

SARA GONÇALVES BARBOSA

**VARIAÇÃO NO TAMANHO DOS GENOMAS REFLETE PROCESSOS
ADAPTATIVOS EM *Coffea* L.**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para obtenção do título de *Magister Scientiae*.

Orientador: Maicon Nardino

Coorientadora: Karla Suemy Clemente
Yotoko

**VIÇOSA - MINAS GERAIS
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

B238v
2022
Barbosa, Sara Gonçalves, 1988-
Variação no tamanho dos genomas reflete processos adaptativos em *Coffea* L. / Sara Gonçalves Barbosa. – Viçosa, MG, 2022.

1 dissertação eletrônica (41 f.): il. (algumas color.).

Orientador: Maicon Nardino.

Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Agronomia, 2022.

Referências bibliográficas: f. 37-41.

DOI: <https://doi.org/10.47328/ufvbbt.2022.181>

Modo de acesso: World Wide Web.

1. Filogenia. 2. Duplicação gênica. 3. Café - África. 4. Café - Ásia. 5. Café - Índico, Oceano, Ilhas do. I. Nardino, Maicon, 1988-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Genética e Melhoramento. III. Título.

CDD 22. ed. 633.73

Bibliotecário(a) responsável: Renata de Fátima Alves CRB6/2578

SARA GONÇALVES BARBOSA

**VARIAÇÃO NO TAMANHO DOS GENOMAS REFLETE PROCESSOS
ADAPTATIVOS EM *Coffea* L.**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para obtenção do título de *Magister Scientiae*.

APROVADA: 23 de março de 2022.

Assentimento:



Sara Gonçalves Barbosa
Autora



Maicon Nardino
Orientador

Às famílias agricultoras, que alimentam o mundo
Dedico

AGRADECIMENTOS

- À CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - pelo suporte financeiro.
- À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realização do curso e pela infraestrutura fornecida.
- À Secretaria do Programa de Pós-Graduação em Genética e Melhoramento nas pessoas de Marco Túlio e Odilon, pelo auxílio e dedicação em todos os momentos.
- Ao Dr. Tiago Olivoto por ter contribuído com a análise dos dados de temperatura da Nasa.
- Ao Prof. Maicon pela orientação serena, pelo acolhimento, e por fornecer conselhos e ensinamentos que contribuíram com o meu desenvolvimento profissional e científico.
- À Prof.^a Loreta pela participação na banca avaliadora e pelas colaborações fundamentais para a melhoria deste trabalho.
- À Karla pelo inestimável apoio e importância em minha trajetória acadêmica e na caminhada da vida. Não existem palavras para expressar o quão importante têm sido seus ensinamentos científicos, as broncas (rs), os conselhos francos para uma vida toda.
- Às colegas do Laboratório de Bioinformática e Evolução pelas contribuições ao desenvolvimento do pensamento científico, pelos riquíssimos diálogos realizados no .
- Aos meus pais por terem me dado a vida e à minha avó que me deu uma nova vida e partiu, mas teve a oportunidade de se orgulhar da educação que me ofereceu, meus mais sinceros agradecimentos. Aos meus sobrinhos Gabriel, Ana Luiza, Yuri e Igor por me motivarem a seguir na ciência e educação. Às minhas tias Silvia e Sylvania e à Bia pelo acolhimento certo. Ao meu irmão Marcos e ao Isaque e Aline por terem me oferecido um lar digno e tornado possível que eu prosseguisse com os estudos.

À Raquel por ter me ensinado que “o inverno nunca falha em se tornar primavera” (Nitiren Daishonin).

Às amigadas de bem perto (Daniel, Cainã, Rô, Maria) que me acompanharam nos momentos de crises e risos e às de longe, (Tercia, Cris Jacob, Henrique) pelos ombros amigos nos momentos de desespero, por todos os momentos divertidos que trouxeram a alegria necessária para seguirmos em frente.

À Mari por todos os cafés e intermináveis prosas sobre ciência, filosofia, feminismo e amores.

À Joane pela franqueza com a qual construímos nossa amizade e por todo o apoio com os desafios da pós-graduação.

À Camila por estar me ajudando em meu processo de autoconhecimento, ao Lairton pelo acolhimento num momento delicado deste trabalho, à Sylvia Franceschini por estar sempre disposta a ajudar a melhorar nossa experiência acadêmica.

A todos os cientistas que me antecederam e geraram os conhecimentos que tornaram possível a realização deste trabalho, especialmente àqueles que acreditam no conhecimento para a superação de nossas mazelas e melhoria da nossa qualidade de vida.

Ao Partido dos Trabalhadores, nas figuras de Dilma e Lula, por terem implementado políticas públicas de democratização do acesso ao ensino superior e estimulado a classe trabalhadora a ocupar este lugar, seu por direito!

À Alexandra Elbakyan e portal Capes pela democratização do acesso ao conhecimento científico. Às políticas públicas de inclusão das filhas e filhos de trabalhadores nas universidades, a todos os militantes que dão seu suor e muitas vezes o sangue na luta por justiça social.

A todas as professoras e professores que atuaram no percurso de minha preparação para a vida e para a academia. Como nos deixou o memorável Paulo Freire, “O educador se eterniza em cada ser que educa”.

“Liberdade pra mim é isso, não ter medo”.

Nina Simone

RESUMO

BARBOSA, Sara Gonçalves, M.Sc., Universidade Federal de Viçosa, março de 2022. **Variação no tamanho dos genomas reflete processos adaptativos em *Coffea* L.** Orientador: Maicon Nardino. Coorientadora: Karla Suemy Clemente Yotoko.

O paradoxo C, ou a ausência de correlação entre o tamanho do genoma e a complexidade dos organismos vem provocando a ciência a explicar tal falta de correlação. Diante disto, alguns autores vêm mostrando que, em alguns grupos, o tamanho do genoma se correlaciona significativamente com algum caráter fenotípico, o que vem sendo interpretado como possível papel do tamanho do genoma na adaptação dos organismos. Neste trabalho, investigamos o sinal filogenético da variação do tamanho dos genomas das espécies diploides de *Coffea* L. que ocorrem naturalmente na África, Ásia e ilhas do Oceano Índico, na costa oriental da África. Nós utilizamos literatura atualizada e bancos de dados na busca dos tamanhos de genoma propriamente ditos, do teor de cafeína e das áreas de ocorrência de cada espécie para determinar a latitude mínima e máxima, além da altitude máxima onde cada espécie ocorre. Estimamos ainda a temperatura mínima das estações frias e a máxima das estações quentes que atingiram cada uma das coordenadas de ocorrência de cada espécie nos últimos dez anos. Nossos resultados mostraram que o tamanho do genoma apresenta sinal filogenético e que este caráter se correlacionou significativamente com o teor de cafeína das diferentes espécies e com as latitudes de ocorrência das mesmas. Encontramos ainda evidências de que, diferentemente de hipóteses anteriores, a produção de cafeína foi herdada dos ancestrais do gênero, ainda que em pequenas doses. Segundo nossos resultados, algumas espécies perderam a capacidade de produzir a substância enquanto outras apresentaram aumentos significativos no teor de cafeína, o que ocorreu independentemente duas vezes, uma no clado de espécies que ocupam regiões da África Central e Ocidental mais próximas ao Equador e outra em espécies que se estabeleceram em ambientes específicos em Madagascar. Nossos resultados sugeriram ainda que a variação do valor-C pode ser considerada um processo adaptativo, desde que investigada em taxa filogeneticamente próximos. Observamos também que mesmo em grupos com tamanhos de genoma reduzidos e

pouco variáveis podem ser detectados eventos raros de aumento do genoma, que podem ter contribuído para a adaptação de algumas espécies a seus ambientes.

Palavras-chave: Métodos filogenéticos comparativos. África, Ásia. Ilhas do Oceano Índico. Duplicação gênica. Perda de genes codificadores.

ABSTRACT

BARBOSA, Sara Gonçalves, M.Sc., Universidade Federal de Viçosa, March 2022. **Variation in genome sizes reflects adaptative processes in *Coffea* L.** Adviser: Maicon Nardino. Co-adviser: Karla Suemy Clemente Yotoko.

The C-paradox, or the correlation absence between genome size and organisms' complexity, has been inducing the science to explain such absence of correlation. Regardless, some authors have been showing that, in some groups, the genome size correlates significantly with some phenotypic character, which they interpreted as a putative adaptative role of the trait. In this work, we investigated the phylogenetic signal of genome size variation in diploid species of *Coffea* L. that naturally occur in Africa, Asia, and the Indian Ocean islands near the east coast of Africa. First, we searched for specific genome sizes, caffeine content, and areas of occurrence in updated literature and databases. Then, based on the geographic coordinates that the individuals occur, we determined the maximum altitude and the minimum and maximum latitude of each species. Finally, we estimated the minimum temperature of the cold seasons and the maximum temperature of the warm seasons of each coordinate to determine the temperature extremes over the last ten years. Our results revealed that the genome size has phylogenetic signal correlates significantly with the caffeine content and extreme latitudes. We also found that the current *Coffea* species inherited caffeine from the genus's ancestral, different from previous findings. According to our results, such ancestral produced a small dose of caffeine. Some species lost the ability to produce the substance. In contrast, others underwent significant increases in caffeine content, which occurred twice, the first in the clade of Central and Occidental Africa species close to the Equator and the second in a specific clade of Madagascar species. Our results also allow us to consider the C-value variation as an adaptive process, provided it is investigated in phylogenetically close taxa. We also saw that even in groups with small and poorly variable genome sizes, it is possible to detect rare events of genome increases, which may have contributed to the adaptation of some species to their environments.

Keywords: Comparative phylogenetic methods. African, Asian. Indian ocean islands.
Genic duplication. Loss of coding genes.

SUMÁRIO

1. INTRODUÇÃO	12
2. MATERIAIS E MÉTODOS.....	14
2.1 Obtenção de dados	14
2.2 Parâmetros descritivos da evolução dos caracteres em estudo	15
2.3 Inferência Ancestral	17
2.4 Correlações entre caracteres	17
3. RESULTADOS E DISCUSSÃO.....	18
4. DISCUSSÃO	33
REFERÊNCIAS.....	37

1. INTRODUÇÃO

A ausência de correlação entre o tamanho dos genomas (quantidade de DNA nos núcleos das células) e a complexidade dos organismos, conhecida como paradoxo do valor-C (Thomas, 1971; Gregory, 2001a), tem provocado a ciência a explicar por que alguns organismos apresentam genomas enormes, enquanto outros parecem ter um genoma enxuto, a ponto de a diferença de tamanho do genoma de plantas ser de até 2400 vezes (Pellicer et al., 2010; Fleischmann et al., 2014; Pellicer and Leitch, 2019).

