RESEARCH ARTICLE

# Nonlinear diversification rates of linguistic phylogenies over the Holocene

**Marcus J. Hamilton** [iD][1,2]*, **Robert S. Walker**[3]

**1** Department of Anthropology, University of Texas at San Antonio, San Antonio, TX, United States of America, **2** Santa Fe Institute, Santa Fe, New Mexico, NM, United States of America, **3** Department of Anthropology, University of Missouri, Columbia, MO, United States of America

\* marcus.hamilton@utsa.edu

## Abstract

The expansion of the human species out of Africa in the Pleistocene, and the subsequent development of agriculture in the Holocene, resulted in waves of linguistic diversification and replacement across the planet. Analogous to the growth of populations or the speciation of biological organisms, languages diversify over time to form phylogenies of language families. However, the dynamics of this diversification process are unclear. Bayesian methods applied to lexical and phonetic data have created dated linguistic phylogenies for 18 language families encompassing ~3,000 of the world's ~7,000 extant languages. In this paper we use these phylogenies to quantify how fast languages expand and diversify through time both within and across language families. The overall diversification rate of languages in our sample is ~0.001 yr$^{-1}$ (or a doubling time of ~700 yr) over the last 6,000 years with evidence for nonlinear dynamics in language diversification rates over time, where both within and across language families, diversity initially increases rapidly and then slows. The expansion, evolution, and diversification of languages as they spread around the planet was a non-constant process.

## Introduction

As the geographic range of the human species expanded throughout Africa and beyond, eventually including the majority of the world's terrestrial environments, human populations and their cultures diversified [1]. Today there are over 7,000 languages around the planet [1,2], more than the total number of mammal species [3]. Moreover, most of the world's current languages belong to agricultural language families [4], and so the majority of current ethnolinguistic diversity has most likely evolved only in the last ten thousand years since the Neolithic [5–7].

While recent research shows that the global biogeographic distribution of linguistic diversity is closely correlated with climatic, environmental, socioeconomic and demographic processes [8–10], the rate at which this diversity evolved and whether it varied over time is unclear [11–14]. Similar questions are asked in the evolutionary dynamics of biodiversity [15]. Under one scenario, rates of linguistic diversification may slow through time as populations compete increasingly for finite resources and space in environments, in which case diversification

eventually slows, saturates and asymptotes toward a quasi-equilibrium; this is the negative density-dependent model. Or, perhaps no stable equilibrium exists because between-group competition, or other mechanisms, continuously drive increasing ethnolinguistic diversification and replacement independent of larger environmental constraints; this is a positive density-dependent model where diversity increases diversification rates. Alternatively, ethnolinguistic diversity may be out of equilibrium because of the relatively recent agricultural revolution, and associated population growth that replaced hunter-gatherer ethnolinguistic diversity over large parts of the planet; this is a non-equilibrium model. If this were the case, increasing post-Neolithic population growth toward the present may be driving novel forms of diversification processes not found in hunter-gatherer populations.

To understand how linguistic diversification played out over time we focus on the structure of linguistic phylogenies. Phylogenetic trees based on linguistic variation have proven to be a powerful analytical tool for reconstructing human population and cultural histories [16–18]. By utilizing a Bayesian statistical approach on the systematic codings of linguistic cognates or phonetic data, these methods have generated dated phylogenies for many of the world's language families, including the large families Austronesian [16], Bantu [17], and Indo-European [18] (Fig 1). These phylogenies allow us to estimate rates of linguistic diversification, show how they vary through time, and compare rates across different language families.

In a simple binary branching process, the number of languages, $N$, will increase exponentially with time, $t$, following the exponential growth function, $N(t) = N_0 e^{rt}$, where $r = b-d$, the net difference in language speciations, $b$ and extinctions, $d$. Starting from the root of a language family, $N_0 = 1$, and solving for the net diversification rate, $r$, we define

$$r \equiv \ln N(t)/t. \qquad (1)$$

If the slope, $r$, of the number of languages, $N$, through time, $t$, is linear in a semi-log plot then the diversification rate is multiplicative and approximately constant, and therefore exponential [19–21] (Fig 2). However, if diversification rates $r$ vary through time, either because of time-varying speciation or extinction rates [22, 23], then the slope $r$ will not be linear on a
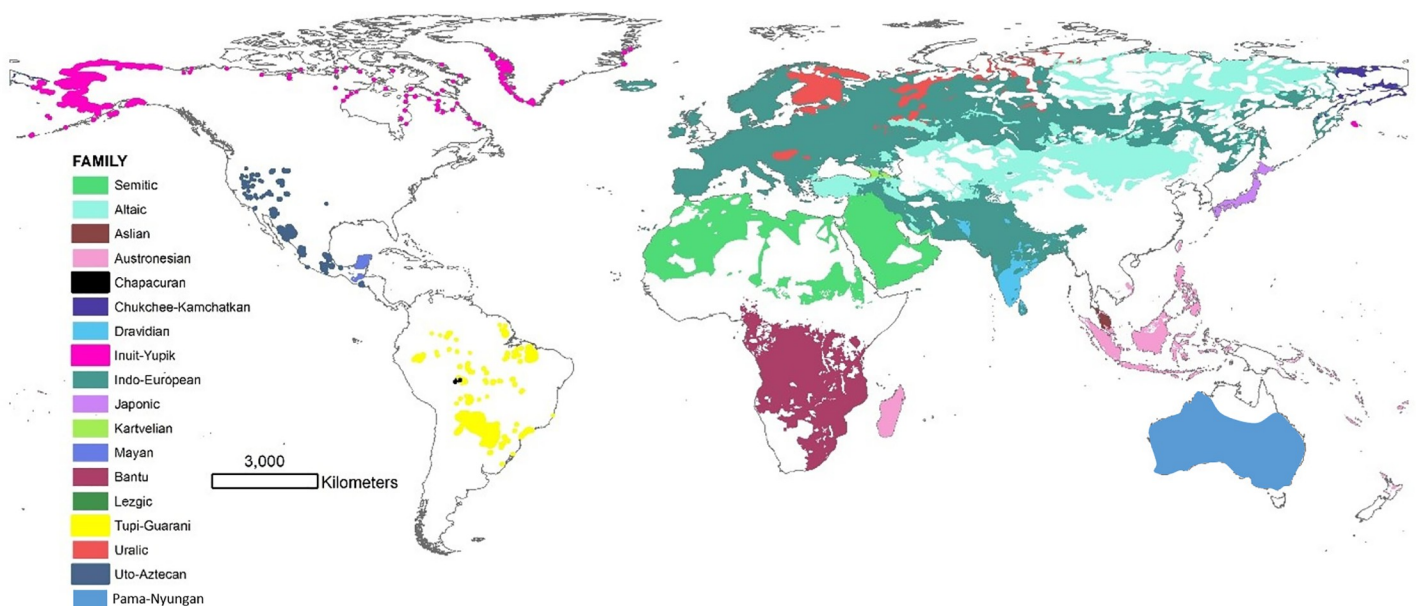


