Research

Check for updates

Author for correspondence:
Menghan Zhang
e-mail: mhzhang@fudan.edu.cn

THE ROYAL SOCIETY
PUBLISHING

Phylogenetic insight into the origin of tones

Baihui Wu[1], Hanzhi Zhang[5] and Menghan Zhang[2,3,4]

[1]State Key Laboratory of Genetic Engineering, Center for Evolutionary Biology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200438, People's Republic of China
[2]Institute of Modern Languages and Linguistics, Fudan University, Shanghai 200433, People's Republic of China
[3]Research Institute of Intelligent Complex Systems, Fudan University, Shanghai 200433, People's Republic of China
[4]Human Phenome Institute and Ministry of Education Key Laboratory of Contemporary Anthropology, Department of Anthropology and Human Genetics, School of Life Sciences, Fudan University, Shanghai 200438, People's Republic of China
[5]Department of Anthropology, University College London, London WC1H 0BW, UK

(iD) BW, 0000-0002-2555-5171; HZ, 0000-0003-1023-8007; MZ, 0000-0002-3505-862X

The origin of tone, also known as *tonogenesis,* has long been a topic of great interest in language evolution and human cognition studies. Several linguistic studies of tonal languages have proposed various hypotheses that tonal origin may be related to different changes of phonological structures. However, such hypotheses have not been quantitatively tested in an evolutionary framework. Here, we conducted phylogenetic comparative analyses to assess the likelihood of different hypotheses of tonogenetic mechanisms across 106 Sino-Tibetan languages, of which approximately 70% are tonal. Our results showed that the presence of tones has a strong phylogenetic pattern and that Proto-Sino-Tibetan languages were most likely non-tonal. Our findings identified that tonal origin was strongly associated with the evolution of specific phonological structures, such as the loss of syllable-final consonants and voice quality on vowels. Furthermore, we found that tonal origin probably did not influence the diversification rates of Sino-Tibetan languages. These findings enabled us to better understand that tone arose as a compensatory mechanism for the structural organization and evolution of languages.

1. Introduction

Human language can be regarded as a complex self-organized system. The nature of its self-organization is demonstrated by the correlated evolution of linguistic properties such as syntactic optimization [1], word formation [2] and sound change [3]. These correlated evolutionary processes are generally motivated by human communicative demands and cognitive development [1,4–6] and shape the structural diversity of languages. Understanding the diversity of linguistic structures and the driving forces of their evolution would provide important insights into language evolution [7]. Worldwide, spoken languages vary considerably in the structures of sound inventories. Every language has phonological inventories of consonants and vowels, but not all contain tones [8–10]. The tone is the use of pitch registers and/or contours to distinguish the lexical meanings and the grammatical forms of words [11,12]. Tones are used to convey attitudinal and emotional information during the communication process [10]. More than half of the approximately 7000 worldwide languages are characterized by contrastive tones [8,13]. Most tonal languages are geographically concentrated in East and Southeast Asia and include Sino-Tibetan, Kra-Dai, Hmong-Mien, Austroasiatic and Austronesian languages [8] (figure 1*a*). From the perspective of historical linguistics, the evolution of tones generally comprises two phases: tonal origin and tonal split. In particular, tonal origin is the process in which tonal contrast emerged from earlier non-tonal contrast (i.e. *tonogenesis* [15,16]), while tonal split is the
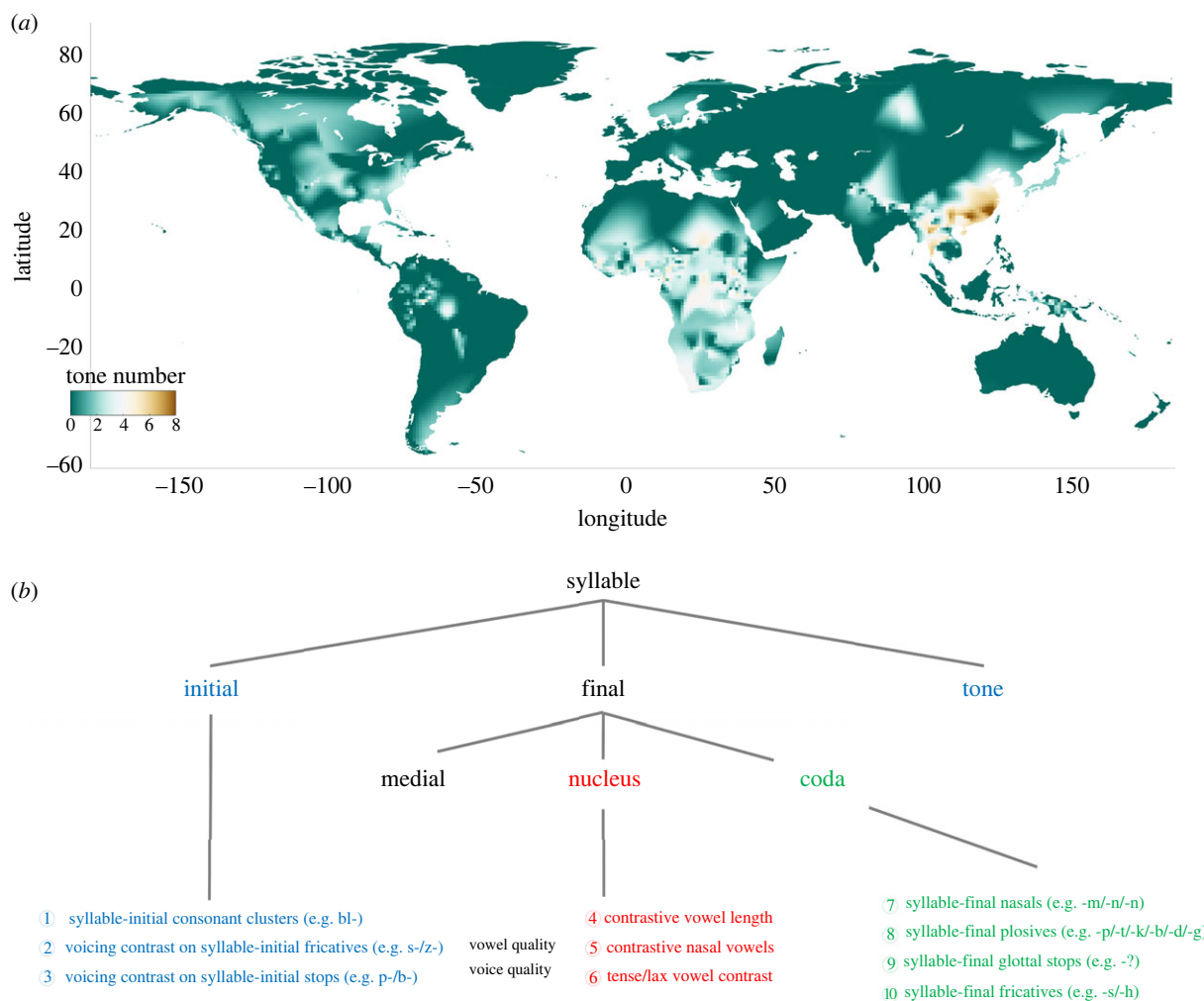
Figure 1. The geographical distribution of tonal contrasts and syllable structures. (*a*) The geographical heatmap of the tonal complexity in the world. The colour bar shows the number of tones of 3756 language samples drawn from Everett *et al.*'s work [14]. Languages with more than eight tones are counted as eight. (*b*) Diagram of a syllable structure in Sino-Tibetan languages. The colours show the models related to the tonal origin grouped by the phonological components of syllable-initial consonants (blue), vowel and voice quality at the syllabic nucleus (red) and syllable-final consonants (green).

process in which the number of tonal contrasts multiplied after the tonal origin. Recent comparative studies have advanced our understanding of the preconditions of tonal splits [10,17]. However, large-scale quantitative studies of tonal origin remain limited due to the lack of comprehensive historical accounts of languages [17]. Note that for tonogenesis, the concept of the tonal origin was equivalent to the emergence of early tonal contrast in this study. A glossary box containing the key linguistic terms can be found in the electronic supplementary material.