Mesmo diante do paradoxo, alguns autores argumentaram que quando o tamanho do genoma se correlaciona significativamente com algum caráter morfológico, ecológico ou funcional ao longo da evolução de um grupo, pode ter tido algum papel adaptativo (Bennet, 1972; Knight and Ackerly, 2002; Achigan-Dako et al., 2008; Yotoko et al., 2011; Razafinarivo et al., 2012; Kang et al., 2014; Du et al., 2017; Khorami et al., 2018; Souza et al., 2019; Gardner et al., 2020; Shao et al., 2021)

O gênero *Coffea* é composto por 124 espécies (Hamon et al., 2017a), cujos genomas variam de 0,48 pg em *Coffea mauritiana* Lam. (1785) a 0,92 pg em *Coffea humilis* A.Chev. (1907) (Noirot et al., 2003; Razafinarivo et al., 2012). Entre os taxa do gênero, existe apenas uma espécie poliploide, *Coffea arabica* L. (1753), produzido pela quebra de barreiras ao fluxo gênico de outras duas espécies (Rimlinger et al., 2020). Com isso, todas as demais espécies atuais são diploides. Chaparro et al. (2015), mostraram que todas as espécies diploides de *Coffea* apresentam basicamente a mesma organização do genoma, que se diferenciou apenas pelo número de constrições secundárias nos cromossomos, número e localização de *loci* de RNA ribossômico, bem como número e distribuição de elementos repetitivos. Hamon et al., (2011); Yuyama et al., (2012) e Dupeyron et al., (2017) atribuíram a variação de tamanho dos genomas de *Coffea* à amplificação de elementos transponíveis.

Denoeud et al., (2014) mostraram que o genoma de *Coffea canephora* Pierre ex A.Froehner (1897), a espécie com o maior teor de cafeína registrado no gênero, inclui diversas expansões de famílias gênicas específicas, entre elas as N-metiltransferases (NMTs), envolvidas na produção de cafeína, além de genes relacionados à defesa como alcaloides e flavonoides, envolvidos na síntese de

compostos secundários. É importante notar que nem todas as *Coffea* produzem cafeína, de modo que Denoeud et al., (2014) e Hamon et al., (2017) sugeriram que a produção da substância teria surgido mais de uma vez por convergência no gênero, assim como surgiu em outros grupos vegetais como nas espécies utilizadas na produção de chá e chocolate (Kang et al., 2014; Xu et al., 2020; Shao et al., 2021).

Diante do exposto, nos perguntamos se o tamanho do genoma estaria correlacionado com a capacidade das espécies de *Coffea* para ocupar ambientes específicos (adaptação ao ambiente) e com o teor de cafeína. apresentado pelas diferentes espécies de *Coffea*. De fato, (Hamon et al., 2017) sugeriram que o alto teor de cafeína presente em algumas espécies as adaptou a ocupar terras baixas próximas ao Equador na África Continental (Oeste e Centro do continente).

Tal correlação precisa ser calculada utilizando métodos filogenéticos comparativos, que requerem uma hipótese filogenética robusta para o grupo em estudo. Utilizando a hipótese filogenética inferida por (Maurin et al., 2007), Razafinarivo et al., (2012) fizeram uma primeira tentativa de estudar o papel das modificações no tamanho dos genomas na adaptação de espécies de *Coffea*. No entanto, a hipótese de (Maurin et al., 2007) como a de outros autores (Davis et al., 2007; Anthony et al., 2010; Nowak et al., 2012) apresenta numerosas politomias basais que dificultaram a aplicação de métodos comparativos e levaram Razafinarivo et al. (2012) a concluir que o tamanho dos genomas não apresenta sinal filogenético no gênero *Coffea*, impedindo o cálculo das correlações. Mais tarde, Hamon et al. (2017) disponibilizaram uma hipótese filogenética bem resolvida para o gênero e a utilizaram para investigar a evolução do teor de cafeína e estudar a distribuição geográfica das espécies de café, sem investigar a correlação dessas características com os tamanhos do genoma das diferentes espécies.

O presente trabalho visou, portanto, a investigação do papel do tamanho do genoma na adaptação das espécies diploides de *Coffea* a diferentes ambientes. Reavaliamos o sinal filogenético do tamanho dos genomas das espécies provenientes dos continentes africano e asiático, bem como das principais ilhas oceânicas da região, tendo como base a filogenia proposta por Hamon et al. (2017). Posteriormente, testamos as correlações filogenéticas do tamanho do genoma com o teor de cafeína, com a resistência a temperaturas mínima e máxima e com a ocupação de latitudes mínima e máxima.

2. MATERIAIS E MÉTODOS

2.1 Obtenção de dados

A hipótese filogenética para as espécies do gênero *Coffea* foi obtida por Hamon et al. (2017) a partir de genotipagem por sequenciamento - GBS (genotyping by sequence). A partir desta árvore, fizemos buscas em bancos de dados e na literatura pelo valor 1C (tamanho do genoma haploide - em picogramas – banco de dados do Kew Gardens - <https://cvalues.science.kew.org/search/angiosperm> - (Pellicer and Leitch, 2019) e teor de cafeína (em % mdb – teor de cafeína na massa seca dos frutos verdes) para todas as espécies presentes na filogenia. Como não havia dados do valor 1C e do teor de cafeína para todas as espécies da filogenia, espécies sem dados para ambos os caracteres foram retiradas da filogenia com auxílio dos pacotes Ape (Paradis and Schliep, 2019), Phangorn (Schliep, 2011), Phytools (Revell, 2012) e Geiger (Pennell et al., 2014) do software R (Team, R. Core, 2021).

Para cada espécie em estudo, verificamos a distribuição geográfica no checklist anual do *Catalogue of Life* - <http://www.catalogueoflife.org/annual-checklist/2019/> - (Govaerts, 2019; Roskov et al., 2019). A partir disso, consultamos o GBIF – *Global Biodiversity Information Facility* - <https://www.gbif.org/species/2895315> - (Edwards James L., 2004) para obter as coordenadas geográficas nas quais cada espécie foi registrada. Para evitar pontos espúrios, só consideramos os relatos de espécimes vivos ou de exsicatas preservadas em herbário cujos acessos disponibilizam as coordenadas geográficas das coletas. Só consideramos os pontos de GPS dentro da faixa latitudinal de 36°N a 36°S e longitudinal de 33°O a 150°L, compreendendo os espécimes relatados nos continentes africano e em suas ilhas, além do continente asiático, considerados como possíveis centros de origem das espécies do gênero.

Uma vez obtidas as coordenadas geográficas nas quais todas as espécies em estudo ocorrem, obtivemos a latitude máxima e mínima de ocorrência de cada espécie. Este valor foi considerado em módulo, pois estávamos interessados apenas na distância até o Equador, região em que as temperaturas e a diversidade de competidoras e predadores costumam ser maiores. Os dados de elevação foram obtidos a partir das coordenadas geográficas, que foram inseridas na versão gratuita do programa Google Earth Pro - <https://earth.google.com/web/> (Lisle, 2006). As

temperaturas máxima e mínima reportadas em cada ponto nos últimos 10 anos nas estações mais quentes e mais frias do ano foram obtidas no portal Nasa Power - <https://power.larc.nasa.gov/data-access-viewer/> (Sparks, 2018). A filtragem dos dados disponíveis no Portal, bem como as estimativas de temperaturas máximas e mínimas foi feita com o auxílio do pacote EnvRtype - https://github.com/gcostaneto/EnvRtype_course/blob/main/MODULE1.md - (Costa-Neto et al., 2021) do software R (Team, R. Core, 2021).

2.2 Parâmetros descritivos da evolução dos caracteres em estudo

Uma vez obtidos os dados de todos os caracteres em estudo (tamanho de genoma, teor de cafeína, latitudes mínima e máxima de ocorrência de cada espécie, elevação máxima e temperaturas mínima e máxima registradas para cada espécie nos últimos dez anos), testamos, com base na hipótese filogenética pré-editada, se estas características apresentam sinal filogenético. Para tanto, utilizamos o programa BayesTraits, versão 3.0.2, disponível em <http://www.evolution.reading.ac.uk/BayesTraits.html> (Pagel et al., 2004) que além de verificar se cada característica apresentou tendências a aumentos ou reduções ao longo da evolução, estabelece se há sinal filogenético, mede a influência dos tamanhos dos ramos e das ramificações nos valores de cada caráter contínuo em estudo e estima seus respectivos estados ancestrais.

O primeiro passo lógico foi plotar os dados das espécies atuais de cada característica em estudo na hipótese filogenética. Feito isto, testamos se cada característica tendeu a aumentos ou reduções ao longo do tempo (modelo B) ou se simplesmente aumentou e diminuiu de acordo com o esperado pelo movimento Browniano (modelo A). A comparação dos dois modelos, bem como a comparação de todos os modelos testados nesse trabalho, foi feita a partir da estimativa do BayesFactor, dada por $2 \times (\log \text{ da verossimilhança marginal do modelo mais complexo} - \log \text{ da verossimilhança marginal do modelo mais simples})$. Qualquer valor acima de 2,0 indica que o modelo mais complexo é o mais apropriado para descrever a evolução do caráter em estudo. O log da verossimilhança marginal foi calculado pela função “stones” do programa Bayestraits. Valores abaixo de 2,0 indicam que o modelo mais simples é o mais apropriado.

Após determinado o modelo de evolução, A ou B, é necessário também estimar os valores dos parâmetros kappa (κ), lambda (λ) e delta (Δ) (Pagel, 1999a, 1999b), que assumem, por *default*, o valor 1,0. Quando o valor dos três parâmetros é 1,0 (o modelo mais simples), a hipótese filogenética que baseia a análise, com seus tamanhos de ramos, descreve acuradamente a evolução dos caracteres em estudo sob o modelo de movimento Browniano (aumentos e reduções aleatórias ao longo do tempo que fazem com que espécies próximas sejam mais parecidas do que espécies distantes). No entanto, valores significativamente diferentes de 1,0 devem ser devidamente estimados e incorporados à análise dos dados para calcular correlações e determinar ancestrais com maior confiabilidade.

Especificamente, $\kappa = 1$ indica que a evolução do caráter em estudo foi gradual e proporcional aos tamanhos de ramos. Valores maiores que 1,0 indicam que ramos mais longos contribuem proporcionalmente mais com as modificações no caráter e os menores que 1,0 comprimem os ramos longos mais que os ramos curtos, indicando a falta de correlação do tamanho do ramo com as modificações do caráter, ou seja, que o caráter não se modificou lenta e gradualmente ao longo do tempo. $\kappa = 0,0$ é consistente com o modelo de equilíbrio pontuado (Eldredge and Gould, 1972), sugerindo que as modificações ocorreram durante o processo de especiação. Em cada caso, é necessário testar, via BayesFactors, se o valor de κ é significativamente diferente de 1, ($\kappa = 1,0$ é o modelo mais simples) e, em caso positivo, se o valor é diferente de zero (nesse caso, $\kappa = 0$ é o modelo mais simples).