**Fig 1. Map of the language families in our sample using polygons from the Ethnologue [2].**
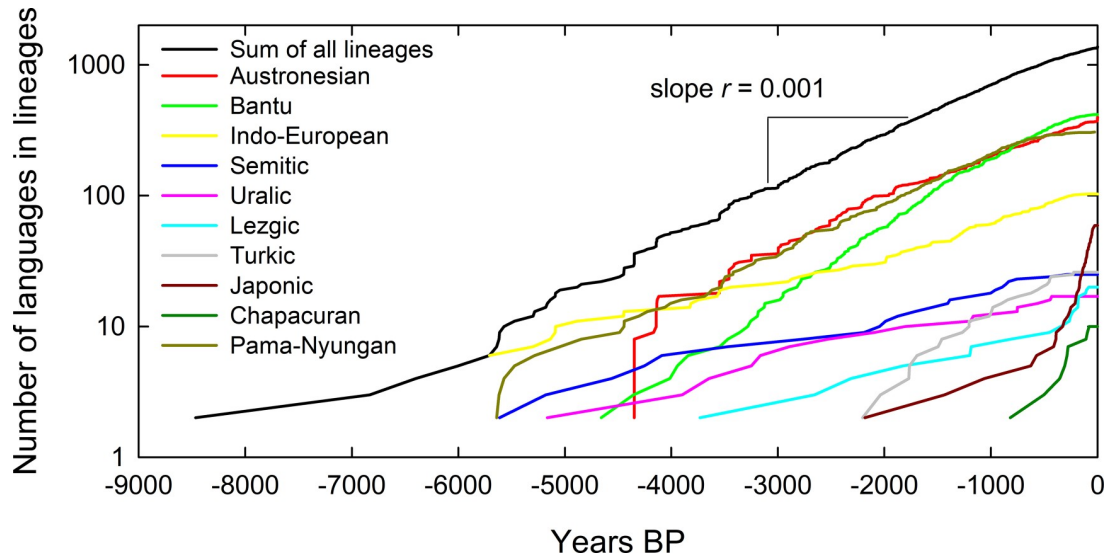
**Fig 2. Lineages-through-time plots for 10 language families and the summed total lineages.** The number of lineages on the y-axis is on a log scale such that the slopes of these lines are estimates of diversification rate *r*. While there is considerable variation in diversification rates within and across language families, the sum of the lineages suggests an approximately constant diversification rate overall.

https://doi.org/10.1371/journal.pone.0213126.g002

semi-log scale. Therefore, *r* is not constant and the diversification rate is time-dependent. To explore the diversification rates we first plot data for all lineages, as well as the sum of all lineages, in Fig 2. Second, for each lineage we then examine the distribution of diversification rates.

The first basic rate of interest for each phylogenetic tree is the overall diversification rate, $r_E$, which is simply the maximum size of the language family (i.e., the number of individual languages), $N_{max}$ of the *i*th tree divided by its' time depth, $T_i$ or

$$r_{E,i} = \ln N_{max,i}/T_i. \tag{2}$$

From the overall diversification rate we then establish a doubling time,

$$t_{2,i} = \ln 2/r_{E,i} \tag{3}$$

which is the number of years it takes for the number of languages in the *i*th tree to double in size (Table 1).

In evolutionary biology decreasing rates of species diversification through time are often interpreted as evidence for negative density dependence where ecological niches become saturated and limit the opportunities for further speciation [24–29]. Speciation and extinction rates often vary through time leading to expanding or contracting biodiversity and are not well fit by simple exponential growth models [15, 25–27, 30–37]. Here, we ask whether linguistic diversity shows similar evolutionary dynamics. As diversification rates are known to vary within large language families [16–18], we calculate diversification rates between successive time periods,

$$r_i(t) = \Delta \ln N_i/\Delta t. \tag{4}$$

Therefore, for each phylogeny we calculate multiple diversification rates and ask how they change over time, and in response to the number of languages. Because the time depth, $T_i$, varies across the lineages between 1,000 to 13,000 years, and the number of languages within