Traditional linguistic surveys have qualitatively described diverse tonogenetic potentials that interact with tones [17]. More specifically, phonetic manifestations of tone interact with various phonological factors and other aspects of linguistic properties, such as syllabic structures [18]. Based on the review by Michaud & Sands [17], these tonogenetic potentials can be categorized into four different components of syllabic structure: syllable-final consonants, vowel and voice quality at the syllabic nucleus, and syllable-initial consonants (figure 1*b*). Among them, the most common one related to tonal origin is the loss of syllable-final consonants [16,17,19–25]. In the typical case of Vietnamese, the loss of syllable-final consonants historically triggered the development of a three-way tonal contrast of rising tone, falling tone and level tone [26]. The contributions of syllable-final consonants

to tonal origin can also be observed in some Sino-Tibetan languages (e.g. Chinese [20,24,27] and Tibetan [28]). Alternatively, the loss of voicing contrast in syllable-initial consonants could be another driving factor to arise in the tonal contrast of language [21,22,29]. Commonly, voiced initial consonants are phonetically associated with low tone, while voiceless consonants are associated with high tone [8,22,30]. Note that the loss of voicing contrasts in syllable-initial consonants can increase the number of tones if a language has already produced tonal contrast from some other sources, such as syllable-final consonants [26]. Moreover, there are intriguing typological characteristics observed in Tibetan languages, in that non-tonal languages are typically characterized by a complicated set of initial consonant clusters, whereas tonal languages do not comprise consonant clusters [31,32]. This suggests that the reduction in complexity of syllabic structure could have a significant influence on tonal origin in Sino-Tibetan languages. In summary, the consonant-based account of tonal origin can be frequently observed in East and Southeast Asian languages and hence be suggested as a common model of tonal origin [20,24,33].

In contrast with consonant-based accounts, other hypotheses of tonogenetic mechanisms suggest that tone arises from vocalic features such as the syllabic nucleus or whole-syllable features such as phonation types. Accordingly, there

are three tonogenetic potentials: vowel length (e.g. long and short vowels) [34], vowel quality (e.g. vowel nasalization) and voice quality (e.g. tense/lax vowel contrasts) of the whole syllable [21]. In particular, vocalic phonation with laryngeal contrasts highlights the tonogenetic attribution of voice quality contrasts of the whole syllable [21]. This phonation may also mediate tonogenetic attributions of syllable-initial consonants [21]. Accordingly, voice quality has been argued to play a central role in tonal origin. Synthesizing these tonogenetic potentials mentioned above, we can roughly outline that the origin of early tonal contrast was induced by the evolved structures of syllables and their phonological systems of languages.

Despite substantial progress in discovering various tonogenetic potentials, it is difficult to formally assess their contributions during the evolutionary process of tonal languages. In linguistics, the historical comparative method is a traditional approach to reconstructing the earlier state of a language based on comparing many words or other linguistic traits across different languages. However, its accuracy and reliability are limited by incomplete language surveys and discontinuous historical records of languages, predominantly in East Asian languages [35]. These limitations present a significant challenge in the large-scale linguistic comparisons of East and Southeast Asian languages such as Sino-Tibetan languages. Alternatively, recent advances in phylogenetic comparative methods (PCMs) provide new opportunities to overcome these limitations. These methods can incorporate phylogenetic uncertainty, control for the effects of shared ancestry and identify evolutionary patterns and generalizations across different taxonomic groups. PCMs have been used to assess several known evolutionary hypotheses of languages worldwide, such as reassessing the coevolutionary changes in linguistic traits [1,6], estimating evolutionary rates [2,36–38] and reconstructing ancestral states [39,40]. Furthermore, PCMs have been applied to cultural evolution, such as social evolution [41], human belief [42], political complexity [43] and music evolution [44].

In this study, we used Bayesian phylogenetic approaches to assess different hypotheses of tonogenetic mechanisms across 106 Sino-Tibetan languages. The Sino-Tibetan language family is the second-largest language family in the world, spoken by approximately 1.5 billion native speakers [45]. Its evolutionary history has recently received wide attention [46–48]. This language family has rich tonal inventories, and approximately 70% of Sino-Tibetan languages are tonal (figure 2). We compiled binary-coded data on the presence/absence of tonal contrast and 10 other linguistic traits that have been reported as tonogenetic potentials comprising syllable-initial and syllable-final consonants, vowel length, vowel quality and voice quality (figure 2; electronic supplementary material, dataset 'Data.xlsx'). We used time-calibrated phylogenies of Sino-Tibetan languages derived from our previous work [46]. Using Bayesian PCMs, we reconstructed the ancestral states of linguistic traits in the Proto-Sino-Tibetan language and investigated the correlated evolution of all 11 binary-coded linguistic traits. We then used Bayes factors (BF) to quantify the evolutionary contributions of tonogenetic potentials. Furthermore, we used speciation and extinction model to examine the effects of each linguistic trait on the diversification rate of Sino-Tibetan languages, especially for the tonal contrast. Here, we assumed that the origin of tonal contrast has little effect on the language diversification rate of Sino-Tibetan languages, owing to

tonogenesis serving as the compensatory mechanism for the phonological simplification [16,51].

