Check for updates

RESEARCH ARTICLE

REVISED A phylogenetic classification of the Je language family

[version 3; peer review: 3 approved, 4 approved with reservations]

Fabrício Ferraz Gerardi[1], Tim Wientzek[1], Ivan Roksandic[2], Jonas Gregorio de Souza[3], Fernando Orphão de Carvalho[4]

[1]University of Tübingen Department of Linguistics, Tübingen, Baden-Württemberg, 72074, Germany
[2]The University of Winnipeg Department of Anthropology, Winnipeg, Manitoba, Canada
[3]Pompeu Fabra University Department of Humanities, Barcelona, Catalonia, Spain
[4]Federal University of Rio de Janeiro National Museum, Rio de Janeiro, State of Rio de Janeiro, Brazil

v3 **First published:** 30 Jan 2025, **5**:29
https://doi.org/10.12688/openreseurope.19346.1
Second version: 27 Mar 2025, **5**:29
https://doi.org/10.12688/openreseurope.19346.2
Latest published: 19 May 2025, **5**:29
https://doi.org/10.12688/openreseurope.19346.3

Abstract

Introduction

This study investigates the Je language family and Macro-Je phylum, addressing a significant gap in previous research by applying quantitative methods to its classification.

Dataset

The study compares a subset of 303 concepts from 14 languages, primarily sourced from Swadesh lists and culturally relevant terms, providing a robust foundation for phylogenetic analysis.

Methods

Bayesian phylogenetic inference and NeighborNet methods were employed to analyze the dataset. These approaches enabled the reconstruction of evolutionary relationships within the Je family, facilitating the identification of language divergence patterns and their historical dynamics.

Results

The analysis reveals well-supported Northern, Central, and Southern subgroups within the Je family, demonstrating clear geographical clustering. The phylogenetic tree aligns with existing hypotheses while offering new insights into the family's structure.

Open Peer Review

Approval Status ✓ ? ✓ ? ? ?

	1	2	3	4	5	6
version 3 (revision) 19 May 2025	✓ view					? view
version 2 (revision) 27 Mar 2025	✓ view	? view	✓ view	? view	? view	
version 1 30 Jan 2025	? view					

1. **Simon J Greenhill**, University of Auckland, Auckland, New Zealand
2. **Joseph Salmons**, University of Wisconsin, Madison, USA
3. **Robin J. Ryder**, Université Paris Dauphine and Crest INSEE, Paris, France
4. **Natalia Chousou-Polydouri**, Institute for Mediterranean Studies, Foundation for Research and Technology—Hellas, Rethymno, Greece
5. **Benedict King**, Max Planck Institute for Evolutionary Anthropology, Leipzig,, Germany

Discussion

The findings were contextualized within pre-Columbian archaeological frameworks, drawing parallels between linguistic divergence and material culture. These connections support the hypothesis that the Je language family's development aligns with distinct cultural and geographical distributions observed in archaeological records.

Conclusion and Future Directions

This study affirms the genetic coherence of the Je family and highlights opportunities for future research, including the incorporation of non-Je languages in the Macro-Je phylum and expanded datasets to refine the understanding of this diverse linguistic group.

Plain Language Summary
This study examines the Je language family using quantitative methods to improve its genetic classification. Researchers analyzed 516 words from 14 languages, mostly from Swadesh lists and culturally important terms, to trace their historical relationships. By applying Bayesian phylogenetic inference and NeighborNet methods, they identified patterns of language divergence and subgroup connections.

The findings align linguistic changes with archaeological evidence, suggesting that the evolution of the Je family reflects historical cultural and geographic patterns. The study confirms the existence of Northern, Central, and Southern subgroups within the Je family, reinforcing previous research while offering new insights. The results also support the genetic unity of the Je family and highlight future research directions, such as including more languages and expanding the dataset for a deeper understanding.

Keywords
phylogenetics, historical linguistics, macro-je

This article is included in the European Research Council (ERC) gateway.

This article is included in the Horizon 2020 gateway.

6. **Sandra Auderset**, University of Bern, Bern, Switzerland
Andrés Pablo Salanova, University of Ottawa, Ottawa, Canada

Please see the article web page for additional reviewers.

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Fabrício Ferraz Gerardi (fabricio.gerardi@uni-tuebingen.de)

Author roles: Ferraz Gerardi F: Conceptualization, Data Curation, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Wientzek T**: Conceptualization, Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Roksandic I**: Writing – Original Draft Preparation, Writing – Review & Editing; **Gregorio de Souza J**: Writing – Original Draft Preparation; **Orphão de Carvalho F**: Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement number: 834050 (Grant agreement No. 834050) (Cross-Linguistic statistical inference using hierarchical Bayesian models)
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Ferraz Gerardi F *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ferraz Gerardi F, Wientzek T, Roksandic I *et al.* **A phylogenetic classification of the Je language family [version 3; peer review: 3 approved, 4 approved with reservations]** Open Research Europe 2025, **5**:29 https://doi.org/10.12688/openreseurope.19346.3

First published: 30 Jan 2025, **5**:29 https://doi.org/10.12688/openreseurope.19346.1

REVISED **Amendments from Version 2**

Minor revisions made to improve clarity and correct typographical errors. Table 1 has been extended to include all doculects mentioned in the text. No changes to the data, analysis, or conclusions have been made.

Any further responses from the reviewers can be found at the end of the article

1 Introduction

This paper represents a first effort to apply phylogenetic classification to the Je language family, addressing a significant gap in Macro-Je studies, as quantitative approaches have been notably absent in previous research. Phylogenetic methods, which are well-established in linguistics, have been widely used to elucidate the classification and dating of various language families, such as Indo-European (Chang *et al.*, 2015), Dravidian (Kolipakam *et al.*, 2018), Alor Pantar (Kaiping & Klamer, 2019), Mixtecan (Auderset *et al.*, 2023) and Tupi-Guarani (Ferraz Gerardi *et al.*, 2023). These methods are particularly suited for testing hypotheses related to prehistory, linguistic evolution, and evolutionary dynamics (Greenhill *et al.*, 2020).

Von den Steinen (1886) introduced the term "Je" to designate this language family based on lexical and cultural comparison (see also Ehrenreich, 1891). He derived the name "Je" from a common collective suffix found in many of the group's names.

The Je language family today comprises about fifteen languages (Hammarström *et al.*, 2024), some with multiple dialects (see, e.g., Gildea & de Castro Alves, 2020: 50), all spoken in Brazil and geographically distributed broadly. Many Je languages have become extinct, and, for most, only limited data is available (Gildea & de Castro Alves, 2010; Nikulin, 2020; Nimuendajú, 2017). Davis (1966), based on phonemic correspondences, proposed internal relations among Apinajé, Timbira, Xavante, Kaingang, and Suyá. This was an important step, as it demonstrated the existence of geographical clusters within the Je family: Northern (Timbira, Apinajé, and Suyá), Central (Xavante), and Southern (Kaingang).

The proposal for the Macro-Je phylum (Mason, 1950; Schmidt, 1926), consisting of the Je family and several smaller families, was supported by studies such as Rivet (1924) (see also Loukotka, 1932; Loukotka, 1937; Loukotka, 1939; Loukotka 1942[1944]), which provided further evidence for this phylum, and included languages such as Ofayé (see also Gudschinsky, 1971), Koropó, Pataxó, Krenak (Borum), Arikapu, Djeoromitxí and Kamakã (see also Mason, 1950; Ribeiro & van der Voort, 2010; Rodrigues, 2009). Boswood (1973) presented lexical evidence supporting the inclusion of Rikbaktsá in this family. Rodrigues (1986) expanded the internal relations within Macro-Je and (Rodrigues, 1999) added data from extinct languages.

The inclusion of Bororo in the Macro-Je phylum, suggested by Guérios (1939), is no longer accepted (Nikulin, 2020). Rodrigues (1986) included Karirí and Guató as small families in the Macro-Je phylum, but considered Pataxó a member of the Maxakalí family. Greenberg (1987) classified Guató, Maxakalí, and Pataxó as part of Macro-Je, adding Chiquitano, Otí, and Jabutí families. Conversely, Kaufman (1990) excluded Otí, Jabutí, and Karirí from Macro-Je but retained Chiquitano (see Nikulin, 2020). For other proposals, see de Castro Alves (2010) and Ribeiro & van der Voort (2010). Nikulin (2020) grouped Maxakalían and Krenakan (perhaps also Kamakanan) in a branch he calls "Transfranciscanan". He also provided thorough evidence for a larger Macro-Je phylum, further suggesting a more distant relationship with other families such as Chiquitano, which he considers to have a common ancestor with Proto-Macro-Je.

Table 1 lists 27 languages considered potential members of the Macro-Je phylum, for which there is some documentation, although limited in some cases. These languages are shown in Figure 1.

Ribeiro (2012: 266) noted the difficulty in conducting a comparative study of Macro-Je due to the monosyllabic nature of many parallels presented, and the lack of systematic sound correspondences in their roots (see also Ribeiro & van der Voort, 2010). Many of the lexical parallels found among the linguistic families attributed to this stock and referred to as 'cognates' may be borrowings, given the high probability of interactions that many populations have had since precolonial times with various other populations (see the dates provided in Section 5). The hypothesis that Macro-Je could result from long-term linguistic contact via ethnogenesis (Dixon & Aikhenvald, 1999) cannot yet be discarded. In this case, similarities found between Je family and some other families / languages, such as Borum and Maxakali, but also Purí and Guató, would be instances of horizontal transfer rather than genetic inheritance (Ramirez *et al.*, 2015).

Evidence from historical-comparative linguistics supports the genetic identity of a 'nuclear' Macro-Je linguistic stock, which includes the Je family along with the Chiquitano (Adelaar, 2008), Jabutí (Ribeiro & van der Voort, 2010), Ofayé and Karajá (Ribeiro, 2011; Ribeiro, 2012) families. Conversely, the evidence for including languages such as Bororo, Yatê, Purí, Guató, Karirí, and Otí in the Macro-Je stock is extremely weak.

Speakers of Macro-Je languages inhabit a vast region that extends longitudinally from the Atlantic coast to the Chiquitano Dry Forest and the Guaporé River, and latitudinally from the lower Tocantins to the northern part of the current state of Rio Grande do Sul (see Figure 1). Currently, they number approximately 90,000 individuals, with the most spoken languages being Kaingáng (around 30,000 speakers), Xavante (around 25,000), and Mebengokre (around 13,500).

Table 1. Likely members of Macro-Je.