O parâmetro δ indica se a evolução do caráter foi mais acentuada na base da filogenia ou nos nós atuais. Quando $\delta = 1,0$, as modificações ocorreram uniformemente ao longo da evolução do grupo. Se $\delta < 1,0$ as modificações mais importantes ocorreram próximo à base, enquanto $\delta > 1,0$ sugere que as modificações mais importantes ocorreram recentemente. Neste caso, é necessário utilizar o BayesFactors para verificar se o valor de δ estimado é diferente de 1,0.

λ é o parâmetro que indica se a filogenia descreve as similaridades entre as OTUs em uma dada característica. Quando $\lambda = 0,0$, o caráter é independente da filogenia e a inferência de estados ancestrais para o cálculo de correlações entre os caracteres é inapropriada. Valores intermediários de λ (menores que 1,0) sugerem que a topologia da árvore superestima a covariância entre as OTUs. Mais uma vez, é necessário testar se os valores obtidos de λ são diferentes de 1,0 ou de 0,0 (por BayesFactors).

Ao final de cada execução do BayesTraits, o programa Tracer, versão 1.7.1 (Rambaut et al., 2018), foi utilizado para verificar se os valores dos parâmetros foram estimados adequadamente. Uma corrida ideal resulta em uma curva estreita de densidade de probabilidade, com desvio padrão menor que a média para todos os parâmetros a serem estimados. Outro indicativo de que o processo correu por tempo suficiente é o ESS, que corresponde a uma estimativa de tamanho amostral (tamanho da corrida). A regra a ser seguida aqui é que o ESS deve ser maior que 200, caso contrário é necessário repetir a análise com mais passos (iterações) na cadeia de Markov.

2.3 Inferência Ancestral

Após a determinação do modelo (A ou B) e dos parâmetros apropriados (κ , δ e λ) para cada característica, foi feita a inferência dos dados faltantes (espécies cujos valores de alguma característica não tenham sido determinados) e dos nós ancestrais com o auxílio da ferramenta BayesTraits, versão 3.0.2. Esta inferência se baseia na topologia da árvore, nos tamanhos dos ramos, nos valores indicados para cada espécie e nos parâmetros estimados.

Como os valores são inferidos por uma Cadeia de Markov que converge para as maiores probabilidades posteriores, o BayesTraits gera um valor médio e um desvio padrão *a posteriori*, inferidos a partir da curva de densidade de probabilidade de cada inferência. Para tanto, utilizamos o *default* do programa BayesTraits 3.2.0, com 1010000 iterações e Burn in de 10000. Em função disto, arbitramos que a modificação de um caráter de um nó interno para o outro (na sequência mostrada na árvore filogenética – Figura 1) foi significativa quando a média posterior inferida para um nó derivado (por exemplo o nó 1 na Figura 1) foi maior que a média mais o desvio do nó ancestral (por exemplo o nó 2 da Figura 1) ou menor que a média menos o desvio do nó ancestral (Bauters et al., 2016).

2.4 Correlações entre caracteres

Após a determinação dos parâmetros K , Δ e Λ , fizemos também a estimativa da correlação entre o tamanho de genoma e os demais caracteres estudados. Para tanto, comparamos, via Bayes Factors (BF), a verossimilhança marginal considerando que a correlação entre eles é igual a zero (modelo mais simples) com a verossimilhança marginal considerando a correlação (modelo mais complexo).

3. RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os locais de ocorrência das espécies em estudo, bem como os nomes científicos e sinônimos aceitos na literatura corrente. Conforme consta na Tabela 2, o tamanho do genoma evoluiu gradualmente e as similaridades entre as OTUs foi superestimada pela filogenia ($\lambda = 0,78$, $dp = 0,12$). A Figura 1 mostra a topologia e os tamanhos de ramos da árvore editada de Hamon et al. (2017). Seus ramos foram coloridos em função do tamanho do genoma, que foi obtido da literatura para parte das espécies atuais e inferido para outras (Tab. 3), bem como para os nós internos (Tab. 4). Setas para cima ou para baixo seguidas da letra grega alpha (α) indicam alterações significativas do tamanho do genoma nos nós internos da árvore. Considerando um genoma ancestral (nó 67) de 0,618 pg, com desvio padrão (dp) de 0,03 (Tab. 4), o genoma apresentou poucas modificações significativas, tendo sido reduzido significativamente no nó 18 para 0,559 pg ($dp = 0,03$), no nó 40 para 0,568 pg ($dp = 0,03$) e no nó 57 para 0,58 pg ($dp = 0,04$). Houve também um aumento significativo no nó 55 (0,698 pg; $dp = 0,03$), seguido de um segundo aumento no nó 54 (0,749 pg, $dp = 0,04$) (Fig. 1, Tab. 4).

Ainda na Figura 1, simbolizamos os valores atuais (Tab. 3) e ancestrais (Tab. 4) do teor de cafeína em círculos coloridos, medida pelo percentual de cafeína na massa seca de frutos verdes (%mdb), que, de acordo com a Tabela 2, evoluiu sob o modelo de equilíbrios pontuados, $\kappa = 0$ (Eldredge and Gould, 1972), sugerindo que o teor de cafeína se alterou durante a especiação das OTUs. As alterações significativas no teor de cafeína dos nós ancestrais foram simbolizadas por setas para cima ou para baixo seguidas da letra grega beta (β). Considerando um ancestral que produzia uma quantidade muito pequena (0,165 %mdb, $dp = 0,11$) de cafeína, foi possível detectar dois aumentos significativos a partir da base, o primeiro no nó 11, para 0,576 %mdb ($dp = 1,99$) e o segundo no nó 60, para 0,547 %mdb ($dp = 0,2$).

O aumento do nó 11 não resultou em espécies atuais com grandes quantidades de cafeína. No entanto, o aumento do teor de cafeína do nó 60 para 0,847 % mdb ($dp = 0,20$), acompanhado por um aumento do tamanho do genoma no nó 55, foi seguido por sucessivos aumentos no teor de cafeína nos nós 52 (1,22 %mdb; $dp = 0,22$), 51 (1,66 %mdb; $dp = 0,21$), 46 (1,96 %mdb; $dp = 0,2$) e 42 (2,4 %mdb; $dp = 0,19$) (Fig.1, Tab. 4), resultando nas espécies com maior teor de cafeína

do gênero: *C. canephora* (2,7 %mdb) e *Coffea brevipes* Hiern (1876) (2,3 %mdb) . Após o aumento no nó 60, houve também aumentos significativos nos nós 57 (0,7 %mdb; dp = 0,19) e 56 (0,9 %mdb; dp = 0,18) (Fig. 1, Tab. 4), que, no entanto, assim como no nó 11, não resultaram em espécies que produzem grandes quantidades de cafeína, talvez por não terem sido acompanhadas de um aumento, mas por uma redução significativa no tamanho do genoma. No nó 58, houve redução no teor de cafeína para 0,18 %mdb (dp = 0,19).

Tabela 1: Unidades Taxonômicas Operacionais (OTUs) em estudo (os mesmos presentes na hipótese filogenética proposta por Hamon et al., 2017), seus nomes específicos (quando disponíveis), sinônimos aceitos na literatura (Govaerts, 2019; Roskov et al., 2019). e local de ocorrência de acordo com checklist anual do *Catalogue of Life*.

OTU	Espécie	sinônimo aceito	País
C. ambodirianensis	<i>Coffea millotii</i> J.-F.Leroy	<i>Coffea ambodirianaensis</i> Portères	Madagascar
C. dolichophylla	<i>Coffea millotii</i> J.-F.Leroy	<i>Coffea dolichophylla</i> J.-F.Leroy	Madagascar
C. millotii	<i>Coffea millotii</i> J.-F.Leroy	<i>Coffea ambodirianaensis</i> Portères/ <i>Coffea dolichophylla</i> J.-F.Leroy	Madagascar
C. abbayesii	<i>Coffea abbayesii</i> J.-F.Leroy	-	Madagascar
C. farafanganensis	<i>Coffea farafanganensis</i> J.-F.Leroy	-	Madagascar
C. richardii	<i>Coffea richardii</i> J.-F.Leroy	-	Madagascar
C. bertrandii	<i>Coffea bertrandii</i> A.Chev.	-	Madagascar
C. resinosa	<i>Coffea resinosa</i> (Hook.f.) Radlk.	-	Madagascar
C. coursiana	<i>Coffea coursiana</i> J.-F.Leroy	-	Madagascar
C. vianneyi	<i>Coffea vianneyi</i> J.-F.Leroy	-	Madagascar
C. montis-sacri	<i>Coffea montis-sacri</i> A.P.Davis	-	Madagascar
C. arenesiana	<i>Coffea arenesiana</i> J.-F.Leroy	-	Madagascar
C. kianjavatensis	<i>Coffea kianjavatensis</i> J.-F.Leroy	-	Madagascar
C. lancifolia	<i>Coffea lancifolia</i> A.Chev.	-	Madagascar
C. tricalysioides	<i>Coffea tricalysioides</i> J.-F.Leroy	-	Madagascar
C. toshii	<i>Coffea toshii</i> A.P.Davis & Rakotonas.	-	Madagascar
C. vatovavyensis	<i>Coffea vatovavyensis</i> J.-F.Leroy	-	Madagascar
C. sakarahae	<i>Coffea sakarahae</i> J.-F.Leroy	-	Madagascar
C. boiviniana	<i>Coffea boiviniana</i> (Baill.) Drake	-	Madagascar
C. bonnierii	<i>Coffea bonnierii</i> Dubard	-	Madagascar
C. homollei	<i>Coffea homollei</i> J.-F.Leroy	-	Madagascar

Tabela 1 (continuação)