2. Materials and methods

(a) Sino-Tibetan language phylogeny and linguistic traits

To examine the tonogenetic models in Sino-Tibetan languages, we performed phylogenetic comparative analyses given phylogenetic trees and tonogenesis-related linguistic traits. We adopted the Sino-Tibetan phylogenies from our previous work [46], which contained 109 language samples across a vast region of East Asia and mainland Southeast Asia. According to the 109 names of these language samples, we manually compiled a total of 11 linguistic traits taken from 10 source bibliographies [49,52–60] (figure 1b; details in the electronic supplementary material, dataset 'Data.xlsx'). These traits were tonal contrast, syllable-initial consonant clusters, voicing contrast on syllable-initial fricatives, voicing contrast on syllable-initial stops, contrastive nasal vowels, contrastive vowel length, tense/lax vowel contrast, syllable-final nasals, syllable-final plosives, syllable-final glottal stops and syllable-final fricatives. We manually coded the presence or absence of each linguistic trait from the descriptions of the phonological system for each language sample, and the dataset was repeatedly checked by the authors. These traits were coded into binary states in which 1 represents a trait that is present in the language and 0 represents a trait that is absent. Note that these traits have been reported to be related to tonal origin in previous linguistic studies [8,17,21]. After quality control, we excluded three languages, Matu, Ngmüün and Marma, which showed proportions of missing data on traits over 50%. Finally, we retained the 106 Sino-Tibetan language samples and 11 linguistic traits. We pruned the original tree structure of 109 Sino-Tibetan languages provided in the previous study [46]. In the previous study, the phylogenies were reconstructed by BEAST2 software (v. 2.4.8) [61] with the Babel package (URL: https://github.com/rbouckaert/Babel). Nine hundred and forty-nine lexical root-meanings were used for phylogenies reconstruction according to the items in the Swadesh 100-word list [62]. The best-fitting model of phylogenetic reconstruction is the covarion model, and we chose one rate category for the gamma rate heterogeneity. Ten thousand Bayesian posterior sampling trees have been generated. Considering the convergence of Markov chain Monte Carlo (MCMC), we retained the final 900 Bayesian posterior sampling trees and generated the consensus tree of 106 Sino-Tibetan languages for further analyses using BayesTrees program (http://www.evolution.reading.ac.uk/BayesTrees.html).

(b) Phylogenetic signals of linguistic traits

Phylogenetic signals show the tendency for closely related languages to display similar linguistic trait values as a consequence of their phylogenetic proximity [63]. Here, we calculated the D-statistic to examine the phylogenetic signals of each linguistic trait on the consensus tree [64]. The common rationale of the D-statistic is that the linguistic traits in phylogenetically related languages should be more similar to each other than those drawn randomly from the given phylogeny. In particular, the D-statistic indicator aims to measure the strength of phylogenetic signals for binary-coded traits given a phylogeny. It scales the observed sum of sister-clade differences by the mean values of the distributions of Brownian motion and random processes [64]. A D-statistic close to 1 indicates a random process of trait evolution, whereas a D-statistic close to 0 indicates that trait evolution follows Brownian motion. A negative D-statistic represents the conservatism of the traits. To obtain the statistical
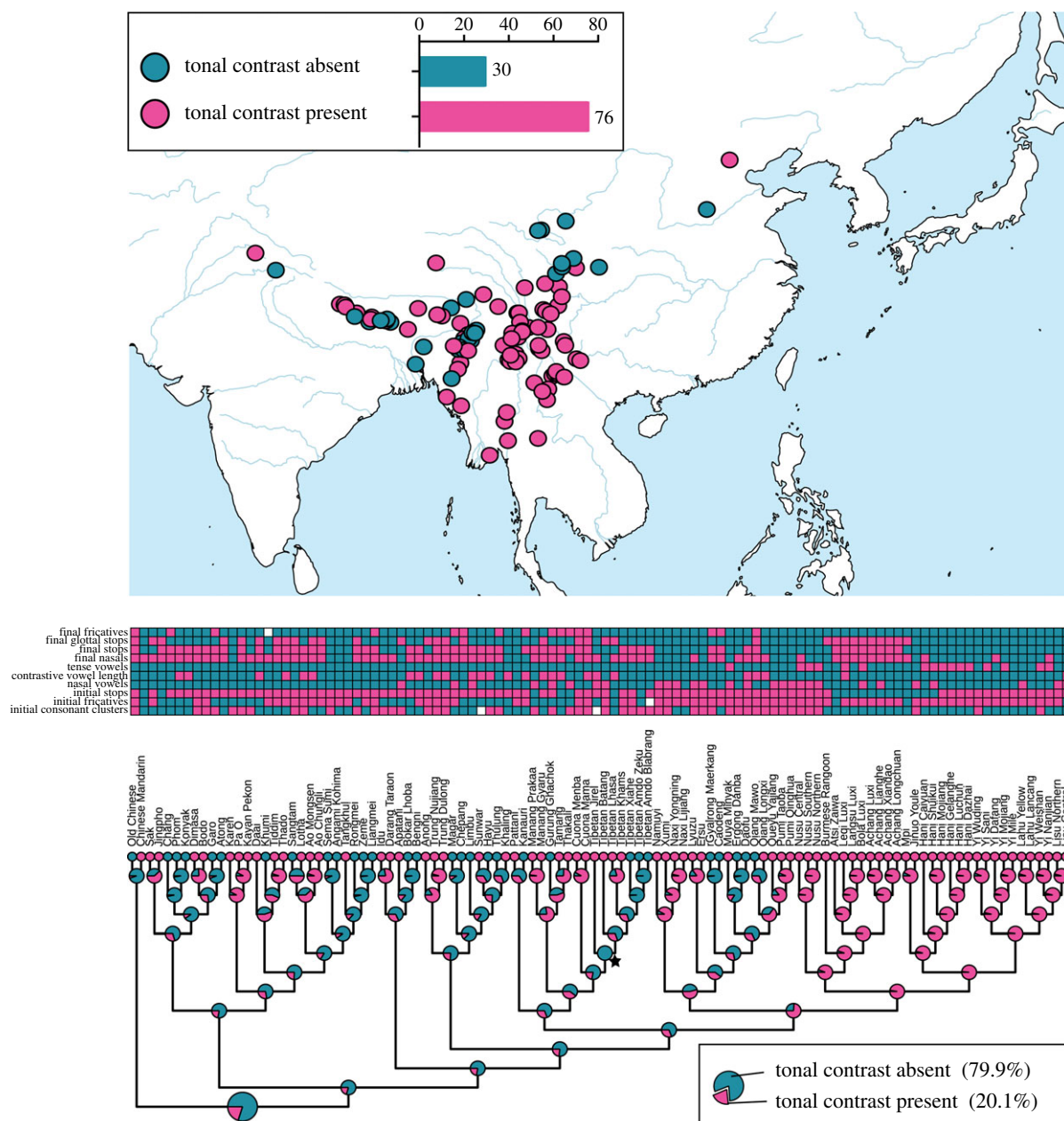
Figure 2. Phylogenetic relationships and geographical locations of 106 Sino-Tibetan languages. Tips on the ultrametric tree denote the presence (hot-pink) or absence (aqua) of the extant tonal contrast. Pie charts at the internal nodes indicate proportional levels of its reconstructed ancestral states. The ancestor of the Tibetan subclade (marked with a star) was fixed to be non-tonal using the Fossil command, based on the consensus from historical documents [49,50]. The squares represent the presence or absence of the other 10 linguistic traits, and the blank refers to the unknown state. The approximate locations of these languages are depicted on the map. Note that it is generally believed that the Old Chinese did not have tone systems. The histogram illustrates the number of tonal languages in the Sino-Tibetan languages.

significance of this indicator, we employed the permutation test of which the number is 1000. The calculation of the D-statistic is implemented by the *phylo.d()* function in the caper package in R.

(c) Phylogenetically and geographically distributed properties of linguistic traits

To compare the phylogenetic and geographical patterns of linguistic traits, we estimated two statistical indicators named the consistency index (CI) [65] and the retention index (RI) [66], which are indices of the fitness of a trait to a tree. For the phylogenetic tree, we used the consensus tree of 106 Sino-Tibetan languages. For the geographical tree, we calculated the great circle distance between pairwise languages using their geographical coordinates of latitude and longitude and then established a geographical tree using the neighbour-joining (NJ) algorithm

[67]. Therefore, the CI and RI of each trait were calculated given the phylogenetic and geographical trees. These computations were implemented by the *CI()* and *RI()* functions in R, respectively.