Doculect	ISO 639	Glottocode	Family
Acroá	acs	acro1239	Central Je
Apinaye	apn	apin1244	Northern Je
Arikapu	ark	arik1265	Jabutí
Borum	kqq	kren1239	Borum
Canela-Kraho	ram	cane1242	Northern Je
Chiquitano	cax	chiq1253	Chiquitano
Djeoromitxí	jbt	djeo1235	Jabutí
Gaviao-Pyhcopji	xri	krik1239	Northern Je
Ingain		inga1254	Southern Je
Kaingang	kgp	kain1272	Southern Je
Karajá	kpj	kara1500	Karajá
Mebengokre (Kayapó)		kaya1330	Northern Je
Koropó	xxr	coro1248	Maxakalí
Kisedje	suy	suya1243	Northern Je
Krenák		kren1240	Borum
Pataxó	pth	pata1261	Maxakalí
Ofayé	opy	ofay1240	Ofayé
Parkateje		timb1254	Northern Je
Panara	kre	pana1307	Northern Je
Rikbakstá	rkb	rikb1245	Rikbakstá
Suyá	suy	suya1243	Northern Je
Tapayuna		beic1238	Northern Je
Timbira		timb1253	Northern Je
Xacriabá	xkrx	xacr1238	Central Je
Xavante	xav	xava1240	Central Je
Xerente	xer	xer1240	Central Je
Xokleng	xok	xokl1240	Southern Je

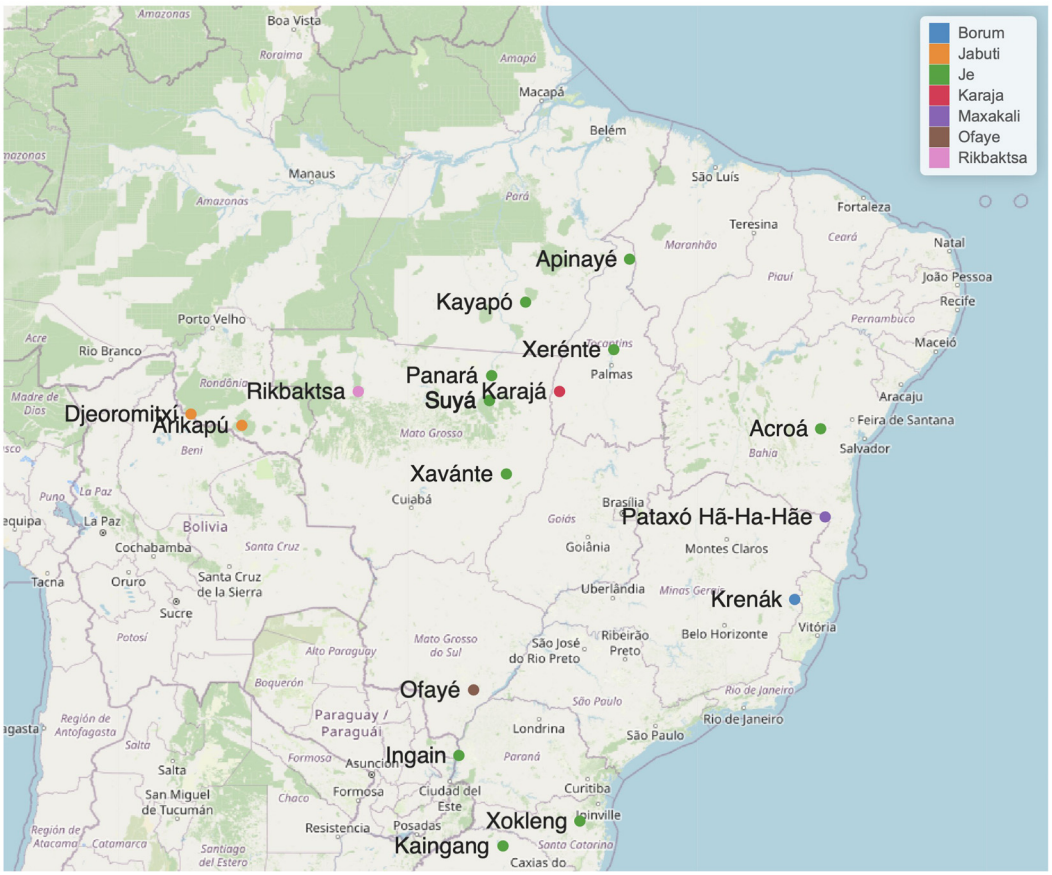


Figure 1. Some Macro-Je languages. The geo-location is based on Hammarström *et al.* (2024).

Typologically, Je languages are characterized phonologically by a rich vowel inventory that includes nasal vowels, which often trigger nasal enhancement phenomena in adjacent segments (de Carvalho & Damulakis, 2015; Pickering, 2010). The languages have a limited number of possible onsets, typically allowing only sequences of a peripheral obstruent (labial or velar) followed by a rhotic (Nikulin, 2020). Je languages tend to be head-final SOV as the most common order. Many Je languages display a rare type of split alignment that has not yet been sufficiently studied. This alignment pattern has been called nominative-absolutive (Gildea & de Castro Alves, 2010; Gildea & de Castro Alves, 2020). Another typical Je feature is the singular-(dual)-plural stem alternation for certain semantic verb types (see Inman & Vuillermet, n.d.), and a somewhat complex system of person marking that interacts with agreement, since most languages are of the dependent marking type. For a more detailed typological profile of the Macro-Je languages, we refer the reader to the works of Salanova (n.d.), Storto (2019: 138–140) (only phonology), Rodrigues (1999) (partially outdated) and Wiesemann (1986).

This paper is structured as follows: Section 2 presents the data used, followed by a presentation of the methods employed in Section 3. The results are discussed in Section 4. The Section on archaeology (5) places Je languages in light of archaeological dates and relates this to the results of our analyses. The paper concludes with final remarks in Section 7.

2 Data

The data used in this study is drawn from Gerardi and Tresoldi 2024, which builds on Nikulin (2020). It consists of 14 languages (given in Table 2) and 516 concepts. The concepts include most Swadesh list items (Swadesh, 1971) along with many culturally relevant concepts related to fields like agriculture, vegetation and fauna. Given the exploratory nature of this study, cognacy judgments were automatically generated using LexStat (List, 2014), a state-of-the-art method implemented in the Lingpy Python package (List & Forkel, 2021). While Rama & List (2019) demonstrated that automated cognacy judgments can be effectively used for preliminary analyses and often produce reliable results, a small subset—approximately 10% of the data—was manually annotated to assess the quality of the automatic annotation. This assessment was conducted by calculating the B-Cubed F-Score (see List (2014) for details) between the automatically and manually assigned cognate classes, resulting in a score of 0.947. This result indicates a high level of accuracy in the automated judgments.

For the analysis described below, we restricted the dataset to concepts that had corresponding forms in at least 50% of the languages, reducing the set to 303 concepts. Table 3 reveals that approximately 38% of the forms across all languages and remaining concepts are missing. Notably, 210 out of the 622 cognate sets in the dataset are singletons—that is, word forms classified as non-cognate with any other word form. However, 72 of those singletons are found in Arikapu, a Jabutí language and the only non-Je language in the sample (Ribeiro & van der Voort, 2010). Conversely, five concepts with forms available across all sampled languages are classified as consisting of a single cognate set: ”FOOT”, ”HEAD”, ”LIVER”, ”MEAT” and ”SLEEP”. The full list of concepts, as well as the results of the automatic cognate detection, the subset used to calculate the F-Score and the results of the analyses discussed below are available online (Ferraz Gerardi *et al.*, 2025).

3 Method

In recent years, the field of computational historical linguistics has experienced substantial growth, with phylogenetic methods adapted from bioinformatics gaining prominence for language classification and dating (see Greenhill (2023) for a

Table 2. Languages used in this study. Glottocodes, latitude and longitude taken from Hammarström *et al.* (2024).

Doculect	ISO 639	Glottocode	Latitude	Longitude
Apinaye	apn	apin1244	-6.11	-47.63
Arikapu	ark	arik1265	-12.49	-62.73
Canela-Kraho	ram	cane1242	-6.11	-45.13
Gaviao-Pyhcopji	xri	krik1239	-5.94	-46.75
Ingain	-	inga1254	-24.62	-54.25
Kaingang	kgp	kain1272	-27.77	-52.54
Kisedje	suy	suya1243	-11.52	-53.07
Mebengokre	txu	kaya1330	-7.77	-51.67
Panara	kre	pana1307	-10.58	-53
Parkateje	-	timb1254	-	-
Tapayuna	-	beic1238	-	-
Xavante	xav	xava1240	-14.30	-52.44
Xerente	xer	xer1240	-9.59	-48.26
Xokleng	xok	xokl1240	-26.92	-49.59

Table 3. Coverage and cognate statistics.

Concepts	303
Languages	14
Words	2638
Coverage	0.622
Cognate sets	622
Singletons	210

comprehensive list of studies). This study seeks to classify the languages in our sample (see Table 2) using Bayesian phylogenetic inference. For the analyses, the data was exported to the NEXUS format (Maddison *et al*., 1997). In this format, the data is represented as a matrix where each column corresponds to a cognate class, and each row represents a language. The resulting binary representation uses '1' to indicate the presence of a cognate class in a language, '0' for its absence, and a special character for missing data in a concept-language pair. Following Hoffmann *et al*. (2021), ascertainment correction was applied.

3.1 NeighborNet

To quantify and visualize conflicting signals within the dataset, we employed δscores and Q-residuals—indices that measure the tree-likeness of the data (Bryant & Moulton, 2004; Gray *et al*., 2010). δ-scores range from 0 to 1, with higher values indicating increased conflict, while Q-residuals, adjusted for scaling effects, provide a more direct measure of divergence from a strict tree structure (Gray *et al*., 2010). Lower Q-residual scores suggest a closer fit to a strict tree. Averaging both scores over the tips yields a measure of overall tree-likeness. These metrics, along with the NeighborNet visualization, were generated using SplitsTree4 (Huson & Bryant, 2006). Although direct comparison of δ-scores and Q-residuals across different datasets and the evaluation of their significance may not be feasible, low values for these metrics validate the application of methods such as Bayesian phylogenetics, which assume a strict tree structure (Gray *et al*., 2010). The plot of δ-scores against Q-residuals for each language is given in Figure 2 and the NeighbourNet is given in Figure 3.

3.2 Bayesian analyses

Phylogenetic reconstruction was performed using BEAST2 (Bouckaert *et al*., 2014), testing various binary covarion models (Huelsenbeck, 2002). Under the binary covarion model, each lexical item, represented as a binary character, can transition between different evolutionary states, such as being subject to change or remaining invariant over time. This flexibility reflects the reality that certain words may evolve at different rates or may be more prone to changes in some languages compared to others, providing a more nuanced and accurate representation of lexical evolution in a phylolinguistic analysis. We tested models that differed in whether each concept was assigned an independent substitution rate ('per-concept rates'), allowing for variation in mutation rates across concepts, or a single shared substitution rate ('shared rates') across all concepts, as well as in whether a strict or relaxed clock was employed. A strict clock assumes a constant rate of evolution across all branches of the tree, while a relaxed clock allows for rate variation, accommodating differing evolutionary dynamics across branches and over time. For all four resulting models, we used a Birth-Death Skyline prior (Stadler *et al*., 2013). As recommended by one of the reviewers, we included calibration points to include temporal estimates. A uniform root prior ranging from 500 to 6000 BP was applied, as it facilitated convergence based on prior experience. The Xerente-Xavante split was calibrated with a lognormal distribution (mean in real space: 200, standard deviation: 0.05), following Maybury-Lewis (1974). The Northern-Southern split was assigned a normal prior (mean: 1900, standard deviation: 100), and the Northern-Central split a normal prior (mean: 1200, standard deviation: 100), both informed by archaeological evidence (see Section 5).

To identify the best-fitting model, we used Nested Sampling (Skilling, 2006), which is available as a package for the BEAST2 software (Russel *et al*., 2018), to calculate the log marginal likelihood for each model. Model selection was performed using the Bayes Factor, which compares the models' likelihoods and identifies the model that best explains the data (Hoffmann *et al*., 2021).

We conducted tree inference on the model selected as the best fit based on the Bayes Factor analysis. We performed 10^7 MCMC iterations, sampling a tree every 10^3 iterations, resulting in a posterior sample of 10^4 trees. From these, we obtained a maximum clade credibility tree (MCC) based on common ancestor heights using TreeAnnotator version 2.7.4 (Heled & Bouckaert, 2013) with a 50% burn-in.

To quantify the quality of the inferred tree topology, we calculated the Generalized Quartet Distance (Pompei *et al*., 2011) to the Glottolog tree (Hammarström *et al*., 2024) using the Quartet package in R (R Core Team, 2023; Sand *et al*., 2014; Smith, 2019). The Glottolog tree was pruned to include only the languages in our sample. Because Mebengokre, Kaingang, Kisedje and Canela-Kraho are represented as internal nodes rather than tips in the Glottolog tree, we used proxy languages within these nodes to complete the tree for comparison. The Glottocodes of the proxies were kaya1331 (Txucarramae), cent2143 (Central Kaingang), yaru1258 (Yaruma) and krah1246 (Krahô), respectively. Lower GQD values indicate greater topological similarity between the inferred tree and the Glottolog tree.