**Table 1. Language families analyzed in this study.**

| Family | N sample | N Ethno | N Glotto | Root age | $r_E$ Ethno | Doubling time | $r_E$ Glotto | Doubling time | Source |
|---|---|---|---|---|---|---|---|---|---|
| Kartvelian | | 5 | 6 | 5500 | 0.00029 | 2390 | 0.00032 | 2127 | [38] |
| Dravidian | | 85 | 80 | 4500 | 0.00034 | 2039 | 0.00034 | 2056 | [73] |
| Mayan | | 35 | 33 | 6500 | 0.00055 | 1260 | 0.00054 | 1288 | [39] |
| Lezgic* | 20 | 9 | 10 | 3730 | 0.00059 | 1175 | 0.00062 | 1122 | [40] |
| Inuit-Yupik | | 11 | 12 | 4000 | 0.0006 | 1155 | 0.00062 | 1115 | [38] |
| Uralic | 17 | 39 | 48 | 5300 | 0.00069 | 1005 | 0.00073 | 948 | [41] |
| Indo-European* | 103 | 449 | 584 | 8700 | 0.0007 | 990 | 0.00073 | 946 | [18] |
| Aslian* | | 18 | 19 | 4000 | 0.00072 | 963 | 0.00074 | 941 | [42] |
| Semitic | 25 | 78 | 97 | 5614 | 0.00078 | 889 | 0.00081 | 850 | [43] |
| Uto-Aztecan | | 61 | 69 | 5000 | 0.00082 | 845 | 0.00085 | 818 | [44] |
| Japonic* | 59 | 12 | 15 | 2182 | 0.00114 | 608 | 0.00124 | 558 | [45] |
| Bantu | 409 | 668 | 721 | 4800 | 0.00136 | 510 | 0.00137 | 505 | [17] |
| Austronesian | 398 | 1268 | 1276 | 5230 | 0.00137 | 506 | 0.00137 | 506 | [16] |
| Chapacuran* | 10 | 5 | 12 | 1039 | 0.00155 | 447 | 0.00239 | 289 | [46] |
| Tupi-Guarani | | 51 | 48 | 2500 | 0.00157 | 441 | 0.00155 | 447 | [47] |
| Chukchee-Kamchatkan | | 5 | 5 | 1000 | 0.00161 | 431 | 0.00161 | 430 | [38] |
| Turkic | 26 | 41 | 44 | 2206 | 0.00168 | 413 | 0.00172 | 404 | [48] |
| Pama-Nyungan | 306 | 285 | 248 | 5671 | 0.00099 | 695 | 0.00097 | 712 | [74] |

In the Family column, "*" indicates that the tree includes both language and dialect data. Sample sizes are given for language numbers in the available phylogenies, including dialects where applicable, and the total number of extant languages within families from both the Ethnologue [1] and the Glottolog [2]. Diversification rates are estimated from Eq 1 using only extant languages (natural-logged) divided by estimated root age in years using both the Ethnologue and Glottolog data.

lineages varies from 5 to 1,268, we standardize the time depth and size of each language family, so $N_i(t)' = N_i(t)/N_{i,\max}$ and $t_i' = t_i/T_i$. Therefore, $N_i(t)'$ is the proportion of the total number of languages at time, $t$, in the $i$th language family and time, $t_i'$, is a constant fraction of the time depth (or root age) of the $i$th language family (see S1 and S2 Figs). To calculate the rate of diversification of the trees we then divided the normalized sizes of the lineages in to ten bins of $0.1t'$ each, thus making the number of measureable diversification rates across the lineages directly comparable. Therefore, we are measuring the diversification rate over each consecutive 10% increment in time within each phylogenetic tree, regardless of the actual time depth of the tree thus making the diversification rates directly comparable across trees. All data and code used in the analysis is included in a "README.txt" file attached to this paper (S1 File).

## Results

### Phylogenetic uncertainty

In the following results we do not report phylogenetic uncertainty as the data we used to calculate estimates of diversification rates come from published consensus trees, not the raw data originally used by each study. As such, we cannot model the statistical error associated with the robustness of each tree stemming either from missing data, or lack of information on the relationships between languages. The degree of statistical error likely varies widely across families and through time as sample sizes vary, and the older the phylogeny, the greater the likelihood for multiple sources of error to compound (i.e., incompleteness and uncertainty). Moreover, different language families have been subject to different levels of research intensity by different scholars resulting in unavoidable measurement error across the data set.

## Lineages through time

Lineages-through-time plots visualize the growth of lineages in a phylogeny [20]. Fig 2 shows the 10 phylogenies in our dataset that have topologies (colored lines) and the sum of all lineages across the 10 phylogenies (black line). Fig 2 shows that the diversification rate for the sum of all lineages $r = 0.001$ yr$^{-1}$ (regression: $F_{1,377} = 229{,}736.79$, $r^2 = 99.4\%$, $p < 0.001$). However, Fig 2 also shows that diversification rates vary both within and across language families. Some rates slow with time, others increase, while others remain stable. For example, diversification rates seem to decrease with time within the Austronesian, Pama-Nyugan, and Bantu lineages. Diversification rates in the Lezgic and Japonic lineages seem to increase through time, and while the sum of all lineages combined seems to be relatively constant, there is a slight decrease through time. Table 1 presents the diversification rates and doubling times for all lineages.

