

REVISTA DE ARQUEOLOGIA

Volume 39 No. 3 Setembro - Dezembro 2026

ARTIGO

A TRANSMISSÃO CULTURAL DE GRUPOS ASSOCIADOS À TRADIÇÃO TUPIGUARANI EXPLICADA ATRAVÉS DA CLADÍSTICA¹

Glauco Constantino Perez*, Astolfo Gomes de Melo Araujo**, Mercedes Okumura***, Ethan E. Cochrane****

RESUMO

Métodos provindos da Biologia Evolutiva vêm sendo aplicados à Arqueologia para compreender mudanças culturais. Este estudo utilizou a cladística para analisar 523 atributos cerâmicos de catorze sítios Tupiguarani em São Paulo, investigando padrões de transmissão e dispersão cultural. A análise de máxima parcimônia, associada ao Non-metric Multidimensional Scaling (NMDS), ou Ordenação Multivariada Não-Métrica, revelou agrupamentos regionais, distinguindo litoral e interior, além de diferenças entre as bacias do Tietê e do Paranapanema. Os índices de Consistência (0,54) e Retenção (0,38) indicam herança filogenética, enquanto comparações entre sítios datados e não datados sugeriram cronologias inéditas. Os resultados apontam diferenciações internas entre os sítios arqueológicos a partir de um ancestral comum.

Palavras-chave: Arqueologia evolutiva; Teoria da transmissão cultural; Tradição tupiguarani; Estatística multivariada; São Paulo; Sudeste do Brasil.

¹ Todos os autores foram presentes de maneira igual na concepção da pesquisa e do artigo, aquisição, análise e interpretação dos dados, construção das versões preliminares, auxílio na redação das diversas versões do artigo, revisão crítica do conteúdo e responsabilidade pela integridade do trabalho.

* Professor Doutor do Departamento de História da Universidade Estadual de Maringá, Paraná.

E-mail: glauco1113@gmail.com - autor correspondente. ORCID: <https://orcid.org/0000-0002-1472-3360>.

** Professor Doutor Livre Docente do Museu de Arqueologia e Etnologia da Universidade de São Paulo, coordenador do Laboratório Interdisciplinar de Pesquisas em Evolução, Cultura e Meio Ambiente (LEVOC - MAE-USP). E-mail: astwolfo@usp.br. ORCID: <https://orcid.org/0000-0002-0349-1226>.

*** Professora Doutora do Departamento de Genética e Biologia Evolutiva e coordenadora do Laboratório de Estudos Evolutivos Humanos (LEEH - USP). E-mail: okumuram@usp.br. ORCID: <https://orcid.org/0000-0002-1894-6430>.

**** Professor Doutor da Faculdade de Artes, Departamento de Antropologia, Universidade de Auckland, Nova Zelândia. E-mail: e.cochrane@auckland.ac.nz. ORCID: <https://orcid.org/0000-0001-9027-0634>.

THE CULTURAL TRANSMISSION OF GROUPS ASSOCIATED WITH THE TUPIGUARANI TRADITION EXPLAINED THROUGH CLADISTICS

ABSTRACT

Methods derived from Evolutionary Biology have been applied to Archaeology to understand cultural changes. This study used cladistics to analyze 523 ceramic attributes from 14 Tupiguarani sites in São Paulo, investigating patterns of cultural transmission and dispersal. The maximum parsimony analysis, combined with Non-metric Multidimensional Scaling (NMDS), revealed regional clusters, distinguishing coastal and inland areas, as well as differences between the Tietê and Paranapanema river basins. The Consistency (0.54) and Retention (0.38) indices indicate phylogenetic inheritance, while comparisons between dated and undated sites suggested new relative chronologies. The results point to differentiation among the archaeological sites from a common ancestor.

Keywords: Evolutionary archaeology; Cultural transmission theory; Tupiguarani tradition; Multivariate statistics; São Paulo; Southeastern Brazil.

LA TRANSMISIÓN CULTURAL DE LOS GRUPOS ASOCIADOS A LA TRADICIÓN TUPIÍ-GUARANÍ EXPLICADA POR LA CLADÍSTICA

RESUMEN

Los métodos provenientes de la Biología Evolutiva han sido aplicados a la Arqueología para comprender los cambios culturales. Este estudio utilizó la cladística para analizar 523 atributos cerámicos de 14 sitios tupí-guaraníes en São Paulo para evaluar los patrones de transmisión y dispersión cultural. El análisis de máxima parsimonia, asociado al Escalamiento multidimensional no métrico (*Non-metric Multidimensional Scaling*, NMDS), reveló agrupamientos regionales al distinguir litoral e interior, además de diferencias entre las cuencas de los ríos Tietê y Paranapanema. Los índices de consistencia (0,54) y retención (0,38) revelan herencia filogenética, mientras que las comparaciones entre sitios datados y sitios no datados sugirieron cronologías relativas inéditas. Los resultados señalan diferenciaciones internas entre los sitios arqueológicos a partir de un ancestro común.

Palabras clave: Arqueología evolutiva; Teoría de la transmisión cultural; Tradición tupí-guaraní; Estadística multivariada; São Paulo; Sureste de Brasil.

INTRODUÇÃO

Desde o final do século XIX, a Arqueologia tem se dedicado à análise das transformações na cultura material em suas dimensões temporal e espacial. Tal problemática tem constituído o cerne das reflexões teóricas que fundamentaram o estabelecimento da disciplina ao longo do século XX (Araujo, 2019). Nesse contexto, a abordagem filogenética surge como um método para explicar os processos de mudança biológica e cultural, possibilitando a reconstrução das relações de descendência e transmissão inerentes ao ser humano.

Originariamente concebida no âmbito da Biologia Evolutiva, a Filogenia compreende o estudo sistemático das relações evolutivas entre organismos, baseado na análise comparativa de dados morfológicos, comportamentais e moleculares, com vistas à formulação de hipóteses sobre sua história evolutiva. Essa lógica analítica é passível de transposição para a investigação das dinâmicas culturais, permitindo inferir padrões de transmissão e modificação em linhagens artefatuais mediante a aplicação de princípios filogenéticos análogos aos empregados nos estudos biológicos (O'Brien; Lyman, 2003).

Embora os métodos filogenéticos tenham sido originalmente desenvolvidos no âmbito da Biologia Evolutiva (O'Brien; Lyman, 2003), sua aplicação em contextos arqueológicos não implica uma transposição direta dos mecanismos da herança genética para a cultura material. Neste trabalho, a cladística é empregada como uma ferramenta analítica formal para explorar padrões estruturados de similaridade, persistência e diferenciação entre conjuntos artefatuais, e não como um modelo estrito de descendência genealógica. Reconhece-se, portanto, que a transmissão cultural envolve processos intencionais, contingentes e historicamente situados, distintos daqueles observados na transmissão biológica, ainda que ambos compartilhem propriedades formais passíveis de modelagem comparativa (O'Brien; Lyman, 2003).

Assim, o uso da cladística neste estudo tem caráter heurístico e exploratório, visando identificar padrões de herança, recorrência e homoplasia (aparecimento de caracteres semelhantes sem relação de ancestralidade) em atributos cerâmicos associados à Tradição Tupiguarani, sem pressupor que tais padrões correspondam a linhagens culturais fechadas ou a trajetórias evolutivas unilineares. A análise privilegia a detecção de estruturas de similaridade que possam ser posteriormente interpretadas à luz de dados espaciais, cronológicos, históricos e linguísticos.

Para compreender os processos de seleção natural e deriva cultural descritos por Dunnell (1978), o principal método adotado naquele programa de investigação foi a seriação, a qual foi revisitada por Lipo (2001, 2005), que propôs uma nova interpretação dos resultados gráficos obtidos por meio dessa técnica, diferenciando-se um pouco dos modelos anteriores de investigação a partir desse método (Chmyz, 1973; Chmyz *et al.*, 1976; Meggers; Evans, 1970). Nessa perspectiva, a Arqueologia tem incorporado outros métodos provenientes da Biologia, como a cladística, testes de similaridade, agrupamentos e análises de acúmulo de espécies, que oferecem respaldo para novas interpretações dos conjuntos artefatuais na perspectiva da Arqueologia Evolutiva (Cochrane, 2004). Neste artigo, exploraremos o potencial da cladística e dos testes de agrupamento para a análise dos sítios associados à Tradição Tupiguarani no estado de São Paulo, aprofundando o conhecimento sobre a variabilidade cultural dos sítios em questão.

A variabilidade na cultura material tem sido objeto de amplo debate entre arqueólogos e a perspectiva teórica da Arqueologia Evolutiva, que permite explicá-la a partir de três situações possíveis: 1) a adaptação paralela entre indivíduos que convergem sob pressões seletivas semelhantes (homoplasia); 2) a ancestralidade comum entre indivíduos

que partilham de herança cultural com as gerações futuras (homologia) e os processos de ramificação cultural que são gerados a partir disso; e 3) a transmissão genética ou cultural horizontal por meio de interação intergrupar (*blending*) (Crema; Kerig; Shennan, 2014; O'Brien *et al.*, 2012). O'Brien *et al.* (2012) entendem que tem crescido o uso dessas abordagens para explicar a mudança da cultura material, uma vez que a comunidade científica reconhece que a transmissão de traços culturais pode ser estudada de modo análogo à transmissão genética, observando os sistemas envolvidos nesses processos. Neste artigo, consideramos que a mudança cultural decorre tanto do surgimento por inovação quanto do contato interétnico (Dunnell, 1978), e que semelhanças e diferenças culturais podem ser explicadas pelos processos de ramificação e mistura ocorridos ao longo do tempo entre comunidades distintas (Crema; Kerig; Shennan, 2014).

As críticas quanto à aplicação do método filogenético à história cultural fundamentam-se no conhecido e pouco adequado argumento de que “*objects do not breed*” (Cochrane, 2004; O'Brien; Lyman, 2003; O'Brien *et al.*, 2012). De fato, objetos são inanimados e, portanto, não se reproduzem biologicamente. Contudo, eles são produzidos e replicados com diferentes graus de fidelidade ao longo das gerações, pois “*toolmakers and the like are cultural descendents of other toolmakers*” (O'Brien *et al.*, 2012, p. 25). Nesse sentido, apesar de escolhas particulares estarem envolvidas, os indivíduos criam seus próprios objetos e ensinam outros a reproduzi-los por meio de mecanismos de transmissão culturalmente circunscritos, como imitação, emulação, aprendizado etc., o que assegura a perpetuação tanto dos objetos quanto dos modos de produção cultural ao longo do tempo. Dessa forma, diferentemente da transmissão genética, a transmissão cultural ocorre por meio de processos sociais como ensino-aprendizagem, imitação, emulação e participação ritual, os quais são, em maior ou menor grau, intencionais e sujeitos a escolhas individuais e coletivas (Shennan; Bentley, 2008). Além disso, tais mecanismos não permanecem fixos no tempo, podendo ser modificados em função de contextos ambientais, sociais e históricos específicos (Cochrane, 2004). Dessa forma, a aplicação de métodos filogenéticos à cultura material exige cautela interpretativa e explicitação de seus pressupostos e de suas limitações. Para tanto, a cladística não é utilizada aqui como um modelo explicativo totalizante da mudança cultural, mas como um procedimento analítico capaz de organizar grandes volumes de dados morfológicos, decorativos e tecnológicos, permitindo a identificação de padrões de continuidade, convergência e recombinação de traços culturais. A interpretação desses padrões depende, necessariamente, de sua articulação com informações externas à matriz de dados, como cronologia, espacialidade e contexto histórico.

Assim, a observação dos processos de ensino e aprendizagem são fundamentais para entender a formação das tradições culturais. Willey (1945) definiu tradição como a persistência de tecnologias ao longo das gerações. Mais recentemente, em Teoria da Arqueologia, tradições podem ser compreendidas como linhagens tecnológicas relacionadas, refletindo a transmissão e a persistência de práticas por meio da replicação e continuidade hereditária (O'Brien *et al.*, 2012).