4 Results

The NeighborNet (Figure 3) and δ-scores and Q-residuals grouped by subgroup (Figure 2) indicate that most conflicting signals were found for Arikapu and parts of the Southern Je subgroup, whereas Northern Je (including Panará) and Central Je languages exhibit the least amount of conflicting signal.

Unsurprisingly Arikapu—a Jabutí language and the only non-Je language in the sample—deviates most from a tree-like structure. However, the elevated Q-residuals observed for the Southern Je languages cannot be conclusively attributed to a potential dialect continuum and may instead result from the limited amount of data available for this branch. Table 4 summarizes the results of model selection, ranking models by their log-marginal likelihoods. The differences between the highest log-marginal likelihood and those of other models are reported as log Bayes Factors. According to Kass & Raftery (1995), a log Bayes Factor of more than 20 indicates strong support for a model. The results show a clear preference for relaxed clock

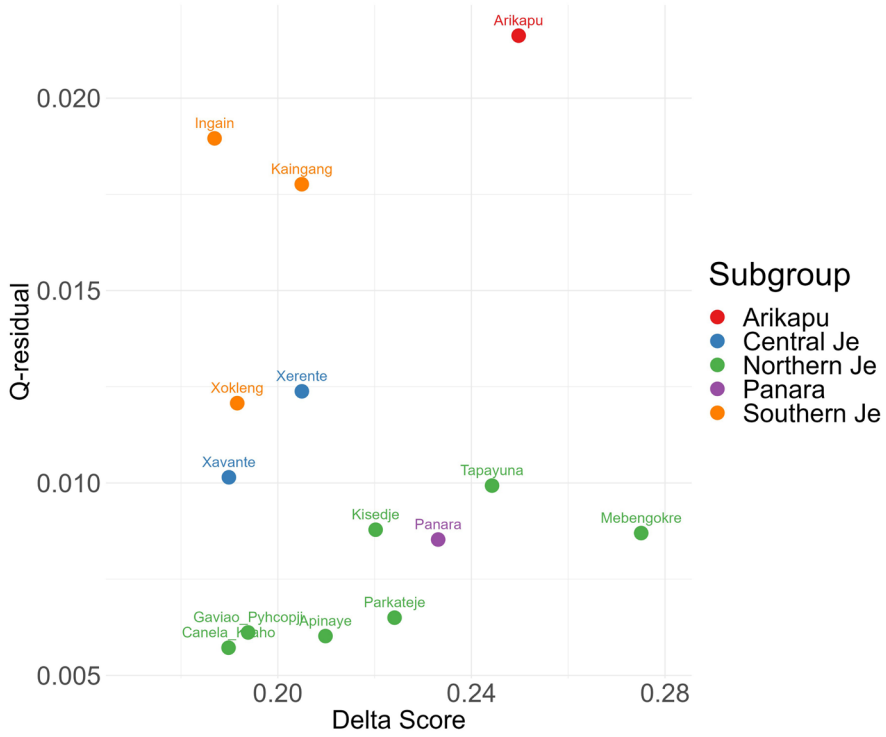


Figure 2. δ-scores and Q-residuals.

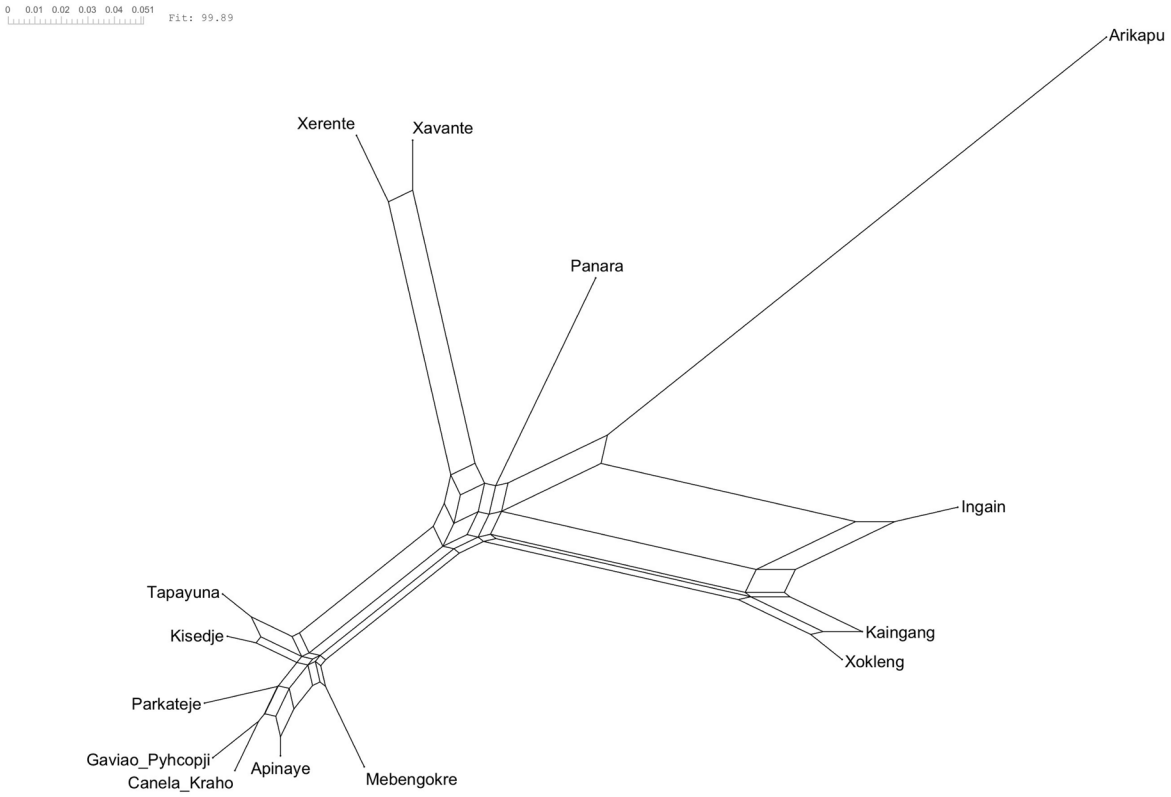


Figure 3. NeighborNet. The reticulations, the ”boxes” between the branches, indicate the conflicting signal in the data.

models over strict clock models. Additionally, models without per-concept substitution rates and therefore having fewer parameters, are favored. Following model selection, we performed tree inference using the best-fitting model, identified as the relaxed clock model using shared rates.

The relative order of splits in the inferred tree (Figure 4) aligns well with the pre-Columbian history of Je-speaking groups, as reconstructed through archaeological findings. The MCC tree (Figure 4) of the inferred phylogenetic tree was visualized using the FigTree software (Rambaut, 2010). The numbers on the nodes correspond to their posterior probabilities and the blue bars show the 95% highest posterior density estimate of the split timings. Notably, the inferred tree closely aligns with the established classifications, as quantified by a GQD (Generalized Quartet Distance) to the gold-standard tree of 0.124. In comparison to the findings of Häuser *et al.* (2024), where the GQD was calculated for trees inferred from datasets with expert cognacy judgments, this value is in the low range.

Table 4. Nested Sampling results (SD = standard deviation of the log-marginal likelihood estimate, BF = Bayes Factor).

Model	Log-Marginal Likelihood	SD	(log) BF
shared rates, relaxed clock	**-1967.84**	**2.76**	-
shared rates, strict clock	-1991.87	2.81	24.03
per-concept rates, relaxed clock	-2019.11	4.26	51.27
per-concept rates, strict clock	-2032.06	4.41	64.22

The first split separates Arikapu (Jabuti) from the Je languages, as expected, followed by a split that separates the Southern group. Subsequently, the Northern and Central languages diverge, consistent with archaeological data (see below).

5 Archaeology

The earliest material culture complex commonly associated with Macro-Je speakers is the Una tradition, distinguished by small, simple ceramic vessels found in open-air settlements and rock shelters in central Brazil (Dias-Jr., 1975; Robrahn-González, 1996). Given its dates and persistence until relatively recently, the Una tradition does not necessarily mark the initial split of the Macro-Je languages, but might have been adopted by various Macro-Je and Je-speaking populations. It does, however, seem to be related to the transition to more sedentary settlements, as well as with the transition to plant cultivation. The earliest dates come from Minas Gerais state, where Una ceramics are associated with early maize and other domesticated plant remains in rock shelters (3490 ± 120 BP, 2 σ cal BP 4078-3406 [SI-2788]) (Bird *et al.*, 1991). After ca. 2700 BP, this tradition is found throughout the *cerrado* (central Brazilian savanna) from Goiás to Mato Grosso (2390 ± 60 BP, 2 σ cal BP 2699-2147 [Beta-78256], and 2360 ± 70 BP, 2 σ cal BP 2698-2133 [SI-4068]) (Barbosa *et al.*, 1982; Wüst, 1998).

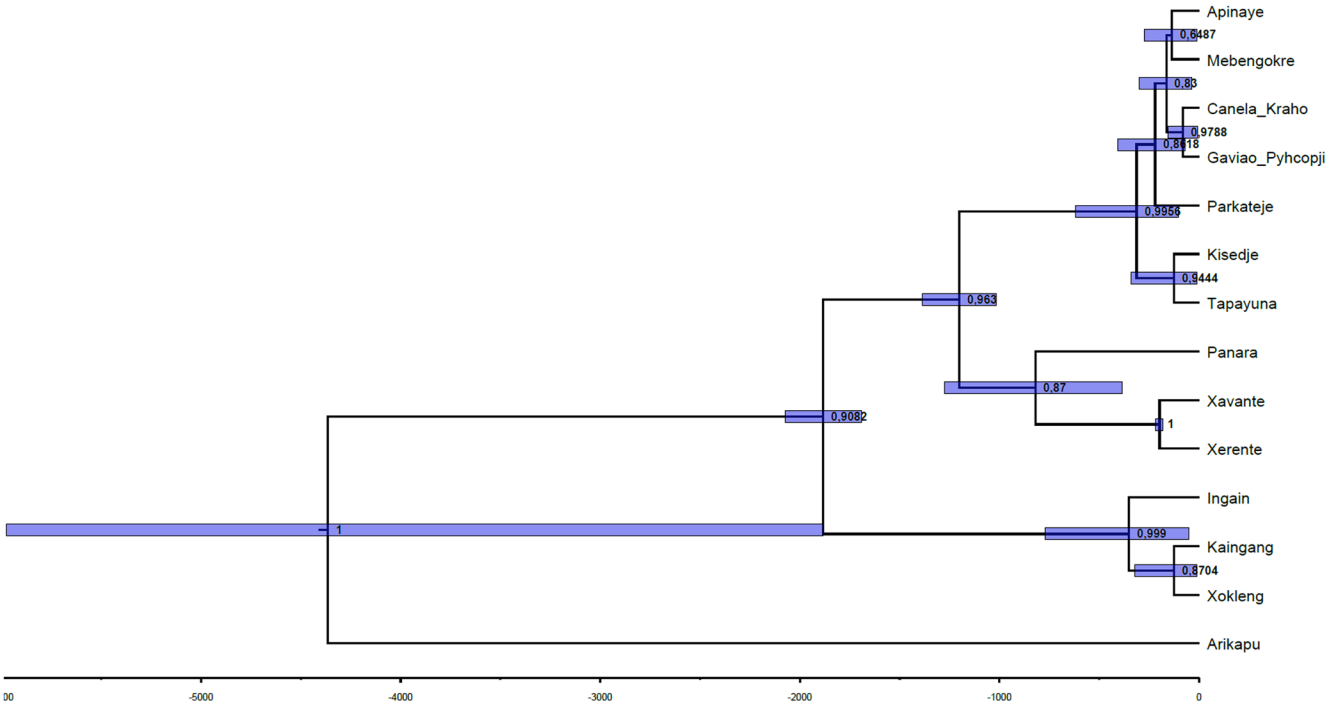


Figure 4. Topology of the tree with shared substitution rates and relaxed clock model.