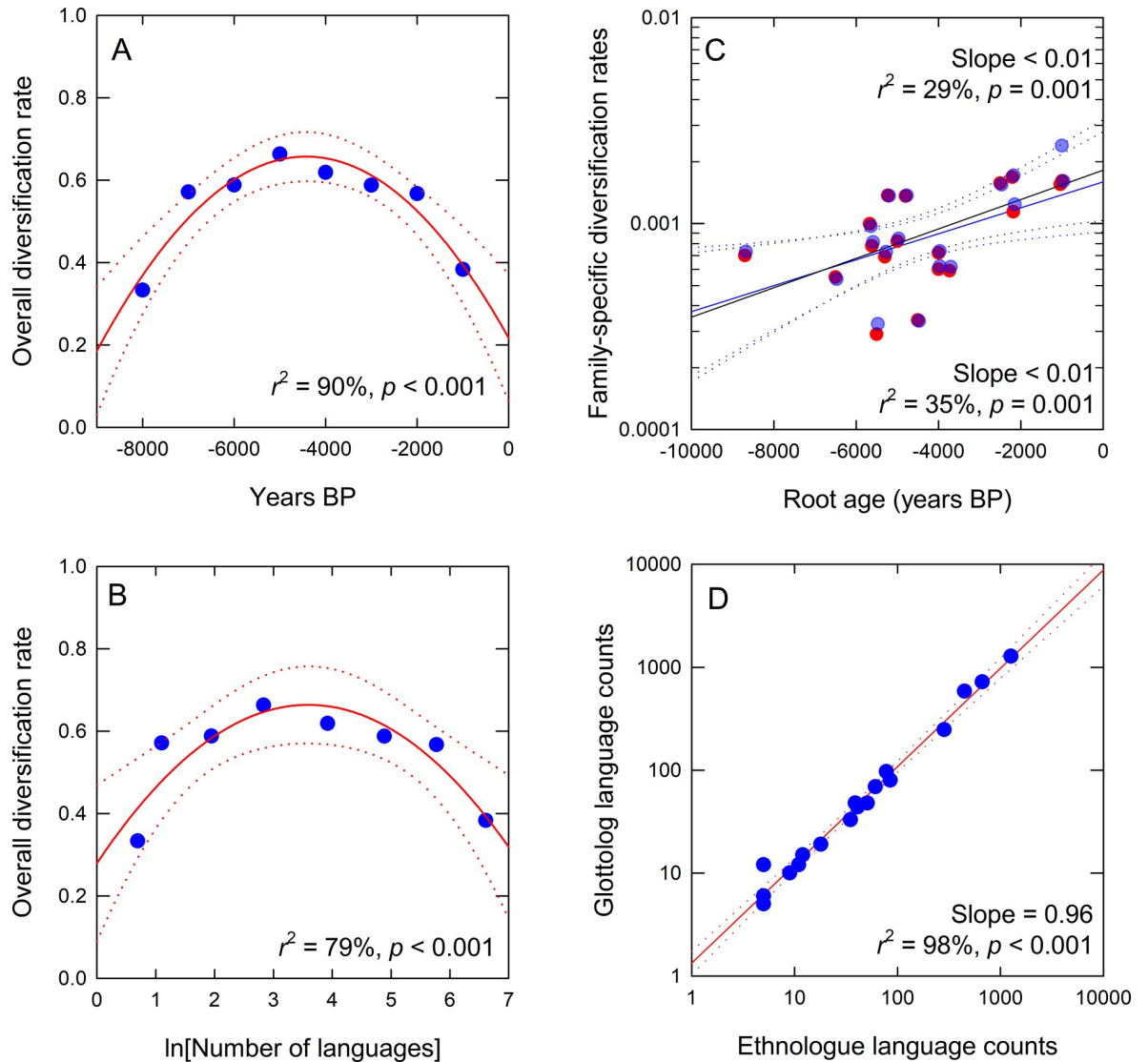## Sum of all lineage diversification rates

To examine how overall diversification rates change through time we estimate the change in the sum of all lineages per millennia by averaging the data over 1,000 year bins and calculating the change over time. While the black line in Fig 2 is relatively linear on a semi-log scale, Fig 3A shows that the diversification rate changes nonlinearly over time as it is well-fit by a quadratic function (regression: $F_8 = 22.69$, $r^2 = 90\%$, $p < 0.01$). Diversity initially increases (i.e., $b > d$) until an inflection ~5,000 BP after which diversity decreases toward the present (i.e., $b < d$). Fig 3B shows that the diversification rates are highly nonlinear with the number of languages, as shown by a quadratic function fit to the logged data, therefore capturing the skew (regression: $F_9 = 9.45$, $r^2 = 79\%$, $p = 0.02$): Diversification rates initially exhibit strong positive density-dependence, where diversification rates increase rapidly with low number of languages, reaching maximum rates at 12 languages, thereafter declining steadily with increasing languages, i.e., negative density-dependence.

## Diversification rates across lineages

Table 1 presents the diversification rates for all 18 lineages for which we have data. The mean rate of diversification across the full sample is 0.001 yr$^{-1}$ (bootstrapped 95% confidence interval = 0.0004–0.001), or a doubling time of ~770 years, similar to the estimated rate from the summed lineages. Using both the Ethnologue and Glottolog data, Fig 3C shows that the diversification rates of phylogenetic trees increases exponentially through time (regression: Ethnologue, $F_{18} = 12.6$, $r^2 = 43\%$, $p = 0.003$; Glottolog, $F_{18} = 16.4$, $r^2 = 49\%$, $p < 0.001$;) by ~20% per millennia, indicating that linguistic diversification is not constant through time. Results do not differ across the data sets, and Fig 3D shows that the data sets are highly correlated.