Técnicas cladísticas geram hipóteses de parentesco relativo de táxon (indivíduo), e na Arqueologia estas podem ser aplicadas a conjuntos de tipos de artefatos. Além disso, esses métodos têm sido empregados nos mais diversos estudos, por exemplo a partir de parentesco linguístico (Atkinson; Gray, 2005), análise de manuscritos (Spencer *et al.*, 2003), crenças religiosas (Basava; Zhang; Mace, 2021), contos de fadas (Tehrani, 2013), digestão de lactose em adultos (Holden; Mace, 1997) e parentesco bantu da evolução social humana (Opie *et al.*, 2014). No Brasil, técnicas cladísticas foram utilizadas em estudos de registros rupestres, destacando a comparação de semelhanças gráficas

entre sítios arqueológicos das regiões Norte e Nordeste do país (Matos, 2019; Valle, 2003, 2006, 2012).

A TRADIÇÃO TUPIGUARANI

A Tradição Tupiguarani foi desenhada nos moldes do Programa Nacional de Pesquisas Arqueológicas (Pronapa), que tinha por objetivo organizar sítios arqueológicos seguindo determinado método que os agrupava pelas suas semelhanças e diferenças dentro de uma definição peculiar de “Tradição Cultural” (Araujo, 2015; Araujo; Okumura, 2021). Esses moldes de definição foram adaptações apresentadas por Willey e Phillips (1958), que estavam preocupados em fornecer referenciais de organização dos sítios a partir de parâmetros de temporalidade, formas e espacialidade.

Em geral, o que entendemos hoje por Tradição Tupiguarani é o resultado de um agrupamento de sítios que apresentam características semelhantes, em uma determinada temporalidade, ao longo de toda a espacialidade da América do Sul (Figura 1). Corrêa (2014) delinea a historiografia e a construção do termo associado à Tradição Tupiguarani compilando dados etnográficos, linguísticos, etno-históricos e arqueológicos. A partir desses parâmetros, temos o jargão arqueológico de Tupiguarani definido por Meggers (1967), com definições feitas com base em tipologias percebidas a partir do registro arqueológico. Assim, naquele período, ficou estabelecido que esse termo estaria associado aos sítios cerâmicos que apresentassem pintura policroma (vermelha ou preta sobre engobo branco ou vermelho), fora da região delimitada pela Bacia Amazônica, com técnicas de acabamento plástico, predominando sobre superfícies alisadas, corrugadas, unguladas ou escovadas, e engobo. Somadas a essas características, reconhece-se a presença de enterramentos secundários, com urnas, machados de pedra polida, tembetás, lascas, talhadores e abrasadores (Brochado, 1981; Brochado *et al.*, 1969; Chmyz *et al.*, 1976; Corrêa, 2014).

As origens e as rotas de expansão desses grupos têm sido debatidas desde o início do século XX, inicialmente sugerindo uma origem amazônica com subseqüentes migrações em duas ondas principais (Brochado, 1984; Meggers, 1979; Meggers *et al.*, 1988; Mello; Kneip, 2005; Métraux, 1927). De acordo com essa visão, a Amazônia Central teria sido o berço dessas culturas, e a partir daí emergiram duas rotas de expansão populacional: uma em direção à foz do Amazonas, estendendo-se para o leste e ocupando toda a costa atlântica do Brasil até aproximadamente a divisa entre os estados de São Paulo e Paraná; e outra rota seguindo para o sul, passando pelo Uruguai, pela bacia do Rio Paraná e pela bacia do Rio da Prata, chegando a povoar o sul do Brasil, o Uruguai e o norte da Argentina (Corrêa, 2014).

Estudos recentes que comparam evidências arqueológicas com a linguística e história apresentaram modelos específicos às teorias pioneiras de Brochado (Bonomo *et al.*, 2025; Gerardi *et al.*, 2023; Noelli, 1996, 2008; Souza; Noelli; Madella, 2021).

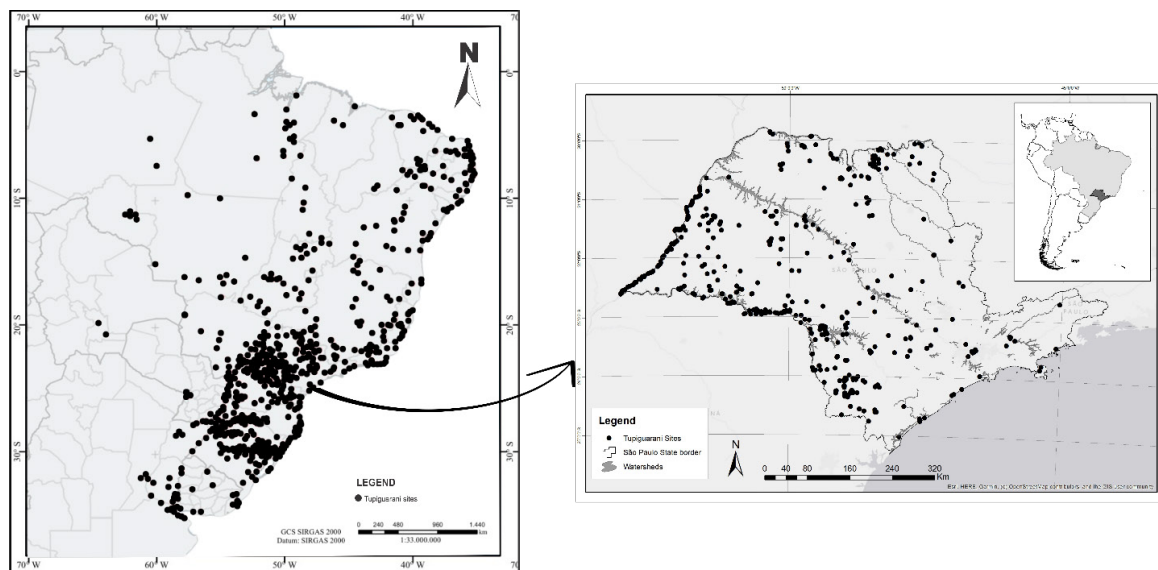
Há cerca de dois mil anos, grupos humanos indígenas produziram numerosos sítios arqueológicos nas regiões Nordeste, Sudeste e Sul do Brasil, exibindo cerâmica e outros artefatos associados à Tradição Tupiguarani, o que sugere que essas comunidades introduziram práticas agroflorestais que incluíam o cultivo de plantas domesticadas, o estabelecimento de assentamentos mais permanentes e a adoção de técnicas de modificação da paisagem, como a criação de terras pretas antropogênicas (Maezumi *et al.*, 2018; Souza; Mateos; Madella, 2020). As vasilhas cerâmicas serviam não apenas para cozinhar e armazenar alimentos, mas também como urnas funerárias, evidenciando usos secundários desses recipientes (Prous, 1992). Dado o cultivo de mais de uma centena de espécies vegetais, incluindo mandioca e milho, esses grupos provavelmente viviam

em áreas densamente povoadas, uma conclusão apoiada pelo tamanho de seus sítios arqueológicos (Albuquerque, 2008; Noelli, 2025).

A maioria das pesquisas tem se concentrado nas origens e nas primeiras dispersões da Tradição Tupiguarani, com menos estudos voltados para a diversidade regional desses grupos. Nesse sentido, este artigo busca compreender melhor as relações culturais entre sítios arqueológicos associados à Tradição Tupiguarani no atual território do estado de São Paulo, incluindo possíveis padrões de transmissão cultural entre esses grupos humanos. Essa abordagem regional é justificada pelo fato de que o estado de São Paulo, assim como o Paraná, foi tradicionalmente proposto como ponto de encontro das duas ondas de dispersão do modelo de Brochado (1984), configurando uma área de fronteira para essas populações (Perez, 2018; Perez; Araujo; Okumura, 2020).

Dado o potencial de cerâmicas arqueológicas para inferir padrões de transmissão cultural no passado e a área do atual estado de São Paulo, delimitada por grandes rios como o Paranapanema, o Rio Grande e o Rio Paraná, essa região pode ser considerada ideal para aplicação de métodos cladísticos, com o objetivo de entender melhor a história evolutiva e as relações de ancestralidade-descendência, assim como identificar as características cerâmicas que são mais significativas para estabelecer os padrões filogenéticos.

Figura 1. Pontos pretos indicam a distribuição espacial geral dos sítios arqueológicos associados à Tradição Tupiguarani na América do Sul e no estado de São Paulo.



Fonte: Adaptado de Corrêa (2014). Elaborado por Thaianne Lima e os autores (2025).

AMPLIANDO A DISCUSSÃO - FILOGENIA E TRANSMISSÃO CULTURAL

Na Biologia Evolutiva, a filogenia, ou filogênese, refere-se ao estudo das relações evolutivas entre organismos e seus sistemas de transmissão de informações, geralmente por meio do sequenciamento de dados moleculares, o que permite traçar a história evolutiva desses grupos. Esse mesmo princípio pode ser aplicado aos estudos de transmissão cultural, permitindo que as relações culturais sejam investigadas com base nos mecanismos filogenéticos, de forma análoga ao que ocorre nos estudos genéticos (O'Brien; Lyman, 2003). Mace e Holden (2005) defendem o uso de métodos filogenéticos para estudar a diversidade cultural humana, traçando paralelos com a evolução biológica. Os autores testam como os métodos comparativos filogenéticos explicam a

interdependência cultural e as hipóteses sobre história cultural, diversificação e adaptação ao analisar características culturais em uma estrutura de árvore, semelhante às árvores genealógicas ou linguísticas.

Segundo O'Brien *et al.* (2012), a transmissão cultural pode ocorrer de forma vertical, de pais para filhos, análoga à transmissão genética, mas também pode ocorrer no sentido inverso, dos filhos para os pais. Além disso, a transmissão pode ser horizontal, entre indivíduos da mesma geração, ou oblíqua, entre pessoas não relacionadas e de diferentes gerações.

Embora as linhagens culturais apresentem persistência e continuidade hereditária, o traço filogenético nelas pode ser difícil de rastrear, principalmente devido à sua natureza mais abstrata. Sob essa perspectiva, a evolução cultural apresenta ritmo diferente da evolução biológica; mesmo que ambos os casos envolvam a transmissão de traços filogenéticos, esse processo pode levar à redução da diversidade cultural e ao surgimento de formas híbridas. Em outras palavras, ambos os processos evolutivos incluem reticulação e misturas durante seu desenvolvimento (O'Brien *et al.*, 2012, Tëmkin; Eldredge, 2007).

Historicamente, O'Brien e Lyman (2003) informam que a construção de cladogramas, ou o uso da cladística, se desenvolveu na década de 1940, com o entomologista alemão Willi Henning, sendo este um dos principais métodos da reconstrução filogenética usado na biologia, e desde então são usadas convenções para se referir aos braços formados nas ramificações desse tipo de representação. A cladística é um corpo de princípios e técnicas para classificar formas vivas (táxon) assumidas como relacionadas através da descendência comum. Seu produto consiste em diagramas ramificados, chamados de cladogramas, ou mais simplesmente árvores, que retratam de forma hierárquica o parentesco percebido entre as taxas em estudo.

À medida que o número de taxa e caracteres em uma análise aumenta, existem exponencialmente mais árvores e ramificações possíveis desses cladogramas. Essas ramificações podem ser entendidas como potenciais hipóteses da relação entre taxa; devido a essa complexidade da análise computacional, um software deve ser utilizado para explorar as árvores possíveis e escolher aquelas que atendem aos critérios determinados pelo pesquisador na sua análise. Riede, Hoggard e Shennan (2019) argumentam que os arqueólogos precisam desenvolver e aplicar sistemas de classificação (taxonomias) mais estruturados e baseados em teorias para a cultura material, a fim de comparar e integrar efetivamente suas descobertas com dados genéticos, permitindo, em última análise, uma compreensão mais profunda da história humana e da evolução cultural.