OTU	Espécie	sinônimo aceito	País
C. andrambovatensis	<i>Coffea andrambovatensis</i> J.-F.Leroy	-	Madagascar
C. leroyi	<i>Coffea leroyi</i> A.P.Davis	-	Madagascar
C. perrieri	<i>Coffea perrieri</i> Drake ex Jum. & H.Perrier	-	Madagascar
C. ankaranensis	<i>Coffea ankaranensis</i> J.-F.Leroy ex A.P.Davis & Rakotonas.	-	Madagascar
C. sahafaryensis	<i>Coffea sahafaryensis</i> J.-F.Leroy	-	Madagascar
C. dubardii	<i>Coffea dubardii</i> Jum.	-	Madagascar
C. heimii	<i>Coffea heimii</i> J.-F.Leroy	-	Madagascar
C. mogenetii	<i>Coffea mogenetii</i> Dubard	-	Madagascar
C. tetragona	<i>Coffea tetragona</i> Jum. & H.Perrier	-	Madagascar
C. jumellei	<i>Coffea jumellei</i> J.-F.Leroy	-	Madagascar
C. tsirananae	<i>Coffea tsirananae</i> J.-F.Leroy	-	Madagascar
C. humblotiana	<i>Coffea humblotiana</i> Baill.	-	Comores
C. pervilleana	<i>Coffea pervilleana</i> (Baill.) Drake	-	Madagascar
C. ratsimamangae	<i>Coffea ratsimamangae</i> J.-F.Leroy ex A.P.Davis & Rakotonas.	-	Madagascar
C. augagneurii	<i>Coffea augagneurii</i> Dubard	-	Madagascar
C. humberitii	<i>Coffea humberitii</i> J.-F.Leroy	-	Madagascar
C. mauritiana-M	<i>Coffea mauritiana</i> Lam.		Madagascar; Ilhas Maurício; Ilhas Reunião
C. mauritiana-R	<i>Coffea mauritiana</i> Lam.		Madagascar; Ilhas Maurício; Ilhas Reunião
C. bernardiniana	<i>Coffea macrocarpa</i> A.Rich.	<i>Coffea bernardiniana</i> J.-F.Leroy	-
C. macrocarpa	<i>Coffea macrocarpa</i> A.Rich.	<i>Coffea bernardiniana</i> J.-F.Leroy	Ilhas Maurício
C. myrtifolia	<i>Coffea myrtifolia</i> (A.Rich. ex DC.) J.-F.Leroy	-	Ilhas Maurício
C. brevipes	<i>Coffea brevipes</i> Hiern	-	Camarões; Congo; Gabão; Zaire

Tabela 1 (continuação)

OTU	Espécie	sinônimo aceito	País
C. canephora	<i>Coffea canephora</i> Pierre ex A.Froehner	-	Angola; Benin; Bornéu; Cabinda; Camarões; República Centro Africana; Congo; Gabão; Gana; Guiné; Guiné- Bissau; Costa do Marfim; Ilhas de Sotavento Britânicas. Libéria; Nigéria; Senegal; Sudão; Tanzânia; Togo; Uganda; Ilhas de Barlavento; Zaire
C. congensis	<i>Coffea congensis</i> A.Froehner	-	Camarões; República Centro Africana; Congo; Gabão; Zaire
C. spNGongo2	ND	-	-
C. spCongo	ND	-	-
C., pNGongo3	ND	-	-
C. dewevrei	<i>Coffea liberica</i> var. <i>dewevrei</i> (De Wild. & T.Durand) Lebrun	-	

Tabela 1 (continuação)

OTU	Espécie	sinônimo aceito	País
C. liberica	<i>Coffea liberica</i> W.Bull	-	Ilhas de Andamão; Angola; Benin; Cabinda; Camarões; República Centro africana; Chade; Congo; Gabão; Gana; Guiné; Ilhas do Golfo da Guiné; Costa do Marfim; Ilhas de Sotavento; Libéria; Nigéria; Seychelles; Sudão; Togo; Uganda; Ilhas de Barlavento; Zaire
C. spKoto	ND	-	-
C. humilis	<i>Coffea humilis</i> A.Chev.	-	Costa do Marfim; Liberia; Serra Leoa
C. stenophylla	<i>Coffea stenophylla</i> G.Don	-	Guiné; Golfo da Guiné; Costa do Marfim; Liberia; Serra Leoa
C. kapakata	<i>Coffea kapakata</i> (A.Chev.) Bridson	-	Angola; Congo
C. eugenioides	<i>Coffea eugenioides</i> S.Moore	-	Burundi; Quênia; Ruanda; Sudão; Tanzânia; Uganda; Zaire
C. heterocalyx	<i>Coffea heterocalyx</i> Stoff.	-	Camarões
C. costatifructa	<i>Coffea costatifructa</i> Bridson	-	Tanzânia

Tabela 1 (continuação)

OTU	Espécie	sinônimo aceito	País
C. racemosa	<i>Coffea racemosa</i> Lour.	-	KwaZulu-Natal; Moçambique ; Canal I de Moçambique ; Zimbábue
C. pocsii-1	<i>Coffea pocsii</i> Bridson	-	Tanzânia
C. pseudozanguebariae	<i>Coffea pseudozanguebariae</i> Bridson	-	Quênia; Tanzânia
C. salvatrix	<i>Coffea salvatrix</i> Swynn. & Philipson	-	Malawi; Moçambique ; Tanzânia; Zimbábue
C. charrieriana-1	<i>Coffea charrieriana</i> Stoff. & F.Anthony	-	Camarões
P. benghalensis	<i>Coffea benghalensis</i> B.Heyne ex Schult.	<i>Coffea benghalensis</i> B.Heyne ex Schult.	Assão; Bangladesh; Leste do Himalaia; Índia; Miyanmar; Nepal; Tailândia; Vietnã
P. travancorensis	<i>Coffea travancorensis</i> Wight & Arn.	<i>Psilanthus travancorensis</i> (Wight & Arn.) J.-F.Leroy	Índia; Sri Lanka
P. ebracteolatus	<i>Coffea ebracteolata</i> (Hiern) Brenan	<i>Psilanthus ebracteolatus</i> Hiern	Benin; Burkina; Camaões; Gana; Guiné; Costa do Marfim; Liberia; Nigéria; Serra Leoa; Togo

Tabela 1 (continuação)

OTU	Espécie	sinônimo aceito	País
P. mannii	<i>Coffea mannii</i> (Hook.f.) A.P.Davis	<i>Psilanthus mannii</i> Hook.f.	Benin; Cabinda; Camarões; República Centro Africana; Congo; Gabão; Gana; Guiné; Golfo da Guiné; Togo; Costa do Marfim; Liberia; Nigéria; Serra Leoa; Zaire Quênia; Somália
C. rhamnifolia	<i>Coffea rhamnifolia</i> (Chiov.) Bridson	-	Quênia; Somália

Tabela 2: Melhor modelo e parâmetros evolutivos kappa (κ), delta (δ) e lambda (λ) calculados para a evolução de tamanho do genoma (TG), teor de cafeína (CAF), latitude mínima (L_mín) latitude máxima (L_máx), altitude máxima (Alt_Máx), temperatura mais alta nos dias mais quentes dos últimos 10 anos (T_Máx) e temperatura mais fria nos dias mais frios nos últimos 10 anos (T_Mín).

Caráter	Modelo	κ	δ	λ
TG	A	1,00	1,00	0,78
CAF	A	0,00	1,00	1,00
Lat_Máx	A	1,00	1,00	0,66
Lat_Mín	A	1,00	1,00	1,00
Alt_Máx	A	1,00	1,00	0,00
T_Máx	A	1,00	1,00	0,65
T_Mín	A	2,52	1,00	0,35

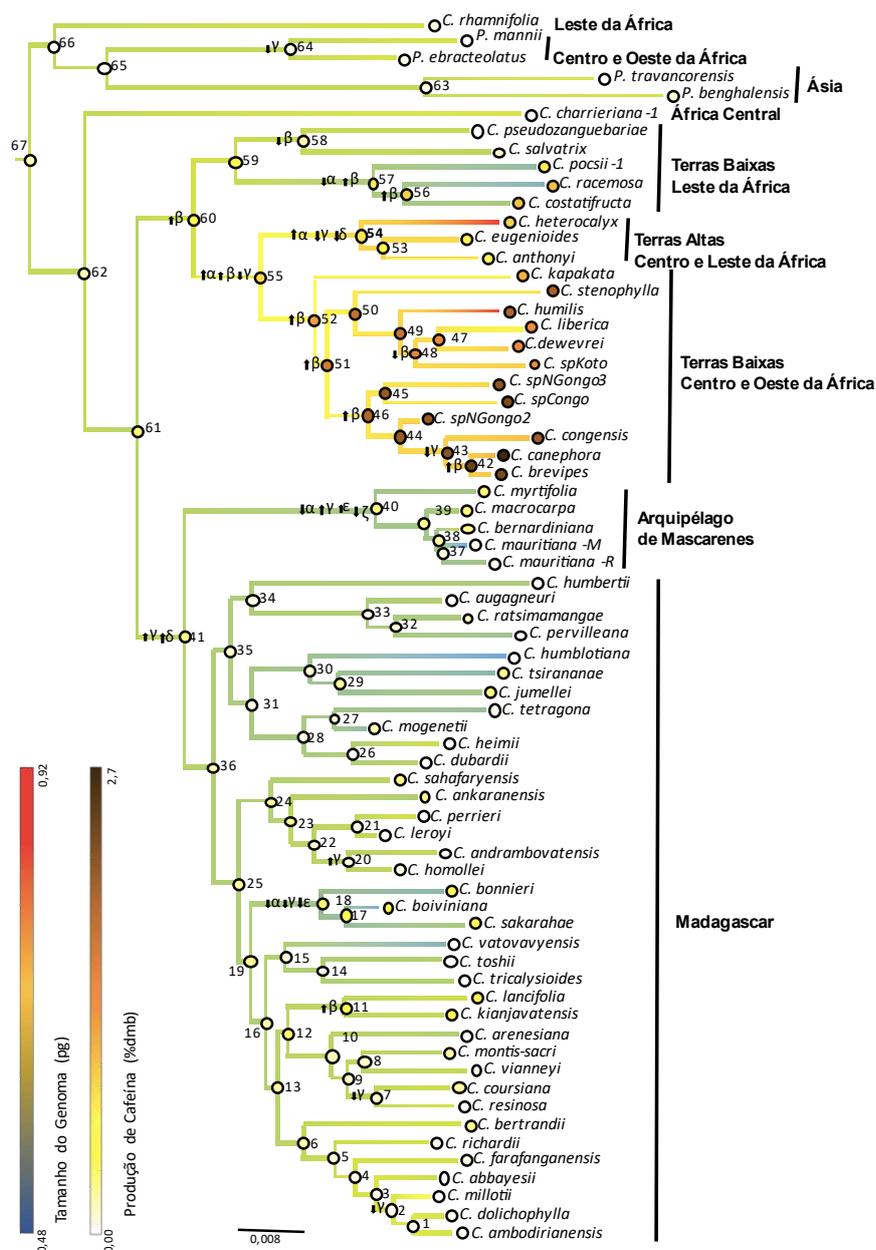


Figura 1. Inferência de valores ancestrais de tamanho de genoma e quantidade de cafeína na inferência filogenética obtida por HAMON (2017). Os ramos da árvore foram coloridos em função do tamanho do genoma inferido para cada espécie e cada nó ancestral (ver métodos). Os círculos em cada OTU e cada nó interno correspondem à quantidade inferida de cafeína. Setas para cima indicam aumento e setas para baixo reduções significativas do tamanho do genoma (α), produção de cafeína (β), latitude mínima (γ), latitude máxima (δ) e temperatura mínima (ϵ) e temperatura máxima (ζ) dos pontos em que as diferentes espécies ocorrem.