The Southern Je branch is clearly associated with the Taquara/Itararé Tradition, based on its spatial distribution, material culture (small, simple ceramic vessels, sometimes with dark burnishing and plastic decoration), burial practices (funerary mounds, sometimes involving cremation), and chronology, all showing a persistence until early colonial times (Corteletti & Iriarte, 2018; de Mello Araujo, 2007; Noelli & De Souza, 2017). The earliest secure dates for this Tradition reveal a similar horizon of initial settlement for the three southern states, Paraná, Santa Catarina and Rio Grande do Sul - respectively, 1941 ± 35 BP (2 σ cal BP 1980-1741 [LACUFF-150057]), 1840 ± 40 (2 σ cal BP 1827-1610 [CAMS-53915]) and 1810 ± 85 (2 σ cal BP 1890-1431 [SI-813]) (De Masi, 2001; Parellada, 2016; Schmitz & Brochado, 1972). The dates associated with this archaeological dispersal to the south, later than those of the Una tradition, are coherent with the next split of the tree, which separates the Southern and the Central-Northern Je branches (Figure 4).

Around the first millennium AD, a new archaeological complex spreads throughout central Brazil and part of the Atlantic coast. Characterized by circular villages with central plazas, urn burials, and large ceramic vessels, this complex is known as the Aratu Tradition (Robrahn-González, 1996; Wüst & Barreto, 1999). The geographical distribution, similarities with ethnographic ring villages, and chronology confirm that this tradition is related to the Northern and Central Je groups. The origins appear to be in Goiás, with a date of 1220 ± 50 BP (2 σ cal BP 1259-960 [Beta-99031]), spreading to the coast of Bahia and Espírito Santo by 1080 ± 90 BP (2 σ cal BP 1175-741 [SI-542]) and 1076 ± 36 (2 σ cal BP 1050-807 [SI-2347]), respectively (Calderón, 1969; Perota, 1979; Silva *et al.*, 1997). It is important to notice that the presence of this culture in the coast implies that it was not a purely Je phenomenon, but might have been adopted by speakers of other Macro-Je branches or even by unrelated groups. However, our tree suggests that it was probably a retention shared by both Northern and Central Je speakers (Figure 4).

6 Discussion

The results presented here provide additional support for the existence of geographical clusters within the Je language family, as previously suggested by Rodrigues (Rodrigues, 1986; Rodrigues, 1999) and Nikulin (Nikulin & Salanova, 2019: 178). The results also highlight with high certainty that Arikapu is not a Je language.

The inclusion of Panará in a clade with the Central Je languages, specifically Xavante and Xerente, is not entirely unexpected given its geographic location – Panará is usually classified as a separate branch. Historically, this area was in close proximity to the regions where Akwe languages were spoken before their separation around 1815 (Maybury-Lewis, 1974). In fact, de Carvalho (2016) identified a loanword from an Akwe language in Panará, this being the only known instance of "contact" between these groups. These languages were historically spoken in the southern part of the Brazilian state of Tocantins and in northern Goiás, particularly along the middle course of the Tocantins River.

Nikulin (2020) does not include Panará in the Septentrional branch, instead proposing a separate branch comprising languages spoken in the Goyaz region (see also Nikulin & Salanova, 2019), which form a larger clade together with Panará. Within this branch, Panará is the southernmost language, and Nikulin (2020) acknowledges that part of its lexicon is of unknown origin (see also de Carvalho, 2016). These facts do not contradict the results in the tree presented in Figure 4; on the contrary, together they offer evidence regarding the internal classification.

It is worth noting that while the time estimates in the tree align closely with the priors set in the model, they are expected to be reasonably reliable. Excluding individual calibration points resulted in estimates for the omitted timings that remained consistent with those from the full analysis. This consistency indicates that the inferred dates are not solely driven by specific priors but reflect genuine signal in the data.

7 Conclusion

The phylogeny presented here does not introduce new findings as far as classification is concerned, but reinforces previous classifications of Je family, highlighting established geographical clusters of languages as sub-families and raising questions for further inquiry. The novelty of this research concerns the dating of the phylogeny, which agrees with archaeological data discussed in Section 5.

In future research, it will be essential to expand the database to attempt a more comprehensive classification that includes a greater number of (non-Je) languages, broader coverage, and manual cognate annotation. Additionally, incorporating not only lexical data but also morphological and syntactic data could be highly valuable, as these elements evolve at different rates and might offer crucial insights into the evolution of a language family (Hübler & Greenhill, 2022).

The geographical clusters (northern, central, and southern groups) provide a realistic model for understanding the spread of Je languages, but they must account for archaeological data, including those for non-Je groups, which are more scarce. While the Macro-Je hypothesis remains valid, further comparative research is required in order to establish a better groundwork for understanding the branching of the phylum. More information about non-Je languages is now available and will be included in future analyses.

Data availability

Zenodo: A phylogenetic classification of the Je language family: Supplementary data. https://doi.org/10.5281/zenodo.15041553 (Ferraz Gerardi *et al.*, 2025). The supplementary data is also available on https://github.com/TGH-2020/Je_Phylogeny.

This project contains the following underlying data:

- data

  - concepts.csv: a list of all concepts that have been used in the study, along with the coverage and the number of cognate sets per concept

  - auto_cogids.tsv: a list of all word forms and corresponding cogids assigned by automated cognacy judgements and used in the study

  - subset_for_metrics_calculation.csv: a subset of the data with columns for automatically and manually annotated cognates (used to assess the quality of automatic annotations)

- trees: files to run all model variations with BEAST, as well as the output from the BEAST runs

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

References

Adelaar WFH: **Relações externas do macro-jê. O caso do chiquitano.** In: S V T de A P L Telles & A S de Paula (eds.), *Topicalizando macro-jê.* Nectar Recife, Braz, 2008; 9–28.

Auderset S, Greenhill SJ, DiCanio CT, *et al.*: **Subgrouping in a 'dialect continuum': a Bayesian phylogenetic analysis of the Mixtecan language family.** *J Lang Evol.* 2023; **8**(1): 33–63.
Publisher Full Text

Barbosa AS, Schmitz PI, Stobhaeus A, *et al.*: **Projeto médiotocantins: monte do carmo, go. Fase cerâmica pindorama.** *Pesquisas: Antropologia.* 1982; **34**: 49–92.

Bird RM, Dias O, de Carvalho ET: **Subsídios para a arqueobotânica no Brasil: o milho antigo em cavernas de Minas Gerais, Brasil.** *Revista de Arqueologia.* 1991; **6**(1): 15–31.
Publisher Full Text

Boswood J: **Evidências para a inclusão do aripaktsá no filo macro-jê.** *Série Lingüística.* 1973; **1**: 67–78.
Reference Source

Bouckaert R, Heled J, Kühnert D, *et al.*: **Beast 2: a software platform for Bayesian evolutionary analysis.** *PLoS Comput Biol.* 2014; **10**(4): e1003537.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Bryant D, Moulton V: **Neighbor-net: an agglomerative method for the construction of phylogenetic networks.** *Mol Biol Evol.* 2004; **21**(2): 255–265.
PubMed Abstract | **Publisher Full Text**

Calderón V: **A fase Aratu no recôncavo e litoral norte do estado da Bahia.** In: PRONAPA (ed.), *Museu Paraense Emílio Goeldi*. 1969; 161–168.

Chang W, Cathcart C, Hall D, *et al.*: **Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis.** *Language.* 2015; **91**(1): 194–244.
Reference Source

Corteletti R, Iriarte J: **Recent advances in the archaeology of the Southern Proto-Jê people**. 2018; 1–11.
Publisher Full Text

Davis I: **Comparative Jê phonology.** *Estudos Lingüísticos: Revista Brasileira de Lingüística Teórica e Aplicada.* 1966; **1**(2): 10–24.
Reference Source

de Carvalho FO: **On the development of the Proto-Northern Jê Rhotic in Panará historical phonology.** *Anthropological Linguistics.* 2016; **58**(1): 52–78.
Reference Source

de Carvalho FO, Damulakis GN: **The structure of Akroá and Xakriabá and their relation to Xavante and Xerente: a contribution to the historical linguistics of the Jê languages.** *LIAMES: Línguas Indígenas Americanas.* 2015; **15**(1): 17–46.
Reference Source

de Castro Alves F: **Evolution of alignment in Timbira.** *International Journal of American Linguistics.* 2010; **76**(4): 439–475.
Publisher Full Text

De Masi MAN: **Pescadores coletores da costa sul do Brasil.** *Pesquisas: Antropologia.* 2001; **57**: 1–136.
Reference Source

de Mello Araujo AG: **A tradição cerâmica Itararé-Taquara: características, área de ocorrência e algumas hipóteses sobre a expansão dos grupos Jê no sudeste do Brasil.** *Revista de Arqueologia.* 2007; **20**(1): 9–38.
Reference Source

Dias O Jr: **Pesquisas arqueológicas nas grutas do Brasil.** *Anais do 10º Congresso Nacional de Espeleologia.* 1975; 61–71.
Reference Source

Dixon RMW, Aikhenvald A: **The Amazonian languages.** Cambridge University Press, 1999.
Reference Source

Ehrenreich P: **Die einteilung und verbreitung der Völkerstämme Brasiliens nach den gegenwärtigen stande unserer kenntnisse.** *Petermanns Mitt.* 1891; **37**: 81–89, 114–124.
Reference Source

Ferraz Gerardi F, Tresoldi T, Aragon CC, *et al.*: **Lexical phylogenetics of the Tupí-Guaraní family: language, archaeology, and the problem of chronology.** *PLoS One.* 2023; **18**(6): e0272226.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Ferraz Gerardi F, Wientzek T, Gregorio de Souza J, *et al.*: **A phylogenetic classification of the Je language family: supplementary data. (version v1).** *Zenodo.* 2025.
http://www.doi.org/10.5281/zenodo.15041553

Gerardi FF, Tresoldi T: **CLDF dataset derived from MaJeLeD (Macro-Je Lexical Database) (v0.1-pre).** *Zenodo.* [Data set]. 2024.
http://www.doi.org/10.5281/zenodo.12304366

Gildea S, de Castro Alves F: **Nominative-absolutive: counteruniversal split ergativity in Jê and Cariban.** In: Spike Gildea & Francesc Queixalós (eds.), *Ergativity in Amazonia.* John Benjamins Amsterdam, 2010; 159–199.
Publisher Full Text

Gildea S, de Castro Alves F: **Reconstructing the source of nominative-absolutive alignment in two Amazonian language families.** In: Jóhanna Barðdal, Spike Gildea & Eugenio R. Luján (eds.), *Reconstructing syntax.* Brill, 2020; 47–107.
Publisher Full Text

Gray RD, Bryant D, Greenhill SJ: **On the shape and fabric of human history.** *Philos Trans R Soc Lond B Biol Sci.* 2010; **365**(1559): 3923–33.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Greenberg JH: **Language in the Americas.** Stanford University Press, 1987.
Reference Source

Greenhill SJ: **Language phylogenies: modelling the evolution of language.** In: *The Oxford Handbook of Cultural Evolution.* Oxford University Press, 2023.
Publisher Full Text

Greenhill SJ, Heggarty P, Gray RD: **Bayesian phylolinguistics.** In: Richard D. Janda, Brian D. Joseph & Barbara S. Vance (eds.), *The handbook of historical linguistics.* Wiley Online Library, 2020; **2**: 226–253.
Publisher Full Text

Gudschinsky SC: **Ofaié-xavánte, a Jê language.** In: Sarah C. Gudschinsky (ed.), *Estudos sôbre línguas e culturas indígenas.* Brasília: Summer Institute of Linguistics, 1971; 1–16.
Reference Source

Guérios RFM: **O nexo linguístico bororo-merrime-caiapó.** *Revista do Círculo de Estudos Bandeirantes.* 1939; **2**(1): 61–74.
Reference Source

Hammarström H, Forkel R, Haspelmath M, *et al.*: **Glottolog database 5.0**. 2024.
Publisher Full Text

Häuser S, Jäger G, Rama T, *et al.*: **Are sounds sound for phylogenetic reconstruction?** 2024.
Reference Source

Heled J, Bouckaert RR: **Looking for trees in the forest: summary tree from posterior samples.** *BMC Evol Biol.* 2013; **13**: 221.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hoffmann K, Bouckaert R, Greenhill SJ, *et al.*: **Bayesian phylogenetic analysis of linguistic data using BEAST.** *J Lang Evol.* 2021; **6**(2): 119–135.
Publisher Full Text

Hübler N, Greenhill SJ: **Modelling admixture across language levels to evaluate deep history claims.** *J Lang Evol.* 2022; **7**(2): 166–183.
Publisher Full Text

Huelsenbeck JP: **Testing a covariotide model of DNA substitution.** *Mol Biol Evol.* 2002; **19**(5): 698–707.
PubMed Abstract | **Publisher Full Text**

Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies.** *Mol Biol Evol.* 2006; **23**(2): 254–267.
PubMed Abstract | **Publisher Full Text**

Inman D, Vuillermet M: **Singular-plural verb stem alternation: uncovering global and local drivers of typological variation.** *Linguist Typol.* (n.d.) to appear/forthcoming.