Neste artigo, apresentamos o método da construção de cladogramas, com foco nos aspectos históricos dos traços culturais. Cochrane (2004) argumenta que as hipóteses geradas pela organização de dados em árvores podem indicar mudanças na uniformidade e na riqueza das categorias em estudo, sem perder o enfoque interpretativo do fenômeno em questão. Nesse sentido, a tentativa de explicar as relações históricas entre caracteres selecionados, com a elaboração dos cladogramas, tem como objetivo específico rastrear a transmissão cultural.

MATERIAIS

Para esta análise, foram selecionados catorze sítios arqueológicos associados à Tradição Tupiguarani dispersos pelo território do estado de São Paulo, cujas coleções arqueológicas estão armazenadas no Museu de Arqueologia e Etnologia da Universidade de São Paulo (MAE-USP), localizado em São Paulo/SP, no Centro Regional de Arqueologia Ambiental Mário Neme (CRAMN/MAE-USP), em Piraju/SP, e no Museu Histórico da Colonização de Pereira Barreto, em Pereira Barreto/SP.

A seleção dos sítios analisados fundamentou-se em quatro critérios principais. Em primeiro lugar, optamos por contextos arqueológicos cujas coleções estavam acessíveis à pesquisa, possibilitando a observação direta e a sistematização detalhada dos atributos cerâmicos analisados. Em segundo lugar, privilegiamos conjuntos suficientemente amplos e representativos, de modo a garantir uma amostragem estatisticamente consistente, evitando vieses decorrentes de coleções pouco expressivas ou fragmentárias. Ainda, as coleções dos sítios selecionados deveriam conter datas para que o estudo pudesse ser compreendido a partir de um contexto cronológico. Por fim, análises exploratórias preliminares, como a Análise de Componentes Principais (PCA), indicaram que esses sítios apresentavam um comportamento coerente e esperado em relação à variabilidade geral da amostragem, reforçando sua adequação para a aplicação de métodos filogenéticos (Perez, 2022, 2023, 2024). Dessa forma, a escolha dos sítios privilegiou a representatividade regional e cronológica das ocupações Tupiguarani no atual território paulista.

O recorte espacial adotado neste estudo, correspondente ao atual território do estado de São Paulo, não é interpretado como uma unidade cultural pretérita nem como um limite efetivo das dinâmicas de ocupação Tupiguarani no passado. Trata-se, antes, de um recorte analítico, condicionado pela disponibilidade de coleções arqueológicas sistematicamente estudadas, com acesso direto aos materiais e aos dados cronológicos comparáveis. Embora grandes rios como o Tietê, o Paranapanema e o Rio Grande sejam utilizados como referenciais espaciais, estes não são tratados como barreiras culturais, mas como eixos potenciais de circulação, interação e transmissão de conhecimentos técnicos. Assim, as análises aqui apresentadas não pretendem esgotar a compreensão da expansão Tupiguarani em escala macrorregional, mas explorar padrões internos de similaridade e diferenciação observáveis dentro de uma amostra regional controlada.

MÉTODOS: CONSTRUINDO CLADOGRAMAS

Tolstoy (2008) entende a cladística como associada à Biologia Evolutiva, sendo um conjunto de princípios e técnicas para classificar formas vivas (táxons) que se assume estarem relacionadas por meio de descendência comum. Seu produto consiste em diagramas ramificados, chamados cladogramas, ou mais simplesmente árvores, que retratam de forma hierárquica o parentesco percebido entre os táxons em estudo. A partir dessa abordagem, grupos de táxon são organizados com base no fato de compartilharem, ou de seus ancestrais compartilharem, os mesmos caracteres. A lógica da cladística é entender que os conjuntos organizados em nichos representam o esperado em uma história evolutiva que se diverge ao longo das gerações e que os traços individuais de cada conjunto são vistos como inovações adquiridas pelas gerações descendentes, em oposição aos traços observados nos antepassados. Sendo assim, O'Brien *et al.* (2012) entendem que o objetivo da cladística é compreender que nem toda similaridade filogenética entre dois ou mais taxa tem a mesma funcionalidade para uma reconstrução filogenética.

Segundo O'Brien e Lyman (2003), os dados devem passar por ajustes (polarizações/ rotações) provocados pela ordenação desses dados com a presença de grupos irmãos ou *outgroups*. A ausência dessa rotação dificulta a ordenação, tornando impossível fazer afirmações precisas sobre os padrões que se pretende identificar. Além disso, Tolstoy (2008) observa que o objetivo da conversão dos dados para formar uma árvore é encontrar conjuntos ou grupos taxonômicos de caracteres que resultem em árvores mais curtas (ou ótimas). Ele explica que o comprimento, o tamanho dos braços e a quantidade de ramificações observados nas árvores são definidos pelo número de passos, ou seja,

pelas aquisições ou perdas de traços culturais ao longo da evolução, a partir de um ancestral comum. Essas árvores orientadas são chamadas de árvores parcimoniosas.

Assim, a cladística agrupa taxa com base em uma relativa semelhança filogenética com hipotéticos ancestrais comuns (O'Brien; Lyman, 2003). Em estudos de filogenética, temos que dois taxa comuns são mais semelhantes e relacionados se dividirem um ancestral comum do que um terceiro que não compartilha esse ancestral, ou o compartilha de forma mais distante, e a evidência de um ancestral comum. Essa lógica de compartilhamento de um ancestral comum faz com que o táxon seja definido como novo ou derivado. Ainda, é possível inferir que dois taxa tenham um ancestral comum se eles exibirem um carácter derivado diferentemente do terceiro táxon que não apresente tal evidência (Cochrane, 2004, 2015; Garcia-Rivero; O'Brien, 2014; O'Brien; Lyman, 2003; O'Brien *et al.*, 2012).

O'Brien e Lyman (2003), Garcia-Rivero e O'Brien (2014) e Cochrane (2015) resumem em quatro etapas a criação de uma árvore filogenética a partir da cladística, sendo:

1. Gerar uma matriz de dados que mostre os caracteres separados a partir de taxa;
2. Estabelecer direções (polarizações/rotações) da mudança evolutiva entre os estados de cada carácter analisado. Um dos métodos para se fazer isso é estabelecendo análises com *outogroup*;
3. Construir um diagrama ramificado (cladograma) em que se relacionam cada um dos taxa para unir os caracteres derivados;
4. Construir um cladograma que seja consistente com o maior número de caracteres e, portanto, uma árvore parcimoniosa.

Após a elaboração da árvore parcimoniosa, O'Brien e Lyman (2003) sugerem alguns cálculos da robustez para avaliar os dados. Isso significa que é necessário apresentar o quanto bem uma árvore filogenética poderia retratar a realidade, e comparada à árvore mais parcimoniosa elaborada. Os autores sugerem o cálculo do Índice de Consistência (IC), que representa o número de homoplasias existentes no conjunto de dados analisados. Isso quer dizer que esse número representaria a quantidade de características comuns entre indivíduos, mesmo que não apresentassem um ancestral comum. Esse índice pode variar de zero (0) a um (1), sendo que zero (0) representa uma homoplasia completa (apresentam grande quantidade de características comuns), e um (1) não existe homoplasia na amostra (não apresenta qualquer afinidade entre os indivíduos). Apesar disso, esse cálculo está diretamente relacionado à quantidade de caracteres inseridos na matriz analisada, e o valor do IC não é completamente associado à quantidade de homoplasias, mas apresenta a relação entre a quantidade de caracteres analisados e as homoplasias existentes na amostra, fazendo com que esse índice não seja suficiente para comparação entre pesquisas com a mesma abordagem metodológica (Collard; Shennan; Tehrani, 2006; Crema; Kerig; Shennan, 2014; Garcia-Rivero; O'Brien, 2014; O'Brien; Lyman, 2003).

O'Brien e Lyman (2003) ainda propõem o uso do Índice de Retenção (IR), que mede o ajuste dos caracteres sob análise em relação à proporção entre a sinapomorfia aparente e a sinapomorfia real da amostra. A sinapomorfia é um caractere (uma característica ou traço) derivado, ou seja, uma novidade evolutiva e compartilhada por um grupo de organismos. Quando se mede o grau de similaridade representado pelas sinapomorfias, o IR fornece uma indicação de quão unidos e relacionados estão os taxa. Dessa forma, é possível identificar um possível ajuste do que está sob representação no cladograma e a realidade da transmissão cultural no mundo real. Por isso, o IR mede o número de semelhanças em um conjunto de dados de uma determinada árvore. Esse índice é insensível tanto à presença de caracteres derivados, que estão presentes em apenas um único táxon, quanto ao número de caracteres ou taxa empregados. Os índices variam de zero (0), que indica falta de ajuste entre uma árvore e o conjunto de dados sob análise, e um (1), que representa um ajuste perfeito.

Segundo Garcia-Rivero e O'Brien (2014), o IR tem sido um índice usado para comparar pesquisas em que são envolvidos números de táxon semelhantes. Isso quer dizer que por meio da verificação do IR, em conjunto com a elaboração de cladogramas parcimoniosos, seria possível verificar se os dados sob análise apresentam ajuste aos dados biológicos ou culturais. Para esses autores, os IRs associados aos conjuntos de análise com taxa biológicos médios apresentam mínimo e máximo entre 0,35 e 0,94, respectivamente. Já os valores correspondentes para os IRs culturais são 0,42 e 0,78. (Collard; Shennan, Tehrani, 2006; Crema; Kerig; Shennan, 2014). Crema, Kerig e Shennan (2014) entendem que valores abaixo dessa faixa raramente são relatados na literatura arqueológica e, portanto, poucos estudos exploraram suas implicações, mas Cochrane e Lipo (2010) obtiveram valores de IR muito baixos (0,35 e 0,49) para seu estudo de cerâmica Lapita e explicam que o resultado identificado naquela pesquisa seria evidência de transmissão horizontal. Além disso, Collard, Shennan e Tehrani (2006), ao investigarem os processos de mistura, comércio e trocas em populações humanas, também relatam um baixo valor para o IR (0,46), o qual descrevem como um possível efeito de processos de reticulação após eventos iniciais de propagação e ramificação diagnosticados ao longo da pesquisa. Essas observações são valiosas para a interpretação final dos dados obtidos nesta pesquisa.

Tehrani e Collard (2002) oferecem outra maneira para interpretar os IC e os IR gerados a partir das relações possibilitadas pela etnogênese e filogênese de um conjunto de dados. Os autores apresentaram um estudo de caso em que o IC da amostra seria de 0,68, o que indicaria a quantidade de caracteres que são homólogos entre os taxa e, por correspondência, suas similaridades. Assim, o IC de 0,68 sugeriria que ~70% da amostra apresente filogênese, enquanto ~30% corresponderia às homoplasias ou à etnogênese. Por filogênese, Tehrani e Collard (2002) percebem a evolução como resultado de uma subdivisão progressiva dos conjuntos culturais, que ocorrem à medida que as populações se dividem e dão origem a novas populações. Desse modo, a filogênese cultural é análoga à filogênese biológica e possibilita associação entre padrões culturais aos dados genéticos e/ou linguísticos. Já por etnogênese, os autores entendem os processos caracterizados por contato, empréstimo, mistura, comércio e troca, tanto de ideias e práticas quanto de objetos entre populações que são contemporâneas, de forma semelhante ao fluxo gênico de uma espécie (Tehrani; Collard, 2002).

