HJ., Eshel G., Narechania A., Burns JA., DeSalle R., Schierwater B. Phylogenomics and the first higher taxonomy of Placozoa, an ancient and enigmatic animal phylum. *Frontiers in Ecology and Evolution* 2022; **10**: 1–18.

Tonzo V., Papadopoulou A., Ortego J. Genomic data reveal deep genetic structure but no support for current taxonomic designation in a grasshopper species complex. *Molecular Ecology* 2019; **28**: 3869–3886.

Vanzolini PE., Williams EE. The vanishing refuge: A mechanism for ecogeographic speciation. *Papéis Avulsos de Zoologia* 1981; **34**: 251–255.

Vitt LJ., Shepard DB., Caldwell JP., Vieira GHC., França FGR., Colli GR. Living with your food: geckos in termitaria of Cantão. *Journal of Zoology* 2007; **272**: 321–328.

Wang IJ. Examining the Full Effects of Landscape Heterogeneity on Spatial Genetic Variation: A Multiple Matrix Regression Approach for Quantifying Geographic and Ecological Isolation. *Evolution* 2013; **67**: 3403–3411.

Wang IJ. Topographic path analysis for modelling dispersal and functional connectivity: Calculating topographic distances using the topoDistance r package. *Methods in Ecology and Evolution* 2020; **11**: 265–272.

Wegener JE., Gartner GEA., Losos JB. Lizard scales in an adaptive radiation: variation in scale number follows climatic and structural habitat diversity in *Anolis* lizards: Scale Variation in *Anolis* Lizards. *Biological Journal of the Linnean Society* 2014; **113**: 570–579.

Weir BS., Cockerham CC. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 1984; **38**: 1350–1370.

Werneck FP., Nogueira C., Colli GR., Sites Jr JW., Costa GC. Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *Journal of Biogeography* 2012; **39**: 1695–1706.

Wiens JJ. What Is Speciation and How Should We Study It? *The American Naturalist* 2004; **163**: 914–923.

Wojcieszek JM., Simmons LW. Evidence for stabilizing selection and slow divergent evolution of male genitalia in a millipede (antichiropus variabilis): genital evolution in a millipede. *Evolution* 2012; **66**: 1138–1153.

Wojcieszek JM., Simmons LW. Divergence in genital morphology may contribute to mechanical reproductive isolation in a millipede. *Ecology and Evolution* 2013; **3**: 334–343.

Zamora-Camacho FJ., Reguera S., Moreno-Rueda G. Bergmann's Rule rules body size in an ectotherm: heat conservation in a lizard along a 2200-metre elevational gradient. *Journal of Evolutionary Biology* 2014; **27**: 2820–2828.

Zhao Q., Zhang H., Wei J. Climatic niche comparison across a cryptic species complex. *PeerJ* 2019; **7**: 1–18.

Zink RM., Klicka J., Barber BR. The tempo of avian diversification during the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 2004; **359**: 215–220.

Zozaya SM., Higgie M., Moritz C., Hoskin CJ. Are Pheromones Key to Unlocking Cryptic

Lizard Diversity? *The American Naturalist* 2019; **194**: 168–182.