Linguistic diversification rates across families appear to increase by nearly an order of magnitude through time (Fig 3C). Taken at face value, diversification rates show an exponential increase over time with a slope of 0.001. As such, the doubling time of diversification rates is about 7,000 years. However, this time dependency must be, at least in part, due to an unavoidable "pull of the present" effect [20,21]. This is because language families that arose recently have had less time to undergo extinction than older families. Moreover, our sample does not include many of the world's smaller families and isolates. We can assume that these slow-growing languages would populate the lower right of Fig 3C to form a wedge shape to the graph and flatten the slope. For these reasons, an overall average diversification rate, as calculated in the previous paragraph, is a good estimate of the rate at which language spread and diversified around the world over the Holocene.
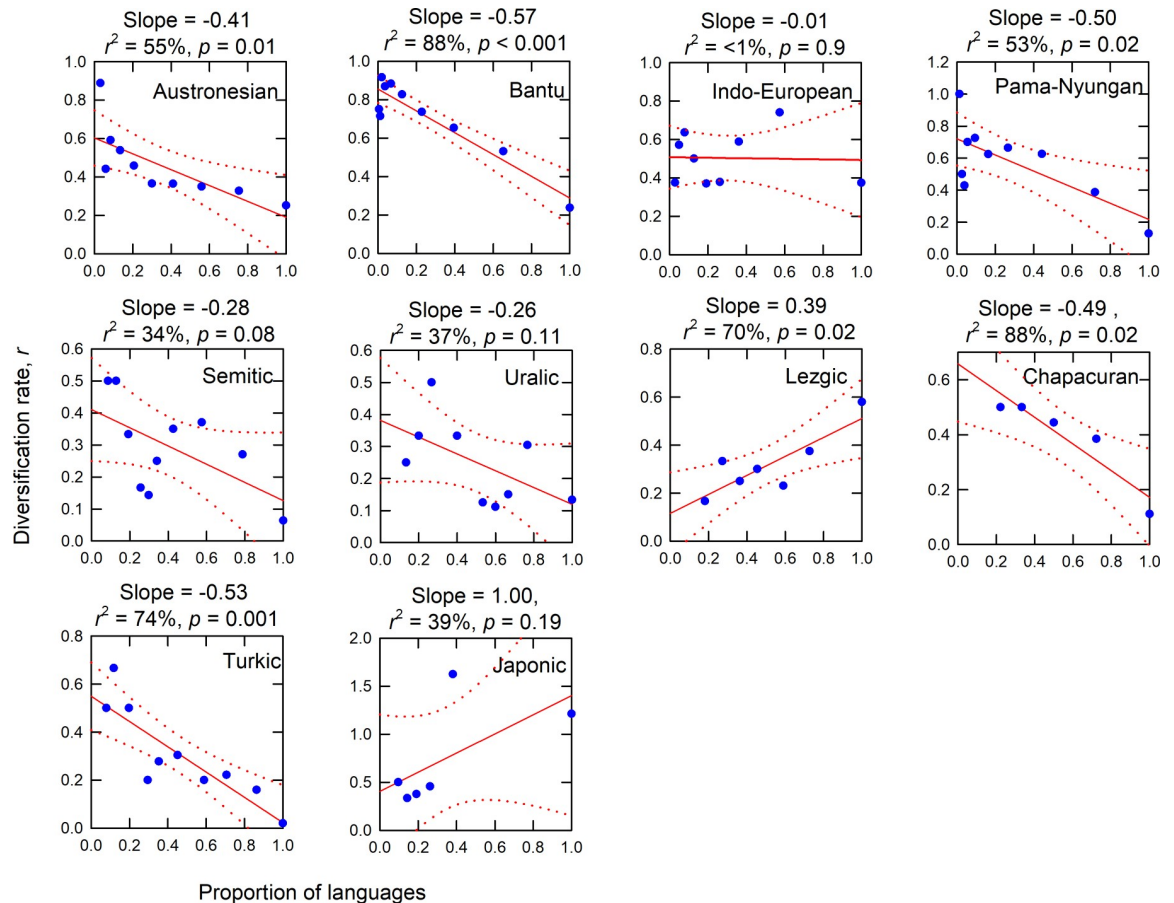
**Fig 3. Diversification and diversification rates within and across lineages.** A) Overall diversification rates over time fit with a quadratic function (red dotted lines are 95% confidence intervals around the slope). B) Overall diversification rates and the total number of languages, with a quadratic function fit to logged data. C) Diversification rates across lineages by their root ages fit by OLS regression. D) Correlation between the Glottolog and Ethnologue language estimates.

https://doi.org/10.1371/journal.pone.0213126.g003

## Diversification rates within lineages

To examine the nature of diversification rates within lineages we recognize three basic forms of potential density-dependence; exponential growth (i.e., constant diversification rates over time), slope = 0; negative density-dependence, slope < 0, where diversification rates slow over time; and positive density-dependence, slope > 0, where diversification rates increase over time. Fig 4 shows the same diversification rates for the nine lineages plotted as a function of the changing number of languages within lineages (statistics of each are given in the title of each panel). Seven of the ten lineages show evidence of negative density-dependence (i.e., slopes > 0): these are the Austronesian, Pama-Nyungan, Bantu, Turkic, Semitic, Uralic, and Chapacuran phylogenies. The Lezgic lineage and Japonic indicate positive density-dependence

**Fig 4. Different forms of density-dependent diversification rates among lineages.**

(slopes > 0). Indo-European exhibits exponential growth independent of density (slope = 0). S2 Fig in the online Supplementary Material show similar results by time.

## Discussion

Our results indicate that rates of linguistic diversification are not constant through time, neither as a whole, nor within most of the individual lineages. Overall rates of diversification across all lineages initially increase for the first few thousand years of the Holocene and thus exhibit positive density-dependence where the actual rate of diversification itself increases with each new language. After the inflection point ~5,000 BP, the diversification rate then decreases, thus exhibiting negative density-dependence, likely reflecting increased competition and saturation of populations on landscapes as languages expand. We find strong evidence for similar patterns within individual lineages where diversification rates are most rapid during the earliest stages of diversification (and at the smallest number of languages), indicating that as new lineages are born, they diversify in punctuated bursts. Lezgic exhibits positive density-dependence, where, similar to the initial stages of overall diversity in the early Holocene, the rate of diversification increases through time toward the present and continues to diversify (i.e., diversity begets diversity). Although the slopes of the Semitic and Uralic are negative, and Japonic is positive, they are not significantly different from 0 and so should be viewed as inconclusive sources of evidence of density-dependent diversification. Indo-European shows

no time-dependent trend and so diversification seems to be constant over time. A clear implication of these results is that empirically, the process of language diversification is not well captured by the Yule model, a commonly chosen prior in Bayesian phylogenies, which models diversification as a stochastic, constant branching process [49–52]. However, the extent to which tree structures are robust to the choice of priors is a matter of debate [52,53].

The average diversification rate $r_E$ of 0.001 $yr^{-1}$ over all lineages means that the doubling time of the world's linguistic diversity is on the order of ~700 years. This diversification is driven by the underlying exponential growth rate of populations as they expanded from their origin [8–11]. When measured ethnographically, growth rates of natural-fertility human populations are often in the range of $r \sim$ 2–4% $yr^{-1}$ [54,55]. Human populations therefore have the capacity to grow 20–40 times faster than languages diversify. However, the long-term growth rate of global human populations over the last 6,000 years, up until the industrial revolution, is also on the order of 0.001 $yr^{-1}$ [56–58]. This suggests that language diversification and population growth rate on a global level are closely coupled and at near equilibrium levels (i.e., growth rates are only slightly above zero).