Para tornar essa discussão mais explicativa, Boyd *et al.* (1997) detalham os modelos sobre a evolução cultural no que diz respeito à filogênese e etnogênese empregando quatro modelos para descrever esse processo evolutivo, sendo: 1) "*cultura como espécie*", em que a cultura é vista como uma unidade coerente e isolada de influências externas e, portanto, é derivada inteiramente da descendência com modificações – neste aspecto é considerada apenas a filogênese da cultura ignorando a etnogênese; 2) "*cultura como sistema hierárquico integrado*", que entende a cultura como um conjunto de elementos centrais e tradicionais e elementos periféricos. Boyd *et al.* (1997) sugerem que os sistemas centrais seriam resultado da descendência com modificação, então, estes podem existir por longos períodos, apesar das altas taxas de interação entre membros de populações adjacentes; em contraste a isso, os elementos periféricos seriam resultado da difusão e rápidas mudanças ocorridas – a filogênese tem maior participação na mudança cultural do que a etnogênese; 3) "*conjunto e muitas unidades coerentes*", neste modelo, nega-se a existência de um núcleo cultural e postula-se que as culturas correspondem a aglomerados em constantes mudanças herdadas de diferentes fontes e diversas maneiras. Neste caso, as tradições individuais poderiam ser relativamente longas temporalmente e se originar através de linhagens, mas a própria cultura seria efêmera, e as relações entre elas, reticulares – neste caso, existe uma atuação maior da etnogênese na mudança da cultura

em relação à filogênese; por fim: 4) é o modelo representado por uma versão extrema e hipotética da etnogênese, que os autores nomeiam “coleção de entidades efêmeras”, e não há tradições culturais, uma vez que os indivíduos sempre tomam decisões imparciais sobre como agir com base em tentativa e erro e nas alternativas às quais estão expostos.

Diante desses aspectos, a utilização dos IC e IR oferece parâmetros fundamentais para avaliar a força dos agrupamentos observados e o grau de homoplasia presente nos dados. Embora tais índices mensurem exclusivamente a distribuição dos caracteres no cladograma, sua interpretação conjunta com informações cronológicas e espaciais dos sítios arqueológicos permitem discutir a dinâmica de continuidade e interação cultural entre as populações analisadas, isto é, a etnogênese e a filogênese. Nesse sentido, ao adotar árvores parcimoniosas, bem como a combinação de IC e IR com dados contextuais de espaço e temporalidade, pode-se perceber os possíveis episódios de convergência ou contato horizontal entre os indivíduos analisados.

O uso de árvores parcimoniosas vem em detrimento de aplicações em outros estudos semelhantes, o que permite a comparabilidade direta das pesquisas e garante que os resultados aqui obtidos possam ser avaliados em continuidade a esse corpo literário. Ainda, essa abordagem se adequa aos dados morfológicos e tecnológicos aqui empregados, nos quais o número de caracteres informativos e a possibilidade de homoplasias podem dificultar a implementação de outros tipos de modelos probabilísticos, como o bayesiano. Por fim, dada a natureza exploratória e preliminar apresentada neste contexto de pesquisa, que visa testar a viabilidade do método para a cerâmica Tupiguarani, a parcimônia se mostra apropriada por sua simplicidade interpretativa, sem excluir a possibilidade de que estudos venham a empregar outras formas de análise.

Para este trabalho, utilizamos o software PAUP* versão 4,0¹ (Swofford, 2001) para a construção dos cladogramas. A opção por esse software se deu por se tratar de um programa consolidado no campo de estudos filogenéticos, com ampla documentação e suporte de manuais e fóruns especializados online, bem como por conta da possibilidade da comparabilidade direta com estudos anteriores na mesma área de atuação, como os desenvolvidos por Cochrane (2004). Assim, a escolha do software é favorecida pela confiabilidade técnica das análises e a continuidade de diálogos com a literatura existente.

O procedimento se inicia com a elaboração de uma matriz de dados binários com uso da ferramenta de software Excel, e a partir do uso de um aplicativo de edição de arquivos de texto se adiciona uma extensão #NEXUS. Esse procedimento é fundamental para a compatibilidade entre os dados da matriz e o software PAUP (Perez, 2025). Nossas análises também foram conduzidas no ambiente R e obtivemos os mesmos resultados apresentados neste artigo (veja o material suplementar para o código para a elaboração em R).

Por se tratar de análise feita a partir de dados binários na matriz, e para auxiliar nas interpretações das árvores cladísticas, utilizamos ainda uma Ordenação Não-métrica Multidimensional NMDS (*Non-metric Multidimensional Scaling*), ou Ordenação Multivariada Não-métrica, para representar a dimensionalidade deste tipo de análise, a qual preserva as relações entre as amostras baseando-se em similaridades e distâncias entre os pontos distribuídos em um plano cartesiano (Holland, 2008; Rabinowitz, 1975). Ainda, aplicamos uma representação conhecida como Árvore Geradora Mínima (*Minimum Spanning Tree – MST*), isto é, uma estrutura matemática que possibilita a visualização e a simplificação das relações de correspondência entre as amostras para entender possíveis conexões

¹ Homepage PAUP. Disponível em: <https://paup.phylosolutions.com/>. Acesso em: 3 out. 2023.

e semelhanças entre elas. A partir da *MST*, é possível investigar associações entre as variáveis e transformar os dados em uma representação gráfica em um espaço bidimensional, sendo que as distâncias entre os pontos refletem a proximidade ou similaridade entre as categorias analisadas (Chandy; Lo, 1973; Grygorash; Zhou; Jorgensen, 2006). Essas análises foram feitas a partir do software Past versão 4.03 (Statistical analysis software) – escolhido por ser um software de código livre que apresenta uma interface simples para manipulação. Nesse sentido, os resultados alcançados são reforçados pelos dados e pelas distribuições representadas entre os pontos e as linhas nos gráficos gerados (Figuras 4 e 5).

APLICANDO O MÉTODO – OS DADOS

Conforme mencionado, foram selecionados catorze sítios arqueológicos que atendessem às condições apresentadas anteriormente. O Quadro 1 organiza os dados dos sítios envolvidos nesta análise e tem como base os estudos de Perez (2022, 2023, 2024). Nota-se que os sítios Prema Tupi, Cardoso, De Rosa e Dourado 2, embora não apresentem datações, foram incorporados na análise com o intuito de uma possível aproximação temporal que pudesse ser estabelecida para eles.

Quadro 1. Sítios usados para análise cladística e estatística. Todas as idades de radiocarbono neste quadro foram calibradas usando CalPal, versão 2022.9, curva INTCAL 2020 (Weninger; Jöris, 2008).

Nome do sítio	Município	Bacia Hidrográfica	Idade AP (Método)	Idade Cal 14C	Idade Calêndrica (aproximada)	Referência
Almeida	Tejupá	Paranapanema River	1500±50 (TL) 930±100 (TL) 560±60 (TL) 470±50 (TL)		EC 440 EC 1000 EC 1400 EC 1500	Scatamacchia (1981)
Ary Carneiro	Suzanápolis	Tietê River	2200±? (TL)		250 AEC	Maranca, Silva e Scabello (1994)
Bianco	Itapeva	Paranapanema River	1630±180 (OSL) 1320±140 (OSL) 930±60 (TL) 810±50 (TL) 770±30 (¹⁴ C) 760±60 (TL) 750±80 (OSL) 590±60 (OSL) 500±40 (TL)	697±28	EC 390 EC 690 EC 1090 EC 1200 EC 1250 EC 1250 EC 1260 EC 1420 EC 1510	Este artigo
Cardoso	Arealva	Tietê River	-		-	Moraes(2007)
De Rosa	Ibitinga	Tietê River	-			Moraes(2007)
Dourado 2	Promissão	Tietê River	870±60 (TL)		EC 1140	Este artigo
Fonseca	Itapeva	Paranapanema River	1220±120 (¹⁴ C) 1190±120 (TL) 1110±110 (TL) 1100±100 (TL) 1010±100 (TL) 970±100 (TL)	1115±158	EC 820 EC 780 EC 860 EC 860 EC 960 EC 1000	Araujo (2001)

continua...

Quadro 1. Continuação

Nome do sítio	Município	Bacia Hidrográfica	Idade AP (Método)	Idade Cal 14C	Idade Calêndrica (aproximada)	Referência
Gramado	Brotas	Tietê River	400±9 (TL) 190±20 (TL)		EC 1600 EC 1800	Moraes(2007)
Lambari II	Casa Branca	Rio Grande River	1085±130 (TL)		EC 920	Moraes(2007)
Mata da Figueira	Candido Mota	Paranapanema River	340±35 (TL)		EC 1660	Morais (2001)
Mineração	Iguape	South Coast	560±60 (¹⁴ C) 550±50 (¹⁴ C) 520±60 (¹⁴ C) 450±60 (¹⁴ C)	574±57 572±53 548±68 449±90	EC 1376 EC 1378 EC 1402 EC 1501	Scatamacchia e Uchôa (1993)
Nunes	Piraju	Paranapanema River	880±90 (TL) 879±90 (TL)		EC 1100 EC 1100	Morais (1985)
Pajeu	Candido Mota	Paranapanema River	875±90 (TL) 340±35 (TL)		EC 1130 EC 1670	Moraes (2007)
Prema Tupi	Rio Claro	Tietê River	-		-	Perez (2012)

Fonte: Adaptado das referências bibliográficas pelos autores.

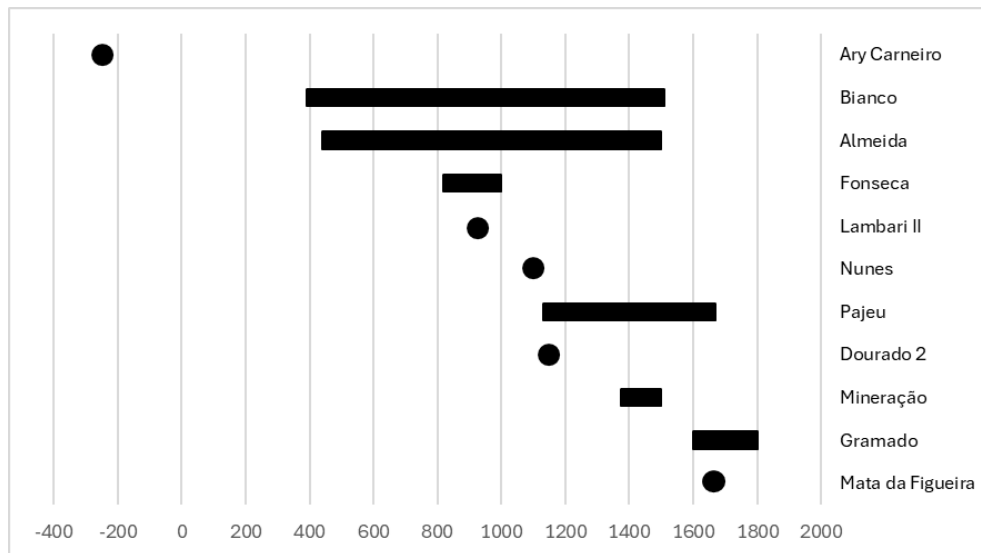
A Figura 2 apresenta um diagrama com uma linha do tempo de sítios arqueológicos. O eixo horizontal indica o tempo, variando de aproximadamente 400 AEC. a 2000 EC. O eixo vertical lista sítios arqueológicos. As barras horizontais pretas denotam períodos durante os quais sítios específicos foram ocupados, enquanto os pontos pretos representam datas isoladas, como resultados de datação por radiocarbono. O Sítio Ary Carneiro é o sítio mais antigo, datando de cerca de 250 AEC, enquanto Bianco mostra longas ocupações começando antes do ano 390 EC. e continuando até aproximadamente 1510 EC. O sítio de Almeida apresenta uma duração notável de 440 EC a 1500 EC. Fonseca apresenta um período de ocupação mais curto, de 820 EC a 1000 EC. Lambari II e Nunes aparecem com datas isoladas e períodos de ocupação mais curtos. Sítios mais recentes, incluindo Dourado 2, Mineração e Gramado, foram ocupados entre 1200 e 1800 EC. Mata da Figueira apresenta uma única data por volta de 1660 EC., indicando um evento específico. O hiato cronológico entre o sítio Ary Carneiro e os demais ainda não pode refutado, dado que a datação do sítio Ary Carneiro se deu durante os trabalhos sistemáticos no oeste paulista entre os anos 1970 e 1990 durante a atuação do Projeto Oeste Paulista, coordenado pela Dra. Silvia Maranca. Atualmente, não é possível realizar uma contraprova dessa data, pois o sítio encontra-se submerso pela Barragem de Pereira Barreto/SP.