(d) Ancestral state reconstruction

We employed ancestral state reconstruction [68] to estimate the states of linguistic traits and show their posterior probabilities at the internal nodes given the phylogeny. Reconstruction using the 'Most-Recent-Common-Ancestor' method was implemented by the *addMRCA* command in *BayesTraits* v. 3.0.2. We used the MCMC method, where the number of iterations to run the chain was 100 000 000 with a sampling frequency of 1000 per iteration. The first iterations of 50 000 000 were discarded as burn-in. We also rescaled the branch length by 0.0001 for Sino-Tibetan trees because the branch length is recommended with a mean of 0.1. The sample of 900 phylogenetic trees was adopted from previous

work [46] with considerations of phylogenetic uncertainty. Based on the consensus from historical documents [49,50], we identified the ancestor of the Tibetan subclade to be non-tonal. Thus, we fixed this node using the Fossil command. We reconstructed the ancestral states of all linguistic traits in the internal nodes of phylogenetic trees and showed the mean values of posterior probabilities (see electronic supplementary material, figure S3).

(e) Evolutionary correlations among linguistic traits

We used the *Discrete* algorithm [69] to examine the evolutionary correlation between two binary-coded linguistic traits by comparing the fit of two continuous-time Markov models. One of these is an independent model in which the two traits evolve independently on the tree. The other is a dependent model in which the evolution of two linguistic traits is correlated with each other. In other words, the state of one binary trait influences the transition rate of the other trait in the dependent model. We used the reversible-jump Markov chain Monte Carlo (RJ-MCMC) method [70] to perform the *Discrete* algorithm. The RJ-MCMC method visits each model in proportion to their posterior probabilities across the phylogenetic tree, together with estimating the posterior distribution of the rate parameters. To compare these two models, we used the stepping-stone sampler [71] to estimate the marginal likelihood of each model and converted it into log BF as follows:

$$\log \text{BF} = 2\log\left(\frac{L_{\text{Ind}}}{L_{\text{Dep}}}\right),$$

where a value of log BF greater than 2 indicates positive evidence for the dependent model over the independent model. A log BF greater than 5 and less than 10 indicates strong evidence for the dependent model, while a log BF greater than 10 indicates very strong evidence for the dependent model [72]. We set a prior exponential distribution of all parameters with a mean of 10. We set the number of chain iterations to 100 000 000 and the burn-in to 50 000 000. The stepping-stone sampler [71] was used with 100 stones and ran each stone for 1000 iterations. A sample of 900 phylogenetic trees was adopted for uncertainty. The branch length was also rescaled by 0.0001 for Sino-Tibetan trees. Each run was repeated five times in both dependent and independent models for stability considerations, and the median value was chosen. This computational process was implemented by the *Discrete* package in *BayesTraits* v. 3.0.2 [69,70] (http://www.evolution.reading.ac.uk/BayesTraitsV3.0.2/BayesTraitsV3.0.2.html).

(f) Directions of trait evolution

We examined the transition rate between pairwise linguistic traits whose evolutionary correlation was satisfied with the dependent model in the RJ-MCMC method. We calculated the posterior mean value of each transition rate. To examine the statistical significance of transition rates, we calculated the prior odds and posterior odds based on the frequency for all the potential models in the prior expectation and those in the posterior distribution of RJ-MCMC samples.

To compute prior odds, we used Stirling numbers [70] to calculate the total number of potential evolution models. For eight parameters, the number of combinations with unrestricted transition rates is 4140. With transition rates set to zero, we multiplied the number of possible positions that zeros can take with the number of combinations that the remaining non-zero rates can take to obtain 17 006 restricting models. Together with 4140, the total number of possible combinations of transition rates is 21 146 (see electronic supplementary material, table S6). Based on the RJ-MCMC samples, we can compute posterior odds, and the posterior odds can be divided by the prior odds to calculate BF to examine how the transition rates are significantly biased to zero. BF < 1 indicates evidence against the model, 1 < BF < 3 indicates positive support, 3 < BF < 10 indicates substantial evidence,

10 < BF < 30 indicates strong evidence, 30 < BF < 100 indicates very strong evidence and BF > 100 indicates decisive evidence [42].

(g) Examination of the effects of linguistic traits on the diversification rate

To examine whether the evolution of linguistic traits affected the speciation and extinction rates, we performed binary state speciation and extinction (BiSSE) [73] and hidden state speciation and extinction (HiSSE) [74] tests. These tests fit various trait-dependent models and trait-independent models. HiSSE incorporates hidden/unmeasured factors that could affect the estimation of diversification rates of the observed trait.

All the phylogenetic linguistic traits were tested by 24 different models that were employed in Beaulieu *et al.* [74]. These 24 models incorporate four BiSSE models, four trait-independent models (CID models), and 16 HiSSE models in which a hidden factor associated with the observed state affects the diversification rate. The dual transitions were disallowed. Due to high incidences of false-positive results in SSE models, we compared the BiSSE and HiSSE models with considerations of hidden states. The consensus tree was forced into an ultrametric tree, due to which the 'Old Chinese' language was deleted. All the models were evaluated based on the Akaike information criterion (AIC) scores. Detailed parameter settings of 24 different models can be found in the electronic supplementary material, table 'Diversification.xlsx'. The analyses were implemented by the *hisse.old()* and *hisse.null4.old()* functions in the *hisse* package v. 2.1.6 in R.

3. Results

(a) Phylogenetic signals and ancestral states of tones and other linguistic traits

Following the recommendations of Evans *et al.* [75], we first calculated the phylogenetic signals for all the 11 linguistic traits using the D-statistic to determine how well the linguistic trait would fit a phylogenetic tree. The results of the D-statistic showed strong phylogenetic signals for all linguistic traits except for the contrastive vowel length (a median D estimate for 10 traits is 0.1 and D for contrastive vowel length is 0.582, P(D = 0) = 0.024, see electronic supplementary material, table S2). Thus, this linguistic trait was excluded from all further analyses. Among the remaining traits, three traits exhibited phylogenetic conservatism with a negative D-statistic, including the voicing contrast on syllable-initial fricatives, voicing contrast on syllable-initial stops and tense/lax vowel contrast (see electronic supplementary material, table S2). These traits with negative D-statistics suggested the strong tendency of Sino-Tibetan languages to retain the phylogenetic clustering of these traits. Moreover, we performed ancestral character reconstruction and showed that the Proto-Sino-Tibetan language was probably non-tonal supported by a posterior probability of 79.9% (figure 2; electronic supplementary material, table S3). This result supports the conjecture of historical linguists [76]. In addition, the ancestral states of the other nine linguistic traits were also in line with traditional historical linguistics knowledge (see electronic supplementary material, figure S3 and table S3).