Tabela 3: Valores de tamanho dos genomas (TG) em picogramas (pg), teor de cafeína (CAF) em percentual de massa seca do fruto verde (%mdb), latitude mínima (L_mín) e máxima (L_máx) em graus decimais e temperaturas mínima (T_mín) e máxima (T_máx) em graus Celcius (°C) registradas nos últimos 10 anos para cada OTU em estudo. Os valores foram obtidos na literatura e bancos de dados. Constatam em negrito os valores das médias e desvios posteriores (desvio padrão, dp) inferidos neste estudo.

OTU	TG	dp	CAF	dp	L_mín	dp	L_máx	dp	T_mín	dp	T_máx	dp
<i>C.ambodirianensis</i>	0,65	0,06	0,03		16,06	2,87	23,09	4,59	7,24	3,00	38,31	2,36
<i>C.dolichophylla</i>	0,66	0,06	0,03		16,02	2,69	23,19		6,89	2,85	38,31	2,24
<i>C.millotii</i>	0,71		0,00		13,22		23,00		4,77		37,72	
<i>C.abbayesii</i>	0,63		0,00		21,22		24,84		6,95		41,56	
<i>C.farafanganensis</i>	0,66		0,05		21,22		23,14		6,95		36,74	
<i>C.richardii</i>	0,64		0,03		14,10		21,39		6,95		36,74	
<i>C.bertrandii</i>	0,61		0,26	0,17	21,22		25,15		6,95		36,74	
<i>C.resinosa</i>	0,64		0,00		13,20		21,38		2,85		36,74	
<i>C.coursiana</i>	0,63		0,21	0,15	13,26		23,18		4,77		38,15	
<i>C.vianneyi</i>	0,62		0,04		21,22		21,38		6,95		36,74	
<i>C.montis-sacri</i>	0,60		0,22	0,16	21,38		21,42		13,06		36,74	
<i>C.arenasiana</i>	0,56		0,00		18,88		21,38		6,76		36,74	
<i>C.kianjavatensis</i>	0,62		0,70		21,22		21,39		6,95		36,74	
<i>C.lancifolia</i>	0,63		0,70		17,93		21,38		6,95		36,74	
<i>C.tricalysioides</i>	0,59	0,07	0,00		12,57		24,51		5,19		37,72	
<i>C.toshii</i>	0,59	0,07	0,00		18,85		23,00		4,77		35,26	
<i>C.vatovavyensis</i>	0,52		0,01		21,38		21,40		13,06		36,74	
<i>C.sakarahaie</i>	0,58		0,41	0,25	15,72		24,78		3,34		42,65	
<i>C.boiviniana</i>	0,50		0,40	0,25	12,25		21,38		2,85		41,44	
<i>C.bonnierii</i>	0,53		0,36	0,23	12,52		21,38		6,95		36,74	
<i>C.homollei</i>	0,61		0,06		16,86		21,38		6,95		36,74	
<i>C.andrambovatensis</i>	0,61	0,06	0,00		21,38		21,53		7,44		36,74	
<i>C.leroyi</i>	0,62		0,02		14,30		23,25		6,95		40,26	
<i>C.perrieri</i>	0,64		0,00		14,28		25,02		5,06		42,38	
<i>C.ankaranensis</i>	0,59		0,27	0,19	12,45		21,38		6,95		37,72	
<i>C.sahafaryensis</i>	0,58		0,29	0,19	12,65		21,38		13,06		36,74	
<i>C.dubardii</i>	0,60		0,00		12,38		21,38		6,95		41,80	
<i>C.heimii</i>	0,63		0,00		12,37		21,38		2,85		37,72	
<i>C.mogenetii</i>	0,55		0,20	0,15	12,50		14,22		8,86		36,17	
<i>C.tetragona</i>	0,54		0,03		13,77		21,38		6,95		41,44	
<i>C.jumellei</i>	0,57		0,29	0,20	12,52		22,66		6,95		41,66	
<i>C.tsirananae</i>	0,53		0,28	0,19	12,17		21,38		6,95		37,72	
<i>C.humblotiana</i>	0,49		0,00		13,13		21,13	7,75	23,05		30,19	
<i>C.pervilleana</i>	0,56		0,00		12,38		21,38		6,95		41,68	
<i>C.ratsimamangae</i>	0,58		0,19	0,14	12,34		21,38		13,06		37,72	
<i>C.augagneuri</i>	0,59		0,00		12,48		21,38		6,95		36,74	

Tabela 3 (continuação)

OTU	TG	dp	CAF	dp	L_mín	dp	L_máx	dp	T_mín	dp	T_máx	dp
C.humbertii	0,58	0,08	0,00		20,17		24,88		6,52		42,79	
C.mauritiana-M	0,48		0,04		20,33		21,13		13,36		32,73	
C.mauritiana-R	0,56	0,05	0,04		19,62	2,63	20,82	4,02	12,52	5,40	32,12	2,02
C.bernardiniana	0,62		0,26	0,18	19,53	2,23	20,75	3,20	12,48	5,14	32,02	1,68
C.macrocarpa	0,59		0,30	0,19	19,20	2,60	20,62	4,10	19,50		30,63	
C.myrtifolia	0,54		0,34	0,22	17,90	4,08	20,38	6,11	19,50		30,63	
C.brevipes	0,76		2,30		0,57		5,53		11,74		37,43	
C.canephora	0,78		2,70		0,05		23,00		-0,76		43,78	
C.congensis	0,77		2,06		0,05		4,45		13,56		41,11	
C.spNGongo2	0,74	0,05	2,12		3,50	2,05	10,79	3,70	8,34	3,21	40,50	2,06
C.spCongo	0,74	0,07	2,27		5,38	3,22	11,27	5,59	8,13	4,45	40,35	3,39
C.spNGongo3	0,73	0,07	2,18		5,38	3,23	11,02	5,64	8,16	4,13	40,55	3,26
C.dewevrei	0,70		1,47	0,24	4,29	2,67	9,29		8,42	3,86	39,84	2,77
C.liberica	0,70		1,42		1,08		9,33		7,25		39,76	
C.spKoto	0,76	0,07	1,31		5,00	3,10	10,00	5,35	8,37	4,31	40,03	3,19
C.humilis	0,92		1,98		4,50		7,66		11,07		39,40	
C.stenophylla	0,64		2,00		5,33		7,75		10,55		41,06	
C.kapakata	0,66		1,12		9,98		12,30		4,75		41,54	
C.eugenioides	0,74		0,54		0,18		5,93		6,44		42,91	
C.anthonyi	0,72		0,60		1,73		4,70		11,12		39,76	
C.heterocalyx	0,91		0,88		3,88		5,25		9,87		35,84	
C.costatifructa	0,58		1,00		7,80		9,04		13,12		40,53	
C.racemosa	0,52		1,00		17,16		28,13		5,37		48,02	
C.pocsii-1	0,54		0,72	0,25	6,14		6,88		12,69		38,61	
C.pseudozanguebariae	0,62		0,02		3,12		7,07		11,86		39,62	
C.salvatrix	0,61		0,11		16,49		19,97		4,95		46,11	
C.charrieriana-1	0,62	0,10	0,00		10,44	6,00	16,02	9,41	8,99	11,75	40,57	6,71
P.benghalensis	0,62	0,13	0,09		9,74	5,98	15,60	9,08	8,44	10,51	40,50	6,20
P.travancorensis	0,62	0,12	0,00		9,49	5,76			8,37	9,71		
P.ebracteolatus	0,62	0,11	0,04		5,11	3,41	11,63	6,21	8,57	5,57	40,60	3,88
P.mannii	0,62	0,10	0,00		0,07		8,52		10,01		41,06	
C.rhamnifolia	0,61	0,10	0,06		0,43		3,73		15,77		43,13	

Tabela 4: Valores posteriores estimados, com seus respectivos desvios padrão (dp) para os nós ancestrais da árvore representada na Figura 1 do tamanho dos genomas (TG) em picogramas (pg), teor de cafeína (CAF) em percentual de massa seca do fruto verde (%mdb), latitude mínima (L_mín) e máxima (L_máx) em graus decimais e temperaturas mínima (T_mín) e máxima (T_máx) em graus Celcius (°C).

Nó	TG	dp	CAF	dp	L_mín	dp	L_máx	dp	T_mín	dp	T_Máx	dp
1	0,66	0,37	0,03	0,19	16,08	0,97	22,89	3,43	7,19	2,66	38,14	1,88
2	0,65	0,04	0,02	0,20	16,08	1,03	22,80	3,11	7,15	2,53	38,27	3,42
3	0,65	0,04	0,01	0,20	17,34	1,15	22,83	3,13	7,16	2,37	38,27	1,73
4	0,64	0,04	0,05	0,20	17,64	1,30	22,69	2,98	7,35	2,37	38,09	1,83
5	0,63	0,04	0,09	0,20	17,21	1,37	22,32	2,84	7,20	2,37	37,87	1,73
6	0,62	0,03	0,20	0,20	17,38	1,40	22,31	2,67	7,42	2,29	37,69	1,61
7	0,62	0,04	0,12	0,20	15,75	1,39	21,75	3,14	6,98	2,60	37,40	1,96
8	0,61	0,04	0,13	0,20	18,53	1,31	21,65	3,14	7,33	2,47	37,19	1,86
9	0,61	0,03	0,14	0,20	17,61	1,31	21,94	2,92	7,24	2,50	37,42	1,91
10	0,60	0,03	0,15	0,21	17,63	1,47	21,94	2,85	7,39	2,39	37,47	1,72
11	0,61	0,04	0,58	0,20	18,49	1,70	21,63	3,00	7,12	2,11	37,28	1,89
12	0,60	0,03	0,33	0,21	17,30	1,30	21,77	2,68	7,41	1,69	37,66	1,56
13	0,60	0,03	0,24	0,20	16,97	1,19	21,79	2,50	7,48	1,62	37,64	1,52
14	0,59	0,04	0,02	0,19	16,47	1,70	22,07	3,02	6,75	2,06	37,30	1,81
15	0,59	0,03	0,07	0,20	16,83	1,43	21,93	2,61	7,80	1,89	37,60	1,63
16	0,60	0,03	0,19	0,21	16,38	1,17	21,53	2,52	7,65	1,62	37,79	1,56
17	0,55	0,04	0,38	0,20	13,35	1,26	21,93	3,04	4,18	1,69	39,68	1,88
18	0,56	0,04	0,34	0,20	13,72	1,48	21,63	3,13	5,36	1,98	39,04	1,77
19	0,59	0,03	0,25	0,21	15,36	1,17	21,11	2,45	7,69	1,65	38,00	1,42
20	0,61	0,04	0,04	0,19	16,96	1,30	21,80	3,06	7,16	1,69	37,90	1,79
21	0,62	0,04	0,05	0,19	14,55	0,98	22,56	3,13	6,68	1,29	39,40	2,00
22	0,61	0,03	0,09	0,20	15,37	1,30	21,74	2,96	7,40	1,66	38,50	1,71
23	0,60	0,03	0,21	0,21	14,78	1,35	21,68	2,89	7,65	2,33	38,26	1,66
24	0,59	0,03	0,24	0,20	14,68	1,40	21,33	2,60	7,56	2,14	38,11	1,64
25	0,59	0,03	0,24	0,21	14,90	1,22	20,94	2,48	7,65	2,00	38,11	1,46
26	0,59	0,04	0,02	0,19	12,69	1,47	20,01	3,21	7,45	2,59	38,86	1,93
27	0,57	0,04	0,09	0,20	12,80	1,31	19,06	3,22	7,74	2,63	38,21	1,78
28	0,58	0,03	0,06	0,21	12,99	1,45	19,48	2,90	7,86	2,38	38,37	1,75
29	0,56	0,04	0,23	0,19	12,93	1,90	20,94	3,22	8,41	2,69	37,93	1,90
30	0,56	0,04	0,10	0,20	13,12	1,86	20,66	3,05	8,35	2,65	37,60	1,88
31	0,58	0,03	0,10	0,21	13,50	1,52	20,23	2,70	8,01	2,18	38,23	1,54
32	0,58	0,04	0,07	0,20	12,60	1,45	20,92	3,36	8,51	3,22	38,59	2,07
33	0,58	0,04	0,05	0,20	12,78	1,54	20,86	3,39	8,41	3,03	38,35	1,92
34	0,59	0,03	0,06	0,20	14,05	1,60	20,43	2,61	7,93	2,25	38,50	1,58
35	0,59	0,03	0,12	0,21	13,78	1,39	20,02	2,49	7,89	2,09	38,27	1,45
36	0,59	0,03	0,22	0,21	13,92	1,28	19,96	2,37	7,99	2,07	38,23	1,38
37	0,56	0,04	0,08	0,20	19,66	0,81	20,25	3,40	12,70	4,47	32,48	1,91
38	0,56	0,04	0,21	0,20	19,51	0,81	20,37	3,26	12,71	4,26	32,51	1,95