A Figura 3 ilustra a localização dos sítios arqueológicos analisados no estado de São Paulo, juntamente com as principais bacias hidrográficas, incluindo as do Rio Grande, do Rio Tietê e do Rio Paranapanema.

Em estudos prévios (Perez, 2012, 2022, 2023, 2024), foram sistematicamente descritos os aspectos tecnológicos, morfológicos, decorativos e dimensionais das coleções cerâmicas provenientes dos catorze sítios arqueológicos listados no Quadro 1. A partir dessa caracterização, os atributos foram convertidos em um sistema binário, registrando presenças (1) e ausências (0), resultando em uma matriz com 523 caracteres (Perez, 2025 – observar a Matriz final 2.1), da mesma maneira que é exemplificado

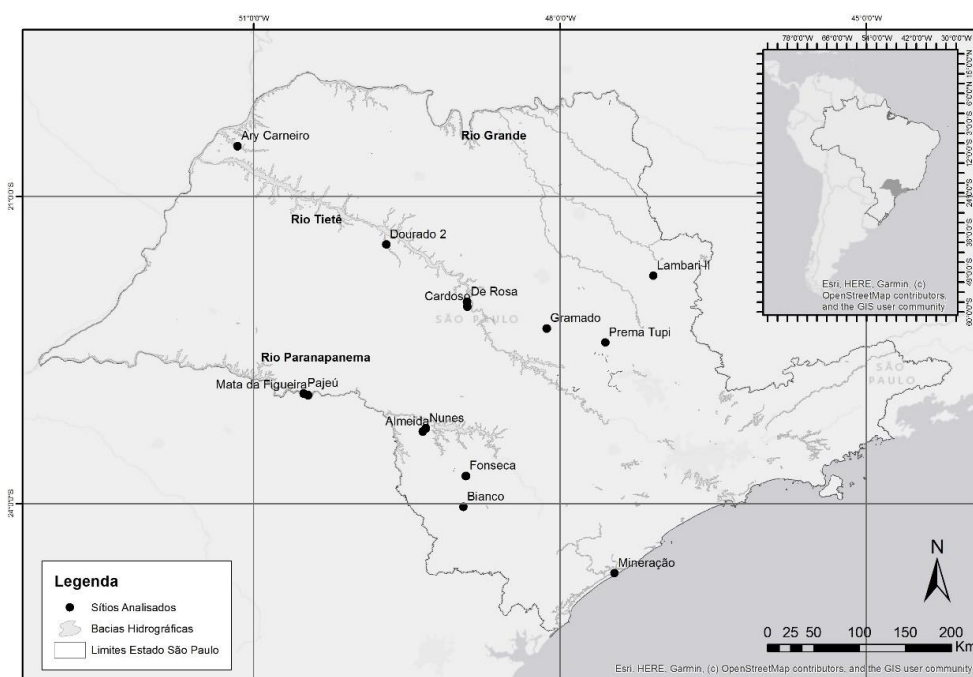
no Quadro 2 (Perez, 2025 – observar a Matriz final 2.2). Os atributos considerados abrangem tratamentos e acabamentos de superfície, grafismos e pinturas, morfologias de borda, além da identificação de temperos e antiplásticos, reunindo, assim, um conjunto de informações tecnológicas, morfológicas e decorativas que serviram de base para a comparação filogenética entre os sítios.

Figura 2. Diagrama ilustrando a distribuição temporal das datas para os locais analisados. Pontos e linhas pretos apresentam a distribuição espacial aproximada das datações apresentadas no Quadro 1. Valores negativos representando anos AEC. e valores positivos representando anos EC.



Fonte: Os autores (2025).

Figura 3. Mapa da distribuição espacial de cada um dos sítios envolvidos nesta análise.



Fonte: Os autores (2025).

Converter os atributos em matriz é um procedimento amplamente utilizado em análises cladísticas aplicadas à cultura material (Cochrane, 2004, 2015; Cochrane; Lipo, 2010). Nesta opção, por variáveis binárias, buscou-se reduzir distorções associadas a diferenças no tamanho das amostras entre os sítios e enfatizar o potencial de compartilhamento de traços, independentemente de sua frequência relativa em cada conjunto. Nesse contexto, a ausência de um atributo não deve ser interpretada como evidência de “transmissão negativa” ou de proibição cultural, mas apenas como um estado informacional do conjunto analisado, padrões de persistência de um traço cultural e recorrência de atributos. O objetivo da codificação binária é identificar padrões de simultaneidade de informações, e não medir a intensidade ou a predominância de determinados atributos.

De maneira que se possa exemplificar essa conversão binária, temos os caracteres representados pelas características no Quadro 2, que apresenta o aspecto analisado, uma abreviação do caractere, o próprio caractere descrito, e na coluna associada ao nome do Sítio, a ausência (0) ou a presença (1).

Quadro 2. Um exemplo da disposição das características analisadas e a conversão para o sistema binário. Os dados estão em inglês para facilitar o alcance da reprodutibilidade das informações apresentadas neste artigo (Perez, 2025 – Matriz final 2.2).

Aspecto analisado	Abreviação	Caracteres	Nome do sítio
Tempero ou antiplástico	GG	Grog/Caco Moído	1
	MN	Mineral/Mineral	1
	CH	Charcoal/Carvão	0
	SH	Shell/Concha	0
	CHMN	charcoal, mineral/Carvão, mineral	1
	GGMN	Grog, mineral/Caco moído, mineral	1

Fonte: Os autores (2025).

Após a conversão dos dados e a inclusão da extensão #NEXUS, a matriz está pronta para a inserção no software PAUP, quando as árvores cladísticas poderão ser geradas. O processo começa com uma análise heurística, e o software é calibrado para gerar até dez mil árvores filogenéticas, e aquela com menor comprimento da árvore, isto é, a mais parcimoniosa, deverá ser escolhida. Neste procedimento, o software usa o método “Tree-Bisection and Reconnection” (TBR) para gerar árvores adicionais e identificar possíveis combinações com ramos mais curtos; e se alguma delas for identificada, o próprio software realiza a substituição na árvore parcimoniosa, e o processo continua automaticamente até que nenhuma outra árvore possa existir dentro da simulação verificada.

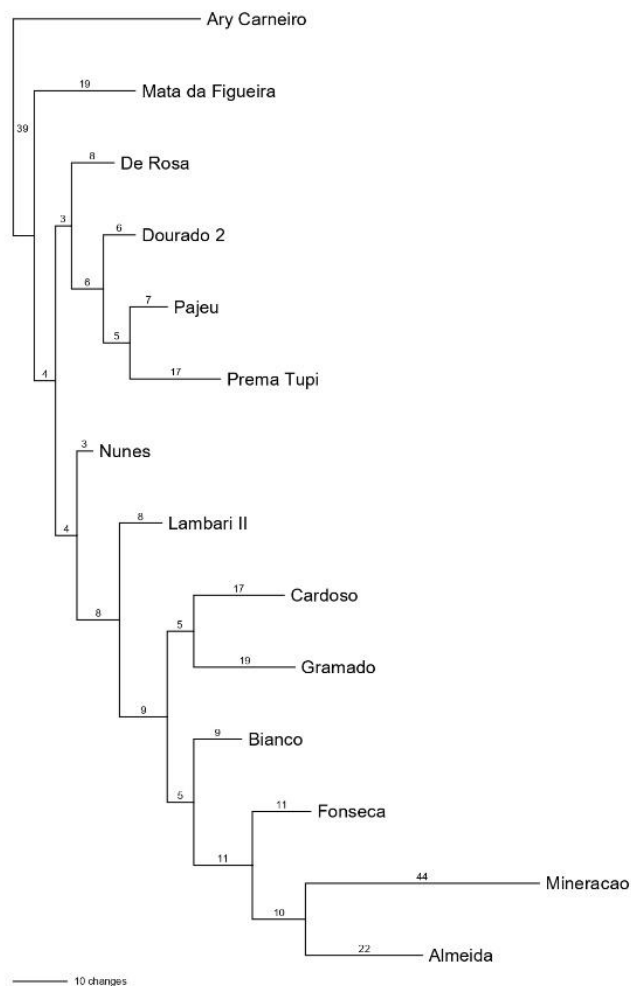
A seguir, são apresentados os resultados tanto da elaboração cladística quanto da aplicação do método de NMDS e da MST. Todos esses processos serão particularmente importantes para a discussão dos resultados obtidos.

RESULTADOS

Os resultados apresentados para esta configuração envolveram os 523 caracteres, dos catorze sítios selecionados. A árvore representada foi escolhida pelo próprio software (sendo a de comprimento 299) com IC de 0,5485 e RI de 0,3807. Considerando as etapas

de elaboração da árvore descrita, temos a representação gráfica das relações filogenéticas entre os sítios analisados expressos na Figura 4.

Figura 4. Árvore Cladística com simulação dos contatos existentes entre os catorze sítios associados à Tradição Tupiguarani para o atual estado de São Paulo.



Fonte: Os autores (2025).

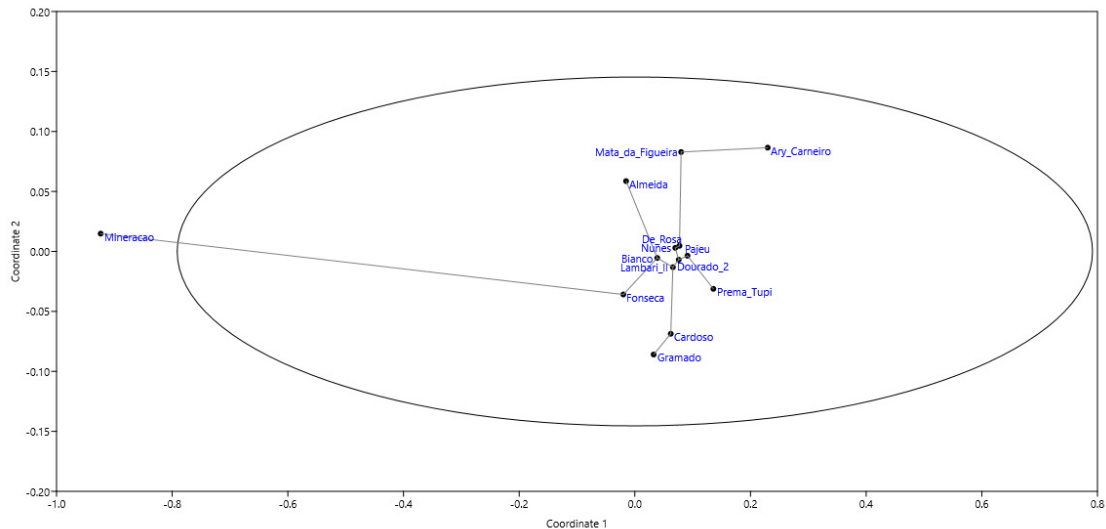
Neste cladograma (Figura 4), temos o sítio Ary Carneiro como *outlier* escolhido propositalmente, por ser o sítio com idade mais recuada, seguido por uma ramificação do sítio Mata da Figueira e outra ramificação que separa os sítios De Rosa, Dourado 2, Pajeú e Prema Tupi (exceto o sítio Pajeú, todos os outros estão na bacia do Rio Tietê); seguida por uma ramificação com o sítio Nunes, localizado no Rio Paranapanema, e dos sítios Lambari II, Cardoso, Gramado da bacia do Rio Tietê; e outra ramificação formada pelo sítios Bianco, Fonseca, Mineração e Almeida (exceto o Mineração, que está no litoral sul, todos os outros são alocados na bacia do médio-alto Paranapanema). Na Figura 4, é sugerida uma possível separação litoral-interior do estado de São Paulo, bem como é indicada a diferenciação entre os sítios da bacia do Rio Tietê e do Rio Paranapanema.