(b) Evolutionary relationships between tonal contrast and other linguistic traits

To identify the evolutionary relationships between tonal contrast and other linguistic traits, we then used a Bayesian
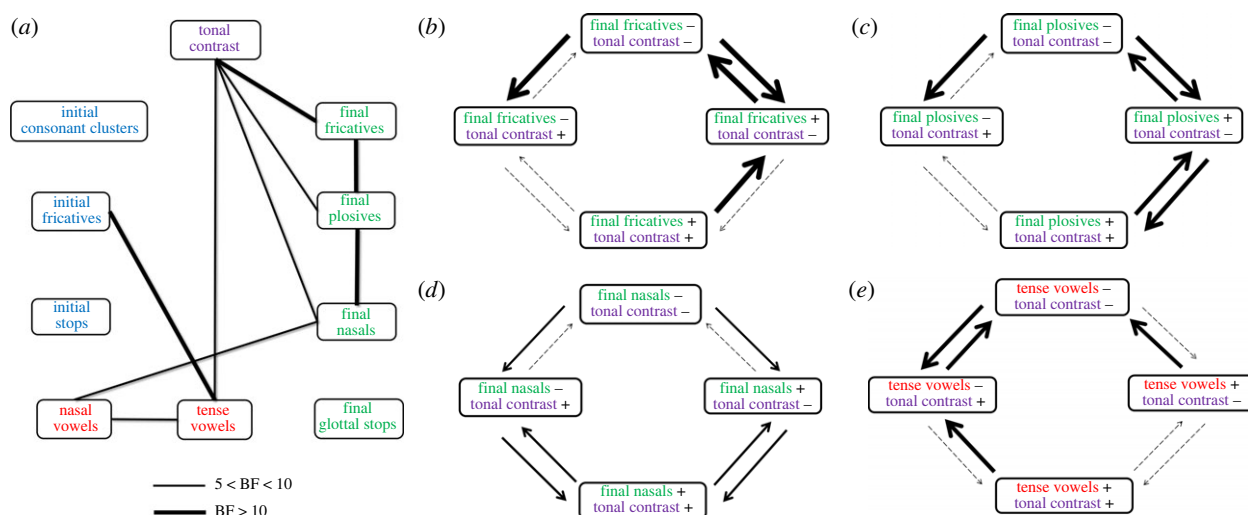
Figure 3. Summary of evolutionary dependencies among linguistic traits and their inferred transition rates. (*a*) All paired coevolutionary linguistic traits (defined as BF > 5) are shown with line width proportional to BF values. Texts in different colours show different groups of traits related to syllable-initial segments, the nucleus and the coda (see details in electronic supplementary material, table S4). (*b*–*e*) Inferred transition rates between states of tonal contrast and related traits. Arrows show the directions of the evolutionary transitions, and arrow widths are proportional to inferred transition rates (see electronic supplementary material, table S5). Dashed arrows indicate transition rates are not significantly different from 0. Detailed values can be found in the electronic supplementary material, tables S7–S10.

Table 1. Log BF between tonal contrast and the other nine linguistic traits.

	linguistic trait	the median value of log marginal likelihood		
		independent model	dependent model	log BF
⑩	syllable-final fricatives	−99.41	−87.30	24.22
⑥	tense/lax vowel contrast	−94.00	−90.14	9.72
⑧	syllable-final plosives	−112.08	−107.85	8.47
⑦	syllable-final nasals	−116.63	−113.41	6.44
①	syllable-initial consonant clusters	−119.47	−119.06	0.83
⑤	contrastive nasal vowels	−120.86	−121.25	−0.77
③	voicing contrast on syllable-initial stops	−102.24	−103.23	−1.98
⑨	syllable-final glottal stops	−113.63	−115.67	−4.07
②	voicing contrast on syllable-initial fricatives	−109.98	−112.35	−4.74

phylogenetic method implemented in the program *BayesTraits* v. 3.0.2 (figure 3; electronic supplementary material, table S4). The log BF was used to compare the difference in the marginal likelihoods between dependent and independent evolutionary models for each pair of linguistic traits (table 1). The comparisons showed that the tonal origin was evolutionarily correlated with the specific traits of voice quality and syllable-final consonants. These traits are the tense/lax vowel contrast (log BF = 9.72) and three syllable-final consonants: final fricatives (log BF = 24.22), plosives (log BF = 8.47) and nasals (log BF = 6.44). By contrast, there was no correlation between tonal origin and syllable-initial consonants. These results supported the views of traditional linguists on trigger mechanisms of tonal origin that both voice quality [77] and syllable-final consonants [17,19,20,23–25] played important roles in tonal origin.

According to the evolutionary correlations between tonal contrast and the four linguistic traits, we further investigated their evolutionary directions in Sino-Tibetan languages, and the results of the *BayesTraits* analysis are summarized in figure 3. In general, we found both common and specific patterns. Commonly, we found that the transitions from the state of both correlated trait and tonal contrast absent to the state of the correlated traits absent and tonal contrast present existed in all four patterns (figure 3*b*–*e*). This suggests that tonal contrast should arise as a compensatory mechanism for the loss of syllable-final consonants and voice quality. Specifically, we found various patterns of evolutionary directions, indicating the different coevolutionary histories of linguistic traits. In particular, the transition rates estimated in the evolutionary direction analyses were statistically significant evidence that the loss of syllable-final fricatives could precede the emergence of tonal contrast, as the syllable-final fricatives present and tonal contrast absent were reconstructed for the Proto-Sino-Tibetan language (figure 3*b*; electronic supplementary material, figure S3). This was consistent with the evolutionary mechanism that the loss of syllable-final consonants served to induce tonal origin. It was noted that the bivariate state with syllable-final fricatives absent and tonal contrast present should be evolutionarily stable, as the transitions away from this state were both inferred to be zero (see electronic supplementary

material, table S7). According to the syllable-final plosives present and tonal contrast absent estimated in the Proto-Sino-Tibetan language, there were two paths for the emergence of tonal contrast resulting from the evolution of syllable-final plosives. One was that the tonal contrast appeared with syllable-final plosives present, and the other was that the emergence of tonal contrast followed the loss of syllable-final plosives (figure 3c; electronic supplementary material, figure S3). By contrast, the absence of both tonal contrast and syllable-final nasals was an unstable state where the Sino-Tibetan languages maintained the presence of tonal contrast, syllable-final nasals or both (figure 3d). In addition, the inferred evolutionary directions illustrated that the transitions of tense/lax vowel contrast from absent to present were impossible, as these transition rates were estimated to be zero (figure 3e; electronic supplementary material, table S10). This implied that tense/lax vowel contrast is easy to lose. Under this circumstance, the origin of tonal contrast can also be used to compensate for the loss of tense/lax vowel contrast. Accordingly, all these results provided quantitative evidence that the tonal origin should be a compensatory mechanism for the loss of specific syllable-final consonants and voice quality.