Kaiping GA, Klamer M: **Systematic Bayesian phylogenetics reveal dialect chain breakup in the Timor-Alor-Pantar language family.** *OSF.* 2019.
Publisher Full Text

Kass RE, Raftery AE: **Bayes factors.** *J Am Stat Assoc.* 1995; **90**(430): 773–795.
Publisher Full Text

Kaufman T: **Language history in South America: what we know and how to know more**. In: D. L. Payne (ed.), *Amazonian linguistics: studies in Lowland South American languages*. Austin: University of Texas Press, 1990; 13–73.
Reference Source

Kolipakam V, Jordan FM, Dunn M, *et al.*: **A Bayesian phylogenetic study of the Dravidian language family.** *R Soc Open Sci.* 2018; **5**(3): 171504.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

List JM: **Sequence comparison in historical linguistics.** In: *Sequence Comparison in Historical Linguistics*. Düsseldorf University Press, 2014.
Publisher Full Text

List JM, Forkel R: **Lingpy. a python library for quantitative tasks in historical linguistics [book library, version 2.6.8].** Leipzig: Max Planck Institute for Evolutionary Anthropology, 2021.
Reference Source

Loukotka C: **La familia lingüística mašakali (estudios de lingüística comparada americana).** *Revista del Instituto de etnología de la Universidad nacional de Tucumán.* 1932; **2**: 21–47.

Loukotka C: **La familia lingüística coroado.** *J Soc Am.* 1937; **29**(1): 157–214.
Reference Source

Loukotka C: **A língua dos patachos.** *Revista Arquivo Municipal.* 1939; **55**: 5–15.
Reference Source

Loukotka C: **Klassifikation der sudamerikanischen sprachen.** *Zeitschrift für Ethnologie.* 1942[1944]; **74**: 1–69.

Maddison DR, Swofford DL, Maddison WP: **Nexus: an extensible file format for systematic information.** *Syst Biol.* 1997; **46**(4): 590–621.
PubMed Abstract | **Publisher Full Text**

Mason JA: **The languages of South America.** In: Julian Steward (ed.), *Handbook of South American Indians*. (Smithsonian Institution Bureau of American Ethnology Bulletin), chap. 6, Washington: Government Printing Office, 1950; **143**: 157–317.

Maybury-Lewis D: **Akwe-Shavante society**. New York: Oxford University Press, 1974.

Nikulin A: **Proto-macro-jê: um estudo reconstrutivo.** Universidade de Brasília. (Doctoral dissertation), 2020.
Reference Source

Nikulin A, Salanova AP: **Northern Jê verb morphology and the reconstruction of finiteness alternations.** *Int J Am Linguist.* 2019; **85**(4): 533–567.
Publisher Full Text

Nimuendajú C: **Mapa etno-histórico do brasil e regiões adjacentes.** Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN) (ed.). 3rd edition. Reprint of the original 1917 map, updated and republished by IPHAN. Brasília, Brazil: IPHAN, 2017.
Reference Source

Noelli FS, De Souza JG: **New perspectives on the archaeological cartography of the jê in Southern Brazil.** *Bol Mus Para Emílio Goeldi Cienc Hum.* 2017; **12**(1).
Publisher Full Text

Parellada CI: **Paisagens transformadas: a arqueologia de povos jê no paraná, sul do Brasil.** *Rev Museu Arqueol Etnol.* 2016; **27**: 158–167.
Publisher Full Text

Perota C: **Sítio monsarás**. 1979.

Pickering WA: **A fonologia xavante: uma revisitação.** Universidade estadual de Campinas. (doctoral dissertation), 2010; **256**.
Reference Source

Pompei S, Loreto V, Tria F: **On the accuracy of Language Trees.** *PLoS One.* 2011; **6**(6): e20109.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

R Core Team: **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna, Austria, 2023.
Reference Source

Rama T, List JM: **An automated framework for fast cognate detection and Bayesian phylogenetic inference in computational historical linguistics**. In: Anna Korhonen, David Traum & Lluís Màrquez (eds.), *Proceedings of the 57th annual meeting of the Association for Computational Linguistics.* Florence, Italy: Association for Computational Linguistics, 2019.
Publisher Full Text

Rambaut A: **Figtree v1.4.4**. 2010.

Ramirez H, Vegini V, de França MCV: **Koropó, puri, kamakã e outras línguas do leste Brasileiro.** *LIAMES: Línguas indígenas Americanas.* 2015; **15**(2): 223–277.
Publisher Full Text

Ribeiro ER: **Prefixos relacionais como evidência históricocomparativa: os casos chiquitano e jabutí.** *Línguas e Culturas Macro-Jê.* 2011; **2**: 105–120.
Reference Source

Ribeiro ER: **A grammar of karajá**. University of Chicago. (Doctoral dissertation): 2012; x+291.
Reference Source

Ribeiro ER, van der Voort H: **Nimuendajú was right: the inclusion of the Jabutí language family in the Macro-Jê stock.** *International Journal of American Linguistics.* 2010; **76**(4): 517–570.
Publisher Full Text

Rivet P: **Langues Américaines iii: langues de l'amérique du sud et des antilles**. In: Antoine Meillet & Marcel Cohen (eds.), *Les langues du monde.* Paris: Société Linguistique de Paris, (Coll. Ling. Paris), 1924; **16**: 639–712.

Robrahn-González EM: **Os grupos ceramistas pré-coloniais do Centro-Oeste brasileiro.** *Revista do Museu de Arqueologia e Etnologia.* 1996; **6**: 83–121.
Publisher Full Text

Rodrigues AD: **Línguas brasileiras: para o conhecimento das línguas indígenas**. Edições Loyola, 1986.
Reference Source

Rodrigues AD: **Macro-jê**. In: R. M. W. Dixon & A. Aikhenvald (eds.), *The Amazonian languages.* Cambridge: Cambridge University Press, 1999; 165–206.

Rodrigues AD: **A case of affinity among Tupí, Karíb and Macro-Jê.** *Revista Brasileira de Linguística Antropológica.* 2009; **1**(1): 139–167.
Reference Source

Russel PM, Brewer BJ, Klaere S, *et al.*: **Model selection and parameter inference in phylogenetics using nested sampling.** *Syst Biol.* 2018; **68**(2): 219–233.
PubMed Abstract | **Publisher Full Text**

Salanova AP: **The Macro-Jê languages**. In: Patience Epps & Lev Michael (eds.), *Amazonian languages. An international handbook.* In preparation. Berlin: Mouton de Gruyter. N.d

Sand A, Holt MK, Johansen J, *et al.*: **Tqdist: a library for computing the quartet and triplet distances between binary or general trees.** *Bioinformatics.* 2014; **30**(14): 2079–2080.
PubMed Abstract | **Publisher Full Text**

Schmidt W: **Die Sprachfamilien und Sprachenkreise der Erde.** Helmut Buske, 1926.
Reference Source

Schmitz PI, Brochado JJP: **Datos para una secuencia cultural del estado de rio grande do sul, Brasil.** *Gabinete de Arqueologia Publicações.* 1972; **2**: 1–20.

Silva R, Rubin JCR, Viana SA: **Resgate arqueológico: sítios gengibre e lourenço.** Universidade Católica de Goiás, 1997.

Skilling J: **Nested sampling for general bayesian computation.** *Bayesian Anal.* 2006; **1**(4): 833–859.
Publisher Full Text

Smith MR: **Quartet: comparison of phylogenetic trees using quartet and split measures**. R package version 1.2.5. 2019.
Publisher Full Text

Stadler T, Kühnert D, Bonhoeffer S, *et al.*: **Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV).** *Proc Natl Acad Sci U S A.* 2013; **110**(1): 228–233.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Storto L: **Línguas indígenas: tradição, universais e diversidade**. Campinas: Mercado de Letras, 2019.
Reference Source

Swadesh M: **The origin and diversification of language: edited post mortem by joel sherzer**. Chicago: Aldine, 1971.

von den Steinen K: **Durch Zentral-Brasilien**. Brockhaus, 1886.
Reference Source

Wiesemann U: **The pronoun systems of some Jê and Macro-Jê languages**. In: Ursula Wiesemann (ed.), *Pronominal systems.* Tübingen: Gunther Narr Verlag, 1986; 359–380.
Reference Source

Wüst I: **Continuities and discontinuities: archaeology and ethnoarchaeology in the heart of the Eastern Bororo territory, Mato Grosso, Brazil.** *Antiquity.* 1998; **72**(277): 663–675.
Publisher Full Text

Wüst I, Barreto C: **The ring villages of central Brazil: a challenge for Amazonian archaeology.** *Latin American Antiquity.* 1999; **10**: 3–23.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ? ✔ ? ? ? ✔

Version 3

Reviewer Report 03 June 2025

https://doi.org/10.21956/openreseurope.22021.r54231

© **2025 Auderset S et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Sandra Auderset
Department of Linguistics, University of Bern, Bern, Switzerland
Andrés Pablo Salanova
Department of Linguistics, University of Ottawa, Ottawa, Ontario, Canada

This paper presents a fairly standard application of Bayesian phylogenetics to the Jê language family. This family has not been studied with those methods prior and the study is innovative in this regard. There are however some inconsistencies and open questions that should be addressed.

Section 1:
The introduction is somewhat confusing and it is not always clear how it relates to the study at hand. Much time is spent discussion various proposals regarding what is to be included in a Macro-Jê phylum, but the study itself focuses on subgrouping within the Jê family.
The authors bring up several issues in the classification of (Macro-)Jê languages, but it is not always clear whether they think those issues are specific to this family or how their study addressed them. The authors state that "The hypothesis that Macro-Je could result from long-term linguistic contact via ethnogenesis (Dixon & Aikhenvald, 1999) cannot yet be discarded.", suggesting perhaps that this family (or area?) developed differently than others. But the comparison among Macro-Jê languages has proceeded in much the same way as with other language families, and the results don't indicate anything unusual about their history. If anything, scarcity of data and sloppy work are responsible for the slow progress in our understanding of this language family. The problem of disentangling contact from inheritance is ubiquitous in classification and historical linguistics more broadly and not unique to this family. The other issue that caused some confusion on our part concerns the sound correspondences or purported absence thereof.

The overview of the typology of Jê languages includes a couple of claims that need rephrasing or further explanation.

- "Many Je languages display a rare type of split alignment that has not yet been sufficiently studied. This alignment pattern has been called nominative-absolutive". This is a

phenomenon that is exclusive to some Northern Jê languages, that is understood in outline as a pattern that arises from restructuring of a complex construction.
- "a somewhat complex system of person marking that interacts with agreement, since most languages are of the dependent marking type". It is not clear to us what they are referring to here.

As suggested by other reviewers, the Archaeology section (5) should be moved up to the introduction or in any case it needs to appear before the results as it is used for the calibration points.