Para corroborar a interpretação da árvore filogenética, temos a verificação a partir de estudos da NMDS dos dados, bem como da *MST*, que também considerou os 523

caracteres representados no sistema binário, sendo os mesmos aspectos tecnológicos, morfológicos e decorativos utilizados para a construção da Figura 4.

Para essa discussão, foi gerada a Figura 5.

Figura 5. Gráfico com representação dos sítios sob análise Ordenação Não-métrica Multidimensional - NMDS (Non-metric Multidimensional Scaling) e conectados a partir de uma Árvore Geradora Mínima (*Minimum Spanning Tree*).



Fonte: Os autores (2025).

O gráfico de representação NMDS e *MST* (Figura 5) apresenta o sítio Mineração em uma extremidade e o Ary Carneiro em outra. Ao centro, há uma conexão entre os demais sítios que perpassa pelos sítios Mata da Figueira, De Rosa, Dourado 2, Pajeú, Prema Tupi e continua para os sítios Branco, Nunes, Lambari II, Almeida e segue para os sítios Cardoso, Gramado, Fonseca e Mineração. Nesse gráfico, não se escolhe um *outlier*, o próprio programa organiza espacialmente os indivíduos sob análise. Essa distribuição representa as conexões possíveis entre os sítios e segue a lógica da Figura 4 – uma maior conexão entre os sítios da bacia do Rio Tietê e do médio-alto Paranapanema.

Para interpretar a relação entre as Figuras 4 e 5, é fundamental considerar os aspectos de espacialidade (Figura 3) e temporalidade (Figura 2), o quais são essenciais para compreender possíveis conexões que poderiam indicar semelhanças e diferenças entre os sítios arqueológicos analisados, dadas as proximidades espaciais e temporais entre eles. Sob essa perspectiva, destaca-se a posição geográfica do sítio Ary Carneiro, localizado no extremo oeste do estado de São Paulo (na Figura 4 foi considerado o *outlier*, e na Figura 5 está à direita) e, associado à datação mais recuada, em contraste com o sítio Mineração, situado no litoral sul, no extremo leste do estado de São Paulo (na Figura 4, está na última ramificação, e na Figura 5 está na extrema esquerda da figura). Além disso, na Figura 5, verifica-se uma concentração dos sítios localizados no Paranapanema (Nunes, Branco, Fonseca, Almeida) em oposição aos sítios do Tietê (Gramado, Cardoso, Prema Tupi e Dourado 2) e reflete o mesmo padrão observado na Figura 4.

De modo geral, o desenho dessa árvore filogenética (Figura 4) segue os mesmos parâmetros de distribuição observados no gráfico NMDS e *MST* (Figura 5), comparados à distribuição espacial do Mapa (Figura 3). Isto quer dizer que o teste envolvendo a NMDS

apresenta a mesma configuração da distribuição filogenética dos sítios no território do estado de São Paulo.

Essas representações gráficas evidenciam a dinâmica de espacialização entre os contextos de interior e litoral, destacando padrões significativos de interação e diferenciação entre os sítios arqueológicos. A partir do Sítio Ary Carneiro, no extremo oeste do estado, os sítios localizados no centro do estado de São Paulo apresentam uma notável homogeneidade nas características analisadas; a princípio, sugerindo uma partilha de informações culturais, tecnológicas ou de práticas sociais que provavelmente refletem uma interação mais próxima, ou ainda um fluxo constante de trocas ao longo do território e do tempo. Essa homogeneidade contrasta com o sítio Mineração, situado no litoral, que demonstra maior distância em relação aos sítios do interior, tanto espacial quanto culturalmente.

Em relação à cronologia demonstrada, os sítios arqueológicos analisados apresentam escalas temporais distintas, variando desde ocupações relativamente breves (como nos sítios Ary Carneiro, LambariII, Nunes, Dourado 2 e o Mata da Figueira), de algumas gerações, até contextos com evidências de uso prolongado ao longo de vários séculos (como nos sítios Bianco, Almeida, Fonseca, Pajeu, Mineração e Gramado). Esses conjuntos devem ser compreendidos como palimpsestos arqueológicos, nos quais materiais de diferentes momentos históricos podem coexistir em um mesmo contexto espacial. Esta análise não busca reconstruir sequências cronológicas finas nem eventos sincrônicos específicos, mas investigar padrões estruturais de recorrência e persistência de atributos cerâmicos ao nível dos conjuntos agregados, isto é, coleções arqueológicas formadas e disponíveis nos diversos museus consultados. Reconhece-se que essa opção metodológica implica uma simplificação do fenômeno temporal, assumida conscientemente como parte de uma abordagem exploratória voltada à identificação de tendências gerais do fenômeno de transmissão cultural.

De modo geral, esta análise confirma o que tem sido discutido na atualidade sobre a distribuição dos sítios Tupiguarani no estado de São Paulo. Noelli (2025) entende que potencialmente essa distribuição e as conexões espaciais entre os sítios aqui analisados se dariam pela razão de todos serem considerados Tupiniquins, ou seja, sítios que se apresentam dispersos desde Itaquaquecetuba até a Foz do Rio Tietê, no extremo oeste paulista. Esses sítios Tupiniquins estariam distribuídos ainda no médio e alto vale do Rio Paranapanema, e em oposição aos sítios do baixo Paranapanema que não foram considerados nesta amostragem, mas que segundo Noelli (2025) seriam caracterizados como Guarani (típicos). Temos a divergência do sítio Mineração, que está localizado em área Tupiniquim (Noelli, 2025), mas nesta organização, o sítio se apresenta como *outlier* na discussão. Já na ordenação apresentada por Noelli (2025), são considerados aspectos linguísticos e históricos, mas não são considerados dados arqueológicos. Assim, este artigo vem para colaborar com essa discussão ampliando as inferências a respeito da distribuição populacional associada à Tradição Tupiguarani no contexto paulista. De fato, a projeção de uma situação colonial (“Tupiniquim”, “Guarani” etc.) para um passado mais remoto demanda a exploração dos dados arqueológicos por meio de procedimentos rigorosos, como o que apresentamos aqui.

DISCUSSÃO

Neste artigo, apresentamos a aplicação do método cladístico ao contexto cultural, com foco em caracteres associados à Tradição Tupiguarani em sítios arqueológicos do estado de São Paulo. A abordagem foi empregada para produzir inferências acerca de uma das questões mais significativas da arqueologia: a mudança na cultura material

e possíveis interações entre sítios arqueológicos. Nesse sentido, a cladística mostra-se uma ferramenta possível para a formulação de hipóteses sobre a dispersão dessa população, permitindo explicações a partir das relações evolutivas entre os conjuntos analisados.

O'Brien *et al.* (2012), ao retomarem os estudos do paleontólogo George Gaylord Simpson na década de 1960, oferecem um exemplo ilustrativo para compreender hipóteses de similaridade e relações evolutivas. Os autores lembram que gêmeos monozigóticos não são idênticos porque se parecem, ao contrário, são semelhantes porque compartilham uma história comum, ou seja, a semelhança decorre da ancestralidade compartilhada (O'Brien *et al.*, 2012). De modo análogo, a cladística auxilia a desvendar histórias comuns, sejam elas de organismos, de sítios arqueológicos ou de objetos da cultura material, como fragmentos cerâmicos de coleções arqueológicas, porque houve necessariamente um mecanismo de transmissão e herdabilidade.

No caso analisado, deve-se considerar a longa duração da ocupação Tupiguarani no território paulista e o fato de que, mesmo dispersos em amplas áreas, esses grupos mantiveram intenso contato cultural entre regiões distintas. Os sítios envolvidos na análise apresentam datações que variam aproximadamente entre 250 AEC e 1800 EC (Quadro 1). Durante a elaboração da matriz de dados, observou-se uma proporção significativa dos caracteres que apresentam estados invariantes entre os sítios analisados, e essa homogeneidade dos dados da matriz não constitui uma limitação inesperada, mas reflete a existência de uma base técnica, decorativa e morfológica amplamente compartilhada entre os grupos associados à Tradição Tupiguarani. Embora caracteres invariantes reduzam a resolução filogenética, eles reforçam a interpretação de uma tradição tecnológica comum, sobre a qual se desenvolvem variações locais e regionais. Além disso, os testes de MSC indicam homogeneidade entre os indivíduos da amostra, o que reforça a identificação de uma Tradição Tupiguarani distribuída por todo o território paulista (Perez, 2022, 2023, 2024).

Ao mesmo tempo, a árvore filogenética mostra que alguns sítios são mais parecidos entre si do que outros. Isso abre um panorama muito mais amplo e rico para futuras abordagens que explorem a variabilidade interna da Tradição Tupiguarani, mostrando a presença de esferas de maior ou menor interação dentro do território. No entanto, não se espera que tais esferas de interação sejam fixas, posto que necessariamente houve um processo evolutivo durante a expansão dos grupos Tupi no território. Fases iniciais de expansão implicam em esferas de interação diferentes das fases intermediárias, quando os vales de diferentes rios (Paranapanema, Tietê, Grande, Mogi-Guaçu) serviam como rotas de transporte, bem como das fases finais, quando o território todo já estava ocupado.

Essa abordagem também possibilitou propor um posicionamento cronológico para os sítios Prema Tupi, Cardoso, e De Rosa, cuja cronologia era até então desconhecida e podem ser correlacionadas através da semelhança das coleções dos sítios como Pajeú, Lambari II e Gramado, indicando provavelmente que essas ocupações sejam posteriores ao século X, como são esses sítios. Esse exercício de correlação favorece a compreensão sobre os períodos de ocupação e sobre os processos de interação cultural entre os grupos humanos que habitaram essas áreas, compatíveis com a cronologia média dos sítios da Tradição Tupiguarani em São Paulo (Corrêa, 2024).

Quanto aos resultados dos IC e IR, obtivemos um IC de 0,54, o que significa que cerca de 54% dos caracteres são consistentes com uma hipótese evolutiva sem homoplasia, refletindo traços herdados de um ancestral comum. Esse valor se alinha a resultados frequentemente observados em análises de dados arqueológicos, nos quais a complexidade cultural envolve tanto heranças comuns quanto inovações locais. O IR, por sua vez, apresentou valor de 0,38, considerado baixo segundo Crema *et al.* (2014)

e Cochrane e Lipo (2010). Tais valores reduzidos de IR podem ser interpretados como resultado de transmissão horizontal ou convergência funcional. Contudo, em contextos arqueológicos de longa duração, como o aqui analisado, a homoplasia também pode refletir a persistência, reativação ou retomada de traços ancestrais compartilhados, associados a uma memória técnica ou simbólica comum. Exemplos desse processo podem ser observados no sítio Fonseca, com idades distribuídas em um intervalo de cerca de 220 anos, e no sítio Pajeú, que abrange aproximadamente 530 anos (Quadro 1 e Figura 2).

A cerâmica, enquanto categoria artefactual altamente plástica, está sujeita a ciclos de inovação, esquecimento e retomada, de modo que a recorrência de determinados atributos em sítios distintos não implica necessariamente contato intergrupar recente. Assim, a homoplasia observada pode ser interpretada como resultado combinado de transmissão horizontal, convergência funcional e continuidade histórica de tradições técnicas de longa duração. Esse padrão é reforçado pelas análises de *MST*, que evidenciam proximidade e interconexões entre sítios dos vales do Tietê e do Paranapanema. Quando integrados, os resultados da árvore filogenética, do gráfico de NMDS e os modelos interpretativos de Boyd *et al.* (1994) ampliam a resolução interpretativa. Para esses autores, as culturas podem ser entendidas como sistemas hierárquicos compostos por elementos centrais (herdados de uma tradição comum) e periféricos (resultantes de variações ou contatos). Essa abordagem oferece ferramentas conceituais para discutir a etnogênese na cultura material da Tradição Tupiguarani em São Paulo.