(c) Tonal contrast and Sino-Tibetan language diversification

To examine whether the tonal contrast affected the diversification rates of Sino-Tibetan languages, we employed two kinds of state-dependent branching process models: BiSSE [73] and HiSSE [74]. Among the 24 models tested, CID-4 (character-independent) with two parametric restrictions was the best-fitting model with the minimum AIC values (AIC = 1674.049, see details in the electronic supplementary material, table 'Diversification.xlsx'). Under this model, mean diversification rates (speciation-extinction; $r = 0.00155$ events per year) were equal between Sino-Tibetan lineages with and without tonal contrast (electronic supplementary material, Dataset 'Diversification.xlsx'). This result indicated that the presence of tonal contrast could not significantly affect the diversification rates across the phylogenetic tree of Sino-Tibetan languages. In other words, the gain or loss of tonal contrast could have been independent of Sino-Tibetan diversification. In contrast with the tonal contrast, three other linguistic traits could have accelerated the diversification of Sino-Tibetan languages, including the presence of tense/lax vowel contrast, contrastive nasal vowels and syllable-final plosives, while the presence of syllable-final nasals could have decelerated the diversification of Sino-Tibetan languages (see electronic supplementary material, figure S4 and Dataset 'Diversification.xlsx'). In addition, we found that the influence of syllable-final plosives and syllable-final nasals was not direct on the diversification rates but mediated by hidden traits that were unobserved. In summary, the gain or loss of different linguistic traits played different roles in shaping the evolution of Sino-Tibetan languages.

4. Discussion

Studying the correlated evolution among linguistic traits is essential for understanding the evolutionary mechanisms underlying language diversity [7]. Explaining tonal origin is an intriguing topic and remains a mid- to long-term goal in studies of language evolution and human cognition. Here, we employed Bayesian PCMs to examine the tonogenetic potentials mentioned in different linguistic hypotheses and investigated the mechanisms of tonal origin in the Sino-Tibetan language family. Our results showed that tonal contrast was phylogenetically patterned, and the Proto-Sino-Tibetan language was probably non-tonal. The results of evolutionary correlation analyses quantitatively showed that tonal origin was strongly associated with evolved phonological structures such as loss of final consonants and voice quality on vowels. Furthermore, inspired by evolutionary biologists, we wanted to explore the role of tonal contrast in language diversification. We identified that tonal origin had little influence on the diversification rates of Sino-Tibetan languages.

In Sino-Tibetan languages, the tonal contrast exhibited phylogenetic patterns that suggested the tendency for closely related languages to display similar tonal patterns as a consequence of their phylogenetic proximity [63]. This finding was consistent with the known linguistic conclusion drawn from cross-language comparisons that phylogenetically related languages showed more shared values [16,26,78]. Moreover, the areal diffusion [79] of linguistic traits cannot be ignored on account of the geographical influence of language sampling bias. Accordingly, to compare the phylogenetic and geographical patterns of linguistic traits, we calculated consistency (CI) [65] and retention (RI) [66] indices for 10 traits based on the Bayesian phylogenetic tree and the geographical NJ tree, respectively. These two indices showed that the phylogenetic dependence of all traits was stronger than their geographical dependence, except for the voicing contrast on syllable-initial fricatives (see electronic supplementary material, figure S2). In addition, our results showed that the Proto-Sino-Tibetan language was non-tonal, which is consistent with previous studies of historical linguistics that the tone system could be of late origin [20].

We found that tonal origin was correlated with evolved phonological structures such as loss of final consonants and voice quality on vowels. In linguistics, tonal origin can be explained by two distinctive models named the segmental-driven model and the laryngeal-based model [21]. The segmental-driven model suggests that the tonal origin is induced by evolved syllabic segments, while the laryngeal-based model emphasizes the essential contribution of phonation to tonogenesis. In this study, our findings reconciled the conflict between these two models. In particular, the loss of syllable-final consonants, including plosives, fricatives and nasals, [16,17,24] could have induced the emergence of tonal contrast, which was supported by Haudricourt's claims on tonal origin observed in Vietnamese [19] and Chinese [20,24,26]. From the phonetic perspective, specific syllable-final consonants may affect the pitch of preceding vowels. Then, the pitch difference can shape the tonal contrast when the syllable-final consonants disappear [21]. For example, from Old Chinese to Middle Chinese, the 'departing' tone appeared in syllables ending in /-h/, which had originated in /-s/ before the tone arose [26]. In addition, tonal contrast can linguistically be generated from the tense/lax vowel contrast in Sino-Tibetan languages [77]. The lax vowel in a syllable has a lower tone, while the tense vowel exhibits together with a higher tone [21]. However, their evolutionary pathway remains uncertain according to the ambiguous ancestral state of tense/lax vowel contrast in the Proto-Sino-

Tibetan language with a posterior probability close to 50% (see electronic supplementary material, figure S3), and no state was stable in the transition rate estimation (figure 3e). In summary, we quantitively identified that tonal contrast arose as a compensatory mechanism for the structural organization and evolution of languages which was consistent with previous linguistic studies [16,51].

Since tonal origin could be a compensatory mechanism for the evolution of syllabic structures, it may not serve as a primary driving force for changing diversification rate of languages. However, such suggestions have not been examined rigorously. In this study, we found that the tonal origin indeed contributed little to shaping the diversification rates of extant Sino-Tibetan languages. By contrast, four linguistic traits are correlated with the Sino-Tibetan language diversification rates, including tense/lax vowel contrast, contrastive nasal vowels, syllable-final nasals and syllable-final plosives. Both the presence of tense/lax vowel contrast and contrastive nasal vowels accelerated the diversification rates, especially in the Nusu clade (see electronic supplementary material, figure S4). In the Burmish clade, the presence of syllable-final plosives accelerated the diversification rates, while the presence of syllable-final nasals decelerated the diversification rates. Our results showed that linguistic traits played various roles in shaping the evolutionary processes of specific language subclades.

In addition, our study illustrates how PCMs can be applied in quantitatively investigating the evolutionary mechanism of tonal origin in Sino-Tibetan languages. Different from the comparative method in historical linguistics, PCMs provide quantitative insights into the evolutionary process of linguistic traits controlling for shared language history [80]. These methods solved the well-known Galton's problem of drawing inferences from cross-cultural data [80–82]. The problem can be described as making statistical estimates when the elements sampled are not statistically independent. By overcoming this difficulty, PCMs have been recently used, for example, to examine the lineage-specific trends in word-order universals [1], to explore the origins of sortal numeral classifiers [6] and to assess the cognitive development of colour terms [36]. Therefore, these quantitative methods are powerful tools that have been widely employed to help us understand the language evolution given phylogenies [83] controlling for the genetic relatedness of samples. Overall, PCMs facilitate us to better understand that tone arose as a compensatory mechanism for the structural organization and evolution of languages.

References

1. Dunn M, Greenhill SJ, Levinson SC, Gray RD. 2011 Evolved structure of language shows lineage-specific trends in word-order universals. Nature 473, 79–82. (doi:10.1038/nature09923)

2. Blasi DE, Moran S, Moisik SR, Widmer P, Dediu D, Bickel B. 2019 Human sound systems are shaped by post-Neolithic changes in bite configuration. Science 363, eaav3218. (doi:10.1126/science.aav3218)

3. Hruschka DJ, Branford S, Smith ED, Wilkins J, Meade A, Pagel M, Bhattacharya T. 2015 Detecting regular sound changes in linguistics as events of concerted evolution. Curr. Biol. 25, 1–9. (doi:10.1016/j.cub.2014.10.064)

4. Futrell R, Mahowald K, Gibson E. 2015 Large-scale evidence of dependency length minimization in 37 languages. Proc. Natl Acad. Sci. USA 112, 10 336–10 341. (doi:10.1073/pnas.1502134112)

5. Pagel M, Meade A. 2018 The deep history of the number words. Phil. Trans. R. Soc. B 373, 20160517. (doi:10.1098/rstb.2016.0517)

6. Cathcart CA, Hölzl A, Jäger G, Widmer P, Bickel B. 2020 Numeral classifiers and number marking in Indo-Iranian: a phylogenetic approach. Lang. Dyn. Change 11, 273–325. (doi:10.1163/22105832-bja10013)

7. Bailes R, Cuskley C. 2023 The cultural evolution of language. In The Oxford handbook of cultural evolution (eds JJ Tehrani, J Kendal, R Kendal). Oxford, UK: Oxford University Press.