It might be the case that competition for limited resources forces local populations to diverge ethnolinguistically, resulting in some positive density dependence. Like species, linguistic diversification experiences punctuated bursts as languages go through splitting events [59,60]. Rates of new word gain are also faster in larger populations [61,62], although others have argued for faster linguistic evolution in smaller populations [63]. A combination of larger populations and more competition could therefore fuel continuing ethnolinguistic diversification, but it remains an open question as to whether competition alone is sufficient to account for this dynamic. Islands are ideal test cases to establish whether ethnolinguistic diversity continues to increase once equilibrium population sizes have been reached. Indeed, linguistic diversity on islands has been shown to increase with time since initial settlement [61,64].

Language families spread and diversified over the Holocene at nonlinear, density-dependent rates. Like other forms of biodiversity, ethnolinguistic diversity may have continued to expand were it not for recent extinction events associated with globalization [65–68]. Interestingly, the data we present here suggests that the current language extinction crisis is part of a larger process of linguistic replacement that began in the mid-Holocene, ~5,000 years ago with the expansion of agricultural language families, now driven by the expansion of global languages central to the expansion of global trade networks.

## Material and methods

### Linguistic phylogenies

We searched for the keywords "Bayesian linguistic phylogeny" to find studies that use Bayesian phylogenetic methods to attach known calendar dates to nodes of language trees as a means to estimate root ages. We found 19 such linguistic phylogenies (Table 1). Of these we were able to obtain 10 of the actual consensus tree files with the highest posterior probabilities either from online supplementary material or by contacting corresponding authors. Unfortunately, we were not able to obtain error estimates for the consensus trees from the data made available.

Character-based phylogenetic methods using cognate codings of basic vocabulary words or phonetic data are useful for inferring the internal classifications and divergence dates of recent language family expansions. These methods are superior to traditional techniques in several ways by allowing for different rates of change between cognate sets and between different lineages, and by explicitly taking into account available archaeological dates and historical events to infer divergence dates [69]. While alternative phylogenetic methods exist [70–72], there is

no competing alternative for our purposes because our analyses require a systematic approach that produces time-dated phylogenies.

The three largest language families in our sample are Austronesian [16], Bantu [17], and Indo-European [18]. Of these Indo-European has proven the most divisive with some arguing for the Kurgan expansion with homelands in the Pontic steppe dating 6,000 years ago versus others supporting the Anatolian expansion of farming around 8,700 years ago [10,18,68]. We opt for the latter following our inclusion criteria of Bayesian methods. Five families in our sample are Eurasiatic language families with estimated dates taken from Pagel and colleagues' study of ultraconserved words [38]. These families are some of the oldest in our sample, and therefore generally considered the most controversial, but removing them from the sample yields only a slightly faster average diversification rate of 0.00104 yr$^{-1}$. The remaining language families in the sample are more recent expansions (Table 1), hence their divergence dates are likely more accurate.

## Analyses

Lineage-through-time plots were created with the *ape* package in R. The sum of the lineages line in Fig 2 is created by simply aligning the 10 available phylogenies together using absolute time as in the graph and adding all of the lineages together. How exactly each phylogeny relates to others at the basal node is not necessary information for the lineage-through-time plot since each additional lineage is simply counted at each point in time.

For within-family analyses of diversification rates as a function of time, we used the *R* package *TESS* [36]. This is a stochastic-branching model that flexibly fits time-varying speciation and extinction rates and accounts for incomplete sampling [37]. See Supplementary Material for complete details, *R* code, and tree data. The assumption of constant diversification rate *r* is a reasonable assumption for some of these language families so we made this assumption for each family in the full sample of 19. Diversification rates were estimated with Eq 1 using the total (natural-logged) number of languages in a family from the Ethnologue [1] and the Glottolog [2], along with the estimated root age of the tree from the original studies (Table 1). Doubling times were calculated as in Eq 3 [24]. While 4 of the phylogenies include dialects, we calculate diversification with counts of languages only.

## Supporting information

**S1 Fig. Standardized proportion of languages and phylogenetic age for the language families used in the paper.** These are the data standardized by the number of languages in a language family and the longevity of each language family.
(JPG)

**S2 Fig. Diversification rates as a function of normalized time for each language family.** These panels show the diversification rates within each language family over standardized time (i.e., the relative longevity of each language family).
(TIF)

**S1 File. The data and R code used in the paper.** All data and code used in this paper are available in this README.txt file.
(TXT)

## Author Contributions

**Conceptualization:** Marcus J. Hamilton, Robert S. Walker.

**Writing – original draft:** Marcus J. Hamilton, Robert S. Walker.

# References

1. Gordon RG. Ethnologue: languages of the world. Dallas, TX: SIL International; 2005.

2. Hammarström H, Forkel R, Haspelmath M. Glottolog 3.4. Jena: Max Planck Institute for the Science of Human History; 2019.

3. Wilson DE, Reeder DM. eds. Mammal species of the world: a taxonomic and geographic reference. Baltimore, MD: Johns Hopkins University Press; 2005.

4. Hammarström H. A full-scale test of the language farming dispersal hypothesis. Diachronica 2010; 27, 197–213.

5. Renfrew C. World linguistic diversity. Scientific American 1993; 270, 104–111.

6. Bellwood P. The dispersals of established food-producing populations. Current Anthropology 2009 5, 621–626.

7. Diamond J, Bellwood P. Farmers and their languages: the first expansions. Science 2003; 300, 597–603. https://doi.org/10.1126/science.1078208 PMID: 12714734

8. Gavin MC, Rangel TF, Bowern C, Colwell RK, Kirby KR, Botero CA et al. Process-based modelling shows how climate and demography shape language diversity Global Ecol and Biogeog 2017; 26(5), 584–591

9. Derungs C, Köhl M, Weibel R, Bickel B. Environmental factors drive language density more in food-producing than in hunter–gatherer populations. Proc Roy Soc B 2018; 285(1885), 20172851

10. Currie TE, Mace R. Political complexity predicts the spread of ethnolinguistic groups Proc Natl Acad Sci, 2009; 106(18), 7339–7344. https://doi.org/10.1073/pnas.0804698106 PMID: 19380740

11. Nichols J. Linguistic diversity and the first settlement of the New World. Language 1990; 66,475–521.

12. Nettle D. Linguistic diversity of the Americas can be reconciled with a recent colonization. Proc Natl Acad Sci USA 1999; 96, 3325–3329. https://doi.org/10.1073/pnas.96.6.3325 PMID: 10077683

13. Pagel M. The history, rate and pattern of world linguistic evolution. In: Knight C, Hurford JR, Studdert-Kennedy M, editors. The evolutionary emergence of language: social function and the origins of linguistic form. Cambridge: Cambridge University Press. 2000. pp. 391–416.