Garcia-Rivero e O'Brien (2014) oferecem um aporte adicional ao interpretar esse tipo de padrão com a expressão "*core tradition*", ou tradição central conservadora, caracterizada por elementos culturais que mantêm relações históricas quando restritos a determinados limites geográficos. Nesta perspectiva, os autores reforçam o entendimento de que o padrão demonstrado a partir da Figura 4 representa um processo de mudança cultural que se deve a contatos externos e também a um processo de diferenciação ao longo do tempo, a partir de um ancestral comum. A homogeneidade revelada pelas análises de NMDS corrobora essa interpretação, indicando o uso sistemático de conhecimento compartilhado e de estruturas ideológicas comuns entre os sítios que representam as populações estudadas, como ilustrado pela proximidade entre os sítios Nunes, Lambari II, Cardoso, Gramado, Pajeú e Prema Tupi.

Importa ressaltar que a árvore cladística não fornece evidências diretas dos processos etnogenéticos, mas o IR permite identificar padrões de diferenciação e homoplasia que sugerem trajetórias de interação cultural. Assim, os dados aqui apresentados dialogam com modelos de ocupação do sul e sudeste brasileiro, segundo os quais São Paulo constitui uma zona de encontro no contexto da expansão Tupiguarani. Nossas análises sugerem que, mais do que simples ponto de passagem, a região se configurou como espaço de diferenciação interna, em que os atributos e características das cerâmicas produzidas nesses locais foram compartilhados, disseminados e reelaborados localmente.

CONCLUSÕES

Em síntese, a transmissão de informações dentro do sistema cultural responsável pelo que chamamos de Tradição Tupiguarani foi provavelmente mais ampla e repleta de meandros do que a representação da árvore cladística isolada poderia ilustrar a partir das suas ramificações. As diferenças observadas entre sítios da costa e do interior, e entre distintas bacias hidrográficas, podem ser explicadas por dinâmicas variadas de ocupação e interação entre os diferentes grupos humanos. O panorama resultante aponta para um processo intrincado e complexo de ocupação e transmissão cultural no território do atual estado de São Paulo.

Os resultados não contradizem modelos clássicos como os de Métraux, Meggers, Brochado, Noelli ou, mais recentemente, Bonomo sobre as ocupações na região sul/sudeste do Brasil, mas acrescentam elementos ao debate, como as conexões existentes entre os sítios dos vales do Tietê e Paranapanema, e o distanciamento dos sítios da costa. As árvores cladísticas sugerem a predominância de padrões filogenéticos de diferenciação temporal a partir de um ancestral comum, permeados por homoplasias interpretáveis como fruto de interações e transmissões horizontais do conhecimento. Todas as considerações apresentadas sugerem que as inferências aqui idealizadas não fornecem explicações únicas sobre a dispersão de traços culturais pelo território paulista, mas o método da cladística surge como um convite à reflexão sobre os diferentes modos de transmissão cultural a partir da variabilidade nos traços culturais analisados.

Por fim, cabe ressaltar que as inferências apresentadas neste estudo devem ser compreendidas como hipóteses exploratórias, condicionadas pelo recorte espacial, temporal e metodológico adotado. A cladística, tal como empregada aqui, não oferece explicações únicas ou definitivas sobre a dispersão dos grupos Tupiguarani, mas constitui uma ferramenta indispensável para organizar a variabilidade artefactual e levantar questões sobre os modos de transmissão cultural. Investigações futuras, com ampliação da amostra espacial, refinamento cronológico e integração sistemática de dados linguísticos e históricos, poderão aprofundar e testar as hipóteses aqui delineadas.

AGRADECIMENTOS

Agradecemos ao Museu de Arqueologia e Etnologia da Universidade de São Paulo, ao Centro Regional de Arqueologia Ambiental Mário Neme e ao Museu Histórico da Colonização Pereira Barreto. GCP À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP, GCP Bolsa de Pós-doutorado no País, Processo 2021/04562-0 e Bolsa BEPE, Processo 2023/03591-1. AGMA, Projeto Temático Processo 2019/17743-2. MO, Jovem Pesquisador Processo 2018/23282-5, SPRINT 2020/00900-5 e Bolsa produtividade 308856/2022-8; EC: Education New Zealand Grant, UoA Project 3721948. “As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do(s) autor(es) e não necessariamente refletem a visão da FAPESP”.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBUQUERQUE, Marcos. Recipientes cerâmicos de grupos Tupi, no nordeste brasileiro. In: PROUS, André; LIMA, Tania (eds.). *Os ceramistas Tupiguarani 1*. Brasília, DF: Iphan, 2008. p. 55-78.
- ARAUJO, Astolfo G. M. *Teoria e método em arqueologia regional: um estudo de caso no Alto Paranapanema, estado de São Paulo*. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2001.
- ARAUJO, Astolfo G. M. On vastness and variability: cultural transmission, historicity, and the Paleoindian record in Eastern South America. *Anais da Academia Brasileira de Ciências*, v. 87, n. 2, p. 1239-1258, 2015.
- ARAUJO, Astolfo G. M. *Por uma arqueologia cética: Ontologia, epistemologia, teoria e prática da mais interdisciplinar das disciplinas*. Curitiba: Appris, 2019.
- ARAUJO, Astolfo G. M.; OKUMURA, Mercedes. Cultural taxonomies in eastern South America: Historical review and perspectives. *Journal of Paleolithic Archaeology*, v. 4, n. 28, 2021.
- ATKINSON, Quentin; GRAY, Russell D. Curious Parallels and Curious Connections-Phylogenetic Thinking in Biology and Historical Linguistics. *Systematic Biology*, v. 54, p. 513-526, 2005.

- BASAVA, Kiran; ZHANG, Hanzhi; MACE, Ruth. A phylogenetic analysis of revolution and afterlife beliefs. *Nature Human Behaviour*, v. 5, n. 5, p. 604-611, 2021.
- BONOMO, Mariano *et al.* Ancient Tupinambá and Guaraní large-scale movements in the South American lowlands. *Antiquity*, p. 1-18, 2025.
- BOYD, Robert *et al.* Are cultural phylogenies possible? In: WEINGART, Peter *et al.* (ed.). *Human by Nature: Between biology and the social sciences*. Mahwah (US): Lawrence Erlbaum Associates Publishers, 1997. p. 355-386.
- BROCHADO, José P. Desarrollo de la Tradición Cerámica Tupiguaraní (A.D. 500-1800). In: Simpósio Nacional de Estudos Missionários, 1., 1971. *Anais [...]*. Santa Rosa, 1981. p. 76-156.
- BROCHADO, José P. *An ecological model of the spread of pottery and agriculture into Eastern South America*. Tese (Doutorado) – University of Illinois at Urbana-Champaign, Champaign (US), 1984.
- BROCHADO, José P. *et al.* *Arqueologia Brasileira em 1968. Um relatório preliminar sobre o Programa Nacional de Pesquisas Arqueológicas*. Belém: Museu Paraense Emílio Goeldi, 1969.
- CHANDY, Kaniyantra M.; LO, Tachen. The capacitated Minimum Spanning Tree. *Networks*, v. 3, n. 2, p. 173-181, 1973.
- CHMYZ, Igor *et al.* Terminologia arqueológica brasileira para a cerâmica. *Cadernos de Arqueologia*, v. 1, n. 1, p. 119-148, 1976.
- CHMYZ, Igor. *Pesquisas paleoetnográficas efetuadas no Vale do rio Paranapanema, Paraná-São Paulo*. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 1973.
- COCHRANE Ethan E. *Explaining cultural diversity in ancient Fiji: the transmission of ceramic variability*. Tese (Doutorado) – Universidade do Hawai, Hawai (US), 2004.
- COCHRANE, Ethan E. Phylogenetic Analysis of Polynesian Ritual Architecture Suggests Extensive Cultural Sharing and Innovation. *The Journal of the Polynesian Society*, v. 124, n. 1, p. 7-46, 2015.
- COLLARD, Mark; SHENNAN, Stephen J.; TEHRANI, Jamshid J. Branching, blending, and the evolution of cultural similarities and differences among human populations. *Evolution and Human Behavior*, v. 27, n. 3, p. 169-184. 2006.
- CORRÊA, Ângelo A. *Pindorama de Mboia e Ikararé: continuidade e mudança na trajetória das populações Tupi*. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2014.
- CREMA, Enrico R.; KERIG, Tim. SHENNAN, Stephen. Culture, space, and metapopulation: a simulation-based study for evaluating signals of blending and branching. *Journal of Archaeological Science*, v. 43, p. 289- 298, 2014.
- DUNNELL, Robert C. Style and function: a fundamental dichotomy. *American Antiquity*, v. 43, n. 2, p. 192- 202. 1978.
- GARCIA-RIVERO, Daniel; O'BRIEN, Michael J. Phylogenetic Analysis Shows That Neolithic Slate Plaques from the Southwestern Iberian Peninsula Are Not Genealogical Recording Systems. *Plos One*, v. 9, n. 2, 2014.
- GERARDI, Fabrício F. *et al.* Lexical phylogenetics of the Tupí-Guaraní family: Language, archaeology, and the problem of chronology. *Plos one*, v. 18, n. 6, 2023.

- GRYGORASH, Oleksandr *et al.* In: IEEE INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELLIGENCE (ICTAI'06), 18., 2006, Arlington (US). *Anais [...]*. Arlington: IEEE Computer Society, 2006. p. 73-81.
- HOLDEN, Clare; MACE, Ruth. Phylogenetic analysis of the evolution of lactose digestion in adults. *Human Biology*, v. 69, n. 5, p. 605-628, 1997.
- HOLLAND, Steven M. Non-metric Multidimensional Scaln (MDS). *Department of Geology, University of Georgia*, p. 30602-2501, 2008.
- LIPO, Carl P. Community Structures among Late Mississippian Populations of the Central Mississippi River Valley. In: HUNT, Terry L.; LIPO, Carl P.; STERLING, Sarah L. (ed.). *Posing Questions for a Scientific Archaeology, Scientific Archaeology for the Third Millenium*. Westport (US): Bergin and Garvey, 2001. p. 175-216.
- LIPO, Carl P. The Resolution of Cultural Phylogenies Using Graphs. In: LIPO, Carl P. *et al.* (ed.). *Mapping Our Ancestors: Phylogenetic Methods in Anthropology and Prehistory*. New York (US): Aldine de Gruyter, 2005. p. 89-108.
- MACE, Ruth; HOLDEN, Clare J. A phylogenetic approach to cultural evolution. *Trends in ecology & evolution*, v. 20, n. 3, p. 116-121. 2005.
- MAEZUMI, S. Yoshi *et al.* The legacy of 4,500 years of polyculture agroforestry in the eastern Amazon. *Nature plants*, v. 4, n. 8, p. 540-547. 2018.
- MARANCA, Silvia; SILVA, Andrea L.; SCABELLO, Ana M. Projeto Oeste Paulista de Arqueologia do Baixo e Médio Vale do Rio Tietê: síntese dos trabalhos realizados. *Rev. Do Museu de Arqueologia e Etnologia*, n. 4, p. 223-2269, 1994.
- MATOS, Francisco A. S. *Entre semelhanças gráficas e ambientais: as recorrências das representações antropomórficas pintadas pré-históricas entre as regiões do Cariri Ocidental-PB, Parque Nacional do Catimbau-PE e Seridó Oriental-RN*. Tese (Doutorado) – Universidade Federal de Pernambuco, Recife, 2019.
- MEGGERS, Betty J. *Considerações gerais. Programa Nacional de Pesquisas Arqueológicas resultados preliminares do primeiro ano 1965-1966*. Belém: Museu Paraense Emílio Goeldi, 1967.
- MEGGERS, Betty J. Climatic Oscillation as a Factor in the Prehistory of Amazonia. *American. Antiquity*, v. 44, n. 2, p. 252-66, 1979.
- MEGGERS, Betty J.; EVANS, Clifford. *Como interpretar a linguagem da cerâmica. Manual para a Arqueologia*. Washington, DC: Smithsonian Institution, 1970.
- MEGGERS, Betty J. *et al.* Implications of archaeological distributions in Amazonia. *Separata dos Anais da Academia Brasileira de Ciências*, p. 275-94, 1988.
- MELLO, Antonio A. S.; KNEIP, Andreas. Diálogo Linguística – Arqueologia: origem e dispersão dos povos Tupi-Guarani. In: Congresso da Sociedade de Arqueologia Brasileira, 13., 2015. *Anais [...]*. Campo Grande, 2005.
- MÉTRAUX, Alfred. *Migrations historiques des Tupi-guarani*. Paris: Librairie Orientale et Américine, 1927.
- MORAES, Camila A. *Arqueologia tupi no Nordeste de São Paulo: um estudo de variabilidade artefactual*. Dissertação (Mestrado) – Universidade de São Paulo, São Paulo, 2007.
- MORAIS, José L. Aerofotoarqueologia: um estudo de caso no Projeto Paranapanema. *Revista do Museu Paulista*, v. 30, p. 99-114, 1985.