8. Yip M. 2002 Tone. Cambridge, UK: Cambridge University Press.

9. Nazzi T, Cutler A. 2019 How consonants and vowels shape spoken-language recognition. Annu. Rev. Ling. 5, 25–47. (doi:10.1146/annurev-linguistics-011718-011919)

10. Gordon MK. 2016 Phonological typology. Oxford, UK: Oxford University Press.

11. Welmers WE. 2018 African language structures. Berkeley, CA: University of California Press.

12. Hyman L, Oesterreicher W. 2001 Tone systems. In Language typology and language universals: an international handbook, vol. 2 (eds M Haspelmath, E König, W Raible), pp. 1367–1380. Berlin, Germany: Walter de Gruyter.

13. Maddieson I. 2008 Glides and gemination. Lingua 118, 1926–1936. (doi:10.1016/j.lingua.2007.10.005)

14. Everett C, Blasi DE, Roberts SG. 2015 Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. Proc. Natl Acad. Sci. USA 112, 1322–1327. (doi:10.1073/pnas.1417413112)

15. Teeranon P. 2012 Initial consanant voicing perturbation of the fundamental frequency of oral vowels and nasal vowels: a controversial case from

Ban Doi Pwo Karen. *MANUSYA: J. Human.* **15**, 39–59. (doi:10.1163/26659077-01502003)

16. Matisoff JA. 1973 Tonogenesis in southeast Asia. *Consonant Types Tone* **1**, 71–96.

17. Michaud A, Sands B. 2020 Tonogenesis. In *Oxford research encyclopedia of linguistics* (ed. Aronoff Mark). Oxford, UK: Oxford University Press. (doi:10.1093/acrefore/9780199384655.013.748)

18. Wang WSY. 1967 Phonological features of tone. *Int. J. Am. Ling.* **33**, 93–105. (doi:10.1086/464946)

19. Przyluski J. 1924 Les langues austroasiatiques. *Les langues du monde, Collection linguistique* **16**, 385–403.

20. Haudricourt AG. 1954 Comment reconstruire le chinois archaïque. *Word* **10**, 351–364. (doi:10.1080/00437956.1954.11659532)

21. Thurgood G. 2002 Vietnamese and tonogenesis: revising the model and the analysis. *Diachronica* **19**, 333–363. (doi:10.1075/dia.19.2.04thu)

22. Hombert JM. 1978 A model of tone systems. In *Elements of tone, stress intonation* (ed. DJ Napoli), pp. 129–143. Washington, DC: Georgetown University Press.

23. Alves M. 1995 Tonal features and the development of Vietnamese tones. *Work. Papers Ling.* **27**, 1–13.

24. Haudricourt AG. 1954 *De l'origine des tons en vietnamien*. Paris, France: Imprimerie Nationale.

25. Ratliff M. 2015 Tonoexodus, tonogenesis, and tone change. In *The Oxford handbook of historical phonology* (eds P Honeybone, J Salmons), pp. 245–261. Oxford, UK: Oxford Academic.

26. Kingston J. 2011 Tonogenesis. *Blackwell Comp. Phonol.* **4**, 1–30.

27. Pulleyblank EG. 1978 The nature of the Middle Chinese tones and their development to Early Mandarin. *J. Chinese Ling.* **6**, 173–203.

28. DeLancey S. 2003 Lhasa Tibetan. *Sino-Tibetan Lang.* **3**, 270–288.

29. Enfield NJ. 2005 Areal linguistics and mainland Southeast Asia. *Annu. Rev. Anthropol.* **34**, 181–206. (doi:10.1146/annurev.anthro.34.081804.120406)

30. Maddieson I. 1997 Combining frication and glottal constriction: two solutions to a dilemma. *J. Acoust. Soc. Am.* **102**, 3135. (doi:10.1121/1.420642)

31. Zhang LS. 1987 The consonant system of middle-old Tibetan and the tonogenesis of Tibetan. PhD thesis, University of California, Berkeley, CA.

32. Hu F. 2012 Tonogenesis in Lhasa Tibetan: towards a gestural account. *Consonant Clusters Struct. Complexity* **26**, 231. (doi:10.1515/9781614510772.231)

33. Pulleyblank EG. 1999 Chinese traditional phonology. *Asia Major.* **12**, 101–137.

34. Svantesson JO (ed.). 2001 Tonogenesis in Southeast Asia–Mon-Khmer and beyond. In *Proc. of the Symp. Cross-Linguistic Studies of Tonal Phenomena*. Institute for the Study of Languages and Culture of Asia and Africa, Tokyo University of Foreign Studies.

35. Mace R, Holden CJ. 2005 A phylogenetic approach to cultural evolution. *Trends Ecol. Evol.* **20**, 116–121. (doi:10.1016/j.tree.2004.12.002)

36. Haynie HJ, Bowern C. 2016 Phylogenetic approach to the evolution of color term systems. *Proc. Natl Acad. Sci. USA* **113**, 13 666–13 671. (doi:10.1073/pnas.1613666113)

37. Hübler N. 2022 Phylogenetic signal and rate of evolutionary change in language structures. *R. Soc. Open Sci.* **9**, 211252. (doi:10.1098/rsos.211252)

38. Macklin-Cordes JL, Round ER. 2020 Re-evaluating phoneme frequencies. *Front. Psychol.* **11**, 570895. (doi:10.3389/fpsyg.2020.570895)

39. Carling G, Cathcart C. 2021 Reconstructing the evolution of Indo-European grammar. *Language* **97**, 561–598.

40. Phillips J, Bowern C. 2022 Bayesian methods for ancestral state reconstruction in morphosyntax: exploring the history of argument marking strategies in a large language family. *J. Lang. Evol.* **7**, 1–15. (doi:10.1093/jole/lzac002)

41. Opie C, Shultz S, Atkinson QD, Currie T, Mace R. 2014 Phylogenetic reconstruction of Bantu kinship challenges main sequence theory of human social evolution. *Proc. Natl Acad. Sci. USA* **111**, 17 414–17 419. (doi:10.1073/pnas.1415744111)

42. Basava K, Zhang H, Mace R. 2021 A phylogenetic analysis of revolution and afterlife beliefs. *Nat. Hum. Behav.* **5**, 604–611. (doi:10.1038/s41562-020-01013-4)

43. Currie TE, Greenhill SJ, Gray RD, Hasegawa T, Mace R. 2010 Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* **467**, 801–804. (doi:10.1038/nature09461)

44. Savage PE, Brown S, Sakai E, Currie TE. 2015 Statistical universals reveal the structures and functions of human music. *Proc. Natl Acad. Sci. USA* **112**, 8987–8992. (doi:10.1073/pnas.1414495112)