Section 2:
The data that was used as input for the Bayesian phylogenetic analysis is not provided anywhere. It's stated that it's extracted from Gerardi & Tresoldi 2024 and a list of the 330 concepts and languages is provided. However, it would helpful to have the actual extracted data set in the supplementary materials.
The authors should also provide more details regarding the cognate-coding or the checking thereof, as this determines to a large extent how comparable this study is to previous qualitative ones based on the comparative method and sound correspondences. The data in Gerardi & Tresoldi 2024 is already annotated for cognacy. Was that cognate coding carried out automatically? The authors say it was based on Nikulin 2020, which suggests it was done manually based on the comparative method.
Looking at Gerardi & Tresoldi 2024, it seems that the data was at least partially morpheme-segmented, but the cognate coding is done per concept not per morpheme. Partial cognates are sometimes treated as the same (cf. set 41 for ALLIGATOR), sometimes as different (cf. sets 17 and 15 for ABANDON). All of these methodological issues should be clarified.

Section 3:
A burn-in of 50% seems very unusual (it's usually around 10%). Could the authors add a brief explanation of why it is so high in this case?
Another question that needs clarification is why the Glottolog tree is used as the 'gold-standard' for assessing the quality of the MCC tree of the Bayesian phylogenetic analysis, especially as the tree had to be modified with respect to tips and nodes.

Section 4:
Here also it would be good to add some clarifications regarding the interpretation of the results.
- "indicate that most conflicting signals were found for Arikapu and parts of the Southern Je subgroup, whereas Northern Je (including Panará) and Central Je languages exhibit the least amount of conflicting signal." → But Mebengokre and Tapayuna (Northern Je) exhibit higher Delta scores than the rest of the languages.
- "However, the elevated Q-residuals observed for the Southern Je languages cannot be conclusively attributed to a potential dialect continuum and may instead result from the limited amount of data available for this branch." → We don't fully understand why high Q-residuals need to be interpreted as resulting from a dialect continuum.
- "The relative order of splits in the inferred tree (Figure 4) aligns well with the pre-Columbian history of Je-speaking groups, as reconstructed through archaeological findings." → This seems a bit circular, as the archaeological findings were used for the calibration points. Also given the difficulties of associating material culture with specific languages, this should phrased more carefully (see also below on Section 5).
- "Excluding individual calibration points resulted in estimates for the omitted timings that

remained consistent with those from the full analysis." → Does this mean that only one calibration point is removed at a time? What happens if you remove all calibration points?

The tree in Figure 4 includes several features that go against the consensus on the classification of Jê languages. A summary of what one can claim about the history of Northern Jê based on traditional methodologies can be found in Section 7 of Salanova & Nikulin 2020. We briefly summarize the main points of departure from said consensus:
(1) Parkatêjê should grouped together with Canela-Krahô and Gavião-Pyhcopji,
(2) Panará should be closer to the northern languages than to the central ones,
(3) Khĩsêtjê + Tapayuna should be closer to Mẽbêngôkre + Apinajé.
A forth issue that might be worth discussing although this is has not reached broad consensus is Nikulin's (2020) proposal that Khĩsêtjê + Tapayuna are closer to Mẽbêngôkre than Apinajé is, based on shared retentions.

Section 5:
The way the archaeological data is interpreted is somewhat wishful: in general terms, the article speaks of splits among branches of the phylum, when archaeological evidence is about expansion. In the case of the Una and the Taquara/Itararé traditions, the expansion is equated with the Macro-Jê and the Southern Jê splits, respectively (the former with reservations). In the case of Aratu, it has to be assumed that there are two branches, split several centuries earlier, that expanded simultaneously with the same material culture.

General comments:
In general, the clarity of the paper could be improved with consistent use of spellings of language and family names. Both are variously spelled with and without diacritics – the latter probably out of convenience. The figures would also benefit from more informative and explicit figure captions.

Some minor issues with terminology/presentation:

- Krenák is in the map, but not the table – or rather it is in the table as Borum, which is also the name of the family(?)
- for Kayapó and Mẽbêngôkre there is an ISO code txu, but it's not given in the table
- Table 1: would work better if grouped by family
- Figure 1: overlaps of language names and points could be fixed, points could be larger
- Figure 3: color by proposed subgroup like the plot for easier comparison
- Fig 4: shouldn't this be shared rates tree with relaxed clock, as this was the best model?

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Partly

Is the argument persuasive and supported by evidence?

Yes

If any, are all the source data and materials underlying the results available?
Partly

Does the research article contribute to the cultural, historical, social understanding of the field?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: SA: historical linguistics, phylogenetics; APS: historical linguistics, Jê languages

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Reviewer Report 30 May 2025

https://doi.org/10.21956/openreseurope.22021.r54391

© 2025 Greenhill S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Simon J Greenhill
Biological Sciences, University of Auckland, Auckland, New Zealand

I approve the article for indexing.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: language phylogenies, human prehistory, language evolution

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 2

Reviewer Report 03 June 2025

https://doi.org/10.21956/openreseurope.21626.r53106

© 2025 Kim R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ronald I Kim
[1] Adam Mickiewicz University in Poznań, Poznań, Poland
[2] Adam Mickiewicz University in Poznań, Poznań, Poland

The article under review marks a significant contribution to the rapidly growing field of computational phylogenetic linguistics as well as to the study of the native languages of South America, specifically present-day Brazil. It unquestionably deserves to be indexed and will be read with interest and profit by those working on those fields, as well as historical linguists in general.

Although I am hardly an expert on computational phylogenetics, the methodology adopted in the study seems both sound and justified, and the authors explain their choices in a mostly clear manner. The results obtained are inherently plausible, although whether they in fact constitute support for phylogenetic groupings is another matter (see below). The discussion of archeological evidence is welcome and appears to support the "real-world" events underlying the diversification of the Je family.

I am in agreement with the assessment of the previous reviewers, particularly Prof. Salmons, who rightly points out the value of this study for dating the Je family and its subgroups and recommends stressing that point in the conclusions. I also concur with him that diffusion via language contact must always be considered, particularly when dealing with "culturally relevant concepts related to fields like agriculture, vegetation and fauna" which have notoriously spread far and wide across multiple language families, e.g. in prehistoric Eurasia or East Asia. This does not in itself vitiate the authors' analysis, but it would be a good idea to address his question whether the cognacy rates for the Swadesh list and "culturally relevant concepts" are similar, and also to what extent the archeological evidence in section 5 is consistent with "genetic" linguistic subgroups (i.e. a "clean break" in phylogenetic terms), a steadily diversifying dialect continuum, or some combination thereof.

A more serious issue is Prof. Salmons's comment regarding the comparison of monosyllabic roots. In addition to Ringe's 1992 monograph, see the same author's 1999 article on the relative ease of matching CVC roots. Given that many of the relevant roots in the Je languages are simple CV syllables and that many of the consonant inventories are relatively small, there is a significant risk that automated cognacy judgments will incorrectly identify accidental resemblances as genuine cognate sets, making it all the more difficult to demonstrate greater-than-chance comparisons across these languages. Once again, this does not invalidate the results of the present study, but the authors are urged to address this concern in a footnote, perhaps adding that the cognacy sets obtained by their methodology remain to be (in)validated by the tried and true Comparative Method, which (at least for this "traditional" scholar) remains indispensable to establishing the reality of language families.

The article is well written, and the English is for the most part idiomatic and stylistically appropriate. A few exceptions are listed below.

Discussion
the Macro-Je language phylum's development > the development of the Macro-Je language phylum

Introduction (p. 4)
paragraph 3: geographically distributed broadly > with broad geographic distribution
Meridional > South
paragraph 4: The proposal for the Macro-Je phylum > The proposal of a Macro-Je phylum
(Rodrigues, 1999) > Rodrigues (1999)
paragraph 5: Maxakali > Maxakalí
paragraph 6: the Macro-Je phylum, for which > the Macro-Je phylum for which
paragraph 7: many parallels presented, > many claimed cognates,

p. 5
Table 1: change Setentrional > North
change Meridional > South
Map: why accents on Xavánte, Xerénte?

Introduction (p. 6)
paragraph 10: Je languages tend to be head-final SOV as the most common order. > Je languages tend to be head-final and have basic SOV order.
a somewhat complex system > a complex system
(only phonology) > (phonology only)
paragraph 11: no mention of Section 6?

Results (p. 7)
paragraph 2: Unsurprisingly Arikapu > Unsurprisingly, Arikapu

Results (p. 9)
paragraph 2: Additionally, models without per-concept substitution rates and therefore...
> Additionally, models without per-concept substitution rates, and therefore...
paragraph 3: The MCC tree (Figure 4) of... > The MCC tree of... [since you just referred to Figure 4]
their posterior probabilities and > their posterior probabilities, and

Archaeology, paragraph 1: Una tradition > Una Tradition (2x)

p. 10
Archaeology, paragraph 2: The earliest secure dates for this Tradition > The earliest secure dates for this tradition
later than those of the Una tradition, > later than those of the Una Tradition,
paragraph 3: It is important to notice that the presence of this culture in the coast > It is important to note that the presence of this culture on the coast
it was probably a retention: how so, and how can a culture be a "retention"? Please explain.

Discussion, paragraph 1:
The results also highlight > The results also demonstrate
paragraph 3: Septentional branch, > Northern branch,

Conclusion, paragraph 1:
The phylogeny presented here does not introduce new findings as far as classification is concerned, but reinforces... > The phylogeny presented here reinforces...
which agrees with archaeological > with agrees with the archaeological
[In line with the comments above, add another sentence stressing the importance of this agreement for the value of your computational results.]

Reference cited:
refer to 1

References

1. Ringe, Don. 1999. How hard is it to match CVC-roots? Transactions of the Philological Society 97(2): 213–244.

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by evidence?

Yes

If any, are all the source data and materials underlying the results available?

Yes

Does the research article contribute to the cultural, historical, social understanding of the field?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: historical linguistics, sociolinguistics, language contact, Indo-European, languages of Eurasia

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 15 May 2025

https://doi.org/10.21956/openreseurope.21626.r53104

© 2025 King B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Benedict King
[1] Max Planck Institute for Evolutionary Anthropology, Leipzig,, Germany
[2] Max Planck Institute for Evolutionary Anthropology, Leipzig,, Germany

The study is a fairly standard phylogenetic analysis of a language family. It's been done well and is clearly written. I have a few points about that analysis, in particular how the calibrations have been implemented. My main reason for selecting approval with reservations is that the position of Panara in the Southern group might be affected by the implementation of the calibrations. Since this is perhaps the most interesting result of the phylogeny, I think it's worth digging into a little more.

- In section 3.2 there is the line "testing various covarion models". It appears that only the standard BEAST implementation of the covarion model is used. What is tested is relaxed vs strict clock and partitioned versus unpartitioned site models.
- The 2 site rate setups used represent extremes: either a single rate for all concepts, or every concept having a different rate (303 extra parameters!). Usually binning by the number of cognate sets (as in the main analysis of Heggarty et al. 2023 for example) provides the best model fit. In an analysis with 303 concepts it seems unlikely that a single rate for all concepts is a good fit.
- The archaeology section should come earlier, to justify the calibration points used in the analysis. This section claims that the archaeological dates are "coherent" with the tree, but isn't this the evidence used to build the tree?
- I am concerned that the calibrations are not set up in the most logical way, and that there may be an interaction with the tree topology affected Panara. I noticed that in the uncalibrated trees the position of Panara is highly uncertain and switches between North-Central and Southern branches. The justification for the North/Central-South calibration is for an "archaeological dispersal to the South", but it's not clearly described how this should be implemented in a phylogenetic analysis. The calibration is a standard common ancestor node prior at the moment, i.e. it defines the most-recent common ancestor of Je languages. I find it hard to justify a calibration prior without monophyly constraints in language phylogeny. How can you be sure of a link between archaeological culture and language but not be sure that the languages are monophyletic? I find it much more logical to implement calibrations as "originate priors". This means that the parent node rather than the common ancestor is calibrated. In this case, if there is evidence that the Taquara culture represents the Southern dispersal of the Southern Je languages, then an originate prior would be appropriate to calibrate the time that Southern languages begin independent evolution. This has the advantage that the sister group and therefore the actual node calibrated can be estimated in the phylogeny. I don't know if Panara should be included as descending from this Southern migration or not, but it sounds like perhaps not.
- From the description in the next paragraph, it sounds like there is no justification for the calibration on North-South group. If this cultural complex is a shared retention, how can it be used to calibrate the common ancestor of the North-Central grouping?
- Briefly summarise Maybury-Lewis for the calibration on Xerente-Xavante. The standard

deviation of 0.05 seems way too precise, that's about 18 days which is shorter than I took to do this review (sorry about that).
- It is not mentioned in the paper if the Bayesian analysis results were checked for proper mixing and convergence and if more than one run was performed. The results do not include the trace log files and only a single tree file.
- It is mentioned that testing for removing calibration points was performed, but no details given and the files aren't shared
- Another possibly missing detail is that an ascertainment bias correction was used to correct for the absence of all-absent cognate sets.
- In Figure 4 label the 3 groupings. Some indication of which nodes were calibrated might also be good.
- Why is Glottolog considered the gold standard tree for comparison? What is the point of the Bayesian analysis if it isn't an improvement?
- In the abstract it says the Je family and Macro-Je phylum were investigated. The latter is not really true since only one non-Je language was included, essentially as an outgroup.
- The plain language summary says 516 words were analysed, but only 303 were used in the end.
- In the conclusion it says "Raising questions for further inquiry". What are these questions?