Tabela 4 (continuação)

Nó	TG	dp	CAF	dp	L_mín	dp	L_máx	dp	T_mín	dp	T_Máx	dp
39	0,57	0,04	0,27	0,19	19,26	0,98	20,39	3,31	12,58	4,47	32,66	1,99
40	0,57	0,04	0,31	0,21	17,92	1,67	20,16	3,29	12,55	4,36	33,24	2,04
41	0,60	0,03	0,28	0,21	13,08	1,61	18,65	2,38	8,29	2,02	38,15	1,46
42	0,76	0,04	2,40	0,19	0,68	0,93	11,64	3,69	8,10	3,23	40,48	2,04
43	0,75	0,04	2,18	0,20	1,34	1,23	11,01	3,40	8,00	3,20	40,69	2,03
44	0,74	0,04	2,11	0,20	3,31	1,02	10,94	3,15	8,37	2,94	40,51	1,84
45	0,73	0,04	2,13	0,19	4,62	1,44	10,99	3,17	8,39	2,77	40,38	1,85
46	0,73	0,03	1,96	0,21	4,41	1,29	10,94	3,18	8,26	2,75	40,37	1,79
47	0,74	0,04	1,47	0,19	3,71	1,41	9,89	3,57	8,52	2,91	40,05	2,04
48	0,74	0,04	1,52	0,20	4,18	1,41	9,80	3,27	8,44	2,85	40,01	2,01
49	0,75	0,04	1,77	0,19	4,41	1,46	9,78	3,37	8,35	2,81	40,07	1,92
50	0,73	0,04	1,81	0,20	4,97	1,57	10,34	3,03	8,49	2,85	40,19	1,94
51	0,72	0,03	1,67	0,21	5,27	1,45	10,94	2,97	8,16	2,74	40,30	1,75
52	0,71	0,04	1,22	0,22	5,62	1,58	10,92	2,80	8,37	2,86	40,31	1,80
53	0,75	0,04	0,65	0,20	2,29	1,53	8,11	3,44	8,77	3,11	40,18	1,99
54	0,75	0,04	0,78	0,20	3,12	1,67	8,26	3,39	8,66	3,10	39,80	1,98
55	0,70	0,04	0,85	0,20	6,03	1,73	11,15	2,68	8,44	2,56	40,21	1,72
56	0,57	0,04	0,90	0,19	10,35	1,87	14,56	3,57	9,07	3,62	41,81	2,08
57	0,58	0,04	0,71	0,19	9,59	1,95	13,76	3,48	9,19	3,47	41,34	2,10
58	0,62	0,04	0,19	0,19	9,17	2,09	13,56	3,23	8,57	2,76	41,14	1,93
59	0,63	0,03	0,49	0,21	8,92	1,98	13,78	2,62	8,61	2,44	40,64	1,76
60	0,64	0,03	0,55	0,20	8,66	1,80	13,82	2,71	8,56	2,10	40,02	1,54
61	0,62	0,03	0,33	0,20	10,47	1,77	16,08	2,26	8,50	1,84	39,17	1,36
62	0,62	0,03	0,16	0,20	9,20	2,22	15,02	2,14	8,52	1,75	39,50	1,41
63	0,62	0,05	0,06	0,19	8,83	2,47	14,48	3,85	8,62	6,30	40,32	2,48
64	0,62	0,04	0,04	0,20	3,97	2,00	11,68	3,27	8,76	3,77	40,53	1,96
65	0,62	0,04	0,06	0,20	6,74	2,49	13,09	2,47	8,82	2,00	40,29	1,69
66	0,62	0,03	0,08	0,21	7,13	2,45	13,31	2,21	8,73	1,66	40,19	1,53
67	0,62	0,03	0,17	0,11	7,91	2,47	13,74	2,06	8,14	1,70	40,02	1,42

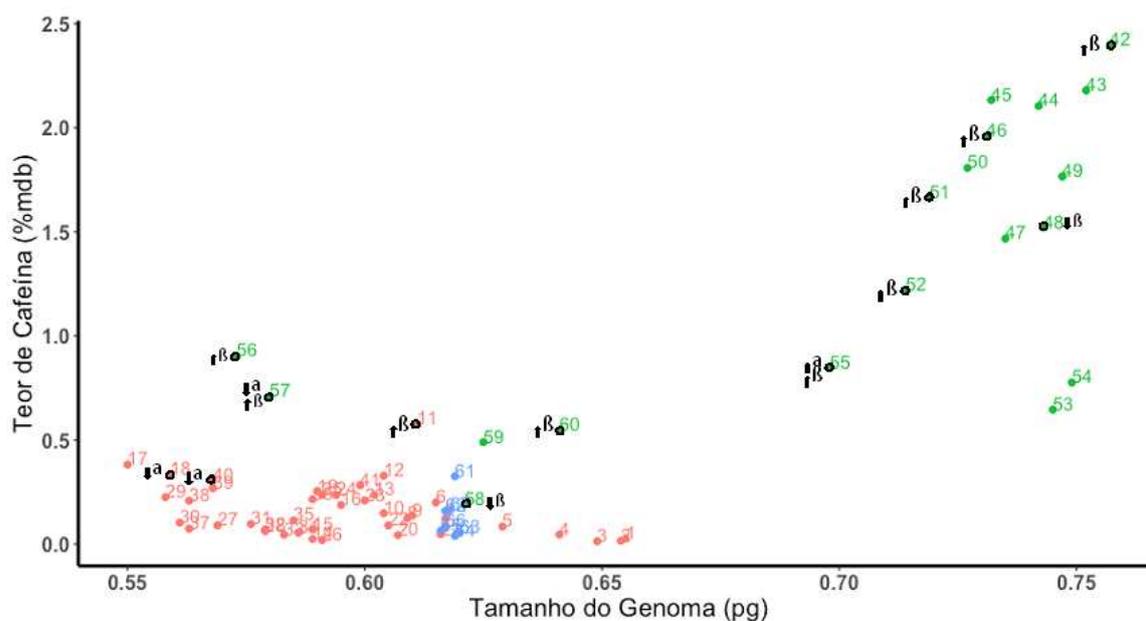


Figura 2. Regressão entre o tamanho do genoma (pg) e o teor de cafeína (%mdb) inferidos nos nós internos (1-67) da árvore filogenética mostrada na Figura 1. Pontos em vermelho representam os nós internos 1 a 41, todos descendentes do nó 41. Pontos em verde representam os nós internos 42 a 60, todos descendentes do nó 60. Pontos em azul representam os nós ancestrais ao nó 41 e o nó 60. As setas com as letras α e β mostram os pontos de aumento ou redução significativos do tamanho do genoma e do teor de cafeína. Os pontos de aumento e redução foram destacados com um círculo preto.

A Figura 2 mostra a regressão entre tamanho do genoma e teor de cafeína estimadas para os ramos ancestrais da Figura 1, que se correlacionaram positiva e significativamente ($R = 0,318$; $BF = 3,09$). As letras alpha (α) e beta (β) simbolizam os aumentos e reduções do tamanho do genoma e da cafeína, respectivamente.

A Figura 3 mostra a regressão entre o tamanho do genoma e as latitudes mínimas (Fig. 3A) e máximas (Fig. 3B) inferidas para os ramos ancestrais mostrados na Figura 1. Ambas as correlações foram negativas e significativas ($R = -0,236$; $BF = 2,88$, $TG \times L_{\text{mín}}$; $R = -0,182$; $BF = 2,74$, $TG \times L_{\text{máx}}$). Os nós mais basais da filogenia apresentam genomas pequenos e ocupam regiões de latitude intermediária, exceto pelo nó 64 que apresentou redução significativa da latitude mínima de ocorrência das espécies. Os nós ancestrais das espécies insulares (nó 41) também apresentam genomas pequenos (redução significativa nos nós 18 e 40) e habitam regiões de maiores latitudes (aumento da latitude mínima nos nós 40 e 41 com posterior redução nos nós 18, 7 e 2; e aumento da latitude máxima no nó 41), enquanto os nós descendentes do nó 55 apresentam genomas maiores (aumento no

nó 54, com redução no nó 57) e habitam regiões de menores latitudes (redução nos nós 55, 54 e 43).