45. Simons G, Fennig C. 2018 *Ethnologue: languages of the world*, 21st edn. Dallas, TX: SIL International.

46. Zhang M, Yan S, Pan W, Jin L. 2019 Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature* **569**, 112–115. (doi:10.1038/s41586-019-1153-z)

47. Sagart L, Jacques G, Lai Y, Ryder RJ, Thouzeau V, Greenhill SJ, List JM. 2019 Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proc. Natl Acad. Sci. USA* **116**, 10 317–10 322. (doi:10.1073/pnas.1817972116)

48. Zhang H, Ji T, Pagel M, Mace R. 2020 Dated phylogeny suggests early Neolithic origin of Sino-Tibetan languages. *Sci. Rep.* **10**, 20792. (doi:10.1038/s41598-020-77404-4)

49. Sun Hongkai HZ. 2007 *Zhongguo de yuyan* 《中国的语言》. Beijing, China: Commercial Press. [In Chinese].

50. Jiangping K. 2017 Multi-dimensional viewpoints on the origin of Tibetan tones/藏语声调起源的多维视角. *J. Chin. Lin. Monogr. Ser.* **27**, 105–129.

51. Matisoff JA (ed) 1999 Tibeto-Burman tonology in an areal context. In *Proc. of the Symp. 'Crosslinguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics, 10–12 December, 1998*. Tokyo, Japan: Tokyo University of Foreign Studies, Institute for the Study of Languages and Cultures of Asia and Africa.

52. Namkung J. 1996 *Phonological inventories of Tibeto-Burman languages: Sino-Tibetan etymological dictionary and thesaurus project, center for Southeast Asia studies*. Berkeley, CA: University of California.

53. Huang Bufan DQ. 1992 *Zangmianyuzu yuyan cihui* 《藏緬語族語言詞匯》 [*A Tibeto-Burman Lexicon*]. Beijing, People's Republic of China: Central Institute of Minorities.

54. Hongkai S. 1991 *Zangmianyu yuyin he cihui* 《藏缅语语音和词汇》 [*Tibeto-Burman phonology and lexicon*]. Beijing, People's Republic of China: Chinese Social Sciences Press.

55. Yongsui L, Ersong W. 1986 *Haniyu jianzhi (brief description of the hani language)*. Beijing, People's Republic of China: Minzu Chubanshe.

56. Manson K. 2007 *Pekon Kayan phonology*. Chiang Mai, Thailand: Payap University.

57. Donohue M, Hetherington R, McElvenny J, Dawson V. 2013 *World phonotactics database*. Canberra, Australia: Australian National University, See https://www.research.ed.ac.uk/en/publications/world-phonotactics-database.

58. So-Hartmann H. 2008 *A descriptive grammar of Daai Chin*. London, UK: University of London, School of Oriental and African Studies.

59. Singha D. 2003 The phonology & morphology of Dimasa. MA dissertation, Assam University, Silchar, Assam, India.

60. Vittrant A, Watkins J, Matisoff J. 2019 *The mainland Southeast Asia linguistic area*. Berlin, Germany: De Gruyter.

61. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014 BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, e1003537. (doi:10.1371/journal.pcbi.1003537)

62. Swadesh M. 1955 Towards greater accuracy in lexicostatistic dating. *Int. J. Am. Ling.* **21**, 121–137. (doi:10.1086/464321)

63. Keck F, Rimet F, Bouchez A Franc A. 2016 phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol. Evol.* **6**, 2774–2780. (doi:10.1002/ece3.2051)

64. Fritz SA, Purvis A. 2010 Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conserv. Biol.* **24**, 1042–1051. (doi:10.1111/j.1523-1739.2010.01455.x)

65. Kluge AG, Farris JS. 1969 Quantitative phyletics and the evolution of anurans. *Syst. Biol.* **18**, 1–32. (doi:10.1093/sysbio/18.1.1)

66. Farris JS. 1989 The retention index and the rescaled consistency index. *Cladistics: Int. J. Willi Hennig Soc.* **5**, 417–419. (doi:10.1111/j.1096-0031.1989.tb00573.x)

67. Saitou N, Nei M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

68. Pagel M, Meade A, Barker D. 2004 Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* **53**, 673–684. (doi:10.1080/10635150490522232)

69. Pagel M. 1994 Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. Ser. B* **255**, 37–45. (doi:10.1098/rspb.1994.0006)

70. Pagel M, Meade A. 2006 Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* **167**, 808–825. (doi:10.1086/503444)

71. Xie W, Lewis PO, Fan Y, Kuo L, Chen MH. 2011 Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* **60**, 150–160. (doi:10.1093/sysbio/syq085)

72. Gilks WR, Roberts GO. 1996 Strategies for improving MCMC. In *Markov chain Monte Carlo in practice* (eds WR Gilks, S Richardson, D Spiegelhalter), pp. 89–114. Boca Raton, FL: Chapman and Hall/CRC.

73. Maddison WP, Midford PE, Otto SP. 2007 Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710. (doi:10.1080/10635150701607033)

74. Beaulieu JM, O'Meara BC. 2016 Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601. (doi:10.1093/sysbio/syw022)

75. Evans CL, Greenhill SJ, Watts J, List JM, Botero CA, Gray RD, Kirby KR. 2021 The uses and abuses of tree thinking in cultural evolution. *Phil. Trans. R. Soc. B* **376**, 20200056. (doi:10.1098/rstb.2020.0056)

76. Haudricourt AG. 1961 Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient. *Bullet. de la Société de Linguist. de Paris* **56**, 163–180.

77. Jiahua Y. 1981 Hanzangyu shengdiao de qiyuan he yanbian (On the tonogenesis and tonal development of Sino-Tibetan languages). *Ling. Res.* **02**, 2–7.

78. Haudricourt AG. 1971 New Caledonia and the Loyalty Islands. *Curr. Trends Ling.* **8**, 359–396.

79. Maddieson I. 2013 *Voicing in plosives and fricatives: the world atlas of language structures online*. Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology. See https://wals.info/chapter/4 (accessed 16 May 2022).

80. Mace R, Pagel M, Bowen JR, Gupta BKD, Otterbein KF, Ridley M, Schweizer T, Voland E. 1994 The comparative method in anthropology [and comments and reply]. *Curr. Anthropol.* **35**, 549–564. (doi:10.1086/204317)

81. Roberts S, Winters J. 2013 Linguistic diversity and traffic accidents: lessons from statistical studies of cultural traits. *PLoS ONE* **8**, e70902. (doi:10.1371/journal.pone.0070902)

82. Ladd DR, Roberts SG, Dediu D. 2015 Correlational studies in typological and historical linguistics. *Annu. Rev. Ling.* **1**, 221–241. (doi:10.1146/annurev-linguist-030514-124819)

83. Gray RD, Greenhill SJ, Ross RM. 2007 The pleasures and perils of Darwinizing culture (with phylogenies). *Biol. Theory* **2**, 360–375. (doi:10.1162/biot.2007.2.4.360)

84. Wu B, Zhang H, Zhang M. 2023 Phylogenetic insight into the origin of tones. Figshare. (doi:10.6084/m9.figshare.c.6700046)