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
Partly

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Phylogenetics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 14 May 2025

https://doi.org/10.21956/openreseurope.21626.r52806

© 2025 Chousou-Polydouri N. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Natalia Chousou-Polydouri
[1] Institute for Mediterranean Studies, Foundation for Research and Technology—Hellas, Rethymno, Greece
[2] Institute for Mediterranean Studies, Foundation for Research and Technology—Hellas, Rethymno, Greece

The article presents a phylogenetic analysis of Je languages (including one non-Je Macro-Je language) using a large lexical dataset and automatic cognate detection. The article is interesting and provides a much-needed step for the classification of this family. However, several methodological choices need to be clarified and motivated, while other methodological options need to be tested additionally. Finally, some results and discussion points may need to be modified accordingly.

Methodological Choices

Coding method

The coding method needs to be further clarified. If I understand well, these are not cognate classes per se, but partial cognate classes that only include items that belong to the same concept. These have been named with different terms in the literature, such as within-concept cognate sets, or cognate sets without items that have undergone semantic shift, or root-meaning (or root-concept) sets. Finally each question mark is indicating a missing language-root-concept triplet, not a concept-language pair in the NEXUS format.

Evolutionary model

It is unclear why only covarion models were tested. Why not a simple binary model or a stochastic Dollo? (even though I can see arguments against a stochastic Dollo, especially in a preliminary analysis). The covarion model was developed for a particular use in evolutionary biology (modeling rate variation due to changes in secondary structure during macroevolution) and it is at best unclear what it would represent in linguistic evolution. The authors' explanation on the appropriateness of the covarion model is inaccurate. First, different words (I assume here roots or concepts are intended) evolving at different rates (a reasonable assumption) should be modeled by rate heterogeneity across sites, not by the covarion model (which is indeed the case in the article). Second, according to the covarion model as used, words that are "prone to change in some languages more than in others", essentially means that the same root appears and disappears relatively rapidly for the same concept for some languages. This is a highly unrealistic situation and does not corresponding to any known linguistic phenomenon.

Tree prior

The choice of tree prior needs to be motivated and further clarified. Also, additional tree priors need to be tested. There are different parameterizations of BDSKY. Which exactly was used and

why? Skyline priors are especially suited for populations dynamics and epidemiology. Why not use a plain birth death tree prior in this study (possibly with incomplete sampling)? All parameters of the model chosen need to be explained (i.e. what they represent in linguistic evolution) and also the specific prior distributions used for each parameter need to be described.

Root prior
The uniform prior used seems like a rather restrictive prior for the root of the tree, especially since it doesn't allow any older dates than 6000. Also, the reasoning given for choosing this prior (to improve convergence) is puzzling. Priors are chosen to reflect prior beliefs and they have to be justified. What exactly were the convergence issues and which other prior distributions were tested? Furthermore, in Figure 4, the bar corresponding to the 95% confidence interval for the root age is going all the way to 6000 years before present, further indicating that the root prior is too restrictive.

Calibrations
general comments
The archaeological section needs to be moved to the introductory section of the paper, before the methods and analysis sections. It is not contextualizing the research, it is used as an assumption in order to calibrate the tree.
The thought process from the archaeological facts to the derivation of calibrating distributions needs to be expanded.
Finally, all details regarding the calibrations need to be given in the Methods section. It is not clear exactly how these calibrations were implemented. Did you constrain specific monophyletic groups? Were they crown or stem calibrations? Also, the distribution choice and the values chosen need to be justified. why was the normal distribution chosen here, while the lognormal in the Xerente-Xavante case? if I understand well, in all cases you only have dates that indicate that a split cannot be younger than this date.

Xerente-Xavante split
The key facts used to derive the calibration from the citation listed need to be repeated here, as well as an explanation of how the calibration was derived (i.e. the thought process).

Northern-Southern and Northern-Center splits
Here the archaeological facts are presented in detail, but it is unclear how they led to the particular calibrations. Why was the normal distribution chosen here, while the lognormal in the Xerente-Xavante case? if I understand well, in all cases you only have dates that indicate that a split cannot be younger than this date.

Analysis
You should report how many runs were conducted and how many among these converged. A 50% burn-in is a really high percentage and indicates that you probably have not reached convergence. In any case longer runs are needed for reliable results.

Results and Discussion
general comment
Throughout the results and discussion section it needs to be clear what is assumed and what is a result of your analysis. For example, were there any subgroups that were assumed to be monophyletic during the calibration process?

Another case is the end of the archaeology section (section 5). It seems here that you are interpreting the archaeology based on your tree, while you have used the archaeological data to calibrate it. It needs to be clear what is assumed in order to derive these calibrations, both in terms of which cultural tradition is associated with which branch, as well as which branches (if any) are assumed to be monophyletic in the analysis.
In the conclusion (section 7), the reasoning seems circular. The archaeological data were used to calibrate the phylogeny, so you can't interpret the dating as fitting with the archaeology. If the dating was done with linguistic data (e.g. through old sources), then a comparison with archaeological dates would be appropriate.

non-tree-like signal of Arikapu
This is not a self-evident result, as indicated at the end of page 7. If anything, the default expectation would be that Arikapu would be widely separated from everything else, since it is a more distant relative. It would not have necessarily more conflicting signal. Instead, I think this may be because of unsure/erroneous cognate judgements between Arikapu and the rest of the languages, exactly because of the distant relationship. Automated cognate judgements for Arikapu need to be evaluated separately to see if they are significantly different from manual cognate judgements and/or all Arikapu judgments should be reviewed by hand.

Bayes Factors comparisons
The text here (end of page 7 and beginning of page 9) is contradictory and also does not agree with Table 4. On the table the model with shared rates is highlighted (presumably as the chosen one), but in the text you mention using per-concept rates. Please clarify which model was used for the final analysis.
Additionally, the way that BF factors are reported in table 4 is confusing. Positive logBF should be associated with the chosen model, not with the non-chosen ones. A 4X4 table with all comparisons would be better.

MCMC tree
The calibrated splits should be annotated on the tree, as well as any clades that are used in the results and discussion (e.. Norther, Central etc).

Minor corrections
- Rikbaktsa is mispelled two times in Table1
- page 6, first column, second paragraph. The sentence starting "Je languages tend to be head-final" is ungrammatical. Maybe there is a "with" missing between head-final and SOV?
- Inman & Vuillermet is published, see www.degruyterbrill.com/document/doi/10.1515/lingty-2023-0040/html?lang=en
- page 10, first column, last paragraph. It is unclear what Akwe languages refers to. Is it a cover term for Xavante and Xerente? Without more context, this section's argument is not entirely clear.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Partly

Is the work clearly and cogently presented?
Partly

Is the argument persuasive and supported by evidence?
Partly

If any, are all the source data and materials underlying the results available?
Yes

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: phylogenetics, historical linguistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 03 May 2025

https://doi.org/10.21956/openreseurope.21626.r52803

© 2025 Ryder R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Robin J. Ryder
[1] Université Paris Dauphine and Crest INSEE, Paris, France
[2] Université Paris Dauphine and Crest INSEE, Paris, France

The authors reconstruct the history of the Je languages, using phylolinguistics methods on lexical data.

This paper is an interesting and relevant addition to the literature and is clearly written. The revised version follows the state of the art; the statistical methods and conclusions seem robust. The authors used automatic cognate detection to transform the lexical data into binary data: their results tend to show that this procedure is appropriate for this data set. The online supplementary material includes the necessary XML files to reproduce the studies, as well as the trees sampled from the posterior distribution.

I recommend accepting this paper. I only have a small number of minor comments:

* You could check that the results are stable when you restrict the data to only concepts in the Swadesh list. The rationale is that those concepts are assumed to change at a slower rate and to be less susceptible to borrowing, thus giving more reliable signal on the deep history.
* Although we are given details of the MCMC, no diagonistics are given. Please specify how you checked the MCMC has converged and mixed well (i.e. that the number of iterations is large enough).
* The calibration of the Xerente-Xavante split has mean 200 and standard deviation 0.05. That is a very low sd; is it maybe a typo?
* The subfamilies are sometimes referred to as Northern and Southern, sometimes as Septentrional and Meridional. Please use consistent names. Also, there is a missing p in Septentrional in Table 1.
* I was a bit confused by the different lists of languages. Table 1 is the most complete list, yet it doesn't include Canela-Kraho and Parkateje (both used in the phylogenies) nor Kayapo and Krenak (both present on the map of Figure 1). Additionally, the code for Mebengokre is different between Table 1 and Table 2.
* I was very glad to read the last paragraph of section 6 on reconstructing calibration: this check strengthens your conclusions. But the paragraph is vague. You should include the details of the reconstructed ages (for example with a figure such as those in Pellard, Ryder & Jacques 2024 or in Ryder & Nicholls 2011, two papers in which we argue for that practice).

Refer:Pellard, T., Ryder, R., & Jacques, G. (2024). The family tree model. The Wiley Blackwell companion to diachronic linguistics.

References

1. Ryder R, Nicholls G: Missing Data in a Stochastic Dollo Model for Binary Trait Data, and its Application to the Dating of Proto-Indo-European. *Journal of the Royal Statistical Society Series C: Applied Statistics*. 2011; **60** (1): 71-92 Publisher Full Text

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by evidence?

Yes

If any, are all the source data and materials underlying the results available?

Yes

Does the research article contribute to the cultural, historical, social understanding of the field?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: phylolinguistics, stochastic modelling, Bayesian inference, computational statistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 14 May 2025

Tim Wientzek-Paul

We thank the reviewer for the helpful and constructive comments. Since the paper has been accepted, we have limited the revisions to minor corrections and clarifications. We have checked MCMC convergence using Tracer 1.7.2 (Rambaut et al., 2018), and most parameters show effective sample sizes above 200, in line with the rule of thumb in Hoffmann et al. (2021). Regarding the calibration prior, the mean of 200 refers to real space, while the standard deviation of 0.05 is specified in log space, as indicated in the manuscript. We have revised the terminology for the subfamilies to consistently use “Northern” and “Southern” throughout. We also added the missing languages to Table 1 and clarified that “Mebengokre” and “Kayapó” refer to the same variety. While we hope that our manual concept selection and cognate coding helped minimize the impact of borrowing, we agree that restricting analyses to Swadesh concepts could be a valuable robustness check, and we will consider this in future work. Similarly, we appreciate the suggestion to report reconstructed calibration ages in more detail and will take this into account in upcoming projects. **References:** Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*. **syy032**. doi:10.1093/sysbio/syy032 Hoffmann, K., Bouckaert, R., Greenhill, S. J., & Kühnert, D. (2021). Bayesian phylogenetic analysis of linguistic data using BEAST. *Journal of Language Evolution*, *6*(2), 119–135. https://doi.org/10.1093/jole/lzab005

Competing Interests: No competing interests were disclosed.