- MORAIS, José L. *Projeto Paranapanema. Arqueologia da Paisagem: Cenas do Paranapanema Paulista (da Pré-História ao Ciclo do Café). Inventário*. São Paulo, 2001.
- NOELLI, Francisco S. As hipóteses sobre o centro de origem e rotas de expansão dos Tupi. *Revista de Antropologia*, v. 39, n. 2, p. 7-53, 1996.
- NOELLI, Francisco S. *The Tupi Expansion, in The Handbook of South American Archaeology*. New York: Springer, 2008.
- NOELLI, Francisco S. *Não há colonialismo sem tekoóba: uma arqueologia das relações e da materialidade entre Tupiniquim e Portugueses na Capitania de São Vicente, Brasil (1502-1700)*. Tese (Doutorado) – Universidade de Lisboa, Lisboa (PT), 2025.
- O'BRIEN, Michael J.; LYMAN, R. Lee. *Cladistics and archaeology*. Salt Lake City: The University of Utah Press, 2003.
- O'BRIEN, Michael J. *et al.* Cultural Cladistics and the early prehistory of North America. In: PONTAROTTI, Pierre (ed.). *Evolutionary biology: mechanisms and trends*. New York: Springer, 2012. p. 23-42.
- OPIE, Christopher *et al.* Phylogenetic reconstruction of Bantu kinship challenges main sequence theory of human social evolution. *Proceedings of the National Academy of Sciences*, v. 111, n. 49, p. 17414-17419, 2014.
- PEREZ, Glauco C. *Relatório de análise do material do Sítio Arqueológico Prema Tupi – Rio Claro/SP*. Não Publicado. 2012.
- PEREZ, Glauco C. *Arqueologia paulista e o marcador cerâmico como delimitador de fronteira étnica: um estudo das regiões sul e oeste do estado de São Paulo*. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2018.
- PEREZ, Glauco C. *Relatório parcial para Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, 2022. O estabelecimento de redes de Transmissão Cultural entre os grupos Tupiguarani no Estado de São Paulo (Processo FAPESP - 2021/04562-0)*. Relatório Parcial. Laboratório Interdisciplinar de Pesquisas em Evolução, Cultura e Meio Ambiente – LEVOC. Museu de Arqueologia e Etnologia., USP. FAPESP. São Paulo. 2022. Disponível em: <https://doi.org/10.5281/zenodo.14731875>. Acesso em: 13 ago. 2026.
- PEREZ, Glauco C. *Relatório Parcial para a Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP – Projeto O estabelecimento de redes de Transmissão Cultural entre os grupos Tupiguarani no estado de São Paulo (Processo FAPESP – 2021/04562-0)*. Relatório apresentado à FAPESP em detrimento do pedido de prorrogação de bolsa. 2023. Disponível em: <https://doi.org/10.5281/zenodo.14732070>. Acesso em: 13 ago. 2026.
- PEREZ, Glauco C. *Relatório Parcial para a Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP – Projeto O estabelecimento de redes de Transmissão Cultural entre os grupos Tupiguarani no estado de São Paulo (Processo FAPESP – 2021/04562-0)*. Relatório apresentado à FAPESP em detrimento do pedido de prorrogação de bolsa. 2024. Disponível em: <https://doi.org/10.5281/zenodo.14732191>. Acesso em: 13 ago. 2026.
- PEREZ, Glauco C. *Matriz PAUP para estudos de Cladística em sítios Arqueológicos do estado de São Paulo - dados binários e de frequência*. Associado ao Projeto O estabelecimento de redes de Transmissão Cultural entre os grupos Tupiguarani no estado de São Paulo (Processo FAPESP – 2021/04562-0). 2025. Disponível em: <https://doi.org/10.5281/zenodo.14732328>. Acesso em: 13 ago. 2026.
- PEREZ, Glauco C.; ARAUJO, Astolfo G. M.; OKUMURA, Mercedes. Novas observações sobre a distribuição dos grupos ceramistas pré-históricas no oeste e sul do Estado de São Paulo – Brasil.

- Revista Noctua*, p. 87-157, 2020. Disponível em: http://fundacaoparanabuc.org.br/arquivo/c6cbb_Artigo_6_AstolfoAraujo_Noctua_2020_1.pdf. Acesso em: 13 ago. 2026.
- PROUS, André. *Arqueologia Brasileira*. Brasília, DF: Editora UnB, 1992.
- RABINOWITZ, George B. Na introduction to nonmetric Multidimensional Scaling. *American Journal of Political Science*, v. 19, n. 2, p. 343-390, 1975.
- RIEDE, Felix; HOGGARD, Christian; SHENNAN, Stephen. Reconciling material cultures in archaeology with genetic data requires robust cultural evolutionary taxonomies. *Palgrave Communications*, v. 5, n. 1. 2019.
- SCATAMACCHIA, Maria Criscina. M. *Tentativa de caracterização da tradição Tupiguarani*. Dissertação (Mestrado) – Universidade de São Paulo, São Paulo, 1981.
- SCATAMACCHIA, Maria C. M.; UCHÔA, Dorah P. O contato euro-indígena visto através de sítios arqueológicos do Estado de São Paulo. *Revista de Arqueologia*, v. 7, n. 1, p. 153-173, 1993.
- SHENNAN, Stephen J.; BENTLEY, Alexander. Style, Interaction, and Demography among the Earliest Farmers of Central Europe. In: O'BRIEN, Michal J. (ed). *Cultural Transmission and Archaeology. Issues and Case Studies*. Washington, D.C.: Society for American Archaeology Press, 2008. p. 164-177.
- SOUZA, Jonas G.; MATEOS, Jonas A; MADELLA, Marco. Archaeological expansions in tropical South America during the late Holocene: assessing the role of demic diffusion. *PLoS One*, v. 15, n. 4. 2020.
- SOUZA, J. G.; NOELLI, Francisco S.; MADELLA, Marco. Reassessing the role of climate change in the Tupi expansion (South America, 5000–500 BP). *Journal of the Royal Society Interface*, v. 18, n. 183, 2021.
- SPENCER, Matthew *et al.* Analyzing the Order of Items in Manuscripts of The Canterbury Tales. *Computers and the Humanities*, v. 37, n. 1, p. 97-109. 2003.
- SWOFFORD, David L. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Sunderland: Sinauer Associates, 2001. 1 CD-ROM.
- TEHRANI, Jamshid J. The Phylogeny of Little Red Riding Hood. *PLoS One*, v. 8, n. 11, 2013.
- TEHRANI, Jamshid J.; COLLARD, Mark. Investigating cultural Evolution through biological phylogenetic analyses of Turken textiles. *Journal of Anthropological Archaeology*, v. 21, n. 4, p. 443-463, 2002.
- TËMKIN, Ilya; ELDREDGE, Niles. Phylogenetics and material cultural Evolution. *Current Anthropology*, v. 48, n. 1, 2007.
- TOLSTOY, Paul. Barkcloth, Polynesia and cladistics: an update. *The Journal of the Polynesian Society*, v. 117, n. 1, p. 15-57, 2008.
- VALLE, Raoni B. M. *Gravuras pré-históricas da área arqueológica do Seridó Potiguar/Paraibano: um estudo técnico e cenográfico*. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Recife, 2003.
- VALLE, Raoni B. M. Gravuras rupestres do Seridó Potiguar e Paraibano, um estudo técnico e cenográfico, novos aportes. In: SIMPÓSIO DE POVOAMENTO PRÉ-HISTÓRICO DAS AMÉRICAS, 2., 2006. *Anais [...]*. São Raimundo Nonato, 2006.

- VALLE, Raoni B. M. *Mentes graníticas e mentes areníticas: fronteiras geo-cognitiva nas gravuras rupestres do Baixo Rio Negro, Amazônia Setentrional*. Tese (Doutorado em Arqueologia) – Universidade de São Paulo, São Paulo, 2012.
- WENINGER, Bernhard; JÖRIS, Olaf. A ^{14}C age calibration curve for the last 60 ka: the Greenland-Hulu U/Th timescale and its impact on understanding the Middle to Upper Paleolithic transition in Western Eurasia. *Journal of human evolution*, v. 55, n. 5, p. 772-781, 2008.
- WILLEY, Gordon R. Horizon styles and pottery traditions in Peruvian archaeology. *American Antiquity*, v. 10, n. 1, p. 49–56, 1945.
- WILLEY, Gordon R.; PHILLIPS, Philip. *Method and Theory in American Archaeology*. Tuscaloosa (US): University of Alabama Press, 1958.

Material Suplementar 1. Código de software R

```

##Sao Paulo, Brazil, Ceramic Phylogenetic Analyses
##Ethan Cochrane and Glauco Perez, 29/1/2025

#Users will need to set their own working directory with setwd()
setwd("C:/Users/ecoc205/OneDrive - The University of Auckland/Current
Research/Brazil/Glauco2023visit/SP Ceramics Article/R_cladistics")
getwd()
rm(list=ls())

# Libraries
library(readxl)
library(ape)
library(TreeTools)
library(phangorn)
library(TreeSearch)

#Import data from Excel
SPceramic_all_raw <- as.matrix (read_excel("SPceramic_all_29.1.25.xlsx",
                                         sheet = 1, #first sheet of file
                                         range = "B1:TD15", #all character data & column

heads
                                         col_types = "text"))

#Reading taxon names
sites <- unlist(read_excel("SPceramic_all_29.1.25.xlsx", sheet = 1, range =
"A1:A15")) #taxon names, first cell is column name, e.g., sites
rownames(SPceramic_all_raw) <- sites

#Convert raw data matrix to PhyDat file
SPceramic_all <- MatrixToPhyDat(SPceramic_all_raw)

##Tree Analyses All Data
#Maximum Parsimony using TBR rearrangements
MPtree <- MaximizeParsimony(SPceramic_all, ratchIter = 100,tbrIter = 100)
length(MPtree) #outputs number of equal length trees in MPtree
MPtree_Ary_Carneiro_root <- (RootTree(MPtree, "Ary_Carneiro"))
con_MPtree <- Consensus(MPtree_Ary_Carneiro_root, p = 0.5) #Do not have to
create consensus tree if n of trees=1
con_MPtree_lengths <- acctran(con_MPtree, SPceramic_all) #Calculates the number
of character changes on each branch
plot.phylo(con_MPtree_lengths, type = "phylogram")
add.scale.bar(x = 2, y = 5, length = 10, cex = 0.8)

#Tree Stats
CI(con_MPtree, SPceramic_all)
RI(con_MPtree, SPceramic_all)

#Neighbor Joining algorithm
NJtree <- NJTree(SPceramic_all, edgeLengths = TRUE)
length(NJtree)
NJtree_Ary_Carneiro_root <- (RootTree(NJtree, "Ary_Carneiro"))
con_NJtree <- Consensus(NJtree_Ary_Carneiro_root, p = 0.5)
plot(con_NJtree)

CI(con_NJtree, SPceramic_all)
RI(con_NJtree, SPceramic_all)

#Comparing NJ and MP trees
parsimony(con_NJtree, SPceramic_all)
parsimony(con_MPtree, SPceramic_all)

```