14. Foley R. The evolutionary ecology of linguistic diversity in human populations. In: Jones M, editor. Traces of ancestry: studies in honour of Colin Renfrew, Cambridge: McDonald Institute. 2004. pp. 61–71.

15. Morlon H, Potts MD, Plotkin JB. Inferring the dynamics of diversification: a coalescent approach. PLoS Biol 2010; 8, e1000493. https://doi.org/10.1371/journal.pbio.1000493 PMID: 20927410

16. Gray RD, Drummond AJ, Greenhill SJ. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. Science 2009; 323,479–483. https://doi.org/10.1126/science.1166858 PMID: 19164742

17. Grollemund R, Branford S, Bostoen K, Meade A, Venditti C, Pagel M. Bantu expansion shows that habitat alters the route and pace of human dispersals. Proc Natl Acad Sci USA 2015; 112, 13296–13301. https://doi.org/10.1073/pnas.1503793112 PMID: 26371302

18. Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, et al. Mapping the origins and expansion of the Indo-European language family. Science 2012; 337,957–960. https://doi.org/10.1126/science.1219669 PMID: 22923579

19. Thieme HR. Mathematics in population biology. Princeton NJ: Princeton University Press; 2003.

20. Nee SC, May RM, Harvey P. The reconstructed evolutionary process. Philos Trans Roy Soc London Ser B 1994; 344, 305–311.

21. Nee SC, Holmes EC, May RM, Harvey PH. 1994 Extinction rates can be estimated from molecular phylogenies. Philos Trans Roy Soc London Ser B 1994; 344, 77–82.

22. Ricklefs RE. Reconciling diversification: Random pulse models of speciation and extinction. The American Naturalist 2014; 184, 268–276. https://doi.org/10.1086/676642 PMID: 25058286

23. Mayr E. Ecological factors in speciation. Evolution 1947; 1, 263–288.

24. Weir JT. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. Evolution 2006; 60, 842–855. PMID: 16739464

25. Phillimore AB, Price TD. Density-dependent cladogenesis in birds. PLoS Biol 2008; 6, e71. https://doi.org/10.1371/journal.pbio.0060071 PMID: 18366256

26. Rabosky DL, Lovette IJ. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution 2008; 62, 1866–1875. https://doi.org/10.1111/j.1558-5646.2008.00409.x PMID: 18452577

27. Rabosky DL. Diversity-dependence, ecological speciation, and the role of competition in macroevolution. Annu Rev Ecol Evol Syst 2013; 44, 481–502.

28. Rundell RJ, Price TD. 2009 Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. Trends Ecol Evol 24, 394–399. https://doi.org/10.1016/j.tree.2009.02.007 PMID: 19409647

29. Lande R. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am Nat 1993; 142, 911–927. https://doi.org/10.1086/285580 PMID: 29519140

30. Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, et al. The delayed rise of present-day mammals. Nature 2007; 446, 507–512. https://doi.org/10.1038/nature05634 PMID: 17392779

31. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds in space and time. Nature 2012; 491, 444–448. https://doi.org/10.1038/nature11631 PMID: 23123857

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

33. Morlon H, Parsons TL, Plotkin JB. Reconciling molecular phylogenies with the fossil record. Proc Natl Acad Sci USA 2011; 108, 16327–16332. https://doi.org/10.1073/pnas.1102543108 PMID: 21930899

34. Soria-Carrasco V, Castresana J. Patterns of mammalian diversification in recent evolutionary times: global tendencies and methodological issues. J Evol Biol 2011; 24, 2611–2623. https://doi.org/10.1111/j.1420-9101.2011.02384.x PMID: 21955145

35. Stadler T. Mammalian phylogeny reveals recent diversification rate shifts. Proc Natl Acad Sci USA 2011; 108, 6187–6192. https://doi.org/10.1073/pnas.1016876108 PMID: 21444816

36. Höhna S, May MR, Moore BR. TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. Bioinformatics 2016; 32, 789–791. https://doi.org/10.1093/bioinformatics/btv651 PMID: 26543171

37. May MR, Höhna S, Moore BR. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. Methods Ecol Evol 2016; 7, 947–959.

38. Pagel M, Atkinson QD, Calude AS, Meade A. Ultraconserved words point to deep language ancestry across Eurasia. Proc Natl Acad Sci USA 2013; 110(21), 8471–8476. https://doi.org/10.1073/pnas.1218726110 PMID: 23650390

39. Atkinson QD. From species to languages: a phylogenetic approach to human prehistory. PhD Dissertation. Auckland NZ: University of New Zealand; 2006.

40. Kassian A. Towards a formal genealogical classification of the Lezgian languages (North Caucasus): testing various phylogenetic methods on lexical data. PloS ONE 2015 10, e0116950. https://doi.org/10.1371/journal.pone.0116950 PMID: 25719456

41. Honkola T, Vesakoski O, Korhonen K, Lehtinen J, Syrjänen K, Wahlberg N. Cultural and climatic changes shape the evolutionary history of the Uralic languages. J Evol Biol 2013 26, 1244–1253. https://doi.org/10.1111/jeb.12107 PMID: 23675756

42. Dunn M, Kruspe N, Burenhult N. Time and place in the prehistory of the Aslian languages. Hum Biol 2013; 85, 383–400. https://doi.org/10.3378/027.085.0318 PMID: 24297234

43. Kitchen A, Ehret C, Assefa S, Mulligan CJ. Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. Proc. R. Soc. B 2009; 276, 2703–2710. https://doi.org/10.1098/rspb.2009.0408 PMID: 19403539

44. Dunn M, Greenhill SJ, Levinson SC, Gray RD. Evolved structure of language shows lineage-specific trends in word-order universals. Nature 2011; 473, 79–82. https://doi.org/10.1038/nature09923 PMID: 21490599

45. Lee S, Hasegawa T. Oceanic barriers promote language diversification in the Japanese Islands. Journal of Evolutionary Biology 2014; 27, 1905–1912. https://doi.org/10.1111/jeb.12442 PMID: 24953224

46. Birchall J, Dunn M, Greenhill S. A combined comparative and phylogenetic analysis of the Chapacuran language family. International Journal of American Linguistics 2016; 82, 255–84.