Reviewer Report 19 April 2025

https://doi.org/10.21956/openreseurope.21626.r52808

© 2025 Salmons J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Joseph Salmons
[1] University of Wisconsin, Madison, USA
[2] University of Wisconsin, Madison, USA

This article provides a welcome contribution to the classification of the Je languages and will be of

interest to comparative linguists generally. It deserves to be out there, and I simply have some suggestions for clarifying and strengthening it.

The authors rightly stress the value of this paper for dating developments in the family, and dating is often a particularly challenging issue — like the large literature on this for Indo-European shows. That said, it's also important that we have independent converging evidence for this picture of the family. First, the phylogenetic evidence is consistent with traditional comparative work. Second, it appears consistent with the archeological record. These three perspectives together give us greater confidence in the overall picture. You might stress that positive rather than concluding with the negative that you don't present 'new findings'.

My main concern is an issue noted in passing in the manuscript and underscored by the first reviewer, that many of the forms used are monosyllabic so that accidental correspondences come into play. In fact, a lot of the forms include only a single consonant, a CV syllable. Ribeiro (2012), cited on this, is more explicit about this than the paper is. The reader should have more details about this issue and how important it is. To go beyond Ribeiro: At least some of these languages have small phoneme inventories, like 13 or 14 consonants, where the modal number of consonants reported across languages is 21 (Gordon 2016: 45, drawing on Maddieson 1984). This means there's a high likelihood of accidental correspondences, even assuming perfectly systematic sound correspondences and leaving aside borrowings. This presumably risks overstating the relatedness of these languages. Campbell (1988) and Ringe (1992) provide two different discussions of this issue for comparative linguistics.

A second issue is language contact. Yes, borrowing surely plays a role here, especially as the subgrouping lines up with geography, so that many sisters would have often been in contact with each other longer than with more distant relatives. 'Culturally relevant concepts' are included in the dataset here, and those are often seen as prone to borrowing, while the traditional claim was that Swadesh items were less likely to be borrowed. (The Leipzig-Jakarta list and other more recent lists have better empirical foundations, but Kossmann 2013 shows how complex the issue is.) Are the cognacy rates similar for the Swadesh and cultural sets? This is an important area for future work, but some initial evidence would be helpful here.

Finally, you talk about using 516 items, but 303 is the actual number of comparisons used. I would change that for clarity while noting that you started from a larger set.

Refer:

1. Campbell, Lyle. 1988. Review article on Greenberg 1987. Language. 64(3). 591-615.
2. Gordon, Matthew K. 2016. Phonological typology. Oxford: Oxford University Press.
3. Kossmann, Maarten. 2013. The Arabic influence on northern Berber. Leiden: Brill.
4. Maddieson, Ian. 1984. Patterns of Sounds. Cambridge: Cambridge University Press.
5. Ringe, Donald A. 1992. On calculating the factor of chance in language comparison. Philadelphia: Transactions of the American Philosophical Society 82(1).

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the

topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
Yes

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: historical and comparative linguistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 14 May 2025

Tim Wientzek-Paul

We are grateful to the reviewer for taking the time to read our piece and contributing with observations and suggestions. The main concern voiced by the reviewer targets the short (CV and CVC) shape of most of the compared lexical elements.

He brings up the well-known fact that putative cognate sets of this kind can be problematic when advancing claims of common ancestry among languages, as a prevalence of short morpheme shapes leaves one unable to exclude chance as the main factor explaining the attested similarities. Further, the reviewer notes that similar concerns were expressed in Ribeiro (2012), a work cited in our paper. Despite the cogent nature of these observations, they are, as far as an evaluation of the present paper goes, entirely beside the point and misplaced, as we explain now.

Although the hypothesis of a Macro-Jê family, a group of languages and families related to the Jê linguistic family, remains essentially open as far as its 'definitive' nature and limits are concerned (see e.g., Rodrigues 1999; Ribeiro & van der Voort 2010; Nikulin 2020), the same is not true for the Jê language family, which constitutes the focus of the paper. Indeed, Davis (1966), even having at his disposal much less extensive data than is currently available (his comparative Jê is limited to 112 cognate sets, among lexical and grammatical morphemes), describes the languages in question as "obviously related" (Davis 1966: 10). See also Ribeiro & van der Voort 2010: 549).

Unsurprisingly, the discussion in Ribeiro (2012), mentioned by the reviewer, concerns the status of Macro-Jê, not Jê, as a family (see Ribeiro 2012: 266). The existence of a Jê language family has been undisputed since the middle of the 20th century. The only non-Jê, Macro-Jê language in the study is Arikapu, a member of the Jabuti family used as an outgroup. The language, along with its closer relative Djeromitxi, is also clearly related to the Jê family, as shown by shared morphological aberrancies involving cognate sets supported by regular correspondences (Ribeiro & van der Voort 2010: 533); paradigmatically structured morphology (Ribeiro & van der Voort 2010:563-4) and shared (inherited) homophonies (e.g., Proto-Jê ŋrɛ 'dance', ŋrɛ 'egg' : Arikapu ɾɛ̃ 'dance', ɾɛ̃ 'egg'; Ribeiro & van der Voort 2010: 562-3). That said, we agree with the general gist of the reviewer's focus on the nature of cognate judgements underlying our analysis. An absolutely necessary follow-up to the present paper will implement a manual checking of the automatic cognation judgements on which our contribution was based. The same applies, of course, to the other source of similarities: borrowing/diffusion.

References: Davis, Irvine. 1966. Comparative Jê phonology. *Estudos Lingüísticos: Revista Brasileira de Lingüística Teórica e Aplicada* 1 (2): 10-24. Ribeiro, Eduardo R. 2012. A Grammar of Karajá. Doctoral Dissertation, The University of Chicago. Ribeiro, Eduardo R. and Hein van der Voort. 2010. Nimuendajú was right: the inclusion of the Jabutí language family in the Macro-Jê stock. *International Journal of American Linguistics* 76 (4): 517-570. Rodrigues, Aryon D. 1999. Macro-Jê. In: R. M. W. Dixon and Alexandra Aikhenvald (eds.) *Amazonian Languages*. Cambridge: CUP: 165-206.

Competing Interests: No competing interests were disclosed.

Reviewer Report 12 April 2025

https://doi.org/10.21956/openreseurope.21626.r52440

© 2025 Greenhill S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Simon J Greenhill
[1] Biological Sciences, University of Auckland, Auckland, New Zealand
[2] Biological Sciences, University of Auckland, Auckland, New Zealand

I think the authors for their considered response to my comments and suggestions. I thnk this paper is now suitable for indexing

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: language phylogenies, human prehistory, language evolution

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 February 2025

https://doi.org/10.21956/openreseurope.20937.r50634

© **2025 Greenhill S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Simon J Greenhill
[1] Biological Sciences, University of Auckland, Auckland, New Zealand
[2] Biological Sciences, University of Auckland, Auckland, New Zealand
[3] Biological Sciences, University of Auckland, Auckland, New Zealand

This paper presents a phylogenetic analysis of the Je languages. The language family is interesting, and the analysis looks to follow best practices. Overall the paper is well-constructed and I only have few suggestions to improve the paper (of which I consider #1 and #2 the most critical).

I recommend acceptance after these revisions.

1. No evaluation of cognate accuracy.
The authors apply an automated cognate detection algorithm to identify cognates, making the argument that these automated methods perform well.

However, automated cognates are not perfect, and not as widely accepted as the current manuscript suggests. This is especially the case when -- as the paper notes -- there has been (a) reported problems identifying cognates due to monosyllabic words, and (b) high levels of borrowing.

How can the authors show us that the cognates they are basing their analyses off are remotely plausible? I would like to see a more careful evaluation on a subset of the data: are there some cognates identified in the literature using traditional methods? do the automated ones recapitulate these? even better, .is there a cognate-coded subset that can be used as a formal test of how well the automated approach is at capturing true cognates? Failing this, can the authors present some of the identified sets and convince the reader that they are plausible?

Adding the identified cognate sets into the supplementary material would also help readers evaluate their performance. (I commend the authors on having everything else in the supplement).

2. Why not date the language family?

Despite using a method that estimates time-dated phylogenies, the analysis does not add any time information. The paper would be much more impactful if it added date calibrations into the analysis. The authors state that they can't add time as there is only one calibration -- Xavante & Xerente ~200B.P. -- but this does not meet the criteria from Maurits et al.

I disagree. First, the Maurits et al. article is overstating the case quite substantially, and they are basing their recommendations on a very well-documented and well-studied language group (Uralic): quite a different situation to Macro-Je. Second, the section on archaeology goes on to describe a range of other plausible dates that look like ready calibrations to me e.g. Southern Je, ~1800-2000 B.P. and the Aratu tradition ~1200B.P.

3. The authors suggest that the low delta/q-residual scores in the southern Je clade indicate a possible dialect chain. An alternative is that there is something different about these data that make the scores higher. This could be higher levels of missing data, or could be fewer identified cognate sets due to some linguistic process. Can these alternatives be ruled out?

4. The paper opens with “Comprises about 15 languages”...but table 1 lists far more which threw me on first reading. To avoid reader confusion, I recommend introducing Macro-Je in the introduction before focusing in on the Je subfamily. Consistency with terms would help as well (e.g. Macro-Je = family, Je = subfamily).

5. The paper opens by talking about phylogenetic methods. This makes the paper seem like it'll be a technical evaluation of these tools, but this is not the case. It also seems like the only reason for doing this analysis is to apply phylogenetic methods rather than shedding light on human prehistory. Can the opening paragraph reframe around the question of Je/Macro-Je rather than the methods? The paper will get more readers this way.

6. The term 'partitioned' is used in the tables, but this is not explained in text. I would recommend a term like "per-word rates" or defining this when you describe the models.

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by evidence?

Partly

If any, are all the source data and materials underlying the results available?

Partly

Does the research article contribute to the cultural, historical, social understanding of the

field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: language phylogenies, human prehistory, language evolution

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 18 Mar 2025

Tim Wientzek-Paul

Dear Simon, we appreciate your valuable feedback which we consider to have improved the paper. We have followed your recommendations, and the article has been revised accordingly.

1) No evaluation of cognate accuracy.

This is an excellent suggestion. To assess the reliability of the automated cognate detection, we manually annotated a subset of the data and calculated a B-Cubed score. The results indicate that the automatic method produces highly plausible cognacy judgments. Additionally, we have included the identified cognate sets in the Supplementary Material to allow readers to evaluate them directly.

2) Why not date the language family?

Following your recommendation, we have repeated the Bayesian analysis using calibration points. This resulted in a slight reordering of splits, with Southern Je now diverging from Proto-Central-Northern Je instead of Central Je splitting first (a split with a very low posterior value in the un-dated tree). The updated results are reflected in the revised Figure 4 and Table 4.

3) Low delta/q-residual scores in the Southern Je clade.

We investigated alternative explanations for the higher Q-residuals observed for the Southern Je clade and found that there were indeed higher levels of missing data than for all other clades. We adjusted our interpretation of the results accordingly.

4) Macro-Je and Je confusion

We agree that the previous terminology was inconsistent and potentially confusing. To improve clarity, we now refer to Macro-Je as a “phylum” and Je as a “family” throughout the manuscript.

5) The paper opens by talking about phylogenetic methods.

We have chosen not to address this point.

6) The term 'partitioned' is used in the tables, but this is not explained in text.

Thank you for pointing this out. We have replaced the terms "partitioned" and "unpartitioned" with the more intuitive labels "per-concept" and "shared" when describing substitution rates.

Competing Interests: No competing interests were disclosed.

Reviewer Response 28 Mar 2025

Simon Greenhill

The authors have handled my comments and suggestions appropriately and I have no further concerns.

Competing Interests: None.