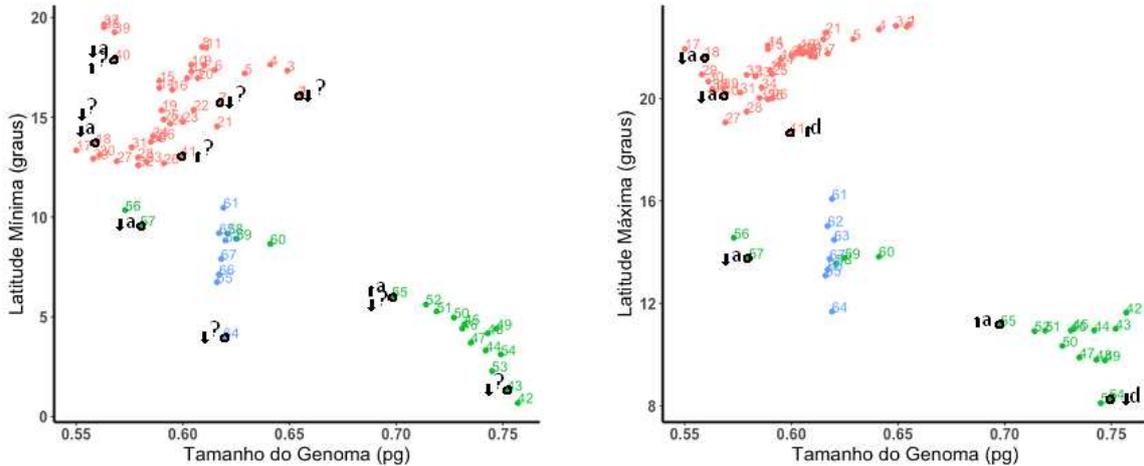


Figura 3. Regressão entre o tamanho do genoma (pg) e a latitude mínima (A) e máxima (B) inferidos nos nós internos (1-67) da árvore filogenética mostrada na Figura 1. Pontos em vermelho representam os nós internos 1 a 41, todos descendentes do nó 41. Pontos em verde representam os nós internos 42 a 60, todos descendentes do nó 60. Pontos em azul representam os nós ancestrais ao nó 41 e o nó 60. As setas com as letras α , γ e δ mostram os pontos de aumento ou redução significativos do tamanho do genoma, da latitude mínima e máxima, respectivamente. A latitude foi considerada em módulo (não importando se as espécies foram encontradas ao norte ou sul do Equador). Os pontos de aumento e redução foram destacados com um círculo preto.

A estimativa de parâmetros da altitude máxima registrada nos pontos onde foram encontradas as espécies de *Coffea* resultou em valores de κ e δ não significativamente diferentes de 1,0. No entanto, o parâmetro λ não foi diferente de zero, o que significa que a altitude onde as plantas se encontram hoje não apresentou sinal filogenético, pelo menos na hipótese filogenética utilizada como base, de modo que não é possível inferir os estados ancestrais ou calcular correlações com os demais caracteres.

A temperatura máxima enfrentada pelas espécies apresentou um valor de 40°C (dp = 1,4) no nó 67 e sofreu apenas uma redução significativa, no nó 40 (33,24°C, dp = 2,0). Já a temperatura mínima, que era de 8,13°C (dp=1,7) no nó 67, aumentou significativamente no nó 40 (12,5°C, dp = 4,4) e reduziu significativamente no nó 18 (5,4°C, dp = 2,0). Apesar de tanto temperaturas mínimas quanto máximas terem apresentado sinal filogenético (λ diferente de zero), a correlação com o tamanho do

genoma não foi significativa. Além disso, o $\kappa > 1,0$ das temperaturas mínimas sugere que o caráter apresentou mais modificações nos ramos mais longos, o que é particularmente notável entre os nós 40 e 41, quando a temperatura mínima suportada passou de 8,3 °C para 12,5 °C.

4. DISCUSSÃO

Os resultados do presente trabalho sugerem que o tamanho do genoma, ao contrário do que havia sido concluído por Razafinarivo et al. (2012) com base em uma filogenia contendo muitas politomias basais (Maurin et al., 2007), apresentou sinal filogenético, tendo como base a filogenia apresentada por Hamon et al. (2017), de modo que foi possível inferir o tamanho do genoma dos nós internos da árvore, bem como calcular a correlação deste caráter com a produção de cafeína, temperaturas extremas, altitudes e latitudes mínimas e máximas.

Hamon et al (2017), também utilizando métodos filogenéticos comparativos, mostraram que teores altos e consistentes de cafeína só foram encontrados em um clado que ocorre em ambientes úmidos e próximos ao Equador nas regiões central e ocidental da África. Estes autores observaram ainda que dentro da linhagem de Madagascar há produção moderada de cafeína em um pequeno grupo de espécies, o que os levou a sugerir que houve evolução convergente da síntese de cafeína no gênero em resposta a estímulos ambientais. Nossos resultados, no entanto, nos permitem argumentar que apesar de algumas espécies não produzirem cafeína (Tab. 3), o caráter foi herdado do ancestral do gênero (nó 67), que produzia uma pequena quantidade, aqui estimada em 0,165 %mdb (com desvio padrão de 0,1094), que não mudou significativamente ao longo da evolução do grupo, mas foi perdido em algumas espécies e apresentou dois incrementos significativos, o primeiro no nó 11, dentro da linhagem de Madagascar e o segundo no nó 60, nas regiões de latitude mais baixa, portanto próximas ao Equador. Nossa inferência de parâmetros evolutivos revelou que as alterações no teor de cafeína apresentam $\kappa = 0,0$, padrão de equilíbrios pontuados (Eldredge and Gould, 1972), sugerindo que os eventos de especiação dentro do gênero podem ter sido acompanhados, pelo menos em alguns nós ancestrais, por eventos de duplicação de cópias dos genes responsáveis pela biossíntese de cafeína. Encontramos ainda que o teor de cafeína apresentou correlação positiva e significativa com o tamanho do genoma e que genomas maiores que 0,7 pg (Fig. 2) só foram encontrados a partir do nó 55 (Fig. 1), no qual ocorreu um segundo aumento significativo do teor de cafeína, acompanhado de um aumento significativo do tamanho do genoma.

Estes achados sugerem que as múltiplas cópias de genes pertencentes às famílias gênicas envolvidas na produção de cafeína e síntese de compostos

secundários encontrados em *C. canephora* (Denoeud et al., 2014b) ocorreram ao longo da evolução do clado ao qual esta espécie pertence. Nossos resultados nos permitiram inferir que, começando no nó 60 (Fig. 1), as duplicações aumentaram significativamente o teor de cafeína (e possivelmente de outros compostos) nos nós 55, 52, 51, 46 e 42, que agrupa *C. canephora* e *C. brevipes*, as duas espécies com o maior teor de cafeína do gênero (Tab. 3). Cabe ressaltar que as famílias gênicas mencionadas por Denoeud *et al.* (2014) estão envolvidas no combate a inimigos naturais (Mithöfer and Boland, 2012), o que coincide com o fato de que as espécies com os maiores teores de cafeína ocorrem nos locais de menores latitudes (já ressaltado por Hamon et al. 2017), onde enfrentam uma maior diversidade de plantas competidoras (Pacheco Bustos et al., 2008; Gillman et al., 2015) e herbívoros (Salazar and Marquis, 2012).

O tamanho do genoma também se correlacionou significativamente com as latitudes mínimas e máximas inferidas nos nós ancestrais da Figura 1 tomando como base a distribuição das espécies atuais (ver métodos). As figuras 3A e 3B mostram padrões muito similares, com os nós descendentes do nó 55 (em verde, à direita do gráfico) nas posições de maior genoma e menor latitude; e os descendentes do nó 41 (em vermelho, à esquerda do gráfico) nas posições de menor genoma e maior latitude.

No nó 41, ancestral de todas as espécies insulares (Madagascar, Mascarenes e outras ilhas do oceano Índico) presentes na filogenia, houve aumento tanto da latitude mínima quanto máxima ocupada pelas espécies em relação às ancestrais. Estes aumentos se justificam pelo fato de que tais ilhas se encontram mais ao sul que a distribuição das espécies de café na África continental. Madagascar é reconhecida como um dos maiores *hotspots* de diversidade do Planeta, com uma enorme quantidade de espécies endêmicas, que ocorrem apenas na ilha; e micro endêmicas, que ocorrem em regiões limitadas dentro da ilha. A alta diversidade geral se justifica tanto pela antiguidade da ilha, que esteve conectada à Gondwana, quanto pelos sucessivos eventos de migração a partir massas continentais em volta (África, Índia, Austrália e Antártica) que ocorrem até a atualidade (Yoder and Nowak, 2006). A ilha apresenta ainda uma série de ambientes distintos que ocorrem lado a lado e que favorecem a radiação e o micro endemismo das espécies que chegam até ela, o que justifica a detecção de aumentos e reduções significativas nas latitudes máximas e mínimas (nós 40, 20, 18, 7 e 2).

A Figura 1 deixa claro que houve apenas uma introdução bem sucedida de *Coffea* em Madagascar, que irradiou numa grande quantidade de espécies, que também se dispersaram para ilhas vizinhas. Anthony et al., (2010), baseados em sequências plastidiais, sugeriram uma radiação rápida dessas espécies insulares que teriam deixado a África há menos de 500 mil anos, idade da ilha mais jovem da região, Grande Comores, que abriga a espécie *C. humblotiana*.

Dentro do grupo de espécies insulares chama a atenção um único aumento significativo no teor de cafeína do nó 12 (0,33 %mdb) para o nó 11 (0,576 % mdb), também registrado por Hamon et al. (2017), e que foi acompanhado, assim como ocorreu no nó 60, por um aumento não significativo do tamanho do genoma (de 0,604 pg no nó 12 para 0,611 pg, no nó 11), reforçando a hipótese de que o aumento no número de cópias responsáveis pela síntese de cafeína pode estar relacionado com o aumento no genoma.

De acordo com nossos resultados, a altitude máxima observada em cada espécie não apresentou sinal filogenético, impossibilitando a busca por correlações com o tamanho do genoma. As temperaturas máximas e mínimas registradas nos pontos que as espécies ocorreram, apesar de apresentarem sinal filogenético, não foram correlacionadas com o tamanho do genoma. Em termos de mudanças significativas, encontramos apenas que as espécies do arquipélago de Mascarenes (nó 40) apresentam a menor amplitude de temperaturas entre as espécies de café (menores máximas e maiores mínimas), muito provavelmente pela restrição territorial do arquipélago. No nó 18, detectamos também uma redução na temperatura mínima, acompanhada de uma redução no tamanho do genoma e da latitude mínima onde as espécies ocorrem.

Nosso estudo demonstrou que no gênero *Coffea*, que contém numerosas espécies com genomas relativamente parecidos, o aumento do conteúdo 1C contribuiu, muito provavelmente por ter permitido a expansão do teor de cafeína e possivelmente de outros genes associados à defesa, com a adaptação das espécies que descendem do nó 55 a ambientes mais próximos ao Equador na África Central e Ocidental. É possível afirmar, que pelo menos neste clado, as correlações entre tamanho de genoma com teor de cafeína e latitude sugerem um papel adaptativo do tamanho do genoma.