47. Michael L, Chousou-Polydouri N, Keith B, Donnelly E, Meira S, Wauters V, et al. Bayesian phylogenetic classification of Tupí-Guaraní. Línguas Indígenas Americanas 2015; 15, 193–221.

48. Hruschka DJ, Branford S, Smith ED, Wilkins J, Meade A, Pagel M, et al. Detecting regular sound changes in linguistics as events of concerted evolution. Current Biology 2015; 25, 1–9. https://doi.org/10.1016/j.cub.2014.10.064 PMID: 25532895

49. Steel M, McKenzie A. Properties of phylogenetic trees generated by Yule-type speciation models. Mathematical Biosciences. 2001; 170, 91–112. https://doi.org/10.1016/S0025-5564(00)00061-4 PMID: 11259805

50. Aldous DJ. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. Statist Sci. 2001; 16, 23–34. https://doi.org/10.1214/ss/998929474

51. Pinsky M, Karlin S. An introduction to stochastic modeling. Academic press; 2010.

52. Blum MGB, François O. Which Random Processes Describe the Tree of Life? A Large-Scale Study of Phylogenetic Tree Imbalance. Syst Biol. 2006; 55, 685–691. https://doi.org/10.1080/10635150600889625 PMID: 16969944

53. Sarver BAJ, Pennell MW, Brown JW, Keeble S, Hardwick KM, Sullivan J, et al. The choice of tree prior and molecular clock does not substantially affect phylogenetic inferences of diversification rates. PeerJ 2019; 7, e6334. https://doi.org/10.7717/peerj.6334 PMID: 30886768

54. Lutz W, Qiang R. Determinants of human population growth. Phil Trans R Soc Lond B 2002; 357, 1197–1210.

55. Hamilton MJ, Walker RS, Kesler DC. Crash and rebound of indigenous populations in lowland South America. Scientific Reports 2014; 4, 4541. https://doi.org/10.1038/srep04541 PMID: 24687096

56. Thomlinson R. Demographic problems: controversy over population control. Encino, CA: Dickenson Publishing Company; 1975.

57. McEvedy C, Jones R. Atlas of world population history. Harmondsworth, UK: Penguin Books Ltd, 1978.

58. Klein Goldewijk K, van Drecht G. HYDE 3: current and historical population and land cover. In: Bouwman AF, Kram T, Klein Goldewijk K, editors. Integrated modelling of global environmental change: An overview of IMAGE 2.4, Bilthoven, Netherlands: Netherlands Environmental Assessment Agency. 2006. pp. 93–111.

59. Atkinson QD, Meade A, Venditti C, Greenhill SJ, Pagel M. Languages evolve in punctuational bursts. Science 2008; 319, 588–588. https://doi.org/10.1126/science.1149683 PMID: 18239118

60. Thurston WR. Processes of change in the languages of north-western New Britain. Canberra: Australian National University; 1987.

61. Bromham L, Hua X, Fitzpatrick TG, Greenhill SJ. Rate of language evolution is affected by population size. Proc Natl Acad Sci USA 2015; 112, 2097–2102. https://doi.org/10.1073/pnas.1419704112 PMID: 25646448

62. Wichmann S, and Holman EW. Population Size and Rates of Language Change. Hum Biol 2009; 8(2+3).

63. Nettle D. Is the rate of linguistic change constant? Lingua 1999; 108, 119–136.

64. Gavin MC, Sibanda N. The island biogeography of languages. Global Ecology and Biogeography 2012; 21, 958–967.

65. Whalen DH, Simons GF. Endangered language families. Language 2012; 88,155–173.

66. Maffi L. Linguistic, cultural, and biological diversity. Annu Rev Anthropol 2005; 34,599–617.

67. Krauss M. The world's languages in crisis. Language 1992; 68, 4–10.

68. Nettle D, Romaine S. Vanishing voices: the extinction of the world's languages. Oxford: Oxford University Press; 2000.

69. Greenhill SJ, Gray RD. Austronesian language phylogenies: myths and misconceptions about Bayesian computational methods. In: Adelaar A, Pawley A, editors. Austronesian historical linguistics and culture history: a festschrift for Robert Blust. Canberra: Pacific Linguistics. 2009. pp. 375–397.

70. Donohue M, Denham T. Farming and language in island Southeast Asia: reframing Austronesian history. Curr Anthropol 2010 51, 223–256.

71. McMahon A, McMahon R. Language classification by numbers. Oxford: Oxford University Press; 2005.

72. Nakhleh L, Warnow T, Ringe D, Evans SN. A comparison of phylogenetic reconstruction methods on Indo-European dataset. Transactions of the Philological Society 2005; 103, 171–192.

73. Kolipakam V, Jordan FM, Dunn M, Greenhill SJ, Bouckaert R, Gray RD, et al. A Bayesian phylogenetic study of the Dravidian language family. Royal Society Open Science 2018 5(3), 171504. https://doi.org/10.1098/rsos.171504 PMID: 29657761

74. Bouckaert RR, Bowern C, Atkinson QD. The origin and expansion of Pama–Nyungan languages across Australia. Nat Ecol Evol 2018; 2, 741–749. https://doi.org/10.1038/s41559-018-0489-3 PMID: 29531347