Segundo Oliver et al., (2007), ancestrais de genomas pequenos tendem a gerar descendentes de genoma também pequeno, enquanto ancestrais de genomas

grandes podem gerar descendentes de tamanho grande e pequeno. Esta teoria explica o fato de a grande maioria das espécies de *Coffea* apresentar tamanhos de genoma similares. De fato, as espécies insulares, que respondem por grande parte das espécies existentes, apresentam pouca variação, com tamanhos de genoma muito próximos aos inferidos para os nós mais basais da filogenia, e uma única redução significativa no nó 40, ancestral das espécies das ilhas Mascarenas. No entanto, no nó 60 parece ter havido um evento singular de aumento de genoma, que, apesar de não ter sido detectado como significativo sob nossos critérios (ver métodos), permitiu o aumento significativo do teor de cafeína, que presumivelmente depende do aumento de cópias gênicas responsáveis pela síntese da molécula. A partir do nó 60, e após um segundo aumento do tamanho do genoma, este sim significativo no nó 55, a produção de cafeína aumentou sucessivamente culminando com as espécies de maior teor de cafeína encontradas no gênero.

Em termos de contribuição para a discussão do paradoxo do valor-C (Thomas, 1971; Gregory, 2001) , resultados reforçam que a investigação do valor adaptativo dos tamanhos de genomas deve se ater a grupos filogeneticamente próximos, uma vez que a avaliação em categorias taxonômicas mais elevadas (famílias, ordens, classes, filos), o tamanho do genoma se comporta como um caráter neutro dada a enorme variação e a falta de correlação entre tamanho do genoma e complexidade. Fica claro também que, mesmo em grupos com tamanhos de genoma pequenos e pouca variação, podem ser detectados eventos de ampliação do genoma que podem estar intimamente relacionados à adaptação de algumas espécies a seus respectivos ambientes.

REFERÊNCIAS

- Achigan-Dako, E. G., Fuchs, J., Ahanchede, A., and Blattner, F. R. (2008). Flow cytometric analysis in *Lagenaria siceraria* (Cucurbitaceae) indicates correlation of genome size with usage types and growing elevation. *Plant Systematics and Evolution* 276, 9–19.
- Anthony, F., Diniz, L. E. C., Combes, M. C., and Lashermes, P. (2010). Adaptive radiation in *Coffea* subgenus *Coffea* L. (Rubiaceae) in Africa and Madagascar. *Plant Systematics and Evolution* 285, 51–64.
- Bauters, K., Asselman, P., Simpson, D. A., Muasya, A. M., Goetghebeur, P., and Larridon, I. (2016). Phylogenetics, ancestral state reconstruction, and a new infrageneric classification of *Scleria* (Cyperaceae) based on three DNA markers. *Taxon* 65, 444–466.
- Bennet, M. D. (1972). Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 181, 109–135.
- Costa-Neto, G., Galli, G., Carvalho, H. F., Crossa, J., and Fritsche-Neto, R. (2021). EnvRtype: A software to interplay enviromics and quantitative genomics in agriculture. *G3: Genes, Genomes, Genetics* 11.
- Davis, A. P., Chester, M., Maurin, O., and Fay, M. F. (2007). Searching for the relatives of *Coffea* (Rubiaceae, Ixoroideae): The circumscription and phylogeny of *Coffeae* based on plastid sequence data and morphology. *American Journal of Botany* 94, 313–329.
- Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., et al. (2014a). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345, 1181–1184.
- Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., et al. (2014b). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345, 1181–1184.
- Du, Y. P., Bi, Y., Zhang, M. F., Yang, F. P., Jia, G. X., and Zhang, X. H. (2017). Genome size diversity in *Lilium* (Liliaceae) is correlated with karyotype and environmental traits. *Frontiers in Plant Science* 8, 1303.
- Dupeyron, M., de Souza, R. F., Hamon, P., de Kochko, A., Cruzillat, D., Couturon, E., et al. (2017). Distribution of *Divo* in *Coffea* genomes, a poorly described family of

- angiosperm LTR-Retrotransposons. *Molecular Genetics and Genomics* 292, 741–754.
- Edwards James L. (2004). Research and Societal Benefits of the Global Biodiversity Information Facility. *BioScience* 54, 485–486.
- Fleischmann, A., Michael, T. P., Rivadavia, F., Sousa, A., Wang, W., Temsch, E. M., et al. (2014). Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. *Annals of Botany* 114, 1651–1663.
- Gardner, J. D., Laurin, M., and Organ, C. L. (2020). The relationship between genome size and metabolic rate in extant vertebrates. *Philosophical Transactions of the Royal Society B: Biological Sciences* 375, 20190146.
- Gillman, L. N., Wright, S. D., Cusens, J., McBride, P. D., Malhi, Y., and Whittaker, R. J. (2015). Latitude, productivity and species richness. *Global Ecology and Biogeography* 24, 107–117.
- Gould, S. J., and Eldredge, N. (1972). “Punctuated equilibria: an alternative to phyletic gradualism,” in *Models in Paleobiology*, 82–115.
- Gregory, T. R. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma, *Biological Review* 76, 65-101.
- Hamon, P., Duroy, P. O., Dubreuil-Tranchant, C., Mafra D’Almeida Costa, P., Duret, C., Razafinarivo, N. J., et al. (2011). Two novel Ty1-copia retrotransposons isolated from coffee trees can effectively reveal evolutionary relationships in the *Coffea* genus (Rubiaceae). *Molecular Genetics and Genomics* 285, 447–460.
- Hamon, P., Grover, C. E., Davis, A. P., Rakotomalala, J. J., Raharimalala, N. E., Albert, V. A., et al. (2017a). Genotyping-by-sequencing provides the first well-resolved phylogeny for coffee (*Coffea*) and insights into the evolution of caffeine content in its species: GBS coffee phylogeny and the evolution of caffeine content. *Molecular Phylogenetics and Evolution* 109, 351–361.
- Kang, M., Tao, J., Wang, J., Ren, C., Qi, Q., Xiang, Q. Y., et al. (2014). Adaptive and nonadaptive genome size evolution in Karst endemic flora of China. *New Phytologist* 202, 1371–1381.
- Khorami, S. S., Arzani, K., Karimzadeh, G., Shojaeiyan, A., and Ligterink, W. (2018). Genome size: A novel predictor of nut weight and nut size of walnut trees. *HortScience* 53, 275–282.

- Knight, C. A., and Ackerly, D. D. (2002). Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. *Ecology Letters* 5, 66–76.
- Lisle Richard. J. (2006). Google Earth a new geological resource. *Geology Today* 22, 29–32.
- Maurin, O., Davis, A. P., Chester, M., Mvungi, E. F., Jaufeerally-Fakim, Y., and Fay, M. F. (2007). Towards a phylogeny for *Coffea* (Rubiaceae): Identifying well-supported lineages based on nuclear and plastid DNA sequences. *Annals of Botany* 100, 1565–1583.
- Mithöfer, A., and Boland, W. (2012). Plant defense against herbivores: Chemical aspects. *Annual Review of Plant Biology* 63, 431–450.
- Noirot, M., Poncet, V., Barre, P., Hamon, P., Hamon, S., and de Kochko, A. (2003). Genome size variations in diploid African *Coffea* species. *Annals of Botany* 92, 709–714.
- Nowak, M. D., Davis, A. P., and Yoder, A. D. (2012). Sequence Data from New Plastid and Nuclear COSII Regions Resolves Early Diverging Lineages in *Coffea* (Rubiaceae). *Systematic Botany* 37, 995–1005.
- Oliver, M. J., Petrov, D., Ackerly, D., Falkowski, P., and Schofield, O. M. (2007). The mode and tempo of genome size evolution in eukaryotes. *Genome Research* 17, 594–601.
- Pacheco Bustos, A., Pohlan, H. A. J., and Schulz, M. (2008). Interaction between Coffee (*Coffea arabica* L.) and Intercropped Herbs under Field Conditions in the Sierra Norte of Puebla, Mexico. *Journal of Agriculture and Rural Development in the Tropics and Subtropics* 109, 85–93.
- Pagel, M. (1999a). Inferring the historical patterns of biological evolution. *Nature* 401, 877–884.
- Pagel, M. (1999b). The Maximum Likelihood Approach to Reconstructing Ancestral Character States of Discrete Characters on Phylogenies. *Systematic Biology* 48, 612–622.
- Pagel, M., Meade, A., and Barker, D. (2004). Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53, 673–684.
- Paradis, E., and Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528.
- Pellicer, J., Fay, M. F., and Leitch, I. J. (2010). The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society* 164, 10–15.

- Pellicer, J., and Leitch, I. J. (2019). The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytologist* 226, 301–305.
- Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., Fitzjohn, R. G., et al. (2014). Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30, 2216–2218.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67, 901–904.
- Razafinarivo, N. J., Rakotomalala, J. J., Brown, S. C., Bourge, M., Hamon, S., de Kochko, A., et al. (2012). Geographical gradients in the genome size variation of wild coffee trees (*Coffea*) native to Africa and Indian Ocean islands. *Tree Genetics and Genomes* 8, 1345–1358.
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3, 217–223.
- Rimlinger, A., Raharimalala, N., Letort, V., Rakotomalala, J. J., Crouzillat, D., Guyot, R., et al. (2020). Phenotypic diversity assessment within a major ex situ collection of wild endemic coffees in Madagascar. *Annals of Botany* 126, 849–863.
- Salazar, D., and Marquis, R. J. (2012). Herbivore pressure increases toward the equator. *Proceedings of the National Academy of Sciences of the United States of America* 109, 12616–12620.
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics* 27, 592–593.
- Shao, C., Li, Y., Luo, A., Wang, Z., Xi, Z., Liu, J., et al. (2021). Relationship between functional traits and genome size variation of angiosperms with different life forms. *Biodiversity Science* 29, 575–585.
- Souza, G., Costa, L., Guignard, M. S., Van-Lume, B., Pellicer, J., Gagnon, E., et al. (2019). Do tropical plants have smaller genomes? Correlation between genome size and climatic variables in the Caesalpinia Group (Caesalpinioideae, Leguminosae). *Perspectives in Plant Ecology, Evolution and Systematics* 38, 13–23.
- Sparks, A. (2018). nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *Journal of Open Source Software* 3, 1035.
- Thomas, C. A. (1971). The Genetic Organization of Chromosomes. *Annual Review of Genetics* 5, 237–256.

- Xu, Z., Xu, Z., Pu, X., Gao, R., Demurtas, O. C., Fleck, S. J., et al. (2020). Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biology* 18, 63.
- Yoder, A. D., and Nowak, M. D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Review of Ecology, Evolution, and Systematics* 37, 405–431.
- Yotoko, K. S. C., Dornelas, M. C., Togni, P. D., Fonseca, T. C., Salzano, F. M., Bonatto, S. L., et al. (2011). Does variation in genome sizes reflect adaptive or neutral processes? new clues from Passiflora. *PLoS ONE* 6. e18212.
- Yuyama, P. M., Pereira, L. F. P., dos Santos, T. B., Sera, T., Vilas-Boas, L. A., Lopes, F. R., et al. (2012). FISH using a *gag* -like fragment probe reveals a common *Ty* 3-*gypsy* -like retrotransposon in genome of *Coffea* species. *Genome* 55, 825–833.